ABSTRACT

Title of Document: SOCIO-TECHNICAL TRANSITION AS A CO-EVOLUTIONARY PROCESS: INNOVATION AND THE ROLE OF NICHE MARKETS IN THE TRANSITION TO MOTOR VEHICLES

Alicia K. Birky, Doctor of Philosophy, 2008

Directed By: Professor Matthias Ruth
School of Public Policy

Significant reductions in greenhouse emissions from personal transportation will require a transition to an alternative technology regime based on renewable energy sources. Two bodies of research, the quasi-evolutionary (QE) model and the multi-level perspective (MLP) assert that processes within niches play a fundamental role in such transitions. This research asks whether the description of transitions based on this niche hypothesis and its underlying assumptions is consistent with the historical U.S. transition to motor vehicles at the beginning of the 20th century.

Unique to this dissertation is the combination of the perspective of the entrepreneur with co-evolutionary approaches to socio-technical transitions. This approach is augmented with concepts from the industry life-cycle model and with a taxonomy of mechanisms of learning. Using this analytic framework, I examine specifically the role of entrepreneurial behavior and processes within and among firms in the co-evolution of technologies and institutions during the transition to motor vehicles.

I find that niche markets played an important role in the development of the technology, institutions, and the industry. However, I also find that the diffusion of the

automobile is not consistent with the niche hypothesis in the following ways: 1) product improvements and cost reductions were not realized in niche markets, but were achieved simultaneously with diffusion into mass markets; 2) in addition to learning-by-doing and learning-by-interacting with users, knowledge spillovers and interacting with suppliers were critical in this process; 3) cost reductions were not automatic results of expanding markets, but rather arose from the strategies of entrepreneurs based on personal perspectives and values. This finding supports the use of a behavioral approach with a micro-focus in the analysis of socio-technical change

I also find that the emergence and diffusion of the motor vehicle can only be understood by considering the combination of developments and processes in multiple regimes, within niches, and within the wider technical, institutional, and ecological complex (TIEC). For the automobile, the process of regime development was more consistent with a fit-stretch pattern of gradual unfolding and adaptation than one of niche proliferation and rapid regime renewal described in the literature.

SOCIO-TECHNICAL TRANSITION AS A CO-EVOLUTIONARY PROCESS: INNOVATION AND THE ROLE OF NICHE MARKETS IN THE TRANSITION TO MOTOR VEHICLES

By

Alicia K. Birky

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor Matthias Ruth, Chair
Professor David Kirsch
Professor Nathan Hultman
Professor Emeritus George Dieter
Leon Clarke, PhD
Elizabeth Malone, PhD

Dedication

To Hannah. If my work makes the world that you inherit even the tiniest bit better, then I will consider it an unqualified success.

Acknowledgements

There are many people without whom I would never have completed this research. To start at the beginning, I would like to thank Rob Sprinkle for convincing me that a PhD in policy would allow me to influence change far more effectively than other choices I was considering. I am indebted to Peter Brown and Herman Daly for teaching me the importance of context and challenging me to reconsider my pre-analytic vision. To them I owe the inspiration for my approach. I am also grateful to Leon Clarke of the Joint Global Change Research Institute for directing my inquiry along the path of innovation and providing initial funding to get underway.

I would like to thank my advisor, Matthias Ruth, for his support and encouragement throughout this process, despite life's many interruptions along the way. I am grateful to all the members of my committee for their invaluable input. Their guidance and contributions greatly improved the product. I am also grateful for the invaluable assistance of Mac Destler and Karen Logan in navigating university procedures and submitting paperwork.

I would also like to thank my family and friends for their unflagging support, for nudging me along with the frequent "how's that dissertation coming?" but never believing I wouldn't finish. I am deeply indebted to my husband, John Maples, not only for his emotional and physical support, but also for his professional insight. Our many discussions helped clarify my thinking and provided fruitful directions for inquiry. Finally, I would like to thank my daughter, Hannah, for reminding me why it all matters and suffering through with me.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAA	American Automobile Association
AFV	alternative fuel vehicle
ALAM	Association of Licensed Automobile Manufacturers
AMCMA	American Motor Car Manufacturers Association
AML	American Motor League
ASME	American Society of Mechanical Engineers
AT-PZEV	advanced technology partial ZEV
CAFE	Corporate Average Fuel Economy
CARB	California Air Resources Board
CO_2	carbon dioxide
EPA	U.S. Environmental Protection Agency
EPAct	Energy Policy Act
ESB	Electric Storage Battery Company
EVAA	Electric Vehicle Association of America
EVC	Electric Vehicle Company
GHG	greenhouse gas
GM	General Motors
GMAC	General Motors Acceptance Corporation
ICE	internal combustion engine
IPF	innovation possibilities function
IR&D	internal research and development
kph	kilometers per hour
LEV	low emission vehicle
MLP	multi-level perspective
mph	miles per hour
NACC	National Automobile Chamber of Commerce
NAAM	National Association of Automobile Manufacturers
OPEC	Organization of the Petroleum Exporting Countries
PHS	Public Health Service
PNGV	Partnership for the Next Generation Vehicle
PZEV	partial ZEV
QE	quasi-evolutionary
QEM	quasi-evolutionary model
R&D	research and development
rpm	revolutions per minute
SAE	Society of Automobile Engineers
SNM	strategic niche management
SULEV	super ultra-low emission vehicle
TEL	tetraethyl lead
TIC	Techno-institutional complex
TIEC	Technical, institutional, and ecological complex
ULEV	ultra-low emission vehicle
VMT	vehicle miles of travel
ZEV	zero emission vehicle

1.0 Introduction

Steering an economy with positive feedbacks into the best of its many possible equilibrium states requires good fortune and good timing – a feel for the moments when beneficial change from one pattern to another is most possible.

-- W.B Arthur, 1990

1.1 Statement of the Problem

For the last twenty years, the international community has engaged in negotiations to stabilize the atmospheric concentration of greenhouse gases (GHGs). Yet scientific concern over GHG emissions arose at least fifty years ago. As Revelle and Suess (1957) succinctly stated, "[h]uman beings are carrying out a large scale geophysical experiment of a kind that could not have happened in the past nor be produced in the future. Within a few centuries, we are returning to the atmosphere and the oceans the concentrated organic carbon stored in sedimentary rock over hundreds of millions of years."

This experiment was enabled by the development of energy technologies and infrastructures that have vastly improved the human condition, even while threatening human and ecosystem health on local, regional, and global scales. While the energy intensity of the U.S. economy has declined, total energy use continues to rise and has tripled since 1950. During that time, transportation has accounted for around one quarter of the nation's energy consumption, though that fraction has risen slightly to 28% (EIA, 2007). Because the U.S. transportation system is almost entirely dependent on petroleum, this sector is currently responsible for about 28% of the nation's total GHG emissions. Transportation contributes the largest fraction of U.S. carbon dioxide (CO_2) emissions from the combustion of fossil fuels by end use sector at 33%. Light vehicles generally used for personal transport (automobiles and light trucks) alone account for

20% of U.S. CO_2 emissions from the combustion of fossil fuels (U.S. EPA, 2007a). In addition to being a significant contributor to emissions, the personal transportation system poses unique problems for control because of the extremely large number and wide distribution of sources, as well as the fact that vehicle purchasing decisions involve more than rational cost considerations.

Carbon dioxide emissions from personal vehicles are a function of carbon content of the fuel, total fuel consumption, and engine properties.[1] Consumption is in turn a function of vehicle fuel efficiency and vehicle miles of travel (VMT). This lends four approaches to addressing the issue of greenhouse gas emissions from personal transportation: efficiency, substitution, travel demand reduction, and carbon capture and sequestration. Technological options addressing efficiency and substitution are generally considered necessary components of any greenhouse gas reduction strategy.

Improving the efficiency of vehicles slows consumption of oil, reduces annual greenhouse gas emissions, and allows more time to develop and deploy alternative technologies. Substitution of alternative fuels displaces demand for petroleum and can reduce GHG emissions, depending on the selected alternative fuel and its production process. While there are renewable fuels and end-use technologies with lower life-cycle GHG emissions that currently are technologically feasible, they pose challenges in terms of relative cost, performance, safety, and infrastructure. Meanwhile, petroleum remains plentiful and relatively inexpensive in the near term. Should the availability or price of petroleum become problematic, due to depletion or socio-political factors, higher fuel prices might stimulate more fuel efficient vehicle technologies and use patterns, thereby

[1] Carbon is emitted as carbon dioxide, carbon monoxide, soot (black carbon), methane, and various organic compounds, depending on engine properties.

slowing the growth of GHG emissions. However, liquid fuels for transportation could be obtained from other plentiful fossil resources (e.g. unconventional petroleum, coal, and natural gas) and used with little alteration of the existing vehicle technology paradigm that has evolved over a period of more than 100 years (Johnson et al., 2003). Yet if road transport remains the dominant mode for personal mobility, significant reductions in emissions will require widespread transition to an alternative technology regime based on renewable energy sources.

In 1973, OPEC production cuts and the Arab oil embargo resulted in a dramatic increase in petroleum fuel prices as well as shortages at gasoline stations.[2] In an effort to decrease U.S. dependence on foreign oil, Congress passed the Energy Policy and Conservation Act (EPAct) of 1975. Since then, there have been numerous federal and state policies directed at improving vehicle efficiency and developing alternatives to petroleum fuels. Despite these efforts, personal transportation remains almost entirely dependent on petroleum and improvements in fuel economy have largely stalled. This failure raises the following questions for policy:

- What conditions would foster the transition of the U.S. personal transportation system from the existing socio-technological regime based on petroleum to one which produces lower carbon emissions?
- Can the development of alternate technology alone enable this transition or are other factors critical in stimulating a transition?

[2] In 1973, Egypt, Syria, and the Arab members of OPEC embargoed shipments of oil to countries that supported Israel during the on-going Yom Kippur War. They specifically targeted the United States and the Netherlands, the latter having supplied arms to Israel and allowed U.S. use of Dutch airfields.

- What are the policy leverage points for facilitating a transition and steering the system to a desired outcome?

1.2 Background: Policy Approaches

Policy approaches to improve the energy efficiency of U.S. light vehicles and promote the development and use of alternative fuel can be grouped into four categories: research and development (R&D), market incentives, regulations, and what I will call niche approaches. These categories relate to explanations for the lack of a market for such technologies and the underlying theoretical assumptions of these explanations. This section discusses these underlying assumptions and demonstrates that niche approaches have emerged as the only policies that attempt to address the underlying dynamics of technical transitions.

1.2.1 Research and Development (R&D)

The first explanation for that lack of a market for alternative technologies is that they simply are not mature enough and that additional research and development (R&D) is needed to overcome cost and performance issues. Government sponsored R&D is justified because private industry generally under-invests in innovation and in particular under-invests in research to solve social problems that promise little profit or in technologies with potentially high payoff but high risk. R&D policy therefore directs innovation toward the solution of specific socially significant problems and may also determine the search space. Federal support of R&D includes research programs at federal laboratories, research grants to other research institutions, and public-private partnerships such as the Partnership for the Next Generation Vehicle (PNGV) established in 1993 and FreedomCAR which replaced PNGV in 2002.

The underlying assumption, then, is that technical and cost hurdles are the only barrier to market development. The new technology will compete successfully once it is developed to the point where it is competitive, which is to say it provides equivalent service at lower cost or better service at an appropriate marginal cost.

1.2.2 Market Incentives

Proponents of market incentives argue that alternative technologies should not be required to compete on the same cost basis because the price signals are incorrect. Market incentives are recommended to correct for various market failures – public goods, externalities, inadequate or asymmetric information, and natural monopolies – that lead to inefficient markets or inequitable distribution of welfare. Tax incentives to fuel producers have been used to stimulate production of ethanol and bio-diesel as well as stimulate further research. Federal and state fuel taxes may also be restructured to decrease the relative cost of alternative fuels, on an energy content basis, to the consumer. Tax incentives and disincentives have also been applied to vehicles, such as the federal 'gas guzzler' tax imposed on the most inefficient new cars. Recently, federal and state income tax incentives have also been offered for the purchase of hybrid electric vehicles and alternative fuel vehicles (AFVs). However, experience has shown that market incentives may need to be extreme to influence consumer behavior.

1.2.3 Regulations

Both R&D and market policies fundamentally rest on the assumptions of neo-classical economics, with the understanding that true markets violate these assumptions in a limited sense. However, experience has shown that these policies, even when combined, are not necessarily sufficient to insure that new technologies are deployed and adopted. The problem is that neo-classical economics relies on static analysis and has little to offer in terms of explaining the dynamic processes involved in transitioning from one equilibrium state to another.

The establishment of regulations or standards can circumvent the difficulties posed by the dynamics of transitions by ‘forcing’ innovation in a particular direction and ensuring deployment. This policy approach may be justified by the existence of market failures but is often criticized as being economically inefficient; there is no guarantee that the cost of meeting the regulation is equivalent to the value of the benefits. Therefore, regulations are most often recommended when market failures lead to fairness or ethical issues which are difficult to place in monetary terms, such as the value of ecosystem services or human life.

Historic examples of regulations related to vehicle fuel consumption include the federal Corporate Average Fuel Economy (CAFE) standards and California’s Zero Emission Vehicle (ZEV) mandates. Such regulations alter the product’s design specifications by establishing functional standards, e.g., for fuel economy or tailpipe emissions. Manufacturers must determine the best technological solution to meet the politically established requirements while simultaneously satisfying consumer demands. Thus, policy ‘forces’ innovation toward the solution of a particular problem, but does not define the search space.

However, the politically established functional requirements may conflict with those of the market; existing or near-term technology may be unable to satisfy both at reasonable cost. If manufacturers are unable to find an innovative solution, product cost rises or performance falls as measured by consumer preferences.[3] In this case, manufacturers may seek to modify the regulations or delay (or even halt) implementation, as has occurred with the California ZEV mandates. Another possibility is that complex system dynamics may over time result in the innovation of new products that do satisfy both consumer preferences and the regulations as written, but this evolution may undermine the goals of the regulations. For example, although vehicle manufacturers have generally met the CAFE standards, the overall fleet fuel economy has fallen since 1987 due to changes in product offerings and an increasing light truck market share. Meanwhile, average weight and horsepower have risen appreciably

1.2.4 Niche Approaches

Where regulations tend to treat dynamic processes as a black box, niche approaches aim to influence them directly. Niche approaches may be seen as arising from a theoretical model that departs drastically from neo-classical economics. This model asserts that the difficulties faced in transitioning to alternative fuels arise because the transportation system is a complex dynamic system where positive feedbacks reinforce the dominance of the existing paradigm. The existence of these feedbacks – learning, adaptive expectations, network effects, and economies of scale – pose significant challenges to the development and adoption of promising new technologies.

In general, a niche can be defined as a distinct application domain, small in scale and scope, characterized by specific functional requirements (Hoogma et al., 2002; Raven, 2005). Because of the special functional requirements of the niche, users are

---

[3] Under CAFE, fuel economy standards are established for a manufacturer's entire sales fleet. Thus, if the cost of meeting the standard exceeds consumers' willingness to pay, manufacturers must adjust prices to encourage the purchase of more efficient vehicles over less efficient ones, thus lowering profits, or opt to pay fines established by the regulations.

willing to accept higher cost or lower performance on other non-essential functionalities compared to those attained with any other products available in the market. Niche approaches may be divided into two categories of policies, the first relating to experimental niches and the second to market niches. Analogous to Hoogma's (2000) 'technological niche,' the term experimental niche refers to non-commercial exploration of a technology in a temporary 'societal experiment' where it is protected from market selection pressures and from the engineering and use practices established by the existing regime. Market niches involve commercial production and occur naturally but may also be created by policies that provide varying degrees of protection from the selection environment.

Explorations that create experimental niches include demonstration and pilot projects. These experiments serve as a test-bed where production and use are investigated in an environment approaching real-world conditions. The goals of these projects vary. In the most limited sense, they are used to test production processes and product performance, improve them, and reduce costs. In the broader sense recommended by an emerging policy approach known as strategic niche management (SNM), experimental niches can be used to stimulate both behavioral and technical change as a stepping stone to a socio-technical transition. According to Elzen et al. (2003, p. 177), "[r]adical change can initially begin at the fringes, in so-called niches, from which it spreads in conjunction with other developments. SNM is directed at overcoming barriers to broader diffusion by exploiting niche dynamics." SNM identifies two key processes in such regime shifts: learning and institutional (or societal) embedding. Therefore, the goals of SNM are to investigate user needs and to simultaneously adapt both the technology and the social environment in which it is produced and used. By altering agent (producer and user) perceptions and behavior toward the new technology, it becomes 'embedded' in a new rule set.

Market niche policies guarantee a small market for new technologies. Examples include federal fleet and fuel provider mandates for AFVs under EPAct and local fleet projects under the Clean Cities program. Because of the structure of the California ZEV mandates, we may also consider that this policy intended to create a market niche for electric vehicles. Only a percentage of vehicles sold by each manufacturer were required to meet the ZEV criteria of zero tailpipe and zero evaporative emissions, which, at the time the regulations were promulgated, could only be met by the battery-powered electric vehicle.

One aim of market niche creation is to overcome the "chicken-and-egg" problem posed by network effects: manufacturers are unwilling to produce the alternative technology until there is market demand and an available fuel infrastructure, while fuel providers are unwilling to invest in infrastructure until there is demand for the fuel. While establishment of a stable market niche initiates commercial manufacturing, substantial emission reductions can only occur through widespread adoption. Thus, the expectation is that processes within the niche market will facilitate development of a viable mass market.

1.3 Theoretical Basis of Niche Approaches: An Overview

Niche approaches have emerged as the only policies that target the dynamic processes underlying transitions in complex socio-technical systems. These approaches clearly assert that processes within niches facilitate such transitions. According to Elzen et al. (2003), “Technological niches appear a necessary component of a regime renewal. They help to create a pathway to a new regime without which there will not be a new regime. Niches act as stepping stones.” The assertion of the fundamental role of niches in technological transitions, or the ‘niche hypothesis,’ is supported by two streams of research which arise from the tradition of evolutionary economics. This section provides a brief overview of this research, summarizes the basic assumptions, and discusses some difficulties with this theoretical basis. These findings are then used to construct a series of specific research questions that could help illuminate the over-arching issue presented in section 1.1. Chapter 2 presents a more complete discussion of the research literature.

1.3.1 The Quasi-Evolutionary Model

The approach of SNM introduced in section 1.2.4 emerged from a body of research involving the ‘quasi-evolutionary’ (QE) model of innovation which arises from the tradition of evolutionary economics. Evolutionary economics traces to Schumpeter (1939, 1942), but its application to innovation, which forms the basis of research on technological transitions, is based primarily on the work of Nelson and Winter (1977) and Dosi (1982, 1988).

Evolutionary economics is a behavioral approach that departs from static analysis and rational behavior and instead focuses on learning processes and adaptive behavior. Drawing from the analogy with biological evolution, innovation is viewed as a complex interaction among firms, the economy, and society that is characterized by three processes (Witt, 1992; Metcalfe, 1994; Malerba, 2002): variation, or the generation of novelty; selection through social, political, and market mechanisms; and the retention of successful variations through imitative and conservative behavior, routines, heuristics, and culturally learned behavioral patterns.

A main theme that arises from evolutionary theory is the cumulative nature of knowledge and learning which leads to path dependence. Thus, technological innovation proceeds along relatively well defined paths or trajectories. This is not to say that technical progress is deterministic, but rather that past innovation makes improvements in certain directions easier. In fact, the self-reinforcing nature of cumulative effects combined with the stochastic element of variation results in an inherent unpredictability of evolutionary systems.

However, according to the QE literature, evolutionary descriptions of technological progress lack a sufficient description of the sociological dimensions of technological regimes (Schot, 1992; Kemp, 1994). Therefore, a body of research has combined evolutionary theory with sociological and historical co-evolutionary models. In the resulting QE theory, a socio-technical regime is defined as "the whole complex of scientific knowledges, engineering practices, production process technologies, product characteristics, skills and procedures, and institutions and infrastructures that make up the totality of a technology" (Rip and Kemp, 1998). The existence of technological trajectories, then, is accounted for by the fact that existing technologies are embedded in broader technical systems, production practices and routines, consumption patterns, engineering and management belief systems, and cultural values – collectively referred to as rules (Kemp et al., 1998). In regime transitions, then, "both the technology and the system in which it is produced and used change through a process of co-evolution and mutual adaptation" (Kemp et al., 1998).

The key feature of the QE model is the "focus on the way variation and selection processes are partly independent and yet coupled" (Schot et al., 1994). Heterogeneous actors link variation and selection in three ways (Schot, 1992). First, entrepreneurs may anticipate selection and adjust the search heuristics used in innovation (*ex ante* selection). Second, institutional links between actors in the two processes create a 'technological nexus' which mediates between social and market requirements (selection) and technological opportunities and constraints (variation). Third, actors may create a niche to protect variation against harsh selection.

Raven (2005), based on Geels (2004), notes that the stability found in regimes results from the alignment of various rules, *e.g.* user preferences, automobile design, and government regulations. This stability is dynamic; rules do change, with or without erosion of stability. However, the alignment and embedded nature of rules that provides stability also results in resistance to change, and change is therefore incremental. Incremental change within a regime results in 'regime optimization' while radical innovation results in 'regime renewal' or transition (Elzen et al., 2003).

In the QE model, new technologies gain momentum when the dominant technology cannot adapt to a changing selection environment or when they are protected against a harsh selection environment through niche creation. Niches allow the technology to be developed in terms of user needs and to achieve wider use through learning processes and simultaneous adaptations in the selection environment. A technological regime shift is conceptualized as a process of niche proliferation or niche branching as the new technology is adapted to new niches defined by new geographical regions or new application domains.

1.3.2 The Multi-level Perspective

Geels (2002, 2004, 2005a, and 2005b) presents a second model of technical change based in the evolutionary tradition. In the multi-level perspective (MLP), socio-technical systems of innovation consist of a nested hierarchy of three heuristic levels: (1) the meso-level consists of socio-technical regimes, (2) the micro-level consists of technological niches; and (3) the macro-level is formed by the socio-technical landscape. The MLP stresses that system innovation occurs when simultaneous processes acting on multiple dimensions and levels "link up and reinforce each other" (Geels, 2005a).

The MLP conceptualizes the innovation process through three phases. In the first phase, novelties emerge within the context of the existing regime, in part due to landscape developments. Entrepreneurs experiment with designs in an attempt to determine users' requirements, and a variety of technical forms may compete with each other. In the second phase, the novelty is adopted in niche markets which serve as incubators. The innovation develops a technological trajectory and improves due to learning processes, while users explore new functionalities. This phase results in the articulation of user preferences and the development of a dominant design. The third phase is characterized by widespread diffusion of the innovation and competition with the established regime.

Two patterns of transition emerge from the interplay between the levels in the perspective. The 'technological substitution route' has a technology push character, where radical innovations emerge in niches and break into mainstream markets when they have sufficient momentum and when landscape developments put pressure on the existing regime. The 'wider transformation route' has a demand-pull character, where the existing regime becomes unstable due to persistent problems or landscape changes. Simultaneous changes on multiple dimensions stimulate actors to experiment with new technical options. After a prolonged period of experimentation and strategic maneuvering, a particular option may acquire dominance and lead to the creation of a new socio-technical regime. In both routes, three factors contribute to the wide diffusion of a new, radical, technology: 1) niche markets that act as incubators; 2) the penetration of the novelty into new application domains, leading to 'niche-accumulation;' and 3) circumstances at the regime or landscape level that create windows of opportunity.

1.3.3 Summary

Both the QE and MLP approaches integrate technological and sociological factors to describe the dynamic processes underlying technological transitions. Both approaches also stress the role of niches – distinct application domains that are small in scale and scope. The MLP describes a socio-technological regime transition as a process of 'niche accumulation,' while the QE model describes it as a process of 'niche branching' or 'niche proliferation.' Based on the QE literature, I identify five fundamental assertions of the role of niches in regime transitions:

1) Within the niche, manufacturers learn about users' needs and users learn about product performance through "learning-by-interacting."
2) Manufacturers improve the technology in terms of those needs and achieve lower costs through "learning-by-doing."
3) As a result of these learning processes, the new technology becomes embedded in a new rule set which reflects the adaptation of institutions. Both the technology and the system in which it is produced and used are altered in this co-evolutionary process

linking variation and selection which arises from the interaction of agents in the technological nexus.

4) Growth of niche markets (or niche branching or niche proliferation) allows manufacturers to reduce costs through economies of scale.
5) These changes facilitate diffusion into wider markets.

The MLP and Raven (2005) add regime dynamics and instability as a significant factor in the growth (or lack of growth) of niches. Instability may be due to three sources:

1) Success of niche technology.
2) External forces – changes at the landscape level – which influence the direction and rate of variation by altering regime agents' perceptions and expectations or altering the resources available for variation; or alter the selection environment by altering prices or consumers' preferences.
3) Internal regime forces – difficulties with the expanding scale or scope of the existing regime or achievement of technological or economic limits of existing technology – which lead regime actors to view the new technology as a solution.

1.3.4 Theoretical Issues

Ultimately, a theory of transitions must be based on a complete theory of innovation that includes existing regimes as well as emerging technologies. While early research on the quasi-evolutionary (QE) model began developing a robust theory of technical change, the literature has stressed the role of niches and recent work has shifted to a focus on SNM. I believe this leaves important aspects of the QE theory of innovation under-developed, leaving promising opportunities for policy unexplored.

First, the focus on the coupling between variation and selection leads to an exclusive emphasis on learning-by-doing (learning curve effects) and learning-by-interacting with users. This yields an incomplete description of the innovative process behind product and process improvements that are critical for wide diffusion.

Second, the technological nexus that links variation and selection is discussed as a key aspect of the process of embedding, yet remains an underdeveloped concept. Who are the relevant agents? How do new networks form and how are they sustained? Are elements of existing networks important? If so, how are they altered?

Third, the focus on a single dominant regime neglects interactions with other existing regimes which may help or hinder market growth (Raven, 2005). New technologies may address problems in multiple regimes or be based on fundamental innovations with applications that cut across regimes. This relates back to the first and second issues, in that the process of innovation is more complex than is reflected in QE theory. It involves multiple learning mechanisms and networks that tie the innovators to a relevant knowledge base. These networks may link seemingly unrelated industries, particularly in the early phases of a technology's life cycle.

Fourth, as Raven (2005) notes, the focus on the operation of niches in competition with a stable regime results in a simplistic description of niche-regime interaction. It neglects the dynamic processes within stable regimes (incremental change) that may either hinder or assist the diffusion of the new technology. In particular, it neglects regime response to internal and external forces of change which cause instability. Such challenges may be interpreted as opportunities for innovation, either for change within the regime or for radical new technologies that lead to a complete regime transition. This leads to the last issue.

The dichotomy of regime optimization (incremental change) versus renewal (radical change and transition) is overly simplistic for complex technologies. It places policy focus on transformation of the socio-technical regime in its entirety, which requires a complete alteration of both producer and user interpretations of technology and the related institutions. Meanwhile, it discounts innovations which make only part(s) of the existing regime obsolete but could still radically reduce environmental impacts. More significantly, it neglects intermediate or transitional technologies that may form a bridge to a new regime, such as hybrid and plug-in hybrid vehicles as a pathway to full electric or fuel cell vehicles. In such a transition, only part(s) of the regime may be transformed at a given time. This deficiency is interesting, because early works on QE theory highlighted the significance of compatibility issues: new technologies that may be easily embedded in the existing production, use, and institutional systems will diffuse more rapidly. According to Kemp (1994), "[o]nly in a few respects did radically new products constitute a radical break with the past, which suggests that the term 'radical' is somewhat misleading. Radical innovations often combined the new with the old (or even combined older technologies) and often rightly so because this helped the product to survive the initial harsh market selection and establish itself in the market place." Thus, transitions can occur as a more gradual unfolding and adaptation of the existing regime than the path described by 'regime renewal.'

1.3.5 Research Questions

These findings from the literature lead to the following research questions related to the assumptions underlying the niche hypothesis, the role of the technological nexus in niche dynamics and the diffusion to wider markets, and regime dynamics and stability:

- Is the role of niche markets as described in the literature historically accurate for the transition from horse-drawn to motorized vehicles?
    - Were learning-by-interacting with users and learning-by-doing the primary processes involved in improving the performance and cost of early motor vehicles?
    - Did growth in niche markets facilitate economies of scale?
    - Is niche branching, niche proliferation, or niche market growth an accurate description of the diffusion of the gasoline motor-vehicle into wider markets?
- Is the linkage between selection and variation described by the technological nexus an accurate account of the interactions that produced the technological and institutional adaptations that facilitated wider diffusion of the automobile?
    - Who were the agents involved and from what regimes were they drawn?
    - How did elements of the existing regime(s) influence their identities and roles? How were these identities and roles altered?
    - What formal institutions were developed and what contributed to their stability and longevity?
    - How were the variation and selection environments altered? Did this co-evolution occur within the context of niches?

    - Was the selection-variation linkage the most significant network contributing to technological and institutional adaptations?
- Were other factors related to niche-regime interaction important in facilitating the transition to motor vehicles?
    - In what ways were early motor vehicles compatible with the existing regime(s)? How did this compatibility evolve and eventually diverge?
    - Did regime instability contribute to the transition? If so, what were the causes of this instability?
    - What role did unexpected exogenous events play in existing and emerging regime instability and the adaptation of institutions?

1.4 Approach

To answer the research questions posed above, I review the historical transition from horse-drawn to gasoline-powered motor-vehicles within a co-evolutionary framework of innovation based in evolutionary and quasi-evolutionary theory. However, in order to account for the issues discussed in section 1.3.4, this framework is constructed using contributions from two additional bodies of research: the industrial lifecycle model (Abernathy, 1978; Abernathy et al., 1983; Abernathy and Utterback, 1978; Utterback and Abernathy, 1975; Anderson and Tushman, 1990) and mechanisms of learning in innovation. In this fashion, I attempt to take advantage of the explanatory power of some of the technical details found in these approaches while avoiding their shortcomings in describing the sociological dimensions of technical change. This framework is not proposed as a new model of innovation or a complete theory of technological transitions, but rather uses concepts and insights from several bodies of theory to provide the structure and vocabulary to trace the historic development of motor vehicle technology and its diffusion to wider markets. This section provides a brief overview of this analytic framework which is developed in Chapter 2.

1.4.1 Analytic Framework

First, I take from Malerba (2002) the perspective of the sectoral system of innovation and production which is defined as “a set of new and established products for specific uses and the set of agents carrying out market and non-market interactions for the creation, production and sale of those products.” This sector operates within a socio-technical system with a technological and social history, all of which is embedded in the natural environment. Therefore, the actions and interactions of the agents are shaped by institutions. Institutions in this context are analogous to QE’s rules and are defined to include any form of social construct or constraint on human interactions. Viewing the sectoral system within an institutional context then approaches the concept of a socio-technical regime, but the focus is clearly on the agents involved in innovation and production. In addition, the focus remains on system dynamics: “Over time, a sectoral system undergoes processes of change and transformation through the co-evolution of its various elements” (Malerba, 2002).

Second, I take from industrial lifecycle models the concept that technologies, firms, industries, and markets typically progress through a series of phases: emergent, transitional, specific, and senescent (Utterback and Abernathy, 1975; Abernathy and Utterback, 1978). Although the phases lie on a continuum with indistinct boundaries, the character of a sector at a given time can be related generally to this typology. The emergent phase is characterized by new technologies, niche markets, rapid innovation, and a focus on product improvement. The specific phase is characterized by dominant technologies, mass markets, incremental innovation, and a focus on process improvement. A dominant design is a (temporarily) stable description of the functional requirements of the technology and the product attributes used to fulfill them. The specific phase thus exhibits the stability of a socio-technical regime that results from the alignment of institutions (rules) and a mature technology.

I borrow from Anderson and Tushman (1990) the idea that innovation is cyclic; after an industry attains the specific phase, incremental innovation is punctuated by periods of 'ferment.' This is represented by movement from the specific phase into a new transitional phase. However, where Anderson and Tushman describe these periods as initiated by stochastic technological breakthroughs, I allow for the possibility that non-technical phenomena (social, political, economic) may create instability which presents opportunities for innovation or even forms the basis of the innovation of new production processes or usage norms. Thus, the phase typology gives structure to the dynamics of niche-regime interactions and the central role of stability. In addition, the cyclic aspect of this framework accounts for transition within an existing regime which is understood as a redefinition of the product's functional requirements and the realignment of the technology through the emergence of a new dominant design.

I then add a richer description of the learning mechanisms in innovation, relying primarily on Malerba (1992), but incorporating the work of various other researchers. This provides the vocabulary to explore the processes underlying product and production process improvements that facilitate the movement from niche to wider markets. I identify four learning mechanisms: internal research and development, learning-by-doing, learning-by-interacting, and spillovers. The boundaries of these classifications are not necessarily distinct since entrepreneurial activities often fall within more than one class. In addition, the ability of entrepreneurs to learn by each mechanism is influenced activities in other classes. All four sources are expected to contribute to the advance of a given technology at any point in time. However, industry lifecycle models predict that the relative importance of these mechanisms will depend, among other things, on the industry's and technology's phase of development.

This conceptual framework has the following advantages:

1) It places focus on the behavior of entrepreneurs (agents of change) which allows examination of learning processes within and among firms and is consistent with the fundamentally behavioral nature and micro-focus of evolutionary theory. According to Kemp (1994), "the relationship between firm behaviour and technological regime shifts is a relatively underresearched area."
2) It allows further exploration of the network of agents involved in variation (the entrepreneurs) and selection (consumers, policy makers, and society at large) and the two-way linkage between the two processes through the technological nexus. According to Schot (1992), "The activities of a technological nexus result in a learning process, which takes place both in the firm, a place where an important part of the variation process is located, and in its environment."
3) It allows exploration of the full range of learning processes underlying innovation and transitions to determine whether learning-by-doing and learning-by-interacting with users are indeed the appropriate focus for policy.

4) It allows exploration of the dynamic processes of innovation and transition *within* a regime in response to internally and externally generated challenges. This in turn determines how regime instability will influence the growth of niches.

1.4.2 Data

As developed in this research, innovation is a synthetic process whereby entrepreneurs combine existing artifacts and threads of thought with personal inspiration and insight to create a new whole that is more than the sum of its individual and, to some extent, preexisting components (Rycroft and Kash, 1999). This dissertation applies that very process to a study of the history of innovation in the automotive industry in order to develop insights on the sources of technological progress and transitions to new technological regimes. It is largely an historical study that presents no new data, but rather presents a new organization and re-analysis of the work of many very able historians and analysts. In defense of this approach, I borrow the words of Flink (1970, p. 9): "…in the interest of economy of research and in the belief that the writing of history must be a cumulative process based upon considerable trust in one's predecessors, I have not hesitated to deal synthetically, in the main, with those aspects of the history of the automobile industry that have already been examined in detail by others and to eschew firsthand reexamination of the various industry sources upon which previous studies have been primarily based." By using a variety of such sources, I hope to minimize selection and interpretation biases.

1.5 Significance of the Dissertation

Addressing climate change will require a transition to alternative transportation technologies. Constructing effective policies to facilitate such socio-technical transitions requires an understanding of the underlying mechanisms and processes of innovation and the mutual adaptation of technology and institutions. Therefore, this research adds to the policy debate on stimulating innovation and diffusion of new technologies to address climate change.

Prior analysis of the transition to motor vehicles has followed two general approaches. The first approach describes transition as a process of logistic substitution where old and new technologies compete for market share beginning with the initial introduction of the new innovation (Grübler, 1998); Nakićenović, 1986). While the data supports a logistic substitution, this aggregate representation masks the underlying process of change. The second approach, found in the MLP literature, attempts to remedy this issue by examining the dynamics between niches, socio-technical regimes, and the wider socio-technical landscape (Geels, 2005b). However, by focusing on interactions between the regime and the wider socio-technical context, this approach still has an aggregate representation of the components of sectoral systems of innovation and production and is ill-structured to examine processes within that system. In other words, innovation processes found within and between firms in the regime remain inside the black box.

This research takes a unique approach and examines socio-technical change from the perspective of the entrepreneur. Thus, this research adds to the literature on socio-technical transitions by examining specifically the role of entrepreneurial behavior and learning processes within and among firms in the historical co-evolution of technologies and institutions during the transition to automobiles. As a result of this unique approach, I show that the conception of innovation in theories of socio-technical transitions is overly simplistic and would benefit from a richer description of mechanisms of learning such as the one used in this research. I also show that firm-specific histories and rules constraining interaction shaped the strategies and adaptive capabilities of firms. These strategies, and thus firm-specific rules, had a large impact on the development and diffusion of the technology. This finding supports the use of a behavioral approach with a micro-focus in the analysis of innovation and socio-technical change.

1.6 Organization of the Dissertation

This dissertation is organized as follows. Chapter 2 reviews the analytical literature on evolutionary approaches to innovation and socio-technical transitions. This section identifies the fundamental assumptions and assertions of the literature as well as some shortcomings. It then reviews several strands of innovation research that may be used to illuminate and improve on these shortcomings. In the final section of this chapter I synthesize these approaches to formulate a framework for analysis. Chapters 3 through 5 then review the historic development of gasoline motor vehicles and places this history within the framework developed in chapter 2 to analyze the development of the personal transportation regime based on this technology. These chapters are roughly chronological but overlapping. Chapter 6 provides a summary of the findings and discusses the implications for theories of socio-technical transitions and policies intended to foster technological innovation and transitions. I conclude this chapter by outlining areas for future research.

2.0 Innovation and Technological Transitions

…the scope for policy is not to optimise with respect to some objective function… but rather to stimulate the introduction and spread of improvements in technology. At the core of this approach are complexity, cognitive limitations, and the role of organizations…

-- Metcalfe, 1994

2.1 Introduction

Although the role of innovation in the economy has long been recognized, the underlying mechanisms of learning and progress are less understood, particularly as they relate to transitions. The study of innovation is complicated by the fact that it does not fall neatly into a single academic discipline. According to Smits (2002), "the discipline of innovation studies is not a firmly integrated theoretical bastion," but rather "can be typified as an evolving (inter)discipline that finds itself at the crossroads of sociological and historic scientific and technological research, economic innovation studies and policy studies." The literature on innovation is therefore vast with wide variation in goals, scope, and approach.

The earliest research in technical progress considered the role of innovation in economic progress and international competitiveness. This research therefore focused on macroeconomic impacts and treated the innovative process as a "black box." Other areas of innovation research have been concerned with economic progress in developing countries, innovation market failures, firm decision-making and strategy, firm survival, industrial and institutional change, explaining the success or failure of specific technologies, and most recently, advancing technologies for mitigating climate change. As these bodies of research have grown, frameworks for studying innovation have evolved from static, neoclassical economic models to complex system dynamic models.

The evolutionary approach to innovation is a subclass of dynamic system theories which is distinguished by a behavioral theory of the firm and a focus on learning and adaptation (Metcalfe, 1994).

Because a transition is by definition dynamic – a process of moving from one state to another – the study of technological transitions has built on these later models. This dissertation is concerned primarily with the assertions of two streams of research on technological transitions that have emerged from the evolutionary approach: the quasi-evolutionary model (QEM) and the multi-level perspective (MLP).

This chapter presents a selective review of the literature on innovation and technological transitions. In sections 2.2 and 2.3 I review the origins and significant contributions of evolutionary systems theories of innovation as it relates to understanding innovative processes. In section 2.4 I review the literature on the QEM and the MLP. This review presents the assertions of these models in regard to the role of niches in technological transitions and discusses a number of theoretical difficulties. In section 2.5 I review two additional bodies of innovation research – the industry life-cycle model and mechanisms of learning in innovation – that I then use in conjunction with the QEM and MLP to construct an analytical framework that accounts for these difficulties. This framework is summarized in section 2.6 and serves as a guide for analysis of the history of the transition to automobiles.

2.2 From Neoclassical Economics to System Theories

Systems theories of innovation arose largely due to dissatisfaction with the neoclassical economic view in handling two separate classes of research: management of research and development in industrialized countries and economic growth in developing countries (Sahal, 1981). This dissatisfaction rests on four main criticisms of neoclassical economics: the focus on static analysis; the focus on aggregate phenomena; the treatment of innovation as an exogenous process; and the simplistic characterization of firm behavior based on rational agents.

Section 2.2.1 describes the neoclassical conception of innovation and introduces the shortcomings of this approach in studying the process of technological change. Section 2.2.2 presents the earliest dynamic model of innovation and the progression of understanding that led to the development of complex systems models of innovation, which are discussed in section 2.2.3.

2.2.1 Neoclassical Conception of Innovation

In neoclassical economics, technology is conceptualized by means of the production function, which represents the relationship between combinations of inputs of production and the resulting output. Theoretically, the production function specifies all technically feasible modes of production under existing knowledge (Sahal, 1981). The production function is typically illustrated by a convex isoquant, which represents the various combinations of inputs required to produce a fixed amount of output. Technological advance is then defined as an increase in productivity such that fewer inputs are required to produce the same quantity of output, thus shifting the production function toward the origin.

Schumpeter (1939) defines innovation as the "setting up of a new production function." Solow (1957) defines technical change as a "…shorthand expression for *any kind of shift* in the production function. Thus slowdowns, speedups, improvements in the education of the labor force, and all sorts of things will appear as 'technical change.'" In other words, innovation allows the economy to produce at a given level of output using fewer resources (*e.g.*, capital and labor).[4]

While developed theoretically at the level of the individual firm, research applying the production function typically considers the aggregate function of an entire industry or nation. Although this concept has provided insight for macroeconomic policy, it leads to a number of theoretical and practical difficulties, the full details of which are beyond the scope of this discussion (Sahal, 1981). What is important to note is that a static analysis of the impact of technical change on the aggregate output of an entire sector or economy provides little insight into the internal drivers and dynamics of that change, and is therefore of little use to those seeking to manage or influence innovation (Grübler, 1998). For this purpose, we must examine the innovative *process,* i.e. the activities involved in *setting up* a new production function.

Unfortunately, neoclassical economic theory has little to say about the process of moving from one production function to another. At worst, innovation is treated as an entirely exogenous process and a new production function is simply imposed. At best, innovation is treated as a response to an exogenous change in data, namely the price of labor or capital. Attempts to explain technical change within neoclassical growth models rely on the familiar assumptions of rational, maximizing behavior and static equilibrium. The resulting theory of induced innovation introduces the innovation possibilities function (IPF) which represents the envelope of all alternative production functions attainable with available innovative resources (time and skill) (Ahmad, 1966). In

[4] More recently, researchers of economic development have shifted focus from the mechanics of the production function to how innovation creates a "learning society," that in turn indirectly improves an economy's productivity. In this view, the most important product of the innovative process is expansion of the knowledge base (Conceiçãoa et al., 2003).

response to exogenous changes in factor prices, firms are assumed to choose a level of investment in research that will maximize expected profits, possibly within the constraints of a research budget.

Researchers have pointed out difficulties with the theory of induced innovation, including theoretical criticisms of the IPF, the single focus on productivity enhancements through reductions in capital and labor inputs, and the assumption of profit maximization as the basis of entrepreneurial effort. While a complete review of these shortcomings is beyond the scope of this discussion, it is worthwhile to note two points. First, the IPF is independent of history; the returns to research effort in a given time period are independent of what research was pursued and what production techniques were used in the past (Nordhaus, 1973). Second, the assumption of profit maximizing behavior, even when uncertainty and risk are considered, is overly simplistic. In reality, research is often under-taken by individuals and organizations not motivated by profit. Further, entrepreneurs face cognitive and resource limitations when selecting from an enormous range of possible research projects with uncertain prospects. According to Nelson and Winter (1977), the assumption of profit maximizing behavior "suggests an unrealistic degree of inevitability and correctness in the choices made, represses the fact that interpersonal and interorganizational differences in judgment and perception matter a lot, and that it is not at all clear ex-ante, except perhaps to God, what is the right think to do."

In addition, shifting focus from static analysis to dynamic processes alone is not sufficient. A macro-level conception of technological progress based on aggregate phenomena and macro-level innovation policies are ill-suited to address research questions such as those studied here. Research has shown that rates of innovation vary greatly among economic sectors and industries, and policies must therefore be designed to influence the sector of interest. Effective policies must consider the nature of current and appropriate institutional structures, which in turn depends on the underlying technologies themselves (Nelson and Winter, 1977; Nelson and Langlois, 1983).

Further, neoclassical theory of innovation focuses on productivity measures, and essentially assumes that product characteristics are fixed. However, in manufacturing firms as much as 90% of research and development (R&D) effort and 75% of R&D results relate to changes in product characteristics or development of new products. In addition, a number of econometric studies have indicated that a large proportion of productivity growth is attributable to changes in the *quality* of inputs (Sahal, 1981).

2.2.2 Dynamic Models of Innovation

The need for a more satisfying conceptualization of the dynamic process of innovation has led to the development of a large number of system dynamic models. Schumpeter (1939, 1942) was perhaps the first economist to emphasize the role of the dynamic process of innovation in the economy, a process that "incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one." While most economists stressed the price equalizing effect of competition for homogenous goods, Schumpeter (1942) instead asserted that the process of "Creative Destruction" was "the essential fact about capitalism" and "the powerful lever that in the long run expands output and brings down prices."

Schumpeter described three distinct processes of innovative activity: invention, innovation, and diffusion. *Invention* creates new ideas that may be scientific, technological, or even organizational. While the creative process may be stimulated by economic and social conditions, the product of invention is only a potential with no realized economic or social value. It is the entrepreneur's role[5], through *innovation*, to create economic value, "to reform or revolutionize the pattern of production by exploiting an invention or, more generally, an untried technological possibility for producing a new commodity or producing an old one in a new way, by opening up a new source of supply of material or a new outlet for products, by reorganizing an industry and so on" (Schumpeter, 1939). From this definition, innovations can by classified as either product innovations (producing a new commodity) or process innovations (producing an old commodity in a new way).[6] The entrepreneur temporarily occupies a monopolistic position in the market and realizes economic profits. In the final process of *diffusion*, imitators bring competition to the market, bidding down the price, eroding the entrepreneurs' profits and moving the market toward economic equilibrium (Schumpeter, 1939 and 1942; Langlois, 2002; Godin, 2005).

Consistent with Schumpeter's description, one of the first frameworks used to study the innovative process followed a sequential linear model that traced the flow of information and the time progression from scientific research and invention, to technological innovation and development, to production and market diffusion. The linear model underpins the "science-driven" or "technology-push" theory of innovation that holds that invention creates the possibility for new products or processes, and thus drives innovation. It is not clear when the framework was first articulated, though some

[5] In this discussion, the term 'entrepreneur' refers to all innovating units, including individuals, teams, or entire firms.

[6] The distinction between process and product innovations depends on the unit of analysis. One entrepreneur's product innovation may be applied in the manufacture of an existing product. Thus the new product becomes a process innovation for the adopting entrepreneur (Abernathy and Utterback, 1978). Often the two entrepreneurial units are the same, as when new process equipment is developed within the user firm or industry (Pavitt, 1984).

attribute the origin of linear model to Schumpeter's "trilogy" of invention, innovation, and diffusion. However, Schumpeter himself clearly refuted linear dependence and causality by asserting that innovation "is possible without anything we should identify as invention and invention does not necessarily induce innovation" (Schumpeter, 1939). According to Godin (2005), "[t]he precise source of the linear model remains nebulous, having never been documented." Godin suggests that it developed in stages over a period of decades, and "is a theoretical construction of industrialists, consultants and business schools, seconded by economists."

The linear model has been heavily criticized beginning as early as the 1960s (Godin, 2005). A main criticism of the model is its independence from the economic and social environment. According to Dosi (1982), "extreme forms of technology-push approaches, allowing for a one-way causal determination (from science to technology to the economy) fail to take into account the intuitive importance of economic factors in shaping the direction of technical change." Economic, institutional, and social factors all serve to focus innovative activities on certain technological problems and to select certain trajectories to the exclusion of others (Rosenberg, 1976; Dosi, 1982).

Griliches (1957) and Schmookler (1962, 1966) found empirical evidence of the importance of market demand in stimulating inventive activities. This research has been used to support the "demand-pull" theory of innovation, which holds that unmet consumer needs drive investment in invention and innovation – 'necessity is the mother of invention.' However, demand-pull theory assumes the "*possibility* of *knowing a priori*… the *direction* in which the market is 'pulling' the inventive activity" (Dosi, 1982; emphasis in original). This requires that consumers' *needs* are effectively articulated as *demands* in the market via price and quantity signals, and that they are interpreted by a responsive industry (Mowery and Rosenberg, 1979; Freeman, 1996; Dosi, 1982). There is little doubt that the market plays an important role in stimulating innovative activity by presenting problems, opportunities, and targets that focus the search process. Freeman (1982) lists market forecasting among the factors that distinguish successful attempts at innovation. However, researchers find insufficient evidence to conclude that consumer needs articulated through market signals are the *primary* drivers of innovation (Mowery and Rosenberg, 1979; Dosi, 1982). Indeed, the entrepreneurs of major innovations of the last century, including the telephone, computer, and laser, often underestimated the importance of and market for their products (Rosenberg, 1996). The ultimate range of application for these products depended on unforeseeable advances in the technologies, the development of complimentary technologies, and co-evolutionary processes. Market demand, then, is more appropriately considered a *necessary*, but not *sufficient*, condition for successful innovation.

Further, Dosi points out the “general weakness of market mechanisms” in selecting the direction of technological progress, especially in the early phase of an industry’s development, and the inability of demand-pull theories in explaining why and when certain innovations occur (Dosi, 1982). Freeman (1994) asserts that the empirical evidence of innovations characterized as demand led were “actually minor innovations along established trajectories.” Meanwhile technology push patterns have been shown to be more characteristic of the early stages of innovation in numerous products (Freeman, 1994).

In 1991, Rosenberg claimed, "Everyone knows that the linear model of innovation is dead." Yet according to Godin (2005), the linear model survives, "despite regular criticisms… due to statistics. Having become entrenched…, the linear model functioned as a 'social fact'. Rival models, because of their lack of statistical foundations, could not easily become substitutes."

2.2.3 Systems View of Innovation

In reality, innovation is a complex dynamic process that is neither independent of the economy nor purely demand driven. Indeed, the history of individual innovations often is characterized by an iterative and co-evolutionary socio-cultural process whereby firms respond to demand, supply, and institutional forces (Mowery and Rosenberg, 1979). Innovation, then, is more accurately viewed as a complex interaction among firms, the economy, and society that links "*potential* users with new developments in science and technology" (Freeman, 1996, emphasis added). The systems approach to innovation is therefore founded on the idea that innovation cannot be understood solely by examining independent decision-making at the level of the firm. Rather, researchers must consider the interaction of firms with their environment, which is taken to include not only customers and suppliers, but also the broader factors shaping firm behavior. This includes the social and cultural context; institutional and organization frameworks; and infrastructures (Smith, 2000).

By definition, a system is composed of a group of independent yet functionally interrelated elements that together comprise a unified whole. These interdependent elements or subsystems are linked by a network of feedbacks. A system's characteristics then derive from the functional properties of the whole, which are more than the simple aggregation of its subsystems or parts. In the systems view, “technology is best understood in terms of its functional attributes” (Sahal, 1981). Among its advantages, the systems view clearly maintains a focus on dynamics in terms of change in functional performance. In addition, it emphasizes the continuous and cumulative nature of innovation, as improvements in functional performance build upon prior achievements.

Complex systems models highlight the systemic nature of innovation and are characterized by multidirectional links among the sources and stages of technical change at any point in time. Such a model blurs the lines between types and phases of research and innovation, and instead stresses the exchange and exploitation of new knowledge by researchers, engineers, designers, manufacturers, suppliers, and users. While there is wide variation in the focus and goals of system dynamics studies, these models of innovation share the following important characteristics (adapted from Freeman, 1996; Soete and Arundel, 1993). These characteristics are discussed further in section 2.3.2 in relation to evolutionary systems.

1) Multidirectional links at the same point in time between the stages of technical change;
2) Complex interactions of firms within the industry, with supporting industries, and with seemingly unrelated industries;
3) Complex interactions with the socio-economic environment;
4) Central role for knowledge and assimilation through learning;
5) Cumulative and self-reinforcing processes;
6) Increasing complexity over time in technology, industry, and innovative process.

There has been a recent surge in systems approaches within the innovation literature, with authors describing our society as increasingly complex and nonlinear (Rycroft and Kash, 1999) with a trend toward “a system with strongly linked fuzzy components" Smits (2002). This thinking, however, is not new. Rosenberg (1963) referred to the industrialization of the 19th century as "involving, not only growing specialization, but also growing complexity and differentiation." Abernathy and Utterback (1978) recognized the importance of understanding the production process as "a system of linked productive units" and noted that "[m]ajor change at one level works its way up and down the chain…"

Systems themes can be found in studies of the role of innovation in economic progress, how firms manage innovation and survive industry transitions, and the forces behind sector transitions. In considering the challenges to firms and policymakers managing innovation, Smits (2002) asserts that "[i]n the network society…innovating in chains, networks and systems becomes more and more important." Conceiçãoa et al. (2003) consider the role of innovation for promoting growth in developing countries, and assert that "…innovation should be understood as a broad social and economic activity: it should transcend any specific technology, even if revolutionary..." Dijkema et al. (2006) assert that fostering environmental sustainability requires a shift in our "techno-economic paradigm." Rather than focusing on corporations, production plants, or individual technologies, they claim that sustainability research must consider innovation in the entire lifespan of a product; the entire production network; and the stakeholders and decision processes, including business organization, public institutions, and policy and regulatory frameworks.

Most relevant, however, is the research in sectoral systems of innovation. Ultimately, technologies provide societal functions, such as transportation. The need for a social function is satisfied, not by a single technology or industry, but rather by a complex system of social, technical, and institutional elements, including supply networks; related industries; physical infrastructures and maintenance industries; government, industry, and research institutions; markets; and society (Unruh, 2000; Geels, 2005b). How such a system is defined – its elements, its boundaries, and its linkages and interactions with the external environment – varies greatly among researchers. Unruh (2000) refers to a techno-institutional complex (TIC); Nelson and Winter (1977) refer to a technological regime; Geels (2002) defines a socio-technical regime embedded in a socio-technical landscape. For the purpose of discussing evolutionary theory in general, I will refer to the socio-technical system and its environment. The socio-technical system is composed of heterogeneous elements of three general types: physical artifacts; human actors and organizations; and a complex set of constraints on human interactions with each other and with the physical elements in the system (Geels, 2004). Various authors refer to these constraints as institutions or rules, which include formal constraints such as legislation, economic rules, and legal contracts and informal constraints such as social conventions, moral codes, and shared knowledge and beliefs.

2.3 Evolutionary Systems Theory of Innovation

The evolutionary approach to innovation represents a special class of complex systems theories and adds an additional and distinctive feature: "its adoption of a behavioural theory of the firm and its focus upon learning processes and adaptive behavior" (Metcalfe, 1994). Thus, the shift away from static analysis to dynamic processes – from 'being' to 'becoming' – and from a macro to a micro focus is captured in evolutionary approaches. There is wide variation in the scope and methods found in evolutionary economics, but all focus on the *process* of economic change which is endogenously generated (Witt, 1992; Metcalfe, 1995). In fact, endogeneity is a central hypothesis, and Witt (1992) defines evolution as the *self-transformation* of a system over time. According to Metcalfe (1995), "evolution means two things: the gradual unfolding of phenomena in a cumulative and thus path-dependent way; and quite separately, a dynamics of system behaviour which creates change and emerging structure from variety in behaviour."

Evolutionary economics discards the global objective function, well defined choice sets, maximizing behavior, rational choice, and the global production function. Instead, it "places emphasis on cognitive dimensions such as beliefs, objectives and expectation, in turn affected by previous learning and experience and by the environment in which agents act" (Malerba, 2002). Drawing from Witt (1992), Dosi (1997), Ruttan (1997), Metcalfe (1995), and Malerba (2002), the methodological building blocks central to an evolutionary approach can be summarized as:

1) The emphasis on dynamics: becoming, rather than being.
2) Behavioral, or micro, focus: what agents do, and why.
3) Imperfect understanding and bounded rationality, broadly defined.
4) Heterogeneity of agent knowledge, perception, behavior and learning.
5) Continuous creation of novelty or variety in technologies, products, processes, firms and organizations due to the local search for innovation.
6) Interactions of agents that collectively act as selection mechanisms, reducing variety and generating system inertia and continuity, including:
    a) imitation of the practices of other agents;
    b) conservative behavior;
    c) culturally learned interpretation patterns, world views, and paradigms;

d) satisficing economic behavior;

e) market selection.

7) Aggregate phenomena are emergent properties of non-equilibrium interactions and heterogenous learning, and have a 'metastable nature', i.e. they persist longer than the processes that generate them, but disappear eventually.

2.3.1 Basic Evolutionary Concepts

Evolutionary approaches entail three processes based on the biological analogy: variation, selection, and retention. These basic evolutionary concepts are summarized in this section. A more complete discussion is included in section 2.3.2 elaborating the characteristics of evolutionary systems models.

The generation of variety in biological systems is a process of mutation caused by random errors in replication. In evolutionary economics, the generation of variety or novelty is the "outcome of human creativity and of the discovery of new possibilities for action" (Witt, 1992). Some researchers have interpreted the biological analogy rather literally, asserting that "[v]ariation is driven by stochastic technological breakthroughs" (Anderson and Tushman, 1990; Tushman and Rosenkopf, 1992) that result from research by "trial and error" (Raven, 2005).[7] This interpretation is unnecessarily restrictive and inconsistent with the main goal of evolutionary theory to explain endogenous processes. In one of the foundational evolutionary works, Nelson and Winter (1977) conceptualize innovation as "purposive, but inherently stochastic." Metcalfe (1994) asserts that "it is central to the evolutionary approach that the creation of diversity is neither blind nor

[7] Tushman and Rosenkopf (1992) interpret variation as a blind process driven by chance events (stochastic technological breakthroughs), while technological advance resulting from the "puzzle-solving actions of many organizations learning-by-doing" is treated as a retention process. In this research, I interpret variation to include all processes that result in new or enhanced products or processes, regardless of whether the advance is incremental or 'radical.'

unbounded." Instead, innovative behavior is "intentional and guided" (Metcalfe, 1994) which results in trajectories of development.

Innovation – the accumulation of knowledge and its application to new domains – is the result of a fundamental characteristic of the restless capitalist economy: the search for profit opportunities and the pursuit of competitive advantage. Witt (1992) proposes two additional motivations for agents to experiment with new technologies: preference for novelty and frustration with the status quo. Preference for novelty explains a basic rate of innovation, while frustration explains innovation in response to crises. It is notable that rational choice models cannot explain such emotional motivations as frustration. Witt, therefore, adopts an adaptive aspiration and satisficing model for decision-making.

In biology, selection refers to the Darwinian concept of 'survival of the fittest.' In evolutionary economics, variations in technologies (products and processes) and organizations are selected through social, political, and market mechanisms. Most evolutionary research focuses on market selection and stresses social and political selection factors only for 'non-market' sectors such as public education and defense.[8] Market selection entails the introduction of technologies and entry of new firms; the elimination of unprofitable technologies and firms; and changes in the relative importance or market position of surviving technologies and firms. Selection reduces variety, as "competition consumes its own fuel" (Metcalfe, 1994). Because survival or 'fitness' of a firm or technology is judged relative to the average of its rivals (the

---

[8] See, for example, Nelson and Winter (1977) and Dosi (1997).

population), evolutionary systems do not guarantee optimal outcomes; any optimum that is achieved is only local (Witt, 1992; Metcalfe, 1994).

However, markets represent only one mode of selection. According to Metcalfe (1995), "[a]ny framework in which agents interact in order to choose between competing patterns of behaviour has selective properties." Of particular note are selection environments internal to innovating organizations. Agents learn from experience, anticipate future market selection, and use this information to choose from a vast array of alternative research projects and approaches. Thus technological variation and selection are linked.

In biological systems, the final evolutionary process of retention is accomplished through the replication of successful variations through reproduction. In economic systems, retention occurs through imitative and conservative behavior; the development of routines; culturally learned interpretation patterns; world views and paradigms. Most evolutionary theories of innovation treat retention as a cognitive process within the minds of researchers and engineers (Raven, 2005). For example, Nelson and Winter (1977) describe a broadly defined 'technological regime' as a cognitive concept "relating to technicians' beliefs about what is feasible or at least worth attempting." Co-evolutionary processes result in a core technological framework or paradigm which is shared by the entire community of technological and economic actors (Kemp, 1994) and which guides innovation. In this vein, the majority of evolutionary research focuses on routines, rules of thumb, and search heuristics within firms which lead to the establishment and maintenance of technological trajectories. This rather limited view of retention is a major criticism addressed by the QEM and MLP approaches.

As discussed in the following sections on learning and cumulative effects, evolutionary systems are inherently path dependent and unpredictable, which is at odds with deterministic tools used for policy analysis. "…evolutionary economic processes are essentially open ended and unpredictable." (Metcalfe, 1995, p. 28) According to Ruttan (1997), the evolutionary approach has not "become a productive source of empirical research," possibly because simulation methodology "lends itself to the easy proliferation of plausible results. At present, the evolutionary approach must be regarded as a 'point of view' rather than as a theory." Dosi (1997), however, disagrees, asserting that "an interpretation of technological dynamics which significantly relaxes the commitments to equilibrium, rationality and inter-agent homogeneity is straightforwardly born by the current evidence, and is also beginning to generate formalized theoretical tales." Nonetheless, taking an evolutionary perspective requires that one abandon the ideas of determinacy and prediction because "[t]heories of evolution, whether biological or social, are not predictive ones – they are retrospective and historical" (Holling,1994).

2.3.2 Characteristics of Evolutionary Systems Models

The characteristics of systems models introduced in section 2.2.3 have implications unique to evolutionary models of innovation. A discussion of these characteristics therefore illuminates the distinctive features of evolutionary theory and the development of technological paradigms and trajectories.

2.3.2.1 Multi-Directional Links Among Phases

Since the stages of technical change are connected by multidirectional links at a given point in time, the classic phases of development – invention or basic research; innovation or applied research and product development; and diffusion – are not clearly defined and cannot be isolated for analysis. Notably, feedback from the diffusion process (selection) is one of the critical elements in technology development, and therefore, innovation and diffusion are "inseparable and mutually reinforcing" (Metcalfe, 1994; see also Metcalfe, 1995).

Nelson and Winter (1977) point out that many uncertainties with a new innovation "cannot be resolved until an innovation actually has been tried in practice." When first introduced, new technologies are often poorly developed in terms of performance (Kemp, 1992) and sometimes even functional definition, such that determining users' needs and improving the technology are critical to diffusion. This is supported by Freeman's (1982) finding that successful innovations are frequently characterized by greater interaction with and attention to feedback from users.

In addition, information from the selection environment feeds back into the variation process by influencing the kinds of research that firms believe are likely to be profitable in the future. In this process of *ex ante* selection, entrepreneurs attempt to anticipate future market selection and use such forecasts as a factor in selecting from a large number of possible research choices. According to Dosi (1982), the "economic and social environment affects technological development in two ways, first selecting the 'direction of mutation' (i.e. selecting the technological paradigm) and then selecting among mutations, in a more darwinian manner (i.e. the *ex post* selection among 'schumpeterian' trials and errors)."

However, recognizing the importance of feedbacks in the innovative process and the indistinct nature of the stages of research and development does not mean that phase characterizations are not useful. In section 2.5.1 I review the industry life-cycle model which describes how industries, products, and markets mature over time. This model provides a guide for studying the dynamics of established technological regimes and for evaluating the changing relative role of various agents and learning mechanisms as a technology and industry evolve. This evaluation should allow the identification of opportunities to introduce alternative technologies.

2.3.2.2 Complex Firm and Industry Interactions

The unit of analysis used in innovation studies ranges from an individual assembly line, to a complete production department within a firm, to an entire firm, entire industry, entire sector, or entire national system. A central theme of the evolutionary systems view is that the "innovation activities of firms involve a wider range of other institutions supplying the knowledge and skills which underpin the efforts of firms" (Metcalfe, 1994). Systems models recognize that innovation and production are influenced by complex interactions within and between firms in the industry, with supporting industries, and with seemingly unrelated industries. Malerba (2002) describes innovation as a 'collective process.' Therefore, systems models stress the importance of networks, both formal and informal, between firms and all institutions producing knowledge, skills and artifacts within the technology system, as well as user-supplier interactions. According to Freeman (1996), "[f]eedback loops and interdependencies can be important at every stage, so that networking and cooperation between research institutions and firms should be continuously encouraged." As a result, we cannot study a single product, firm, or industry in isolation.

According to Abernathy and Utterback (1978), examining a single product type and firm is "of little use in understanding innovation. Technological change causes these terms to change their meaning, and the very shape of the production process is altered." In addition, management of operations and management of innovation within a firm or industry are intrinsically linked and co-evolve with product and process changes. Whipp and Clark (1986) define strategic innovation as "changes in technology and forms of work organization at all levels, which includes boards of directors or the various interfaces between a company and its suppliers or potential customers." In their analysis of a British auto manufacturer's innovation project in the 1970s, the authors include the overall organization of the firm, and the processes and politics involved in order to analyze the firm's *capacity to undertake design and innovation*.

The unit of analysis for a technology must include all the characteristics of the manufacturing unit that produced it, which may include several firms, often not even in the same industry, depending on the level of diversification or fragmentation. This idea traces back to Rosenberg (1963), who recognized the interdependence of apparently unassociated industries in 19^{th} century manufacturing. "For many analytical purposes it is necessary to group firms together on the basis of some features of the commodity as a final product; but we cannot properly appraise important aspects of technological developments… until we give up the Marshallian concept of an industry as the focal point of our attention and analysis. These developments may be understood more effectively in terms of certain functional processes which cut entirely across industrial lines..."

Rosenberg points out that products from apparently disparate industries can share fundamental production technologies and processes, and therefore suffer from common problems and limiting conditions. This 'technological convergence' creates synergies that are of critical importance for the innovation and diffusion of production technologies and techniques. Innovations that are developed as a solution for a narrow range of problems in a single industry can be generalized for wide ranging applications. Raven (2005) finds that applications of new technology in multiple regimes help explain the development of niches for bioenergy technologies. Mowery and Rosenberg (1998, p. 5) assert that the inter-sectoral flow of new technologies is a "fundamental characteristic of 20th-century innovation in the U.S. economy."

Further, while emerging and growing industries are characterized by high rates of internal innovation, mature industries may experience significant productivity growth resulting from external sources of innovation. Under these conditions, the degree of vertical integration in an industry has significant implications for a firm's ability to appropriate new production technologies.[9]

It is significant to note that the interactions among firms and industries are multidirectional. According to Rosenberg, advances in machine tools used for the manufacture of firearms, sewing machines, and bicycles, made possible the rise of the automobile industry. As the industry developed, solutions to problems in automobile design were incorporated into and improved the performance of the specialized machine tools used in their manufacture. These improvements were then generalized into and again revolutionized the machine tool industry.

---

[9] According to Rosenberg, in the late 1900s, increasing vertical disintegration of individual industries was accompanied by technological convergence in a larger group of industries. He attributes the high degree of machinery specialization to the *combination* of vertical disintegration and technological convergence. Stigler (1951) theorized that vertical disintegration is generally a feature of growing industries, while increasing vertical integration is typical of declining industries. At the birth of a new industry, firms must produce their own input materials, design and possibly manufacture their own specialized equipment, and recruit skilled labor. When the young industry attains a certain size, it becomes profitable for other firms to assume and become specialists in these tasks. As the industry declines, it can no longer support independent firms for these functions and surviving firms must re-assume them.

2.3.2.3 Complex Interactions with the Environment

A growing body of innovation research expands on the role of the environment in which innovation occurs, and the constant interplay among technology, society, and institutions. According to Smith (2000), systems approaches to innovation are founded on the idea that innovation cannot be understood purely in terms of independent decision-making at the level of the individual, firm, or industry. Rather, innovation involves complex interactions between these entities and their environment. Therefore, large technological systems that are a product of decades of innovative efforts cannot be understood by analyzing the existing set of technological artifacts, but must be viewed "as complex systems of technologies embedded in a powerful conditioning social context of public and private institutions" (Unruh, 2000).

Evolutionary systems theories of innovation follow Schumpeter's definition of innovation as the creation of value by exploiting change, but expand the notion of *what* change is relevant. Innovation occurs when entrepreneurs exploit a new invention or changes in any element of the socio-technical system or its environment. In addition to scientific and technical knowledge, such change may include relative prices, demographic makeup, consumer tastes and practices, cultural meanings, physical and institutional infrastructures, regulatory frameworks, and even geopolitics (Dijkema et al., 2006). For example, successful energy technologies and policies simultaneously must satisfy multiple and often conflicting economic, environmental, and security goals. While technologically feasible alternatives to conventional oil currently do exist, their development and diffusion is impeded by economic (too costly), social (too inconvenient, too disruptive, too dangerous, *etc.*), and environmental (too polluting) considerations

(Holdren, 2006). Cowan and Hultén (1996) identify six "extraordinary events" that could facilitate the transition from internal combustion engines: (1) a crisis in the technology, (2) regulations, (3) technological breakthroughs, (4) changes in consumer preferences, (5) niche markets, and (6) new scientific results. I interpret these events, along with shifts in social values, as *innovation opportunities* that could be exploited by entrepreneurs.

Further, innovation creates ripples of change throughout the elements of the socio-technical system and its environment. For example, the introduction of a new technology can alter consumers' expectations and even social structures, as witnessed by the automobile in the first half of the last century and the computer in the latter half. In the words of Schon (1967), "…technological innovation belongs to us less than we belong to it. It has demands and effects of its own on the nature and structure of corporations, industries, government-industry relations, and the values and norms that make up our idea of ourselves and of progress. We do not remain the same throughout." Success changes the game (Gharjedaghi, 2006). Thus, technological innovation and changes in the socio-technical system and its environment are co-evolutionary processes, marked by positive feedbacks, associated path-dependent behavior, and potential lock-in (Unruh, 2000).

The multi-directional linkages between the environment and the process of innovation are a major issue in the quasi-evolutionary model and the multi-level perspective discussed in section 2.4. As these approaches point out, most evolutionary studies include interactions with the environment in a very limited fashion. Economic and social factors are incorporated as 'focusing devices' which entrepreneurs factor into the process of determining which technical problems and solution spaces to pursue – i.e.

the direction of variation (Rosenberg, 1976; Dosi, 1982). Little to no attention is paid to the more co-evolutionary interactions involved in aligning product designs and behavioral rules – engineering perspectives, production processes, usage norms, cultural meanings, and world views.

Note that with co-evolutionary models of innovation, it is impossible to attribute simple linear causality. Changing environmental conditions provide opportunities and focus innovation in particular directions at a given time, but at the same time, the introduction, diffusion, and use of new technologies alter the environment. This is particularly true for a transition to a new socio-technological paradigm. This inability to attribute causality is discussed further in section 2.3.2.6.

2.3.2.4 Knowledge and Learning

In the latest economic thinking, the ability to continuously generate knew knowledge through learning is the ultimate determinant of a firm's success and a nation's long-run prospects; knowledge is the new competitive resource (Rycroft and Kash, 1999; Dijkemaa et al., 2006). Evolutionary approaches in economics and innovation are fundamentally behavioral theories that focus on learning processes and adaptation (Metcalfe, 1994). Dosi (1982, 1988) describes innovation as the process of solving technological problems using information drawn from previous experience and formal (codified) knowledge as well as tacit (uncodified) capabilities of the entrepreneurs. The current state of a particular technology can be considered as a multidimensional frontier that represents the current trade-offs among technological and economic factors (Dosi, 1982). Problem solving activity along the current trajectory can be represented by the search for changes in the trade-offs among the technological and economic variables that the *entrepreneurs consider relevant*. Since the entrepreneurs determine what variables are relevant, they define the search space based on their *knowledge and interpretation* of the technical problems and possibilities, economic realities, user needs, and social constraints.

Dosi uses the term "technological paradigm" to encompass the definition of the relevant problems as well as the set of procedures and specific knowledge required for their solution. According to Metcalfe (1994), "The concept of a paradigmatic framework for a technology also provides a natural unit of analysis for the technology policy maker in terms of the knowledge base, theories, facts and concepts, which define the technology and the institutions in which its development is taking place." Following Dosi (1988), a knowledge base is the set of information inputs, knowledge, and *capabilities* that inventors or entrepreneurs draw on when looking for innovative solutions to problems. A knowledge base has a distinct character along a continuum from tacit and specific to public and universal. Tacit knowledge and capabilities are ill defined, uncodified, unpublished and difficult to express. However, it is to some degree shared by, and therefore specific to, inventors and entrepreneurs with common experience.

Pavitt (1984) provides empirical evidence that the knowledge base is a key determinant of the direction of innovation, finding that the innovative choices of industrial firms in the U.K are constrained by their "existing range of knowledge." In addition, an organization's ability to assimilate and exploit new knowledge from external institutions, such as advances in basic science or competitors' production processes, is a critical determinant of the rate of innovation. Cohen and Levinthal (1989, 1990) assert that this "absorptive capacity" is a function of the existing knowledge base, as well as investments in research and development (R&D). This assertion is supported by Lieberman (1984), who finds statistical evidence that R&D expenditures increase the learning rate in the chemical industry.[10]

2.3.2.5 Cumulative Processes

Let us return to Dosi's (1982, 1988) description of innovation as the process of solving technological problems, where the direction of innovation depends on what specific problems are addressed (the search space) and what information is used for their solution. The determination of the search space and techniques for solution is influenced by two internal forces (Rosenberg, 1963). The first is a *reactive* force, driven by the technology itself, as advances "tend to *create* their own future problems, which compel further modification and revision." The second is *self-reinforcing* and relates to the origination and expansion of the entrepreneur's knowledge base. These forces, which I discuss in turn, are cumulative processes that result in path dependence and potential lock-in. As a result, the history of a technology matters and the details of the development path and diffusion process for each technology are unique.

The reactive force arises because "complex technologies create internal compulsions and pressures which, in turn, initiate exploratory activity in particular directions" (Rosenberg (1963). In any production process, the component elements vary in their ability to exceed current performance (*e.g.* capacity, quality, cost), and one particular element may therefore be considered the limiting factor of production. This element may be related to any aspect of firm or industry operations, including supply

[10] Lieberman (1984) finds a statistically significant correlation between the reduction in product price and both cumulative output and an interaction term for R&D investments and cumulative output, while controlling for other sources of cost reductions.

lines, input materials, product design, machinery, plant layout, firm organization, distribution, marketing, *etc.* Innovation efforts are commonly directed at improving this limiting factor. However, such an improvement may exceed the capacity of the next most restrictive factor, possibly even causing a production failure, thereby stimulating further innovations. Consequently, the direction of technical change depends on the technology's cumulative history.

The concept of the limiting factor of production can be generalized to the entire socio-technical system, and the factors may include economic (*e.g.* input prices), political (*e.g.* regulations), social (*e.g.* cultural norms), and environmental conditions. These focusing devices help entrepreneurs select a subset of technological problems from a large set of possible directions for development (Rosenberg, 1976). For example, congestion could be considered a limiting factor for the automobile market. Although commuters could switch to alternative modes of mass transit, those inside the automotive industry would be motivated to maintain the existing dominance of personal vehicles. One possible solution is so-called "intelligent transportation," where vehicle spacing and speed are automatically controlled.

The viewpoint of the agents is of particular significance in this example. Those inside the mass transit industry see congestion as an imperative to develop bus or rail innovations. Those in the information technology industry see an opportunity for telecommuting innovations such as telecommuting centers, satellite broadband, or web-based teleconferencing. This brings us to consideration of how the entrepreneur determines not only what problems on which to focus but also the solution space – the applicable scientific principles, material properties, and technological options.

As asserted in the previous section, the entrepreneur determines the search space and solution techniques based on his existing knowledge and skills. Since technological innovation and changes in the socio-technical system and its environment are co-evolutionary processes, these factors are determined by the socio-technical history. As a technological regime develops, industry organizations arise and academic disciplines emerge. Such organizations create self-sustaining networks, develop standard approaches, and train ‘like-minded’ professionals (Unruh, 2000). Unions and user groups arise, as do trade and enthusiast publications. The perspectives of these organizations become an integral part of the knowledge base and technological framework shared by an entire community of technological actors. These groups reinforce expectations and influence the political process. Eventually, a technological system becomes an integral part of daily life, leading to the emergence of behavioral norms, customs and rituals.

Since the entrepreneurs’ skills, knowledge base, and social view are a product of their experience, the socio-technical system, and its environment, the search space is to some extent circumscribed by the existing technological paradigm. Except perhaps in early in a technology’s life-cycle, innovation seldom consists of a radical breakthrough, but rather results from the synthesis of numerous minor advances (Sahal, 1981). Often, a standard design may emerge that persists for decades while innovation continues incrementally. This overall configuration of the technology remains relatively unchanged, serving as a technological ‘guide-post’ that points in the “general direction of technical advance” and partially determines the boundaries of further development (Sahal, 1981). Thus, the state of the technology and its history determine both the industry’s perceived needs *and* the tools available for change. As a result, innovation follows a particular technological trajectory: a pattern of solution of *selected* technological problems, based on *selected* scientific principles and on *selected* material technologies.

For the most part, these innovative choices are made within firms and are constrained by their existing range of experience and knowledge. As a result, "technical change is largely a cumulative process specific to firms" (Pavitt, 1984). With investments in innovation and experience in production, the entrepreneur continues to develop the technology and his capabilities along the existing technological trajectory. According to Rosenberg (1963), "in technological change…one thing often leads to another – not in a strictly deterministic sense, but in the more modest sense that doing some things successfully creates a capacity for doing other things." In addition, firms and research organizations develop a corporate culture including search heuristics, rules of thumb, and standard operating procedures based on past activities, successes and failures. Gharjedaghi (2006) identifies three strategic behaviors of firms that maintain the existing trajectory. Firms are likely to imitate the strategies of the most successful firms, stick with their own past strategies if they were successful (inertia), and may possibly take their past strategy to a greater extreme (sub-optimization), believing it will solve new problems that arise. The forces behind imitation, inertia, and sub-optimization are the result of economic as well as institutional factors. For example, the firm may have accumulated large investments in specialized equipment and expertise that makes moving in new directions expensive and time consuming. This is especially true in mature industries where efforts have been focused on process innovations; productivity and flexibility are often at odds (Abernathy, 1978).

Of course, the technological trajectory is not entirely self-determined. Rosenberg (1994) refers to the sequence of technological development as a "'soft determinism,' in which one historical event did not rigidly prescribe certain subsequent technological developments, but at least made sequences of technological improvements in one direction easier." As discussed, economic, political, social, and environmental factors serve to focus innovative activity in a particular direction at any one time. However, these factors often stimulate progress along an existing technological trajectory, as when intelligent transportation systems are proposed as a solution to congestion. The accumulation of knowledge, artifacts, and societal experience with the predominating technology tend to "lock" society and industry into the existing technological paradigm. According to Dosi (1982), radical shifts along new technological directions "emerge either in relation to new opportunities opened-up by scientific developments or to the increasing difficulty in going forward on a given technological direction" for technological, economic, or other reasons. Thus a shift of paradigm may be by design (learning and unlearning) or out of frustration (Gharjedaghi, 2006). Note also that the focusing factors and difficulty going forward actually may be due to success of the existing technological paradigm, which has altered elements of the socio-technical system and its environment and possibly posed unforeseen problems as it expanded in scale and scope. Thus cumulative processes may also overcome system lock-in by forcing a shift in paradigm, either in terms of our understanding of the nature of reality (philosophy), our analytical and technological approach, or both (Gharjedaghi, 2006).

2.3.2.6 Increasing Complexity

The increasing complexity exhibited by evolutionary systems models derives from two sources: the continual generation of variety (a distinctive feature of evolutionary systems) and the co-evolution of technology, supporting industries and organizations, society, and behavioral rules, *i.e.*, the socio-technical system and its environment. Evolutionary models of innovation therefore demonstrate increasing complexity in any and possibly all of the elements of the system, including the product; its manufacture, sale and use; markets; and the process of technological innovation. However, evolutionary systems involve both positive and negative feedbacks, and *selection* processes reduce both the variety of products and the number of firms in an industry as a dominant design emerges. For example, Simons (1995) and Klepper and Simons (1997, 2005) find evidence that the 'shakeout' dynamic that reduces the number of firms in an emerging industry is largely due to a process of continual technological change. In mature markets, the introduction of a new competing innovation can also begin a new process of 'creative destruction' that decreases the market for the established technology, the number of firms producing it, and eventually possibly eliminating the industry altogether.

How this last phase occurs for complex socio-technical systems is exactly the question addressed in this research. In the foregoing sections, I have discussed the complexity of firm interactions with other firms, other industries, research institutions, suppliers, users and the firm's socio-economic environment. These interactions make it difficult to attribute linear causality, and make such relationships less meaningful (Rycroft and Kash, 1999). However, the increasing complexity of the product itself, its production process, and the environment in which innovation occurs has a host of implications regarding the process and sources of innovation, as well as the outcome of market selection.

In general, a complex product or system is composed of many interconnected parts, resulting in a whole that is complicated and difficult to understand. According to Rycroft and Kash (2002), "[c]omplex technologies are those that cannot be understood in detail by an individual expert, and cannot be precisely communicated among experts across time and distance." Hodbay (1998) uses the term 'complex' to "reflect the number of customised components, the breadth of knowledge and skills required and the degree of new knowledge involved in production," and defines complex products and systems as "high cost, engineering-intensive products, systems, networks and constructs." Hodbay is concerned with a specific class of products that are manufactured in small batches or as "one-of-a-kind" goods, and does not consider the automobile to be 'complex'. I argue, however, that his definition does apply to automotive technology.

Consider that today's automobile consists of multiple subsystems – powertrain, body, interior, emission control, cabin environmental control, navigation, stability, safety, *etc*. – each of which requires specialized engineering expertise. Further, the interaction of these subsystems can result in unexpected behavior, such that the integration of subsystem designs is also a specialized task. No single engineer has complete *detailed* knowledge in all of these areas. Therefore, no single individual is capable of undertaking innovation in all areas of automotive design. Building on Rycroft and Kash (2002), I define a complex technology or system as one that consists of many, interacting and interdependent subsystems, the behavior of which cannot be predicted by isolated analysis of any one subsystem, and the whole of which cannot be understood in detail by an individual expert nor precisely communicated among experts across time and distance.

The (increasing) complexity of a technology, its production process, and its production and use systems has significant implications for those trying to influence innovation. Rycroft and Kash (1999) describe the development of technologies and their manufacturing processes as a progression from craft production of a simple technology, where innovation is undertaken by individuals; through mass production of intermediate technologies, where innovation is undertaken by in-house R&D units; to what the authors' call 'synthetic production' of complex technologies, where innovation occurs through organization. Rycroft and Kash define 'innovation by synthesis' as "creating new and enhanced technological products and processes with previously unattained performance by combining components, knowledge, and capabilities in ways that deliver synergism."

According to the authors, "[a]ll complex technologies manifest a process of coevolution between the technologies and the organizational networks that produce and use them. Embedded in these organizational networks is a capacity for creativity and innovation that in earlier times was provided by inventors and entrepreneurs like Edison and Ford." Therefore, successful innovating organizations are networks that facilitate rapid learning and purposefully organize to "learn to learn." In these organizations, teams of experts from every commercialization activity, including suppliers and users, work *concurrently* to innovate new technologies. Complexity of product and production process means that components and subsystems are "synthesized by widely dispersed, rapidly evolving, and often obscure organizations." Thus, the focus and targets for technology analysts, strategists, and policy makers become difficult to identify.

Increasing complexity means that production and use systems are increasingly nonlinear, and the process of innovation is altered by emphasizing continuous learning and adaptation, celebrating speed to market rather than efficiency, and shifting key decision-making roles to groups rather than individuals. Furthermore, according to Tushman and Rosenkopf (1992), the more complex the product, the greater role for non-technical factors in influencing technological evolution, and the "less likely the selection process reflects economic efficiency."

2.4 Technological Transitions

Much of evolutionary theory is concerned with innovation within existing technological regimes and the forces that reinforce technological trajectories and regime stability. Although this stability is dynamic, it resists radical change and impedes the introduction and diffusion of alternative technologies. In general, evolutionary research provides only suggestions of how new trajectories and paradigms emerge. Most describe two factors that provide new opportunities for 'radical' change: 1) scientific or technological breakthroughs and 2) difficulty moving forward on the current trajectory. These difficulties include looming theoretical limits (real or perceived); increasing marginal costs of improvements in the current technology; problems posed by the increasing scale or scope of the existing paradigm; and economic or institutional factors.

While the bulk of research in evolutionary economics has neglected a full exploration of the emergence and development of new trajectories, two strands of research specifically target these transitions: the quasi-evolutionary model (QEM) and the multi-level perspective (MLP). Therefore, this research addresses the core question of how radical technologies emerge, compete, and gain momentum despite the barriers and myopia presented by the existing technological system (Schot et al., 1994). The QEM and the MLP arose due to dissatisfaction with the description of the sociological dimensions of technological regimes found in most evolutionary theories of technological progress (Schot, 1992; Kemp, 1994). Therefore, this body of research has combined evolutionary theory with sociological and historical co-evolutionary models. Because of its focus on managing transitions, it has also built on constructive technology assessment, which holds that "during the course of technological development, choices are constantly being made about the form, the function, and the use of that technology and, consequently, that technological development can be steered to a certain extent" (Schot, 1992).

Sections 2.4.1 and 2.4.2 review the literature on the quasi-evolutionary and multi-level perspective theories of technological transitions. Section 2.4.3 summarizes the assertions of these two bodies of theory regarding the role of niches in the development and diffusion of new technologies, or what I have called the niche hypothesis. Section 2.4.4 then raises some theoretical issues with QE theory and the niche hypothesis that give rise to the set of research questions posed in chapter 1.

2.4.1 The Quasi-Evolutionary Model

The quasi-evolutionary (QE) literature identifies three main criticisms of 'mainstream' evolutionary theory that arise from an inadequate description of the sociological dimensions of technology. To some extent, this critique describes errors of emphasis and depth rather than errors of omission. This section describes these deficiencies in evolutionary theory, the elements of the QE model that seek to rectify these deficiencies, and the QE description of technological transition.

2.4.1.1 Elements of the Quasi-Evolutionary Model

According to the QE literature, evolutionary theory fails to adequately account for the sociological dimensions of technology as it relates to regime stability (retention mechanisms), the selection environment, and the linkage between variation and the selection environment. First, evolutionary theory stresses the importance of a core technological framework which encompasses the shared engineering beliefs and expectations of the community of technological and economic actors involved in innovation (see sections 2.3.2.4 and 2.3.2.5). This framework forms the basis for competition and guides the search for improvements in process efficiency and product performance. Thus evolutionary theory conceptualizes the technological paradigm or regime, which serves as a retention mechanism, as a cognitive concept that relates to technologists perceptions regarding the relevant problems and the solution spaces worth pursuing (Nelson and Winter, 1977). While QE theory acknowledges the importance of shared engineering beliefs, it asserts that "there is a clear *socioeconomic* dimension involved in the stability of search activities and the patterns of technological change" (Kemp, 1994).

QE theory therefore re-conceptualizes retention as a process of embedding that is more in line with the co-evolution of a techno-institutional complex described by Unruh (2000). Innovation choices depend "not just on the prevailing interpretative framework of engineers," but also on the embedding of technologies in "engineering practices, production plants and organizational routines," as well as the embedding of products in consumption and usage patterns and in supporting and complimentary systems (Kemp et al., 1998). Thus, while QE theory uses the term 'regime' first introduced by Nelson and Winter (1977), it redefines the technological regime as "the rule-set or grammar embedded in a complex of engineering practices, production process technologies, product characteristics, skills and procedures, ways of handling relevant artifacts and persons, ways of defining problems – all of them embedded in institutions and infrastructures" (Rip and Kemp, 1998, p. 338). The term regime is used in a sense similar to that of a political or regulatory regime in that it refers to rules as a set of commands and requirements and also as roles and practices that guide the research activities of firms, the solution spaces explored, and the strategies of actors. The focused nature of socio-technical change is "accounted for in large part by the embedding of existing technologies in broader technical systems, in production practices and routines, consumption patterns, engineering and management belief systems, and cultural values – much more than it is by engineering imagination" (Kemp et al., 1998).

Second, while evolutionary theory recognizes that the selection environment includes social and political mechanisms, these interactions generally are discussed only in non-market sectors such as government services. Thus, market interactions are the main focus for both *ex ante* and *ex post* selection (see section 2.3.2.1). However, social and political conditions are included peripherally as 'focusing devices' in *ex ante* selection; entrepreneurs' perceptions of these conditions may focus attention to particular research directions (problems and solution spaces) that are likely to be profitable at a given time (see sections 2.3.2.3 and 2.3.2.5).

The view of the selection environment in QE is closer to Metcalfe's (1995) assertion that "[a]ny framework in which agents interact in order to choose between competing patterns of behaviour has selective properties." Therefore, the selection environment includes institutional and geographical factors as well as economic markets (Schot, 1992). New technologies must become embedded in existing socio-technical system, making non-market interactions a significant factor in adoption and diffusion. Thus, "[t]he selection environment is defined by the capital outlays, physical infrastructure, supplier-user linkages, production routines, skills, technical standards, government rules, norms, people's preferences and beliefs" (Kemp, 1994). These existing institutions, linkages, and patterns of behavior are determined by the socio-technical history and are therefore aligned with existing technologies. According to Kemp (1994), "[t]he key problem for new technologies to become incorporated into the socioeconomic system is that of compatibility… New technologies that can easily be embedded in the production system and people's ways of life will diffuse more rapidly than technologies which require the replacement of capital goods, a new infrastructure, different skills, new ideas about production and consumption, and regulatory changes."

However, the embedding of a new technology alters the environment such that technology and the wider socioeconomic system co-evolve in a process of mutual adaptation. This brings us to QE's third and most significant criticism: evolutionary theory of technical change lacks a complete understanding of the coupling between variation and selection processes. As discussed above, evolutionary theory fails to adequately account for the influence of social factors on variation via *ex ante* selection. In addition, evolutionary theory neglects how the introduction of new technology changes the selection environment through the process of embedding. To realize successful design, 'heterogeneous engineers' must be involved in overcoming technical, economic, and social obstacles. As a result, they often must design elements of the environment in which the technology is marketed, sold, and used – i.e. the selection environment (Schot, 1992). As a result, social, political, and economic factors are embedded in technology development, and successful variations change the selection environment. As a technology comes into wider use, new institutions arise and old ones adapt. Consumer tastes, lifestyles, and habits also adapt. As user feedback is incorporated into product improvement, new functions emerge and the technology and consumer tastes co-evolve. While the selection environment has a certain momentum that tends to block some variations and encourage others, it can be altered by the introduction and embedding of new technology. Therefore, the distinguishing characteristic of the QE model is its focus on "the way variation and selection processes are partly independent and yet coupled" by the regime (Schot et al., 1994).

This coupling occurs through *ex ante* selection, but also through interactions involving institutional linkages that QE theory calls the 'technological nexus.' The active, boundary-spanning nexus role is carried out by a broad range of agents who link selection and variation by translating information from both realms; shaping interactions; and harmonizing social and market needs with the results of scientific and technological research. A number of actors within and external to firms shape the process of variation either through the direct formulation of research objectives and search heuristics or by attempting to influence the process from the outside (*e.g.*, government regulators and environmental activists). However, actors serving in the nexus role translate the requirements of the selection environment into objectives for technological development and also impose the requirements of technology on the selection environment. These activities result in a learning process that occurs simultaneously in firms and in the environment (Schot, 1992).

QE theory asserts that the co-evolutionary nature of technology and the socio-economic environment in which it is produced and used results in regime stability and gives rise to trajectories of technological development. According to Kemp (1994), "the dominance of particular trajectories is related to the 'dynamic scale and learning effects' from which prevailing technologies have benefited and the adaptation of the 'selection environment' to the old technological regime" (Kemp, 1994). Thus, cumulative learning effects and alteration of the selection environment via the coupling of variation and selection together serve as a retention mechanism.

2.4.1.2 Quasi-Evolutionary Theory of Transitions

While co-evolutionary forces lead to stability in technological regimes, this stability is dynamic. Innovation continues along the current technological trajectory, to some extent constrained by the regime as described above. Elzen et al. (2003) refer to these incremental improvements within the existing regime as 'regime optimization.' On the other hand, 'regime renewal' refers to radical innovations that depart from the existing trajectory and require a change in the architecture and rules of the existing regime. The distinguishing feature of regime renewal is that it requires considerable change in "the behavior of various actors in relation to various relevant technologies" (Elzen et al., 2003). For example, these changes may include new interpretation of the function of technology, new usage norms, and new forms of ownership. While incremental innovations in regime optimization may accumulate to yield significant improvement in system function, regime renewal has much greater potential to yield societal benefits. The question then arises as to how radical innovations emerge and diffuse despite the inertia presented by regime stability.

Much innovation research has asserted that radical innovations depend on the opening of new technological and economic opportunities by new scientific insights or advances in engineering and material technology. However, novelty may emerge due to pressing technological needs that arise from bottlenecks in development due to the growth in technological systems, from pervasive shifts in consumer preferences, or from external crises such as war which produce new demands or create resource scarcities. In addition, looming theoretical limits to current technology, real or perceived, may stimulate the search for innovation along a new trajectory (Kemp, 1994).

Schot et al. (1994) list two possible routes for new technologies to compete and gain momentum despite the "myopia of the existing technological system." First, the existing system may lose viability due to a changing selection environment which presents new challenges that the existing technology cannot meet or can only meet at excessive marginal cost. Second, new technologies may "use a niche that protects them against too harsh selection and provides space to grow." A niche can be defined as a distinct application domain, small in scale and scope, characterized by specific functional requirements (Hoogma et al., 2002; Raven, 2005). Because of the special requirements of the niche, users are willing to accept higher cost or lower performance on standard functionalities compared to existing products available in established mass markets.[11] This first route, which I will refer to as 'regime instability,' remains relatively unexplored within the QE literature. Meanwhile, a predominant theme found in the QE literature is the critical function of early market niches in the development and eventual diffusion of radical technologies. The remainder of this section describes the basic tenets of this 'niche hypothesis.'

According to Rip and Kemp (1998, p. 338), novelty arises within existing regimes "starting at the micro-level of local practices." These radical technologies are relatively immature at the time of introduction, require improvement, and are only able to compete in specialized markets or market niches. Kemp (1994) asserts that "market niches may be an important stepping stone for the further evolution of radically new energy technologies. It helps suppliers better to understand user needs, to identify and overcome critical problems, to achieve cost reductions in mass production, and, perhaps most important, to build a constituency behind the new product to sustain a new technological trajectory." In reviewing the historical evidence of transitions, Kemp et al. (1998) conclude: "Apart from demonstrating the viability of a new technology and providing financial means for further development, niches helped to build a constituency behind a new technology, and to set in motion interactive learning processes and institutional adaptations – in management, organization and the institutional context – that are all-important for the wider diffusion and development of the new technology."

A brief review of Strategic Niche Management (SNM), a policy recommendation that emerges from the QE literature, helps illuminate the assumptions regarding processes

[11] In contrast to niche markets, mass markets are understood to be large in scale and scope. Thus, products provide functions with wide appeal and demand is sufficient to allow high volume production. Competition in mass markets is based mainly on cost and quality.

that occur within niches. Kemp et al. (1998) define SNM as “the creation, development, and controlled phase-out of protected spaces for the development and use of promising technologies by means of experimentation.” SNM seeks to create linkages between users and other actors to stimulate learning and facilitate institutional adaptation. Critical to this process is the alignment of agent expectations; the articulation of technological characteristics, policies, cultural and psychological meanings, the market (users and their needs), and production and maintenance needs; and the formation of networks. The four aims of SNM are:

1) To articulate the changes in technology and institutional framework necessary for economic success;
2) To learn more about the technical and economic feasibility and environmental benefits of the technology;
3) To stimulate further development of the technology, achieve cost efficiencies in mass production, promote development of complementary technologies and skills, and stimulate changes in social organization important to wider diffusion; and
4) To build a constituency behind the technology.

Niches, then, serve as temporary protected spaces that allow the technology to be developed in terms of user needs and to achieve wider use through learning processes and adaptations in the selection environment. The technology may then be adapted to new niches which are defined by new geographical regions or new applications. Thus, a technological regime shift is conceptualized as a process of niche proliferation or niche branching. “Eventually, [the protected] technology might compete head-on with the dominant technological option in a part of its market or markets" (Hoogma et al., 2002, p. 4).

While the QE conversation concentrates on niche development as an independent process, there are two notable exceptions to which we will return later. First, Kemp (1994) comments that "[o]nly in a few respects did radically new products constitute a radical break with the past, which suggests that the term 'radical' is somewhat misleading. Radical innovations often combined the new with the old (or even combined older technologies) and often rightly so because this helped the product to survive the initial harsh market selection and establish itself in the market place." Thus, new paradigms emerge and develop within the previous one. Second, Kemp et al. (1998) refer in passing to the importance of niche-regime interactions and developing regime instability: "The processes of niche formation occur against the backdrop of existing technological regimes… The success of niche formation is, therefore, linked to structural problems, shifts and changes within the existing regime(s). The ultimate fate of processes of niche formation depends as much on successful processes within the niche as on changes outside the niche: it is the coincidence of both developments that gives rise to niche development patterns."

2.4.2 The Multi-level Perspective

The multi-level perspective (MLP) compliments the QE model and arose from a similar perspective of the sociological shortcomings of evolutionary theory. The MLP integrates the findings from various literatures, including evolutionary theories, the sociology of technology, and the history of technology, into a set of analytical concepts to study the complex dynamics of socio-technical change. In the MLP, technological transitions are seen as evolutionary processes, where evolution is both a process of variation, selection and retention, and also a process of unfolding and reconfiguration. This follows from Metcalfe (1995), who states that "evolution means two things: the gradual unfolding of phenomena in a cumulative and thus path-dependent way; and quite separately, a dynamics of system behaviour which creates change and emerging structure from variety in behaviour."

The MLP builds on the concept from actor-network theory that technical systems are configurations or networks of linked heterogeneous elements (technical and social) where the linkages among elements provide stability. However, the MLP distinguishes among three types of elements: systems, consisting of resources and material aspects; actors involved in maintaining and changing the system; and the rules which guide actors' perceptions and behavior. The concept of a regime in the MLP then closely follows the QE model but the term 'socio-technical regime' is used to refer to the "semi-coherent set of rules carried by different social groups." The MLP identifies three types of rules: cognitive, normative, and formal/regulative.[12] Rules provide structure for and coordinate human interactions yet leave room for perception and strategic action (Geels, 2004). The alignment of rules and the linkage of elements results in system stability and trajectories of development. Thus in a regime, the rules of all relevant social groups (users, engineers, producers, policy-makers, *etc.*) are aligned with a particular technological system or artifact (*e.g.* personal automobiles) (Raven, 2005). The regime thus serves as a selection and a retention mechanism (Geels, 2002).

---

[12] The MLP distinguishes rules from institutions, the latter referring specifically to public organizations.

Despite this stability, the system remains dynamic and interactions between elements are multidirectional. As illustrated in Figure 2-1, while rules shape human action, they are also carried and reproduced by humans and can therefore be changed by human action. Similarly, artifacts are created by human action but also shape further action. Rules are both embedded in and shaped by artifacts. This multidirectional interaction allows for co-evolution and endogenous change.

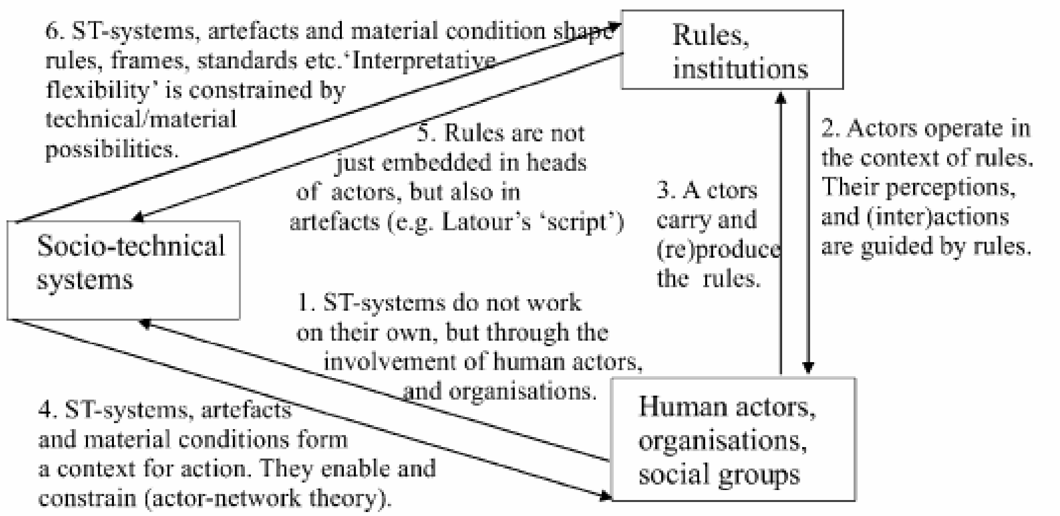


Figure 2-1 Interaction of System Elements
(reproduced from Geels, 2004)

To analyze technological transitions, the MLP (Geels, 2002, 2004, 2005a and 2005b) conceptualizes a nested hierarchy of three heuristic levels: (1) the meso-level consists of socio-technical regimes, where cognitive rules are embedded widely in the knowledge base, engineering practices, corporate structures, manufacturing processes, product characteristics, *etc*.; (2) the micro-level consists of technological niches; and (3) the macro-level is formed by the socio-technical landscape. The socio-technical landscape provides an external structure or context, containing the material aspect of society (*e.g.*, cities, factories, and highways) and heterogeneous factors such as oil prices, economic growth, wars, cultural values, and environmental problems. The transition to a new system involves crucial interactions between the socio-technical regime (the meso-level) and the two other levels.

Regimes are embedded within the landscape which undergoes slow change and provides ‘gradients’ for the technological trajectories of regimes. As in the QE model, innovation within the regime is incremental. Radical innovations emerge in niches in the context of the specific problems, rules, and capabilities of the existing regimes and landscapes. These innovations require protection because they “have relatively low technical performance, are often cumbersome and expensive” (Geels, 2002). Niches insulate the novelty from the selection environment represented by the regime and provide space for learning processes (learning-by-doing, learning-by-using, and learning-by-interacting) and for building social networks, such as supply chains and user-producer relationships.

Thus, the MLP conceptualizes innovation as a process involving three phases (Geels, 2005a). First, novelties emerge in within the context of the existing regime, in part due to landscape developments. Entrepreneurs experiment with designs in an attempt to determine users’ requirements, and a variety of technical forms may compete with each other. Second, the novelty is adopted in niche markets which serve as incubators. Within the niche market, the innovation develops a technological trajectory and improves due to learning processes, while users explore new functionalities. This phase results in the articulation of user preferences and the development of a dominant design. Third, the innovation diffuses into wider markets and competes with the established regime.

However, successfully breaking out of niche markets depends not only on processes within the niche but also on developments within the regime and landscape. Technological transitions therefore occur when simultaneous processes on multiple dimensions, both technological and social, align and reinforce each other (Geels, 2002, 2004, 2005a). In particular, when the activities of social groups are 'mis-matched,' the resulting tension provides windows of opportunity for radical innovations to break into wider markets. Geels (2004) identifies five sources of tension and misalignment:

1) Changes at the landscape level that put pressure on the existing regime and cause internal restructuring.
2) Technical problems internal to the existing regime, such as bottlenecks, diminishing returns to innovative activities, and looming (real or perceived) theoretical limits.
3) Negative externalities of existing regime.
4) Changing user preferences due to negative externalities, cultural changes, changes in relative prices, policies, or endogenous change.
5) Strategic and competitive games between firms.

Two patterns of transition emerge from the interplay between the levels in the MLP. In the 'technological substitution route:'

> …the existing sociotechnical regime is initially relatively stable, characterised by incremental developments. Radical innovations emerge in niches… Breakthrough of the novelty in mainstream markets may occur when the novelty has acquired enough internal momentum, and when landscape developments put pressure on the existing regime… This may lead to creative destruction and the downfall of established firms. The breakthrough triggers all kinds of adaptations and transformations in the regime, leading to an 'era of ferment.' Hence, this route has a technology push character. Once a new sociotechnical regime has been formed, the dynamic shifts back to incremental change (Geels, 2004).

In the 'wider transformation route:'

> …the regime becomes unstable and opens up in an early phase, because of persistent problems or landscape changes. Simultaneous changes occur on multiple regime dimensions, e.g., policy, user preference, technology, infrastructure, culture, etc. The loosening up of the existing regime stimulates actors to experiment with other technical options. There is a prolonged period of experimentation, and strategic maneuvering. Such a period of heating up is eventually followed by a period of 'cooling down,' i.e., narrowing down the number of technical options. A particular technical option may come to be seen as 'universal,' push other options out of the market and acquire dominance. This is complemented by the creation of a new sociotechnical regime (Geels, 2004).

Geels (2002) clearly identifies niche accumulation as the primary mechanism leading to wider diffusion of radical technologies. "The step from niche to regime-level does not occur at once, but gradually, as radical innovations are used in subsequent application domains or market niches." However, the penetration of new niches is not necessarily easy, and involves "experimentation, learning processes, adjustments and reconfigurations" (Geels, 2002). If the specific characteristics of the innovative technology lead to the articulation of new functionalities, the novelty may penetrate new application domains without competition from the existing technology. In a stable regime, this process of 'niche branching' allows the technology to develop within niches where the new functionality is especially suited and 'below the surface' of the existing regime. In the absence of new functionalities, the novelty must conquer new niches by battling head-on with existing technology. In an unstable regime, problems in the regime or landscape present opportunities for the new technology to serve as a solution, effectively creating new niches. New functionalities may be articulated early in the process as regime actors search for new options and niche actors present their innovation as a solution. Thus, the novelty may link up with and reinforce processes in many dimensions, leading to transformation and co-evolution early on (Raven, 2005).

In the case of unstable regimes, Geels (2002) identifies a second mechanism of diffusion: technological add-on and hybridization. In this process, new technologies are merged with existing ones to solve particular problems by forming a symbiosis. Thus, the two technologies are used in conjunction with as few regime modifications as possible. However, the hybrid technology may gradually evolve into a new form defined by the functional capabilities of the novelty. This is mechanism is consistent with a more general 'fit-stretch' pattern often found in the co-evolution of a new technology's form and function as illustrated in Figure 2-2. The fit-stretch pattern satisfies the issue of compatibility raised by Kemp (1994; see section 2.4.1.1). New technologies are likely to diffuse more rapidly when both form and function fit closely with the existing regime. However, continuous development eventually leads to the articulation of new functionalities and new technical designs tailored specifically to these functions.

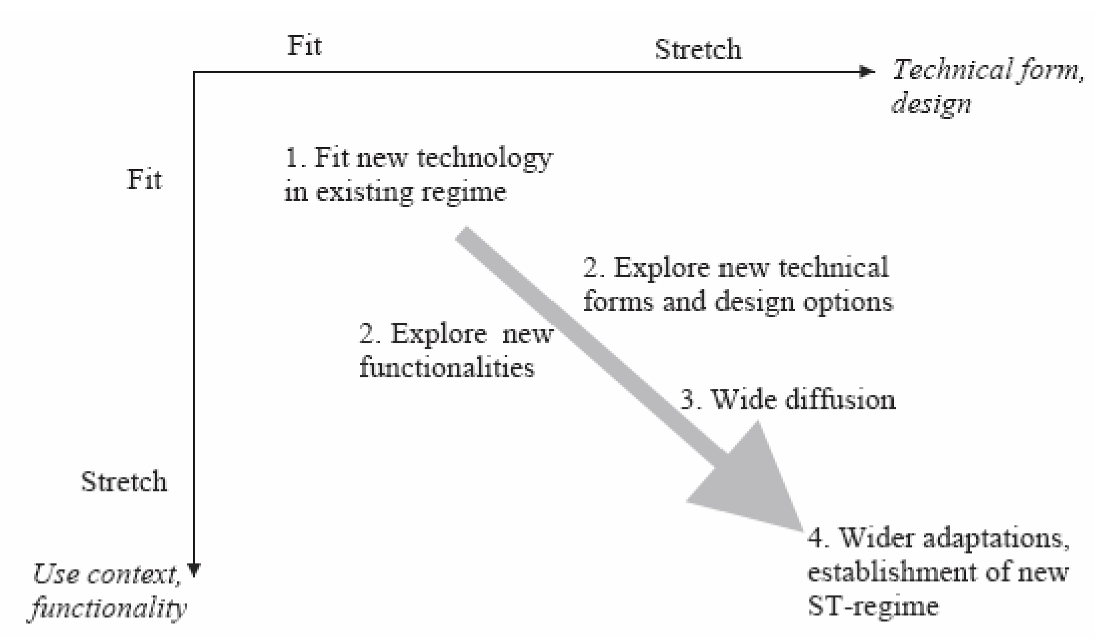


Figure 2-2: Fit-Stretch Pattern in the Co-evolution of Form and Function
(reproduced from Geels, 2005a)

Geels (2002) identifies a third and final mechanism of diffusion: "new technologies may break out of niches by riding along with growth in particular markets." Though Geels does not identify the forces behind such growth, I will posit that growth of a particular niche may be due to landscape development or endogenous changes in consumer tastes and preferences.

In summary, the distinct contribution of the MLP is the inclusion of two factors in the transition to new socio-technical regimes: 1) the socio-technical landscape as a constraint and an opportunity; and 1) the stability of the existing regime and resulting niche-regime dynamics. The MLP conceptualizes technological transitions as gradual reconfigurations involving a series of adaptations that take place "on all dimensions of the socio-technical regime (*e.g.* markets, user groups and user practices, technologies, production networks, policies)" Geels (2002). Three primary factors contribute to the wide diffusion of a new, radical, technology: 1) niche markets that act as incubators; 2) 'niche-accumulation;' and 3) circumstances at the regime or landscape level that create windows of opportunity. The MLP has identified two transition routes distinguished by the initial stability or instability of the existing regime. Within each route, the transition path is also defined by articulation of new functionalities for the novelty or the lack thereof.

2.4.3 The Niche Hypothesis

Both the QEM and MLP stress the role of niches – distinct application domains that are small in scale and scope – in the emergence, development, and diffusion of radical innovations. The transition to a new technological regime, defined and stabilized by a set of rules shaping the interactions of social and technical elements, is described as a process of niche branching, niche proliferation, and niche accumulation. Based on the QE literature, I identify five fundamental assertions of the role of niches in regime transitions, which I will call the niche hypothesis:

6) Within the niche, manufacturers learn about users' needs and users learn about product performance through "learning-by-interacting."
7) Manufacturers improve the technology in terms of those needs and achieve lower costs through "learning-by-doing."
8) As a result of these learning processes, the new technology becomes embedded in a new rule set which reflects the adaptation of institutions. Both the technology and the system in which it is produced and used are altered in this co-evolutionary process linking variation and selection which arises from the interaction of agents in the technological nexus.
9) Growth of niche markets, niche branching, or niche proliferation allows manufacturers to reduce costs through economies of scale.
10) These changes facilitate diffusion into wider markets.

The MLP and the work of Raven (2005) add regime dynamics and instability as a significant factor in the growth (or lack of growth) of niches. Instability may arise from three sources:

4) Success of niche technology.
5) External forces – changes at the landscape level – which:
    a. influence the direction and rate of variation by altering regime agents' perceptions and expectations or altering the resources available for variation; or

b. alter the selection environment by altering prices or consumers' preferences.

6) Internal regime forces – difficulties with the expanding scale or scope of the existing regime or achievement of technological or economic limits of existing technology – which lead regime actors to view the new technology as a solution.

2.4.4 Theoretical Issues

The niche hypothesis has led to a specific policy recommendation within the QE literature called strategic niche management (SNM). The literature has thus shifted focus to SNM, which I believe leaves important aspects of the QE theory of innovation under-developed. Section 1.3.4 listed four related theoretical issues with the QEM and the niche hypothesis which I repeat here.

First, the focus on the coupling between variation and selection leads to an exclusive emphasis on learning-by-doing (learning curve effects) and learningbyinteracting with users. This yields an incomplete description of the innovative process behind product and process improvements that are critical for wide diffusion.

Second, the technological nexus that links variation and selection is discussed as a key aspect of the process of embedding, yet remains an underdeveloped concept. Who are the relevant agents? How do new networks form and how are they sustained? Are elements of existing networks important? If so, how are they altered?

Third, the focus on a single dominant regime neglects interactions with other existing regimes which may help or hinder market growth (Raven, 2005). New technologies may address problems in multiple regimes or be based on fundamental innovations with applications that cut across regimes. This relates back to the first and second issues, in that the process of innovation is more complex than is reflected in QE theory. It involves multiple learning mechanisms and networks that tie the innovators to a relevant knowledge base. These networks may link seemingly unrelated industries, particularly in the early phases of a technology's life cycle.

Fourth, as Raven (2005) notes, the focus on the operation of niches in competition with a stable regime results in a simplistic description of niche-regime interaction. It neglects the dynamic processes within stable regimes (incremental change) that may either hinder or assist the diffusion of the new technology. In particular, it neglects regime response to internal and external forces of change which cause instability. Such challenges may be interpreted as opportunities for innovation, either for change within the regime or for radical new technologies that lead to a complete regime transition. While the MLP begins to address regime dynamics, it only does so in relation to the presentation of opportunities for innovations that fall outside the regime. The consideration of change within the regime leads to the last issue.

The dichotomy of regime optimization (incremental change) versus renewal (radical change and transition) is overly simplistic for complex technologies. It places policy focus on transformation of the socio-technical regime in its entirety which requires a complete alteration of both producer and user interpretation of the technology, its function, and the related organizations and rules. Meanwhile, it discounts innovations which make only part(s) of the existing regime obsolete but could still radically reduce environmental impacts. More significantly, it neglects intermediate or transitional technologies that may form a bridge to a new regime, such as hybrid and plug-in hybrid vehicles as a pathway to full electric or fuel cell vehicles. In such a transition, only part(s) of the regime may be transformed at a given time. This deficiency is interesting, because early works on QE theory highlighted the significance of compatibility issues: new technologies that may be easily embedded in the existing production, use, and institutional systems will diffuse more rapidly. According to Kemp (1994), "[o]nly in a few respects did radically new products constitute a radical break with the past, which suggests that the term 'radical' is somewhat misleading. Radical innovations often combined the new with the old (or even combined older technologies) and often rightly so because this helped the product to survive the initial harsh market selection and establish itself in the market place." Thus, transitions can occur as a more gradual unfolding and adaptation of the existing regime than the path described by 'regime renewal.' This suggests that the concept of transition as gradual reconfiguration and the mechanism of hybridization introduced by Geels (2002) should be more prominent in the formalization of a model of innovation and transition.

2.5 Additions to Theory

In order to develop a structured framework for analysis that addresses the theoretical issues raised above and allows exploration of the research questions, I draw from two additional themes within the evolutionary systems of innovation literature. Section 2.5.1 reviews the industry life-cycle model which provides a structure for analyzing regime dynamics, specifically the emergence of regimes, the stability of existing regimes and innovation within established regimes. In section 2.5.2, I introduce a more complete description of the learning mechanisms involved in innovation that will provide the nomenclature for analyzing the learning processes involved in improving the performance and cost of early motor vehicles. I then develop a conceptual framework in section 2.6 that will be used to explore the research questions posed in section 1.3.5. This framework begins with the sectoral system of innovation which focuses more directly on the entrepreneurs and the linkages that facilitate learning and innovation. This provides the perspective to analyze innovation processes within and between firms, the both the development of the technological nexus, and the learning mechanisms involved in transitioning the automobile from niche markets to a regime. Each of these approaches used in constructing this framework has its limitations. However, by integrating them, I hope to take advantage of their explanatory power while compensating for their shortcomings.

2.5.1 Industry Life-Cycle Model

As discussed in section 2.3.2.1, the classic phases of development – invention, innovation, product development, and diffusion – are not clearly defined in systems models of innovation. However, a number of researchers have noted general patterns in the development of technologies, industries, production processes, and markets. The number of stages of development identified by these researchers varies, as does their nomenclature. This section provides an overview of the concepts found in the literature consolidated into four phases: emergent, transitional, specific, and senescent. While development might be more accurately considered along a continuum, the phases described here and summarized in Table 2-1 provide a useful structure and vocabulary for studying technological progress.

Let us consider first the early, or *emergent*, phase of a new technology that begins with the first practical application of a new idea. This phase is characterized by radical innovation, flexible production capacity, and the emergence of new firms (Dosi, 1982; Freeman, 1996; Utterback and Abernathy, 1975; Abernathy and Utterback, 1978). New knowledge is generated by directed research activities and is largely tacit, existing as uncodified knowledge of the entrepreneurs (Grübler et al., 1999a and 1999b; Cowan et al., 2004). At this stage, primary importance must be attributed to the institutions which produce and direct the accumulation of knowledge, the existence of risk-taking actors or 'product champions', and the exchange of knowledge among R&D units within firms and external research institutions (Dosi, 1982; Kantrow, 1980). Because no supporting industries exist, entrepreneurs must either use off-the-shelf parts and process equipment or manufacture their own custom parts and equipment. Thus, craft production techniques and a high degree of vertical integration are typical early in the emergent phase.

Table 2-1 Phases of Technological Development

	Phase			
	Emergent	**Transitional**	**Specific**	**Senescent**
Innovation	Innovation is rapid, product oriented and radical. Primary importance of risk-taking product champions. Focus on defining users' needs and reducing risk.	Refinement of users' needs. Focus shifting to production cost and product quality. Innovation becoming more process oriented.	Innovation is incremental and process oriented.	Innovation stalls.
Primary Source of Innovation	Interaction with users, suppliers, competitors and manufacturers of complimentary products. Spillovers.	Learning-by-doing. Interacting with users and suppliers.	IR&D. Interacting with suppliers.	Learning-by-interacting with users.
Supporting Network	Industry produces its own supplies and process equipment.	Emerging support industries and institutions. Vertical disintegration.	Complex web of supporting and complementary industries and institutions.	Supporting industries unable to survive. Vertical re-integration.
Knowledge Base	Uncodified knowledge of entrepreneurs.	Institutional knowledge is increasingly codified.	Institutional knowledge of sector and supporting network, codified in plants, capital equipment, and operating procedures.	Institutional knowledge of individual firms.
Production System	Craft	Craft → Mass	Mass → Synthetic	Synthetic
Product	Variety of product configurations. Competition based on product differentiation.	Emerging dominant design. Competition based on product differentiation.	Products proliferate around dominant design. Competition based on product price and quality.	Products tailored to niche markets with special requirements or those requiring heritage designs.
Market	Niche	Growing	Established	Niche

After the feasibility of the product has been demonstrated it is still unable to compete on a cost basis with existing technologies in mainstream markets. However, it may have performance advantages in specific uses which lead to competitive application in niche markets. In fact, Christensen (1997) finds that the very attributes that make disruptive technologies uncompetitive in mainstream markets may actually count as positive attributes in initial markets.[13] Within the niche market, various technical approaches and product configurations may be competing. At this point, there is a great deal of uncertainty about the product and its potential market. Therefore, entrepreneurs must interact with suppliers, customers, manufacturers of complementary products, and competitors to reduce these uncertainties (Afuah and Utterback, 1997). Freeman (1982) finds that successful attempts at innovation are distinguished by greater attention to the education of users, to publicity, to market forecasting and selling, and to the understanding of user requirements. However, Christensen (1997) points out that no one, not market researchers or even users, can predict what the early markets and uses for a new technology will be. Rather, through a process of experimentation with real products, customers discover these uses at the same time as producers do. These assertions, drawn from empirical observations, are consistent with processes of learning and the mutual adaptation of variation and selection described in the niche hypothesis. Cowan et al. (2004) assert that cooperation among firms and geographic clustering are likely to enhance innovation during this phase since the transmission of tacit knowledge often requires face-to-face contact.

---

[13] Christensen (1997) distinguishes technologies that disrupt or redefine the performance trajectory from those that sustain the trajectory.

Next, during the *transitional* phase of a product's lifecycle, a dominant design emerges as consumers' needs and product features, design, and components become standardized. Innovation slows and technological change becomes incremental and focused on cost reduction and quality enhancements, often through process improvements. Firms compete mainly on product differentiation (Afuah and Utterback, 1997). Knowledge is codified and technology is largely embodied in production plants, equipment and products. New knowledge is generated primarily through learning-by-doing and learning-by-using such that and user and supplier interactions are of primary importance (Abernathy and Utterback, 1978; Grübler et al., 1999a). Malerba (1992) finds empirical support for these assertions, but also finds that advances in science and technology contribute to firms' efforts at product differentiation. Technological advances still convey temporary monopolistic advantages, and possibly long-run oligopolistic positions, which serve as an incentive for further innovation (Dosi, 1982).

Afuah and Utterback (1997) name a third *specific* phase where products proliferate around the dominant design and even more emphasis is given to process innovation. Knowledge is embodied in capital equipment (process technology) and engineering personnel (Abernathy et al., 1983). There is little product differentiation – the product design becomes increasingly standardized – and the industry is primed for the next round of innovation. A new innovation may render the old technology obsolete, pushing it into decline, or senescence.

Abernathy (1978) and Abernathy et al. (1983) refer to the specific phase as industry maturity, but note that the evolutionary process is not "self-contained or irreversible". A product design can be conceptualized as a combination of functional requirements and the product attributes that fulfill them. Changes in market preferences and technological innovations (both internal and external to the industry) can alter the relative importance of the product's parameters and attributes, possibly adding new functional requirements. Changes in market preferences may be due forces outside the industry, including changes in the prices or attributes of complementary or substitute products, shifts in cultural values, governmental actions, economic or social crises, *etc.* Such interactions with the environment were discussed in section 2.3.2.3. Firms must develop new design concepts through product innovation in order to remain competitive in the new market reality.

Mature firms that have experienced past success may be at a disadvantage at this point. Concentration on productivity is often associated with a decrease in innovative flexibility due to capital investments in highly specialized equipment that makes retooling expensive and time consuming (Abernathy, 1978). Radical product innovation is disruptive and competence destroying. In addition, these firms' past experience shapes their current thinking (see sections 2.3.2.4 and 2.3.2.5). For both cognitive and strategic reasons, incumbent firms' innovative efforts are shaped by their existing customers, and they therefore are often slow to develop innovations that appeal to new customers (Christensen, 1997; Klepper and Simons, 1997). Thus, innovations that result in new uses for a product by a new market segment provide opportunities for new entrants. If technical progress is rapid, the performance of the new product in serving existing functions may eventually provide an opportunity for these new firms to compete with the incumbents, possibly leading to the decline of industry leaders.

Gharajedaghi (2006) identifies three hierarchical forces of this nature that can transform past success into future failure: inertia, sub-optimization, and change of the game. Inertia and sub-optimization were introduced in section 2.3.2.5. In the third force, change of the game, a firm's past success essentially solves the major historic technological challenge and fundamentally changes the nature of the problem – the industry has come to the end of the current trajectory. Because of its triumphant success, the incumbent firm may have difficulty seeing or accepting the new direction of development. The end result is that established firms are likely to attempt 'retrofitting,' altering secondary or component attributes before pursuing fundamental technological changes. Meanwhile, new entrants have the advantage of starting from scratch with completely new product designs and process equipment. However, there is renewed uncertainty and risk in new designs as users' needs and technological solutions are redefined.

Changes in the functional requirements of a product and the resulting round of innovations represent a reversal of the process of maturity, or what Abernathy calls a phase of 'de-maturity'. De-maturity leads to an increase in the variety of products available in the market and possibly an increase in the number of firms in the industry. This introduces the idea that a product and industry may experience cycles of maturity and reinvention, moving between the specific and transitional phases. Anderson and Tushman (1990) propose a cyclic evolutionary model of technological change in which a stochastic technological breakthrough initiates an era of innovative ferment which generates a variety of product configurations. Competition among these variations culminates in the selection of a single dominant design. The era of ferment is then followed by a period of incremental technical progress which preserves the standard architecture until a new innovation initiates another cycle of variation, selection, and retention.

The final or *senescent* phase of technological development begins when a new innovation begins to dominate and the existing technology experiences declining market share. Eventually, the old technology once again finds application only in niche markets where it still holds some performance advantage or where heritage design is required. As markets shrink, supporting industries become economically unviable. Remaining firms in the final product industry must begin manufacturing their own parts and become vertically integrated. Innovation slows or stops, is tailored to the needs of niche users, and occurs primarily through learningbyinteracting with these users. Eventually, even the remaining firms are unable to survive and the industry ceases to exist.

2.5.2 Mechanisms of Learning

Systems models of innovation assert that technical progress is dependent on knowledge and learning. In this section, I identify four mechanisms of learning, each of which has a distinct character and knowledge base. The boundaries of these classifications should be interpreted as being rather ‘fuzzy,’ since entrepreneurial activities often fall at the boundary or within more than one class (Clarke et al., 2006). While all four sources are expected to contribute to the advance of a given technology at any point in time, their relative importance may shift as an industry matures. Various researchers’ theories on the relationship between learning mechanisms and industry phase are included in the discussion. Also, the ability of an entrepreneur to learn via any one mechanism is expected to depend on efforts in other mechanisms, the depth of all relevant knowledge bases, and on how tightly the entrepreneur is linked into research, production, and use networks.

2.5.2.1 Internal Research and Development

The distinguishing characteristic of research and development (R&D) is the institutionalized, guided search for new knowledge, generally aimed at solving particular problems – what Malerba (1992) calls learning by searching. The goal may be to expand fundamental sciences (so called 'basic research'), apply existing or new scientific findings to a technological problem (so called 'applied research'), to develop original products, or improve existing product design or production. R&D is generally systematic, progressing through problem definition, theory development, data acquisition, model construction, and bench testing, with results used iteratively to refine the process.

R&D is undertaken by private firms, government agencies and labs, universities, and non-profit organizations. I use the term 'internal R&D' (IR&D) to refer to research investments within a single firm or industry, as distinguished from research in other firms, institutions and industries (depending on the unit of analysis), which I classify as 'spillovers.'

Investments in IR&D expand a firm's internal knowledge base and directly produce technological innovations. In addition, the research effort enhances a firm's ability to identify, assimilate, and exploit knowledge, new findings, and innovations developed outside the firm (Cohen and Levinthal, 1989 and 1990). In other words, IR&D develops the firm's capacity to access and utilize external knowledge bases (spillovers), what Cohen and Levinthal call 'absorptive capacity'.

R&D undertaken by research organizations is likely to play the greatest role during the emergent phase of technological development when innovations are more radical. During this phase of development, researchers are forging new frontiers, so knowledge is largely tacit, residing primarily in human capital. However, during this phase, manufacturers are focused on determining user's needs and their relative preferences for the various product attributes that can meet these functional requirements. New firms therefore are likely to delay establishing R&D units until the emergence of a dominant design reduces the market risk.[14] In such cases, IR&D is expected to contribute to incremental product and process innovations later in the product's life cycle. With complex products such as automobiles, component subsystems may undergo repeated cycles of radical innovation. However, these innovations may originate from IR&D as well as from supplier and external R&D. Thus, with increasing complexity, learning-by-interacting with suppliers and the exploitation of inter-industry spillovers become increasingly important. This in turn emphasizes the significance of absorptive capacity.

2.5.2.2 Learning-by-Doing

According to Rycroft and Kash (1999), U.S. networks that successfully innovate complex technologies have emphasized experience over theory. The concept of learning-by-doing evolved out of the long held idea that individual performance improves as a worker gains experience with a particular task (Argote and Epple 1990). As individuals become more proficient with a task, it seems likely that a manufacturing team would, on average, become more productive. This concept of organizational learning-by-doing can be extended to include experience with product designs, manufacturing processes and

[14] In some mature 'science based' industries, such as pharmaceuticals and biochemicals, established firms engage in continuous new product R&D. In such cases, industry and product maturity cycles do not align.

machinery (learning-by-using), and material inputs. Innovations from learning-by-doing include changes in any of these factors, as well as changes in plant layout, firm organization, inventory control, and supply and distribution networks and methods. The results of these changes include increased labor and equipment productivity, reduced material costs, improved plant capacity or uptime, decreased manufacturing defects, improved product quality, and reduced time to market.

In learning-by-doing, process and product knowledge can be codified in texts, databases, operating procedures, and organizational structure and can be embodied in production plants, equipment, and products. However, a significant amount of knowledge is still tacit, existing in human capital and organizational culture. Thus, employee turnover and corporate reorganization can account for institutional 'forgetting' (Argote and Epple, 1990).

While learning-by-doing is often referred to as moving up the 'learning curve,' it is important to distinguish between the process of learning-by-doing and the studies that provide statistical evidence of the process. Early researchers recognized that learning-by-doing could have significant implications for manufacturing as direct-labor personnel gained experience performing repetitive assembly-line tasks. Wright (1936) is generally credited as the first to publish findings that per unit manufacturing labor hours decreased with cumulative production. Early applications focused on military applications (Alchian, 1963; Rapping, 1965), but thousands of studies in diverse industries have examined the relationship between labor productivity, per unit costs, or average costs, and cumulative production, capital equipment, or some other measure of experience. This relationship, variously termed the learning curve, progress function, or experience curve, has proven to be quite robust. Dutton and Thomas (1984) review 100 studies and find statistical evidence that a doubling in cumulative production is associated, on average, with a reduction in production costs to 80% of the initial value (a progress ratio of 0.80).

While experience curves provide a methodologically tractable method of modeling endogenous technological change, care must be taken in interpreting these statistical studies. The literature on progress rates demonstrates high variation both across *and within* industries, firms, processes and products (Dutton and Thomas, 1984; Hirsch, 1952; Hirsch, 1956; Joskow and Rose, 1985). According to Dutton and Thomas (1984) "progress curves are aggregate empirical descriptions of a process, and they mask its underlying dynamics." Organizational learning is more complex than individual workers improving their manual facility. Improvements in coordination, scheduling, plant organization, production processes, and manufacturing technologies can increase labor productivity, decrease production costs, decrease time to market, or improve product quality. Therefore, learning-by-doing encompasses experience gained at all levels in the organization, including engineers, managers and even sales and marketing staff (Hirsch, 1952; Baloff, 1966; Yelle, 1979; Montgomery and Day, 1985; Argote and Epple, 1990). This learning may or may not be represented in the learning curve, depending on how progress and experience are measured.

At the same time, reductions in both direct-labor hours and unit costs may be realized through other factors that are correlated with time and therefore with cumulative output, such as increasing capacity, cumulative investments in internal R&D, and progress in other industries and basic science (Ghemawat, 1985; Day and Montgomery

1983). Therefore, so-called experience curves may conflate learning effects with economies of scale and the influx of new knowledge that is both internally and externally generated.[15] A number of studies that control for these variables find that measures of cumulative experience are still statistically significant (Rapping, 1965; Hirsch, 1952; Joskow and Rose, 1985; Lieberman, 1984; Zimmerman, 1982).

Finally, the blanket application of simple experience curves assumes that all firms and industries are capable of attaining the average progress rate without any effort; it treats progress as a "free good". However, the variation in progress rates found in the literature supports the idea that firms must expend resources to exploit knowledge from the environment. In a study of progress in the chemical industry, Leiberman (1984) found statistical evidence that R&D expenditures steepened the learning curve. In addition, organizational structure and culture can either thwart or facilitate learning. Individual learning is not enough; continuous learning must occur across the organization; "the ability to integrate individual learning into system-wide learning" is an important factor in failure or success (Rycroft and Kash, 1999).

As I review the history of the automobile, I will attempt to distinguish among learning-by-doing, other sources of knowledge and learning, and economies of scale. The influence of organizational structure and corporate structure on learning will be examined and changes in these factors will be considered as learning-by-doing.

---

[15] In his foundational research, Wright (1936) attributed the reduction in labor costs to "improvement in proficiency of a workman with practice" as well as to economies of scale, such as spreading the tooling and set up time over larger production volumes, the opportunity for greater tooling, and the "ability to use less skilled labor as more and more tooling and standardization of procedure is introduced."

2.5.2.3 Learning-by-Interacting

Learning-by-interacting can be classified in two categories: learning from users and learning from suppliers. Learning-by-interacting represents a fundamental aspect of the evolutionary dynamic: outcome of selection processes must feed back into the next round of variety generation. While the idea of 'selection processes' is typically interpreted narrowly as market mechanisms, here were refer to any framework in which manufacturers and suppliers or users interact in order to choose between competing patterns of behavior (Metcalfe, 1995).

Feedback from users is particularly important during the emergent phase of a product and industry. During this phase, such feedback establishes the product's functional requirements as well as users' preferences for product attributes that fulfill them. Even after the emergence of a dominant design, firms must continuously interact with users due to the ever-present potential for shifts in user preferences. However, Hayes and Abernathy (1980) caution that over-reliance on market analysis tends to lower the overall level of product innovation and leads to imitative rather than innovative product design:

> It may be useful to remember that the initial market estimate for computers in 1945 projected total worldwide sales of only 10 units. Similarly, even the most carefully researched analysis of consumer preferences for gas-guzzling cars in an era of gasoline abundance offers little useful guidance to today's automobile manufacturers in making wise product investment decisions. Customers may know what their needs are, but they often define those needs in terms of existing products, processes, markets and prices. Deferring to a market-driven strategy without paying attention to its limitations is, quite possibly, opting for customer satisfaction and lower risk in the short run at the expense of superior products in the future.

In the earliest phase of development, entrepreneurs are building one-of-a-kind prototypes and are likely manufacturing their own parts and equipment, or modifying off-the-shelf parts to fit a new use. As the product and its market evolve, it becomes profitable to purchase supplies, parts, and equipment from specialists who can profit from economies of scale and in-depth knowledge and experience. Therefore, mature manufacturing industries typically are mutually supported by a large network of suppliers of raw materials, parts, subassemblies, process equipment, and complementary products. When these specialists support multiple industries, they serve to diffuse knowledge and innovation across sectors. For this reason, firms that follow a strategy of backward integration may isolate themselves from the R&D of independent suppliers who now view them as a competitor. Such a strategy may eliminate supplier power and supply chain bottlenecks and therefore reduce costs in the short run. However, it may also slow innovation by diverting the company's resources from its core business and can lock the firm into outdated process technology (Hayes and Abernathy, 1980). At the same time, firms should not leave innovation of supplies and process technologies entirely to their manufacturers; rather, the interaction must involve information flows in both directions with mutually supporting research programs.

2.5.2.4 Spillovers

Spillovers refer to innovations that originate outside the sectoral system – in other industries, institutions or other countries. Spillovers may involve either a transfer of knowledge (knowledge spillovers) or economic benefits (rent spillovers). Rent spillovers are realized when an innovation in another (supplier) industry results in decreased costs or increased quality or performance of the material inputs or the capital equipment used by the subject industry. If the entire cost saving is not passed on to consumers of the final good, the manufacturer of that product realizes some profit from the innovation. In this research, we are primarily concerned with knowledge spillovers which lead to technological progress in the automotive industry.

However, the distinction becomes somewhat problematic with a complex product and production system. Here, the relationship between manufacturers and suppliers is dynamic and more complicated than a simple market transaction. Suppliers may provide entire subsystems which may be jointly designed. Even ‘simple’ improvements in material inputs, like new metal alloys, may require iterative problem-solving by the supplier and manufacturer to develop product-unique specifications and appropriate product designs, manufacturing processes, and machine tools. Therefore, rent spillovers, knowledge spillovers, and learning-by-interacting with users become difficult to differentiate.

Knowledge spillovers from other manufacturing industries, scientific research institutions, and technological research institutions outside the sector require that the organization be capable of exploiting knowledge from outside the firm. While some externally generated innovations may be directly used in existing products (direct spillovers), others (indirect spillovers) must be adapted and existing product designs may require modifications. Thus, manufacturers must be able to identify and synthesize promising innovations. This capacity is facilitated by an active research department and effective linkages with the knowledge base beyond the individual firm (Cohen and Levinthal, 1989).

2.6 Analytical Framework and Approach

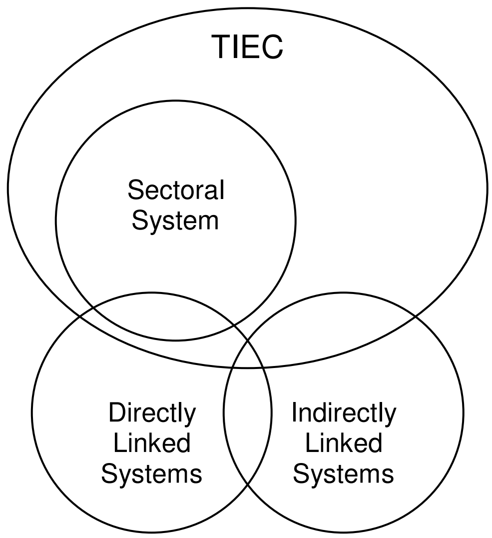
The research questions posed in Chapter 1 follow from the assertions of the QEM and MLP theories of technological transitions, mainly those regarding the role of niche markets and the technological nexus. Section 2.4.4 identified a number of shortcomings of these theories which could lead to incomplete or flawed conclusions. Therefore, in this section I integrate the main concepts of the QEM and MLP with concepts from the additional research reviewed in section 2.5 to construct an analytical framework that accounts for these difficulties. In this manner, I hope to take advantage of the explanatory power of each body of research while compensating for any theoretical and practical oversights. This dissertation does not assert that this framework represents a new model of innovation or a complete theory of technological transitions. Rather, this conceptual framework provides the structure and vocabulary to trace the historic development of motor vehicle technology and its diffusion to mass markets. However, because of the uniqueness of this approach, the findings of this research have important implications for the theory of socio-technical transitions.

Because of the multidirectional interactions among elements in complex systems and between the system and its environment, determining the system boundaries for study is a difficult task. While the QEM and MLP take a technological regime as the unit of analysis, it is not entirely clear prior to analysis what actors, organizations, and artifacts are relevant. Because the research questions posed in chapter 1 address the coupling between variation and selection identified in QE theory, both agents involved in supply (innovators and producers) and demand (users) must be considered. However, the research questions also specifically address the innovative processes involved in improving the performance and cost of early motor vehicles. Thus, the analytic framework must adequately account for the behavior of entrepreneurs (agents of change) and how they gather, interpret, and respond to information from the selection environment. Therefore, I take the perspective of the entrepreneurs and trace their actions and interactions through history. This viewpoint provides the structure to analyze both the development of the technological nexus and the learning mechanisms involved in transitioning the automobile from niche to wider markets.

The conceptual framework used here is based on a systems view where "technology is best understood in terms of its functional attributes" (Sahal, 1981). The system of interest provides the societal function of personal transport. To begin constructing this framework, I appeal to the sectoral system of innovation and production described by Malerba (2002) as "a set of new and established products for specific uses and the set of agents carrying out market and non-market interactions for the creation, production and sale of those products. A sectoral system has a knowledge base, technologies, inputs and an existing, emergent and potential demand. The agents composing the sectoral system are organizations and individuals (e.g. consumers, entrepreneurs, scientists)." Within the sectoral system, "[a]gents are characterized by specific learning processes, competencies, beliefs, objectives, organizational structures and behaviors. They interact through processes of communication, exchange, co-operation, competition and command, and their interactions are shaped by institutions (rules and regulations). Over time, a sectoral system undergoes processes of change and transformation through the co-evolution of its various elements."

Thus, the personal transportation system of innovation and production includes agents – individuals, firms, and units within the firms – that are carrying out market and non-market interactions for the innovation, design, production, marketing, sales, service (maintenance and repair), purchase, and use of new and established products for the provision of personal transport. This system, then, includes manufacturers of automobiles and automobile parts (direct suppliers) and automobile buyers and users. In addition, it includes physical artifacts – factories, equipment, distribution and service facilities, product and parts inventories, and stock in use. The relationship of the agents to the physical artifacts and the interaction of agents are shaped by a set of institutions. Institutions in this context are defined to include both organizations and rules (cognitive, normative, and regulative). They include any form of social construct or constraint on human interactions, including organizational constructs, such as governmental bodies and research and industry organizations; formal constraints, such as legislation, economic rules, and legal contracts; and tacit constraints, such as shared perceptions, beliefs, social conventions, and moral codes. In general, I will use the term institutions in this manner, resorting to the terms ‘rules’ and ‘organizations’ when the distinction is significant to the discussion. Thus, the sectoral system describes all three of the types of elements found in complex systems as identified by Geels (2004) – human actors and organizations, physical artifacts, and rules.

Geels (2004) asserts that the sectoral system of innovation inadequately considers users and this criticism falls within the broad problem that evolutionary theories generally neglect the sociological aspects of systems of production and innovation. This can be remedied by considering that the sectoral system is an integral part of a larger socio-technical system with a technological and social history, all of which is embedded in the natural environment. In other words, the sectoral system may be considered a subsystem of a larger complex system composed of actors, physical artifacts, and rules. Therefore, this analysis examines the innovative behavior of automobile manufacturers, automotive suppliers, and the manufacturers of precursor technologies embedded in a technical, institutional, and ecological complex (TIEC) as illustrated in Figure 2-3.



Sectoral System: Individuals, firms, and units in firms or other institutions engaged in the creation, production, sale, purchase, and use of products for the provision of a societal function (e.g. personal transport).

Directly Linked Systems:
Supporting industries and institutions.
Sectoral economic markets.
Natural systems that provide raw materials and absorb wastes.

Indirectly Linked Systems:
Apparently unrelated industries, institutions, and markets.
Remainder of ecosystem.

TIEC: also consists of higher level institutional constructs (e.g. legislation, cultural norms, etc.) and "non-linked" systems (industries and institutions) connected through diffuse associations.

Figure 2-3: The Sectoral System of Production and Innovation

While some elements of the TIEC are analogous to the socio-technical landscape in the MLP, there are important conceptual differences. Geels (2002) defines the socio-technical landscape as the macro-level which provides an external structure or context that is relatively 'hard,' which is to say that the landscape is difficult to influence and that change occurs slowly. It contains the material aspect of society, such as cities, factories, and highways, and heterogeneous factors such as oil prices, economic growth, wars, cultural values, and environmental problems. In contrast, the TIEC consists of both macro- and meso-level elements. For example, some of the elements of the landscape here are considered as part of the regime's system of production or as linked systems, such as factories and highways respectively. Meanwhile, the MLP's landscape does not explicitly account for the interaction between the regime under study and other meso-level systems. By incorporating overlapping (linked) systems, the TIEC accounts for interactions between regimes that are somewhat independent yet inter-related. In addition, meso-level systems that are not directly linked to the system under study may still influence it through diffuse associations within the TIEC. A fuller discussion should clarify these important distinctions between the MLP's landscape and the TIEC.

Like the socio-technical landscape in the MLP, the TIEC includes regionally or nationally established rules and higher level social constructs (*e.g.* cultural values and culturally learned patterns of behavior), among other elements. Thus, while the observable structure and linkages of the system are described by Figure 2-3, the institutions determine the interactions that occur through these links, *i.e.* the processes of communication, exchange, co-operation, competition and command. In dynamic systems modeling terminology, the institutions (rules) are analogous to the equations that determine the flow of energy, materials, and information. While a majority of rules are imposed by the TIEC, some rules exist that are specific to the sectoral system such as perceptions that are unique to the system's knowledge base and standard operating procedures specific to firms or industries.

Within the TIEC, the sectoral system is directly linked to supporting industries, markets, government and private institutions, and the natural systems that provide raw materials and absorb material and energy wastes. These directly linked systems are in turn linked to their own network of supporting industries, institutions, and natural systems, providing an indirect link from these 'apparently unrelated' institutions to the sectoral system of interest. For example, the sectoral system has a knowledge base that resides with the manufacturers, suppliers, universities, government agencies, trade associates, *etc.*, and is networked through the TIEC with knowledge bases in other sectors and industries. As a specific example, consider that members of the Society of Automotive Engineers (SAE) may also be members of the American Society of Mechanical Engineers (ASME), which includes engineers employed in other sectors and industry. And because all planetary ecological systems (water, air, land) are interconnected, there is a direct or indirect link between the sectoral system and the entire global ecosystem.

The network of linkages among systems and the TIEC allows for the feedback dynamics that are a crucial characteristic of evolutionary change and accounts for the two-way linkage between variation and selection identified in the QEM. Over time, the system undergoes processes of change and transformation through the co-evolution of its various elements and elements of the TIEC. For example, the increasing scale of automobile use in the U.S. contributed to local air quality problems by the 1970s. This change in a directly linked ecological system (atmospheric processes) led to a social change in the TIEC – increased public concern for the health effects of air pollution. In response, the actions of private and government agents led to the creation of new institutions for the control of vehicle emissions: the Clean Air Act and the U.S. Environmental Protection Agency (EPA). In effect, the success of the automobile had radically altered the selection environment, and this change led to new requirements for innovation. Actors within the government served the nexus role of interpreting the new demands of the selection environment and imposing them on the variation that occurred within firms. Actors within firms also participated in the nexus through negotiations and collaboration with government agents to establish standards and testing procedures.

Clearly, the structure of the TIEC and its subsystems is dynamic as new institutions and linkages emerge and others disappear. Meanwhile, the overlap in systems is only partial. For example, while the EPA has responsibility for promulgating and enforcing vehicle emissions standards, other divisions within the agency fulfill functions unrelated to motor vehicles. At the same time, the EPA has no jurisdiction over establishing new vehicle fleet fuel economy standards, though it is tasked with establishing the testing methods manufacturers use to measure the fuel economy of new models to verify compliance with these standards.[16]

Further, some elements of the TIEC can influence the sectoral system with no apparent feedback in the reverse direction, while other elements have no identifiable connections to the sectoral system. Although there is no traceable link between these systems and the sector of interest, they are connected through associations, often diffuse, with the TIEC. These elements can be considered the sector's environment and are somewhat analogous to the socio-technical landscape described by Geels (2002), though they are not necessarily all as 'hard' and unchangeable, nor are they necessarily at the macro-level. Dynamics within these systems can effect changes within the TIEC that are exogenous to the sectoral system, but which exert influence on it. This dynamic, in effect, changes the boundary conditions for the system under study. For example, U.S.

[16] The National Highway Safety Administration is tasked with reviewing and establishing the Corporate Average Fuel Economy (CAFE) standards. The fuel economy of each manufacturer's new vehicles, averaged over its entire sales fleet, must meet or exceed the CAFE standard, or the manufacturer must pay fines.

involvement in both World Wars – events with no causal links to the transportation sector – required, among other things, the re-allocation of productive resources to the war effort. The actions of governmental agents with no prior relationship to the transportation industry curtailed the production of automobiles. In addition, automobile manufacturers became involved in the design and production of military supplies such as aircraft engines, creating new linkages to a formerly independent industry.

The actions of the agents and the process of change are studied within a cyclic framework, illustrated in Figure 2-4, based on the industry life-cycle model described in section 2.5.1 and elements of the cyclic model of innovation described by Anderson and Tushman (1990). The sectoral system (technology, industry, firms and markets) progresses through four phases: emergent, transitional, specific, and senescent. The general characteristics of these phases are summarized in Table 2-1. The specific phase is characterized by the emergence of a dominant product design, which represents a stable description of the functional requirements of the technology and the product attributes used to fulfill them. In addition, the specific phase entails stable production processes and organizational forms, which to some extent are embodied in rules. Stability in this case is understood to be dynamic, where technological change occurs incrementally without causing discontinuous disruptions to production processes and usage norms.

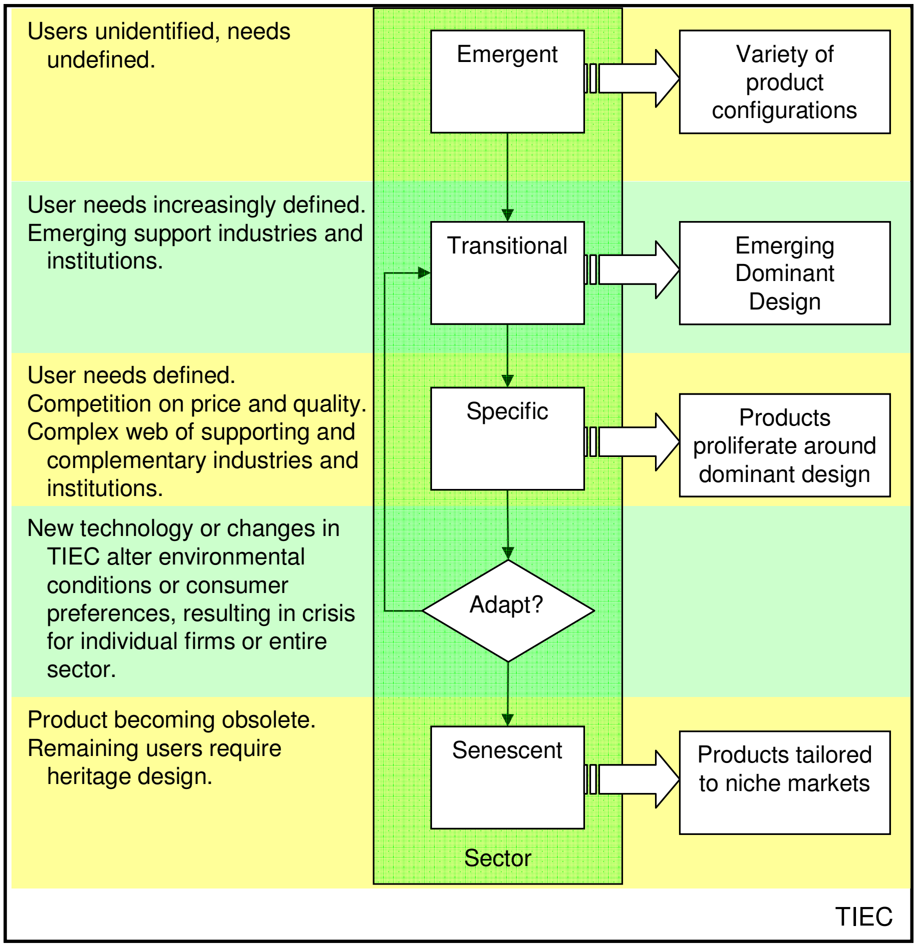


Figure 2-4 Cyclic Co-evolutionary Framework

I borrow from Anderson and Tushman (1990) the idea that innovation is cyclic; after an industry attains the specific phase, incremental innovation may be punctuated by periods of 'ferment.' This is represented by movement from the specific phase into a new transitional phase. Anderson and Tushman, however, describe these periods as being initiated by stochastic technological breakthroughs (see also Tushman and Rosenkopf, 1992). In the framework presented here, cyclic evolutionary change may be triggered by *any* change in the TIEC – technological, economic, social, institutional, or environmental – which may be brought about through co-evolutionary interactions between the transportation sector and other elements of the TIEC or by causes entirely exogenous to that sector. Thus, through the TIEC, the cyclic framework incorporates the sociological dimensions of technological change found in the QEM. Non-technical phenomena (social, political, economic) may create instability which presents opportunities for innovation or even forms the basis of the innovation of new production processes or usage norms. Changes that trigger innovative ferment may alter the firm's operating environment or result in a redefinition of users' needs and therefore the product's functional requirements in an evolutionary or discontinuous fashion.

Entrepreneurs, within or outside the sector's existing firms, can exploit the change in the TIEC to create economic value by modifying existing products and processes or developing new ones. Existing firms in the industry must adapt through changes in their products, processes, and organizational structures or face obsolescence and senescence. Successful innovation eventually moves the sector toward another specific phase awaiting the next round of innovation. Finally, when an innovation that can fulfill the new functional requirements arises outside the existing technological paradigm, the entire sector enters a phase of senescence while the success of the innovation gives rise to a new sectoral system. Note that in this framework, the phases of development are a continuum with indistinct boundaries and the sector's characteristics may not align perfectly with those listed in Table 2-1. Further, the system may not attain the characteristics of the specific phase as diagramed in Figure 2-4 before another change in the TIEC triggers additional innovation and transition.

In the cyclic model proposed here, changes in the TIEC that trigger a shift to a new transitional phase or to the senescent phase can be placed in one of two categories: 1) new scientific discoveries, technological innovations, or institutional arrangements that arise outside the existing technological paradigm and stimulate the development of competing technology; or 2) difficulties advancing along the existing technological paradigm due to problems created by its success, its expanding scale and scope, theoretical limits, or exogenous factors. These two classifications are conceptually consistent with Geels' (2005a) 'technological substitution' and 'wider transformation' routes; Gharjedaghi's (2006) shift of paradigm by design or out of frustration; and Dosi's (1982) theory that radical shifts along new technological directions "emerge either in relation to new opportunities opened-up by scientific developments or to the increasing difficulty in going forward on a given technological direction" for technological, economic, or other reasons. In both cases, the change in the TIEC may manifest itself in one of three ways: 1) through the market via altered consumer preferences for product attributes, the price of input materials or waste disposal (*i.e.* the cost of the factors of production), or competition with other technologies; 2) through other institutional changes such as taxation or regulation; or 3) through direct environmental effects such as the impact of weather patterns on production processes or end use.[17]

The cyclic evolutionary model accounts for the increasing complexity of products, production systems, and support networks through the cumulative effects of innovation through successive cycles. Also, change may occur at any level – component, subsystem, or system – while the other levels remain stable. For example, with a complex technology like the automobile which is produced in a complex industry structure, there can be cycles of radical innovation and emerging and senescent markets

[17] For the case of the current personal transportation system, most environmental effects are manifested through input prices, taxes, and regulation. If one considers the increasing frequency and severity of storms, a case might be made for climate change impacts on end use. Other sectoral systems, notably agriculture and renewable energy systems, obviously have large direct environmental effects which are accounted for in the general model.

for component technologies, even while the basic architecture of the consumer product that supplies personal transport remains unchanged. However, the functional requirements it fulfills and the attributes used to do so may be altered significantly. When this is the case, a new dominant design emerges. Similarly, while the industry structure of automobile manufacturers may appear stable over a long period of time, their organization, production processes, and interactions may be radically altered, along with the network of supporting industries and their relationships with manufacturers. Thus, the network of linkages and the rules guiding these interactions may change while the observable structure of the system appears unaltered.

The phase typology of the cyclic model provides a structure for analyzing the emergence of new technologies, the development of a socio-technical regime, and the dynamics of niche-regime interactions, while also stressing the central role of stability. The QE process of embedding that occurs with movement from the emergent phase to the specific phase (and the movement from niche to mass markets) results from the co-evolution of the technology, the sectoral system, and the TIEC. As shown in Table 2-1 and Figure 2-4, this co-evolution results in the emergence of a complex web of complimentary and supporting industries, institutions and infrastructure. Using concepts from the MLP, a socio-technical regime is understood to exist when a technology and sector have achieved the specific phase and when the stability of that specific phase is supported by the alignment of institutions, including rules, and existing physical infrastructure with the dominant product design, production processes, and organizational forms. This alignment confers advantages to the existing technology over alternatives. Changes in the TIEC result in a misalignment that erodes stability and provides opportunities for an emergent technology to compete with the existing technology, possibly pushing it into senescence as described above.

Alternately, such a misalignment could lead to technological transition within an existing regime. While new functional requirements arising from the misalignment may be met by a radically new technology, they may also be met by a new product design that emerges within the regime. What distinguishes transition *within* the regime from transition to a new regime is that internal transition involves a great deal of adaptation versus replacement – of firms, industries, knowledge bases, infrastructures and institutions. This is not to say that there is *no* replacement, as some firms or even entire supporting industries may be eliminated by the transition and some new infrastructure and institutions may be required, but internal transition is, overall, less disruptive to the regime. The regime and the technology remain identifiable. Transition to a new regime, or 'regime renewal,' as described by QE theory, involves the simultaneous and relatively abrupt change to both a radically new technology and to a new rule set. It requires considerable change in "the behavior of various actors in relation to various relevant technologies," such as new usage norms and new forms of ownership (Elzen et al., 2003). The description of innovation in the cyclic model then allows for the transition to a new regime to occur as a gradual unfolding or reconfiguration of the existing regime as described by Geels (2002). In internal transition, a new product design may integrate radically new technologies with the existing technology in a manner which allows for compatibility with existing rules and infrastructure. Over a longer time period, this may lead to the gradual emergence of a new regime through co-evolutionary adaptations.

This analytic framework will be used to review the early history of motor vehicles. The analysis will focus on the behavior of entrepreneurs (agents of change) and explore the development of the technology, the establishment of the network of agents involved in variation (the entrepreneurs) and selection (consumers, policy makers, and society at large) and the two-way linkage between the two processes through the technological nexus role. In particular, I will attempt to identify the interactions that produced the technological and institutional adaptations that facilitated wider diffusion of the automobile. This includes identifying significant innovations and, using the vocabulary introduced in section 2.5.2, exploring the learning processes underlying these innovations. I will also attempt to generally delineate the sector’s phases of development based on the emergence of dominant designs, mass markets, and regime stability, recognizing that the boundaries of these phases are not necessarily distinct.

3.0 Emergence

Technological innovations which may have market application, lie fallow until markets can be identified or created.
-- J.M. Utterback and W.J. Abernathy, 1975

Prior to the development of the motor vehicle, personal transportation was provided by what may be considered a socio-technical regime based on draft animals, primarily horses. This regime had existed for thousands of years and was characterized by a set of rules and aligned industries for the production, ownership, use, and maintenance of vehicles (carts, wagons, carriages, *etc.*) as well as the breeding, housing, care, and training of draft and riding animals. These rules also included the construction, maintenance and usage of urban roads. The regime consisted of two separate but linked systems of production and innovation – vehicles and draft animals – that were characterized by incremental progress.

However, the ancient Greeks first conceived of and experimented with alternative sources of power as early as the second or first century BC, thus marking the *invention* of the motor vehicle. It took some two thousand years for this invention to become an innovative product with practical application and economic significance. This chapter reviews the history of the technologies involved in this innovation and includes the emergent phase of motor vehicles. Section 3.1 discusses the development of the underlying technologies for motor vehicles powered by steam engines, electricity, and gasoline internal combustion engines. Section 3.2 reviews the historical development of manufacturing processes for volume production and standardized interchangeable parts, two necessary components of mass production. Section 3.3 relates the history of automobile manufacturing during the emergent phase, which concludes around the 1900. Section 3.4 provides the context of developments in the technical, institutional, and ecological complex (TIEC).

Section 3.5 then synthesizes and analyzes this history and context within the framework presented in section 2.6. I identify two initial niche markets for motor vehicles – local urban transport and long-distance pleasure touring. I determine that misalignments in multiple regimes combined with several developments in the TIEC explain why the motor vehicle industry emerged in these markets when it did. I find that spillovers and learning-by-doing appear as the primary mechanisms of innovation in a cumulative and synthetic process during this era. I also identify four examples of the emerging technological nexus role which links variation and selection.

3.1 Technologies for Motive Power

3.1.1 Steam Powered Vehicles

The concept of steam power dates back to ancient Greece, but it was another 1700 years before the technological possibilities were fully grasped. During the Renaissance, the revival of interest in classical approaches to understanding the physical world led to two significant developments: the scientific method, a process of discovery relying on empirical evidence and mathematics; and a new mechanical philosophy which held that all natural phenomena can be explained by physical causes. In 1680, Sir Isaac Newton envisaged a steam carriage and predicted that steam power would revolutionize transportation. The earliest patents on practical applications of the principles behind steam power were obtained shortly thereafter, and engines for pumping came into widespread use in England during the first half of the 18^{th} century. James Watt's revolutionary inventions between 1762 and 1782 brought the application of steam power to other uses. Watt described a steam locomotive in a patent in 1784 but never constructed one and, in protection of his patents, actually hindered further development of steam engines using high-pressure.

Nicholas Joseph Cugnot of France constructed the first motorized vehicle in 1769, a three-wheeled gun tractor powered by a two-cylinder steam engine. The vehicle was inefficient and slow and Cugnot was driven into exile by the French war ministry which had funded the project. Although Napoleon later brought Cugnot back to France to continue his work, the project was never completed. In 1801 and 1802, Oliver Evans and Richard Trevithick built high pressure ‘road engines’ to pull heavy loads on tracks in the U.S. and England, respectively. In 1825, the first public railway opened in England and successful steamboats applied Watt engines in both countries early in the century (Nevins and Hill, 1954).

By the 1830s, stationary steam engines were being used to power U.S. threshing machines, sawmills, and gristmills. Portable steam engines that were mounted on wheels and could be pulled by horses came into use on farms in the 1850s. They became particularly popular for ‘belt work,’ mainly threshing, in the 1870s, though advertising was required to overcome some farmers’ well-founded fears of boiler explosions. The machines were more economical than horses, since they required no veterinary care and wood for fuel and sperm oil and beef tallow for lubrication were readily available. Experimentation with steam-powered plows began in 1856, but these trial and error efforts resulted in machines that were far too heavy for practical use. Wik (1951) blames this failure in part on the absence of steel tools and lack of access to schools of agricultural engineering, though overly ambitious design goals also contributed. The first successful self-propelled, or traction, engine was introduced in 1873 when an Ohio manufacturer added a gear attachment to a portable engine.[18] The traction engine was still dependent on horses for steering until the 1882 introduction of a self-steering model. Steam power use in agriculture reached a zenith in 1910 at around 3.6 million horsepower. Manufacturing of steam farm engines peaked at 10,000 in 1913 but was essentially abandoned by 1925, having been replaced by the gasoline tractor (Wik, 1951).

Attempts to use steam engines on U.S. roads initially failed because the vehicles were too heavy to negotiate the country's poor roads. Sylvester Roper was the first American to produce lightweight steam cars, some of which weighed no more than 500 pounds. Roper built around ten experimental road steamers using coal-fired boilers between 1860 and 1895. Following Roper's death in 1896, George Eli Whitney, who had worked occasionally in Roper's shop, built his first steam carriage using kerosene to heat the boiler. Whitney founded the Whitney Motor Wagon Company in 1897 and produced and sold a few automobiles weighing about 800 pounds for $1500 and up (May, 1990).

In 1898, the Stanley brothers, Francis and Freelan, built a 600-pound steamer using a lightweight engine built by J.W. Perry & Sons and a light, high-pressure boiler of their own design. Frances won a race in the steamer in October, 1898, and Freelan drove one over unpaved rocky roads to the summit of Mount Washington the next year, marking the first time an automobile had accomplished that feat. Encouraged by their success and the interested it garnered, the brothers formed the Stanley Brothers Motor Carriage Company in late 1898 and began production of 100 steamers. The company and

[18] The steam traction engine, or road locomotive, was used to plow, pull heavy loads, or provide power at a chosen location. It was used extensively in agriculture and is the forerunner of the steam tractor. Though powerful and self-propelled, it was heavy, slow, and poorly maneuverable.

the Stanleys' patents were promptly bought out and the firm was renamed Locomobile. The patents were later sold back to the brothers who began building an improved version of the steamer in 1901. A specially built Stanley automobile set a speed record of 127.6 miles per hour (mph) in 1906. A year later, the racer became airborne at a speed of 197 mph, seriously injuring the driver, and the brothers abandoned attempts to set speed records (May, 1990).

More than 100 U.S. manufacturers produced steam cars based on the Stanley-Locomobile design, though most firms were short-lived. In 1901, Rollin H. White began producing a car based on the Stanley but he incorporated a 'semi-flash' boiler and a condenser. White also gradually increased the working steam pressure, nearly doubling it by 1910. White sold 9,122 vehicles before abandoning steam power in 1911. The market for steam-powered automobiles began a rapid decline around 1910, but the Stanleys continued to make improvements to their automobiles, shifting from gasoline to kerosene, adding weight and strength, and, in 1915, adding a condensing system. The brothers' reputation allowed them to continue selling several hundred cars per year. Many analysts have asserted that the Stanleys could have sold many more vehicles if they had shifted to mass production techniques and more aggressive advertising. However, the brothers preferred to stick to traditional craft techniques, believing it produced higher-quality workmanship, and maintained annual production below 650 cars. The brothers sold the company in 1917, but Stanley steamers were produced until 1924, with cumulative production exceeding 14,000 vehicles (May, 1990).

Steam-powered automobiles were designed to operate using gasoline or kerosene, which was widely available at rural general stores in the 1890s. And in an emergency, steam boilers could use nearly any fuel. Prior to the incorporation of condensers, steam cars also required frequent refills with clean, preferably soft, water, which could be obtained in many areas at public watering troughs that were made available for horses. Steam engines could easily provide adequate and economical power, provided a wider range of operation than electric motors, and were less mechanically complex than gasoline internal combustion engines (ICE). They did not stall, permitted smooth transmission of power to the wheels, and were easier to manufacture. However, the boiler required sufficient time to heat up, then required frequent attention during operation and posed serious risks of explosion. The development of the flash boiler overcame many of these difficulties, but steam car technology reached a plateau in development after 1901, just as the gasoline automobile began rapid improvements (Flink, 1970, p. 236).

Although steam-powered automobiles did not ultimately succeed in the market, technological developments in this arena made significant contributions to the development of gasoline ICE vehicles. The designers of the steam traction engine had developed basic solutions to the problems presented by transmitting power from the engine to the wheels, changing speed, and steering. In addition, the first gasoline ICE engines adopted from steam power the concept of a 'doubling acting' cylinder: the practice of applying expanding gas to one side of the piston to move it in one direction, then applying another charge to the other side to return it to the original position (Nevins and Hill, 1954).

3.1.2 Electric Powered Vehicles

Experimentation with electricity traces back at least to the ancient Greeks who first observed properties we now call static electricity. In 1800, Alessandro Volta created the first chemical battery, consisting of a pile of discs of copper and zinc separated by brine-soaked cardboard. The first batteries for practical use, developed in the 1830s, used a liquid electrolyte which was prone to spillage, making them impractical for portable or mobile applications. Dry cell batteries were invented near the end of the 19th century and replaced the liquid electrolyte with a paste. Both the wet and dry cell batteries were permanently drained when the chemical reaction was spent. In 1859, Gaston Planté invented the lead-acid battery, the first battery that could be recharged by reversing the current through it. In 1881, Camille Alphonse Faure developed an improved lead-acid battery with better performance that was also easier to produce in quantity. In the 1880's, massive storage batteries were being used in the first electrical power distribution systems to augment power at peak usage, smooth current output, boost output at remote substations, and provide power during outages.

Public transportation in cities was initially provided by the horse-drawn omnibus which was replaced by the horse car.[19] These vehicles were limited in their ability to negotiate steep grades and contributed to problems with animal wastes, odor, flies and disease. Humans risked infection with tetanus, while close quarters spread infectious diseases rapidly among urban horse populations. A major outbreak of an equine infectious disease in 1872 proved devastating to most eastern cities. In the 1880s, entrepreneurs experimented with steam-powered cable cars with little success.

[19] The horse car looked much like its successor, the cable car, but was pulled by horse.

Electrification of urban transport in the U.S. began in 1888 with the installation of Frank J. Sprague's electric streetcar system in Richmond, Virginia. The system used a generator and overhead power distribution, which minimized the energy losses from friction that plagued steam-powered cable systems. Electric streetcar technology diffused rapidly, and by 1901, $2 billion had been invested in 15,000 miles of track. By 1902, 97% of urban street railway systems were electric. Because of inherent economies of scale, the industry was soon concentrated in the hands of large firms. These firms were often granted long-term or perpetual franchises by local authorities, leading to public controversy over control of city streets and government corruption.

Construction of new lines was terminated abruptly with the Panic of 1907 (see section 4.5.1), after which only a few systems were extended. Though safe and comfortable, streetcars required high initial investments and suffered from system inflexibility. While the systems were capable of supporting heavy use, few urban centers generated such high traffic. By 1917, the industry was caught between rising costs and rigid fare structures. Passenger miles reached a peak in 1923 and the industry went into decline (Hilton, 1969).

The earliest U.S. experiments with electric powered personal vehicles were undertaken by Andrew L. Riker, who built an electric tricycle in 1884. He formed the Riker Electric Motor Company in 1888 to produce electric motors then established the Riker Motor Vehicle Company in 1898 to manufacture electric cars. William Morrison of Des Moines, Iowa built the first successful four-wheeled electric car in the U.S. in 1891 and drove it on the streets of Chicago the following year (Rae, 1959). Morrison apparently did not pursue commercial production of the vehicle. The Pope

Manufacturing Company, the nation's largest bicycle manufacturer until 1896, began producing both electric and gasoline ICE cars in 1897. Within two years, Pope had built 500 electric and 40 gasoline ICE vehicles and was the leading U.S. automobile manufacturer.

In 1894, Henry G. Morris and Pedro G. Salom of Philadelphia produced an electric car called the Electrobat consisting of a wooden-wheeled wagon to which masses of lead-plate batteries and electric motors were attached. The two had spent ten years developing the vehicle with the support of the Electric Storage Battery Company (ESB). In 1895, the Electrobat II was awarded a gold medal for excellence in design by the judges of the Chicago *Times-Herald* race. However, the Electrobats were not capable of the long-range travel required to complete the race.

Morris and Salom created the Electric Carriage and Wagon Company in 1896 and the first experimental vehicles were ready the next year. However, the two engineers felt that untrained drivers were not capable of servicing and maintaining electric vehicles and decided instead to operate them as a fleet of public cabs in New York City beginning in March, 1897. By this time, Morris and Salom had relinquished ownership and managerial control to executives of the ESB. Six months later, ESB's Isaac Rice formed the Electric Vehicle Company (EVC) with the intention of expanding the cab service to include one hundred vehicles, a monumental task for a nascent industry employing craft production methods. By January, 1899, forty-five electric cabs were in regular service, and as many as forty more were leased under long-term contracts (Kirsch, 2000).

Of the three automobile types, the electric car most closely resembled horse-drawn vehicles in terms of performance and styling. It appears that electric vehicle manufacturers conceived of the automobile as a direct replacement for the horse-drawn buggy, providing exactly the same functionality but without the disadvantages (feeding and stabling requirements, manure production, flies, odor, *etc*.). According to Flink (1970, p. 238), the Woods Motor Vehicle Company eschewed the production of an automobile capable of touring, and instead planned to sell electric cars as “fine carriages in all variety of styles” for “gentlemen’s private stables.”

Battery powered electric vehicles could only travel a short range between recharging (initially 20 miles), recharging was time consuming (two to three hours), and operation was more expensive than steam and gasoline internal combustion alternatives. Batteries yielded poor hill climbing ability and their weight made negotiating mud and snow difficult. In addition, the batteries were unable to withstand the punishment of traveling on rough roads and rapidly deteriorated. However, the electric automobile was superior to horse-drawn, steam-driven, and gasoline internal combustion vehicles in terms of noise, odor, and ease of operation. It especially appealed to women who had difficulty with shifting gears in the early transmission found in internal combustion cars. Around the turn of the century, the general consensus was that electric vehicles were the best choice for travel in and around town. Enthusiasts were optimistic that improvements in battery technology and recharging or battery exchange infrastructure would expand their use to even enable touring.

In the summer of 1899, news out of Europe cast a shadow on public optimism regarding the electric vehicle. Electric cab companies in London and Paris announced the discontinuation of service, claiming difficulties retaining employees. Shortly afterward, the Automobile Club of France released discouraging initial results from tests of battery endurance. These reports coincided with efforts in the U.S. by William C. Whitney to corner the electric vehicle market. The financier and transportation magnate had cobbled together New York's Metropolitan Street Railway and had begun electrifying the system. In the process, he had also obtained interest in several of the city's electricity providers. In 1899, Whitney gained control of both the EVC and its parent company, the ESB, through stock purchases on the open market. Shortly afterward, he negotiated the purchase of Pope's automobile manufacturing enterprises. Whitney intended to create an electric vehicle industrial combination that would supply electric vehicles to a nationwide organization of regional companies that would in turn operate cab fleets and sell electric vehicles to the public. The new holding company formed a manufacturing subsidiary named the Columbia and Electric Vehicle Company from the production facilities acquired from Pope and the EVC. Operating companies were established in New York, Atlantic City, Boston, Philadelphia, and Chicago.

Editorials in a leading trade publication, *Horseless Age,* scathingly attacked the plan, asserting that the formation of a motor vehicle trust was premature and only in the interests of the investors. The magazine attacked both the storage battery and the EVC, pointing out the 'inherent weaknesses' of the technology and protesting that the cab scheme stood in the way of progress. Other trade publications admonished the editor of *Horseless Age* for misleading its readers with inaccurate information. By January, 1900, the magazine's position on electric vehicles had softened, but public opinion had already turned against Whitney's enterprise which was derisively named the Lead Cab Trust. Perhaps most significantly, the EVC in 1899 also purchased and attempted to enforce the Selden patent, leading to a number of patent infringement lawsuits against manufacturers of gasoline ICE automobiles that dragged on through 1910 (see section 4.5.2). According to Kirsch (2000, p. 32), the EVC "became synonymous with trust building, stock jobbing, financial manipulation, and legal chicanery."

Despite the damaging press, the Columbia and Electric Vehicle Company became the country's leading motor vehicle manufacturer, producing around 2,000 electric cabs in 1899. Although this expansion in production was admirable, it fell far short of the Lead Cab Trust's ambitious goal to produce 10,000 electric cabs. In addition, the company failed to unify operations or improve the production process, continuing to rely on carriage production techniques. According to Kirsch (2000), EVC customers were reportedly conservative, desiring vehicles that were consistent in appearance with those to which they were accustomed – carriages.

The Lead Cab Trust was plagued by problems regarding finances, production, maintenance, and operations. In December, 1899, an unrelated scheme to monopolize the electric vehicle industry led to a speculative bubble that inflated the stock price of the EVC then dragged it down when the bubble collapsed. In addition, the first vehicles put on the road apparently had been rushed into service and performed poorly due to design defects and improper assembly. Renters in Boston drove the vehicles carelessly and reports of reckless driving were common. Due to inadequate inspection and maintenance, more than half of the vehicles were inoperable after four months of service. Operators in Atlantic City experienced similar problems with maintenance, but were also plagued by an unreliable electric power grid that led to undercharged batteries. On some occasions, overnight power fluctuations resulted in the batteries discharging into the grid to power the town. Despite high demand for service, the electric cabs failed to show sufficient profits. When cab drivers in Chicago refused a profit-sharing plan and staged a strike in February, 1901, the manager chose to liquidate the company claiming the service was unprofitable. Boston followed suit within weeks and the EVC shut down operations in Atlantic City in May. The EVC was reorganized in December, 1901, resulting in the sale of the Philadelphia operation to the ESB. The New York cab fleet became a profitable independent firm and continued offering transportation on a fee-for-service basis until 1905.

Following the failure of the EVC, the electric vehicle market grew modestly, the technology remaining popular with wealthy urban women and in some commercial applications, particularly local delivery trucks. Meanwhile, the market for gasoline ICE automobiles expanded exponentially. Electric vehicles enjoyed a brief revival beginning in 1910, in part due to the efforts of the Electric Vehicle Association of America (EVAA) – a national organization of electricity providers, vehicle manufacturers, and the battery industry. To foster the market for electrics, particularly in commercial fleets, the EVAA pursued programs in education, advertising, standardization, and cooperation. A handful of central station operators established departments to provide technical support and maintenance to electric vehicle owners, subsidize local recharging facilities, and provide delivery and service vehicles for their own use. EVAA efforts resulted in the adoption of two standard charging plugs, one for passenger cars and one for commercial vehicles. With the increasing reliability of power distribution networks and the adoption of standard charging plugs, battery designs coalesced into a general standard by 1912.

However, efforts at cooperation soon foundered. Vehicle owners were unwilling to share operational cost information and prospective customers became suspicious of the data provided. The various factions within the EVAA disagreed on responsibilities, particularly with regard to promoting vehicle sales. Amid slow market growth, cooperation gave way to finger-pointing, and in 1916 the EVAA ceased to exist as a separate entity, having been absorbed by the National Electric Light Association (Kirsch, 2000).

3.1.3 Gasoline Internal Combustion Vehicles

The history of the gasoline internal combustion engine (ICE) automobile begins in Europe during the late nineteenth century. In 1801, French engineer Phillipe Lebon patented a two-stroke engine using electric ignition of a compressed illuminating gas made from coal. The engine was ‘double-acting,’ using explosions on both sides of the cylinder, a practice used in steam engines. Unfortunately, Lebon was assassinated in 1804 and his design was neglected for sixty years before Belgian mechanic Etienne Lenoir was inspired by the Lebon engine. Lenoir added a carburetor for mixing air and a liquid hydrocarbon to form a vapor, but he did not compress it. The design contained no new components, but its construction represented a striking advance over previous engines. Lenoir also benefited from the availability of better materials, machining techniques, and electrical equipment than Lebon had available sixty years earlier. In 1860, Lenoir obtained a French patent for the first commercially successful internal-combustion engine and around 500 Lenoir engines were sold in France, England, and the U.S. over the next five years. Unfortunately, the engine did not prove to be as economical as originally assumed, the electric ignition was unreliable, and the engine overheated quickly.

The same year that Lenoir obtained the patent for his engine, he designed a four-wheeled carriage to be propelled by it using a sprocket chain to transmit power to the wheels, a rod for steering, and a brake. He built a lighter engine that could run at higher speed, but quickly foresaw the desirability of substituting a liquid petroleum fuel and began designing a new engine with the addition of a carburetor. Lenoir built a model automobile that ran twenty-four kilometers in three hours, but there is no evidence that he continued his automotive pursuits, likely because the carriage was heavy and undependable (Nevins and Hill, 1954).

Nikolaus Otto, a young German store clerk, was inspired by reports of the Lenoir engine and began experimenting with engine designs for stationary power. Otto joined forces with Eugen Langen and the two developed an improved engine which used flame ignition, a compressed illuminating gas mixture, and a flywheel to regularize the force of the working stroke. The open cylinder design and free piston avoided the overheating that plagued the Lenoir, although it resulted in noisy operation. But the engine was dependable, used half the volume of gas as the Lenoir, and actually cost less to operate than a small steam engine. Otto obtained a patent in 1866; demand was slow to develop, but by 1872, Langen founded the Deutz Gas-Engine Factory for quantity manufacture. Langen hired engineer Gottlieb Daimler as chief engineer and William Maybach as chief of production. The two engineers improved the engine design, making it simpler, more efficient, and more reliable. Although still cumbersome and noisy, the unit was more compact and cost effective and was widely used in Europe.

In 1872, Englishman George Bailey Brayton, who was living in Boston, invented and patented a two-stroke continuous-compression-cycle engine that operated on liquid petroleum fuel; it was the first engine to use crude petroleum. The thermodynamic cycle of the Brayton engine is sometimes referred to as the Joules cycle, which vaporizes and then compresses the fuel in a separate chamber, feeds it into the cylinder, then burns it in a continuous flame rather than exploding it. The engine was quiet, easy to start, and reportedly as efficient as the Otto-Langen engine. While Europeans preferred the Otto engine, Americans showed a preference for the quieter Brayton engine (May, 1990; Nevins and Hill, 1954).

Otto, however, was not satisfied with the performance of his engine, and was already working on an entirely new design. He returned to a concept he had explored briefly in 1860 but had abandoned: a four-cycle engine. By drawing in a larger charge of gas, the four-cycle is able to produce an explosion powerful enough to carry through two revolutions of the crankshaft. This allows the piston to compress the next charge of gas during the third stroke, an operation that must be performed externally in the two-stroke engine, or not at all. In 1862, French engineer Alphonse Beau de Rochas had also proposed a four-cycle engine, discussing it in some detail, but had not attempted to apply the principles. With the help of Daimler and Maybach, Otto was able to make the concept operational, receiving a patent in 1877. Despite requiring a large flywheel to regularize the motion of the crankshaft, the 'silent Otto' was compact and relatively quiet, and was soon licensed for sale in England and France. By 1882, it was being manufactured in the U.S., and the Brayton soon disappeared.

Otto and Langen showed no interest in applying their invention to transportation, and although Brayton experimented with installing his motor in an omnibus, he was not fully successful. The first account of a successful internal-combustion-powered automobile is attributed to German Siegfried Marcus, who reportedly built two experimental motor carriages in the early 1870s using a four-stroke engine inspired by Beau de Rochas' work. However, Marcus saw no commercial use for his carriage and discontinued working on it. Around the same time in the U.S., patent attorney George Selden became interested in developing a horseless carriage. He first considered using steam power, but quickly turned to internal combustion. After witnessing a demonstration of the Brayton engine in 1879, he drew up plans for a 'road-locomotive' consisting of a four-wheeled vehicle powered by a gasoline engine similar to Brayton's. Selden applied for a U.S. patent, but made minor changes to the application annually which delayed its formal registration until November, 1895. Selden sought financial backing to manufacture his automobile but was unsuccessful. His patent, however, would figure prominently in the future development of the automotive industry (see section 4.5.2).

In Europe, two Germans working independently built the first crude gasoline-powered road vehicles.[20] In 1882, Daimler left the Deutz Gas-Engine Factory and began working on a lighter engine that could be used in a vehicle. Completed in 1883, Daimler's new four-cycle engine weighed around 80 pounds per horsepower and operated at up to 900 revolutions per minute (rpm), compared to the four-cycle Otto's 200. Daimler developed a carburetor that delivered fuel at a suitable ratio of gasoline and air and also developed a system that used a heated platinum tube to ignite the fuel when compression was complete. The motor eventually was water-cooled and used oil for lubrication. The light, compact, and highly efficient engine was the first specifically

[20] For brevity's sake, the term gasoline or gasoline-powered vehicle is used from here on to refer to motor vehicles powered by gasoline internal combustion engines. Although gasoline was often used as a fuel for the boiler of steam engines, the reference should be clear based on context.

designed as an effective power-plant for a motor vehicle. Daimler immediately began experimenting with using the engine in bicycles, four-wheeled carriages, and boats. He received a German patent for a two-speed motorcycle in August, 1885, and for his two-speed motor carriage in October, 1886.

Meanwhile, Karl Benz developed a two-stroke engine in the 1870s then began work on a four-stroke for automotive applications. Using knowledge of the Otto, he developed a 300 rpm, water-cooled engine with an electric ignition. In July, 1886, he successfully applied the engine to power a one-speed tricycle, using a system of belts and chains to transmit the power to the wheels. He later developed a two-speed vehicle and obtained a patent for a sun-and-planet transmission in April, 1887.

The appearance of gasoline-powered road vehicles in Germany was met with a mixture of excitement and alarm. Benz was granted permission to operate his vehicle on city streets, but was initially ordered to not exceed 6 kilometers per hour (kph) in the city and 12 kph in the country. Neither Daimler nor Benz vigorously promoted their automobiles, though Daimler began manufacturing his engine commercially in 1887. Over the next few years, he adapted the engine for motorboats, streetcars, a fire engine, and a small locomotive. In 1889, the French firm Panhard & Levassor began developing a motor carriage using Daimler's engine and began production in 1892. Benz made an alliance with Emil Roger and also began selling vehicles in France, where smooth, well-paved, level roads were more abundant than in the U.S. Thus the gasoline-powered automobile arrived first in France.

Many American entrepreneurs were experimenting with gasoline internal combustion engines and motor carriages at this time, including Henry Ford, Ransom E.

Olds, Elwood Haynes, Charles and Frank Duryea, Hiram Percy Maxim, and countless others. Henry Ford claimed to have begun building gasoline engines as early as 1891, though there is no evidence of experiments with a satisfactory motor until 1893. His first successful test of a motor carriage followed in 1896. Credit for the first successful U.S. gasoline-powered motor vehicle goes to Charles and Frank Duryea, who began manufacturing bicycles in the 1880s. The brothers began experimenting with a gasoline-powered carriage in 1891 and obtained financing for construction in March, 1892. Although Charles developed the initial designs for the vehicle, Frank worked alone over much of the following year, solving problems with the two-cycle engine, ignition, and transmission. Frank therefore deserves much of the credit for their success when the vehicle completed a public trial run in September, 1893.

Frank proceeded to redesign the vehicle with wooden-spoked wheels, rubber tires, a new water-cooled four-cycle engine mounted in the rear, a cranking device, a clutch and gear transmission, and a lever for steering. The economic crisis beginning in 1893 postponed introduction of the vehicle for sale, but in September, 1895, the brothers and backers from Springfield Massachusetts founded the Duryea Motor Wagon Company, the first company organized in the U.S. for the production of automobiles. The company also became the first to produce multiple motor vehicles from a single design when it built and sold thirteen automobiles in 1896 (Nevins and Hill, 1954; May, 1990).

3.2 Volume Production and Standardized Interchangeable Parts

Early automobiles thrilled the public with their speed and were viewed as luxury items for the rich. Conventional wisdom for high profit was to manufacture a small quantity of a highly personalized product (Ford and Crowther, 1922; Rubenstein, 2001).

This meant that demand outpaced supply, and manufacturers could sell at a high per-unit profit. But not all early automobile manufacturers subscribed to this doctrine. With the rise of venture capitalism, industries were discovering the profits that could be made with large production volumes.

However, manufacturing technology at the dawn of the automobile industry was incapable of producing *any* quantity of identical vehicles. The automobile requires parts of unprecedented durability, especially gears, and steel parts must therefore be hardened to withstand such heavy use. Because early machine tools could not work hardened steel, the parts were heat-treated after fabrication. This process often caused warping and the parts then required reshaping through further machining, grinding, or hand filing. In addition, car builders contracted different craftsmen to manufacture various parts. Though highly skilled, these craftsmen did not work to a common gauge. The resulting parts had to be custom filed and fitted for final assembly.

Innovations in machine tools and production methods around the turn of the century allowed a transformation in the automobile industry. While Henry Ford is popularly thought of as the originator of mass production, the Olds Motor Works actually became the first company to mass produce a vehicle when it introduced the curved dash for $650 in 1900. Olds produced about 1,400 vehicles in 1900, 2,100 in 1901, 3,300 in 1902 and about 4,000 in 1903 (Pound, 1934, p. 54; Epstein, 1972, p. 37; Abernathy, 1978; Rubenstein, 2001). By 1904, the Olds plant was producing up to 40 vehicles per day and the Cadillac Company was soon making 30 to 40 cars per day (Epstein, 1972). Henry Ford believed that the desire for vehicle ownership was nearly universal and saw a huge potential market for an inexpensive vehicle, but did not find commercial success until after 1903 (see section 3.3 and chapter 4).

According to Womack et al. (1990, p. 27), the key to mass production was not the moving assembly line, but rather "*the complete and consistent interchangeability of parts and the simplicity of attaching them to each other*" (emphasis in original). This required that the vehicle be *designed* for ease of manufacturing, and that all parts be standardized and made to a common gauge. While this may seem obvious in retrospect, for the 19th century manufacturer, it required entirely new ways of thinking about production. The move to mass production techniques with standardized interchangeable parts did not occur discontinuously, nor did it occur within the automobile industry. Rosenberg (1963) notes:

> The problems of large-scale automobile production involved the extension to a new product of skills and machines not fundamentally different from those which had already been developed for such products as bicycles and sewing machines. Underlying the discontinuity of product innovation, then, were significant continuities with respect to productive processes. The transition to automobile production for the American economy after 1900 was therefore *relatively* easy, because the basic skills and knowledge required to produce the automobile did not themselves have to be "produced" but merely transferred from existing uses to new ones. This transfer was readily performed by the machine tool industry.

Thus, the development of techniques for mass production reviewed in this section was a continuous and cumulative process that occurred over decades and within several industries that superficially appear to be entirely unrelated.

3.2.1 Arms Manufacturing

While the earliest automobile production was a craft industry, large volume production using standard parts was not a novel concept. As early as the end of the 16th century, the Venetian Arsenal employed interchangeable pre-manufactured parts, specialized labor, and an assembly line for shipbuilding. In 1765, French General Jean

Baptiste Vaquette de Gribeauval proposed that muskets could be manufactured economically if they were made from interchangeable parts, with the added benefit of being more easily repaired on the battlefield. However, the ideal of standardized, interchangeable parts was not easy to achieve in practice (Hounshell, 1984).

The U.S. War Department became convinced that a ‘system of uniformity’ was necessary for military progress, and invested a great deal of effort over a 40 to 50 year period to solve the problem of manufacturing interchangeable parts. The Department established the Springfield and Harper’s Ferry Armories in 1794 and 1798, respectively, and also began contracting with private arms-makers for muskets and pistols. In 1798, Eli Whitney proposed that mechanized production would facilitate large volume production, but he never produced the quantity of guns he promised, and none had interchangeable parts. Arms contractor Simeon North had greater success. With his first two contracts for a total of 2,000 pistols, he found that the division of labor resulted in significant time savings and higher quality workmanship. North’s third contract specified that any component part should be able to fit any of the 20,000 pistols ordered. To fulfill the contract, North developed the revolutionary idea of manufacturing to a standard model or gauge rather than a pattern. In addition, he pursued the design and development of special-purpose machinery.

During this time, the Springfield and Harper’s Ferry Armories began phasing out the traditional apprentice system and experimenting with new production and managerial procedures, including division of labor and piece-rate wage systems. By 1819, an elaborate gauging system had been devised. However, it was an array of special-purpose machinery developed by Thomas Blanchard that brought the benefits of mechanization and allowed large volume production of interchangeable parts. Sequentially arranged special-purpose machinery was first used in making rope blocks for the British Royal Navy around 1807. But Blanchard was the first to develop machinery for making irregularly shaped parts, a technology he applied to barrels and gunstocks for the Harpers Ferry and Springfield Armories in 1818-1819. By 1826, Blanchard, with the support of the Springfield Armory, was able to fully mechanize stock manufacturing using fourteen sequential machines. By combining this process with the gauging methods developed by North, an intense commitment to making interchangeable parts, and the financial backing of U.S. government, John Hall of Harpers Ferry Armory was able to design and manufacture a rifle with fully interchangeable parts in 1824. Hall was convinced that his process could be used for production on a large scale that, when combined with "a minute subdivision of Labor," could realize significant economies of scale (Hounshell, 1984, p. 42). However, government appropriations never allowed for more than 1,000 rifles per year at a single armory, which in 1834 were produced at Harpers Ferry and Simeon North's armories.

Hall's methods were transferred to the Springfield Armory, where they were applied to the manufacture of muskets in the 1830s. The government contracted with private arms-makers and freely shared parts, workers, and information on production technology. As a result, the Armory served as a "clearing house for technical information and a training ground for mechanics," thus facilitating the diffusion of the innovative production methods at the heart of what came to be known as the 'armory practice' or the 'American System of Manufactures' (Hounshell, 1984, p. 45). According to a British committee charged in 1854 with studying U.S. manufacturing practices, the American system was characterized by the “adaptation of special tools to minute purposes,” “the ample provision of workshop room,” “systematic arrangement” of manufacturing, “the progress of material through the manufactory,” and the “discipline and sobriety of the employed” (cited in Hounshell, 1984, p. 64).

It is significant to note that the U.S. government was not interested in efficiency, but only in producing uniform parts such that firearms could be easily repaired. Therefore, despite increasing mechanization, the Springfield Armory never decreased the cost of its products. But at the same time that the arms-makers were applying mechanization to solve the problem of uniformity, U.S. clockmakers were increasingly using special-purpose machines and tools to manufacture large quantities of inexpensive clocks. The market-driven clockmakers never aimed at nor achieved interchangeable parts and apparently did not grasp the importance of this innovation for high volume production. Though they were turning out hundreds of thousands of clocks annually by 1850, assembly was slow. With such high production levels, the clockmakers developed novel marketing and financing strategies to maintain demand, including sales on credit and the introduction of new models when demand fell (Hounshell, 1984). Similar strategies would be adopted by automakers some 70 years later.

Beginning around 1840, the increasing requirements of the firearms industry spurred the development of the infant machine tool industry, which had originated in the 1830s with the development of heavy steam-powered machinery for the New England textile mills.[21] The machine tools required for precision manufacturing of

---

[21] According to Rosenberg (1963), the development of the machine tool industry was hindered by the difficulty and expense of transporting heavy equipment. Machinery production therefore remained general-purpose and localized until the development of the railroads. After 1840, a growing number of firms began producing increasingly specialized machinery for a narrow range of applications.

interchangeable firearms parts were further developed through the 1850s and adapted for use in other industries. Rosenberg (1963) describes the machine tool industry as a "reservoir of skills and technical knowledge which are employed throughout the entire machine-using sectors of the economy." As such, it can be viewed as a knowledge base that played a central role in the learning processes responsible for U.S. industrialization. "This role…is a dual one: (1) new skills and techniques were developed or perfected here in response to the demands of specific customers; and (2) once they were acquired, the machine tool industry was the main transmission center for the transfer of new skills and techniques to the entire machine-using sector of the economy."

3.2.2 Sewing Machine Manufacturing

The requirements of sewing machine production stimulated further machining innovations in the latter half of the 19th century, and many of the mechanics hired by the nascent sewing machine industry had experience at the Springfield or private armories (Rosenberg, 1963). The Wheeler and Wilson Manufacturing Company adopted manufacturing techniques similar to those used at the Colt armory, which utilized extensive drop forging prior to machining. By 1862, Wheeler and Wilson was able to produce nearly 30,000 sewing machines annually with uniform parts that the company claimed required no filing during assembly. The company continued to expand with a production goal of 100,000 per year at reduced prices that would allow every family to own one. Its production peaked in 1872 at 174,088 machines (Hounshell, 1984).

In 1858, the Willcox and Gibbs Sewing Machine Company contracted with the Brown & Sharpe machine shop of Rhode Island, which was experienced with making clocks and watches. Despite having no experience with the armory practice, Brown &

Sharpe approached the task by designing a model, special tools, fixtures, jigs and gauges right from the start. Brown & Sharpe's success in developing special-purpose tools for sewing machine manufacture, a process that took eight months, ultimately led the company to begin selling machine tools to the industry as well. Because the company made both process machinery and finished products, it was uniquely positioned to understand the needs of the industry and benefited from in-house testing capabilities (Hounshell, 1984). Rosenberg (1963) credits Browne & Sharpe with producing the most important innovations in the sewing machine industry, including the universal milling machine first sold in 1862.

Henry Leland, the eventual founder of the Lincoln Motor Company, worked as a tool builder at the Springfield and Colt armories and several machine shops, and was hired by Browne & Sharpe in 1872 (Hounshell, 1984). According to Hounshell (1984), Leland's insights and advice were critical in the development of the universal grinding machine and a grinder for production work that were capable of producing a fine finish on hardened steel parts. These new technologies overcame the warping caused by hardening and helped eliminate the need for hand fitting at assembly. In addition, they improved the precision of gauges, and therefore the precision of manufactured parts (Hounshell, 1984).

In addition to contributing to the development of new machine tools, Leland helped institute important organizational changes. Under the armory system, arms manufacturers hired inside contractors who used the company's shop space, power, tools and materials to produce parts at a set piece-rate. The contractor hired, managed, and paid his own workers. The contract system was a hallmark of New England armory practice and contributed to the diffusion of the armory system and new technologies. Leland, however, believed that adopting a strict procedural approach would ensure that sequential operations were more closely followed and that materials would flow more smoothly. Leland eliminated the contract system and instituted a piece-work system that allowed department supervisors to determine pay rates. After a year, these changes reportedly resulted in a 47% reduction in labor costs and a comparable improvement in the quality of work (Hounshell, 1984).

After leaving Browne and Sharpe, Leland founded the Leland and Faulconer Company, which in 1899 designed a machine for production grinding of hardened bevel gears for bicycles (Rosenberg, 1963). The company became Detroit's most successful machine shop and earned a national reputation for machining cylinders and pistons to closer tolerance than other companies. Leland was then hired in 1902 to run the failing Henry Ford Company, which was renamed the Cadillac Automobile Company.[22] The Cadillac and Leland and Faulconer companies largely had the same stockholders, and were officially merged in 1905 (Nevins and Hill, 1954). It is notable that Cadillac was the first to manufacture precision-machined interchangeable automobile parts (Abernathy, 1978).

The I.M. Singer & Company initially built sewing machines using predominantly European craft manufacturing methods. Company officials believed that hand finishing by skilled workers resulted in the highest product quality. Rather than focusing on manufacturing methods, Singer concentrated on a highly successful worldwide marketing

---

[22] Cadillac was purchased by GM in 1909, but Leland retained complete control of management. When the U.S. entered World War I in 1917, Leland left Cadillac and founded the Lincoln Motor Company to produce Liberty airplane engines. The company began producing automobiles in 1920 and was sold to Ford in 1923.

strategy, which included advertising, demonstrations, financing, and establishment of retail stores. Singer's sales tripled between 1855 and 1856, due in part to the first Singer model intended exclusively for home use, changes in the company's sales system, and the initiation of an installment purchasing plan. This last innovation, which Singer's Edward Clark credited to the Wheeler and Wilson Company, allowed customers to advance a percentage of the purchase price to 'hire' the machine then pay monthly installments and eventually own it. Despite growing demand, the company continued to rely on hand fitting and assembly was slow and expensive. Singer's factories began having trouble meeting demand and the company was soon plagued with complaints from customers and sales agents about the quality of the sewing machines.

Around 1863, the now Singer Manufacturing Company began a process of increasing mechanization. Singer hired L.B. Miller, who had two years experience with New England armory practice at the Manhattan Firearms Company, to supervise the development of special-purpose machinery and tools to facilitate manufacturing with interchangeable parts. The transition was rocky, and company officials and inside contractors clashed over the tradeoff between quality and quantity production. But by 1882, Singer was producing sewing machines with interchangeable parts that did not require hand fitting when soft. The parts were stamped with a serial number, hardened, and then refit by hand using matching serial numbers. In order to gain more control over the production process, Singer subsequently eliminated the inside contract system and developed an operations guide that detailed machining instructions and work-flow routes.

3.2.3 Bicycle Manufacturing

Beginning in the late 1880s, the U.S. experienced what contemporaries named the 'bicycle craze.' In 1818, German Karl von Drais invented the first two-wheeled, steerable, human-powered vehicle, commonly called a velocipede and nicknamed a 'dandy horse'. Riders propelled the dandy horse by pushing it along with their feet. The velocipede was popular with wealthy Europeans, but was banned on city streets due to accidents involving pedestrians. Over the next 40 years, European inventors experimented with velocipedes using 2, 3 and 4 wheels, some using treadles and cranks. Around 1863, French inventors developed the first commercially successful bicycle using two wheels, rotary cranks and pedals.[23] The first bicycles were made in 1867 by blacksmith Pierre Michaux who was experienced in making carriage parts. The bicycle consisted of a cast iron frame (later replaced with wrought iron) and wooden wheels surrounded with iron 'tires'. The velocipede rolled easily on newly paved macadam roads in Europe, but the stiff frame delivered a rough ride that earned the nickname 'boneshaker'. The 'ordinary' bicycle appeared around 1870, with a large, wire-spoke front wheel, ball bearings, solid rubber tires, and a hollow-section steel frame. The design provided a smoother ride and was fast but unsafe for the rider, limiting cycling to a sport for adventurous young men. Bicycles and racing became popular among the wealthy, especially in England.

Albert Pope began selling imported English bicycles in the U.S. around 1876, and in 1878 obtained the U.S. patent rights to produce his own version of the ordinary, which he named the Columbia. Although English bicycles were manufactured using European

[23] There were conflicting claims for the French patent rights, but Pierre Lallement filed the first and only U.S. patent claim for the pedal driven bicycle in 1866.

craft techniques, Pope set out to build bicycles with interchangeable parts under contract with the Weed Sewing Machine Company. In 1866, Weed had initially contracted with the Sharps Rifle Manufacturing Company to build its sewing machines. But when Sharps changed ownership and moved to new facilities in 1875, Weed began its own production in the old plant by hiring some former Sharps employees. The Weed Company therefore had extensive experience in armory practice, and the president of manufacturing operations had even been an inside contractor at the Colt armory (Hounshell, 1984). According to Hounshell, the production of bicycles required essentially no new special purpose machinery or techniques over sewing machine manufacturing, though finishing and assembly was considerably more complicated and time consuming.

Pope aggressively marketed bicycles through cycling journals, cycling clubs, and monthly poster contests. Cycling clubs actively lobbied for road improvements. With Albert Pope's assistance, the League of American Wheelmen drafted a petition to Congress, resulting in the birth in 1893 of the Office of Road Inquiry in the U.S. Department of Agriculture. In 1883, he organized the first bicycle trade show, an institution that became an important promotional and sales tool for the bicycle and automobile industries. Pope's patent royalties allowed him to dominate the market, but the Ames Manufacturing Company, under contract to A. H. Overman, posed the first competitive threat to Pope in 1885. Notably, Ames had been involved in arms and machine tool production from the early days of the armory practice.

High-wheeled bicycles were quite successful in the U.S., with sales totaling around 250,000 by 1887 when the chain-driven safety bicycle was introduced from England (Hounshell, 1984). The instability and height of the ordinary bicycle meant riding was a daring feat, but the two same-sized wheels of the safety bicycle made it stable enough for men, women, and children of all ages. Demand for bicycles exploded, in part due to interest in the healthful benefits of outdoor recreation as a remedy for the ills caused by increasing urbanization (Geels, 2005b). In addition, bicycles conveyed to both men and women a level of freedom and mobility they had never before experienced (Hugill, 1982).

As demand grew, an increasing number of arms-makers, sewing machine companies, and even clockmakers began manufacturing bicycles in addition to their main product. According to Rosenberg (1963), there were 27 U.S. firms manufacturing bicycles and tricycles in 1890; by 1900 there were 312. Because bicycle production was a sideline, most of these companies invested in only a minimum of new equipment, purchasing machinery from machine tool companies like Pratt & Whitney and Brown & Sharpe. However, Albert Pope gained control of the Weed Company in 1890, renamed it the Pope Manufacturing Company, and ended production of sewing machines. He then pursued a strategy of backward integration by establishing a cold-drawn steel tubing plant for building frames – the first of its kind in the U.S. – and by purchasing a rubber factory for producing pneumatic tires – an entirely new industry at the time. Pope also established a research department, headed by an MIT graduate, to ensure quality control, improve designs, and perform extensive testing of materials, parts, and bicycles (Rae, 1959; Hounshell, 1984).

The safety bicycle presented new manufacturing challenges in the need for lightness (tubular steel), hardened precision parts (gears and hubs), efficient power transmission (chain), and reduced friction (ball bearings and races) (Rosenberg, 1963).

Machining techniques and tools were gradually improved, and New England manufacturers began departing from the drop-forging methods used in armory practice, increasingly opting to machine parts from solid steel bars (Rosenberg, 1963; Hounshell, 1984).

In the western states, a separate group of bicycle manufacturers arose from the carriage, wagon, toy, and agricultural implement industries. Having no experience with New England armory practice, these firms developed entirely different manufacturing processes. Beginning in the 1890s, these firms reduced machining to a minimum by developing techniques for sheet metal stamping to produce frame joints, crank hangers, hubs and sprockets. The electric resistance welder developed around 1888 by Elihu Thomson proved essential for these production methods. The crank hanger, which holds the pedal axle and the frame tubing that runs to the seat, rear wheel and steering head, was particularly complicated to machine and strength was critical. Stamping techniques resulted in a lighter, sufficiently strong part with significantly less waste, and therefore lower cost. Stamped crank hangers were eventually adopted by the entire industry. In another departure from armory practice, the mechanics in the western bicycle companies remained seated at their machines, while runners brought materials to them. Although New England bicycle manufacturers insisted that their forged and machined hubs and sprockets were far superior to stamped ones, the Western Wheel Works of Chicago overtook Pope as the largest U.S. bicycle producer before the market collapsed (Hounshell, 1984). Stamped or pressed steel would soon become vital to the automobile industry.

3.3 Automobile Manufacturing

By providing an unprecedented level of freedom and mobility, the bicycle stimulated Americans' desire for light, personal transport. It is not surprising, then, that many early automobile manufacturers had experience in the bicycle industry and, therefore, also in arms and sewing machine production. In 1892, Hayden Eames, then in charge of Pope Manufacturing's Tube Department, arranged an interview for Hiram Percy Maxim, an MIT graduate who had successfully mounted a gasoline engine on a Columbia tricycle. After Pope's engineering laboratory had inspected the invention, Maxim was invited to work for Eames in a newly created Motor Carriage Department.

Albert Pope, however, was unimpressed by the noisy gasoline contraption. Though the company continued experimentation with gasoline vehicles, Pope concentrated on developing an electric vehicle which was introduced as the Columbia in 1897. For its introduction, Albert Pope invited reporters to a private showing, allowed them to operate it, and supplied pictures for publications. The elaborate press interview soon became a standard custom in introducing new vehicles to the public (Flink, 1988). Over the next two years, Pope manufactured 500 electric and 40 gasoline vehicles, making it the leading motor vehicle manufacturer of the time. In 1899, the motor vehicle venture was separated from the parent company, taking Eames and Maxim with it, and sold to the Electric Vehicle Company. This move raised cash for Pope's attempt to save the bicycle industry through the consolidation of some 45 firms into a 'bicycle trust.' Two of the firms acquired by Pope's new American Bicycle Company produced electric and steam vehicles. These operations were reorganized as the International Motor Company in 1901 and converted to producing gasoline cars. As the bicycle market continued to collapse, Albert Pope converted more facilities to manufacturing gasoline vehicles and consolidated his automotive ventures as the Pope Motor Car Company in 1903.

Charles and Frank Duryea began manufacturing bicycles in the 1880s. Charles claimed to have been interested in 'horseless carriages' since boyhood, but an announcement in 1890 encouraged him to pursue this interest more seriously: an exhibit at the World's Columbian Exposition in Chicago in 1893 would offer rewards for the best examples of "vehicles propelled by other than animal power." By this time, Charles had contracted the Ames Manufacturing Company of Massachusetts to produce his bicycle, and Frank was employed as a toolmaker at Ames. With Charles' designs and Frank's practical abilities, the two began experimenting with a gasoline-powered carriage (May, 1990). In September, 1895, the brothers and backers from Springfield Massachusetts founded the Duryea Motor Wagon Company, the first company organized in the U.S. for the production of automobiles. The company also became the first to produce multiple copies of a motor vehicle for sale to the public when it built and sold thirteen automobiles in 1896 (May, 1990).

However, Duryea produced only three automobiles in 1897 and failed in 1898. It appears that the Duryeas were unsuccessful at transferring critical elements of armory practice from their experience with Ames, including the organization of operations. May (1990, p. 167) describes the Duryea factory as "a hopeless tangle of machinery and workmen," and attributes the company's failure to the lack of experience in manufacturing a complex product. But availability of startup capital for such a large undertaking may also have played a role (Nevins and Hill, 1954). According to Epstein

(1972), people were skeptical about large-scale production of automobiles and few were willing to invest in the industry. Manufacturers who were interested in high volume production, such as Olds, Cadillac, and Ford, were forced to start on a small scale, turn over capital quickly, and reinvest any profits in expansion. Since these manufacturers largely purchased parts, assembled them into complete vehicles, and sold for cash, the industry was basically financed by the parts manufacturers who were largely engaged in production of parts for other industries (Epstein, 1972).

In fact, the majority of automobile manufacturing prior to 1900 was undertaken as a sideline business by established firms in New England. In 1899, three quarters of the automobiles produced were built in the Northeast (Jarvis, 1972). These firms were well capitalized and had a ready supply of skilled machinists and facilities. In addition to manufacturing infrastructure, the region benefited from relatively well-developed roads and communication systems. Perhaps because of the availability of soft water for steam and electric power for recharging, New England manufacturers specialized in steam or electric automobiles of medium to high price.

A second center of automobile production arose in the Midwest which was the center of carriage and wagon manufacturing, due in part to the abundance of hardwood forests. The region was largely rural, lacked electricity, with the population widely dispersed and connected by some of the nation's worst roads. In addition, the area was prominent in the production of gasoline engines used for rural and marine applications (Flink, 1970). Therefore, the Midwestern manufacturers specialized in gasoline-powered automobiles. But the makers of gasoline cars were characterized by a significant factor other than geography. According to Jarvis (1972, p. 69), "Established firms went for the electric or the steam auto, whereas the small, underfinanced company headed by a man with technical expertise favored the internal combustion engine." The individual entrepreneur was willing to pursue a more radical innovation than the established company. By 1904, the Midwest and the gasoline automobile dominated the industry.

Ransom E. Olds was one such Midwestern entrepreneur and the only automotive pioneer to have early experience in all three sources of motive power. Olds had worked in his father's Lansing, Michigan, machine shop, P.F. Olds & Son, eventually taking on a leadership role in the enterprise. Around 1885, Olds developed a steam engine for business use that employed a gasoline burner. The engine was extremely successful and the business expanded. Olds experimented with applying his steam engine to a motor carriage, eventually linking two 2-hp engines to power a 1,200 pound, four-wheeled vehicle. Although seriously under-powered, the Olds steamer was featured in *Scientific American* in May, 1892, and Olds sold the vehicle to an English patent medicine firm the following year. In 1893, Olds also observed the large number of European-made gasoline internal combustion motor vehicles on display at the Chicago World's Fair and became convinced that gasoline engines were the technology of the future.

Olds immediately began developing a gasoline engine and applied for a patent on a simple and efficient design in 1895. P.F. Olds & Son began production in 1896 and the highly successful engine soon became the company's main product.[24] Olds attached the one-cylinder engine beneath a carriage supplied by a local manufacturer and drove it in a public demonstration in August of 1896. In his application for a patent, Olds stated his practical objectives to "produce a road vehicle that will meet most of the requirements for

[24] In 1897, P.F. Olds & Son was recapitalized and reorganized as the Olds Gasoline Engine Works.

the ordinary uses on the road without complicated gear or requiring engine of great power and to avoid all necessary weight" (May, 1990, p. 360).

Although Olds attempted to assemble automobiles in the P.F. Olds & Son shop using his engines and purchased components, demand for the engines for farm and business use took precedence and he hesitated to divert labor to the new endeavor. Olds obtained outside financial backing to establish the Olds Motor Vehicle Company in 1897 and the board of directors empowered Olds to "build one carriage in as nearly perfect a manner as possible" (Pound, 1934, p. 50). Unfortunately, the funds were insufficient to produce more than a handful of vehicles. Olds obtained more significant funding from banker Samuel L. Smith, but Smith insisted that the engine manufacturing and automobile businesses be consolidated. The Olds Motor Works was established in 1899 with Smith holding a large majority of the stock, though Olds retained presidency of the engine works.[25] A new factory for automobile manufacturing was promptly constructed in Detroit, but Olds, who was in charge of operations, was unable to select a vehicle design. Though he firmly believed that gasoline internal combustion was the motive power of choice, the first vehicle produced was an electric. In the first year, Olds produced a small number of the electrics and some large, expensive gasoline cars.

Though the company's initial plans were to build a fairly expensive automobile with "up to the minute improvements – pneumatic clutch, cushion tires, and electric push-button starter," Olds became convinced that the machine would be too complicated for the still skeptical public. By summer 1900, Olds finally settled on an inexpensive, lightweight, gasoline-powered runabout that could be constructed for $300 and sold for

---

[25] The Olds Motor Works divested of the engine works in 1903.

$650 – the Oldsmobile curved dash. Reminiscing in 1926, Ransom Olds stated that "[his] whole idea in building it was to have the operation so simple that anyone could run it and the construction such that it could be repaired at any local shop" (cited in Pound, 1934, p. 52). The curved dash weighed only 700 pounds, was powered by a 4.5 hp, one-cylinder engine, with a two-speed planetary type transmission and a chain drive.

Olds contracted out the supply of wheels, batteries, and bodies. Demand for the Olds engine, which was manufactured in both Lansing and Detroit, remained high, so Olds contracted the Dodge Brothers and Leland and Faulconer to manufacture engines and transmissions for the curved dash. By spring of the following year, more than 300 orders had been received, and Olds directed the Detroit plant superintendent to gear up to produce 1,000 vehicles over the coming year. But just as production got underway, the plant burned to the ground in early March, 1901. Legend has it that a single curved dash survived the fire, forcing Olds to concentrate solely on this model. In truth, the other vehicles could still have been produced since their plans survived in the company's vaults. However, the accident may have prompted Olds to abandon the larger vehicle models earlier than initially planned.

While a new plant was built on the old site, Olds established temporary assembly facilities and production of the curved dash resumed in April. At the same time, Lansing officials made the Olds Motor Works an attractive land offer, and the firm began constructing a second automobile facility. In 1901, Samuel Smith served as company president and Olds served as vice president and general manager of the Lansing plant, while Smith's son Fred took over operations in Detroit. To keep up with demand, Olds assigned workers a limited number of tasks rather than have them work on assembly of the entire vehicle, and arranged them in a primitive sort of production line. By 1902, the Olds Motor Works was assembling the curved dash at both the Lansing and Detroit facilities with production on the order of 2,500, making the company the largest U.S. automobile producer. By 1904, Oldsmobile sales topped 5,500 (Flink, 1988).

However, a rift began to develop between Ransom Olds and Fred Smith. Concerned about the number of complaints from owners, Smith urged Olds to correct defects in the fragile curved dash. At the same time, Smith began pushing Olds to develop a larger touring car. While Olds apparently agreed with the need for a larger vehicle, he disagreed on the timing and resented Smith's intrusion in his domain – the production side of operations. The relationship became untenable in 1903 when Smith, without Olds' knowledge or authorization, established an experimental engineering shop at the Detroit plant and soon after applied for a patent on improved springs that would be more suited to heavier vehicles. By summer of 1904, Olds had sold all his stock in the Olds Motor Works and resigned from the board.

Olds went on to establish the Reo Car Company, named for his initials, which began production in 1905. The Reo line included a $650 runabout and a 1,500 pound touring car that sold for $1,250. Meanwhile, Smith continued production of the curved dash but also moved forward with the production of a larger touring vehicle as well. Apparently because of Olds Motor Works' reputation as a producer of small cars, the new Oldsmobile did not fare well in the market, and Reo soon outsold Olds. By 1908, Olds Motor Works sales fell to only 1,000 cars per year.

The Dodge brothers' history parallels the Duryeas', but with greater success. Frank and Horace Dodge were both experienced machinists in 1897 when they began manufacturing bicycles in Windsor, Ontario. In 1900, they left the bicycle industry to open a machine shop in nearby Detroit. The Dodge Brothers served a wide variety of clients, and soon became one of the best machine shops in the Midwest, noted for precision machining and for an arrangement of machinery that allowed for ease and speed of manufacturing (Nevins and Hill, 1954). Dodge Brothers began supplying automobile components, notably engines and transmissions under contract to Ransom E. Olds (Hyde, 2003).

In 1903, the Dodge Brothers Company took a huge risk and established an alliance with the fledgling Ford Motor Company. Dodge began supplying chassis – engines, transmissions, and axles – for the Model A, but the brothers were far more than supply contractors. Prior to production, they made improvements to Henry Ford's engine and axle designs (May, 1990). Ford was having difficulty raising startup cash, so the Dodges accepted 10% of the capital stock in the Ford Motor Company in exchange for $7,000 worth of components and a $3,000 bank note. John F. Dodge was named as one of five directors when the company was incorporated (Nevins and Hill, 1954). Although Ford built its own factory in 1904 on Piquette Avenue in Detroit, the company purchased most of its parts. The facility was therefore designed for assembly and housed only general purpose machines (Hounshell, 1984). By 1904, Ford had less than $10,000 invested in equipment, while Dodge had spent more than $60,000 for machinery dedicated to producing Ford automobiles. While Ford was mainly an assembler of components, Dodge was manufacturing nearly a complete automobile except for wheels and a body. By 1905, Dodge was supplying Ford with 400 engines and transmissions per month.

Dodge continued to be Ford's largest supplier until 1913, with Dodge manufacturing around 60% of the total value of Ford automobiles (Hyde, 2003). John Dodge was named vice president of the Ford Manufacturing Company when it was incorporated in 1905 and was named vice president of the Ford Motor Company in 1906.[26] Sometime in 1912, the Dodge Brothers decided to manufacture their own automobile, although they knew that, as a result, Henry Ford would not allow them to continue participating in management decisions. While the arrangement between Ford and the Dodges had been mutually beneficial and the brothers had prospered, annual price negotiations over Dodge parts caused friction. In August of 1913, John Dodge resigned as director and vice president of the Ford Motor Company (Nevins and Hill, 1954). The Dodges were able to rely on their extensive experience and reputation, and when the first Dodge automobile was produced in November, 1914, it was an immediate success.

The influence of Henry Ford himself was of primary significance after the turn of the century. While a complete discussion of Ford's history is left for the following chapter, a brief overview of his activities prior to 1902 completes the discussion of the origins of the automobile industry. As a child growing up on a farm in rural Michigan, Ford showed an interest in and aptitude for all things mechanical. His formal mechanical training began at age 16 when he left the farm for Detroit, where he garnered extensive experience repairing engines. Sometime in the late 1880s, Ford began experimenting with steam and gasoline internal combustion engines for a horseless carriage. While working as chief engineer at Edison Illuminating Company in 1896, Ford built a two-

[26] Henry Ford served as vice president of the Ford Motor Company and banker John Gray served as president. After Gray's death in 1906, Ford was named president and John Dodge replaced Ford as vice president.

cylinder engine, chassis, and body with the assistance of two fellow Edison employees. Built in a garage behind Ford's rented house, the quadricycle weighed only 500 pounds and used bicycle wheels with pneumatic tires. After completing a successful test run in May, 1896, Ford sold it so that he could begin building a new vehicle.

Ford's second and improved horseless carriage was completed in the summer of 1899 with the help of a carriage blacksmith. Though still lightweight, it was substantially stronger than the quadricycle, which had been made almost entirely from wood. In August, Ford resigned from Edison after he and his financial backers had established the Detroit Automobile Company. The company announced ambitious plans for production in 1900, but the shop encountered a number of supply problems. According to Nevins and Hill (1954), the company's car also was probably too expensive and not quite ready for production. Ford apparently wanted to design a better car but was vetoed by the stockholders. Though the Detroit Automobile Company did produce a few vehicles, it soon ground to a halt and was dissolved in January, 1901.

Undaunted, Ford decided to build a racer and use it to gain recognition. His racing successes reinvigorated his investors, who reorganized the old Detroit Automobile Company as the Henry Ford Company in November, 1901. Unfortunately, dissension arose over Ford's preoccupation with racing and his dissatisfaction with the financial arrangements of the new company. The stockholders eventually hired Henry Leland of Leland & Faulconer to complete the design of a commercial car and begin production. In March, 1902, Ford left and Henry Leland took over the company, which was renamed the Cadillac Automobile Company.

Ford then joined forces with a new assistant, C. Harold Wills, who, much like Ford, had keen mechanical instincts and preferred to learn by doing. Wills took charge of designing two new racers, one of which set a speed record in October 1902 by averaging a mile in one minute and six seconds. Ford, however, was no longer entirely focused on racing. Despite his prior business failures, he was determined to build and produce a car to replace the family horse. In the summer of 1902 he began planning an inexpensive motor vehicle that included a novel, vertical engine design that resulted in less vibration, noise and wear. In August, 1902, Ford established a partnership with a coal merchant, Alexander T. Malcomson, to produce a pilot model. The partnership would become the Ford Motor Company in June, 1903, and begin production of the Model A.

3.4 Technical and Institutional Context

3.4.1 Industrialism, Individualism, Capitalism, and the Progressive Movement

By the beginning of the 20th century, increasing industrialism had brought prosperity and a rising quality of life to many. While technological innovations facilitated the exploitation of natural resources and a phenomenal growth in productivity, these developments rested on fundamental philosophies and the institutional arrangements based on them. The original U.S. settlers had come to America to escape religious persecution and were heavily influenced by egalitarian ideals. These ideals shaped the U.S. Constitution and Bill of Rights, which sought to protect individual rights. Thus, the idea of individualism, which stresses human independence and the importance of individual self-reliance and liberty, was a fundamental American ideology.

This individualism was also reflected in the development of American economic doctrine advocating private ownership of property over collective or state arrangements and individual autonomy in economic decision-making. In the 19th century, many Americans believed in strict adherence to these ideals. In such strict free-market economics, or *laissez-faire* capitalism, the state has no responsibility to intervene in markets to maintain a desired wealth distribution or to protect citizens from poverty. In addition, in the *laissez-faire* view, the government should not create legal monopolies or break up *de facto* monopolies.

By the late 19th century, it was clear that not everyone was sharing in the prosperity wrought by industrialism and capitalism. The rise of factory production required an unprecedented number of semi- or unskilled workers, which were drawn from the countryside and from an influx of immigrants. Rural farm culture was eroded while crowded urban slums arose to house this large population of poorly compensated factory workers. The pollution from factories combined with wastes from the slums to create unhealthy living conditions which the wealthy and middle-class soon sought to escape. Although the growing middle-class owed its prosperity to industrialism and capitalism, increasing tensions in the large working class threatened the security of their new-found socio-economic position. Meanwhile, a select few were becoming phenomenally rich as their business enterprises grew to colossal proportions.

A growing faction began to view capitalism as unfair and exploitative and felt that an increasingly corrupt government failed to serve its citizens. Activists began pursuing a wide range of social, economic, and political reforms to cure American society from the problems associated with industrialism. The 'progressive' ideals of these reformers diverged from the strictly individualist view and often leaned toward socialism. Drawn from the urban educated middle-class, the Progressives believed that government should play a role in solving social problems, establishing fairness in economic matters, and improving the welfare of all citizens. The movement sought to remove corruption from government and give people more participation in and direct control over the political process. The Progressive movement was not coherent and many reformers disagreed on issues and approaches, but in general, they sought to regulate business practices, improve working conditions, and address health hazards. Progressive issues included temperance (prohibition), child labor and welfare, safe and sanitary food and drugs, the right to collective bargaining, conservation of natural resources, immigration control, women's suffrage, and trust regulation.

Activists worked at the local, state and national levels, and much of the movement's success can be attributed to publicity generated by muckrakers – writers like Upton Sinclair who detailed the horrors of poverty, urban slums, factory conditions, and child labor. Local efforts attacked political corruption, addressed child welfare, suppressed red-light districts, expanded schools, and constructed playgrounds. At the state level, Progressives promoted minimum wage laws for women, industrial accident insurance, child labor restrictions, and factory regulation. Among the movement's successes at the national level were the Interstate Commerce Act passed in 1887 which addressed issues of railroad abuse and price discrimination; the Sherman Anti-trust Act passed in 1890 authorizing federal action against the formation of trusts and the restraint of trade; the Pure Food and Drug Act passed in 1906 providing for federal meat inspection and prohibiting the manufacture, sale, and transportation of adulterated food products or poisonous patent medicines; the 18th Amendment ratified in 1919 prohibiting the manufacture, sale, and transportation of alcohol; the 19th Amendment ratified in 1920 granting women the right to vote; and the development of the National Park system beginning in 1902.

3.4.2 Emerging Markets and Fuel Infrastructure, 1895-1900

Around the turn of the century, gasoline automobiles were competing with steam and electric cars, each of which had niche markets, as well as horse drawn vehicles and bicycles which had established markets. By 1899, around 30 manufacturers had produced 2,500 motor vehicles. Based on their interpretation of the functional requirements of the automobile and on the product and production technologies with which they were experienced, individual entrepreneurs developed a variety of product solutions that could provide personal motorized travel. Motor vehicles varied in terms of motive power, mechanism of power transmission, location of engine, body style, and a multitude of design characteristics. Consider the great variety within just gasoline cars:

> Nothing was certain, all was in process of experiment. There were offered cars with engines of one, two, three, four, or even six cylinders. The engine could be of either the four-cycle or two-cycle type. The cylinders could be parallel or opposed; they could be set in the chassis either horizontally or vertically; they could be mounted under the body, under the hood, or at the rear of the car near the axle. The engine might be either air-cooled or water cooled. Steering could be by bar, tiller, or wheel. The drive might be by shaft, through bevel gears, or by chain. If by chain, it might be either double or single. The transmission could be either planetary or sliding gear; it might have either 2, 3 or 4 speeds forward. Finally… the ignition might be by either battery or magneto; if batteries were employed, they could be either dry or storage; if magneto, this might be of either high or low tension (Epstein, 1972, p. 87-89).

No doubt the bewildering variety of technical configurations caused many to postpone purchasing a motor carriage. More significant, however, was the fact that automobiles were expensive to purchase and maintain. Parts wore out quickly or broke suddenly. Springs lasted only 1,200-2,000 miles and tires were only warranted for 2,000-3,000 miles (Epstein, 1972). Because parts were not standardized among vehicles nor produced in volume, they were extremely expensive. It is not surprising, then, that the wealthy elite were the first to purchase motor vehicles. Nevins and Hill (1954, p. 197) describe the development of a 'special market' for automobiles among the social leaders of the East, advertised by "lavish activities not unworthy of the Arabian Nights." The son of a banker initiated the era in 1897 with the purchase of a French car, making it "good form" to own an automobile. By 1900, "nearly everyone at Newport owned 'a machine,' and some two or three… Naturally the fashion in the East became the fashion elsewhere."

Among these early owners, some of the men favored gasoline cars while nearly all the women used clean quiet electric vehicles which they used for social calls, shopping, and riding downtown. Many of the wealthiest households could afford to own more than one type of vehicle – a horse and carriage, an electric vehicle for short trips around town for the lady or doctors making late night calls, and a gasoline vehicle for touring the countryside. This created what Kirsch (2000) refers to as the hybrid household. According to Flink (1970, p. 57), *Harper's Weekly* estimated in 1906 that "there are more than 200 persons in New York who have from five to ten cars apiece… The string of vehicles owned by an enthusiast of this class will include two or three touring cars, a pair of racers, a couple of broughams, a runabout, a station-car, and a work car."

Early observers expected the alternative motive powers to share the eventual established market. *Horseless Age* editor Albert Clough commented in 1903 that "experience will furnish the decisive evidence which shall assign each motive power to its appropriate sphere" (cited in Kirsch, 2000, p. 217). Initially, steam and electric vehicles offered distinct advantages over internal combustion engines. Steam engine technology was more mature and more people had experience operating and repairing them, while electric vehicles were quiet, clean, easy to operate, and less mechanically complex because they did not require a clutch and multi-speed transmission. As a result of these advantages, more steam and electric automobiles were produced in the early years of the industry. In 1900, 57 U.S. manufacturers produced 4,192 motor vehicles of which 1,681 were steam, 1,575 were electric, and 936 were gasoline (Flink, 1970; Mowery and Rosenberg, 1998). Locomobile steamers, based on the Stanley patent, accounted for around 5,000 of the 22,000 automobiles produced between 1899 and 1902.

Poor quality intercity roads produced regional isolation and sustained the unique characteristics of local markets (Kirsch, 2000). In 1900 in Chicago, electric vehicles accounted for 65% of motorized vehicles; in 1902 in Portland, Maine, 88% of vehicles were steam; in Cleveland, 49% of the first 100 vehicles registered were internal combustion. According to Flink (1970), about half of the 100 automobiles in Los Angeles in early 1902 were electrics, with the remaining half divided between steam and gasoline. But by 1904 in Cleveland, internal combustion engines accounted for 61% of the vehicles.

Fuel infrastructure played a significant role in the development of these regional preferences. Steam cars were designed to run on kerosene or gasoline and which nearly all rural general stores in the 1890s sold for lighting and stationary gasoline engines. Prior to the incorporation of condensers beginning in 1901, steam cars also required clean, soft water, which in many areas could be obtained at public watering troughs that were made available for horses. Gasoline cars also benefited from the general availability of gasoline, but there were issues of quality, as some retailers diluted the fuel with water or heavier petroleum fractions. In rural areas, gasoline and steam engines also benefited from the relative availability of mechanics experienced with repairing engines used as stationary power sources. Meanwhile, electricity was only available in parts of some urban centers and output was unreliable. Furthermore, connecting to the power supply was complicated by the lack of a standard plug or operating voltage (Kirsch, 2000).

3.4.3 Enthusiast Publications and the Popular Press

The first publication devoted to the automobile, *Horseless Age*, appeared in November, 1895. Enthusiast publications were read by motorists and also served communication among engineers. While hundreds of early press articles revealed a strong skepticism of motor vehicles, others extolled the virtues of automobiles in glowing prose. Consider the following description found in the inaugural edition of *Horseless Age* (cited in Nevins and Hill, 1954, p. 165):

> A pleasing prospect it is, that rises before us in contemplating this array of horseless vehicles! From the gradual displacement of the horse in business and pleasure, will come economy of time and practical money-saving. In cities and towns the noise and clatter of the streets will be reduced, a priceless boon to the tired nerves of this overwrought generation… On sanitary grounds too the banishing of horses from our city streets will be a blessing. Streets will be cleaner, jams and blockades less likely to occur, and accidents less frequent, for the horse is not so manageable as a mechanical vehicle.

The general consensus among proponents was that each option for motive power had its place in the market. Established professional publications like *Electrical World* and *Electrical Engineer* expressed optimism for the future of electric vehicles, citing the maturity and innate superiority of the underlying technology. However, editorials published in *Horseless Age* in 1899 took a stance against attempts by the Electric Vehicle

Corporation (EVC) to monopolize the market, saying that trust-building was premature, would slow technological and market advance, and would only benefit investors. The magazine also reported discouraging results from European electric cab experiments and battery testing and its editorials soon lambasted the technology as well as the EVC. After the EVC plan failed, the editor of *Horseless Age* took a more balanced position and claimed its actions had been in the interests of saving the technology and industry from unscrupulous speculators (Krisch, 2000).

In 1897, Albert Pope pioneered the use of the elaborate press interview for introducing new automobile models. Other manufacturers, including Henry Ford, adopted his tactics, making it common industry practice. The articles produced from these interviews were read more like advertisements than balanced journalism. A reporter for the Detroit *News-Tribune* rode in Ford's prototype vehicle for the Detroit Automobile Company in the fall of 1899. His vivid account appeared on the front page under the headline (cited in Nevins and Hill, 1954, p. 180):

SWIFTER THAN A RACE-HORSE
IT FLEW OVER THE ICY STREETS
Thrilling Trip on the First Detroit-Made Automobile,
When Mercury Hovered About Zero

It was also the popular press which initiated races in the United States, beginning with the Chicago *Times-Herald* race in 1895. Manufacturers also enlisted the press to publicize their vehicles' performance in long distance reliability tests. Alexander Winton was the first to follow this tactic when a reporter accompanied him on his second 800-mile trip from Cleveland to New York City in 1899.

3.4.4 Races, Demonstrations, and Trade Shows

Speed and reliability races captured the public's interest and demonstrated the performance, endurance, and dependability of competing technologies and models. The earliest automobile races were held in France, where twenty-five horseless carriages competed in a race in July, 1894. Of the fifteen that completed the course, the winner was a De Dion steam car which was nearly tied by a gasoline Peugeot; a gasoline Panhard & Levassor finished fourth. In a 727-mile French race held the following year, a Panhard & Levassor made the fastest time, but, as a 2-seater, it was ineligible for the prize which the Peugeot Brothers took. Already, the gasoline internal combustion engine was gaining on the competition. Shortly afterward, *Scientific American* conceded that "[u]pon the whole, it was the lightest vehicles that behaved best on the road, and this fact, now universally established, proves the superiority of gasoline and naphtha over any other motive power at present known" (cited in Nevins and Hill, 1954, p. 139).

The first U.S. reliability race was sponsored by the Chicago *Times-Herald* in November, 1895. Since none of the 89 potential entrants was ready by race day, the race was postponed three times. On the third try, only two contestants were ready: a Duryea and a Benz, both gasoline-powered. The organizers allowed a 92-mile consolation run that the Benz won after the Duryea ran into a ditch to avoid a farm wagon. On the fourth try, six vehicles started the race, which required the competitors to brave 30-degree temperatures and negotiate deep snow that had fallen the day before. Two electric cars and a Benz dropped out after brief runs. Another Benz lasted eight hours before the engine died. Only two completed the 52.4 mile course: the same Benz and Duryea that had completed the consolation run. The Duryea emerged the winner, averaging only 6.66 mph through the snowy streets of Chicago (Nevins and Hill, 1954; Rubenstein, 2001).

Despite the encouraging performance of gasoline cars in early reliability races and demonstrations, Americans showed a definite initial preference for steam. Two years after the Chicago *Times-Herald* race, an industry expert asserted that "the vast majority of people would prefer a smooth-running, reliable steam engine for use as the propelling medium of a pleasure or light business carriage, to the evil-smelling, dangerous, wasteful, and at best uncertain and unreliable engine heretofore chiefly employed for that purpose in motors of recent construction" (Nevins and Hill, 1954, p. 141).

The first U.S. track race was held in Rhode Island in 1896 with five Duryeas and two electric cars competing. An electric won, averaging 26.8 mph and eliciting a cry from spectators of "get a horse!" (Flink, 1970, p. 42). But the pace soon picked up and gasoline-powered cars quickly demonstrated an advantage over the competing options. An early speed race at Grosse Pointe race track near Detroit in October, 1901, held separate competitions for steam, electric, and gasoline cars. The winning cars then participated in an exhibition, with the gasoline car completing one mile in 1 minute, 12.4 seconds compared to 1 minute, 52 seconds for the steam car, and more than 4 minutes for the electric (Rubenstein, 2001). However, in 1906, a Stanley Steamer averaged 127.7 mph in a one mile race at Ormond Beach, Florida (Flink, 1970).

Sanctioned races placed primary emphasis on speed and revealed weaknesses and strengths in technologies and designs (Flink, 1970). However, racing tended to reinforce the definition of the automobile as an expensive toy for the very wealthy and inhibited diffusion into wider markets. Ford recalled (Ford and Crowther, 1922, p. 36):

> When it was found that an automobile really could go and several makers started to put out cars, the immediate query was as to which would go fastest. It was a curious but natural development – that racing idea. I never thought anything of racing, but the public refused to consider the automobile in any light other than as a fast toy. Therefore later we had to race. The industry was held back by this initial racing slant, for the attention of the makers was diverted to making fast rather than good cars. It was a business for speculators.

Meanwhile, long-distance reliability demonstrations proved that the gasoline automobile could successfully negotiate the nation's poor roads, enabling the new sport of touring. In the summer of 1897, Alexander Winton drove one of his company's cars 800 miles from Cleveland to New York City in less than seventy-nine hours. Two years later, he repeated the trip in only forty-eight hours, taking with him a newspaper reported to publicize the trip *en route*. Hoping to prove that his moderately priced, light car was reliable for touring, Ransom Olds picked a tester from his factory to drive the new curved dash Oldsmobile from the Detroit plant to the New York Automobile Show in 1901. Despite extremely muddy roads, the Oldsmobile made the trip in nine days, requiring all of the myriad spare parts the driver had taken along (Flink, 1970).

In the summer of 1903, three separate expeditions completed cross-country trips from San Franscisco to New York City. The first, and perhaps most noteworthy, was undertaken by a Vermont physician, Horatio Nelson Jackson, and a hired chauffer in a Winton automobile without the sanction or assistance of the manufacturer. The two took shifts driving over rough, rock-strewn dirt roads, through mud and creeks, using a block and tackle to dislodge the automobile when it got stuck and frequently patching punctured inner tubes. More than once, the team was stranded for days awaiting replacement batteries, tires, and parts, but Jackson proved that private motorists could undertake long-distance motor trips. Jackson completed the trip in sixty-three days. The two other trips, one in a Packard and the second in an Olds curved dash, were undertaken by the manufacturers while Jackson was *en route*. According to Flink (1970), Olds' unprecedented sales success in 1902 and 1903 was to a large extent attributable to the favorable publicity generated by these accomplishments. Such feats fueled popular expectations that the horseless carriage might actually become accessible to the masses and eventually displace the horse.

In addition to races, automobile manufacturers followed the lead of the bicycle industry and began using trade shows to publicize their new models. The first automobile show was held in February, 1899, in conjunction with the bicycle exhibit at Madison Square Garden. The first independent automobile show was held in November, 1900. American and foreign manufacturers displayed thirty-four models, of which nineteen were gasoline-powered, seven were steam-powered, six were electric, and two combined gasoline internal combustion engines and electric motors. One manufacturer of gasoline vehicles built a wooden hill and hired a stunt driver to demonstrate the the Gasmobile's hill-climbing, braking, and backing abilities (Nevins and Hill, 1954).

3.5 Synthesis and Analysis

This section synthesizes the history of the emergent phase of the automobile industry presented in this chapter and analyzes it within the framework presented in section 2.6. Section 3.5.1 describes the changes in the technical, institutional, and ecological complex (TIEC) resulting from the success of the bicycle that initiated the era of innovative ferment from which the automobile industry emerged. Section 3.5.2 identifies the initial markets for automobiles and describes the functions that motor vehicles provided in these markets. These markets emerged due to multiple factors, including problems with the existing carriage regime, the combination of functional features which provided advantages over existing options, and the articulation of a new functionality. Product design is then described as arising from multiple regimes and benefiting from innovation in multiple industries. Section 3.5.3 identifies spillovers and learning-by-doing as the primary mechanisms of learning and describes innovation as a cumulative and synthetic process during this era. Section 3.5.4 describes the relationship between the agents of change and existing regimes and the role of these agents as heterogeneous engineers. I also identify four examples of the emerging technological nexus role which links variation and selection.

Section 3.5.5 then summarizes the findings of this analysis as they relate to the research questions posed in Chapter 1. I find that the role of niches during the emergent phase of the motor vehicle industry is consistent with the QE and MLP theories in several ways. However, while the QE and MLP approaches focus on a single existing regime, this analysis shows that considering misalignments in multiple regimes along with other developments in the TIEC is critical to understanding why the automobile industry emerged when it did and in the particular niche markets that it did. Further, while these theories identify learning-by-doing and learning-by-interacting with users as the primary mechanisms in learning about user's needs and improving product performance, I find that innovation in this era was highly synthetic and resulted from cumulative progress in products and processes over more than a century of development. Thus, knowledge spillovers were a significant, if not the primary, mechanisms underlying the development of this new innovation. Finally, the activities of the agents of change in an emerging technological nexus support the assertion that, in order to realize successful designs, entrepreneurs must behave as heterogeneous engineers, solving social, economic, and infrastructural as well as technical issues.

3.5.1 Initiation of an Era of Innovative Ferment

Evolutionary theories generally assert that innovation is triggered by new scientific or technical breakthroughs – a new invention. The QE and MLP theories focus on transitions rather than the initial introduction of radical technologies, though they assert that innovations emerge in niches. However, the MLP identifies tensions or misalignments in the existing regime and landscape developments as opportunities for niche branching, whereby the technology then breaks into new niches. The motor vehicle was not an innovation based on a recent scientific or radical breakthrough. Rather, innovation was initiated by two factors arising from the bicycle regime: 1) the bicycle altered perceptions about personal travel; and 2) the bicycle was reaching technological limits. This finding, discussed further in this section, supports the description of innovation in the framework described in section 2.6 where innovative ferment is triggered by any change in the TIEC – technical, institutional (economic, social, etc.), or ecological – which entrepreneurs can exploit to create value.

At the dawn of the motor vehicle industry, the underlying technologies for generating power from steam, electricity, and gasoline did not represent recent inventions. Most notably, practical steam engines were developed around 1800 and were in use in rural America as early as 1830. Self-propelled traction engines were in use in the U.S. beginning around 1873 and Roper was building experimental steam-powered vehicles capable of road travel as early as 1860. Indeed, steam technology was capable of fulfilling the functional requirements for basic personal road transport 20-30 years before the first commercially successful automobiles were built. A handful of entrepreneurs experimented with it, perhaps sensing a latent demand, but the market did not materialize.[27]

In the words of Utterback and Abernathy (1975), "[t]echnological innovations which may have market application, lie fallow until markets can be identified or created." One could speculate that the market may have developed earlier if a particularly determined product champion with exceptional vision and technical abilities had built a functional, light, steam-powered motor carriage and invested in a marketing strategy in line with that of the I.M Singer & Company for sewing machines and Albert Pope for bicycles. But that is not how history unfolded. The automobile industry did eventually pattern its marketing strategy in this manner, using journals, demonstrations, races, and trade shows, but not until after 1895.

Given that the development of a *practical* gasoline engine was nearly simultaneous with the commercial development of motor carriages, it is tempting to conclude that this innovation served as a critical enabling technology – the technology that initiated the 'technology push.' But while the gasoline engine quickly became the motive power of choice, it was a less mature technology compared to steam and not the preferred option at the outset of the industry. Contemporaries did not attribute the development of the automobile to the invention of the gasoline engine. In his reminiscences of 1937, Hiram Percy Maxim (cited in Nevins and Hill, 1954, p. 133) recalls,

---

[27] Latent demand generally refers to desires that consumers cannot satisfy due to a lack of money or information regarding the product's availability. As used here, it refers to desires that consumers cannot satisfy due to product unavailability. As such, it cannot be revealed through market signals (price and quantity).

> As I look back, I am amazed that so many of us began work so nearly at the same time, and without the slightest notion that others were working on the problem. In 1892, when I began my work on a mechanical road vehicle, I suppose there were fifty persons in the United States working on the same idea. Why did so many different and widely separated persons have the same thoughts at the same time?...
>
> It has always been my belief that we all began to work on a gasoline-engine-propelled road vehicle at about the same time because it had become apparent that civilization was ready for the mechanical vehicle… It has been the habit to give the gasoline-engine all the credit for bringing the automobile… In my opinion this is a wrong explanation…
>
> The reason why we did not build mechanical road vehicles before this [1895], in my opinion, was because the bicycle had not yet come in numbers and had not directed men's minds to the possibilities of independent, long-distance travel over the ordinary highway. We thought the railroad was good enough. The bicycle created a new demand which it was beyond the ability of the railroads to supply. Then it came about that the bicycle could not satisfy the demand which it had created. A mechanically propelled vehicle was wanted instead of a foot-propelled one, and we now know that the automobile was the answer.

Like Maxim, many historians credit the bicycle with altering Americans thinking about the possibilities of transport. While the railroad had brought rapid long-distance transport, passengers had to adjust their schedules and routes to rail schedules and were subjected to unfair pricing policies. The bicycle brought both independence and relative speed, which agreed with and reinforced an established and growing cultural focus on the individual. Thus, it was the bicycle craze of the 1880s and 1890s that began to give many entrepreneurial and mechanically minded individuals the same vision. Riders like Maxim began to think of adding mechanical power to relieve the hard physical labor required to travel distances of any consequence. According to Abernathy (1978, p.11):

> Emerging consumer needs, not new technological capabilities, triggered the rapid development of the U.S automobile industry at the turn of the century. A practical steam-powered car could have been produced twenty years earlier… [T]he industry was born from the consumer's desire for a light personal transportation vehicle, a desire stimulated by the bicycle boom of the 1890s. Hitherto, the motorpowered vehicle had been envisioned as a product for the commercial transportation industry, not the consumer.

From the perspective of bicycle's technological paradigm, the success of the bicycle had altered the TIEC in terms of consumer perceptions of personal travel. Consumers wanted better performance in terms of speed and distance than the bicycle, based on human power, could provide. Development of the bicycle was reaching technological limits. By stimulating a demand which it could not fulfill, its very success would ultimately lead to the industry's senescence.[28] It was this change in the TIEC – the developing desire for rapid long-distance personalized transport – that entrepreneurs in the late 1880s began to exploit to create economic value out of the invention of motorized transport that had occurred to Sir Isaac Newton as early as 1680 and had become technologically feasible with the steam engine in 1860. Serendipitously, entrepreneurs now could also exploit the innovation of the practical gasoline engine.

3.5.2 Initial Niche Markets, Product Definition, and Design

Two developments that originated with the bicycle regime – the alteration of perceptions regarding personal transport and approaching technological limits – explain why entrepreneurs began experimenting with motor vehicles when they did. However, these developments do not fully explain why the automobile found commercial success at that time and in the markets that it did. This section identifies four additional factors that contributed to the commercial success of the motor vehicle and to the definition of its function and initial niche markets: 1) difficulties moving forward with the existing carriage regime for providing urban transport; 2) difficulties in the railroad regime: abuses of monopoly power and price discrimination; 3) a growing cultural trend known

[28] Due to later changes in the TIEC and resultant shifts in consumer preferences, the U.S. bicycle industry would eventually be reborn with the functional characteristics of the product redefined strictly for sport rather than practical transport.

as Progressivism; and 4) new cultural values for sport, recreation, and the healthful benefits of escaping urban centers. Thus, the emergence of the automobile industry can only be understood by considering misalignments in multiple regimes along with other developments in the TIEC. This finding provides additional support for the inclusion of technical, institutional and ecological factors among the triggers of innovation. It also supports the assertion that a framework for analysis of transitions should include multiple regimes as discussed in section 2.4.4 and accounted for in the framework described in section 2.6.

With its unstable design, the first commercially successful bicycle, the ordinary, limited cycling to a sport for adventurous young men. Thus, the bicycle craze began with the popularity of racing among the wealthy. After the introduction of the safety bicycle, cycling became accessible to men, women, and children of all physical abilities. The bicycle could be used as a form of clean urban transport for short trips, and while it found some use in this function, for the most part it remained functionally defined as providing sport and recreation. The new middle class escaped the increasingly crowded and polluted cities to enjoy the healthful benefits of cycling in parks and the countryside. However, longer trips and transport of any cargo required too much effort for a 'human-powered' vehicle. Therefore, the bicycle had established new niche markets but for the most part did not compete with horse-drawn vehicles. Yet the bicycle aroused a desire to combine the functionalities of flexibility and speed with long-distance travel. Entrepreneurs interpreted this challenge as an opportunity for a radically new technology – motorized transport, most commonly in the form of four-wheeled motorized carriages and wagons but also as motorized bicycles.

Because of the technology's complexity and the state of industrial production, the first motor vehicles were costly to manufacture. However, the idea of individualism, freedom, speed, and the attendant danger appealed to wealthy young men who were willing to pay for excitement. As a result, the earliest applications of motor vehicles were racing and cross country touring. Much as bicycling originally had been conceived as a sport for the wealthy and daring, the earliest automobiles were functionally defined as entertainment for the rich. Because of their cost, they also conveyed social prestige, as motoring provided a very public demonstration of wealth. By combining the flexibility of bicycles and horse-drawn vehicles with the speed and long distance capabilities of rail, touring provided a new function – rapid, cross country, personal transport – that fulfilled the desire for entertainment, sport, and thrill. In this function, motor vehicles did not compete with rail transport and carriages because those existing technologies could not provide it. However, demonstrations of the technology fueled expectations that motor vehicles might become accessible to the masses, freeing them from the tyranny of the railroads and the limitations of the horse. This niche was served by steam and gasoline-powered motor vehicles.

Entrepreneurs were experimenting with a variety of technological approaches and another market niche emerged for more practical use on shorter trips in cities where animal waste from the use of carriages was creating problems. This market consisted of wealthy businessmen, bankers, their wives, and also doctors. Electric vehicles were easy to operate, quiet and odorless and therefore were best suited to this niche. In this market, electric vehicles competed with carriages, electric rail systems, and bicycles. However, electric vehicles combined the advantageous features of each without the disadvantages.

They were as clean like electric rail and bicycles, flexible like carriages and bicycles, and did not require the physical exertion of bicycles or the stabling and upkeep of horses. Use of electric motor vehicles also did not require investments in costly rail lines. Because electric vehicles were also very expensive, they provided both practical transport and social prestige and were conceived as a direct replacement for fine carriages in urban transport.

The bicycle craze provided more than the beginnings of demand for personal transport and the inspiration for innovation. The earliest motor vehicle designs borrowed heavily from bicycle technology, employing tubular frames and pneumatic tires and transmitting power to the wheels using sprockets and chains. Some designs also followed bicycle architecture, such as Benz's tricycle, Daimler's motorcycles, and Riker's electric tricycle. However, early motor vehicle designs also evolved from carriages and steam traction engines. The materials, techniques, and organization of production arose largely from experience in the bicycle and carriage manufacturing trades, the former benefiting from process innovations developed for the arms and sewing machine industries.

3.5.3 Mechanisms of Learning and Innovation

Like most new technologies, the automobile was immature when introduced and required significant improvements in performance. The QE and MLP theories stress the importance of learning-by-interacting with users and learning-by-doing in this critical process. However, innovation during the emergent phase of the automobile industry was highly synthetic and resulted from cumulative progress in products and processes over more than a century of development. As discussed in this section, entrepreneurs relied on learning-by-doing to synthesize existing technologies from a variety of arenas into an innovative new application. Thus, spillovers were a significant, if not the primary, learning mechanism underlying the development of this new innovation.

Although superficially a radical innovation, the automobile arose from a continuous evolution in product and process technologies (Rosenberg, 1963) as well as consumer demand. Spillovers from the stationary engine and bicycle industries appear to be the most significant mechanism involved in innovation during the emergence of the motor vehicle industry. The automobile industry relied on the continuous, incremental, and cumulative technical progress in production techniques over more than a century in a myriad of seemingly unrelated industries: arms manufacturing, sewing machines, bicycles, machine tools, and steel working. As Rosenberg (1963) points out, the machine tool industry served a critical role in transmitting these spillovers in technical knowledge across industries. This knowledge was embodied in the machines and manufacturing facilities themselves, but also in human capital as vividly illustrated by the careers of Henry Leland, Ransom Olds, and the Dodge and Duryea brothers. This dynamic would continue to play out in the first decade of the 20th century.

Firms that stayed connected to customers and suppliers evolved into new applications of expertise previously developed for existing applications. In this manner, the evolution of specific manufacturing firms from the production of arms to sewing machines to bicycles to automobiles facilitated the transmission and accumulation of knowledge in the form of plants, capital equipment and human capital. This progression is illustrated by the Weed Sewing Machine Company, which contracted with the Sharps Rifle Manufacturing Company, acquired some of its facilities and personnel, and evolved into the Pope Manufacturing Company, which made bicycles then transitioned to automobiles. Thus, certain firms exhibited longevity and success in multiple industries over time based on the cumulative nature of knowledge and learning as well as the firm's willingness and ability to adapt to change – both technological and social.

Successful early automobile manufacturers, then, were "less likely to be individual inventors starting a completely new business than men who added the production of automobiles to an existing operation" (Rae, 1959, p. 8). New firms like the Ford Motor Company that lacked such legacies were largely assemblers of components purchased from existing machine shops and firms in the bicycle and carriage industries, as will be illustrated further in the next chapter. However, many of them benefited greatly from hiring individuals from these industries that brought with them the largely tacit knowledge of armory practice. Unsuccessful firms, like Duryea, were unable to bring process technology into the production of technically successful product innovations.

In general, entrepreneurs were relying on learning-by-doing to synthesize existing technologies from a variety of arenas into an innovative new application. Automotive pioneers were working with limited finances and, in many cases, only in their spare time while working paid jobs or running another business. Epstein (1972, p. 28) asserts that prior to 1900, road vehicles were not 'manufactured,' but were "crude products of home or workshop experiment," built using improvised tools and, wherever possible, borrowing parts from steam engines, bicycles, and carriages. Experimental designs were sometimes abandoned because their complexity made commercial fabrication impossible. According to Epstein, true commercial manufacturing did not begin until after 1900.

Contrary to Rycroft and Kash's (1999) assertion that the creativity in these early years were provided by individual entrepreneurs, it is clear that these innovations were produced by multiple agents who collectively synthesized existing artifacts and threads of thought from wide ranging industries and disciplines with personal inspiration and insight to create a new product or concept. Even the European inventors of gasoline engines and motor carriages generally were inspired by the work of others and worked with an assistant or partner. Etienne Lenoir was inspired by the work of Phillipe Lebon some 60 years earlier; Nikolaus Otto was inspired by Lenoir and collaborated with Eugen Langen, Gottlieb Daimler, and William Maybach. In the U.S., the Duryea and Dodge brothers worked together as a team and Ford relied on a talented team of engineers and metal workers, notably including C. Harold Wills.

Epstein (1972, p. 27-28) describes the American automobile entrepreneurs as experimenters that "both through the copying of foreign designs and through a great deal of more or less original work… drawing on the mechanical, chemical, and electrical developments of decade piled upon decade, succeeded in building practical self-propelled road vehicles…" They worked steadily to improve their designs and construction methods through the "synthesis and skillful application of more or less well-known principles in order to produce results of an order higher than was known before." While already evident, teams of entrepreneurs, innovating through synthesis, would become even more significant in the early 20th century.

3.5.4 Agents and Mechanisms of Change

The QE theory of transitions identifies the technological nexus, where selection and variation are linked, as a critical mechanism in the mutual technological and institutional adaptations that result in a new technology becoming embedded in the wider socio-technical complex. In this section, I identify the emergence of agents and institutions in the nexus role and trace any connections to existing regimes. The activities these agents of change supports the assertion that, in order to realize successful designs, entrepreneurs must behave as heterogeneous engineers, solving social, economic, and infrastructural as well as technical issues.

Automotive entrepreneurs emerged from existing regimes and represented stationary engine manufacturers, machine shop mechanics, and bicycle manufacturers. Contributions were also made by carriage manufacturers, though entrepreneurs from this industry would become more prominent after 1910. Although these agents of change arose from within related existing regimes, they pursued different technical approaches based on experience and geography. A number of established bicycle and sewing machine firms, particularly those in the East, began manufacturing automobiles as a sideline. These firms, most notably Pope Manufacturing, generally developed electric and steam cars. Meanwhile, new firms established by Midwestern entrepreneurs with expertise in stationary engines or machining, such as Ransom Olds, were more likely to pursue less developed and therefore riskier gasoline technology. These entrepreneurs opted to secure outside financing to establish new firms rather than place their existing businesses at risk.

While most innovators began by working in small teams tackling technological problems, their roles quickly evolved into that of heterogeneous engineers involved in promoting the technology, overcoming public perceptions, altering the use environment, and securing financing. Prior to the turn of the century, we see four examples of agents serving the emerging technological nexus role linking variation and selection. This role involves two aspects: interpreting demands from the selection environment into requirements for design, and altering the selection environment, in part by imposing the requirements of technology on the selection environment. Enthusiast publications and the popular press were the first institutions to emerge in this intermediary role. They facilitated communication among engineers and as such shaped the shared perceptions of designers and entrepreneurs. They also altered selection by influencing public opinion and expectations of motor vehicles, the technological options and specific manufacturers. In this role, they were not necessarily a source of unbiased information. Manufacturers also made extensive use of the press beginning with the initiation of the elaborate press interview by Albert Pope.

Albert Pope had a great deal of experience serving in the nexus role for bicycles and translated much of that experience to the marketing of motor vehicles. He was active in cycling journals and cycling clubs that lobbied for improved roads, and after 1900 this movement shifted to the automobile industry and began to bear fruit (see section 4.5.6). This movement represented attempts to impose the requirements of technology on the selection environment. Pope organized the first bicycle trade show to link manufacturers and consumers, a practice adopted by the automobile industry, with the first one held in conjunction with the bicycle show. Trade shows advertised new models and provided manufacturers with feedback on new designs and innovations.

Third, Ransom Olds understood that the public was still skeptical of the technology and believed motor vehicle design therefore needed to be greatly simplified to allow anyone to operate the vehicle and any local shop to repair it. This interpretation of the selection environment resulted in the design of the first mass produced automobile – the quite successful Olds curved dash. Olds also influenced the selection environment through long-distance demonstrations that altered public perceptions that the automobile and long-distance touring were privileges for the wealthy.

Fourth, automakers began using speed and reliability races to influence public opinion of automobiles generally and demonstrate the capabilities of their technical approaches and products specifically. However, races reinforced the image of motor vehicles as toys for the wealthy. Races did more than alter public perceptions and expectations; they also stimulated the interest of investors and helped swayed their opinion in favor of gasoline engine technology, which was most in need of financial backing. By directing resources to particular firms and technologies, racing altered variation as well as the selection environment.

3.5.5 Summary

In the 1890's, the motor vehicle emerged as a technological innovation which, though radical as a synthetic whole, resulted from cumulative progress in products and processes over more than a century of development. The primary mechanisms of innovation in this era were knowledge spillovers and learning-by-doing. Conceptions of motor vehicles and experimentation trace back to at least the 17th century. Therefore, the automobile was not a new invention, but it emerged as an innovation (a commercial product) when five developments in the TIEC provided opportunities that entrepreneurs were able to exploit:

1) the success of the bicycle which altered consumer perceptions about personal transportation and created a demand for performance characteristics (greater speed, distance, and cargo capacity) that the technology was unable to fulfill due to technological limits;

2) difficulties moving forward with the existing carriage regime for providing urban transport;
3) difficulties in the railroad regime: abuses of monopoly power and price discrimination;
4) a growing cultural trend known as Progressivism; and
5) new cultural values for sport, recreation, and the healthful benefits of escaping urban centers.

Only the combination of these developments explains both why the automobile emerged in the 1890s rather than earlier and why it emerged in two specific niche markets. This supports the issue raised in section 2.4.4 that an appropriate analytic framework should account for multiple regimes, i.e. multiple meso-level systems, as well as institutional developments.

Initially, motor vehicles served two niche markets, both consisting of very wealthy consumers, serving the functions of: 1) local urban transport and 2) country touring. Because of the specific requirements of these two niches, the technological solutions best suited to each differed. Electric automobiles were used for short trips within cities and were the technology of choice for wealthy women. They provided practical urban transport, conveyed social prestige, and, because they offered both cleanliness and flexibility, were attractive as a solution to rising problems with the existing animal-powered transport regime in this function. Thus, it was emerging misalignments in *multiple* regimes that provided an opportunity for electric cars to enter the urban niche, where they competed with carriages, electric rail cars, and bicycles.

Steam-powered and gasoline ICE vehicles were used for touring the countryside. This use represented a new function by combining the flexibility of horse-drawn vehicles and bicycles with the speed and long-distance capabilities of rail transport. This function – rapid, cross country, personal transport – provided sport and recreation and conveyed social prestige. Motor vehicles in this application did not compete with the existing technologies for practical long distance transport (rail and carriage). However, to some extent, they did compete with bicycles as wealthy urbanites began driving electric-, steam-, and gasoline-powered motor vehicles in city parks for recreation.

Thus, the role of niches in the emergence of the motor vehicle industry is consistent with the MLP and QE theories in the following ways:

1) Radical technologies emerge in niches;
2) Instabilities or misalignments in existing regimes provide opportunities for radical technologies to serve as a solution;
3) The articulation of new functionalities creates new niche markets;
4) Niches serve to demonstrate technical feasibility relative to niche market needs;
5) Niche markets serve to demonstrate the relative advantages of alternative technological approaches; and
6) Within niches, firms develop technical capabilities, establish reputations, and form networks for financing, manufacturing, and sales.

However, while the QE and MLP approaches focus on a single regime, this analysis shows that considering developments and misalignments in multiple regimes is critical to understanding the emergence of the automobile industry.

Agents and organizations serving in the emerging technological nexus role began to emerge, with noticeable roots in related regimes. Many of these institutions included individuals like Albert Pope that had been active in these related industries and others were patterned after institutions established in these industries. The Good Roads movement began with cycling enthusiasts, and trade shows, press interviews, and racing evolved from customs in the bicycle regime. Agents experienced with stationary engines, notably Ransom Olds and Henry Ford, emerged as important agents interpreting design requirements based on market and infrastructural requirements while also attempting to alter perceptions about gasoline-powered vehicles and the function they could serve. The activities of these agents of change support the assertion that, in order to realize successful designs, entrepreneurs behave as heterogeneous engineers, solving social, economic, and infrastructural as well as technical issues.

The turn of the century approximately marks the end of the emergent phase, which is characterized by a variety of product designs and configurations, and the beginning of the transitional phase, where products begin to converge on a standard design. While there is no sharp delineation between the emergent and transitional phase, the year 1900 marks the introduction of the first mass produced automobile, the Olds curved dash. The Olds Motor Works produced about 1,400 vehicles that year and was producing up to 40 vehicles per day by 1904. Several manufacturers would soon achieve similar production volumes. Thus, ‘true’ commercial production began around 1900 (Epstein, 1972).

4.0 Transition and Onset of the Specific Phase

"…the motorcar business is a fusion of three arts – the art of buying materials, the art of production, and the art of selling."
-- Sorensen, 1956, p. 93

By around 1900, the automobile industry was entering the transitional phase, where products begin to converge on a standard design. This phase is particularly critical to understanding technological transitions because it also involves the movement from niche to wider markets. Two developments are critical to diffusion into wider markets: defining or refining the product's functional requirements to appeal to the needs of a larger customer base and improving the product's performance relative to these requirements. The quasi-evolutionary (QE) and multi-level perspective (MLP) theories assert that these developments occur within initial markets through learning-by-interacting with users and learning-by-doing. Consequently, growth of niche markets, niche branching, or niche proliferation allows manufacturers to reduce costs through economies of scale. During this development, the technology begins a process of embedding that is characterized by the co-evolution of the technology and the selection environment.

This chapter continues the history of the automobile through approximately 1915 in order to identify the technological improvements to both products and production processes that enabled movement into a wider market. Most importantly, I seek to identify the role of niche markets in this transition and the mechanisms of learning involved in these improvements. I also identify the emergence of rules, other institutions, and artifacts which contributed to the onset and stability of the specific phase. This review also traces the continuing development of the technological nexus linking variation and selection, identifies agents serving the nexus role, identifies activities which link variation and selection, and describes the institutional adaptations that co-evolve with the technology. Other institutional factors shaping this evolution, including exogenous events, also are identified.

Section 4.1 reviews the history of the Ford Motor Company, beginning with Henry Ford's struggles as a heterogeneous engineer to obtain financial backing. Both product and process innovations are discussed. While learning-by-interacting with users and learning-by-doing were important aspects of this process, they were not the only or even primary mechanisms of learning. First, Ford's own experiences and perspectives were important in determining users' needs in the rural market segment. Second, learning-by-interacting with suppliers were critical in the product innovations that significantly improved the performance and reliability of Ford's vehicles. Third, the innovative process was highly synthetic and spillovers played a fundamental role in both product and process improvements. By 1910, Ford's product was consistent with a new functional definition that suited the very large rural market. The company's production processes were becoming increasingly aligned with broader developments in the technical, institutional, and ecological complex (TIEC). By 1915, Ford had developed a revolutionary production system that reinforced these trends and became the industrial model for 20th century. This section concludes with a discussion of the role of innovations, institutional factors, and Ford's experiences and perspectives in the company's move toward vertical integration and centralized control as the industry entered the specific phase. This progression would have important consequences for Ford's innovative capabilities in the following era.

Section 4.2 reviews the formation of General Motors, beginning with the Buick Motor Company. Once again, a clear picture emerges of entrepreneurship as a heterogeneous endeavor, with William Durant jointly addressing issues of design, production, management, marketing, and finance. The success of the Model 10 in 1908 testifies to the rising importance of the low-priced market. While the product design during the transitional phase was converging on a dominant design, manufacturers did not know what that design would entail, which presented them with substantial risk. While Ford attempted to influence the selection environment with his interpretation of the automobile and appropriate product attributes, Durant attempted to mitigate the risk presented by the converging dominant design by creating GM, an industrial combination that diversified his product line in terms of technological approaches and market segments. However, by 1910, only Buick and Cadillac were profitable. This can be attributed to two factors: 1) the emergence of a new functional definition that appealed to a wider market and an aligned dominant design which supplanted other technological interpretations and configurations; and 2) manufacturing processes at GM subsidiaries that were not aligned with the TIEC and market needs for the combination of high reliability and low cost. These problems, combined with an aggressive acquisition and expansion strategy, contributed to the transfer of control to a banking syndicate between 1910 and 1915. While the bankers did improve production processes and institute a research program, they also increased the company's misalignment with these emerging conditions. Finally, this section provides a detailed description of the development of the electric self starter 1911. This review reveals that spillovers and learning-by-interacting with suppliers were the most significant mechanisms in this critical innovation.

Section 4.3 reviews Durant's formation of the Chevrolet Motor Company. The firm's failure with expensive touring vehicles and its contrasting success with small inexpensive vehicles again illustrate the increasing importance of the low-price market and the changing functional definition of the automobile. This success also paved the way for Durant to re-take control of GM, which is covered in chapter 5.

Section 4.4 discusses the severe industry shakeout that occurred beginning in 1909. A shakeout is typical as industries approach the specific phase, a dominant design emerges, and firms begin competing based mainly on cost and quality rather than product differentiation. However, the shakeout for the automobile industry was extreme. I identify eight factors that contributed to this phenomenon, several of which are inter-related: 1) increasing economies of scale that necessitated high volume production and high capital investments; 2) the cumulative burden of manufacturing parts for multiple models; 3) process innovations at Ford that were not easy for firms to adopt; 4) the cumulative effects of and increasing returns to innovation; 5) the redefinition of the automobile for mass markets; 6) an emerging dominant design; 7) the saturation of the market niche for very expensive vehicles; and 8) the Panic of 1907 and ensuing recession. Items 5-7 together represent an important feedback mechanism between selection and variation. By contributing to the failure of specific firms, the selection environment influenced variation and further reinforced the emerging dominant design.

Section 4.5 discusses five developments in the TIEC that provide the technical and institutional context for this era: the Panic of 1907; the Selden Patent patent dispute; the emergence of trade associations and their efforts to standardize parts and materials; the expansion and evolution of the autombile market; the activities of motoring clubs;

and efforts to improve the condition of the country's roads. Finally, section 4.6 synthesizes this review and summarizes the findings as they relate to the research questions presented in chapter 1.

4.1 Fordism and the Model T

Despite an early interest in motor vehicles, Henry Ford was not one of the first American entrepreneurs to break into the automotive industry. Ford's first attempt at commercial manufacturing produced fewer than 25 vehicles before being dissolved and the second firm succeeded only after Ford's departure. Ford's third attempt found commercial success with production of the Model A in 1903. By the end of that year, the Olds Motor Works already had produced an estimated 10,800 motor vehicles. Over the next ten years, the Ford Motor Company revolutionized the manufacture of motor vehicles and quickly overtook Olds' production. Using a combination of product and process innovations, Ford achieved his goal of producing a car that every man could afford. The company's production system, which came to be known as 'Fordism,' was so successful at cost reduction that it soon became the industrial model for 20th century production.

This chapter begins in section 4.1.1 with Henry Ford's background, which shaped the perspectives that led him to experiment with motor vehicles and to establish a goal to produce an automobile that could replace the family horse. These perspectives, which aligned with evolving higher-level rules within the TIEC regarding capitalism, mechanistic reductionism, and increased cultural valuation of efficiency, served as a basis for Ford's highly successful innovations between 1900 and 1915. Ford also needed financial backing, but his conception of the automobile was not aligned with the initial product definition presented in chapter 3. This section describes his struggles as a heterogeneous engineer to obtain funding and sway the opinion of financiers and customers regarding the function of the motor vehicle and the attributes necessary to fulfill it. During this process, Ford made use of two institutions described in chapter 3 that were emerging in the technological nexus: the press and racing.

Section 4.1.2 then reviews the product innovation and design process at the Ford Motor Company beginning with the Model A in 1903 and culminating with the Model T in 1908. Ford targeted the rural market with the company's least expensive models, most notably the A, N and T. While the QE and MLP theories place primary importance on learning-by-interacting with users to determine users' needs, Ford relied on his own childhood experience to understand users' needs in the rural market segment. However, he did rely on feedback from dealers and buyers to identify defects as he focused on product innovations to improve reliability. Consistent with the QE and MLP theories, Ford showed a marked preference for learning-by-doing over guided research, though C. Harold Wills did begin metallurgy testing at some point. However, this review also uncovers the critical role of learning-by-interacting with suppliers in the product innovations that significantly improved the performance and reliability of Ford's vehicles during the transitional phase.

In addition, process innovations at Ford discussed in sections 4.1.3 through 4.1.5 once again highlight the highly synthetic nature of innovation and the fundamental importance of spillovers in the product and process innovations that enabled Ford to manufacture a product of both high quality and low cost. While one or the other had been achieved within the industry, it was the combination that allowed Ford to conquer a mass market.

Section 4.1.3 covers the adoption and adaptation of process innovations in plant organization from New England armory practice, which was transferred to the Ford Motor Company through the ‘acquisition’ of human capital – this knowledge was embodied in key personnel with experience in the sewing machine and machine tool industry. Section 4.1.4 discusses a second aspect of Ford’s evolving production process, the management of production, specifically the breakdown of manufacturing and assembly operations into minute tasks. This section places these developments within the context of a wider movement toward ‘scientific management,’ the formulation of which is attributed to Frederick W. Taylor. The question arises as to how much of the innovation at Ford represented original inspiration versus knowledge spillovers transmitted either directly from the work of Taylor or indirectly through shared perceptions of the knowledge base. As with the emergence of the automobile itself, multiple entrepreneurs were working separately yet concurrently on the same problems and were developing a variety of similar solutions. I find that this era of innovative ferment was attributable to pervasive changes in the TIEC and that the role of spillovers in Ford’s innovations was of primary significance.

Section 4.1.5 covers the third and final aspect of Ford’s process innovations, continuously moving assembly. Popular opinion generally regards this innovation as Ford’s most significant achievement and attributes its development to Henry Ford alone. However, this review reveals first that the development of continuously moving assembly at the Ford Motor Company was a highly synthetic process involving the talents and personal inspiration of multiple entrepreneurs and spillovers from multiple industries. In addition, moving assembly was just one component of several which, when combined, created a production system that was more efficient than the simple sum of the parts.

Finally, section 4.1.6 completes the history of the Ford Motor Company through 1915 and discusses the company's progression toward centralized control and vertical integration. Throughout the period covered in this chapter, Henry Ford clashed with his financial backers over strategy. This fact, combined with Ford's predisposition toward individualism and autocracy, led Ford to seek complete financial and managerial independence. Ford was assisted in achieving this goal by an unexpected exogenous event, the Panic of 1907. In addition, the price and availability of raw materials was always a concern, and supply holdups and bottlenecks became an increasing problem as Ford moved toward continuously moving assembly. Therefore, Ford also began a process to bring these supplies under his control. This movement toward centralized control and vertical integration would not be complete until 1918, at which point it had significant implications for the company's innovative capabilities.

4.1.1 Birth of the Ford Motor Company

Born in 1863, Henry Ford grew up in rural Michigan where he found the endless work on the family farm to be tedious. He developed an interest in all things mechanical at an early age and much preferred to tinker at his tool bench. Around the age of 12, Henry saw a steam traction engine moving along the road under its own power. Fascinated by the possibility of an engine being used for a self-propelled vehicle, he grilled the operator about the machine. The man was impressed and allowed Henry to operate the machine over the next few summers.

With his schoolwork finished at the age of 16, Henry Ford left the farm for Detroit. Between 1879 and 1882 he earned the first-hand mechanical experience he had desired all his young life. His first job as an apprentice at the Michigan Car Company, a large Detroit manufacturer of railroad cars, lasted only six days. Henry was discharged after angering the older hands by repairing in half an hour a machine they had spent most of the day trying to fix. Henry then spent nine months as an apprentice in a machine shop before moving on to Detroit's largest shipbuilding firm where he was employed in the engine works. Between 1882 and 1885, Henry worked various jobs, installing and repairing engines for the Westinghouse Engine Company and other employers, finding occasional work in Detroit factories, and lending a hand at his father's farm. He also attended business school for practical commercial training.

In 1887, Henry Ford's father offered him forty acres of wooded land, which Henry accepted so that he might marry Clara Bryant. He set up a sawmill, which provided a source of income and raw materials to build a new home. Sometime afterward, he built a small shop where he did repair work and reportedly began experimenting with steam and gasoline engines, aiming to build one to pull a wagon or plow.[29] Although no record of his work survives, it is clear that he had become interested in the idea of a horseless carriage. As the timber began running out on his property, he became restless; he had no desire to become a farmer as his father wished. Instead, he decided that he needed to learn more about electricity in order to build his horseless carriage. In 1891, the Fords moved to Detroit and Henry began working for the Edison Illuminating Company, where he was soon promoted to chief engineer. Shortly after the

[29] Except for public demonstrations, the timeline of Ford's experimentation and achievements is uncertain due to discrepancies between his accounts and those of friends and colleagues.

move, Ford's experimentation with horseless carriages began in earnest, and by the end of 1893 he reportedly had built a functional one-cylinder gasoline engine.

In 1896, Ford began working on a two-cylinder engine, chassis, and body with the assistance of two fellow Edison employees. Ford's quadricycle, as he called it, weighed only 500 pounds and used bicycle wheels with pneumatic tires. The experimental vehicle completed a successful test run in May, 1896. Ford improved the quadricycle then sold it so that he could begin building a new vehicle. After obtaining financial assistance, Ford began building a second horseless carriage. The quadricycle had been built almost entirely of wooden parts and Ford hired a carriage blacksmith to fashion metal parts for the wheels, seat railings, undercarriage, and steering mechanism. The new vehicle, completed in the summer of 1899, was still lightweight but was substantially stronger and capable of carrying both a driver and passenger.

In August, Ford secured financial backing to establish the Detroit Automobile Company and then resigned from Edison to devote himself full-time to the endeavor. Ford had a test vehicle ready in January, 1900, and the new company announced ambitious plans to produce ten vehicles a month beginning in March. However, Ford wrestled with problems of design while the shop encountered problems with assembly and the workmanship of parts supplied under contract. In addition, the company's car probably was too expensive, not as advanced as some others, and not quite ready for production (Nevins and Hill, 1954). Ford apparently believed the car would not sell and wanted to design a better car but was vetoed by the stockholders. Though it did produce a few vehicles, perhaps as many as 25, production at the Detroit Automobile Company soon ground to a halt and the company was dissolved in January, 1901.

Undaunted, Ford took a new approach. The public "refused to consider the automobile in any light other than a fast toy," so he decided to build a racer and use it to gain recognition (Ford and Crowther, 1922, p. 36). The race car, designed and built with the assistance of Oliver Barthel, was ready in time for an October 10 race sponsored by the Detroit Automobile Club. Ford was one of only three competitors that showed for the ten-mile race held at the Grosse Pointe track near Detroit. After one dropped out due to a leaking cylinder, only Ford and Alexander Winton remained. Winton took an early lead, but Ford was gaining when Winton developed engine trouble. Ford won the race with an average time just under 1 minute and 21 seconds per mile. Later that month, Ford made a credible challenge of Winton's newly established one-mile record of 1:09 before losing his nerve on the turn. Inspired by Ford's racing successes, his investors reorganized the old Detroit Automobile Company as the Henry Ford Company in November, 1901.

Ford hired Barthel full-time, ostensibly to work on the design of a vehicle for commercial production, but Ford instead immediately began plans for a new racer. While the Henry Ford Company's financers supported racing as a vital part of publicity, dissension arose over his preoccupation with it. Ford apparently was extremely dissatisfied with his share in the profits of the new company and felt he would make much more money from racing. As Ford increasingly neglected his duties as chief engineer, the stockholders pressed him to complete design of the commercial car and begin production. Eventually, they brought in the services of Henry Leland of Leland & Faulconer who apparently found fault with Ford's "cut and try" design methods. In March, 1902, Ford left the company. The stockholders gave him the incomplete plans for the racer and $900 and agreed to discontinue using his name. Henry Leland took over the company which was renamed the Cadillac Automobile Company.

Ford joined forces with famous cyclist Tom Cooper, who had ridden with him in the Grosse Point race the previous October, and a new assistant, C. Harold Wills. Wills had completed a four-year apprenticeship as a toolmaker, worked in a machine shop, and studied engineering, chemistry, and metallurgy at night. Much like Ford, Wills had keen mechanical instincts and preferred to learn by doing. He agreed to work for Ford early in the morning and after dinner so that he could keep his full-time job. Wills took charge of drafting and design of two new racers, "The Arrow" and the "999," the latter named after the Empire State Express No. 999, a New York Central train that in 1893 set a world speed record of over 112 mph. The cars were tested in September and the 999 was entered in a five-mile race at Grosse Ponte in October. Apparently both Ford and Cooper were intimidated by the vehicle's power, as neither was willing to drive in the race. Cooper found another cyclist who, though he had never driven an automobile, agreed to try it. The 999 won, setting a speed record by averaging 1 minute and 6 seconds per mile. The relationship between Ford and Cooper, however, had begun to cool. Cooper purchased the 999 and the two parted ways.

Since leaving the Henry Ford Company, Ford was no longer entirely focused on racing. Despite his prior business failures, he became determined to build a car to replace the family horse, and in the summer of 1902 began planning a motor vehicle of moderate cost. His new car included a novel vertical engine design that resulted in less vibration, noise and wear than Olds' horizontal engine. To obtain funds to produce a pilot model, Ford approached successful coal merchant Alexander T. Malcomson, with whom he had become acquainted while working at Edison Illuminating Company. This time, Ford was careful to ensure a profitable stake in the venture. In August, 1902, Malcomson agreed to an equal partnership in developing the pilot model and to establish a corporation to replace the partnership and bring Ford's automobile to market once they had sufficient capital for manufacturing. Ford took charge of engineering and manufacturing, while Malcomson was to handle the financial and commercial aspects of the business. Malcomson assigned James Couzens, his office manager and business advisor, to handle the accounts for the new partnership.

Following the victory of the 999 in October, the design team turned their full attention to developing the new commercial car. Ford provided general design concepts, and Wills helped develop them into specific plans. With the car, work force, and a rented shop on Mack Avenue coming along, Ford and Malcomson made plans for incorporation, but had difficulty finding willing investors. Nonetheless, they began approaching parts makers for contracts. Ford and Malcomson contacted the Dodge Brothers, which was one of the best machine shops in the Midwest, to manufacture the new engines. The Dodges had been building engines for Olds, but faith in Ford's new design and hopes for large future profits led them to contract to supply 650 engines, transmissions, and axles. The Dodges agreed to invest in machinery, tools, and materials in advance of payment, with installments due from Ford beginning in March, 1903. Ford and Malcomson arranged contracts for bodies, wheels, and tires, and were able to show potential investors the promising margin for profits from their automobiles.

Nonetheless, start-up capital remained elusive. The risks of investing in the automobile industry during this still experimental stage were well known, and three-fifths of the companies formed between 1900 and 1908 to produce automobiles would fail. Further, Malcomson was over-extended, debt-ridden, and distracted by his other business. His reputation as an impulsive ‘plunger’ did little to reassure potential investors. With the first payments to the Dodge Brothers and other suppliers looming, Ford and Malcomson were finally able to draw in a few investors. Among them were Malcomson’s cousin, attorneys, friends, and, most significantly his uncle, John S. Gray, a banker to whom Malcomson was already heavily in debt.

The Ford Motor Company was incorporated in June, 1903. In exchange for a $10,500 cash investment, John Gray was named president and granted 10.5% of the company stock. Ford and Malcomson each took 25.5% of the stock in exchange for work completed, plans, machinery and patents. The Dodge brothers were each granted 5% in exchange for $7,000 in materials and a $3,000 note. Couzens invested $1,000 and a $1,500 note and was granted 2.5% of the company. With only $28,000 paid in cash, the Ford Motor Company was born on a shoestring. Gray, Malcomson, Ford, John Dodge, and investor John W. Anderson were named as the board of directors. Henry Ford was named vice-president, Malcomson treasurer, and Couzens secretary. In practice, however, Ford served as president, handling all matters of production, and Couzens served as both treasurer and secretary, handling all business affairs (Nevins and Hill, 1954). Couzens would eventually develop an elaborate but highly efficient business organization. Sorensen (1956, p. 36) designates the first twelve years of the Ford Motor Company (1903-1915) as the Couzens period, giving the enterprising and severe individualist equal credit for the company’s success:

> True, the company had Henry Ford’s name, its product and production were his. There never would have been a Ford car without him. But the Ford Motor Company would not have made Ford cars long without James Couzens. He

> controlled expenditures, organized sales, and set the pattern for business operation. He drove Ford and the production side to produce cars to meet the public's demand. He yelled for plant expansion and drove us from the Piquette Avenue Plant into Highland Park. Everyone in the company, including Henry Ford, acknowledged him as the driving force during this period.

4.1.2 Product Design

This section reviews the product innovation and design process at the Ford Motor Company beginning with the Model A in 1903 and culminating with the Model T in 1908. Ford targeted the rural market with the company's least expensive models and relied on his own childhood experience to understand users' needs in this market segment. Ford focused on improving the reliability of these vehicles, relying on feedback from dealers and buyers to identify defects. This history shows Ford's marked preference for learning-by-doing over guided research, but also uncovers the critical role of learning-by-interacting with suppliers in the product innovations that significantly improved the performance and reliability of Ford's vehicles.

In 1903, the Ford Motor Company began producing the Model A in a rented shop on Mack Avenue in Detroit. The Model A design was simple, lightweight (the runabout weighed 1250 pounds), and efficient. The major novelty of the Model A was Ford's two-cylinder engine, which could produce eight horsepower and up to 30 mph. Ford contracted the Dodge Brothers for the engines, transmissions, and axles; the C. R. Wilson Carriage Company for the wooden body and cushions; the Hartford Rubber Company for tires; and the Prudden Company for wheels. Ford employees assembled four cars at once, aiming to finish 15 per day. The runabout sold for between $750 and $850.

Although advertised as reliable and rugged, in addition to inexpensive, the first Model A's suffered from many design and manufacturing faults. The radiators were defective, the carburetors were inefficient, and the brakes and engines were often assembled hastily and shoddily by the Dodge Brothers. Although serious, these problems were not a great threat to the longevity of the Ford Motor Company since, in 1903, all automobiles were similarly defective, and no one expected them to be dependable. Complaints poured in from dealers, one of whom sent a list of recommended changes. Ford brought the dealer to Detroit and worked through the problems.

Ford initially replaced the faulty carburetor with one from a different supplier, but the replacement was still unsatisfactory. Ford then contacted George Holley, who had garnered a reputation manufacturing a one-cylinder car, motorcycles, and a simple carburetor. Using principles laid down by Ford and Wills, Holley designed a new carburetor that finally solved the problem. Suppliers also marketed their products to improve the vehicle. For example, Peter Steenstrup of Hyatt Roller Bearing admonished Ford that he had failed to honor a promise to use his company's bearings instead of ball bearings and Ford eventually switched. Within 6 months, the Model A design had been reworked (Nevins and Hill, 1954).

Sales were better than expected and the profits allowed for payment of dividends while still financing expansion, all without the need for raising working capital on Wall Street. Since many car-buyers were suspicious of high finance, a Wall Street connection likely would have negatively impacted sales. In 1904, the stockholders approved the construction of a new factory on Piquette Avenue. In the 1904-1905 sales year,[30] Ford offered three models, all somewhat larger in wheelbase than the Model A. The Model C, which replaced the Model A, weighed 1250 pounds, delivered 10 hp, and sold for $800. The new Model F was a touring automobile that weighed 1400 pounds, delivered 12 hp,

[30] The sales year began in the fall. Because early automobiles were open to the elements, Northerners did little driving from Thanksgiving through Easter. Accordingly, few automobiles were sold during this period and sales agents were laid off (Nevins and Hill, 1954).

and sold for $1000. Malcomson insisted on tapping the wealthy market and Ford introduced the Model B which weighed 1700 pounds, incorporated a four-cylinder engine that produced 24 hp, and sold for $2000. Where the Model A, C and F were all chain-driven, the Model B had a rotating-shaft or torque drive. The new models were better than the A but still had many defects. Sales continued to climb and production reached 25 automobiles per day in 1905 (Nevins and Hill, 1954).

Shortly after occupying the Piquette plant in late 1904, Ford began planning two new models: the moderately priced four-cylinder, 15 hp, Model N, which would improve on and replace the C and F, and the six-cylinder, torque-driven Model K, which was priced at $2800. As expected, Malcomson championed the K, bolstered by the industry trend toward larger, more expensive cars.   In 1903, 26% of the automobiles produced sold for under $875 while 12% were over $2775. In 1905, the situation was reversed, with 10% of vehicles selling for under $875 and 25% for over $2775 (Epstein, 1972).

The industry and the Ford Motor Company were moving ever farther away from Henry Ford’s ideal of manufacturing a single automobile of simple, standardized design that could be sold at $500, enabling the masses to replace the family horse. Malcomson seemed justified, however, as many contemporary industry analysts wrote that the country was prospering and too many manufacturers were making inexpensive automobiles while there were too few expensive models available. Henry Ford was not alone in his battle; Olds Motor Works had been the first to produce in volume a lightweight, inexpensive, and highly successful motorcar, yet even Olds’ financial backer pushed for higher-priced models. Unhappy with what he saw as interference with his decision-making, Ransom Olds in 1904 resigned from the company that bore his name.

Ford, however, remained adamant; he wanted to build a car that workingmen could afford.

Although the Model N was not completed in time for the New York Automobile Show in early 1906, Ford showed a specimen without an engine, advertising a price of $500. The response was immediate and overwhelming; Ford's championship of the "car for the ordinary man" was finally vindicated. The Model N incorporated ideas that had been developed for the expensive models B and K, including a torque-drive, yet was light, weighing 1050 pounds, and inexpensive. The engine was located under a front hood, lamps were added to the radiator and dash, and mud guards and polished fenders were added.

Ford planned production of the Model N at 100 per month beginning in July of 1906, with the engines, axles and gears manufactured by the Ford Manufacturing Company at the Bellevue plant.[31] Such a large production volume was previously unheard of and, by this time, Henry Ford was convinced that this feat would require fully interchangeable parts and careful plant organization for progressive flow. Rapid expansion of the automobile industry complicated the acquisition of the necessary machine tools, but production was in full swing by the fall.

Dealers and customers were highly enthusiastic, and advance orders for 1907 production streamed in. Nevins and Hill (1954, p. 323) declared the new Model N "one of the best designed cars yet seen in the United States." Nonetheless, the Model N suffered from defects and production problems. The biggest challenge was manufacturing interchangeable camshafts of sufficient strength that also provided proper

---

[31] See sections 4.1.3 and 4.1.6 for the history of the Ford Manufacturing Company and the Bellevue plant.

timing (Sorensen, 1956). Sales were brisk, but Ford was already looking ahead. Late in 1906, he set up the design room for the next Ford automobile: the Model T.

Since founding the Ford Motor Company, Henry Ford's goal had been to produce an automobile for mass consumption, and each small car model moved him closer to his goal. The car for everyman would have to be rugged, reliable, and easy to operate, maintain, and repair. To produce it in volume, it needed to be designed for ease of manufacturing as well. The comparative simplicity of Ford's Model N and T designs made them easy to understand and use, and also was well adapted to mechanized, mass production techniques. Ease of repair was facilitated by the availability of standardized, precision-machined, yet inexpensive mass-produced parts. The Ford models' lightness and rugged designs meant that they could negotiate the rough roads and weather conditions better than costlier and heavier models and remain structurally sound.

To correct the faults of previous models and adapt the vehicle design to the rugged conditions of American roads, the Model T's body was lifted high above the road. It had stout wheels and springs which negotiated rough terrain and partially relieved the body and chassis from the painful wrenching suffered by most cars. It was offered with two carburetors, one of which Holley designed and then manufactured in a Detroit factory that Ford purchased for him in exchange for a full order of carburetors. Dry cells for ignition had been replaced by a magneto, and the engine and transmission were entirely enclosed. For the first time, the engine cylinder block and crankcase were made in a single casting with a detachable cylinder head, which simplified both design and manufacturing once a suitable gasket was developed. The engine utilized a three-point suspension to eliminate the distortion of the motor base from operation on rough roads, and the transmission was an improved planetary type, both of which had been first introduced by Northern in 1902 (Abernathy et al., 1983, Appendix D).

The rugged, easy to operate transmission was key to the mass appeal of the Model T. The soft metal used in previous transmissions meant that gears were easily stripped by motorists who had little experience with shifting, and many drivers, especially women, had difficulty with the heavy, sticky clutches of previous gasoline-powered automobiles. For the first time in automotive manufacturing, the Model T used vanadium steel gears, crankshafts, axles, and springs (Nevins and Hill, 1954; Sorensen, 1956). According to Abernathy (1978, p. 96), "The Model T was superior as a car in its era because it innovated in the use of a vanadium-steel alloy to achieve a high strength-to-weight ratio and could be powered by a moderate-sized … engine." Vanadium steel provided more than twice the tensile strength of the steel it replaced and yet was easier to machine.

However, consistent with Rosenberg's (1963) thesis of technological convergence and the reactive forces behind innovation, the use of vanadium steel required improvement in the machine tools used to work on it. According to Sorensen (1956), when test parts for the Model N were forged with the new alloy, it was discovered that better drills and cutting tools would be needed to machine it.

The introduction of vanadium steel appears to have inspired Ford to design the Model T, and certainly determined its final form. The origin of this innovation is a bit of a mystery, as accounts vary. According to Nevins and Hill (1954), Ford wrote in his reminiscences that, in 1905, he picked up an engine part from a French racecar following a wreck at Palm Beach. Noting that it was very light and tough, he had a company mechanic look into its composition and found that it was a French vanadium steel alloy.

Ford then sent for metallurgist J. Kent-Smith of England and subsequently arranged for a steel mill in Canton, Ohio, to experiment with the new alloy. In another account, C. Harold Wills reports that he saw an exhibit of the steel at an engineering conference in 1905 and subsequently had it tested.

The third account by Charles Sorensen[32] is the only one where a Ford official is not claiming for himself credit for pioneering the use of the high strength steel. Kent-Smith had worked extensively with vanadium steel since 1900 and his work was widely published in technical journals. According to Sorensen, Ford executives therefore knew about the alloy from their readings (Nevins and Hill, 1954). Kent-Smith made his first U.S. contacts in Canton, Ohio, and then traveled throughout the country with samples of the new steel. When he visited Ford several times in 1905, company officials met him with great interest. Afterward, Henry Ford told Sorensen, "this means entirely new design requirements, and we can get a better, lighter, and cheaper car as a result of it" (Sorensen, 1956, p. 98). The lack of self-interest in this final account perhaps lends it more credibility than the others, since Henry Ford was known to embellish stories. Further, Nevins and Hill (1954, p. 349) conclude that Kent-Smith "seems to have come to the Piquette plant on his own initiative." However, Sorensen had his own priorities and loyalties and as such was not an unbiased source either. Regardless, in any version of the story, supplier interactions were critical in this major innovation.

Supplier interactions also played a significant role in the development of stamped steel parts for the Model T. Sometime in 1907, William H. Smith of the John R. Keim

[32] Charles Sorensen began working at Ford as a pattern-maker in 1905 and within a year was in charge of the pattern-making department. After Walter Flanders and Max Wollering left Ford in 1908, P.E. Martin served as plant superintendent and Sorensen served as assistant superintendent, though Ford disdained the use of job titles (Sorensen, 1956).

mills in Buffalo showed Henry Ford the stamped steel bell housing of a telephone receiver and suggested making a full-scale axle housing using the same process.[33] Ford was reportedly delighted by the idea of using such parts for both the current Model N the design of the new Model T and sent Sorensen and Wills to tour the plant (Nevins and Hill, 1954). Sorensen subsequently proposed using a stamped steel cover for the Model T crankcase and transmission (Sorensen, 1956). By 1908, the Keim plant was making large quantities of pressed steel parts for Ford cars.

The design process for the Model T is described in fairly vivid detail in the reminiscences of company officials. Henry Ford was clearly the originator of design concepts, which he described to a small engineering team led by Joseph Galamb and Harold Wills. Ford would make rough sketches on a blackboard and Galamb and his assistants would draw up graphic sketches and blueprints. Ford photographed the blackboard to document their progress in case of any patent suits. To facilitate visualization of the design concepts, Sorensen built wooden mockups and castings of parts, often before blueprints were drawn up.[34] These ideas were then presented to a group of machinists to work out details and test them. Meanwhile, Wills took charge of testing the new vanadium steel alloys which were fundamental to the lightweight design. This process of "cut-and-try" design continued for more than a year before the Model T was introduced. Sorensen, however, asserts that it took more than four years to develop

[33] Pressed steel rear-axle housings were first introduced by the Fiat Company around 1907 (Nevins and Hill, 1954).

[34] It is well documented that Ford preferred models over blueprints and several contemporaries assert that he had difficulty reading blueprints. Nevins (1954) finds this unlikely since Ford's machine shop apprenticeship would have taught him to draw and use blueprints.

the new car since previous models served as experimentation on the road to realizing Henry Ford's dream of a car for everyman (Sorensen, 1956; Nevins and Hill, 1954).

The Model T was first advertised in the spring of 1908 with prices beginning at $825. The Model T was the synthesis of various product designs and was consistent with an emerging dominant design: a water-cooled, 4-stroke, 4-cylinder gasoline engine mounted in front of the driver; left hand drive using a steering wheel; shaft drive; magneto; and pneumatic tires. The Model T's only major departure from the industry standard was the use of a planetary-type transmission. Ford assumed that his main market would be farmers who were experienced in repairing farm equipment. Therefore, the Model T came with a 64-page owner's manual that explained how to use simple tools to repair any of the 140 problems likely to occur (Womack et al., 1990). At the same time, Ford trained dealers how to handle stocks of parts and how to do repair work, and also supplied parts lists that included prices. While the new Model T provided unsurpassed simplicity, practicality, dependability and quality at an incredibly low price, it was still fraught with problems. Yet the Model T performed remarkably well in contests, negotiating steep grades, rough roads, mud, and water obstacles. It even performed admirably in cross country speed races. Orders for the new Ford flooded in.

4.1.3 Plant Organization

In order to manufacture a vehicle that could replace the family horse, Ford needed to produce an automobile that was simultaneously high quality and low cost. Process improvements at the Ford Motor Company began with the adoption and adaptation of innovations in plant organization that originated with New England armory practice. This knowledge was embodied in key personnel with experience in the sewing machine and machine tool industry. However, its adaptation to automobile manufacturing and assembly was accomplished through learning-by-doing on the shop floor.

Late in 1905, Henry Ford established the Ford Manufacturing Company to begin manufacturing parts for the new Model N and as a means to gain control of the Ford Motor Company. Ford rented a factory on Bellevue Avenue for the new company and began equipping it to produce engines and small parts for the Model N. According to Nevins and Hill (1954, p. 324), the division of work between the two factories, followed by reorganization and expansion of the Piquette facility, enabled Ford "to take the first steps, 1906-08, toward a production system that within the next decade became not only world famous, but a world force."

In 1906, Ford hired Walter E. Flanders, a Yankee machinist who had worked for the Singer Manufacturing Company and witnessed quantity manufacture firsthand. He had later built tools for the Landis Tool Company, a pioneer in precision automotive grinding. Although Flanders' experience with Singer was not true armory practice, he conveyed to Ford officials and mechanics that complete interchangeability would be absolutely essential for assembly under high volume production. While the Ford Manufacturing Company's plant superintendent, Max Wollering, claimed the idea was not new to him, he conceded that it may have been new to the company. Whether Flanders was responsible for bringing the concept to Ford is unclear. Nonetheless, in 1906, the Ford Motor Company began advertising that it was manufacturing in unprecedented volume and that every part of each Ford car was made to be exactly alike across all 10,000 automobiles produced. Yet the company's factory and contractors had not yet achieved interchangeability (Nevins and Hill, 1954; Hounshell, 1984).

Flanders and Wollering were responsible for rearranging machine tools according to sequential operations rather than according to machine type and they introduced Ford to the idea that productivity gains could be realized through the use of special- or single-purpose tools (Hounshell, 1984). But Flanders brought more than an appreciation of the fundamentals of armory practice; he taught Ford officials that "the motorcar business is a fusion of three arts – the art of buying materials, the art of production, and the art of selling" (Sorensen, 1956, p. 91). The large demand for the Model N allowed Ford to establish monthly delivery contracts with sales agents, which allowed Flanders to plan production for steady output. Late in 1906, when demand began to exceed capacity, he set up a 12-month production plan and established long-term contracts with suppliers that required them to carry inventories, allowing Ford to reduce theirs to only 10 days. This freed up funds that became critical to design and planning for the Model T (Sorensen, 1956).

When Ford gained financial control of the Ford Motor Company in 1907, he consolidated the two Ford companies, bringing the Bellevue equipment and Flanders to Piquette. Flanders and Wollering subsequently reorganized the placement of machines and flow of materials at Piquette, effectively introducing "the fundamentals of an admittedly modern version of New England armory practice to the handful of young mechanics Ford had assembled" (Hounshell, 1984). However, the orderly progression of materials and parts was limited by the plant itself, as was total production. The huge volume of advance orders for the Model T in 1908 outstripped the capacity of the Piquette plant and Ford began planning for the construction of a third plant at Highland Park on the outskirts of Detroit. This time, progressive operation – logical sequencing of operations and machine layout – was an integral part of the facility's design. Construction began at Highland Park in 1908 and the move to the new plant began in January, 1910, when only a quarter of the construction was complete (Nevins and Hill, 1954). Prior to occupying the Highland Park plant, Ford production had utilized fairly standard machinery. But in equipping the new plant, Ford began designing special-purpose tools in earnest (Sorensen, 1956).

As sales of the Model T increased, shipping fully assembled cars by train became wasteful. To make more efficient use of train space, Ford decided in 1909 to establish branch assembly plants and ship knocked-down automobiles. This innovation resulted in lower freight and handling costs and orders were filled more promptly. In addition, stock of automobiles and parts could be held at the branch assembly plants, allowing production to continue in Detroit in months when sales were typically slow (Nevins and Hill, 1954). The first branch assembly plant was built in Kansas City, Missouri, and by 1913, Ford had established branch assembly plants in thirty-one U.S. cities (Flink, 1988).

4.1.4 Management of Production

At the turn of the century, American industry was at a crossroads as traditional craft production was giving way to modern ideas of high volume production. Many products were being manufactured in factories, including watches, clocks, sewing machines, typewriters, and bicycles, but the transition was not yet complete. Production and management systems were not yet aligned with emerging ideas about capitalism and high volume production. Because traditional craft manufacturing practices gave craftsmen complete control of the production process, manufacturers engaged in factory production continually clashed with suppliers and workers in trying to increase the quantity and reduce the cost of their products. The shift to mass production presented new challenges for product and process management, especially with a complex product like the automobile. New techniques were required to manage quality control, inventory, a large labor force, and investments in expensive equipment (Abernathy, 1978).

This section reviews the second aspect of Ford's evolving production process, the management of production, specifically the breakdown of manufacturing and assembly operations into minute tasks. These developments took place within the context of a wider movement toward 'scientific management,' the formulation of which is attributed to Frederick W. Taylor. By reviewing both the development of so-called 'Taylorism' and the management of production at Ford, this section explores the question of how much of the innovation at Ford represented original inspiration versus knowledge spillovers transmitted either directly from the work of Taylor or indirectly through shared perspectives of the knowledge base.

As with the emergence of the automobile itself, multiple entrepreneurs were working separately yet concurrently on the same problems and were developing a variety of similar solutions. Craft production was misaligned with emerging rules within the TIEC regarding capitalism, factory production, mechanistic reductionism, and an increased cultural valuation of efficiency. Thus, this era of innovative ferment was attributable to pervasive changes in the TIEC. Scientific management and the Ford system improved efficiency by taking control of the production process out of the hands of skilled craftsmen, thereby establishing new rules for production that were aligned with the evolving higher-level rules within the TIEC. While the Ford team certainly must be credited with a great deal of inspiration and insight, critical members of that team were either educated in scientific management or were tied into the knowledge base where these ideas were circulating. Therefore, the role of knowledge spillovers in Ford's highly synthetic innovations was of primary significance.

4.1.4.1 Taylorism

Frederick W. Taylor, the son of a wealthy Philadelphia Quaker family, was an educated man who had aspired to law school but dropped out of college when problems arose with his eyesight. The formal education of mechanical engineers was still a new idea, and, despite his parents urging, Taylor opted to follow a more traditional path and became a patternmaker's apprentice. On completion of his apprenticeship in 1878, Taylor began working at Midvale Steel Company, where he quickly rose to gang boss and then foreman. In trying to extract more work from the men, Taylor saw his camaraderie with them turn to hostility. He was convinced that the men could be more productive, yet they claimed to be working as fast as possible. Taylor's observations led him to conclude that management was responsible for most worker inefficiencies. Believing it would defuse the animosity between workers and supervisors, Taylor resolved to study the elements of work and let science impartially determine what constituted a fair day's work (Kanigel, 1996).

Taylor and his contemporaries were heavily influenced by the scientific paradigm of the time based on the ideas of reductionism and mechanism. Methodological reductionism holds that the best way to understand a complex system is to develop an understanding of its component parts. The ideas of reductionism and mechanism were introduced by René Descartes in 1637, who argued that the physical world was like a machine, its pieces like clockwork mechanisms, and that the machine could be understood by taking its pieces apart, studying them, and then putting them back together to see the larger picture. Descartes, however, was a dualist and did not believe that the human mind could be explained in mechanistic terms. The mechanistic view gained favor with the revolutionary successes of Isaac Newton, whose work in mechanics seemed to successfully explain the motion of everything in the heavens and on earth.

Scientists and engineers in the late 19th century applied Newton's laws of motion and the principles of classical thermodynamics to improve the operating efficiency of steam-driven machinery. Taylor had begun studying engineering through a series of correspondence courses at the Stevens Institute of Technology and eventually obtained a Mechanical Engineering degree in 1883 (Kanigel, 1997). Familiar with current scientific and engineering thinking, Taylor believed that a mechanistic approach could be applied to the organization and execution of work in order to improve the efficiency of human labor. By breaking down a job into small, isolated parts, Taylor believed he would be better able to understand the individual tasks and thus be able to control them (Freedman, 1992). Taylor tackled the problem of managing the human worker as he would have approached the design and optimization of a machine.

Taylor began experimenting with 'time' studies around 1880, using a stopwatch to measure the component parts of a task to the hundredth of a minute to determine the "one best way" to complete the task. With contributions from Frank Gilbreth, this methodology later became known as time-and-motion study.[35] Through time studies, Taylor determined the amount of work a first-class laborer could complete in a day and

[35] While Gilbreth's and Taylor's work had similar methodologies, the two differed in philosophies and goals. Taylor was concerned with determining a fair day's work, primarily as a means to increase profits. Gilbreth was concerned with the efficiency of motion and elimination of unnecessary movement purportedly for the workers' welfare (Price, 1990).

used this to develop work standards and pay incentive schemes that would reduce 'soldiering' or loafing. With his approach, which later became known as Taylorism, he sought to deliver more goods at lower prices while paying higher wages.

Through extensive metal cutting experiments, Taylor and his assistants developed customized slide rules that could quickly determine the fastest cutting speed that would simultaneously satisfy a machine tool's ability to exert power and a cutting tool's ability to bear it. The customary rules of thumb traditionally used by skilled machinists could be replaced with scientifically determined, step-by-step, minute-by-minute instructions. Control of the machine shop was taken out of the hands of the workmen and placed completely in the hands of management (Kanigel, 1997). Workers need only be trained in the "one best way" and left to execute their instructions mechanically, no thinking required. The planning of work was left to engineers and managers, who were to apply Taylor's scientific management principles to determine the best method for each task. A Taylorist firm was thus hierarchical, with layers of efficiency engineers and managers supervising highly specialized workers who operated according to minutely detailed instructions.

Taylor believed that his system of work organization and pay incentives would equally divide work between workers and managers, force workers to deal individually with managers' demands, and bring an understanding of the managers' perspective to the worker. Unfortunately, instead of decreasing hostilities between labor and management, Taylorism was demonized by workers and bosses alike. Labor saw it as exploitative; workers had to work harder and faster, while reduced to operating mindlessly like a machine. Meanwhile, bosses resented the higher wages and interference with their management decisions (Kanigel, 1996). Nonetheless, Taylorism effectively removed control of the production process from skilled craftsmen and placed it in the hands of management.

Taylor left Midvale in 1890 to become a management consultant for varied industrial clients. In 1903, Taylor presented an essay titled *Shop Management,* the most complete statement of his management system to date, to a gathering of 350 mechanical engineers in Saratoga, New York. He served as president of the American Society of Mechanical Engineers (ASME) from 1906-1907, where he was asked to reorganize the management of the twenty-five year-old society.

Yet Taylor's ideas were little known outside industrial engineering circles until 1910, when lawyer Louis Brandeis used his ideas to argue before the Interstate Commerce Commission that a wage increase did not necessitate an increase in railroad rates. The Eastern Rate Case stimulated great interest in worker efficiency and 'scientific management,' as Brandeis called it. Taylor published a collection of his articles in 1911 in a book titled *The Principles of Scientific Management* and later became known as "The Father of Scientific Management." The book was re-published as a three part series in the *American Magazine*. Later that year, Taylor's disciples, including Frank Gilbreth and H.L. Gantt, published a number of books on efficiency. Kanigel (1997) reports an incomplete count of thirty-one articles on scientific management in 1910, fifty-eight in 1911, and 220 in 1912. According to Kanigel (1996), Taylor "took strands of thought and practice already present in the late 19th century and wound them into a thick, muscled cable – Taylorism."

4.1.4.2 Scientific Management in the Automotive Industry

The emerging field of scientific management defended specialization based on scientific observations, but specialization of automobile assembly operations began due to happenstance and necessity. The Olds Motor Works was just turning out the first mass produced automobile when its shops were destroyed by fire in 1901. This disaster led Olds to organize a separate final assembly operation, which "demonstrated, perhaps for the first time anywhere, that a major production process could be organized as a series of separate specialized plants" (Abernathy, 1978). Departmentalization had become common and elaborate in other industries, as illustrated by the organization of Standard Oil, and developed along functional lines in the automotive industry. By 1909, Cadillac, Buick, Olds, Studebaker, and other large factories had instituted minute subdivision of both business and production sides, with separate departments for purchasing, time-keeping, cost-accounting, engineering, design, and manufacturing. Cadillac had forty-four different manufacturing departments. This organization required a complex management system for inventory tracking, materials handling, and job routing (Nevins and Hill, 1954).

In 1909, the president of Packard invited Taylor, who by then was considered the nation's leading productivity consultant, to Detroit to speak to company executives. During a four-hour meeting, Taylor advised that managers should remove all brainwork and skilled labor from the shop floor by restructuring every job as dozens of simple repetitive tasks. Packard subsequently instituted a jobs analysis at their factory and had 'Taylorized' the plant by 1913. However, when Taylor returned to Detroit in 1914 to speak to over six hundred superintendents and foreman from various industries, he was informed that “several Detroit manufacturers had anticipated his ideas” on their own (Nevins and Hill, 1954, p. 468).

Nevins and Hill (1954, p. 468) assert that some understanding of Taylor’s theory “slowly percolated through industry” and that Ford employees “had doubtless caught some of his ideas.” Taylor had pioneered the idea of breaking down a task into minute steps around 1880 and according to Kanigel, by 1908 when Ford was planning the Highland Park plant, “this habit of thought was densely part of the American industrial air.” After Taylor published *The Principles of Scientific Management* in 1911, “you could hardly buy a stopwatch in Detroit, so swiftly were they snapped up” (Kanigel, 1997, p. 497). The Ford Motor Company established a time study department perhaps as early as 1912, and time and motion studies were used to establish wage rates and were critical in the design of the chassis assembly line in 1913-1914 (see section 4.1.5).

Rubenstein (2001) concludes that structural changes that led to deskilling the automotive labor force predated Taylorism. Hounshell (1984) is less conclusive, stating that “[t]he Ford Motor Company *could* have been ‘Taylorized’ without Taylor” (emphasis added). Indeed, the automobile industry had faced substantial management challenges early on. The booming demand for automobiles starting in 1904 made it difficult for manufacturers to meet demand and increasing production led to significant problems securing sufficient skilled labor in Detroit. Given his mechanical inclinations from an early age, it is likely that Ford would naturally approach the problems posed by mass production and labor shortages from a mechanistic and reductionist perspective, just as Taylor had. Further, advances in machine tools and high speed steel (which Taylor had helped revolutionize) meant that machines no longer held up production processes;

labor was now the limiting factor in productivity. Nevins and Hill (1954, p. 469) conclude that "[t]he machine process, in short, was generating and perfecting its own procedures. Plant engineers and production superintendents, knowing little of theory but schooled in machine shop, foundry, and assembly room to a firm grasp of practical needs, were creating a system of management to meet them."

Charles Sorensen, who began with Ford as a patternmaker and became an indispensable member of Ford's production team, stresses that Ford did not employ a formal management system and that no one at Ford was acquainted with Taylor's theories. He claims that the articulation of the principles and philosophy of Ford's production system followed its achievement, but his *ex post* description is highly suggestive of Taylorism. According to Sorensen, the Ford system put higher skill into planning, management, and tool building, thereby making jobs easier and making it possible for unskilled workers to earn higher wages (Sorensen, 1956). In 1926, Henry Ford provided a ghost-written article to the *Encyclopaedia Britannica* that listed among the principles of mass production the "analysis of operations into their constituent parts." He went on to say that successfully planning the orderly progression of materials through the shop requires "a careful breaking up of the work into the sequence of its 'operations'" (Encyclopedia Britannica, 1926, p. 822). When *American Machinist* editor Fred Colvin visited Highland Park in early 1913, he was convinced "that motion study has been carefully looked into, whether it is called by that name or not" (cited in Kanigel, 1997, p. 497). Ford described the 1913 establishment of new wage rates using strikingly Taylorist language:

> And then, too, the wages were not scientifically adjusted to the jobs. The man in job "A" might get one rate and the man in job "B" a higher rate, while as a matter of fact job "A" might require more skill or exertion than job "B." A great deal of

> inequity creeps into wage rates unless both the employer and the employee know that the rate paid has been arrived at by something better than a guess. Therefore, starting about 1913 we had time studies made of all the thousands of operations in the shops. By a time study it is possible theoretically to determine what a man's output should be. Then, making large allowances, it is further possible to get at a satisfactory standard output for a day, and, taking into consideration the skill, to arrive at a rate which will express with fair accuracy the amount of skill and exertion that goes into a job--and how much is to be expected from the man in the job in return for the wage. Without scientific study the employer does not know why he is paying a wage and the worker does not know why he is getting it. On the time figures all of the jobs in our factory were standardized and rates set (Ford and Crowther, 1922, p. 125-126).

The highly publicized Eastern Rate Case in 1910 and the publication of *The Principles of Scientific Management* in 1911 marked the height of influence of scientific management, and corresponded with the development of Ford's revolutionary production methods. It therefore is reasonable to theorize that the fundamentals and language of Taylorism had permeated the knowledge base, possibly in tacit form. However, by the time Ford was designing the Model T and the Highland Park plant, some of this knowledge also would have been codified in operating procedures at a number of Detroit's manufacturing plants. Indeed, the field of scientific management grew quickly after 1911 and widespread interest brought an "efficiency craze that reached into home, farm, and office, as well as factory" (Kanigel, 1997, p. 486). When investigators toured Highland Park in 1914, they commented that the Ford had arrived, by trial and error, at the practical results detailed in two new theoretical texts: *The Business Administrator*, by E. C. Jones, and *Installing Efficiency Methods,* by C. E. Knoppel (Nevins and Hill, 1954).

Further, there is ample evidence that Ford officials were networked in the knowledge base. Ford Motor Company's chief tool designer, Oscar Charles Bornholdt, is listed on the membership roster for the ASME in 1907, while Taylor served as president from 1906-1907. Hounshell (1984, p. 223) names Bornholdt, who worked at Ford from

April, 1906, until 1913, among the "backbone of the Ford production team" responsible for tooling up for production of the first Model T.

In 1912, Ford hired Clarence Avery, who had taught Henry Ford's son, Edsel, in high school. Avery was an educated man, having studied at the Ferris Institute and the University of Michigan. According to Nevins and Hill (1954, p. 474), "[h]e read widely, knew the latest European and American advances in engineering, and kept in touch with the ideas of men like Frederick W. Taylor." As discussed in section 4.1.5, Avery played a critical role in developing Ford's new production system and likely originated some of its revolutionary ideas. Sorensen (1956, p. 130) states that it was Avery who redesigned operations and "worked out the timing schedules necessary" to install continuously moving conveyer assembly systems and Avery certainly was responsible for instituting time-motion studies for chassis assembly.

Higher labor productivity at Ford came at a price. The rapid pace of production was set by machines and conveyor systems while the work became minutely prescribed and monotonous. Labor became increasingly dissatisfied, and by 1914, Ford was required to hire 53,000 employees per year to maintain a staff of 14,000 (Ford and Crowther, 1922). Although the workers now required little skill and could be quickly trained, the rising turnover rate would become untenable. In 1895, Frederick Taylor had presented a paper titled "A Piece Rate System" to the Detroit meeting of the American Society of Mechanical Engineers which warned that "MEN WILL NOT DO AN EXTRAORDINARY DAY'S WORK FOR AN ORDINARY DAY'S PAY" (Taylor, 1895, emphasis in original). In conjunction with his management system, Taylor recommended pay increases between 30 and 100% depending on the skill, intelligence, and strength required (Taylor, 1911).

John R. Lee, who came to work for the Ford Motor Company though the purchase of the John R. Keim Mill in 1911, took over personnel and began an effort to systematize the wages. The new structure initiated in October, 1913, established wages based on skill class and granted a 13% factory-wide raise apportioned according to individual worker efficiency. Lee designed other labor reforms to protect workers from potential discriminatory conduct of foremen and to enable steady advancement from an unskilled standing to that of first-class workman. Foremen were required to participate in continuous reviews of worker capacity and job assignment to further ensure proper advancement. In addition, safety measures, such as mechanical guards, railings, and improved lighting, were put in place to protect workers from "external" causes of accidents. Alarm systems were installed to protect them from accidents caused by "internal" or psychological causes like carelessness and boredom (Nevins and Hill, 1954).

In 1914, Ford nearly doubled wages to $5 per day – an increase in the upper range of Taylor's recommendation – and reduced the work day from 9 to 8 hours. Ford claimed that the wage hike was simply profit sharing: consumers benefited from low prices wrought through high productivity; the company was enjoying huge profits and paying dividends and bonuses to executives; it was time for the working man to share in the profits too. When Ford executives were working out the new pay scales, Sorensen (1956) claims he envisaged higher productivity and economies that would result from more-satisfied willing workers. In addition, higher wages made economic and business sense because a more prosperous national workforce would expand the market for automobiles. Whatever the motivation, labor management reforms and increased wages had a profound and immediate impact on the company's retention rate. The following year, Ford hired only 6,508 new employees, most of which were required due to growth rather than turnover (Ford and Crowther, 1922). Only one man was discharged in the six months prior to April, 1916 (Nevins and Hill, 1954). As a result of sweeping labor changes, Ford's employees were more experienced and happier and Henry Ford became a folk hero.

However, Fordism diverged from Taylorism on a significant point. Taylor largely took production hardware as a given, stripped away needless elements, and focused on maximizing the efficiency of work with that hardware through motion studies (Hounshell, 1984; Kanigel, 1997).[36] The studies were then used to determine a fair day's work and establish incentive pay schedules. Meanwhile, Ford engineers redesigned the fundamental production hardware to mechanize work processes to the fullest extent possible. While he may have applied time and motion studies to set up the processes, the machines then set the pace of work. Kanigel (1997) describes Taylorism as the universal case and Fordism the special case: the application of Taylor's system to mass production. Taylor may have provided the tools and the descriptive language, maybe even the mental paradigm, but Ford's team combined them with principles of armory practice and moving assembly and extended all three to develop an entirely new system of continuous production.

---

[36] Taylor did, however, design a number of specialized tools to substitute mechanical devices for skilled labor prior to 1900 (Kanigel, 1997).

4.1.4.3 Summary

The traditional craft system put control of production in the hands of skilled craftsmen and was incompatible high volume, low cost production in factories. Thus, craft production was misaligned with emerging rules within the TIEC regarding capitalism, factory production, mechanistic reductionism, and an increased cultural valuation of efficiency. To improve efficiency, managers needed to take control of the production process out of the hands of skilled craftsmen. This problem was not unique to the automobile industry and multiple entrepreneurs were working separately yet concurrently on the same problems and were developing a variety of similar solutions. An era of innovative ferment in the management of production had arisen in response to pervasive changes in the TIEC.

Scientific management and the Ford system improved efficiency by taking control of the production process out of the hands of skilled craftsmen, thereby establishing new rules for production that were aligned with the evolving higher-level rules within the TIEC. Ford's production team was uniquely able to synthesize and exploit the opportunities provided by this social and institutional context, in part because the management was relatively unfamiliar with the rules of production under the craft system. However, critical members of Ford's production team were either educated in scientific management or were tied into the knowledge base where these ideas were circulating. Therefore, while the Ford team certainly must be credited with a great deal of inspiration and insight, the role of knowledge spillovers in the company's highly synthetic innovations was of primary significance.

4.1.5 Moving Assembly Line, 1908-1915

The third and final aspect of Ford's process innovations, continuously moving assembly, is popularly regarded as his most significant contribution to modern production methods and confers to Henry Ford the status of 'hero entrepreneur.' However, the development of continuously moving assembly at the Ford Motor Company was a highly synthetic process involving the talents and personal inspiration of multiple entrepreneurs and spillovers from multiple industries. In addition, moving assembly was just one component of a production system in which synergies yielded higher cost reductions than the simple sum of the parts.

By 1906, Ford had established a highly innovative atmosphere. After obtaining financial control of the Ford Motor Company in 1907, Ford pursued an atypical fiscal strategy. Contemptuous of money-making and profit-seekers, Ford opted to reinvest profits in the company instead of paying out large dividends and corporate salaries. This internal funding gave the company extraordinary financial stability. In addition, though Flanders had introduced critical elements of armory practice, Ford and his engineers had no rigid ideas about production processes. Sorensen (1956, p. 55) recalls, "The Ford operations and creative work were directed by men who had no previous knowledge of the subject. They did not have a chance to get on really familiar terms with the impossible... There was no one around who had had experience, and if one came along who had done well in any other business we had more of a problem to get him to drop his ideas and fall in with our progressive manufacturing and assembling."

According to Hounshell (1984, p. 220), "Ford allowed an extensive amount of experimentation to be carried out in the factory and a surprising rate of scrapping processes and machine tools when they did not suit the immediate fancy of his production engineers." Ford's young talented engineers "carried out production experiments and worked out fresh ideas in gauging, fixture design, machine tool design and placement, factory layout, quality control, and materials handling." Machinery was organized for sequential operation, increasingly specialized tools and equipment were designed and installed, and the work was progressively broken down into smaller tasks. Each worker became more specialized, performing only a single task. In 1908, a single worker's average time before repeating the same operation was 8.56 hours. By 1913, the task cycle had been reduced to 2.3 minutes (Womack et al., 1990). In 1915, engine assembly at Highland Park was performed in a sequence of 84 tasks on three assembly lines (Rubenstein, 2001).

In trying to reduce the cycle time, moving the workers from stand to stand quickly became a problem. The solution, obvious in retrospect, was a major departure from standard industry practice: move the cars instead of the workers. According to Rubenstein (2001), Henry Ford and four other Ford Motor Company leaders all claim credit for inventing the moving assembly line. Yet Sorensen (1956, p. 129) claims that "Mr. Ford had nothing to do with originating, planning, and carrying out the assembly line. He encouraged the work, his vision to try unorthodox methods was an example to us." Although Henry Ford took a special interest in magneto assembly, it appears that "seminal ideas moved from the bottom to the top" at Ford and that Clarence Avery played the largest single role in developing Ford's new production system (Nevins and Hill, 1954, p. 474). When Avery was hired in 1912, he spent eight months studying every production department and then was assigned to assist Sorensen in production planning. A member of Ford's experimental staff recalled that Avery was "known as pushing the assembly line" and that "it was Avery who put that before the boss" (Nevins and Hill, 1954, p. 474). Others in the Ford experimental staff concur that Avery was the first to see the potential for putting the entire plant on the new system and that he served as the "guiding light" in working out assembly operations.

The idea of the mechanized moving assembly line is perhaps a natural progression from the practice pioneered by western bicycle manufacturers like the Western Wheel Works of Chicago, where machinists remained seated and runners brought materials to them. In addition, the Westinghouse Foundry used a conveyor system to move molds and sand as early as 1890. However, the idea apparently came to Ford officials from outside the metalworking industries. The flour-milling, brewery, meat-packing, and food canning industries were using gravity slides and conveyor systems to facilitate the smooth flow of materials, and all were well publicized.

Late in the 18th century, Oliver Evans had synthesized existing technological elements into an automatic flour mill and use of such sophisticated materials handling technology in the mills in Minnesota were a matter of national and local pride. Late in the 19th century, Edwin Norton developed automatic can-making machinery that combined sequential arrangement of machinery with automatic conveyors. Around the time that Ford implemented moving assembly lines, toolmaker Bornholdt compared the arrangement of Ford's machine tools to the layout of canning machinery. William Klann, head of Ford's engine department, had worked repairing grain elevators and other mechanical conveyors in Detroit breweries and his former employer also made conveyors for foundries. But it was the 'disassembly' lines of the slaughterhouses, described in detail in Upton Sinclair's 1906 book *The Jungle,* that apparently captured the Ford officials' imagination. After touring a Chicago slaughterhouse, Klann reportedly told the factory superintendent P.E. Martin that "[i]f they can kill pigs and cows that way, we can build cars that way and motors that way" (Hounshell, 1984, p. 241).

Experimentation with the moving assembly line was so rapid that the chronology of events found in photo documentation, reminiscences of company officials, and historical writings conflict (Hounshell, 1984). It is clear that Ford added the first moving assembly line in 1913 and improved the design through trial and error experimentation over the next year. Moving assembly lines were completed for magnetos, motors, and transmissions, and by the summer of 1913, the production from these lines threatened to swamp final assembly. It was time to revolutionize this line as well. Clarence Avery and William Klann were responsible for the intricate calculations necessary to develop the sequence and timing of steps and the distribution of parts. They began by timing the men as they assembled fifty chassis at fifty different spots in each of two assembly lines. On average, each chassis took 12.5 man-hours to complete. Next, they pulled a chassis along the floor to test their ideas. "Then after trial and error test to determine speed, optimum part placement, and other technical questions, an assembly line pulled by rope and windlass was set in operation October 1, 1913" (Nevins and Hill, 1954, p. 473).

Even in this rough experiment, the average assembly time fell to 5 hours and 50 minutes. By December, after careful motion-study, the line was extended to 300 feet and assembly time averaged only 2 hours and 38 minutes. Machine power soon replaced hand power on the windlass, an endless chain replaced the rope, and the assembly lines were raised to waist level. With further subdivision of the work, assembly time fell to 1 hour and 33 minutes. Engine assembly time was similarly cut from 600 minutes in the fall of 1913 to 226 minutes by the spring of 1914.

But the moving assembly line can not be credited as the sole source of Ford's unprecedented efficiency. Ford had invested a great deal in developing the Highland Park plant and subsequent moving assembly innovations. According to Nevins and Hill (1956, p. 504):

> A large force was constantly employed in what might be called the creative preparations for mass production; that is, in equipping the factory to carry it on. This spring of 1914 saw fifty-nine men hunched over desks with triangles, compasses, and slide-rules, making drawings for machine tools and fixtures; forty men busy with hammer, saw, and chisel making patterns; and nearly five hundred men at Highland Park and three hundred in outside machine shops forging, casting, and building tools. A regiment of men, in short, was constantly storming the bastions of old-style factory methods.

The combination of fully interchangeable parts, specially-designed single-purpose machinery, sequential placement of men and machines, overhead carriers for materials, slides and moving assembly lines, and the minute breakdown of assembly operations combined into a production system that was more efficient than the simple sum of the parts. Of this system, continuously moving assembly was the crowning jewel: the key to complete coordination and control. According to Hounshell (1984, p. 187), "Although the bicycle industry in the late nineteenth century brought to perfection the American system of manufactures it was unable to solve the problem of finishing and assembly of parts. The Ford assembly line… overcame this problem…"

4.1.6 Centralized Control and Vertical Integration, 1907-1914

Throughout the period covered in this chapter, Henry Ford clashed with his financial backers over strategy. This fact, combined with Ford's predisposition toward individualism and autocracy, led Ford to seek complete financial and managerial independence. Ford was assisted in achieving this goal by an unexpected exogenous event, the Panic of 1907. In addition, the price and availability of raw materials was always a concern and supply holdups and bottlenecks became an increasing problem as Ford moved toward continuously moving assembly. Therefore, Ford also began a process to bring these supplies under his control. This section relates this progression toward centralized control and vertical integration through 1915, though it would not be complete until 1918. Therefore, this trend and its implications for the company's innovative capabilities will be continued in the following chapter.

Like many early automobile manufacturers, the Ford Motor Company began as an assembler of purchased parts, first at a rented facility on Mack Avenue, then at its own plant on Piquette Avenue. In these early days, the financial backers, Malcomson and Gray, wielded significant managerial power. Malcomson owned 25.5% of the voting stock and Gray held 10.5% and served as company president. Ford, who owned 25.5% of the company, was designated vice president and had final say on automobile design, engineering, and production. Though Couzens held a small minority (2.5%) of company stock, he was completely responsible for all business matters, including bookkeeping, correspondence, accounts payable and receivable, supply contracts, advertising, and sales. A determined individualist with a fiery temper, Couzens was not afraid to assert his will over Ford or Malcomson when it came to meeting market demand and making the company profitable.

Despite both being hard-nosed individualists, Ford and Couzens worked well together, regarding themselves as the insiders and Gray, Malcomson, and the rest of the stockholders as the outsiders (Nevins and Hill, 1954). Ford had clashed with Malcomson over managerial and design issues from the beginning. The most significant conflict involved Ford's goal to produce a single, inexpensive automobile, while Malcomson preferred that the company build expensive cars. Late in 1905, Henry Ford established the Ford Manufacturing Company to begin manufacturing major parts for the new Model N, believing that retention and reinvestment of profits from these parts was critical to the success of producing an inexpensive car for the masses. Though Ford, the Dodges, Couzens, and three other Ford Motor Company stock holders were allotted shares in the new company, Malcomson was conspicuously excluded. While Ford's primary goal was to begin the in-house manufacture of the most important parts of Ford cars, it appears that he also created the new company as a means to free himself of Malcomson's influence.

Ford rented a factory on Bellevue Avenue for the new company and equipped it to produce engines, gears, and other parts solely for the Model N. The Dodge Brothers continued to manufacture the engines and parts for the more expensive Model B and Model K. In 1907, the new Ford Manufacturing Company enjoyed large profits from parts for the highly successful Model N, but Ford Motor Company profits fell due to flagging sales of the higher priced models. The saturating market for expensive vehicles and the Panic of 1907 no doubt contributed to the slow sales of the Model K, which represented Ford's first foray into the highest price bracket. Meanwhile, Malcomson had launched his own automobile manufacturing enterprise, the Aerocar Company. With his dividend income dwindling, Malcomson was coerced into selling his now unprofitable shares of the Ford Motor Company in order to keep his new venture afloat.

The bylaws initially adopted for the Ford Motor Company stipulated that no stockholder could sell shares to an outsider without approval from the other stockholders and without offering the other stockholders an opportunity to purchase them. Henry Ford purchased all of Malcomson's stock, and within the next year, three of Malcomson's friends sold their company stock to Ford and Couzens. The Ford Motor Company purchased the Ford Manufacturing Company and, by the fall of 1907, Henry Ford owned 58.5% of the reunified company and Couzens owned 11% (Nevins and Hill, 1954). Shortly after the death of John Gray in 1906, Ford had been named president of the Ford Motor Company and John Dodge had been named vice president. Henry Ford now wielded both managerial and financial control of the firm that bore his name.

With the success of the Model N and the ousting of Malcomson, Ford made his strategy for the future clear, saying that "he was going to produce a four-cylinder automobile, that once it was produced he was going to stick to that standardized design without changing it, that he was going to reach constantly toward a growing volume because it would drastically cut his costs, and that he was going to reduce prices steadily" (Nevins and Hill, 1954, p. 339). In the years that followed, production cost reductions were realized through a combination of high volume production of a single product, the Model T, and a number of process innovations – highly standardized parts, extensive use of specially-designed single-purpose machinery, and moving assembly. The introduction of this last innovation required a timely supply of parts, further reinforcing Ford's desire for control over supplies.

From the birth of the company, the Dodge Brothers had manufactured a major portion of the Ford automobiles, including engines, transmissions, axles, drive shaft assemblies, and drop forgings. While this arrangement had worked to the benefit of both companies, difficulties arose over annual price negotiations. With the inclusion of a huge foundry and a machine shop for manufacturing engines, transmissions, and axles, the Ford plant at Highland Park brought in-house the production capabilities previously provided by the Dodge Brothers. The move from Piquette Avenue to Highland Park began in 1910. With Ford as their only customer, the Dodge brothers must certainly have been concerned about the future of their business. In August of 1913, John Dodge resigned as director and vice president of the Ford Motor Company so that the Dodge brothers could pursue manufacturing their own automobile. Dodge soon ceased supplying parts to Ford (Nevins and Hill, 1954).

Just after the decision on the Selden patent dispute in 1911 (see section 4.5.1), Ford purchased the John R. Keim Mills of Buffalo, New York, which supplied stamped- and drawn-steel components for Ford cars.[37] Following a wildcat strike in 1912, Ford moved pressed steel manufacturing from the Keim plant to Highland Park and converted the Buffalo facility into a branch assembly plant. With the purchase of the Keim Mills, Ford was capable of manufacturing its own axle housings, crankcase and transmission covers, and even bodies. The intellectual capital that Ford inherited from the Keim Mills would also prove invaluable to the development of Ford's ground-breaking mass production techniques: William Knudsen was instrumental in setting up branch assembly plants; William H. Smith, who had first approached Henry Ford with the idea of used stamped steel, was a highly capable engineer; and John R. Lee later became Ford's personnel manager and engineered the company's sweeping labor and wage reforms.

By the close of 1914, Henry Ford was well on his way to bringing under one roof the entire business of automobile manufacturing. He had realized his dream of producing

---

[37] Ford had underwritten the new tools and dies necessary for this production, and the outlay represented a significant risk (Sorensen, 1956).

a car to replace the family horse and, without the influence of Malcomson, he was able to pursue his goal of steadily decreasing the price of that car so that every working man could afford it. By synthesizing a number of process innovations – the manufacturing of fully interchangeable parts, the development and use of special purpose machinery, the sequential placement of men and machines, the minute breakdown of production operations, and continuously moving assembly – he also was well on the way to developing the production system to make that happen. In 1914, sales of the Model T reached nearly a quarter of a million and Henry Ford had plans for expansion on a grand scale.

4.2 The Rise of General Motors

During the period from 1900-1915, the Ford Motor Company emerged as the industry leader with sales in 1915 exceeding half the market for automobiles. In the same time period, another industry giant was forming. But while Ford had arisen from a single firm that, by 1915, was producing a single model, this second giant took a very different form and path. The history of General Motors Corporation (GM) involves numerous automobile manufacturers and a multitude of suppliers. Like Ford, the company's history prior to 1915 is dominated by a single personality whose name, unlike Henry Ford's, is largely forgotten: William C. Durant.

This chapter begins in section 4.2.1 with the origins of GM in the Buick Motor Company and with W.C. Durant, whose experience in the carriage trade did not involve engineering or production, but rather management and primarily sales. Durant's success with Buick supports the notion that entrepreneurs (in this case the team of Durant and his partner Josiah Dort) must act as heterogeneous engineers who jointly address issues of technology, management, marketing, and finance. The Model 10 priced under $1,000 propelled Buick to the lead market position by 1908 and serves as a testament to the rising importance of the low-priced market.

Section 4.2.2 reviews Durant's establishment of GM and subsequent loss of control. While the product design during the transitional phase was converging on a dominant design, manufacturers did not know what that design would entail, which presented them with substantial risk. Ford attempted to influence the selection environment with his interpretation of the automobile and appropriate product attributes, while Durant attempted to mitigate his risk through diversification. However, by 1910, only Buick and Cadillac were profitable. This can be attributed to two factors: 1) the emergence of a new functional definition that appealed to a wider market and an aligned dominant design which supplanted other technological interpretations and configurations; and 2) manufacturing processes at GM subsidiaries that were not aligned with the TIEC and market needs for the combination of high reliability and low cost. This situation is illustrative of the difficulties facing much of the industry, as discussed later in sections 4.4 and 4.5.4. While most of GMs subsidiaries produced mediocre and outdated products using craft production techniques, Buick and Cadillac had high quality models positioned for the upper end of the mass market segment and lower end of the wealthy market segment. However, GM continued to struggle with aligning production methods to new rules in the TIEC.

These problems, combined with an aggressive expansion strategy and an economic slowdown, contributed to the transfer of control to a banking syndicate between 1910 and 1915, the era covered in section 4.2.3. In an effort to restore GM to profitability, the bankers increased the company's misalignment with both the emerging functional definition of the automobile and the rules of manufacturing for the mass market this definition served. By eliminating the Buick Model 10, they eliminated participation in the low-price market. And by maintaining high prices, low volume production, and hand-craft techniques, they adhered to outmoded economic rules. However, the bankers did improve production processes, bringing GM into alignment with the rules of armory practice, and also instituted a research program to replace rules of thumb for engineering. The company was returned to profitability, but a significant loss of market share and stock value placed it in jeopardy.

Finally, section 4.2.4 provides a detailed description of the development of the electric self starter 1911, one of the most significant product innovations for GM and the industry in this era. The electric starter overcame one of the largest disadvantages of the gasoline car and began a process of 'electrifying' the gasoline motor vehicles, making it a technological hybrid that combined advantages of electric and gasoline vehicles. This review reveals that spillovers and learning-by-interacting with suppliers were the most significant mechanisms in this critical innovation.

4.2.1 Buick and William C. Durant, 1902-1908

David Dunbar Buick, an inventor and manufacturer of plumbing supplies, founded the Buick Motor Company in 1902. Buick developed a well-engineered vehicle but exhausted the company's finances doing so. Buick's main asset was a patent on a valve-in-head engine that obtained a better fuel-air mixture and thus produced more power. The company was sold twice, landing in 1903 in the hands of James H. Whiting of Flint, Michigan. Located at a ford in the Flint River surrounded by abundant pine forests, Flint had been founded on the fur trade then later prospered in the lumber industry. About 90% of the original settlers hailed from New England and New York. By the time the forests began disappearing in the 1880s, the city was a manufacturing town of flour, paper, wool, and cotton mills. But its primary business was woodworking and the city's industrial base included an array of woodworking tools and skilled workmen. When the lumber mills closed down, Flint turned to carriage making and became the national center of the industry by the turn of the century.

James Whiting was president of the Flint Wagon Works and looking to enter the automobile industry when he purchased Buick and moved it from Detroit to Flint. Although David Buick and his engineer, Walter Marr, put together a successful prototype of what would be the Buick Model B, they had once again exhausted the company's capital. Yet they had produced no more than forty cars in 1904. Whiting had borrowed heavily from three Flint banks and the city's entire financial community was at risk. Whiting determined he needed a younger man to reorganize and run the company and turned to William C. Durant.

Durant was a highly energetic businessman and a persuasive salesman. At the age of 43, Durant was already a self-made millionaire and the semi-retired head of the country's largest carriage producer, the Durant-Dort Carriage Company. His achievements and his sincere and persuasive personality had earned the high opinion of his neighbors and associates in Flint. In building his carriage business, he had developed a strategy for corporate expansion that he readily transferred to the automotive industry.

Durant's sales philosophy is revealed in his advice to other salesmen: "Do not talk too much. Give the customer time to think. In other words, let the customer sell himself.

Look for a self-seller" (Weisberger, 1979, p. 35). Durant made record sales and new friends of prospective customers by giving them a vision of themselves as improved by the product. The trick was finding a deserving product. Durant had found his first 'self-seller' in 1886, when he purchased the patent rights and manufacturing operations for a simple yet elegant two-wheeled road cart with a novel suspension. His partner, Josiah Dallas Dort, agreed to supervise production while Durant handled sales and financing for the new Flint Road Cart Company. Although the business he had purchased could only turn out two cars per day, within two weeks Durant had garnered orders for the quick delivery of 600 carts. To meet this demand, he subcontracted a Flint competitor, W. A. Paterson, who had excess capacity. The enterprise manufactured around 4,000 carts in its first year (Weisberger, 1979). In order to meet this volume of production, Paterson had to reorganize his plant and operations, thereby introducing the rudiments of mass production techniques to the carriage industry (Pound, 1934). Unfortunately, after the first 1,500 carts were delivered, Durant discovered that Paterson was preparing to market a cart suspiciously similar to his own but at a lower cost. Durant and Dort promptly established their own production facilities and ended the subcontract arrangement.

The existing wagon and carriage industry in Flint followed the seasonal pattern of agriculture: inventories were built up during the spring and summer then sold to farmers in the fall. Durant, however, sought higher volume production, lower prices, and wider markets. To pursue his goals, he systemized sales, developed advertising, and spread production and sales operations over the entire year. Durant was soon swamped with orders and his Flint Road Cart Company became well known. His dealers suggested he begin manufacturing a four-wheeled family buggy for which they saw good prospects.

After his staff designed a new buggy, Durant again approached Paterson with a contract to supply 200 units. This time, Paterson obtained the names of his customers and lured some away with the promise of a cheaper model direct from the factory (Weisberger, 1979). Once again, Durant ended the contract arrangement, this time learning a valuable lesson on the virtues of vertical integration that would shape his business strategy for the next forty years.

Paterson's deceit was not the only impetus for vertical integration. Dependence on suppliers dictated production volumes and costs, which was a worrisome problem in an age of horizontal trusts. In 1900, there were rumors of a trust to control the manufacture of steel axles and wagon tires; in 1901, there were rumors of a combination to control the price of linseed oil used in paint and varnish (Weisberger, 1979). Durant recalled, "My twenty years' experience in the carriage business taught me a lesson. We started out as assemblers with no advantage over our competitors. We paid about the same prices for everything we purchased. We realized that we were making no progress and would not unless and until we manufactured practically every important part that we used" (Rubenstein, 2001, p. 72). Durant proceeded through acquisitions and the founding of new enterprises to bring under his control the manufacture of all significant components and many smaller parts:

> Our plan was to manufacture practically every important part of a buggy, and carrying out this idea, we did not stop until we had controlled or were interested in building a full line of bodies, wheels, axles, forgings, stampings, leather, paint, trimmings and various other items, even whip sockets; but not until our accessory plants were in operation… did we have a product that had no competition in value or price in the country. This gave us control of the business in that line as long as carriages were in demand. (Weisberger, 1979, p. 47)

In addition, Durant's acquisitions provided product diversification according to both price and styling. The two principles of Durant's strategy – vertical integration and product diversification – helped build the country's largest carriage company in terms of production volume. In 1895, the company was renamed the Durant-Dort Carriage Company to better reflect the full range of its products. In seven of ten years beginning in 1892, the company added a major division (Weisberger, 1979). By 1900, Durant-Dort consisted of fourteen branch plants and hundreds of sales agencies with annual sales of over 150,000 carriages (Flink, 1988).

While a more complacent man would have been content with his success, Durant was in search of a new challenge when Whiting approached him to run Buick in 1904. Durant spent two months testing the Buick automobile and decided he had found his next self-seller. Six weeks later after accepting the reins, he traveled to the Automobile Show in New York and obtained 1,108 orders for automobiles – 1,071 more than the company had produced in the previous year – but he also effectively raised significant amounts of capital. Durant had insisted on absolute control of Buick and, within ten months, he had increased the company's capitalization to $1.5 million. Durant recruited new engineering talent to work with Buick, Marr, and Arthur Mason, an expert the previous management had hired from Cadillac.[38] Mason developed a new engine that increased operating speed from 1,800 to 4,000 rpm, which led to the development in 1905 of the Model C, a powerful two-cylinder, 22-horsepower car (Weisberger, 1979).

Durant also created a national sales organization, using Durant-Dort carriage showrooms and seeking out dealers willing to travel into the countryside in New England and the West. One of his New England dealers, Harry Shiland, was a mechanic who personally overhauled each vehicle he sold. He was not impressed with the early Buicks

---

[38] Though David Buick was highly successful with technical innovations, he apparently was quite inept at finances and sank into debt. In 1906, he resigned from the company that bore his name and disappeared into poverty and obscurity.

and wrote a letter of complaint to the company. At Durant's invitation, Shiland toured the plant and provided such accurate criticisms that Durant decided to appoint him director of Buick's service department. Shiland stressed that cars needed to be foolproof in order to market them to doctors and businessmen and he proceeded to build a parts and complaint department on this principle.

When it came to parts delivery, automakers and parts manufacturers were at the mercy of railroad freight dispatchers. Within months after Durant took the helm, Buick experienced a production bottleneck due to delayed delivery of axles from the Weston-Mott Company in Utica, New York. Durant promptly contacted Charles Mott and invited him to relocate or establish a branch plant in Michigan. In the negotiations that ensued, Durant offered to subscribe to 20% of the stock in the new enterprise, donate the site for the new plant, and award the new company all of Buick's axle business (Weisberger, 1979). In 1907, Buick became an owner of the Weston-Mott Axle Company which moved its machinery and operations to Flint, Michigan.

Durant was once again following his successful strategy of vertical integration and product diversification. He directed Buick to develop several models of cars in different price ranges and was particularly interested in entering the lower priced market. In 1908, Buick introduced the Model 10, a lightweight, four-cylinder runabout with 18 horsepower and a simple planetary transmission. The Model 10, fondly known as the White Streak, was fast, sporty in design, easy to drive and maintain, and sold for under $1000. Buick also sold the luxury Model D for $2,500 and two additional models priced between the D and the 10. In only four years, Durant was able to turn Buick into the country's leading automobile manufacturer with production of 8,487 units in 1908 (Pound, 1934; Weisberger, 1979; Flink, 1988).

4.2.2 The Birth of General Motors, 1908-1910

Benjamin Briscoe, head of the Maxwell-Briscoe Motor Company, approached Durant in 1908 with a plan to unite the major firms in the automobile industry into a single company. Like many other businessmen of the times, Briscoe believed that a large industrial 'combination' on the order of U.S. Steel Corporation would protect the automotive industry from the 'destabilizing' effects of unregulated competition. In 1907, manufacturers were competing for an ill-defined and relatively small market for automobiles, with the top 25% of firms accounting for around 75% of sales. Producers financed operations with credit from suppliers and cash advances from dealers and customers, leaving little cushion against a slump in demand. A company faced with ruin might resort to desperate competitive tactics while other unscrupulous new entrants promised quality at prices they could not deliver. The resulting collapses shook buyer and investor confidence, sending ripples through the entire industry (Weisberger, 1979).

Briscoe and Durant, with the backing of J. P. Morgan & Company, began negotiations with Henry Ford, Ransom Olds, who was now making the Reo, and other leading automobile producers. Durant envisioned a holding company that allowed the manufacturers to continue to operate under existing management. Ford and Olds, however, insisted on a cash buyout rather than an exchange of stock, and pulled out of the negotiations. Although Morgan initially supported a merger of Briscoe and Durant's enterprises, the bank soon withdrew. According to Flink (1988), the bankers were concerned by Durant's cavalier attitude, particularly in dealing with Buick's stockholders and also in an offer he had made to buy the Olds Motor Works without reviewing the company's books. They became convinced that Durant was unbalanced when he prophesied that U.S. automobile sales would soon reach half a million units.[39] Durant also insisted he be left in control of the new combination's finances, but the bankers insisted on choosing the officers and directors of the new company. When details of the proposed merger were leaked to the press, the plan died (Weisberger, 1979).

Durant, however, moved forward on his own, founding the General Motors Company (GM) in September, 1908, as a holding company that would purchase companies by exchanging stock in the old firm for stock in the new one. GM first acquired Buick then within months acquired the Olds Motor Works, which was failing. Durant initially left the management in charge at Olds, but replaced them in 1909. He then reportedly brought a Buick White Streak to the Oldsmobile plant, sawed its wooden body into quarters, separated the four parts into a wider and longer profile, and directed the engineers to duplicate the result to create the basic 1910 Oldsmobile (Weisberger, 1979). In 1908, GM also purchased the underfinanced Oakland Motor Car Company which would later be renamed Pontiac. In July, 1909, Durant purchased the Cadillac Automobile Company from Henry Leland, who had taken control of the firm in 1902 when it was the failing Henry Ford Company. Under Leland's guidance, Cadillac had earned international recognition for high quality engineering and the production of precision-machined interchangeable parts. In 1909, the Cadillac was widely viewed as the best high-priced car in the country (Weisberger, 1979). As agreed during negotiations, Durant left Leland in complete control of managing Cadillac.

[39] Durant was, of course, proven correct six years later when sales reached 548,000 in 1914.

Durant also began purchasing parts manufacturers, believing that "[c]ontrolling this enormous volume would make it possible for these accessory plants…to materially reduce costs because of the volume of business from GM which they could depend upon if motor cars and motor trucks, as I was firmly convinced, was [sic] to become important factors in the industrial life of America" (Rubenstein, 2001, p. 73). Shortly after the creation of GM, Durant was approached by Albert Champion, who was manufacturing a spark plug of his own design. Durant was impressed and, as before with the Weston-Mott Axle Company, promised Champion a market for every plug he could make if he moved his operations to Flint. Durant sold his interest in the Champion Ignition Company to GM in 1909, and the firm was later renamed the A.C. Spark Plug Company. Within two years of incorporation, GM acquired at least part interest in manufacturers of electrical components, transmissions, wheels, brakes, engine castings, springs, body panels, and other parts.

Unsure of the future technological direction of the industry, Durant pursued "getting every car in sight, playing it safe all along the line" (Flink, 1988, p. 65). By 1910, GM had acquired thirteen motor vehicle manufacturers and ten parts manufacturers, employed 14,000 workers, and produced one-fifth of the automobiles made in the U.S. (Nevins and Hill, 1954; Weisberger, 1979). Unfortunately, Durant's policy of broad acquisition and loose managerial control quickly turned disastrous. Several of his purchases turned out to be ill-advised, notably the Cartercar Company whose friction drive turned out to be poorly designed, the Elmore Manufacturing Company for its outdated two-cycle engine, and the Heany Lamp Company whose incandescent lamp patent was later revealed to be fraudulent. Meanwhile, of the automobile manufacturers, only the Cadillac and Buick companies turned profits and Buick's fell as Durant's energies were increasingly stretched. GM was a combination of twenty-five independent firms which had no unified system of accounting, purchasing, marketing or engineering research. The company had no cash reserves and little working cash, so when sales dropped due to an economic slowdown in 1910, GM was unable to make the payroll or pay suppliers. GM's accounting system was so inadequate that Durant was unable to demonstrate the company's exact needs to potential lenders.

The fact that only Buick and Cadillac were profitable can be attributed to two factors: 1) the emergence of a new functional definition that appealed to a wider market and an aligned dominant design which supplanted other technological interpretations and configurations; and 2) manufacturing and management systems at GM subsidiaries that were not aligned with the TIEC and needs of this wider market for high volume production, high reliability, and low cost. Most of the companies acquired by GM (Elmore, Oakland, Rainier, Welsh, and Cartercar) produced fairly low volumes of mediocre-quality automobiles using craft production methods and incorporating soon-to-be outmoded technologies like the two-cycle engine and the friction drive. However, in the Buick Model 10 priced from $900, GM had a product positioned for the upper end of the low-priced mass market. Buick sold almost 11,000 Model 10s in 1910 while Ford sold 18,700 Model Ts with prices starting at $680. In the Cadillac Model 30 priced from $1,400 to $3,000, GM had a high quality car positioned for the lower end of the wealthy market segment. The Cadillac also incorporated Delco's new ignition system discussed in section 4.2.4 and offered a closed body style for $2,200.[40] Cadillac sold about 8,000 Model 30s in 1910.

4.2.3 Banker Control of GM, 1910-1915

A banking syndicate came to GM's rescue with a $12.75 million loan that came at a very high price – $15 million in 5-year, 6% notes, and $6 million in stock. GM's stock was placed in a five-year voting trust and, though Durant was appointed a trustee and still held a seat on the board, he was outnumbered by bank representatives. Durant was forced to step down from active management of GM. Banker James Jackson Storrow was appointed as interim president of GM.

Storrow was a Progressive who distrusted one man rule and endorsed administrative decentralization. He appointed managers to track profits and costs for each of the GM divisions and gave them a great deal of independence. In return, he expected teamwork: the managers were to fully share information and resources and not compete with one another.

Charles W. Nash, who had worked his way through the ranks at Durant-Dort, was appointed president and general manager of Buick in 1910. Nash slashed inventories, squeezed more productivity out of the facilities, and compressed costs. With Nash consumed with administrative details, Storrow hired Walter Chrysler, then a superintendent at the American Locomotive Company's Pittsburgh plant, to supervise Buick's plant operations. When Chrysler arrived at Buick in January, 1912, he found the company's production methods were still based on labor-intensive handcraft methods inherited from the carriage industry. He also discovered that Buick had no system to

[40] Closed-body styles became increasingly significant in the specific and second transitional phases as discussed in detail in chapter 5.

determine the costs of production (Hyde, 2003). Relying on his experience building railroad cars, Chrysler instituted changes in management, production processes, and material handling that greatly improved scheduling and productivity, cut costs, and decreased manufacturing time. Concurrent with Ford's experimentation with moving assembly lines, Chrysler introduced a system to push unfinished chassis along tracks through the assembly area (Hyde, 2003). In November, 1912, Nash was promoted to the presidency of GM and Chrysler remained in charge of Buick operations.

In order to reduce the number of defects in GM's automobiles, Storrow decided to replace the rule-of-thumb methods used for design with organized research programs. Arthur D. Little, Inc., developed a plan for a centralized testing and research laboratory. With Storrow's support, the board voted in early 1911 to fund an engineering laboratory for mechanical and electrical testing which was later named the General Motors Research Department. The Department was staffed by nine researchers and their assistants. The first problems they tackled included painting, lubricating, and cutting oil practices at GM plants; materials testing for purchasing departments; and the investigation of new parts and accessories proposed by outside suppliers (Pound, 1934).

Between 1910 and 1915, the banker controlled management of GM instituted tighter fiscal controls and liquidated most manufacturing units, leaving only Buick, Cadillac, GM Truck, Oakland, and Oldsmobile. At the same time, they increased the GM interests in Weston-Mott, A.C. Spark Plug, GM of Canada, and Brown-Lipe-Chapin, which supplied gear-trains. They hired new administrative staff and improved communication and coordination among the units within the combination. GM was soon restored to solvency, but receipts were devoted to debt retirement rather than dividends or growth. In addition, the bankers failed to pursue the low-priced market or invest in production innovations, the two strategies that propelled Ford to the forefront of the industry during this time period. GM manufacturing units essentially adhered to two tenets of traditional manufacturing strategy: 1) that high-priced products yield high per unit profits and thus higher overall profits than large volume production of inexpensive products; and 2) that craft manufacturing techniques yield superior quality over machine-intensive mass production techniques. Thus, in an effort to increase per unit profits, Nash ended production of the Buick Model 10 that was competitive with the Model T and essentially cut Buick sales in half in 1911. Meanwhile, Leland refused to adopt assembly-line techniques at Cadillac, believing they would compromise mechanical precision. Despite improvements made by Chrysler, GM's strategies and production techniques were incompatible with the new functional definition of the automobile that appealed to a wider market and with emerging economic rules. By 1914, the company's market share had fallen from 21% to 8.5% and its stock price was depressed (Flink, 1988).

4.2.4 Development of the Electric Self-Starter, 1911-1912

Prior to the invention of the electric self-starter, motorists had to adjust the spark and turn the engine over by hand, spinning a crank which turned the flywheel. This feat required, according to Pound (1934, p. 271) "the strength of Ajax, the cunning of Ulysses, and the speed of Hermes." Starting a car in the cold was particularly problematic, and the starter problem was one reason that electric vehicles were preferred by some motorists, especially women, since "the whole back-breaking operation was quite beyond the powers of all women save those of Amazonian proportions." From almost the beginning of the industry, gasoline-powered auto-makers had experimented unsuccessfully with air, spring, and electric starting devices (Epstein, 1972). But a random accident was perhaps the catalyst that finally stimulated a solution. In the winter of 1910, the founder of Cartercar stopped to help a female motorist whose car had stalled near Detroit's Belle Island Park. When Byron Carter spun the crank, the engine backfired, throwing the crank in reverse, breaking Carter's arm and jaw. Two Cadillac engineers driving by took Carter to the hospital where he developed pneumonia and died a few weeks later. As legend has it, Henry Leland was horrified and vowed "I won't have Cadillacs hurting people that way" (Cray, 1980, p. 116).[41] He called his engineers to a special meeting and told them that eliminating the hand crank was a top priority. Unfortunately, the working model they developed was far too large to put into an automobile. When the team was unsuccessful at solving the starter problem, Leland turned to Charles F. Kettering.[42]

While at National Cash Register (NCR) in Dayton, Ohio, Kettering had invented a small electric motor to replace the hand crank on cash registers. All earlier attempts had failed, leading experts to the conclusion that it was physically impossible because the electric motor would need to be at least as big as the cash register itself. Kettering, however, perceived that the motor needed to provide only a short burst of power, leading him to develop a small motor with a mechanical clutch.

---

[41] Other versions of the story claim that Leland was a close fried of Carter's and that Leland said either "the Cadillac car will kill no more men if we can help it," or "I won't have Cadillacs hurting any more people that way," both of which imply that the stalled car was a Cadillac (Barach, 2007; Cray, 1980).

[42] Although Cadillac had already purchased electric ignition systems from Kettering's Dayton Electronics Company, Cray (1980) reports that it was Cadillac's assistant sales manager, Earl Howard, who suggested that Leland should contact Kettering. Howard was employed as secretary to the sales manager at National Cash Register when Kettering had developed a compact electric motor to replace the hand crank on cash registers.

Edward Deeds, general manger at NCR, subsequently asked Kettering to assist in building a car from a kit he had purchased. Kettering developed a high-energy spark ignition to replace the weak one supplied with the kit. At that time, the emerging standard ignition system for gasoline-powered cars, including the Cadillac, employed a magneto for spark ignition rather than a dry cell or a battery. Dry cells wore down quickly and had to be replaced. Thanks to development for electric vehicles, batteries were becoming more reliable, but as late as 1911, batteries in gasoline-powered vehicles still wore down in about 500 miles. They could be removed from the car and recharged, but they rapidly deteriorated due to vibration and shock from the rough roads. Kettering simplified the ignition system by placing the induction coils, one for each spark plug, in series in a heat-resistant, armored steel box which reduced the failure rate caused by heat and vibration. He then replaced the vibrators that made and broke the circuit with a single master set of contact points connected to a condenser that drew away excess current, prolonging the life of the points. His system was less susceptible to vibration and arcing, produced a hotter spark, and used less current, which extended battery and component life (Barach, 2007).

Kettering perfected his ignition system and installed it in his Cadillac Roadster, then wrote to Leland about its flawless performance. In September, 1908, Leland sent his chief engineer, Ernest Sweet, to Dayton to test drive Kettering's Cadillac.[43] Sweet was impressed, and Leland ordered electrical ignition sets for the entire production run of model year 1910 Cadillacs, around 8,000 units. But Leland also included a standard magneto ignition system as a backup, just in case. With the contract in hand, Kettering

[43] Ironically, it reportedly was Sweet and another Cadillac engineer who two years later came upon Byron Carter after his fateful accident near Belle Island Park and drove him to the hospital (Barach, 2007).

and Deeds established the Dayton Engineering Laboratory Company, later shortened to Delco, in 1909 to produce the ignition devices and to develop electrical components for the automotive and other industries.

When Leland approached Kettering with the self-starter problem, it took him and moonlighting engineers from NCR little time to develop a solution based on his small motor for the cash register. In fact, according to Pound (1934), Kettering had already been trying to sell his electric self-starter to auto manufacturers who were skeptical of its reliance on a battery they believed would be quickly exhausted. Kettering demonstrated a prototype to Leland in December of 1910 and by August, 1911, had been awarded a patent. Kettering's solution was an integrated starter and ignition system that employed a battery, starting motor, and generator equipped with a clutch and reduction gear. Leland had the electrical self-starter installed in the 1912 Cadillac, but still included a magneto and hand crank as a backup. He also installed a variable speed regulator developed by Kettering that allowed recharging of the battery without removing it from the car and also prevented overcharging, which had plagued earlier attempts at recharging.

Within two years, nearly all manufacturers offered automatic starters on their automobiles (Epstein, 1972). The electric self-starter incorporated one of the most attractive features of the electric vehicle – ease of starting – into the internal combustion engine vehicle and opened the gasoline-powered car market to women. Industry observers of the day referred to this innovation as the 'electrified gas car,' and in truth, it represented a weak hybrid gasoline-electric vehicle (Kirsch, 2000). By 1920, electrical components accounted for 5% of the total cost of a gasoline vehicle. Today, that fraction is at least one third.

In 1915, Delco was purchased by Durant's United Motor Company, which was in turn acquired by GM in 1918. In 1919, GM purchased the remaining three companies that Kettering had founded, the Delco Light Company, Dayton Metal Products, and the Dayton-Wright Airplane Company. With Kettering and his research operations completely enfolded in the corporation, GM commissioned Kettering to set up and direct the General Motors Research Laboratories in Dayton. In 1920, the laboratories were incorporated as the General Motors Research Corporation, and in 1925, the research activities were moved to Detroit.

4.3 The Birth of Chevrolet, 1911-1915

After being forced from active management of GM in 1910, Durant did not sit idle. In 1911, he established three new firms: the Chevrolet Motor Company in Detroit, and the Little Motor Company and Mason Motor Company in Flint, Michigan. He hired Louis Chevrolet, a star on his Buick racing team, to design a new vehicle for Chevrolet and William Little to head the Little Motor Company. He hired Arthur Mason, who had been the Buick engine superintendent, to run the Mason Motor Company which was to produce engines for both the Chevrolet and Little cars. Durant was incensed by Nash's decision to discontinue the Buick Model 10 in 1911 and therefore aimed for the inexpensive market. Priced at $600, the underpowered Little was moderately successful, but the Chevrolet, a ponderous 6-cylinder model with a $2,150 price tag, was a market failure. Louis Chevrolet left the company and Durant consolidated the Chevrolet and Little companies as the Chevrolet Company in 1913, discontinuing the Little nameplate. The new company retained partial ownership of Mason Motors under the assumption that Chevrolet would require only part of Mason's output.

It took Chevrolet's engineers another year to develop a satisfactory new automobile, but in 1914 they introduced two new 4-cylinder models that proved to be spectacularly successful: the $875 Baby Grand touring car and the $750 Royal Mail roadster. Weisberger (1979) provides a vivid picture of Durant's 'family proposition' and of innovation in the early years at Chevrolet. During engineering conferences, Mason, Little, and engineer Alfred Sturt would discuss shop problems, sometimes arguing heatedly. Durant would then step in to calm them down and they would set to work solving the problem. On one occasion, Buick veteran Charlie Wetherald, an incorporator of Mason Motors, heard how the tight-fitting pistons of a new engine produced more power but seized up when hot. Wetherald proceeded to hand file the pistons to a slightly elliptical shape that solved the problem. Another engineer then developed a special cam for the grinder to mechanize and standardized the new process. "We didn't even take the trouble to find out if we could apply for a patent" (Weisberger, 1979, p. 173). Workers simply developed a solution through trial and error and moved on to the next problem.

In 1914, Chevrolet could sell as many cars as it could produce and Durant expanded production and sales operations, establishing new assembly and regional sales facilities nationwide and in Canada. Durant then announced that the 1915 Chevrolet line would include a new model called the 490 – named for the price of Ford's Model T with which it was intended to compete. Arthur Mason was called in to assist in the new design and Durant was impressed with his ideas on components other than motors, including axles, transmissions, and suspension. Aware that Mason was personally in debt and that

Mason Motors was behind schedule and short of cash, Durant decided that Chevrolet should purchase the remaining shares of the company.

Although the price of the 490 came in above the Model T at $550, Chevrolet was flooded with orders. At over 70,000 units, Chevrolet sales in 1916 exceeded even Durant's optimism (Weisberger, 1979). This phenomenal success, contrasted with the failure of the original, powerful and expensive vehicle designed by Louis Chevrolet underscores the growing importance of the low-priced market after 1900. But while Durant had once again found business success, his sights were set higher. He yearned to retake control of GM, which he referred to as 'my baby.' His success with Chevrolet, combined with the conservative behavior of GM's banker regime and Durant's clever stock maneuvering, would provide the means for a dramatic return to GM, which is covered in chapter 5.

4.4 Industry Shakeout

As an industry approaches the specific phase, a dominant design emerges and firms begin competing based on cost and quality rather than product differentiation. At this point, it is typical for an industry to experience shakeout, where the number of firms peaks and begins to decline. The shakeout for the automobile industry began in 1909 and was extreme. I identify eight factors that contributed to this phenomenon, several of which are inter-related: 1) increasing economies of scale that necessitated high volume production and high capital investments; 2) the cumulative burden of manufacturing parts for multiple models; 3) process innovations at Ford that were not easy for firms to adopt; 4) the cumulative effects of and increasing returns to innovation; 5) the redefinition of the automobile for the mass market; 6) an emerging dominant design; 7) the saturation of the market niche for very expensive vehicles; and 8) the Panic of 1907 and ensuing recession. Items 5-7 together represent an important feedback mechanism between selection and variation. By contributing to the failure of specific firms, the selection environment influenced variation and further reinforced the emerging dominant design.

Between 1902 and 1910, the number of automobile manufacturers entering the market generally rose, with a peak entry of 81 firms in 1907 (Simons, 1995; Klepper and Simons, 1997). Market entry declined between 1911 and 1921 then became negligible. The exit rate followed a similar pattern, peaking around 1909. As shown in Figure 4-1, the number of firms in the industry peaked at 274 in 1909 then began a precipitous decline, leaving only 30 firms by 1929.[44] With the reduction in firms, there was also a reduction in product variety. In 1911, 270 producers manufactured 400 different nameplates; by 1942, there were only eight producers manufacturing 17 nameplates (Rubenstein, 2001).

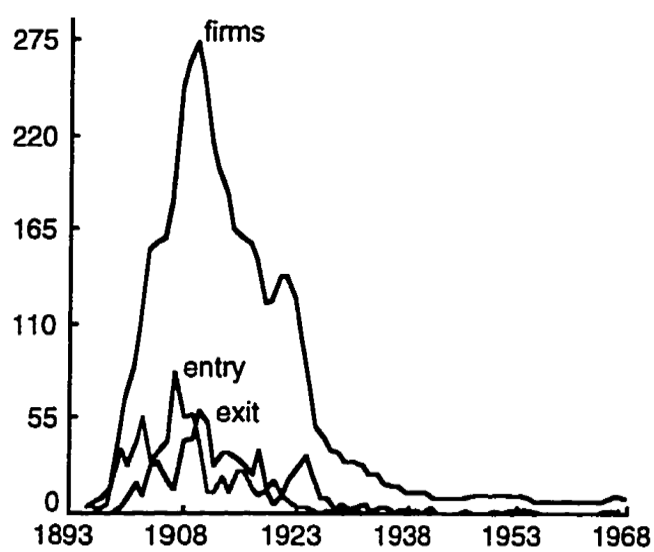


Figure 4-1: Number of Automobile Manufacturing Firms, Entry, and Exit
Reproduced from Simons (1995, p. 100)

[44] Data is from Simons (1995) based on Smith (1968) and excludes small firms that never achieved commercial production.

Around the turn of the century, new automobile manufacturing ventures were launched with very little capital. For example, Ford was incorporated in 1903 with a $150,000 capitalization of which $50,000 was held in company stock. Machinery was assessed at $10,000 and patents at $40,000. The company initially ran at the verge of bankruptcy, owing well more in contracts to suppliers than it could cover until cash advances came in from dealers. Many manufacturers rented modest facilities where they largely assembled parts and supplies bought on contract. By 1911, when Ford was implementing revolutionary mass production techniques and Durant was establishing Chevrolet, the situation was far different. Large factories were now required that could turn out huge quantities of vehicles using a vast array of expensive, specialized equipment. A manufacturer needed established credit, highly trained executives for both production and business offices, and connections in the steel, rubber, nickel and upholstery industries. In 1910, Flanders commented on the difficulty of entry in the *Detroit Saturday Night* (cited in Nevins and Hill, 1954, p. 476):

> The man or concern that would start in the automobile business today must begin, not where others began, but where they are *now*… To equal in quality cars now selling at $700 to $900, it is not only necessary to build them in tremendous quantities, but to build and equip factories for the economical manufacture of every part. You must begin with an experimental laboratory for analyzing and testing metals. The outlay for this department alone, with its necessary adjunct, the machine shop for tool and jig making, would amount to more than the earlier makers ever had invested in factories. Yet the laboratory is an indispensable adjunct to the modern automobile factory. It is the chief means of effecting an economy that cuts dollars from the final prices of the car.

Flanders estimated that, in 1910, it would take an initial investment of at least two million dollars to enter the market with a medium-priced car.

Innovation is among the significant factors influencing firm shakeout, with cumulative and self-reinforcing effects strengthening the position of early entrants

(Simons, 1995; Klepper and Simons, 1997). Process innovations, made almost exclusively by Ford, conferred significant and growing advantages to the market leaders. When innovations are difficult for other firms to adopt, profitable entry becomes increasingly difficult. In 1914, the National Automobile Chamber of Commerce (NACC) instituted a cross-licensing agreement among its members. Though Ford was not a member, the company also shared patents with competitors without assessing royalties. In addition, Ford's mass production innovations were well publicized in trade publications such as the *American Machinist* and a book by Arnold and Faurote published in 1919. Henry Ford also extended an open invitation for all to tour his Highland Park plant. As such, these innovations were accessible to other manufacturers, but not necessarily easy to adopt, as discussed further below.

However, it is extremely difficult to differentiate between the advantages conferred by innovation and those conferred by simple economies of scale. In 1909, *Scientific American* (cited in Flink, 1988, p. 71) wrote that "standardization and interchangeability of parts will have the effect of giving us a higher grade of motorcar at a lower price, but this is dependent in considerable degree upon the production of one model in great numbers and the elimination of extensive annual changes in design that necessitate the making of costly jigs, gauges, and special machinery." The president of the Society of Automobile Engineers (SAE) in 1910 expressed a belief that the lack of inter-company standardization was "responsible for nine tenths of the production troubles and most of the needless expense entailed in the manufacturing of motorcars" (cited in Flink, 1988, p. 71). The mechanical branch of the SAE attempted to level the playing field for the small producer by standardizing parts across manufacturers. Whether wrought through innovation or economies of scale, the falling production costs realized by the largest producers put profit margin pressures on smaller firms, leaving no cushion for a crisis in production or a drop in demand.

In such a climate, corporate strategies were key factors in survival. In 1905, Henry Ford staked the future of the company on the Model N. With its success in 1907, he made his strategy for the future clear, saying that "he was going to produce a four-cylinder automobile, that once it was produced he was going to stick to that standardized design without changing it, that he was going to reach constantly toward a growing volume because it would drastically cut his costs, and that he was going to reduce prices steadily" (Nevins and Hill, 1954, p. 339). Ford's strategy was consistent with *Scientific American's* assessment of the industry published two years later. In the years that followed, reductions in production costs at Ford were realized through a combination of high volume production of a single model and a number of synergetic process innovations – highly standardized parts, extensive use of specially-designed single-purpose machinery, minute subdivision of labor, and automated materials and work flow.[45]

Ford's complete production system certainly yielded more than simple economies of scale and was not necessarily easy for others to adopt, despite Ford's open door policy (Raff, 1991). While many auto manufacturers adopted some of Ford's production methods such as conveyor systems, most still required skilled fitters as late as World War I. Ford's system allowed for a high degree of centralized control, which Henry Ford demanded. Raff (1991) argues that Ford's conveyors and moving assembly lines were a

[45] Ford himself described mass production as entailing the application of seven principles: power, accuracy, economy, continuity, system, speed, and repetition (Nevins and Hill, 1957).

means of coordination and control and that other firms had difficulty imitating these methods because they did not fully understand them. He cites as an example that Studebaker conceived of conveyors only as labor-saving devices. While much was made in the press, both then and through history, of the moving assembly line, the key to Fordism was the *combination* of progressive assembly and the production of *fully* interchangeable (standardized) parts – no fitting, and thus no skilled workers, required.

In addition, Ford's complete focus on a single model was also a critical part of his success, though it would later become a detriment. According to Nevins and Hill (1954, p. 332): "One factor thrusting numerous early companies into bankruptcy was their production of too many chassis. Just before the First World War one of the largest automobile makers said than no company was strong enough to build two chassis every year without going broke." As early as 1906, auto manufacturers had concluded that the annual model change was a curse. By producing multiple chassis and introducing new models every year, manufacturers were saddled with high design and tooling costs and were forced to produce a large variety of parts in small volumes at high cost. The manufacturer faced selling these parts below cost or ruining its reputation. The longer the company was in production, the larger the parts burden became. Ford solved this problem by producing large volumes of a single chassis and making only minor styling changes from year to year. Thus, Ford's production system and strategy placed him in unique position early in the 20th century.

GM later solved the multiple chassis problem differently, building on the SAE's approach, with the introduction of the Pontiac in 1927 which initiated a program to standardize parts across its divisions (see chapter 5). While this strategy would help GM eventually emerge as the dominant firm in the industry, the corporation flirted with financial failure throughout its first decade. This difficulty was due in part to the incomplete transition to the new manufacturing techniques and management systems that Ford exploited so fully. Since it was not clear what technological configuration and price range would prove most successful, Durant pursued a strategy of diversification. He assembled a large combination of automobile and parts manufacturing companies and purchased a wide variety of technologies. This strategy improved GM's chances of owning a product innovation that proved to be a crucial refinement. Unfortunately, it also meant GM was likely to purchase some market failures. In fact, Durant's rash behavior resulted in some spectacularly disastrous, as well as some brilliant, acquisitions.

In addition, Durant failed to integrate the management of the companies he purchased or rationalize the product lines. As a result, GM companies competed among themselves as much as with other firms. And because each division produced a number of different chassis, GM was not able to fully realize the economies of scale afforded by the high volume production of the corporation as a whole. Meanwhile, the parts manufacturing divisions continued to operate as independent companies concerned primarily with their own profitability rather than corporate level issues, as witnessed by the clash of priorities of the Fisher brothers and the GM board discussed in chapter 5. As of 1918, a few profitable divisions were carrying the entire corporation. These struggles were representative of the industry as a whole during this era.

By 1916, the situation for market concentration was further reinforced. According to Rae (1959, p. 133),

> [N]o independent company founded after 1916 managed to survive for more than a few years. This phenomenon occurred not in spite of the increasing use of motor vehicles, but rather because of it. The day of the small-scale producer in

> the automobile industry was over. The successful manufacturer had to be able to produce in quantity and sell in a nation-wide market, and the needed investment in plant, equipment, and sales organization was prohibitively high for a newcomer to raise on the uncertain prospect of being able to break into a highly competitive business in the face of the existing well-established companies.

The one major exception was the Chrysler Corporation, established in 1920. However, the corporation was actually a reorganization of the Maxwell and Chalmers companies, and as such can not be considered a new enterprise.

In summary, economies of scale, a cumulative parts burden, cumulative learning effects, positive returns to innovation, and process innovations that were difficult to adopt served as selection mechanisms by reducing the number of firms in the industry. But other factors contributed to the failure rate of firms. The redefinition of the automobile for the mass market, an emerging dominant design, and saturation of the market for expensive vehicles (see section 4.5.4) also played a significant role in the shakeout phenomenon. Epstein (1972) concludes that "[t]he extraordinary exit figure in 1910, of 26% of all firms engaged in the field, is probably due, above all, to the absolute falling off in the demand for high-priced cars which occurred in that year." Firms that were unwilling or unable to adapt to production of smaller lightweight vehicles were effectively weeded out. In addition, the Panic of 1907 and ensuing recession (see section 4.5.1) occurred just prior to the year in which the largest number of firms exited the industry and likely played a role in the failure of some of these firms, particularly those manufacturing expensive automobiles for the wealthy, who would have been hardest hit by the stock market plunge. Thus, by contributing to the failure of specific firms, a random event and the selection environment influenced variation and further reinforced the emerging dominant design.

4.5 Technical and Institutional Context

This section discusses five developments in the TIEC that provide the technical and institutional context for this era: the Panic of 1907; the Selden Patent patent dispute; the emergence of trade associations and their efforts to standardize parts and materials; the expansion and evolution of the autombile market; the activities of motoring clubs; and efforts to improve the condition of the country's roads.

4.5.1 The Panic of 1907

After a period of economic prosperity, the nation faced an economic slump in early 1907. The economic weakness was due in part to losses from the violent earthquake that shook the San Francisco Bay area on April 18, 1906 and the fires that ravaged San Francisco for the next four days. However, in October, 1907, the economic downturn, an unusually tight money supply, and a failed attempt by F. A. Heinze to corner the stock of United Copper Company combined to trigger one of the most severe bank panics the country had experienced to date. Heinze was both a speculator and a banker whose dealings involved "an intricate network of interlocking directorates across banks, brokerage houses, and trust companies in New York City. Contemporary observers… believed that the close associations between bankers and brokers heightened depositors' anxiety" (Tallman and Moen, 1990). As news of Heinze's financial losses and bank connections spread, depositors, fearing insolvency, rushed to withdraw their savings from banks where Heinze and his associates held prominent positions. The bank run spread to the third largest trust in New York City, Knickerbocker Trust Company, whose president, Charles Barney, was also rumored to have been involved in Heinze's corner attempt. Because Barney also served on the board of directors of Trust Company of America, the second largest trust in New York, it soon was involved as well. The bank run lasted two weeks, during which Trust Company of America reportedly paid out $47.5 million of its nearly $60 million total deposits.

To avert the failure of several large trusts, J. P. Morgan channeled $3 million to Trust Company of America on Wednesday, October 23. J. D. Rockefeller announced his support of Morgan and deposited $10 million with another trust. However, with so many of New York's banks in trouble, the crisis had spread to Wall Street. On Thursday, October 24, no money was available on the New York Stock Exchange floor for the purchase of stock equity despite offers of up to 60% interest. Prices plummeted and the stock market was threatened with total collapse. Less than an hour before the closing bell, J. P. Morgan organized a group of bankers who made a pool of $25 million available to the exchange. The group provided an additional $10 million pool for call loans the following day and other sources provided $2.5 million. Over a period of ten days, the U.S. Treasury deposited a total of $37.6 million in New York national banks and provided $36 million to meet bank runs. In early November, Morgan convinced other trust presidents to provide another $25 million in loans to threatened institutions. The panic finally eased when these presidents formed a consortium to support trust companies facing runs. Though the panic had ended, the nation suffered a deep but brief recession in 1908 with GDP falling about 11%. The stock market suffered more severely, with the Dow Jones Industrial Average falling about 50% from a peak value in January, 1906, before bottoming out in November of 1907.

Bank runs and economic panics were fairly common in the 19th and early 20th century, occurring in 1819, 1837, 1857, 1873 and 1893. The Panic of 1907 underscored the weaknesses of the U.S. financial system and served as a catalyst for banking reform. In May, 1908, Congress passed the Aldrich-Vreeland Act, which eased the credit situation in the short-term and established the National Monetary Commission. In 1912, the Commission recommended adoption of a central banking system to be coordinated by a board of commercial bankers. Congress subsequently crafted the Owen-Glass Federal Reserve Act, enacted in December, 1913. The Act established twelve regional Federal Reserve Banks to act as central banks for national and other member financial institutions. The Banks were not Federal bodies, but were privately owned by the member banks. However, the Act formed a public body, the Federal Reserve Board, to oversee the system rather than leave control in the hands of private bankers. The Act also created a new form of currency, the Federal Reserve Note, which was an obligation of the U.S. Treasury.[46] The tight money supply encountered in 1907 arose in part because seasonal increases in economic activity, primarily driven by agriculture, were not matched by increases in the money supply. The Federal Reserve Banks were authorized to issue Federal Reserve Notes to solve this problem of inelasticity. The Federal Reserve System was also authorized to establish the discount rate – the interest rate the Banks charged to its member institutions. By making loans to commercial banks and issuing currency, the Federal Reserve System was intended to avoid the problems that contributed to the Panic of 1907.

4.5.2 The Selden Patent Dispute

After witnessing a demonstration of the Brayton engine in 1879, patent attorney George Selden drew up plans for a "road-locomotive" consisting of a four wheeled

[46] The Federal Reserve Note now makes up the nation's supply of paper money.

vehicle powered by a gasoline engine similar to Brayton's. Selden applied for a U.S. patent, but made minor changes to the application annually to delay its formal registration until November, 1895, the same month as the first U.S. reliability race which was held in Chicago (Rubenstein, 2001). Although Selden had not invented any component of his design, he claimed it represented an original combination of elements. The patent design was defined in broad terms, covering "the combination with a road locomotive, provided with suitable running gear including a propelling wheel and steering mechanism, of a liquid hydrocarbon gas-engine of the compression type, comprising one or more power cylinders, a suitable liquid-fuel receptacle, a power shaft connected with and arranged to run faster than the propelling wheel, an intermediate clutch or disconnecting device and a suitable carriage body adapted to the conveyance of person or goods" (U.S. Patent Office, 2008).

Selden sought financial backing to manufacture his automobile but ultimately was unsuccessful. Though he never built a single vehicle, he claimed the engine and vehicle were operable and that the patent was basic. Therefore, he asserted, it covered all vehicles built after 1879 that used a compression engine, and no automobiles could be built during the seventeen years that the patent was valid (1895-1912) without his permission.

Lacking the resources to enforce his patent, Selden in 1899 assigned the rights to a group of investors that also bought the Electric Vehicle Company (EVC). Selden and the investors believed that restricting production would maintain high prices and royalties. After the EVC filed several infringement suits between 1900 and 1903, a group of automakers agreed to settle the dispute by recognizing the validity of the patent in exchange for exclusive license to manufacture automobiles under its stipulations. As part of the agreement, a new trade organization called the Association of Licensed Automobile Manufacturers (ALAM) was formed to enforce the license and prevent ‘disreputable’ manufacturers from entering the business (May, 1990). The ALAM leased to its members the right to sell a limited number of automobiles. The ALAM also was charged with deciding what companies should be sued for patent infringement (Rubenstein, 2001). However, since the ALAM reserved the right to *deny* licenses, many industry observers opposed the patent claim on ethical grounds (Nevins and Hill, 1954).

Ford approached the president of the ALAM after forming the Ford Motor Company in 1903 but was told the association likely would not rule favorably on an application because the company was nothing but an ‘assemblage’ plant rather than a true manufacturer. ALAM officials recommended that Ford delay application until the company had its own plant (Nevins and Hill, 1954). A fierce dispute ensued, with Ford publicly challenging the ALAM to take legal action. The ALAM subsequently filed infringement suits against Ford, Ford’s New York sales agent, and a company that had purchased a Ford automobile. The ALAM also sued a foreign auto manufacturer (Panhard et Levassor), and in 1904, an automobile importer.

The five lawsuits were eventually combined into one and litigation dragged out over the next five years. The defense prepared arguments based on two interpretations of the patent: either the patent was broad, covering all engines of the compression type, or it was specific to automobiles utilizing the two-stroke, external compression, Brayton engine. Under the first interpretation, the defense argued, the patent was invalid due to “prior art” since patents, devices, and machines that antedated Selden’s application demonstrated that it did not represent a new invention. Under the second interpretation, they denied any infringement because nearly all gasoline automobile engines employed the four-stroke internal compression Otto cycle.

In early 1905, the ALAM gained control of the lease of Madison Square Garden, where the annual Automobile Show was held, and announced that only ALAM members would be permitted to exhibit at the show. Less than a month later, twenty independent manufacturers formed the American Motor Car Manufacturers Association (AMCMA) and leased nearby property to house its own exhibition. The AMCMA was formed with seven objectives which basically served in the same capacity as the ALAM did for its members: 1) to arrange for public exhibition of members' cars; 2) to promote races; 3) to promote the sale of members' cars; 4) to encourage public interest in automobiles; 5) to establish agencies; 6) to promote good roads; and 7) to exchange technical information. By 1908, the AMCMA consisted of 48 companies.

As the litigation dragged on, the ALAM sent notices demanding payment of royalty fees by owners of imported automobiles and filed an additional seventy suits against independent manufacturers (Nevins and Hill, 1954). The fate of the industry depended on resolution of the case against Ford, which finally was decided in September, 1909, when the judge filed in favor of Selden. Ford had introduced the Model T a year earlier and was liable for royalties running into the millions of dollars on the cars the company had sold since being formed in 1903 (Rubenstein, 2001). Sorensen (1956) states that if the decision had been upheld Ford would have faced financial ruin, though others dispute this claim.

Independent manufacturers rapidly withdrew from the AMCMA and sought licenses from the ALAM; on February 9, the AMCMA disbanded. However, Ford's resolution never faltered, and he promptly appealed the court's decision. In January, 1911, the U.S. Court of Appeals upheld the validity of the Selden patent but ruled that it only covered automobiles built with the modified two-stroke Brayton-type engine Selden had shown in his application. The court therefore found that the patent had not been infringed. With only two years remaining on the now worthless patent, the ALAM decided not to appeal (Nevins and Hill, 1954; May, 1990). The ALAM was disbanded in 1911 and transferred responsibility for promoting and improving the industry to the newly organized Automobile Board of Trade, which in 1914 became the National Automobile Chamber of Commerce (NACC).[47] To prevent a similar costly patent dispute in the future, the NACC instituted a cross-licensing agreement among its members (Flink, 1988).

4.5.3 Trade Associations and Parts Standardization, 1905-1910

In an editorial in 1902, Peter Heldt of *The Horseless Age* suggested that a technical society devoted to automobiles would be best suited to solving technical questions requiring cooperation among industry engineers. Horace Swetland, whose editorials in *The Automobile* had served as the voice of the automobile engineer, also promoted the idea. In 1905, a group of trade journalists founded the Society of

[47] The NACC became the Automobile Manufacturers Association in 1932, and then the Motor Vehicle Manufacturers Association (MVMA) in 1972 (Flink, 1988). When foreign manufacturers began production in the U.S., they joined the MVMA. Due to differences in objectives between these foreign owned companies and the remaining U.S. companies, the 'big three' (Ford, GM, and Chrysler) dissolved the MVMA in 1992 and formed the American Automobile Manufacturers Association (AAMA), a U.S. only lobbying group. When Daimler-Benz purchased Chrysler in 1998, the AAMA was dissolved. Ford and GM now participate in the international organization, the Alliance of Automobile Manufacturers which was created in 1998 (Rubenstein, 2001).

Automobile Engineers (SAE)[48] to serve this function, primarily through the publication of technical articles. Andrew Riker, an automotive engineer who had built electric cars beginning in 1888 and gasoline cars for Locomobile after 1902, served as the first president and Henry Ford served as the first vice president. Beginning with 30 members, the society grew over the next ten years, adding full-time staff and the publication of a technical journal, the SAE Transactions (SAE, 2007; Flink, 1988).

Perhaps the most valuable contribution of the SAE came through work begun by the Association of Licensed Automobile Manufacturers (ALAM). Although the ALAM was organized to enforce the Selden patent and therefore served as an instrument to wield monopolistic market control, in 1905 it established a Mechanical Branch whose aim was to promote technical collaboration among manufacturers. Composed of around 100 engineers and plant superintendents who met regularly to exchange information, the Mechanical Branch initiated a program to standardize automotive parts and materials. It also established a metallurgy lab to test automotive parts and materials.

During the Panic of 1907, funding for the ALAM's Mechanical Branch was cut and never restored. Its work and valuable library was transferred to the SAE in 1910. In the process, the SAE's membership jumped from 310 to 899. Members previously associated with the ALAM who became influential officers in SAE included Coker Clarkson, manager of the Mechanical Branch, and Henry Souther, who had been in charge of ALAM's experimental and testing laboratory. Clarkson, Souther, and SAE president Howard Coffin, who was also vice-president of the Hudson Motor Car

[48] With Orville Wright's backing, the SAE incorporated the aeronautical industry in 1916 and changed its name to the Society of Automotive Engineers to represent any form of self-powered vehicle.

Company, were largely responsible for completing the important work on standardization begun by the Mechanical Branch (Barnes, 1921).

Prior to 1910, automotive manufacturers were using at least 800 different bolts, 800 different sized lock washers, 1,600 different sizes of steel tubing, and some 230 different steel alloys (Barnes, 1921; Epstein, 1972). The ALAM Mechanical Branch and the SAE standardized small parts, including spark plugs, screw threads, nuts, bolts, tubing, and rods, resulting in 16 standard sizes of washers and 210 types of steel tubing. The development of uniform specifications for 50 steel alloys allowed steel makers to supply superior grade materials to the industry (Epstein, 1972; May, 1990).

In general, the SAE represented the interests of small automotive firms whose employees constituted most of its membership. These firms had the most to gain from the standardization of parts because it allowed them to purchase small orders of readily available, standard, small components at much lower prices, essentially helping them realize economies of scale formerly only available to large manufacturers. Unfortunately, standardization ultimately failed to keep the small producer competitive (Flink, 1988). Nonetheless, standardization of parts brought order out of chaos and eliminated waste. And because of the micro scale of these parts, standardization did not stifle innovation in vehicle design (Epstein, 1972).

4.5.4 Market Expansion

At the beginning of the century, popular opinion held that the motor vehicle would never be more than a toy for the rich. Ford (Ford and Crowther, 1922, p. 36) recalled, "No man of money even thought of it as a commercial possibility… [I]n the beginning there was hardly any one who sensed that the automobile could be a large factor in industry. The most optimistic hoped only for a development akin to that of the bicycle." But the market for automobiles quickly shifted from the very wealthy to the lower upper class. According to Nevins and Hill (1954, p. 241) "…as a group, doctors were among the first to make significant use of the automobile." Epstein (1972, p. 95-97) lists the occupation of the first purchasers of automobiles as capitalists, manufacturers, merchants and physicians. Notably, among the purchasers of the first twenty Waverly Electric cars sold in Detroit by 1898 there are six "ladies – wives of the above class, including two wives of bankers." Purchasers of the relatively inexpensive Olds curved dash model between 1900 and 1903 were mostly merchants and physicians, and no capitalists or manufacturers are listed.

Beginning in 1904, the U.S. enjoyed a period of growing prosperity the scale of which had never before been seen. Political reforms and the rise of new industrial methods resulted in an increase in production, incomes, and standards of living. Technological innovation had brought the telephone, electric lighting, the phonograph, modern plumbing, and the bicycle. Automobile production reached 11,000 units in 1903 and doubled in 1904.

The growing popularity of the automobile was accompanied by a shift in the functional definition of the product. According to Epstein (1972), the purposes of the motor vehicle can be classified under four headings: transportation service, sport, personal possession, and social prestige. As 'playthings for the rich,' the earliest automobiles largely fulfilled the desire for sport – the thrill of rapid, cross country transit – and social prestige – a public demonstration of wealth. A writer in the *Independent* declared in 1906, "The man who owns a motorcar gets for himself, besides the joys of touring, the adulation of the walking crowd, and the daring driver of a racing machine that bounds and rushes and disappears in the perspective in a thunder of explosions is a god to the women" (cited in Flink, 1970, p. 64). The social prestige conveyed by automobile ownership was so alluring that by 1908 many Americans were recklessly mortgaging their homes to purchase vehicles they could not afford, leading even to foreclosures (Flink, 1970).

But to the physician who traveled frequently about town, the automobile was utilitarian as well, providing more rapid transport and promising more economical service than a horse and buggy. The initial expense of a moderately priced motor car was comparable to that of a carriage and, although the monthly maintenance costs of early automobiles often turned out to be more than the upkeep of a horse, the automobile could do far more work and was likely to last longer than a hard-used horse (Flink, 1970). The automobile was not willful and unpredictable, nor was it susceptible to heat or fatigue.

By 1905, demand for automobiles outstripped capacity, and cars had gone from a fashionable novelty for the very rich to a necessity for well-to-do business and professional men (Nevins and Hill, 1954). Automobile manufacturing was significant enough for the U.S. Bureau of the Census to publish separate statistics for industry that year. In the introduction to the report, George E. Oller commented, "As a means of amusement, [the automobile's] popularity may fluctuate or decline, but its practical value has been so thoroughly demonstrated that its use will doubtless become more general each year, until it is displaced by some vehicle as much its superior as the automobile is the superior of the horse and wagon" (cited in Jarvis, 1972, p. 64). The automobile was rapidly becoming a utilitarian product that fulfilled the necessary service of transportation, though with initial and repair costs it was still limited to the well-off.

At the turn of the century, farmers generally viewed the automobile with indifference. But by 1904, they became noticeably hostile with the rise of touring. The early touring automobiles were purchased by wealthy people "of a sporting and adventurous turn of mind who sometimes used them ruthlessly on the highways" (Pound, 1934, p. 38-39). Speeding automobiles on dirt roads endangered horses and livestock and raised clouds of dust that settled on crops, houses, barns, and wash hung out to dry (Flink, 1970). At first, farmers could not foresee owning cars themselves, and felt the automobile "flaunted idleness and conspicuous consumption in the faces of hardworking, hard-pinched men" (Nevins and Hill, 1954, p. 397). Some farmers took extreme actions, plowing up roads, stringing barbed wire across roads, threatening to boycott businessmen who purchased cars, and lobbying for resolutions to ban automobiles from state roads. However, most opposition was more balanced, with agricultural publications and organizations lobbying for better legislation and enforcement of existing laws regarding speed and reckless driving (Flink, 1970).

The anti-automobile sentiment among farmers was short-lived as prosperous farmers became ever more familiar with the vehicles and began to envision owning automobiles themselves. A number of manufacturers in the Midwest produced cheap buggy-type cars with a light surrey body, solid rubber tires on carriage wheels, a small one- or two-cylinder engine mounted under the seat, and chain or rope drives. These cars were well-suited to navigating poor rural roads and if they did get stuck were light enough to easily be pushed out (Flink, 1970). However, they suffered from inherent mechanical weaknesses and required frequent repairs. In 1906, Ford introduced the rugged and moderately priced Model N that delivered more comfort, reliability, and performance for a similar price. By 1907, farmers' opposition to the automobile had given way to a substantial rural market and dealers in many medium-sized cities were selling more cars to farmers than to city inhabitants. The American Motor Car Manufacturers' Association reported in 1909 that manufacturers were selling to two large new segments: the farmer and the middle class (Flink, 1970).

Americans were relatively wealthy by European standards, the wealth was more equally distributed, and social class distinctions were less pronounced. As a result, the U.S. constituted a large market for standardized, homogenous products – the perfect conditions for mass production. In addition, much of the population was highly dispersed, creating a large market for a rugged vehicle capable of rapid, long-distance travel. The motor car could bring significant advantages to these rural populations. Prior to the automobile, traveling 20 or 30 miles by horse took a whole day and the farmer was dependent on the railroad to transport his products to the market. With an automobile, the farmer's social radius expanded and he could market directly to a wider area.

Figure 4-2 shows sales of motor vehicles between 1903 and 1916 stratified by three price classes. Even in 1903, low-priced cars were each outselling the medium- and high-price categories. At that time, most of the vehicles in the least expensive class were light, low-horsepower, surrey-styled vehicles, with the Olds curved dash accounting for more than half of these sales.[49] These vehicles tended to rattle apart quickly on rough

[49] Flink (1988, p. 33) refers to the curved dash as "merely a motorized horse buggy." Olds reportedly sold 4,000 vehicles in 1903, while a total of 7,253 vehicles were sold for less that $1,375 (Epstein, 1972).

country roads. By 1910, however, surrey-style design was largely abandoned and the reliability of vehicles in this price class was much improved.

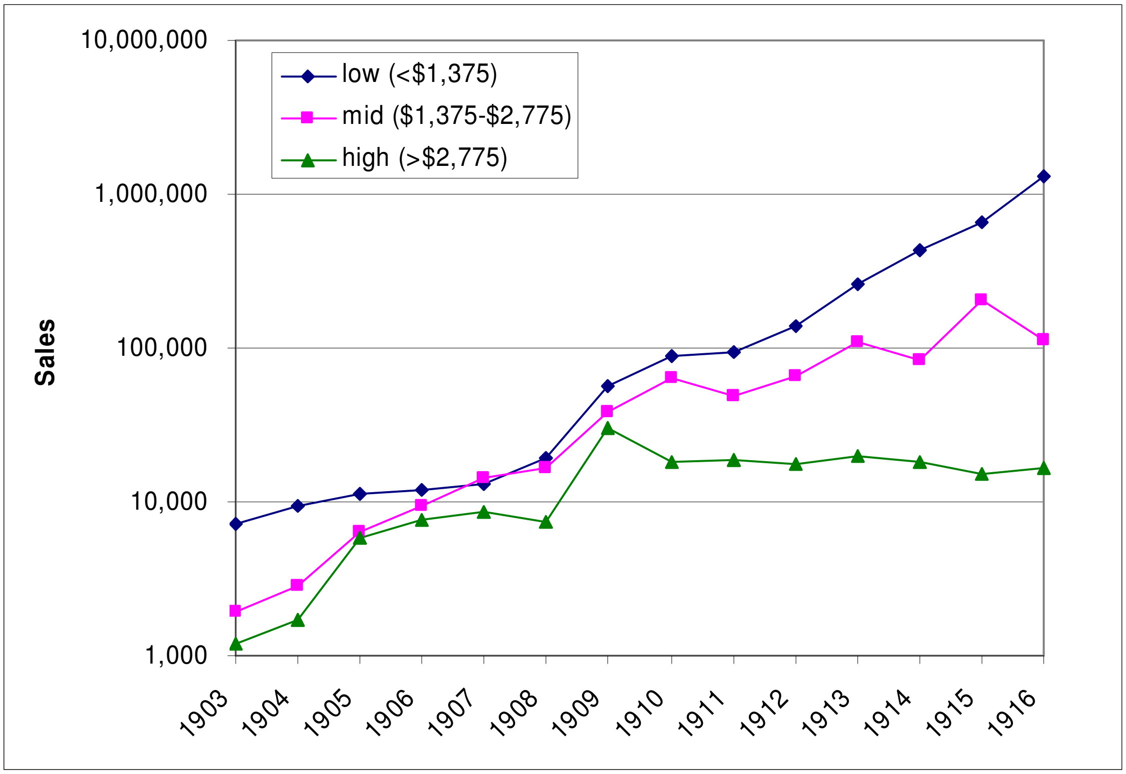


Figure 4-2: Sales by Price Class
Source: Based on data from Epstein (1972, Appendix B)

This development of a new market for automobiles after 1900 represents the onset of wider diffusion of the motor vehicle. The niche hypothesis asserts that growth in initial niche markets yields cost reductions through learning-by-doing and economies of scale, and that these reductions facilitate diffusion to wider markets. However, as the automobile was penetrating a new segment, the market for expensive cars was no longer growing, it was saturating – buyers in this market were not purchasing their first car, but were almost exclusively replacing an older car. The proportion of the market consisting of moderately priced cars (over $1,375) reached a maximum of almost 64% in 1907 then began a steady decline. Similarly, expensive cars priced over $2,775 declined after reaching 31% of the market in 1907. The down-turn in demand for expensive motor vehicles in 1908 may be attributed at least partially to the Panic of 1907 and the ensuing recession (see section 4.5.1). It is likely that buyers of this class of vehicles were hit hardest by the stock market plunge in 1908, but it appears that many of these purchases were delayed rather than foregone. As shown in Figure 4-2, sales of expensive cars spiked in 1909 then stabilized in 1910 at an average of 17,400 per year, which is slightly lower than Epstein's (1972) estimate of replacement demand.

The original niche markets for gasoline automobiles were not only stabilizing, but Figure 4-3 shows that the trend within the high-price category between 1903 and 1908 was, if anything, toward *higher* price, not lower, since vehicles priced above $3,375 were gaining market share over those priced between $2,775 and $3,775. Instead, innovation was improving the performance of these vehicles relative to their functional requirements. According to Epstein (1972), the trend in these cars was toward ever higher horsepower and weight. Thus, while in 1904, the high-priced vehicle was represented by a four cylinder vehicle, by 1916, it was more likely a six-cylinder that produced 60 hp. Similarly, the representative low-priced vehicle had progressed from a light two-cylinder engine to a four-cylinder that produced 30 hp. Thus, while vehicles in all price classes had improved dramatically in performance, the relative size, workmanship, and performance of each class remained comparable over the period.

While demand for the highest price class had stabilized, demand for medium-priced vehicles continued to grow, but at an ever decreasing rate. With a midpoint around $2,000, the cost of these vehicles was still high enough that "their consumption was mainly indulged in, up to 1910-1911, by the very wealthy" (Epstein, 1972, p. 94).[50] Thus, the market for cars priced above $1,375 consisted primarily of manufacturers, capitalists, bankers, merchants and physicians.

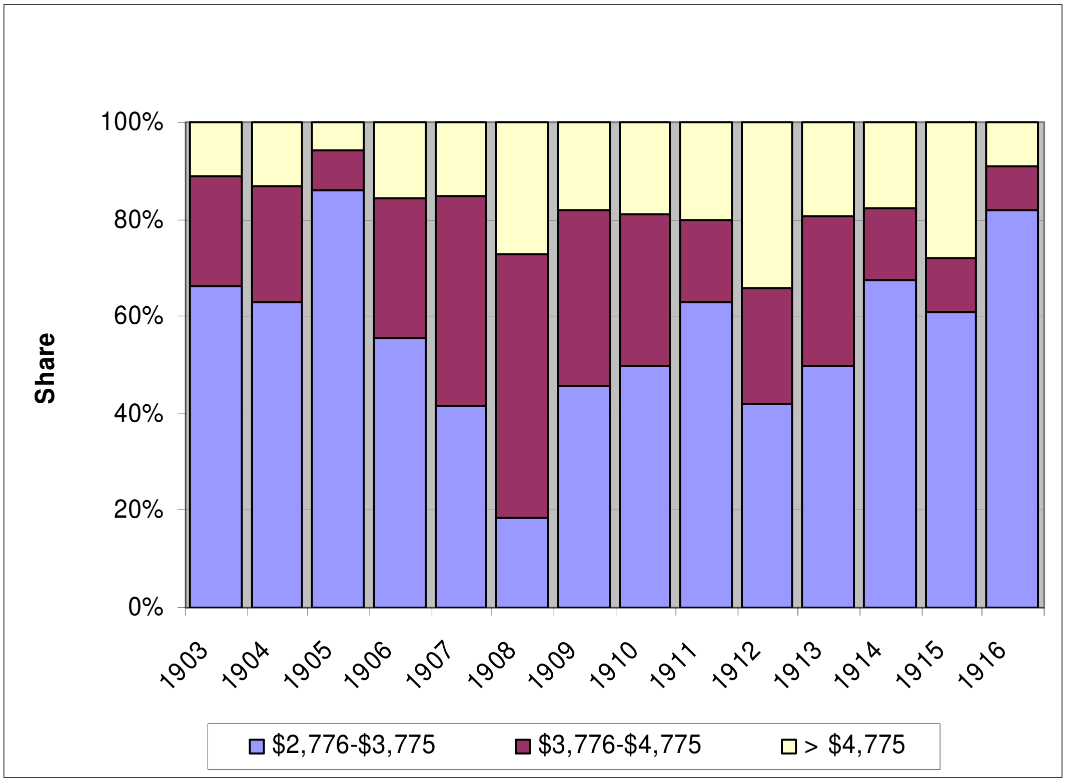


Figure 4-3: Market Share within the High-Price Category
Source: Based on Epstein (1972, Appendix B).

Meanwhile, sales of inexpensive cars experienced the highest rate of change with growth actually accelerating during this period. In 1903, merchants and physicians dominated this market. As shown in Figure 4-4, the trend within this category after 1907 was clearly toward lower price as well as improved performance, with sales under $675 gaining market share on those priced between $675 and $1,375. By 1915, cars selling for

[50] For reference, nominal GDP per capita was only $322 in 1903 and rose to $362 by 1910.

under $1,375 constituted nearly 75% of the market, while those priced above $2,775 amounted to less than 2% of sales (Epstein, 1972).[51]

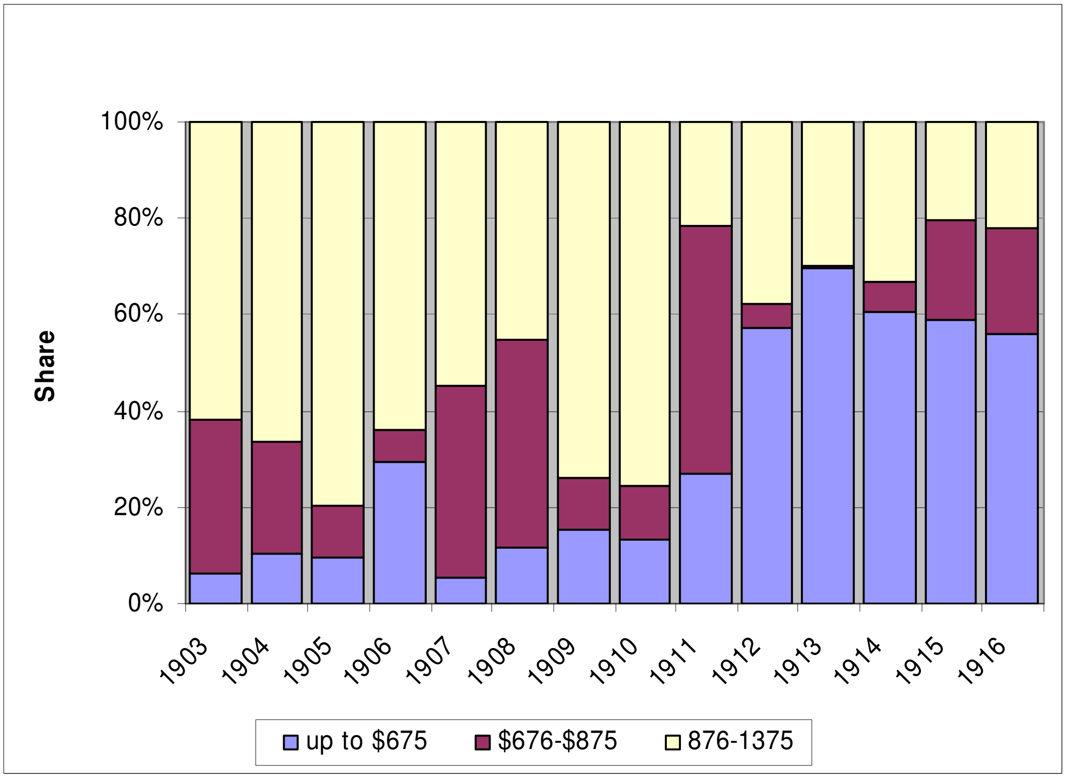


Figure 4-4: Market Share within the Low-Price Category
Source: Based on Epstein (1972, Appendix B).

This pattern is not consistent with the assertions of the niche hypothesis. In fact, while some of the largest manufacturers offered models in both the low- and medium-priced classes, the market for expensive touring cars was for the most part served by different manufacturers than the market for inexpensive cars. The strategy among makers of the most expensive vehicles was not to achieve high volume sales or low price, but rather to maintain high quality through hand craftsmanship, increase power, and sustain high prices and thus high profits. Such firms that were unwilling or unable to adapt to production of smaller lightweight vehicles faced fierce competition within a

[51] Epstein's (1972) data is not adjusted for purchasing power.

stagnating market. Epstein (1972) asserts that the saturation of the high-priced market was the most significant factor behind the extraordinary number of firms that exited the industry in 1910 (see section 4.4).

Henry Leland at the Cadillac Motor Car Company was one of the first manufacturers to make progress in reducing costs and maintaining quality in volume production. However, between 1903 and 1908 the company moved toward production of moderately priced cars of exceptional quality like the Cadillac Model 30 which sold for $1,400 in 1908 (Flink, 1988). While a few manufacturers like Brush attempted to make cars for the lowest price bracket (under $675), their methods of reducing costs often led to poor quality vehicles with significant mechanical weaknesses. In this effort, Ford emerged the leader in pursuing ever lower prices and unprecedented volume production of reliable, high quality, basic transportation, recognizing the vast market attainable with this novel strategy. By 1915, the Ford Model T was priced starting at $390 and accounted for 53% of all automobile sales and roughly 70% of sales under $1,375 (Epstein, 1972; Cain, 2006; Nevins and Hill, 1954). While designed with the rural market in mind, the low price, high reliability and high quality of Ford's vehicles also served the needs of the urban middle class.

In 1900, as many as 57 plants were engaged in manufacturing automobiles, though some were still in the experimental stages. Around 4,000 motor vehicles were manufactured that year, but designs and technology varied dramatically and steam- and electric-powered vehicles accounted for at least 75%. But by 1910, sales of motor vehicles exceeded 180,000 units (Cain, 2006) and the extreme variety of product configurations and the accompanying customer confusion had largely disappeared. A standard design had appeared that served the great majority of the market: a shaft driven, magneto equipped, vehicle powered by a water-cooled, four-stroke, four-cylinder, gasoline internal combustion engine mounted vertically under a front hood, with a steering wheel and a three-speed sliding-gear transmission (Epstein, 1972).

4.5.5 Motoring Clubs, 1901-1910

The first automobile club, the American Motor League, was formed on November 1, 1895, just prior to the Chicago *Times-Herald* race. The organization was premature and failed, but several local automobile clubs in large cities were active by the turn of the century. The members of these early clubs were almost exclusively very wealthy and socially prominent and the clubs emphasized social functions. However, automobile ownership quickly became too widespread to form the basis for an exclusive social club. At least twenty-two local organizations existed by early 1901 and almost 100 U.S. clubs were active by early 1904 (Flink, 1970). Motoring was becoming increasingly utilitarian by 1905, and the majority of motorists were concerned with local issues of motor vehicle legislation, road improvements, and community relations. Clubs in smaller cities and rural areas developed into democratic organizations that charged small dues, drew membership from a wide cross section of the population, and set limited practical goals.

Yet many motorists believed in the need for a national association of local clubs to deal with issues beyond local concerns. Two such affiliations were founded in 1902. In February, eight local clubs founded the American Automobile Association (AAA). In March, the American Motor League (AML) was revived by its original organizers and representatives from two cycling associations: the League of American Wheelmen and the National Cycling Association. The AML was organized as a national association of individuals and aimed to advance the interests of motorists and add to the value of motoring through reports of the mechanical features of automobiles, education, lobbying, defense of motoring rights, and promoting better roads. The AML intended to establish branches throughout the states, but its influence never rivaled that of the AAA.

The AAA was designed as an affiliation of local clubs that would pursue legislative issues, protect the interests of motorists, and promote the construction and maintenance of good roads and highways. In mid-1903, the AAA opened its membership to individual motorists. However, legislative issues were largely local and aggressively pursued by local clubs. Therefore, the AAA accomplished little in the first few years. Dissention erupted among the organization's officers and dissatisfied clubs began to withdraw, believing the AAA delivered little for their dues. The AAA accepted a more limited role for the national organization, granting more power at the local level, and membership began to rise again in 1908. At the same time, the organization established a relationship with the automobile industry by granting three seats on its board of directors to the National Association of Automobile Manufacturers (NAAM) in exchange for a $5,000 contribution. By 1909, the AAA consisted of 225 clubs from thirty state associations with total membership over 15,000 and was firmly established as the voice of the American motorist (Flink, 1970).

Public safety concerns prompted municipalities and states to consider legislation regarding vehicle registration, speed regulation and driver licensing. Automobile clubs actively opposed the first two. Registrations, initiated in New York in 1901, allowed officials to pursue reckless and speeding drivers and to track down owners who evaded vehicle property taxes. Other states followed suit in 1903. Registrations were either annual or perennial and in some states were initially established at the county level, requiring motorist to obtain multiple registrations. Automobile clubs and the NAAM contested the legality of compulsory registration, claiming class discrimination since horse-drawn vehicles were not required to register and display identification tags. However, motorists' main objections were based on fear of punishment – arrest, fines, imprisonment, and damage claim liabilities – for tax evasion, reckless driving, and accidents. Of course, such illegitimate objections were never voiced in court and the constitutionality of small registration fees was repeatedly upheld.

Automobile associations shifted attention to establishing reasonable state laws to avoid multiple and possibly costly fees at the municipal level. Eventually, these associations recognized that registration taxes were often used to improve roads and as such were a benefit to their members. By 1910, thirty-six states had enacted motor vehicle registration laws, and all states had done so by 1915 (Flink, 1970).

4.5.6 Road Improvements, 1880-1916

At the turn of the century, American roads were in poor condition compared to European countries, particularly France where the first automotive industry was born. The U.S. federal government was involved in road-building early in the 19th century, but abandoned this role in 1837 due to the rapid growth of railroads and an economic depression which followed a banking panic.[52] Therefore, constructing and marking roads was left to counties and municipalities and it was common for road conditions to change dramatically at county lines.

[52] The Panic of 1837 was preceded by a speculative bubble which burst in May. An unprecedented number of banks failed and unemployment reached record levels during the ensuing five-year depression.

Bicycle manufacturers, dealers, and enthusiasts began promoting better roads as early as 1880. At the 1893 World's Fair in Chicago, the Columbia Exposition, General Roy Stone organized the National League for Good Roads which established branches in most states. With Albert Pope's assistance, the League of American Wheelmen, a bicycling organization, drafted a petition to Congress, resulting in the birth in 1893 of the Office of Road Inquiry in the U.S. Department of Agriculture. This agency became the Office of Public Roads, which conducted a census of American roads in 1904. The census reported that only 7% of U.S. roads were surfaced, most with just gravel. There was only one mile of improved road per 492 inhabitants, and most of this mileage was of poor quality unsuitable for steady motor vehicle use (Flink, 1970). In truth, there still were few roads passable by automobile.

With the demise of the bicycle industry, leadership for the "good roads movement" shifted to automobile interests around 1903. However, the leaders were largely the same individuals, as bicycle manufacturers entered the automotive industry and bicycle enthusiasts bought automobiles and joined or organized motoring clubs. For example, by 1899, Albert Pope had become one of the country's leading automobile manufacturers with the production of 500 electric and 40 gasoline vehicles (Rae, 1959).

The influence of motor vehicles and the automobile clubs propelled the need for good roads to public consciousness. Prominent political leaders, including President Theodore Roosevelt, attended the National Good Roads Convention in St. Louis in 1903, where resolutions were passed endorsing federal support for highway construction and state aid to counties and municipalities. Activists exploited issues of popular appeal, including the high cost of transporting farm products to the market, which was due in part to high rates charged by the railroads (Flink, 1970; Geels, 2005b). The AAA and the National Grange, representing motorists and farmers respectively, were the dominant influences at the second Good Roads Convention in 1909.

The federal government gradually increased funding for the Office of Public Roads and states increasingly passed legislation designed to improve highways, often funding road construction using revenues from registration fees. By 1910, total mileage of improved roads had increased by 48,266 miles, but this represented only a 1.5 percentage point gain over 1904. More significantly, gravel was being replaced by macadam. Yet these improvements were easily offset by the rising demands of automobile traffic as even macadam broke down rapidly under frequent use. Concrete, first used on a 24-mile stretch of highway to Ford Highland Park plant in Michigan in 1908, proved to be the most durable and economical solution. Some 750 miles of concrete roads were added nationwide in 1912 and 1913. California rapidly adopted the use of concrete to serve the transportation needs of its agriculture, covering 35% of its 2,600 miles of improved roads with concrete by 1914 (Nevins and Hill, 1954; Flink, 1988).

In 1912, a businessman and automobile enthusiast pitched an idea for a continuous, coast-to-coast highway to a group of prominent automotive and supplier executives. Led by automobile interests, the Lincoln Highway Association was organized in 1913, having already garnered pledges for four of the estimated $10 million required for the project. Although it was approved by the Conference of Governors that year, planning for the project was plagued by a multitude of problems, including financing, pressure to change the route, and the inability of states and communities to act harmoniously. The association downgraded its activities and the road was never completed under its original name, the Lincoln Highway. The federal government eventually took over the task in 1923 and used its own system of numbering its routes (Nevins and Hill, 1954).

The breakthrough in road construction began in 1916 when Congress passed the Federal Aid Road Act, appropriating $75 million to be distributed to states in matching grants over a period of five years. The main intent of federal involvement was to improve postal roads for Rural Free Delivery instituted by the Post Office in 1896. To receive matching grants, states had to have "properly constituted highway departments to handle the road-building projects. The result was the modernization of all state activity, as well as a marked increase in it" (Nevins and Hill, 1954, p. 487).

4.6 Synthesis and Analysis

This section analyzes the history of the industry between 1900 and 1915 using the approach developed in chapter 2. This era traces the industry's first transitional phase and includes the onset of the first specific phase around 1910. Because history is cumulative, this analysis also builds on the narrative from chapter 3, which covers the industry's emergent phase.

Section 4.6.1 discusses the emerging dominant design and diffusion of the gasoline automobile into wider markets. The QE and MLP theories assert that, within niche markets, product and process improvements achieved through learning-by-doing and learning-by-interacting with users allow new technologies to define or conquer new niches. This process of niche proliferation then allows further cost reductions through economies of scale, which facilitates diffusion to wider markets. I find that the history of the diffusion of automobiles was not consistent with the assertions of the QE and MLP theories in the following ways. First, movement from niches to a wider market occurred simultaneously with a functional redefinition of the automobile. The needs of this market were not discovered within the niche, nor were product improvements *relative to these needs* achieved in the niche market. Product improvements and cost reductions from both process innovations and economies of scale were all realized within the new market for reliable, rugged, inexpensive vehicles. And because the vast, primarily rural, low-price market can not be accurately described as a market niche, niche branching, niche proliferation, or niche market growth are not accurate descriptions of the diffusion of the gasoline motor-vehicle into mass markets.

Section 4.6.2 discusses the process innovations that allowed motor vehicles to conquer the mass market. I identify four facets of the TIEC that enabled the volume production of automobiles. I find that Ford's innovative production system resulted from a highly synthetic process and relied on knowledge spillovers as well as learning-by-doing. However, Ford's greatest accomplishment was the manufacture of a product of both high quality and low cost, a combination that had not yet been achieved in the industry. Ford's success was accomplished by jointly addressing issues of product design and manufacturing processes. Thus, product and process innovations that served the needs of wider markets originated concurrently with the development of the large rural market and arose from a more complex process than simple learning-by-doing.

In section 4.6.3 I address the mechanisms of learning and innovation during the transitional and early specific phases. I find that the emphasis on learning-by-interacting with users and learning-by-doing found in the QE and MLP literatures is overly simplistic. While these were important mechanisms of learning during this timeframe, spillovers and learning-by-interacting with suppliers were also extremely prominent. Consistent with the assertions of the industry life-cycle model, the first attempts at systematic research within firms were instituted at the onset of the specific phase.

Section 4.6.4 explores the continuing development of the technological nexus linking variation and selection and the ongoing mutual adaptation of the technology and institutions. I find that the technology itself was a major factor in altering the selection environment, first inhibiting then later enhancing diffusion to wider markets. Several new institutions emerged during the transitional phase, including trade organizations that served communication among engineers, connected manufacturers to customers, established industry standards, and lobbied for road improvements and the rationalization of registration laws. I also identify one set of institutions, the Selden patent and the ALAM, that was misaligned with emerging rules in the TIEC, the redefinition of the automobile, and the needs of the mass market. Motoring clubs also emerged in the technological nexus and their role changed as the era progressed. I find that the process of embedding was underway and many of the activities responsible for the co-evolution of technologies and institutions were consistent with QE theory's technological nexus description. However, most of these adaptations were simultaneous with or lagged behind the diffusion of the automobile into wider markets. Thus, success wider markets did not depend on these processes occurring first.

In chapter 2, I asserted that a socio-technical regime exists when a technology and sector have achieved the specific phase and when the stability of that specific phase is supported by the alignment of the technology, institutions, physical infrastructure. In section 4.6.5, I summarize the role of the alignment in the success of the automobile in the mass market. A number of emerging institutions were increasing the alignment of rules and organizations but one set of institutions was misaligned with emerging rules in the industry and the TIEC. I determine that physical infrastructure was not well aligned with the mass use of automobiles at the onset of the specific phase. Thus, while stability was increasing during that phase, the automobile regime was still immature and the technology was not yet "locked in."

Section 4.6.6 summarizes the role of the Panic of 1907 and the ensuing recession in the diffusion of the gasoline automobile to wider markets. By contributing to the downturn in sales of the most expensive vehicles in 1908 and to the high failure rate of manufacturing firms in 1909, event served as a selection mechanism, reducing variety of firms and products. It also supported Henry Ford's efforts gain control of the Ford Motor Company and pursue production of a single inexpensive model, a strategy that was critical to the success of the automobile in the mass market.

Finally, section 4.6.7 summarizes the events during the time period covered by this chapter and the findings of this analysis.

4.6.1 Emerging Dominant Design and Wider Diffusion

The MLP and QE theories assert that new technologies emerge within niche markets. Within these niches, manufacturers learn about users' needs and users learn about product performance through learning-by-interacting. Manufacturers then improve the technology in terms of those needs and achieve lower costs through learning-by-doing. Consequently, growth of niche markets, or the penetration of new niches allows manufacturers to achieve economies of scale and further reduce costs. Once cost and performance are improved, the technology enters wider markets where it competes with existing technologies. This diffusion to wider markets occurs as the technology moves through what I have called the transitional phase to the specific phase. This transition is marked by the emergence of a dominant design. In the specific phase, products proliferate around this design, competing based on price and quality.

By 1910, a dominant design for gasoline motor vehicles was coalescing and the technology had begun conquering what proved to be a mass market consisting of farmers and the growing middle class. The analysis in this section shows that the movement from niches to this mass market occurred simultaneously with a functional redefinition of the automobile. Further, cost reductions were realized within the mass market. Thus, growth in niche markets did not provide the economies of scale that reduced the costs of motor vehicles and facilitated diffusion to wider markets. Further, the needs of the new market for reliability, ruggedness, high quality and low price were not discovered within the niche market, nor were product improvements relative to these needs achieved in the niche market. And because the vast, primarily rural, low-price market can not be accurately described as a market niche, niche branching, niche proliferation, or niche market growth are not accurate descriptions of the diffusion of the gasoline motor-vehicle into wider markets. In fact, niche branching occurred *after* diffusion to the mass market. The gasoline motor vehicle conquered the urban niche occupied by electric vehicles after being vastly improved for use in the mass market. Therefore, the role of niche markets in the transitional phase of the automobile is not consistent with the characterization found in the MLP and QE theories of transitions.

During the emergent phase discussed in chapter 3, individual entrepreneurs borrowed from the product and production technologies with which they were experienced, resulting in a bewildering variety of solutions to the basic engineering problem of providing motorized travel. As the industry entered the transitional phase around the turn of the century, the original function of the motor vehicle as a novelty that provided sport and prestige for the very wealthy began to yield to more practical applications. All of the motorized transport options – powered by electric motors, steam engines, or gasoline ICE – could replace existing technology, but each option was best suited to a narrow application or niche market. Thus, while some initial niches were defined by geographic preferences and infrastructure as discussed in the previous chapter, functional requirements became increasingly important during the transitional phase.

Electric vehicles were easy to operate, quiet and odorless and therefore were best suited to the urban niche market, where they served as a direct replacement for fine carriages by providing both clean practical transport and social prestige. They were especially favored by wealthy women. However, because of their limited range and the lack of refueling infrastructure beyond city limits, electric motor vehicles did not provide any new functionality. Steam-powered and gasoline internal-combustion engine motor vehicles combined the flexibility of bicycles and horse-drawn vehicles with the speed and long distance capabilities of rail to provide a new function – touring – that fulfilled the desire for entertainment, sport, and thrill, while also conveying social prestige. These vehicles could also fulfill the need for local travel in cities, but they were odiferous and extremely noisy, so were not as well suited to this purpose as the electric vehicle.

Because of its consistency with developments in the TIEC – new perceptions about personal travel wrought by the bicycle and rising individualism – flexible, rapid, long-distance travel had wide appeal and thus a large potential market. As early as 1901, *Motor World* reported that with a motor vehicle "you have a method of moving from place to place as tireless as a train, one which for short journeys and cross journeys is as quick as the train, and yet one which is individualistic and independent, hence its charm" (cited in Flink, 1970, p. 102). An article in *Harpers Weekly* in 1909 claimed that one of the chief attractions of the automobile was "the feeling of independence – the freedom from timetables, from fixed and inflexible routes, from the proximity of other human beings than one's own companions; the ability to go where and when one wills, to linger and stop where the country is beautiful and the way pleasant, or to rush through unattractive surroundings" (cited in Flink, 1970, p. 101). Such flexibility and speed allowed Americans to enrich their social experiences. Travel by horse limited social visits to a 10-12 mile radius, the round trip distance a horse could handle in a day. Gasoline- and steam-powered motor vehicles extended that range to 20-30 miles for a day trip, and much further for vacations.

While the state of steam car and electric battery technology stalled in the early 1900s, rapid improvements in gasoline ICE technology partially overcame the problems of noise and ease of operation such that gasoline-powered cars were good enough or better at serving all the niche and larger markets – short in-town trips, pleasure touring, and long distance travel. However, while manufacturers of expensive touring vehicles continued to increase the quality and performance of their vehicles, they did not seek high volume production or low prices. It was the pioneering strategy of two manufacturers to produce simple, lightweight, inexpensive, reliable, and rugged vehicles that ultimately tipped the scales in favor of the gasoline vehicle and expanded the market for these vehicles beyond the very wealthy. While the manufacturers of electric and steam vehicles continued to use craft production techniques to produce expensive, conservatively styled vehicles for the wealthy market, the inexpensive Olds curved dash and Ford Model N appealed to a wider market that included the lower-upper class consisting of physicians and merchants. For these relatively affluent buyers, the automobile was more a utilitarian product than a toy.

These inexpensive vehicles also added another important new market and function – farmers who needed a practical vehicle to transport products to the market, buy supplies in town, and make social visits. In the words of Utterback and Abernathy (1975), during the early phases of a product's life cycle, "[t]he locus for innovation is in the individual or organization that is intimately familiar with needs." Based on his own experiences, Henry Ford was intimately familiar with the farmers' needs and he targeted them when developing his inexpensive vehicles, particularly the Model T introduced in 1908. Following its introduction, Ford's strategy of progressively reducing the price of the Model T combined with rising farm incomes to make the rural market the mainstay of the automobile industry (Flink, 1988). Expansion of the rural market was reinforced by positive feedbacks – as the price of motor vehicles came down, more relatively wealthy farmers purchased them, and the rural populace became ever more familiar with and accepting of the technology.

By 1910, the bewildering variety in motor vehicle design and the accompanying customer confusion had largely disappeared. Sales of automobiles reached 181,000 units that year[53] with 52% of these vehicles selling for under $1,375. The majority of motor vehicles sold were no longer expensive toys for the very rich. The automobile had been functionally redefined as a technology to provide basic, reliable, inexpensive, personal transportation that was capable of negotiating rough rural roads and was easy and inexpensive to repair. In 1903, only 14% of vehicles sold had four-cylinder engines, while the majority of vehicles were powered by one-cylinder engines. In 1910, 84% of vehicles had four cylinders and nearly 7% had six cylinders. The four-wheeled automobile powered by a water-cooled, four-cylinder, gasoline internal combustion engine had emerged as the dominant design. This design – the functional requirements and the technological attributes that met them – was epitomized by the Model T. By 1916, sales of cars priced below $1,375 had expanded to nearly 1.3 million units, constituted 90% of the total market, and exceeded sales of higher priced vehicles by a factor of ten. The Model T alone accounted for around half of sales. Clearly, the vast and growing scale of the primarily rural low-price market can not be accurately described as a market niche.

The product redefinition that opened the large market for inexpensive automobiles directly altered the selection environment, but, combined with the saturation of the market for expensive vehicles, it also had a powerful effect on variation by reducing the total number of firms. Firms that were unwilling or unable to adapt to production of smaller lightweight vehicles were effectively weeded out. This further reduced the variety of products offered and further reinforced the new functional definition and dominant design.

---

[53] See Appendix A.

4.6.2 Process Innovations

The volume production of automobiles, beginning with Olds in 1901 and perfected by Ford around 1913, was enabled by four facets of the TIEC: 1) a broad-based acceptance of the ideals of individualism as expressed in capitalism; 2) the evolution of the American system of manufactures begun in the arms-manufacturing industry and perfected through the sewing machine and bicycle industries; 3) the disintegration of traditional craft society; and 4) mechanistic scientific approaches based in methodological reductionism that led to the development of scientific management.

The section analyzes the process innovations in this era. I find that Ford's innovative production system resulted from a highly synthetic process and relied on knowledge spillovers as well as learning-by-doing. However, Ford's greatest accomplishment was the manufacture of a product of both high quality and low cost. While one or the other had been achieved within the industry, it was the combination that allowed Ford to conquer the mass market. This combination was achieved by jointly addressing issues of product design and manufacturing processes. Thus, the transition to these new production techniques and the resulting cost reductions that enabled market expansion did not originate within niche markets, but rather as Ford was developing the rural market. Further, product and process innovations that served the needs of the large rural market for ruggedness, reliability, high quality, and low initial and repair cost arose from a more complex process than simple learning-by-doing.

The disintegration of traditional craft society was itself a manifestation of the misalignment of old and emerging rules within the TIEC. Capitalism, factory production, and mechanistic reductionism led to an increased cultural valuation of efficiency. The old rules governing the craft system were incompatible with these evolving economic rules and cultural perspectives. To improve efficiency, managers needed to take control of the production process out of the hands of skilled craftsmen. Scientific management and Fordism realized this goal and established new rules for production that were aligned with the evolving higher-level rules within the TIEC.

The companies within GM were struggling with the misalignment of these rules as late as 1914, and the continued use of craft techniques contributed to the company's financial difficulties. However, Ford's production team was uniquely able to synthesize and exploit the opportunities provided by this social and institutional context, in part because the management was relatively unfamiliar with the existing rules of production. Around 1910, Ford's production system was evolving rapidly as the team began experimenting with the addition of another innovation adapted from of the flour milling, canning, brewing, and meat packing industries: mechanical slides and conveyors to move materials through the plant. While the individuals on the team perceived the potential of each element in isolation, no one foresaw the synergistic results of their collective efforts.

Nevins and Hill (1954, p. 366) summarize mass production as the application of "simplification of design, standardization of parts, precision machining, carefully timed speed, continuous motion, and use of the most ingenious labor-saving mechanism." As such, it represents an amalgam of process and product innovations. Ford did not originate any of the process innovations: standardized parts all built to common gauge, precision manufacturing, high volume production, specialized workers performing increasingly minute tasks, control and planning by supervisors and engineers, hierarchical management, specialized machine tools, and continuously moving assembly. And except for continuously moving assembly, Ford was not the first automobile company to apply them. The company was, however, the first to fully integrate them.

In addition, Ford designed the Model T with the requirements of mass production in mind, jointly evolving the product design, machine tools, and production process. Abernathy et al. (1983) list four Ford process innovations introduced in 1907, of which two (stamped crankcase pans and integrally cast cylinder block and crankcase) require component redesign. Ford introduced five innovations in 1908, three of which (magneto integrated into flywheel, detachable cylinder heads, and vanadium steel components) involved product redesign for manufacturing. It appears that Ford was fairly unique in jointly addressing issues of process innovation and product design, and much credit should be give to the highly innovative atmosphere created at the plant.

Rather ironically, the resulting system of production included minutely prescribed procedures. Ford had begun with no rules of production and ended with a complex, detailed, and (eventually) inflexible set of rules that were embodied in standard operating procedures, routines, and capital equipment.

4.6.3 Mechanisms of Learning and Innovation

The QE and MLP theories stress the importance of learning-by-interacting with users and learning-by-doing in improving product performance and decreasing production costs. While I find that these were important mechanisms during the transitional phase, I also find that spillovers and learning by interacting with suppliers were also extremely prominent in the description of product innovations at Ford and GM. Process innovations were achieved through a highly synthetic process in which knowledge spillovers were critical. Consistent with the assertions of the industry life-cycle model, we also see the first attempts systematic research along the lines of internal research and development (IR&D) beginning with the onset of the specific phase.

As true manufacturing began around 1900, competition among Midwest manufacturers of gasoline-powered automobiles was fierce. The technology still was unproven and risky. Firms operated on slim resources, relying on credit from parts suppliers and advance deposits from dealers. Established firms that sought to expand were forced to reinvest profits into improved plant and equipment. Technology and designs changed rapidly, and failure to keep up could mean financial ruin.

Entrepreneurs focused first on product innovations to improve reliability. The detailed description we have of Ford's design efforts shows a reliance on feedback from dealers and buyers to identify defects. The role of the producer-supplier relationship is also prominent in Ford's early history. The Dodge brothers took a great risk in joining with the Ford Motor Company, foregoing all other business and essentially financing the new venture. Ford and George Holley collaborated to design a satisfactory carburetor for the Model A and Ford later bought a plant for Holley to produce carburetors for the Model T. Suppliers were constantly marketing their parts, materials, and methods to manufacturers. Metallurgist J. Kent-Smith initiated contact with Ford and likely provided the motivation to use vanadium steel in the Model T. Peter Steenstrup of Hyatt Roller Bearing pressured Ford to use the company's bearings in the Model A, and William H. Smith of the John R. Keim mills approached Ford about using stamped steel axle housings. Ford was soon using numerous stamped steel parts, and within four years had purchased the Keim plant.

With cumulative learning effects, each Ford model was an improvement of the previous ones. Ford was also learning from experience on the shop floor, working out production improvements through arrangement of men and machines and designing his vehicles to facilitate production on an ever increasing scale. By the time Ford began designing the Model T in 1906, he had a clear idea of the functional requirements of inexpensive car for the working man and the necessary knowledge to design and build it.

Weisberger (1979, p. 65) points out that, at the onset of motor vehicle manufacturing, no one had the will or the resources for "systematic research – defining a problem, accumulating data, testing materials and devices in the laboratory or on the bench, building and studying models. Trial-and-error, which is a costly long-run method of learning but requires almost no head-start capital, was what all the pioneers relied on." However, the distinction between learning-by-doing and organized research and development is extremely difficult to make. Systematic corporate research (IR&D) as we think of it first appeared at GM in 1911 when James Storrow established an engineering laboratory for mechanical and electrical testing. Charles Kettering, who would later refine systematic research at GM, also seems to have progressed systematically from problem definition to theory to prototype in developing the electric ignition and the electric self-starter between 1908 and 1912. However, the earliest efforts at systematic research were conducted by a trade association, the Mechanical Branch of the ALAM, which was formed in 1905 and established a metallurgy lab to test parts and materials. The lab ceased operation when funding was cut during the Panic of 1907.

Ford's innovative efforts in automobile design and production methods were described by contemporaries as "cut-and-try" and definitely fall more along the lines of learning-by-doing. However, these efforts were aimed at the solution of specific problems and did involve construction of mock-ups to assist in working out manufacturing issues prior to building a prototype. In addition, Wills was operating a materials testing laboratory that provided strength test data used to design vanadium steel parts for the Model T prior to its introduction in 1908. In 1910, Walter Flanders commented that starting an automobile business required the establishment of such a laboratory.

However, Ford's cut-and-try approach, undertaken by engineers with no preconceived notions of what was and was not possible, was highly effective at jointly addressing issues of process innovation and product design. In prior attempts to establish a manufacturing firm, Henry Ford had been restricted by investors' understandable focus on near-term profit. This time, Ford's lack of interest in personal wealth allowed for the reinvestment of profits in innovation, resulting in "an extensive amount of experimentation to be carried out in the factory and a surprising rate of scrapping processes and machine tools when they did not suit the immediate fancy of his production engineers" (Hounshell, 1984, p. 220).

The synthetic and heterogeneous nature of the rapid innovations after 1900 is quite clear and innovation by teams of entrepreneurs became even more prominent than during the emergent phase. David Buick and William Marr put together a technically sound motor vehicle, but it took a businessman like William Durant to make it commercially successful. Ford's product and enterprise would not have succeeded without C. Harold Wills, James Couzens, and talented vehicle and tool design teams. For the process innovations during this phase, spillovers remained prominent and were combined in a highly synthetic manner. The ability of the automobile industry to develop and to mass produce high-quality inexpensive motor vehicles relied on continuous, incremental, and cumulative technical progress in production techniques from multiple industries. The roots of mass production began prior to the emergence of the automobile with the developments for firearms, sewing machines, and bicycles. Much of the necessary innovations – machining technologies, sequential organization of production, and application of special purpose machine tools for mass production – had been introduced prior to the turn of the century and were extended by the emergent automotive industry. Then, between 1905 and 1914, Ford engineers collectively refined these techniques and combined them with innovations in the flour milling, canning, and meat packing industries, as well innovations from scientific management, to create a new production system that amounted to more than the sum of the constituent parts.

There are two other examples of the importance of spillovers between 1900 and 1915. First, Kettering's solution to the very significant problem of the automatic starter built upon his experience developing an electric motor to replace the hand crank on cash registers. Second, Walter Chrysler drew on his experience building rail cars when instituting changes in management, production processes, and materials handling at Buick in 1912. Once again, the movement of key entrepreneurs between industries served to facilitate the transfer of knowledge. This, however, did not represent the only mechanism of knowledge transfer, as entrepreneurs were increasingly networked with the knowledge base.

Most, though not all, of the early manufacturers had technical experience or training, and Rae (1959, p. 203) points out that "[t]he number of college-trained engineers in this first generation of automotive entrepreneurs is surprisingly large, considering that engineering degrees were something of a rarity in the United States until the twentieth century." According to Raff and Trajtenberg (1997), "the highest rate of quality change occurred at the very beginning [of the industry's history] (1900-1914). This is undoubtedly the portion of our period in which the greatest proportion of entrepreneurs were engineers or mechanics by training, knowledge spillovers were all-pervasive, and design bureaucracies were shallowest."

Many of these engineers were members of the relatively new American Society of Mechanical Engineers (ASME), which was founded in 1880 "to promote the art, science and practice of mechanical engineering and the allied arts and sciences" (ASME, 2008) Listed among the ASME's original objectives was the collection and diffusion of knowledge. Prior to the turn of the century, the exchange of information specific to the industry was facilitated by several enthusiast publications, including *Horseless Age* and *The Automobile*. In 1905, three industry groups were formed to foster collaboration among automobile engineers: the SAE, the mechanical branch of the ALAM, and the AMCMA, which served the independent manufacturers resisting the Selden patent. After the SAE inherited the ALAM's library and mechanical work in 1910, the SAE consisted of 899 members. The most significant contribution of these groups was the standardization of automobile parts and materials.

4.6.4 Agents and Mechanisms of Change

This section explores the continuing development of the technological nexus linking variation and selection and the ongoing mutual adaptation of the technology and institutions. I find that the technology itself was a major factor in altering the selection environment, first in a negative way as touring offended the sensibilities of the less affluent and aroused the anger of farmers. Later, the proliferation of motor vehicles that conformed to a new utilitarian product definition was a major factor in overcoming these hostilities. I identify several new institutions that emerged during the transitional phase. These included trade organizations that served communication among engineers, connected manufacturers to customers, established industry standards, and lobbied for road improvements and the rationalization of registration laws. I also identify one set of institutions, the Selden patent and the ALAM, that was misaligned with emerging rules in the TIEC, the redefinition of the automobile, and the needs of the mass market. In this timeframe, local and national motoring clubs also emerged in the technological nexus and their role changed from serving to provide entertainment to one of activism as the era progressed. I find that the process of embedding was underway and many of the activities responsible for the co-evolution of technologies and institutions were consistent with QE theory's technological nexus description. However, most of these adaptations were simultaneous with or lagged behind the diffusion of the automobile into the mass market. Thus, success in the mass market did not depend on these processes occurring first. The question of how they contributed to stability and strengthened diffusion will be addressed in section 4.6.5 and again in chapter 5.

As discussed in Chapter 3, races demonstrated the performance, endurance, and dependability of competing technologies and models. At the same time, racing tended to reinforce the definition of the automobile as an expensive toy for the very wealthy and inhibited diffusion into wider markets. Though Ford was not interested in building racing vehicles, he successfully used racing as a means to build recognition and attract financing. This use of racing altered the selection environment by directing funding toward specific entrepreneurs and projects. Ford's efforts to build racers also began building a successful team of entrepreneurs and accumulating critical knowledge.

While races reinforced the public image of motor vehicles as toys for the very wealthy, the *use* of automobiles in this application also inhibited diffusion into wider markets. The very wealthy staged elaborate festivities involving automobiles and recklessly toured the countryside, flaunting their conspicuous consumption in the face of the hard-working middle and lower class and arousing farmers' anger. However, this negative opinion was overcome around 1907 by two factors: rising farm income and the introduction of inexpensive motor vehicles that rural populations and the middle class could afford and in which they found a valuable function. The first of these were light buggy-type vehicles manufactured by Midwestern firms.

However, Olds and Ford pursued a pioneering strategy when they targeted this relatively untapped market by producing a simple, lightweight, reliable, and inexpensive vehicle in large quantities. The Olds curved dash and Ford Model N introduced in 1900 and 1906 respectively were designed specifically for this more utilitarian purpose. The Ford was superior in this respect since it departed from the buggy-style design and its inherent weaknesses that were found in the curved dash. As these vehicles became more common and as their performance was demonstrated and improved, the public's interpretation of the motor vehicle and its function shifted. Ford did not require that these users discover and communicate their needs to him; he had lived the life of a farmer and already knew what they had yet to discover: if properly designed, the motor vehicle could vastly improve their economic and social well-being. Ford had interpreted an unmet demand and imposed it on the technology.

Other activities of manufacturers through emerging industry trade organizations served the nexus role. The ALAM and the AMCMA in particular were conceived as organizations to connect manufacturers to each other and to customers by arranging for public exhibition new vehicles at trade shows; promoting races; promoting the sale of members' cars; encouraging public interest in automobiles; promoting good roads; and facilitating the exchange of technical information. In enforcing the Selden patent, the ALAM claimed to be protecting the industry from disreputable manufacturers that would have created public distrust and a negative opinion of the automobile. But by restricting licenses and controlling production output, the ALAM also sought to maintain high prices, thus affecting variation. The issue was settled by the legal system, and these higher level rules ultimately supported the independent manufacturers of the AMCMA, though this group had already disbanded. After the legal decision in 1911, the ALAM's boundary spanning role was transferred to the Automobile Board of Trade and then, in 1914, to the National Automobile Chamber of Commerce (NACC).

The ALAM's technical work was transferred to another prominent industry organization, the Society of Automobile Engineers (SAE), which served as a linkage among manufacturers and suppliers. While the SAE did not connect manufacturers to customers, it did influence both variation and selection. As a technical society, the SAE represented the shared knowledge base which served as a cognitive constraint on innovation by shaping the perceptions of engineers. But more significantly, the confusion caused by the huge number of parts used by the industry constituted an adverse selection environment which the SAE attempted to alleviate through parts standardization. While the organization succeeded at easing the parts burden and improving the quality of materials, it was unable to make the selection environment significantly more benign for small manufacturers.

Other than the Selden patent dispute, the automobile industry was fairly free to develop without governmental restrictions during the emergent and transitional phase. However, formal constraints on automobile use were emerging in the form of local and state speed and registration laws. Initially, these laws were misaligned with the capabilities of the automobile and served to limit its usefulness and therefore its diffusion. Through motoring clubs, enthusiasts joined forces with manufacturers to remove this element from the selection environment, eventually shifting efforts from opposition to coordination. In this process, the purpose of motoring clubs shifted from organizing social activities for the wealthy elite to activism for all motorists. By 1910, registration laws were more coherent and actually served to create a more benign selection environment since fees were used for road improvements.

The poor condition of roads initially had served to limit the usefulness of the automobile and was a major design consideration for inexpensive vehicles. Once again, motorists, through local clubs and the AAA, joined forces with manufacturers to promote road improvements. Some of this activism evolved from movements begun by bicycle enthusiasts. This movement tied into dissatisfaction with rail rates, particularly among farmers who relied on the railroad to transport agricultural goods to urban and export markets, and farmers joined these efforts around 1909 through the National Grange. The Good Roads movement produced modest results, but improvements were overtaken by the increasing demands of motor vehicle use. Although there was an increase in the mileage and condition of surfaced roads between 1900 and 1915, it was not likely a significant factor in the process of embedding in this timeframe. Federal funding for road improvements was initiated in 1916, but the major breakthrough occurred after WWI.

Early in the development of the automobile, critics had warned that the innovation would not find market success until the roads were improved. However, in this timeframe, construction of roads passable by motor vehicles lagged well behind the development and diffusion of the automobile itself. The capability to navigate rough roads was therefore a main design criterion for those vehicles, like the Model T, that were ultimately successful in the first decade of the 20th century. Road improvements would play a significant role in the co-evolution of the automobile regime and the wider TIEC between 1915 and 1930. Similarly, although a pre-existing dispersed fuel infrastructure was aligned with the gasoline automobile (see section 3.4.2), a dedicated refueling system tailored to the needs of the technology and its widespread use did not begin to appear until 1907 and significant growth began around 1915. Therefore, the topic of infrastructure will be taken up again in chapter 5.

By the beginning of the specific phase (1910), mutual adaptation had resulted in a dominant design that was aligned with existing physical infrastructure (rough roads and a dispersed refueling system) and supporting organizations were actively working to improve that infrastructure as well as legislative rules that created an adverse selection environment. Thus, the process of embedding was underway and many of the activities responsible for the co-evolution of technologies and institutions were consistent with QE theory's technological nexus description. However, most of these adaptations occurred during the transitional phase and were simultaneous with or lagged behind the diffusion of the automobile into the mass market. In addition, they continued and were reinvigorated with the onset of the second transitional phase covered in chapter 5.

4.6.5 Alignment and Stability

In describing the framework for analysis (see chapter 2), I asserted that a socio-technical regime is understood to exist when a technology and sector have achieved the specific phase and when the stability of that specific phase is supported by the alignment of institutions (rules and organizations) and existing physical infrastructure with the dominant product design, production processes, and organizational forms. Further, changes in the TIEC can result in a misalignment that erodes stability and initiates a new transitional phase or provides opportunities for the emergence of a new technology.

In this section, I summarize the role of alignment in the success of the Ford Motor Company and of the automobile in the mass market. I identify a number of emerging institutions that increased the alignment of rules and organizations but also one set of institutions that was misaligned with emerging rules in the industry and the TIEC. Finally, physical infrastructure was not well aligned with the mass use of automobiles at the beginning of the transitional phase but efforts were underway to improve this situation. Thus, while stability was increasing during the first specific phase, the automobile regime was still immature.

The automobile itself was aligned with and reinforced increasing individualism. Meanwhile, scientific management and Fordism were aligned with evolving higher-level rules within the TIEC regarding capitalism, mechanistic reductionism, and increased cultural valuation of efficiency. In the specific phase, the success of the automobile and

Fordism contributed to evolving economy-wide rules on the organization of production and further reinforced these trends. Related alignments supporting the first specific phase include the emergence of engineering institutions, including engineering schools and associations like the SAE and ASME. Parts standardization begun by the ALAM and completed by the SAE also contributed to the stability of the industry by improving the quality of materials and reducing the complexity of the supply environment. It did not, however, have the intended effect of bringing economies of scale to small producers. The Selden patent and the ALAM represented institutions that were misaligned with both the trend toward high volume production and the 1910 functional definition of the automobile. The patent could potentially have blocked or delayed market growth, but this barrier was removed in 1911 with the resolution of lawsuits brought by the ALAM.

The condition of roads was improving during this era through an increase in the mileage of surfaced and the application of new materials with better performance. However, this development lagged behind the diffusion of the automobile, and increased usage quickly overwhelmed road improvements during the transitional and early specific phases. As of 1916, a mechanism was in place for federal funding of road construction, but significant improvements in physical infrastructure did not occur until well into the specific phase. The breakthrough would occur after WWI. And although gasoline was widely available at country stores and other dispersed sources, growth in dedicated refueling infrastructure also began a period of significant growth after the emergence of a dominant automobile design in 1910 as will be discussed in the chapter 5.

In conclusion, the stability of the automobile industry was increasing during the first specific phase. However, it still was somewhat fragile and the automobile regime therefore was still immature. This implies that changing technological trajectory would not likely have represented a major disruption at this point; the technology was not yet "locked in."

4.6.6 Role of Exogenous Events

This section summarizes the role of the Panic of 1907 and the ensuing recession in the diffusion of the gasoline automobile in the mass market. This event likely contributed to the downturn in sales of the most expensive vehicles in 1908 and to the high failure rate of manufacturing firms in 1909. It also supported Henry Ford's efforts gain control of the Ford Motor Company and pursue production of a single inexpensive model, a strategy that was critical to the success of the automobile in the mass market.

The economic downturn and the Panic of 1907, events with no causal link to the motor vehicle sector, altered the selection environment in two significant ways. First, the Panic was coincidentally timed with the saturation of the market for expensive automobiles. It is impossible to know whether the recession played a causal role or served to hasten the stabilization of demand for these cars, but very wealthy buyers of these automobiles would have been among the hardest hit by the 50% drop in the stock market between early 1906 and late 1907. Following on the heels of slowed growth in demand for vehicles priced over $2,775 in 1906 and 1907, the downturn in sales in 1908 would have looked ominous to manufacturers. Firms in this era were operating with slim cash reserves and, for many, a slump in demand spelled financial ruin. It is significant to note that the number of firms exiting the industry peaked in 1909 (Simons, 1995). The Panic of 1907 very likely contributed to the failure of some of the firms that ceased production between 1907 and 1909 and would have influenced the strategies of the surviving firms. In therefore served as a selection mechanism that reduced both the variety of firms and the products offered.

More specifically, the Panic and recession likely contributed to the disappointing demand for the Ford Model K in 1907, Ford's first foray into the highest price bracket, as well as the mid-priced Model B. Serendipitously for Henry Ford, this event occurred as he was planning to wrest financial control of the Ford Motor Company from Alexander Malcomson. Malcomson was known as an impulsive 'plunger' who took speculative risks and was perpetually over-extended. His Ford Motor Company stock was no longer paying dividends and the value of any other of his stock holdings would have plunged with the market. Malcomson's undoing was the result of multiple factors: his own speculation, including the launch of another automobile firm; Henry Ford's maneuvering with the Ford Manufacturing Company; and the economic crisis. Thus, the Panic of 1907 was a contributing factor in providing both the justification and the means for Henry Ford to pursue his unique strategy of producing a single inexpensive model. This strategy was critical to achieving the cost reductions that made the automobile accessible to nearly every American.

4.6.7 Summary

The transitional phase of development for the automobile industry began roughly in 1900 and, by about 1910, the industry was beginning the specific phase. The number of firms in the industry, the number entering, and the number exiting had peaked by 1909 and the total number of firms began a precipitous decline in 1910. By that time, the automobile had been functionally redefined as a technology to provide basic, reliable, inexpensive, personal transportation that was capable of negotiating rough roads and traveling long distances. In 1910, just over half of the vehicles sold were priced under $1,375 and 84% utilized a four-cylinder gasoline internal combustion engine. The four-wheeled automobile powered by a water-cooled, four-cylinder, gasoline ICE emerged as the dominant design.

During the transitional phase, two main factors contributed to the gasoline internal combustion engine's selection over rival technologies and its diffusion to wider markets: 1) the ability of gasoline ICE technology to fulfill a new function with mass appeal; and 2) the strategies pursued by the entrepreneurs. Gasoline vehicles provided a new functionality – rapid, flexible, long-distance, personal travel – that ultimately appealed to a much larger market than the niche markets served by electric vehicles. Success in this market required a low price and therefore large reductions in production cost. Chapter 3 identified five developments in the TIEC that provided opportunities for entrepreneurs to begin developing a commercial motor vehicle. During the transitional phase, the Ford production team was able to synthesize four additional facets of the TIEC to develop innovations for the production of *inexpensive and reliable* automobiles:

1) a broad-based acceptance of the ideals of individualism as expressed in capitalism;
2) the evolution of the American system of manufactures begun in the arms-manufacturing industry and perfected through the sewing machine and bicycle industries;
3) the disintegration of traditional craft society; and
4) mechanistic scientific approaches based in methodological reductionism that led to the development of scientific management.

While process innovations for interchangeable parts originally had been developed to facilitate repairing firearms in the field, the Ford's team perceived that these innovations could also reduce assembly time and manufacturing costs; therefore, they were critical to high volume production.

The important mechanisms of learning during this timeframe included inter-industry spillovers, learning-by-doing, and learning-by-interacting with both users and suppliers. The electric starter was arguably the most significant improvement to gasoline ICE technology and was achieved through inter-industry spillovers as well through spillovers from electric vehicle technology. Throughout the period, the role of suppliers as innovators and collaborators is particularly striking. Meanwhile, learning-by-interacting with users played a significant role in identifying flaws and improving the performance of early motor vehicles, but was less significant in communicating users' needs in terms of functionality. Rather, Olds and Ford relied on their own perceptions and experiences with the needs of the rural market and actively targeted this new application.

Thus, I find that the role of niche markets in the diffusion of the automobile to wider markets is not consistent with the characterization found in the MLP and QE theories of transitions in the following ways:

1) Niche markets did not provide cost reductions through process improvements and economies of scale. Rather, cost reductions were realized simultaneously with a movement to wider markets. Further cost reductions were realized within this mass market.

2) The needs of the primarily rural mass market were not revealed within the niche market; Ford interpreted these needs from his own experience.
3) Product improvements relative to these needs were achieved in the mass market, not the niche market.
4) While learning-by-interacting with users and learning-by-doing were important, spillovers and learning- by-interacting with suppliers were also critical mechanisms of learning.
5) Because niche markets were saturating and the vast, primarily rural, low-price market can not be accurately described as a market niche, niche branching, niche proliferation, or niche market growth are not accurate descriptions of the diffusion of the gasoline motor-vehicle into mass markets.
6) The gasoline motor vehicle conquered the urban niche occupied by electric vehicles after being vastly improved for use in the mass market. Therefore, niche branching occurred *after* diffusion to the mass market.

The very existence of inexpensive and practical motor vehicles did much to alter the perceptions of the public, probably more than any other efforts to influence the selection environment in this era. The emergence of a dominant design provided a culturally shared definition of the technology – a common image that readily came to mind when one heard the term “automobile.” This image was increasingly symbolic, connoting freedom and rugged individualism. Thus, the automobile and the industry were aligned with wider developments in the TIEC and the growth of supporting institutions. Despite these early embedding processes, supporting institutions and infrastructure were still in their infancy. The development of coherent registration laws, roads passable by motor vehicles, and dedicated refueling infrastructure lagged behind the growth in markets for automobiles. This somewhat undermines the argument that the alignment of technology, infrastructure and institutions contributed to diffusion into mass markets. Further, while stability was increasing, it was fragile and the automobile regime therefore was still immature. This implies that the technology was not yet "locked in" and that changing technological trajectory would not likely have represented a major disruption at this point.

5.0 Stability Yields to New Challenges

To keep an industry alive, it should be kept in perpetual ferment.
-- Sorensen (1956, p. 51)

This chapter picks up the history of the industry beginning in 1916, but also includes earlier developments that became important to the industry between 1916 and the late 1920s. As discussed in chapter 4, the first specific phase began around 1910 with the emergence of a dominant design. During the first five years of this phase, the automobile was being embedded in complex systems for production and use, but the stability of this developing regime was still fragile. However, a mass market was well established by 1916, with sales of cars priced under $1,375 approaching 1.3 million units and constituting 90% of the market for automobiles. Cars priced under $675 constituted 51% of the market. The Model T was the best selling car and Ford's market share reached 50%. Henry Ford announced his strategy prior to the introduction of the Model T and proceeded to follow it to the letter: he produced a four-cylinder automobile, he stuck to that standardized design without changing it, he reached constantly toward a growing volume in order to cut costs, and he reduced prices steadily (Nevins and Hill, 1954).

This strategy meshes perfectly with the product and industry life-cycle model's description of the specific phase – standard product design, mass production, a focus on process innovations, and competition based mainly on price and quality. However, for this strategy to be successful, the industry must remain in the specific phase, which requires a static environment. But beginning around 1918, exogenous and co-evolutionary forces altered the technical, institutional, and ecological complex (TIEC) in which the automobile industry was embedded, requiring that firms adapt. In this chapter, I discuss the characteristics of innovation, mechanisms of learning, development of institutions, and the two-way interactions between variation and selection during the first specific and second transitional phases. I examine the mutual adaptation of the technology, institutions, and physical infrastructure and discuss the role of alignment and embedding, as well as unexpected events, in the stability of the developing automobile regime. I identify forces and misalignments that undermined the stability of the first specific phase and triggered a second transitional phase beginning around 1918. By 1926, the automobile had once again been functionally redefined, a new dominant design had emerged, and the industry was beginning a second specific phase.

Section 5.1 relates the history at the Ford Motor Company between 1915 and 1927. Henry Ford continued to follow strategies based on his personal rules, acquiring complete ownership and control of the company, pursuing ever deeper vertical integration and centralized control, and clinging steadfastly to the product and production processes that had been so successful in the prior era. However, the changes in the TIEC that initiated a new transitional phase around 1917 were misaligned with these rules, strategies, artifacts and processes. Adaptation at Ford lagged behind these changes in the TIEC and was costly.

Section 5.2 relates the history of the General Motors Company (GM) between 1915 and 1927. Because of William Durant's haphazard style, GM was plagued with uncoordinated management. But because of the corporation's acquisition strategy and Durant's past blunders, GM was also lucky to have managerial and technical talent in

Pierre du Pont, Alfred Sloan, and Charles Kettering. In addition, business ties to the E.I. Du Pont de Nemours Company gave GM access to that company's war surplus. With these resources, the corporation was able to respond aggressively to the challenges posed by the second transitional phase and turned crisis into entrepreneurial opportunity. By 1927, the corporation had products and systems of management and production that were aligned with the TIEC and structured to sense and adapt to future change.

Section 5.3 relates the history of the Chrysler Motor Company, a late entrant in the industry. While many small automobile manufacturing firms were unable to survive the post-war recession, Walter Chrysler was able to exploit the opportunity this presented by converting the failing Maxwell and Chalmers companies into the third largest manufacturer by 1928. In doing so, he also created an opportunity out of another unexpected event: the influenza epidemic of 1920 which took the lives of the John and Horace Dodge, leading eventually to the sale of the Dodge Brothers in 1928. Chrysler's success was due in large part to his strategy to maintain a position of technological leadership.

Sections 5.4 and 5.5 discuss two significant industry innovations, enclosed steel bodies and consumer financing, which cannot be attributed to a single manufacturer. Inexpensive closed wooden bodies and all-steel bodies were developed by suppliers and were key contributors to the shift in consumer preferences that triggered the second transitional phase. Consumer financing was largely a spillover, and represents a widespread trend in the TIEC that was strengthened by its use in the automobile industry.

Section 5.6 discusses five developments in the TIEC that provide the technical and institutional context for the period from 1915 to the late 1920s: World War I and the post-war recession; on-going efforts to improve the condition of the country's roads; development of a dedicated refueling infrastructure; market saturation and the developing used car problem; and the increasing role of women in automobile purchases.

Finally, section 5.7 synthesizes this historical review and summarizes the findings as they relate to the research questions presented in chapter 1.

5.1 Ford: Inertia and Sub-optimization

Henry Ford successfully gained managerial and financial control of the Ford Motor Company in 1907. Without the influence of Alexander Malcomson, who had owned an equal share of company stock, Ford was free to pursue his strategy of producing a single inexpensive car. The departure of the Dodge brothers from active management in 1913 further empowered Henry. He had begun to bring the supply of raw materials and parts under his control and had plans for massive increases in capacity. His keen engineering instincts and understanding of the rural market were major factors in his past business success. However, Ford's steadfast adherence to his personal beliefs and perceptions and his increasingly autocratic nature became dysfunctional as the industry matured. Henry Ford was not prepared for evolutionary change.

Section 5.1.1 relates the events culminating with Ford's acquisition of 100% of the Ford Motor Company's stock. This move was motivated by Ford's personal predilection for autocracy and was precipitated by friction between Ford and the Dodge brothers over expansion plans. By 1919, Ford had complete financial and managerial control, but had lost important human and financial capital along the way. Thus, Ford's personal beliefs, aspirations, and tactics placed the company in a precarious position just as the disruptions of the second transitional phase gripped the industry. Henry Ford's response to the initial crisis posed by the post-war recession, which is related in section 5.1.2, was shaped again by his beliefs and perspectives. While his personal rules and strategies in earlier phases had been aligned with the TIEC and the market, this was not the case in the second transitional phase. By the beginning of 1922, the Ford Motor Company was no longer in debt, but had lost critical human capital and was not prepared to adapt to the remaining challenges of the transitional phase.

Section 5.1.3 describes the extent of vertical integration achieved at the River Rouge industrial complex, which, by 1923, was a perfectly-timed and efficient industrial machine that converted raw materials into cash in 33 hours. Ford not only controlled the manufacture of nearly every vehicle part, he also secured access to raw materials and operated rail lines and fleets of cargo ships. Although the Ford Motor Company continued to rely on suppliers for some key materials, the company's exploitation of resources and transportation sent a clear signal to suppliers and shippers considering price increases.

Section 5.1.4 examines the stagnation in product design and innovation at Ford. Although improvements were made to the Model T between 1908 and 1926, Ford's incorporation of new technologies lagged behind the rest of the industry. Product innovations and road improvements eliminated many of the requirements that had dictated the Model T's design, turning many of the attributes that had been advantages into disadvantages. As consumer preferences began to shift, they became increasingly misaligned with Henry Ford's personal rules. Although user feedback was available, Ford refused to act on it, imposing his personal rules on the company.

By 1927, the Model T was no longer aligned with the rules or physical artifacts of the TIEC. Section 5.1.5 covers Ford's development of an entirely new four-cylinder car beginning in 1926. Henry Ford took charge of the project himself, continuing with "cut-and-try" methods. Developing a light, low rpm, high power engine proved to be the most significant design problem, which was solved in part by knowledge spillovers when Ford brought in an engineer from the company's aircraft division. Equipment in Ford's plants were so specialized that a complete retooling was necessary to manufacture the Model A. The changeover was costly in terms of lost production, capital outlays and the loss of experienced personnel. In the process, Ford learned that accommodating change required advanced planning in both product design and production processes. More importantly, the changeover revealed the importance of shifting from machinery designed for a single task to more flexible tools that could accommodate changes in design. The rules of production embodied in Fordism and Ford plants were misaligned with the new functional definition of the automobile which required frequent model changes. New rules in the TIEC dictated new rules of production represented by "flexible mass production."

5.1.1 Ousting the Minority Stockholders, 1915-1919

In 1912, Ford's stockholders and directors had agreed to reinvest 15% of the annual profits to expansion of productive resources, including facilities at Highland Park and branch assembly plants nationwide and beyond. By 1914, all sixty acres of the Highland Park tract were covered with Ford Motor Company buildings (Sorensen, 1956) and the plant was producing 1,200 automobiles daily (Nevins and Hill, 1954). Sometime around 1907, Ford had initiated experiments to fulfill a boyhood dream to "lift farm drudgery off flesh and blood and lay it on steel and motors" (Nevins and Hill, 1957, p. 21). Funded through Henry Ford's personal accounts, several tractor models were constructed in the following years, but the designs were unsatisfactory until 1915. Ford made plans to produce the tractor in temporary facilities in Dearborn, but also had more grandiose plans in mind. Due to space restrictions and the availability of water for the steel plant, Highland Park would not be able to accommodate expansion beyond an annual capacity of 500,000, yet Ford believed he should plan for a goal of at least 1,000,000. Developing a new site would accommodate tractor production and would allow Ford to shift some production functions, such as the foundry and power house, from Highland Park. In June, 1915, Ford announced plans for developing an industrial complex along the River Rouge to the southwest of Detroit, where he had quietly purchased nearly 2,000 acres of farm and woodland.

The Dodge brothers, who still owned a minority of Ford Motor Company stock, supported the expansion even though they were primarily interested in receiving generous dividends to help finance their independent automobile manufacturing. They apparently believed that tractor production was its primary purpose and trusted Couzens, as treasurer, to keep spending under control. However, the relationship between Ford and Couzens was deteriorating and there had been noticeable friction since 1909. Ford had always been disparaging of the business office, asserting that it was a necessary evil to be tolerated. He was irritated by Couzens' domineering manner and what he viewed as excessive economizing for the sake of money itself. Now Ford was showing interest in a larger executive role just as Couzens was becoming interested in serving in politics. Further, Couzens was vigorously pro-Ally and was annoyed by Ford's public anti-war remarks. The situation came to a head when Couzens squelched a pacifist editorial and article for the *Ford Times*, the company's organ for advertising. In the ensuing confrontation, Ford overrode Couzens authority to take such an action. On October 12, 1915, Couzens resigned as vice-president and treasurer of the Ford Motor Company.

According to Nevins and Hill (1954, p. 500), "[Ford and Couzens] had provided the inspiration which made [the company] great, and both manifested in growing degree the dictatorial and temperamental qualities which would soon mar and eventually impair its greatness. So long as one balanced the other, and John F. Dodge and others exercised considerable influence, the possibilities of harm to the company were limited." But with both John Dodge and Couzens gone from active management, Henry Ford became increasingly autocratic, making important decisions as though there were no board of directors. On more than one occasion, funds were disbursed for design and construction of facilities before the plans were ratified by or even put before the board (Nevins and Hill, 1954).

In 1916, Ford lowered prices across the board and announced that no special dividends would be paid. He believed this move would increase sales from 500,000 to 800,000 cars per year, and plans were underway to double production capacity. The Dodge brothers opposed these moves, believing the company had attained its goals. In November, 1916, the Dodge brothers filed a lawsuit to stop the diversion of company profits from dividends, thereby delaying expenditures on plant expansion. Development of both Highland Park and the River Rouge was allowed to move ahead in 1917 when Ford posted a $10-million bond to protect the minority stockholders. The courts then forced Ford to pay a special dividend of over $19 million plus interest in February, 1919.

Although Henry Ford received the bulk of this dividend, he resented the interference of the minority stockholders and resolved to rid himself of such constraints.

In December, 1918, Henry Ford resigned as president of the Ford Motor Company, claiming he wanted to devote himself to other organizations, presumably Henry Ford & Son which he had established to manufacture tractors. The board elected Henry's son Edsel to replace him as president. Then in March, 1919, Henry Ford announced he would organize a new firm, owned entirely by the Ford family, to produce a better and lower priced car than the Model T. He claimed the move had been necessitated by the settlement of the Dodge suit which, through dividend payments, had given him large personal resources that he could not reinvest in the Ford Motor Company, which was left without cash reserves for expansion. The new company would allow Ford complete freedom and control.

Needless to say, the minority stockholders of the Ford Motor Company were alarmed. Apparently at Edsel's initiation, Henry Ford's agents began arrangements to obtain their stock, representing 41.5% of the company, with the requirement that if the entire block could not be purchased then none would be. Though the stockholders knew that the agents were acting for Ford and felt the negotiated stock price was below its actual value, they were left with few alternatives. When the transaction was complete in July, 1919, Henry Ford became the first individual to control completely a company the size of the Ford Motor Company. Ford immediately reorganized the company, integrating all Ford enterprises, and allocated the new firm's stock 55.2% to himself, 41.7% to Edsel, and 3.1% to his wife Clara. Although Edsel was named company president, Henry clearly remained in control (Nevins and Hill, 1957; Flink, 1988).

The price of control was high. Many of Ford's ablest executives – C. Harold Wills, John R. Lee, and Norval Hawkins, head of Ford sales – resigned in the spring of 1919, apparently as a result of growing tensions and Ford's desire for absolute authority. William Knudsen followed in 1921 and both he and Hawkins found executive positions at GM. Wills and Lee joined forces to develop and produce a new and technologically advanced automobile, the Wills-Sainte Claire. More significantly, Ford, who despised reliance on bankers and credit, had to borrow to finance the stock buyout.

5.1.2 Financial Crisis, 1919-1921

Payments on the debt, the court ordered dividend, aggressive expansion at the River Rouge, and the acquisition of mines and timber tracts left the company in a delicate financial position. It seemed irrelevant in early 1919 when sales were booming, fueled by pent-up demand from the war years. But in the summer of 1920, the post-war recession sent sales plummeting (see section 5.6.1). Ford cut prices, knowing that he would take a loss with every car sold, but believing he would make up for it with profits from the sale of parts. By October, nearly half of the industry had followed suit. Though Ford's sales rallied temporarily, they sank even further by December. One by one, manufacturers began slowing production then shutting down. By the end of the year, Highland Park was closed. With the plant idle, Ford restructured the company, slashed employment, and sold off obsolescent and excess equipment. Prior to WWI, Ernest Kanzler had worked out schedules at the Fordson tractor plant that minimized supply inventories. As a result, parts were managed on a nearly just-in-time basis, arriving only hours before being used. Ford brought Kanzler to Highland Park to assist Edsel in a program of 'waste elimination' (Nevins and Hill, 1957).

But with payment on the debt looming in early 1922, Henry Ford figured his obligations at $58 million while the company held only $20 million in cash. Ford's financial position was precarious enough that company officials began investigating loan possibilities, though Henry Ford denied he had ever seriously considered that option. Instead, Ford took a calculated gamble and resumed limited production at Highland Park in February, insisting that dealers resume taking deliveries despite the fact that they already held considerable stocks of unsold vehicles. The move shifted the burden of Ford's debts to the dealers, most of who were forced to borrow in order to pay for the consignments (Nevins and Hill, 1957). However, company officials worked with the dealers to secure the financing required. Sales figures also showed that the situation was improving and that the dealers were generally reducing their inventories.

By April, sales were rising steadily and Ford had paid off his debts. His reorganization of the plant during the downtime helped streamline production. But while his ruthless restructuring of the business offices considerably reduced the payroll, it was essentially an inept hatchet job. Because of Henry Ford's disparaging views of the 'parasitic' office force, he "completely missed the opportunity [for an orderly reorganization] which General Motors seized and developed under Alfred P. Sloan, Jr." (Nevins and Hill, 1957, p. 167). By the end of 1921, many of Ford's most talented executives were discharged or had resigned over friction with Henry, who had become increasingly autocratic and impatient with independent individualists. Those who remained were talented and resourceful, but also completely deferential to Ford's wishes.

5.1.3 The River Rouge

From 1915 until the post-war recession hit, the price of raw materials for automobile manufacturing – steel, malleable iron, leather, and glass – more than doubled in price. Ford had long been interested in bringing the price and availability of these vital resources under his control and had begun the process when planning development of the Rouge site. Because of a contract to supply Eagle Boats for the war effort, the U.S. government in 1918 granted Ford priority materials for construction of the blast furnace which Ford had planned anyway. Construction forged ahead despite the Dodge suit (Nevins and Hill, 1957). In addition, the government widened and dredged the river and constructed a turning basin so that the Eagle Boats could be launched from the plant. Further improvements were required to allow passage of deep draft ships but were not completed until 1923.

In 1919, with complete control of the company in Henry's hands and with the war over, Ford began development of the Rouge in earnest. Body making at the site began in August and production reached 800 units per day by November. By the end of the year, the coke plant was in operation. Early in 1920, a sawmill began preparing wood for body making and the first blast furnace went into operation in May. Shortly afterward, the postwar recession hit and most construction was suspended, though work continued on setting up tractor manufacturing, which had ended in Dearborn.

After Ford's quick recovery from the recession, tractor production began at the Rouge in late February, 1921. Soon after, the foundry and power plant began operation – both were the largest of their kind in the world. By 1922, 10,000 workers manned the foundry. When complete, the power plant was capable of generating 240,000 kilowatts and 500,000 hp. In 1924, the Rouge complex covered nearly 1,000 acres, contained forty-four buildings, employed 42,000 workers, and was still growing. The blast furnaces produced 35-50% of the iron used in Model T's and tractors and the foundry produced nearly all the iron, brass, steel, and bronze castings used by Ford factories.

The Rouge realized Ford's objectives to integrate the flow of materials with key production units. A vast array of conveyors moved materials through an apparently congested plant. Traversing the floor was nearly impossible, but each stationary worker had sufficient room to perform his operation. A visitor to the plant in 1922 described the factory as a huge machine where he saw "each unit as a carefully designed gear which meshes with other gears and operates in synchronism with them, the whole forming one huge, perfectly-timed, smoothly operating industrial machine of almost unbelievable efficiency" (cited in Nevins and Hill, 1957, p. 288). The resulting processes also made remarkably efficient use of waste materials like gas from the coke ovens and ore dust from the blast furnaces. In its publicity materials, the company claimed that ore arrived at 8:00 AM on Monday, and by Tuesday afternoon, it was a finished motor *en route* by conveyor to final assembly: raw materials were converted to cash in about 33 hours.

Ford's steel making was expanded, beginning in 1925, to include open hearth furnaces, blooming mills, and rolling mills. Ford had not only become the only automobile manufacturer to have its own steel-making facilities, but beginning in 1921, was also the only one to make its own glass. The Rouge eventually manufactured coatings, lubricants, artificial leather for tops, solvents, gaskets, upholstery, and carpet. Soybean meal was used to manufacture small interior plastic components, including window strips, horn buttons, switches, and gear shifter knobs. Ford eventually manufactured rubber components as well, including tires, mountings, hoses, gaskets, mats, running board covers, and insulation (Rubenstein, 2001).

Ford also secured access to the raw materials for production. Between 1920 and 1923, he purchased over 700,000 acres of timberland, an iron mine plus an additional 2200 acres of iron ore-bearing land, coal mines, and dock facilities. By 1928, Ford's coal mining operations were large enough to sell one quarter of the output to the public. While Ford never produced a significant amount of iron ore and continued to buy from contractors about 80% of the lumber required for the company's automobiles, this exploitation of resources sent a clear signal to suppliers considering price increases. Ford later invested in a Brazilian rubber plantation but the venture was never profitable.

Ford also bought and completely remade the Detroit, Toledo, & Ironton railroad in 1920, the use of which cut shipping time and insulated the company from rail strikes. Ford purchased a fleet of barges and cargo ships in 1924 and, although costs were higher than those of rivals, the operation of the fleet was profitable. In addition, use of the fleet forced other shippers to liberalize their carrying practices and reduce rates.

5.1.4 End of the Model T

Though the Ford Motor Company was committed to making only one automobile, the Model T first introduced in 1908 was not a static design. Between 1913 and 1918, steady improvements were made in response to complaints by dealers and customers, as well as part of constant efforts to accelerate production, reduce costs, and reduce requirements for replacement parts. In 1914, the Model T was modernized to replace the bulb horn, acetylene gas headlamps and straight fenders with a hand operated Klaxon horn, brass electric headlamps powered by the magneto, and curved fenders. More style changes were made in 1917, mainly in response to consumer complaints. The more streamlined appearance and nickel trim was well received. Cadillac introduced the electric self-starter on its 1912 model and most manufacturers had adopted it within two years, but Ford held out until 1919, and then only offered it as an option. Initially, the compression ratio of the engine was actually reduced because the quality of gasoline had been steadily declining, resulting in problems with engine knock at higher compression. However, other improvements in the design and machining processes improved compression ratio, acceleration, and strength.

Model T sales climbed steadily, slowed only temporarily by World War I and the post-war recession. In 1921, sales approached one million. No longer tempered by Couzen's and the Dodge brothers, Ford's increasing tendency toward arrogance and autocratic behavior began to stifle product innovation. When the dealers began to press for changes, they were brought to Highland Park for a meeting as they had been in the previous decade. They complained that the Model T brakes were noisy and difficult to repair; brake shoes wore out quickly; and the planetary transmission was inferior to the now popular three-speed selective transmission. This time, however, their recommendations got a chilly reception. When asked his opinion at the close of the discussion, Ford replied, "I think that the only thing we need worry about is the best way to make more cars" (Nevins and Hill, 1957, p. 389).

Meanwhile, Ford's competitors were slowly closing the price gap. Manufacturers cut prices in all price categories in the summer of 1925 and, the following year, ten firms were offering 27 models of four-passenger cars priced below $1,000. Most notable were the price cuts in closed-body cars, with several now available for less than $1,000. In

1924, dealers reported that somewhere between half and all of their sales represented exchanges of open- for closed-body cars. An analysis completed in 1925 revealed that four fifths of the used vehicles traded in were open-body cars. According to Sloan (1964, p. 162), "[t]he rise of the closed body made it impossible for Mr. Ford to maintain his leading position in the low-price field, for he had frozen his policy in the Model T, and the Model T was pre-eminently an open-car design. With its light chassis, it was unsuited to the heavier closed body, and so in less than two years, the closed body made the already obsolescing design of the Model T noncompetitive as an engineering design." Although Ford did put closed bodies on the Model T beginning in 1924, they made up only 37.5% of Ford's sales. By 1927, closed-body cars represented 85% of the automobile market, but constituted only 58% of Ford's sales compared to 82% of Chevrolet's sales.[54]

New innovations were also making cars safer and more comfortable. For the 1924 model year, four-wheel brakes were offered by 25 firms as standard or optional equipment. Balloon tires that gave a smoother ride were standard on 14 models. The Model T offered neither innovation. In addition, more powerful six-cylinder engines were gaining popularity, growing from less than 20% of the market in 1917 to nearly 40% by 1926. Excluding Ford's sales, six-cylinder engines made up 55% of the market in 1926 and a few were offered at under $1,000 (Epstein, 1972). Though the Model T had been restyled in 1923 and was available in a choice of colors beginning in 1925, it remained a homely and utilitarian vehicle now competing with more stylish and comfortable cars that were also easier to operate.

[54] Nearly 80% of overall GM sales were closed-body cars as early as 1925. But it was the penetration of closed cars into the low-price field occupied by Ford and Chevrolet (and Plymouth beginning in 1928) that posed a serious challenge for the Model T.

During the 1920s, rising prosperity, availability of credit, improved roads, new technologies, the increasing market influence of women, and a saturating market fueled a shift in preferences from simple utility to style, comfort, and power (see section 5.6). Ford could scarcely have been unaware of the trend. Industry publications and company executives alike pointed out the obsolescence of the Model T. Market saturation around 1923 posed a particularly difficult problem for the aging Tin Lizzie. "When first-car buyers returned to the market for the second round, with the old car as a first payment on the new car, they were selling basic transportation and demanding something more than that in the new car. Middle-income buyers, assisted by the trade-in and installment financing, created the demand, not for basic transportation, but for progress in new cars, for comfort, convenience, power, and style" (Sloan, 1964, p. 163). With most buyers trading in and up, an increasing number of inexpensive used cars were competing with new Model T's to fill the demand for basic transportation, and many of the used cars had more options and advanced features than the Ford.

Infrastructural changes also contributed directly to the obsolescence of the Ford. By the 1920s, many roads had improved dramatically and motorists were driving at ever faster speeds. The high road clearance and stiff suspension of the Model T had been designed to navigate rugged rural roads, which it did extremely well, though not comfortably. With improved road surfaces, this functionality became unnecessary and the Ford design actually became dysfunctional. According to Hugill (1982), on smooth roads at high speed "the Model T rolled like a ship in a bad storm." And when Ford finally incorporated four-wheel brakes and balloon tires, innovations that made driving safer and more comfortable on most cars, the combination on the Model T "produced a disastrous front-end shimmy when braking at high speeds" (Hugill, 1982). In addition, the Ford's high center of gravity, necessary on rugged roads, became increasingly unsafe at higher speeds.

In 1924, sales of the Model T reached two million but had stagnated. Both sales and market share were falling by 1926 and a price reduction had no effect. Although Ford could certainly see the trends in the market, he had no tolerance for frills or conspicuous consumption and continued to believe that automobiles were and *should* be service vehicles. Ford vehemently opposed the growing emphasis on style and the trend toward annual model changes:

> It is considered good manufacturing practice, and not bad ethics, occasionally to change designs so that old models will become obsolete and new ones will have to be bought either because repair parts for the old cannot be had, or because the new model offers a new sales argument which can be used to persuade a consumer to scrap what he has and buy something new. We have been told that this is good business, that it is clever business, that the object of business ought to be to get people to buy frequently and that it is bad business to try to make anything that will last forever, because when once a man is sold he will not buy again.
>
> Our principle of business is precisely to the contrary. We cannot conceive how to serve the consumer unless we make for him something that, as far as we can provide, will last forever. We want to construct some kind of a machine that will last forever. It does not please us to have a buyer's car wear out or become obsolete. We want the man who buys one of our products never to have to buy another. We never make an improvement that renders any previous model obsolete. (Ford and Crowther, 1922, p. 148-149)

In December, 1926, Ford was quoted in the *New York Times* defending the Model T's design. "The Ford car is a tried and proved product that requires no tinkering. It has met all the conditions of transportation the world over" (cited in Hounshell, 1984, p. 277). But despite his philosophical objections, there was no escaping the direction of the market. Five months later on May 25, 1927, Ford announced the company would cease production of the Model T and build a new car. The next day, the fifteen millionth

Model T rolled off the assembly line. When production was suspended, Ford had built a total of 15,458,781 Model T's.

5.1.5 The New Model A

Although Henry Ford had repeatedly insisted that the Model T required no 'tinkering,' he recognized that eventually it would need to be replaced. One factor in his delay in introducing a new model apparently was the hope that the next Ford automobile would be as revolutionary for its day as the Model T had been in 1908 (Nevins and Hill, 1957). Since 1920, Ford engineers had experimented intermittently with design concepts for a radical "X" car utilizing an engine with eight or twelve cylinders arranged half facing upward and half downward in an X pattern. But by 1926, the car was still a dream needing years of development to become reality and Edsel believed it was too radical for public acceptance.

Sometime in mid to late 1926, Ford discontinued work on the X-car and gave the order to begin designing a new four-cylinder model. Ford took charge of developing the entirely new car, outlining design objectives of speed, power, and comfort. Integrity was to come first, and Ford wanted to replace stampings and malleable castings with more expensive steel forgings. The body would be somewhat longer than the Chevrolet and, at Edsel's insistence, set lower to the ground than the Model T. Ford no longer had an engineer with the talent of C. Harold Wills, but he had a competent and adaptable team willing to subordinate their ideas to Ford's.

However, the company had no organized engineering laboratories or research facilities and work was carried out in various locations near the production lines where it would be applied. Ford encouraged competition and more than one group of engineers often worked on the same problem. The resulting research maintained a highly practical focus, but lacked careful planning and depth. According to one employee, "No money was being spent… to get highly trained young men out of the universities and technological schools. Expenditures for laboratory equipment and general research facilities were made according to Henry Ford's mood and whim" (Nevins and Hill, 1957, p. 443). And while other manufacturers had sophisticated testing laboratories and proving grounds, Ford used public highways to test its experimental cars. Parts, materials, and even entire vehicles were tested on a pass-fail basis, and no data was recorded. Ford was still embracing a cut-and-try approach.

By the spring of 1927, the first experimental car was ready for road testing; at least twenty more would follow. The biggest problem encountered by the design team was the new engine. Ford insisted it be light, with cylinder bore and stroke only slightly larger than the Model T, and kept at relatively low revolutions-per-minute to prolong life, yet that it deliver at least 40 hp. The experimental engines produced no more than 22 hp until Ford finally brought in Harold Hicks, chief engineer of Ford's aircraft division.[55] Hicks is credited with solving the basic engine problem, though he gave credit to contributions by Edsel and to Henry Ford for his insistence on the combination of lightness and power. The new engine won national acclaim for its quick take-off and remarkable acceleration. Where the Model T produced a top speed of 43 mph, the new Ford engine achieved 65 mph.

[55] Ford had remained interested in aircraft after producing Liberty engines during the war. In 1925, he purchased the Stout Metal Airplane Company, which was making single-engine monoplanes designed by William Stout. Hicks was initially assigned to assist Stout in the development of a larger three-engine aircraft, but replaced him as chief engineer when the tri-motor performed poorly in initial tests.

The engine was the last technical issue to be solved. Edsel had directed the design of a new stylish body and distinctive interior. A selective sliding-gear transmission replaced the Model T's planetary transmission, which had been a major source of complaints. A water pump was added. For finishing touches, Ford added a safety glass windshield and hydraulic shock absorbers. The safety glass was added after Hicks crashed one of the test cars and was hurled through the windshield, badly mangling his arm. The shock absorbers were unheard of in a low-priced vehicle and did more than any other feature to provide passenger comfort. The prototype was complete in August, 1927. But Ford faced another daunting task – retooling to produce the new car.

Every one of the 5,580 parts of the new Ford Model A was entirely new; every Ford plant had to be rebuilt from the ground up. It was the largest and most urgent change-over undertaken in U.S. industry to date. Workers were brought in at Highland Park and the Rouge starting in May, and by July, 17,000 were busy with new construction, electrical work, tool design, and the complete reconfiguration of the factories. Ford once again developed and purchased extremely specialized machine tools. One notable development was a self-indexing automatic welder used to replace bolting on subassemblies; the industry would later adopt this innovation. Final assembly of the Model T had occurred at Highland Park, but for the Model A, it was set up at the Rouge.

The first factory-produced Model A came off the assembly line in late October, 1927, but Ford used this car for testing and inspection. Final changes were made in the main assembly line, and limited production began November 1. Finally, in December, the public got the first look at the new Ford. Though the car was not radical, it included up-to-date innovations, superior workmanship and excellent styling; it was available in four colors, none of them black. The anxiously awaited new Ford was an overwhelming success. But as orders came flooding in, Ford had trouble achieving full production. There were problems with plant layout, training, and the delivery of machine tools. Design refinements were required. And in the changeover, Charles Sorensen was cleaning house, purging the company of “all the Model T son-of-a-bitches” from Highland Park. The loss of seasoned production men, including Clarence Avery and William Klann, was crippling (Sorensen, 1956; Nevins and Hill, 1957; Hounshell, 1984).

Dealers were hard hit by the delays and in March, 1928, Ford established the Universal Credit Corporation to finance time-sales of Ford cars and assist authorized agents in accumulating stocks. After six months, only 100,000 Model A’s were in the hands of customers, but production was finally up. The following year proved to be a resounding success; production was the largest since 1925 and sales were strong.

The changeover to the Model A cost roughly $250 million and taught Ford some hard lessons in accommodating change (Nevins and Hill, 1957; Hounshell, 1984). First, changeover could not be accomplished smoothly without adequate advanced planning in design *and production* of the automobile. Although about 25% of the planning for retooling was accomplished in the six months before shutting down Model T production, this was not nearly sufficient. Second, Ford learned the value of establishing pilot production lines to test new approaches to machine work before tearing out old equipment. Time restrictions had only allowed for one pilot line for the Model A engine block, which proved important in identifying and solving production problems. Third, Ford learned the importance of leaving adequate time for error in planning body die work. It was particularly important for production experts to review new designs since body styles proposed by designers could involve curves too deep for drawing. Fourth, Ford personnel learned the hard way the importance of shifting from machinery designed for a single task to more flexible tools that could accommodate changes in design. For example, the presses used for the Model T were unable to handle the larger sized stampings required by a slight lengthening of the automobile's design in 1926. Later, specialized multi-spindle drill presses were replaced with ones that could accommodate a new arrangement of drills by simply changing the head. These changes to specialized machine tools led to what became known as "flexible mass production." Finally, Ford was reminded that "a new model in the initial stages of production was a sure target for hundreds of proposed changes," just as the Model T had been in 1908 (Hounshell, 1984, p. 294).

5.2 GM: Planning for Change

Between 1910 and 1915, the banker controlled management of General Motors headed by James Jackson Storrow had cut unproductive units, tightened fiscal controls, and restored the firm to solvency. In the process, Charles Nash had ended production of the Buick Model 10, thereby eliminating GM's share of the growing market for low priced vehicles. By 1915, GM's market share had slipped to 8.5%. Meanwhile, William Durant had successfully established the new Chevrolet Motor Company in the inexpensive market dominated by Ford. This section reviews the history of GM from 1915 to 1930. This era is dominated by four key individuals: William Durant, Pierre du Pont, Alfred Sloan, and Charles Kettering.

Section 5.2.1 covers the history of GM from 1915 to 1920. In 1915, William Durant used his success with Chevrolet to regain control of GM, bringing Pierre du Pont to the board of directors. GM's strategy of vertical integration through acquisitions, along with Durant's ultimately destructive actions in the stock market, led to corporate associations and the appropriation of critical human capital. The most significant of these were associations with the E.I. Du Pont de Nemours and Fisher Body Companies and the managerial and technical talents of Pierre du Pont, Alfred Sloan, and Charles Kettering. However, Durant's haphazard management combined with GM's aggressive acquisitions and optimistic expansion left the company in a vulnerable position when the brunt of the post-war recession hit the industry in 1920. That same year, Durant was forced from GM after his second disastrous attempt to prop up the value of GM stock.

Pierre du Pont reluctantly succeeded Durant, serving as president from 1920 to 1923. During this period, covered in section 5.2.2, du Pont and Sloan instituted sweeping management reforms and restructured the product lines. Durant's management had left behind an uncoordinated set of rules for the individual subsidiaries that did not serve the welfare of the corporation as a whole. The new system of management began to establish a coherent set of rules for GM that was codified in corporate structure and policies. Sloan's reforms represented a synthesis of prior managerial experience, knowledge spillovers in financial management from the Du Pont Company, and personal inspiration. The restructured product line began a process of aligning the technology with new consumer preferences. However, du Pont's attempt to compete with Ford using innovative technology failed and, as a result, du Pont stepped down as president.

Sloan succeeded du Pont, ushering in a new era for GM that is covered in section 5.2.3. The failure of du Pont's pet project, the copper-cooled engine, prompted Sloan to adopt new corporate policies that 1) coordinated the functions of and communication between the production divisions and the research labs; 2) distinguished programs for advanced product engineering from those for long-range research; 3) explicitly avoided technological leadership and its inherent risks. However, the failure, combined with Sloan's market segmentation strategy, also led to one of GM's most significant innovations in the second transitional phase – sharing parts across divisions – which allowed GM to offer a greater variety of products while achieving cost reductions through economies of scale. Sloan's increasing reliance on statistical reporting and analysis not only helped with production planning but also provided GM with a wealth of market data. This data may have contributed to GM's ability to align its products to shifting consumer preferences. By 1930, the corporation had institutionalized designing for style and was accommodating annual model changes.

Section 5.2.4 provides a more detailed history of five significant GM innovations between 1915 and 1930: the development of financial controls; the copper air-cooled engine; tetra-ethyl lead (TEL) gasoline additive; Duco paint; and institutionalizing style and annual model changes. GM institutionalized research and development in 1919, and both the copper-cooled engine and TEL were products of internal research and development. However, this review reveals that knowledge spillovers and learning-by-doing were also prominent in the development of financial controls, TEL, and Duco paint. Simple luck also played a role in the breakthroughs that led to the development of TEL and Duco paint.

5.2.1 Durant Returns to GM, 1915-1920

With the end of the GM five-year voting trust approaching in October, 1915, Durant began purchasing GM stock and arranging for voting proxies on stock owned by friends and family. On the scheduled date of the board meeting on September 16, 1915, Durant, Storrow, and Nash met privately. Durant also brought banker Louis Gaveret Kaufman, Pierre du Pont, and du Pont's chief financial assistant, John J. Raskob, to the meeting. Du Pont was one of Kaufman's directors and held GM stock in his personal portfolio. At the private meeting, it became apparent that neither Durant nor Storrow had control of a clear majority of stock. Rather than resort to an official count, the two parties negotiated a compromise. A new board of seventeen members would be elected, with each party selecting seven members. The remaining three positions would be appointed by du Pont and not connected with either party. Du Pont selected Raskob, his brother-in-law, and a du Pont corporate executive. The executives then joined the official board meeting and Storrow announced the details of the agreement. In November, the new GM board of directors elected Pierre du Pont as chairman.

However, Durant's plans were not yet complete. A week after the September GM board meeting, he reorganized the Chevrolet Motor Company in Delaware as a holding company for Chevrolet operations. The new company was capitalized to $20 million with $6.8 million in stock used to raise funds to triple production. Ninety day later, the directors of Chevrolet recapitalized again at $80 million. He then offered to trade five of his shares in Chevrolet for one in GM, a profitable trade for GM stockholders who received no dividends and watched stock prices fall under the banker regime. By the end of 1915, Durant owned around 45% of Chevrolet common stock and 44% of GM. In a desperate bid to maintain control, the pro-Storrow faction proposed a new three-year voting trust, reminding stockholders that the conservative leaders had managed to retire the $15 million debt. The proposal was doomed to failure in the face of the enormous profits to be made from Durant's stock trade.

In May, 1916, Durant announced that he controlled GM with 54.5% of the company's stock. Nash, whose name had appeared as a proposed director on the petition for a new voting trust, resigned. Durant took over as GM president in June. Storrow made plans to start his own automobile manufacturing company, intending to install Nash and Walter Chrysler in charge of the enterprise. Durant, however, enticed Chrysler to stay with GM by promoting him to general manager and president of Buick, raising his salary from $50,000 to $500,000, and promising no interference with his authority.[56] Durant reincorporated the General Motors Company as General Motors Corporation in October, and the production facilities were organized as corporate divisions.

Just months before Durant took control of GM, he formed the United Motors Corporation as a holding company for automotive parts manufacturers. To ensure that he retained control of the company, he arranged for only a small minority of stock, owned by himself and Kaufman, to have voting rights. Within months, United Motors owned five parts and accessories firms, including Hyatt Roller Bearing, Delco Corporation, and the Remy Electric Company. Hyatt's Alfred Sloan was named president of the new corporation.[57]

---

[56] Chrysler was promoted to vice president of GM in charge of operations in 1919 (Hyde, 2003).

[57] Delco, originally Dayton Engineering Laboratories Company, had been formed by Charles Kettering and Edward Deeds in 1909 to build an electric ignition system under contract to Cadillac. See section 4.2.4.

Unfortunately, fortune would not smile on Durant for long. In April, 1917, America entered World War I. Henry Leland petitioned Durant to allow Cadillac to produce Liberty aircraft engines, but Durant refused. Leland resigned in June to form the Lincoln Motor Company.[58] Within a few months, mounting public pressure forced Durant to begin token production of Liberty engines at Buick and Cadillac. Then, despite excellent profits, GM stock began an unexplained slump. Durant proposed to du Pont and Raskob that they form a buying syndicate to purchase large amounts of GM stock and thereby slow the decline in price. When they demurred, Durant began purchasing on the margin himself. As prices continued to fall, Durant found himself at the brink of ruin. In October, he approached the board and requested a $1 million loan. His actions had been intended to support GM and therefore its stockholders, so he felt his request was perfectly reasonable. Fearing legal liability, the board turned him down. However, they were concerned that public knowledge of Durant's predicament would harm the company and instead voted him an annual $500,000 salary, retroactive to 1916.

While this solved Durant's immediate personal financial problems, it did nothing for GM's stock situation. Raskob later proposed to Pierre du Pont that the Du Pont Corporation, which was holding a capital surplus from wartime profits in gunpowder, purchase large blocks of GM and Chevrolet stock. This move would solve GM's stock problem and serve to further diversify Du Pont's business, a process begun before the U.S. entered World War I. As part of the plan, Durant would have to cede financial

[58] The Lincoln Motor Company was initially devoted to building aircraft engines for the war effort. After the war ended, Leland decided to produce a new luxury car line. Lincoln was still in debt from tooling for Liberty engines and retooling for automobile manufacturing when the post-war recession struck. Due to outdated styling, sales of the new car were poor and Lincoln went into receivership. Ford purchased Lincoln in February, 1922, leaving left Henry Leland and his son, Wilfred, in charge. However, Henry Ford insisted that Lincoln be aligned with Ford philosophies and processes. Friction soon arose, and in June, Wilfred was relieved and Henry Leland resigned in protest.

control to du Pont and agree to merge GM, United Motors, and Chevrolet – a move he had repeatedly refused. In pitching the plan to the Du Pont board, Raskob pointed out the prospect of a large and closed market for Du Pont paints, varnishes, and synthetic fabrics. He went on to predict that the Du Pont Company would eventually control the entire GM enterprise.

By the end of January, 1918, a newly created Du Pont American Industries had purchased 24% of GM and a sizeable share of Chevrolet for $25 million. Most of the stock was purchased on the open market with $1.2 million purchased from Durant. Du Pont was now the second largest stockholder in GM and controlled GM's finance committee which Pierre du Pont now chaired (Weisberger, 1979). GM acquired Chevrolet in May and the United Motors Corporation in June, 1918. With the purchase of United Motors, Alfred Sloan became a director of GM and vice president in charge of its accessories division.

Du Pont doubled its investment in GM later in 1918. Over the next two years, Pierre du Pont's control of GM's finances did nothing to slow the firm's growth as it continued a veritable spree of acquisitions, purchasing all or part interest in manufacturers of automobiles, accessories, automotive parts, automobile bodies, tractors, aircraft, tires, rubber, and leather, as well as a steel plant and surplus war plants. Notable acquisitions during this timeframe included the Scripps-Booth Company, Samson Tractor, and Frigidaire, as well as shares of the Fisher Body Company, Goodyear Tire and Rubber, and Dunlop Rubber. GM secured all of Charles Kettering's time by purchasing three additional firms he had founded, the Delco Light Company, Dayton Metal Products, and the Dayton-Wright Airplane Company. With Kettering and his research operations now completely enfolded in the corporation, GM commissioned Kettering to set up and direct the General Motors Research Laboratories, which were subsequently incorporated as the General Motors Research Corporation. GM also established the General Motors Acceptance Corporation (GMAC) and the General Motors Export Division.

By the end of 1919, GM's fixed investments totaled $153.8 million, its payroll was 85,890, and it was carrying huge inventories and their attendant storage and control costs. But with the end of World War I, the future looked bright. Because automobile production had been limited during the war, there was now a shortage of automobiles for sale. The market was clamoring for every car the industry could produce and at just about any price. GM was just one of the many manufacturers scrambling to increase production to fill a backlog of orders.

However, the following year brought about abrupt changes at GM, beginning with the resignation of Walter Chrysler in March. He had given Durant the three years he had promised, and although he loved the man personally, he was unable to tolerate Durant's interference with his management of Buick (Weisberger, 1979). Meanwhile, Pierre du Pont's doubts about Durant's management skills and methods were growing. With the post-war inflation pushing up material and labor costs in late 1919 and early 1920, it also became apparent that GM needed more cash to cover operations and the optimistic expansion plans initiated in 1919. A stock offering failed to generate the needed capital, so Durant and du Pont formed buying syndicates to acquire and temporarily hold GM stock hoping to bid up the price and attract more buyers. When that proved insufficient, du Pont approached J.P. Morgan. The bank agreed to invest in GM and distribute stock to associated banks, but demanded seats on the board in return.

Within weeks of the new relationship, the price of GM stock began to slide. Then a block of 100,000 shares was dropped on the market causing a dive. Durant's sources informed him the sale had come from one of the Morgan partners who now sat on GM's board. The Federal Reserve Board had raised the discount rate in November, 1919, and speculators were liquidating over-inflated stocks. By summer, the nation was entering a recession and the price of GM stock continued its downward spiral. The 1919 boom in automobile demand had been fleeting, based on inflated wartime prices, easy credit, and over-blown expectations. Now the boom had collapsed. Sales plummeted; trainloads of cars stood undelivered. After Ford cut prices as much as 30%, GM and other competitors' sales dropped even further. GM's factory lots were clogged with inventory that tied up $210 million (Cray, 1980). Production was cancelled, factories shut down, and workers were laid off to stem the cash outflow. Yet supplies continued to arrive at warehouses that were already full. The situation was made worse by the lack of accurate inventory figures despite the improvements in coordination and management of the GM divisions since 1910 (Pound, 1934).

Amid the great optimism of 1919, Alfred Sloan had become concerned about Durant's management and the lack of any contingency plan for a drop in demand or a serious problem in production. Sloan had built the Hyatt Roller Bearing Company into an industry force by keeping pace with the expanding and varied needs of the automobile industry. In the process, he had learned invaluable lessons. From Cadillac's Henry Leland he had learned about the necessity of absolute precision and accuracy in mass production. From Ford's C. Harold Wills, who had insisted on price concessions on large volume orders, he had learned that mass production demanded a constant search for improved processes and lower costs. These imperatives required systematic planning, research, and accounting (Rae, 1959). Durant's "management by crony" was anything *but* systematic (Sloan, 1964, p. 27).

With du Pont's encouragement, Sloan had drafted three reports that detailed organizational and management reforms, including strict procedures for approval of capital expenditures, operating policies, an interdivisional billing system, and a plan for centralizing engineering and research functions. Durant had filed them away (Weisberger, 1979; Cray, 1980). Now Sloan's worst fears were materializing; it appeared that GM could actually collapse under a real market shock and that his life savings, tied up in corporate stock, would disappear with it. Banker and former GM president James Storrow had offered him a partnership in the Boston firm of Lee, Higginson & Company and Sloan took a month-long vacation to consider it. He decided he should resign and gradually sell his GM stock for whatever he could get. But he apparently cut his vacation short to return to GM, finding the company in the midst of big changes and decided to wait.

The drop in GM's stock price had put Durant's personal finances in immediate peril once again. As prices fell, he had continued to buy on the margin. Brokers essentially loaned him 90% of the stock price. They then borrowed the balance from bankers, using the stock as collateral. When the stock price fell, so did the value of the collateral, and the banks asked for either more collateral or repayment. Durant had dealt on such a scale that even he did not know how much he owed. In early November, Durant was finally forced to admit his predicament to Pierre du Pont. His assessment of his debts reached $38 million, enough to ruin GM, forty-four brokers, and an unknown number of banks (Cray, 1980). With the fate of GM once again in the balance, du Pont and J.P. Morgan arranged, partially through borrowing, to buy 60% of Durant's stock and clear his debt (Weisberger, 1979). The remainder of his stock was purchased by a newly created holding company that would hold it until prices recovered. Though Durant owned a minority share in the holding company, he did not directly own any GM stock and could not serve as president. As the news spread that GM control was now completely in du Pont and Morgan hands, the stock price began to recover. On December 1, 1920, Durant again left GM; this time he would not return.

5.2.2 Du Pont Control, 1920-1923

Pierre du Pont continued as chair of GM's board of directors and now reluctantly succeeded Durant as president. With the country in a deep recession, it was an unfortunate time to begin a career in the automotive industry. Sales early in 1920 resulted in a net profit for the industry, but by year end, even Ford had closed the Highland Park and River Rouge plants. Though du Pont would only serve as president for two and a half years, the era would see dramatic changes in GM's finances, operations, organizational structure, and product lines. Du Pont attributed much of these changes to GM's new vice president in charge of operations, Alfred Sloan, who had forwarded to du Pont copies of his lengthy reports on administrative reforms.

5.2.2.1 Establishing Coordinated De-centralized Control

At the end of his first week as president, du Pont reviewed Sloan's plan for organization and in late December put it before the Executive Committee with only minor modifications. The committee unanimously approved it and sent it to the board of directors, which also approved it (Sloan, 1964; Cray, 1980). In essence, Sloan believed that GM was too large to operate under Durant's decentralized, 'feudal' system with *laissez faire* doctrines or under Du Pont Company's rigidly centralized management. The new organizational structure developed by Sloan has been compared to military staff-and-line division, but with a system of coordination that resembles government organizations; local managers retained administrative jurisdiction over and responsibility for their divisions (Pound, 1934). Of the plan's origins, Sloan recalled in 1964 (p. 47):

> I cannot, of course, say for sure how much of my thought on management came from contacts with my associates. Ideas, I imagine, are seldom, if ever, wholly original, but so far as I am aware, this study came out of my experience in Hyatt, United Motors, and General Motors. I had not been much of a book reader, and if I had been, I understand that I would not have found much in that line in those days to help; and I had no military experience.

Nonetheless, Sloan's model would eventually become the gospel for business administration students, be copied by nearly every major U.S. corporation, and adapted by foreign companies.

While the central office dictated business policy, division managers were free to develop tactics consistent with corporate strategy. They controlled manufacturing, distribution, and sales; hired their own staff; and made their own purchasing decisions. But du Pont and Sloan established interdivisional committees, staffed with representatives from the manufacturing divisions, to integrate policies regarding products, sales, purchasing and advertising. The committee staff, selected by Sloan, controlled the decision-making process by controlling the information passed down. Committee recommendations went to du Pont for approval, though Sloan gradually took on this responsibility. Many division managers were unhappy with the reorganization and resigned; others were reassigned. Sloan and du Pont then promoted men from within the corporation to replace them, with Sloan selecting men with engineering backgrounds.

Many of the committees were temporary in nature, utilized only until problems were under control and new organizations and policies were in place. Of particular note, Sloan (1964) felt that the General Purchasing Committee could not be deemed an unqualified success for three reasons. First, many parts were already required in large enough quantities by a single division to realize supplier economies of scale and, therefore, low prices. Second, some parts were specific to a single engineering concept and not applicable to other projects. Third, vendors that lost a bid sometimes approached one of the divisions directly afterward, offering a better price, and thereby causing confusion and dissatisfaction with the process. However, the committee successfully demonstrated the feasibility of cooperation among the divisions, proving that representatives could work together for the good of their divisions and for the profit of the stockholders simultaneously. Most significantly, though, the committee was instrumental in beginning a process of materials and parts standardization across the divisions. By 1926, the evolution of this concept would allow GM to challenge Ford's domination of the market.

While the committee structure facilitated coordination, the corporation required a set of financial controls that could reduce the need to administer operations from the top. Forever in search of hard data to support decision-making, Sloan increasingly turned to two financial experts, Donaldson Brown and Albert Bradley. According to Cray (1980, p. 200), Brown and Bradley were "among the first of a new generation of financial specialists, men who cared little about machines and manufacture and even less about such intangibles as excitement or the romance of the industry. Trained as economists and accountants, they were hard-eyed, coldly dispassionate champions of financial analysis and, through analysis, control." Brown had been an assistant to Raskob at Du Pont and succeeded him as treasurer when Raskob was transferred to GM. While serving in this post, Brown introduced the then unusual practice of hiring economists and statisticians to assist in monitoring divisional efficiency and performance. At Raskob's request in January, 1921, Brown was hired as GM's vice president in charge of finance. Sloan and Brown "shared similar views on the value of detailed, disciplined controls in the operation of a business" (Sloan, 1964, p. 118). Brown soon tapped the statistical talents of Bradley, a young associate who had joined GM's financial staff in 1919. Sloan, Brown, and Bradley developed financial instruments to control appropriations, cash, inventory, and production. According to Sloan (1964), Brown and Bradley were largely the architects of the specific forms of these controls.

Prior to 1920, each division had controlled its own finances. Although the financial pressures of that year pushed GM into debt, some divisions had more cash than was required. However, division managers concerned with their own operations were unwilling to transfer money, forcing the corporation to borrow. In fact, the GM executives had no way of knowing just how much cash was held by each division at any given time. Further, transferring cash involved an inter-city transfer by mail that required time. In 1922, GM created a consolidated cash-control system using the newly created Federal Reserve System (see section 4.5.1). Local depository accounts with fixed minimum and maximum balances were established across the country. Deposits were credited to GM and withdrawals were administered by a central financial staff. When balances exceeded the preset maximum, funds were transferred to one of several central reservoir banks. If a division required cash, it applied to headquarters using a requisition system and funds could then be transferred by telegraph within a few hours.

GM attacked the worst of its emergency problems – the inventory – by halting delivery of materials. Suppliers were persuaded to delay billing, essentially financing the company as they had in the early days of the industry. To reduce finished product inventories, du Pont eventually was forced to follow the lead of Ford and other manufacturers and cut prices. The divisions gradually used up their excess inventories, and by June, 1921, Buick, Cadillac, and Oldsmobile were in good shape. Obviously, inventory controls were required to prevent similar problems in the future. Since supply shortages had been common prior to 1920, plant managers placed orders in excess of expected needs resulting in an oversupply when business slowed. New policies restricted purchasing materials to monthly budgets prepared by the division managers based on a four-month forecast of sales. These forecasts were later expanded to include plant investments, working capital, and outstanding product inventory commitments. As vice president in charge of operations, Sloan reviewed these forecasts to approve or modify production schedules. However, division managers had no mechanism for determining the actual inventories held in dealers' hands and could provide only estimates of retail sales and market conditions. Early in 1923, the R.L. Polk & Company began providing the industry with statistics on automobile registrations.

5.2.2.2 Improving and Rationalizing the Product Lines

The 1920 market slump also uncovered engineering deficiencies in some GM automobile designs. As long as demand was in excess of supply, cars of reasonable quality sold themselves and manufacturers were free to concentrate on increasing production. As soon as demand slackened, inferior mechanical performance quickly affected sales and reputation. Buick and Cadillac continued to be profitable, but Chevrolet, Oldsmobile, and Oakland were in need of overhauling. In early 1921, du Pont appointed Sloan to chair a committee tasked with studying the corporation's products and recommending revisions. He also hired consultants to provide an independent evaluation of GM automobiles. As part of the assessment, Du Pont insisted that GM must compete directly with Ford in the low-priced market.

The independent review concluded that the Chevrolet products and reputation were so poor that the line should be terminated. Since Oakland and Oldsmobile were not much better off, they recommended GM change the names of all three lines. Sloan argued that doing so would not improve the models offered and that starting over with no reputation would be harder than improving the current one. In addition, terminating a line would abandon and alienate current owners. Sloan recommended eliminating two newer divisions that had no product recognition nor established dealer organizations: the two-year-old Sheridan and unreleased Scripps-Booth. He further recommended rationalizing the remaining product lines which overlapped in some price categories, resulting in competition with each other, yet left gaps in the pricing structure.

Du Pont agreed to postpone a decision on restructuring the GM lines. However, he insisted that Sloan include in his committee report a plan for using a new air-cooled engine in a vehicle that would compete with Ford. Kettering had been working on the engine since 1918 and Du Pont believed it was the bold stroke that would lure customers away from Ford (Cray, 1980). The air-cooled engine was not a new innovation; Franklin had introduced it in 1898 (Abernathy et al., 1983) and used it in its expensive automobiles for twenty years. Kettering's design, the development of which is detailed in section 5.2.4.2, employed copper fins to conduct heat away from the cylinders and there were many manufacturing problems yet to be solved. Nonetheless, Sloan was forced to incorporate Kettering's copper-cooled engine in new designs for a four-cylinder Chevrolet and a six-cylinder Oakland.

Sloan's plan for rationalizing the GM products limited the number of lines to six price categories within which the divisions would not overlap in price nor compete with one another. Sloan's proposed policy would make GM automobiles competitive, not by *leading* in design or innovation, but by being technologically *equal* to competitors' best models. Sloan proposed that GM cars should be of slightly higher quality and priced just above the competition in each price category. With this strategy, GM would compete with its competitor in a price category based on quality. But it would also compete with its competitor in the next higher bracket based on price, since the GM car would be of just slightly lower quality but priced low enough to appear a good value. Chevrolet would fill the lowest price category and compete with Ford. The remaining divisions in order of increasing price would be Oakland, Buick, Oldsmobile, and Cadillac, with Buick filling two price categories.

Two features of Sloan's strategy are noteworthy. First, Sloan was advocating that GM *not* compete with Ford head-on in the lowest price category, a move that Sloan believed would be "suicidal." Ford essentially had a lock on the first-time buyer with limited financial means (Rubenstein, 2001). Rather, "the strategy we devised was to take a bite from the top of his position" (Sloan, 1964, p. 69), a strategy that was definitely at odds with du Pont's mandate. Though there is no documentation of the origin of Sloan's strategy, an expert industry analysis published in 1913 reported that "at least 15,000 buyers every year would have preferred a car of more stylistic distinction if they could have gotten it for only $100 above the Ford prices" (Parlin and Youker Report, cited in Nevins and Hill, 1954, p. 509). This was particularly important in a market with fewer first-time buyers and more consumers who were trading in and up.

Second, by structuring its product line within classes distinguished by price and quality, GM offered a product that was available to members of every social class. Further, each nameplate was an immediately recognizable statement of the owner's social status. Chevrolet was for first-time buyers and the working man. Oakland (renamed Pontiac in 1926) offered more style and flair for a bit more money. Oldsmobile touted superior engineering. Buick appealed to the upper-middle class – doctors, lawyers, and managers. Cadillac offered luxury and superior craftsmanship for the wealthy owner (Rubenstein, 2001). By trading up from a Chevrolet to an Oakland or Buick, the middle-class owner demonstrated his progress up the social ladder. In marketing a "car for every purse and purpose," GM turned a single mass market for low priced vehicles into a group of niche markets defined by socio-economic status as well as functional requirements. Within each niche, competition was based on price, quality, and options, and GM sought to place itself near the top of each niche in all three.

By 1923, all divisions except for Chevrolet offered new models. The Chevrolet design was seven years old and in need of replacement, but the new design was awaiting the development of the copper-cooled engine. Performance of the new engine in a production vehicle was still questionable, and the managers of the other GM divisions had passed on it. Chevrolet's division manager protested as well, stressing himself to the point of a nervous breakdown. The job of leading Chevrolet passed to its new vice president in charge of operations – William Knudsen. Knudsen had only been with GM one month, having quit the Ford Motor Company in early 1921. Knudsen assured du Pont that he could solve production problems with the copper-cooled engine in time for the 1923 season, but by January had turned out only half of the promised volume. Around 100 of the new vehicles were sold that spring, but by June it was clear the vehicles were faulty and Knudsen recalled the cars. With the costly failure of his pet project, du Pont believed he had become too personally vested in the technology and failed to listen to the misgivings of Sloan, the division managers, and the production engineers (Cray, 1980). Du Pont had successfully led the company through its financial crisis – GM had cleared its debt by June 30, 1922, and the automobile market had rebounded. In May, 1923, Pierre du Pont resigned as president of GM and recommended that Sloan succeed him; the board concurred. Du Pont retained his position as chairman of the board of directors.

5.2.3 The Sloan Era, 1923-1930

Soon after succeeding Pierre du Pont as president, Alfred Sloan cancelled all production plans using the copper-cooled engine, which prompted Kettering to resign in discouragement. Though the move distressed both Kettering and du Pont, Sloan was unwilling to force new innovations on the divisions and to hold back programs while awaiting uncertain developments (Sloan, 1964). Yet Kettering's work was inestimably valuable – he was reportedly close to finding a solution to the problem of engine knock – so Sloan negotiated a compromise with Kettering and the Executive Committee. Sloan believed that the main mistake with the copper-cooled engine project had been the poorly coordinated division of responsibilities among the Executive Committee, the operating divisions, and the Research Corporation. Therefore, GM established an independent pilot operation, under Kettering's sole discretion, to develop, manufacture, and market the six-cylinder copper-cooled engine and automobile. Chevrolet would continue to work out problems with the four-cylinder model. A reluctant Kettering agreed to stay on. After two additional years of development, the project eventually faded away.

The failure of the copper-cooled engine, which is discussed in detail in section 5.2.4.2, was a significant factor in shaping the future strategy and organization of GM. According to Flink (1988, p. 233),

> The copper-cooled engine was doomed by a lack of coordination and cooperation on the project among GM units. Production problems were the inevitable result of a lack of communication between Kettering's… laboratory, where the engine was designed, and the factories… where the car that the engine powered was to be produced. Design problems arose also, because it had not been foreseen that numerous other components would have to be redesigned to accommodate to the light weight of the new engine.

Sloan (1964, p. 70) characterized the problem as "one of conflict between the research organization and the producing divisions, and of a parallel conflict between the top management of the corporation and the divisional management." In 1964 (p.94), he recalled that the failure "taught us about the value of organized cooperation and coordination in engineering and other matters. It showed the need to make an effective distinction between divisional and corporate functions in engineering, and also between advanced product engineering and long-range research." At the same time, Sloan recognized that long-range research of the type undertaken in Kettering's lab would only be of benefit to the corporation and stockholders if the manufacturing divisions were able to exploit the results of that research. To facilitate coordination and cooperation on engineering matters among the divisions as well as between the divisions and the labs, Sloan established a General Technical Committee consisting of chief divisional engineers and top research staff in the Research Corporation.

While the experience with the copper-cooled engine may have catalyzed Sloan to clarify functional distinctions and improve relationships and coordination within the firm, it also encouraged him to adopt a strategy that tended to inhibit technological innovation. As early as 1921, Sloan (1964, p. 64) asserted that "the primary object of the corporation… was to make money, not just to make motor cars… The problem was to design a product line that would make money." Under du Pont, GM had sought to "meet Ford more or less head on with a revolutionary car design." Sloan instead had proposed policies to "produce a line of cars in each price area, from the lowest price up to one for a strictly high-grade quantity-production car." GM would not produce any vehicles in small quantities; the price categories would not leave wide gaps in the product line; and there would be no duplication by the divisions in the price categories. The idea was to secure the greatest advantage that quantity production could provide. Sloan (1964, p. 66, emphasis in original) asserted that this policy "was valid if our cars were at least equal in design to the best of our competitors in a grade, so that it was not *necessary* to lead in design or to run the risk of untried experiments. Certainly I preferred this concept to an irrevocable commitment to replace the then standard Chevrolet with a revolutionary car." The same concept carried through to production efficiency, advertising, selling, and servicing – none need be better than that of GM's best competitor. According to Cray (1980, p. 558), Sloan insisted that "GM is a production company, not a research-oriented company."

Under Sloan, GM pushed forward with the redesign of a new Chevrolet model. A booming market in 1923 had resulted in decent sales, but a brief recession in 1924 underscored the need to redesign Chevrolet's outdated model. While U.S. automobile sales dropped 12%, GM's sales fell 28% and represented nearly half of the industry's decline. Half of GM's decline was attributed to Chevrolet, where sales dropped 37%. Rival Ford's sales slipped only 4% (Sloan, 1964). Meanwhile, Dodge, Chrysler, and Hudson had increased sales and market shares by introducing new models and affordable closed-body designs (Freeland, 2000). In 1925, Chevrolet introduced the new K Model which corrected a number of defects with the previous Chevrolet. The successful model did not include radical changes, but among the improvements boasted a Duco lacquer finish that had been introduced on GM's Oakland the previous year. The K Model was also available in two closed styles, using Fisher bodies, which sold for 40-57% more than the roadster. By 1927, closed bodies constituted 82% of Chevrolet's production and 85% of industry-wide production.

Consistent with his new product strategy, Sloan also moved to fill a gap in GM's lineup between the low-priced Chevrolet and the more expensive Olds. The entirely new model, which was to incorporate a new six-cylinder engine, was likely to take some sales from the four-cylinder Chevrolet as well as the more expensive Olds, but at least those sales would stay within the corporation. In order to minimize the impact on Chevrolet's economies of scale, Sloan decided the new model should use as many Chevrolet body and chassis parts as would fit the new design. As it turned out, the failed copper-cooled engine project had given GM some valuable experience with dual-purpose chassis and bodies for two different engines. Although the new GM model would be manufactured and sold by Oakland, Sloan assigned its initial development to Chevrolet due to its use of Chevrolet parts. The introduction of the new Oakland Pontiac in 1926 marked the first time that GM physically coordinated the manufacture of vehicles made by different divisions; this innovation had sweeping implications for the company. GM finally had a competitive response to Fordism:

> Physical co-ordination in one form or another is, of course, the first principle of mass production, but at that time it was widely supposed, from the example of the Model T, that mass production on a grand scale required a uniform product. The Pontiac, co-ordinated in part with a car in another price class, was to demonstrate that mass production of automobiles could be reconciled with variety in product… If the cars in the higher-price classes could benefit from the volume economies of the lower-price classes, the advantages of mass production could be extended to the whole car line. (Sloan, 1964, p. 158)

While working out improvements to the GM product lines, Sloan also continued to integrate and coordinate the operation of the divisions, relying again on financial analysts Donaldson Brown and Albert Bradley as discussed in detail in section 5.2.4. Sloan was once again forced to address inventory control when a brief recession in mid-1924 led to an oversupply in dealer stock and forced GM to drastically but belatedly cut production. It was clear that GM needed a more effective way to forecast retail sales and to quickly respond to market changes. A comprehensive study of the total automobile market begun by Bradley in 1923 was put toward establishing estimates of consumer demand in each price category. This estimate was then used to set production levels for the coming year, allowing for mid-course corrections based on dealers' reports of sales and inventories. Rather than stay committed to a four-month production schedule determined in advance, GM could now respond quickly to market conditions using up-to-date actual data.

The organizational structure put in place by Sloan required decentralization with coordinated control. GM could only leave the execution of operational procedures to the divisions if corporate management had a means to monitor and assess the effectiveness of these operations. For Sloan and his analysts, the key was another financial control based on rate of return. Section 5.2.4.1 provides a detailed discussion of these innovations, but a brief summary is presented here. Accounting procedures were first standardized across the divisions. Brown then developed standards for sales and manufacturing expenses using past performance adjusted for future plans. Monthly reports of actual operating results, based on uniform return on investment analysis, were used to determine each division's competitive position in the corporation and to troubleshoot when a division's performance deviated from expectations.

However, comparisons with past performance were complicated by year-to-year variations in production volume. In 1925, Brown developed a new concept called the 'standard volume' based on the average expected production over a number of years and a long-term return on investment goal. Unit costs and prices were developed using the standard volume, allowing year-to-year cost comparisons that were unaffected by fluctuating volume. According to Flink (1988, p. 234), GM's conservative estimate of demand used to establish the standard volume ensured "high rates of return in a market assumed to be both saturated and technologically mature."

With record sales in 1926, GM began to reap the rewards of its reorganization and the re-invigoration of its product line. While Ford clung to the inexpensive, utilitarian Model T, GM marketed fifty models of cars and trucks ranging in price from $525 to $4,485 – "a car for every purse and purpose." In addition, Kettering's lab had found a solution to the engine knock problem and profits from the growing market for Ethyl gasoline, sold in partnership with Du Pont, helped close the gap with Ford (Cray, 1980) (see section 5.2.4.3). However, GM still faced three challenges in the post-war era: 1) market saturation; 2) the growing used car problem; and 3) shifting consumer tastes for comfort and style. The corporation's differentiated product line and consumer financing (see section 5.5) helped with these issues, but GM devised one more innovation that jointly addressed these challenges and propelled the company to the industry forefront – the annual model change. Again, a brief summary of this innovation is presented here, while a more detailed description is given in section 5.2.4.5.

Since the birth of the industry, style had been forced to adapt to mechanical and mass production requirements. Body and chassis design were largely two separate processes, resulting in an awkward fit and appearance. Closed-body cars were especially tall and ungainly and the typical automobile's high center of gravity was becoming increasingly dangerous. In 1926, Sloan and Lawrence Fisher hired Harley J. Earl, who designed custom car bodies for Hollywood celebrities, as a consultant for the design of a new GM automobile. The new La Salle, introduced in 1927, was the first mass produced car designed by stylists. Impressed with the result, Sloan hired Earl to direct a newly established Art and Color Section and institutionalized the process of designing for style.

GM had no policy for regular model changes, which were undertaken as needed in response to the market and competition. Because of the new emphasis on style and the need to motivate owners to trade up, GM's constant upgrading resulted in the introduction of annual models beginning in 1923. By the 1930s, GM had recognized this unavoidable fact and was planning for regularized change. Much of GM's ability to accommodate change is attributed to William Knudsen, who was appointed vice president of Chevrolet in 1922 and promoted to president and general manager of Chevrolet in 1923. Knudsen built a decentralized organization that could accommodate change and expansion. Although Knudsen arranged machinery and operations for sequential production as at Ford, he replaced single-purpose machine tools with standard or general-purpose machines that could be adapted to accommodate product design changes. Highland Park had been extremely efficient, but also inflexible; in contrast, Chevrolet production had been designed to accommodate change and marked the beginning of flexible mass production.

5.2.4 Significant Innovations

5.2.4.1 Financial Controls

Despite the 1921 improvements in inventory control at GM designed by Donaldson Brown and Albert Bradley, a demand slump in mid-1924 led to an oversupply in dealer stock and forced GM to drastically but belatedly cut production. It was clear that GM and the industry as a whole needed a more effective way to forecast retail sales and, more importantly, a way to quickly respond to market changes. In 1923, Sloan had tasked Bradley with a comprehensive study of the total automobile market using the available sales data and Bradley's concept of a "pyramid of demand" (Sloan, 1964, p. 139). The intent was to determine the potential market by class, the effect of a price reduction on total demand, the competitive relationship of new and used cars, and when the market would reach saturation – the point when most new car purchasers were replacing their old automobile rather than buying their first car. This study had, for the first time, demonstrated the relationship between demand and income distribution.[59] This statistical work was now put toward developing a scientifically based procedure for determining production schedules.

In 1924, GM established official corporate-wide estimates of consumer demand in each price category based on actual sales over the previous three years and a general business outlook. This estimate was then used to set production levels for the coming year, though mid-course corrections could be made, within tooling constraints, as updated information became available. To improve the reliability of these estimates and facilitate corrections, GM had dealers report new and used cars sales and inventories every ten days. Division managers compared the reports to their monthly forecasts and adjusted purchases and production as necessary; at the end of the month, the reports were used to update forecasts for the following month. Rather than stay committed to a four-month production schedule determined in advance, GM could now respond quickly to market conditions using up-to-date actual data.

Sloan described GM's fundamental management problem as finding the key to decentralization with coordinated control. With new procedures in place for controlling operations, GM could leave the execution of these procedures to the divisions – but only if corporate management had a means to monitor and assess the effectiveness of these operations. For Sloan and his analysts, the key was another financial control based on rate of return. Sloan had first applied the principle of return on investment while at United Motors, which was his first experience managing multiple divisions that

[59] Bradley later correlated automobile sales and overall economic activity, finding that when national income was rising, car sales increased at an even faster rate; when national income fell, sales decreased at a faster rate. When better statistics were available, he was able to show a close correlation between sales and personal income (Sloan, 1964).

separately manufactured diverse products. "By placing each division on its own profit-making basis, I gave the general office a common measure of efficiency with which to judge the contribution of each division to the whole… [T]he general principle of rate of return as the measure of worth of a business… was fundamental in my thinking about management problems" (Sloan, 1964, p. 47-50). This principle had governed the thinking of du Pont's people and, therefore, GM's Finance Committee since 1918.

Brown developed a method for computing return on investment that could be used to evaluate investment decisions and the performance of the divisions. The idea was not to maximize the short-term return on capital, but rather to achieve the "highest return consistent with attainable volume in the market. The long-term rate of return was to be the highest expectation consistent with a sound growth of the business" (Sloan, 1964, p. 141). Brown developed standards of performance for business and manufacturing expenses using past performance with adjustments for future plans. Actual performance was then compared to these standards. To facilitate inter-divisional comparisons, accounting practices were standardized throughout the corporation. Each division completed monthly reports of operating results, and was informed of its competitive position in the corporation based on uniform return on investment analysis. These reports were used to troubleshoot when a division's performance deviated from expectations.

However, production volume varied from year to year, making comparison to past performance difficult. In 1925, Brown developed a new standard for overall performance which related a long-term return on investment goal and the average expected production or average plant utilization over a number of years – the 'standard volume.' Unit costs and prices were developed using the standard volume, allowing year-to-year cost comparisons that were unaffected by fluctuating volume. With prices stabilized in this way, when demand and production exceeded the standard volume, GM realized windfall profits. Of course, when sales fell short of the standard volume, profits fell below the rate of return goal. According to Flink (1988, p. 234), GM used a conservative estimate of demand to establish the standard volume and reinvested only a conservative portion of actual profits in expansion, thereby "guaranteeing the safety of invested capital and ensuring high rates of return in a market assumed to be both saturated and technologically mature."

5.2.4.2 Copper Air-Cooled Engine

The first air-cooled engine was introduced by Franklin in 1898 (Abernathy et al., 1980) and offered on the company's expensive automobiles. Air-cooling theoretically offers several advantages over water cooling: lighter weight and therefore higher fuel economy for the same horsepower; simpler design and construction by eliminating the radiator, hoses and ducts; and elimination of problems from freezing in winter and overheating in summer. Air-cooled engines of the day utilized fins cast on the exterior of the cylinders. The fins were cooled by air pushed over the engine by the movement of the car. Although this worked fine for small engines, larger ones overheated, resulting in 'burned' valves, sticking pistons, and a loss of power (Leslie, 1979; Flink, 1988). Therefore, by 1910, most automobile engines were water-cooled.

In 1918, Kettering began working on what Flink in 1988 (p. 232) called "the last attempt by an American automobile manufacturer to pioneer to the stage of production a truly radical engine design." In an attempt to solve the overheating problem in air-cooled engines, Kettering experimented with using copper, which has superior thermal properties compared to iron. If successful, the engine would realize a higher compression ratio and therefore achieve greater fuel economy and power for the same displacement. Since copper was too soft and too expensive for use in constructing the engine cylinders, the fins had to be attached to the iron cylinders, a difficult proposition. Kettering's team at Delco experimented with casting the copper fins directly onto the cylinders, attaching the fins with molten zinc, and brazing them on with a torch. All failed until 1921 when a member of the team developed a special electrical furnace that could successfully do the brazing. After working out production methods for the cylinders and fins, Kettering's team constructed a four-cylinder prototype engine and tested it in a Chevrolet chassis using a fan to boost the air speed over the fins. This experiment apparently was only marginally successful (Leslie, 1979).

With the drop in automobile sales in 1920, Pierre du Pont became convinced that it was critical for GM to enter the lowest-priced market. He believed that Kettering's new engine was the key to producing a lightweight, inexpensive vehicle that could compete directly with Ford's Model T. Ford had cut the price of the Model T in September, 1920, and sales had picked up enough for Ford to resume production. Meanwhile GM was still working off inventories, hoping that sales would rebound with warmer weather in spring. Just as Kettering began making headway on attaching the copper fins to the cylinders in January, 1921, du Pont and GM's Executive Committee announced their intent to develop the air-cooled engine for use in a low-priced car made by Chevrolet. The GM executives also decided to develop a six-cylinder air-cooled engine that could be used in a more expensive model produced by Oakland, whose sales had faltered.

The manufacturing division managers were eager for new designs that would boost their sales and initially appeared supportive of plans for the new models. However, when Chevrolet manager K.W. Zimmerschied was informed in February, 1921, that his division would begin manufacture of the air-cooled model in August, he objected. He had already made improvements in Chevrolet's water-cooled engine and had designed a new body. He wanted to postpone production of the air-cooled engine another year.

By May, Kettering had test cars of both the four- and six-cylinder models in operation. The Executive Committee authorized the creation of a small pilot manufacturing section, capable of producing up to twenty-five cars per day, at Kettering's laboratory facility in Dayton. The research lab thus had been given responsibility for the development and the initial production of the new car, but the manufacturing divisions were responsible for mass production. Zimmerschied was skeptical about the engineering design and wanted to know who was advisor to whom on issues of production. Meanwhile, Kettering worried that the divisions would alter his designs.

Kettering assembled a fleet of test cars, including several competitors' cars with conventional engines, and set out for a road test in July, 1921. Although the road conditions were grueling and none of the cars survived intact, Kettering was satisfied with the performance of his engine (Leslie, 1979). Because Chevrolet still had inventory to work off and because Oakland's manager, George Hannum, was more receptive than Zimmerschied, the Executive Committee decided Oakland would receive the first new model and Chevrolet's four-cylinder model would follow in early 1922. Production of Oakland's water-cooled car was to cease in December, 1921, and production of the air-cooled car would begin in February, 1922. In November, Oakland received a car for final testing – the first tests performed by personnel outside of the research staff. The results surprised the Executive Committee: Hannum informed du Pont that the car was not ready for production. He believed another six months of development was required to make the necessary changes to eliminate problems with overheating. He also informed du Pont that Oakland would bring out a new water-cooled line in December to bridge the time until the air-cooled model was ready. The Executive Committee remained committed to the new engine, but decided to delay production of the air-cooled Oakland until at least June, 1923.

Kettering believed that the division managers were resisting his innovation for other than technical reasons, perhaps feeling the technology had been forced on them because of Kettering's close relationship with du Pont. Nonetheless, he resolved to address the issues revealed by the Oakland tests. His team focused as much on problems with the frame, axle, and transmission as those with the engine itself (Leslie, 1979). The pressure was now on developing the four-cylinder model for Chevrolet. Meanwhile, Alfred Sloan decided to prepare a second line of defense by establishing a parallel effort in the Chevrolet division to improve the existing water-cooled design. The team responsible for manufacturing both cars was now in direct competition with Kettering's research team.

In March, 1922, du Pont replaced Zimmerschied as general manager of Chevrolet and assigned Knudsen as vice president of operations at Chevrolet. Knudsen recommended that GM immediately put the air-cooled engine into small level production for technical and commercial testing. Production of the Chevrolet four-cylinder was scheduled for September. The next trial in May, 1922, was encouraging and Kettering pushed for immediate production. But sales of traditional water-cooled cars finally were on the rise and division managers openly questioned the wisdom of deploying a new technology under such conditions. Kettering confided to du Pont that he believed the divisions were trying to thwart his innovation. Sloan argued for caution in deploying a new innovation during a time of peak sales. September came, but production was not underway. The executive management remained committed to the new engine but decided that production of the water-cooled Chevrolet should continue and that the new air-cooled model would be sold as an option. The new model was to be unveiled at the New York automobile show in January, 1923.

Knudsen reassured du Pont that he would have 500 of the new engines ready for the start of production in 1923, but he was able to produce only 250 in time for the New York automobile show where the new model was "the sensation of the show" (Sloan, 1964, p. 85). Spring sales were booming, but production of the new vehicle soon encountered problems of technical malfunctions and miscommunication. Chevrolet produced 759 of the automobiles, but 239 were scrapped by the production team. Of the 500 delivered to the sales organization, only 100 were sold to retail customers with the remainder either driven by factory representatives or held in inventory. Complaints poured in from field representatives, dealers, and customers. The engine overheated and lost power, even in cool weather. The design was partly at fault: the air entered the engine housing from the bottom and was expelled at the top. But as it moved over the engine, it warmed before ever reaching the cylinders which were the hottest part of the engine. Customers also complained about excess noise, clutch problems, wear on cylinders, carburetor malfunctions, axle break downs, and fan-belt trouble (Leslie, 1979).

In addition to design flaws, many of the problems were the result of miscommunication, poor coordination, and shoddy manufacturing or assembly, though the division engineers complained that it was impossible to "produce the car within the limits set by Dayton, and rightly or wrongly demonstrated by their sample car" (Leslie, 1979). While Kettering remained convinced that there was organized resistance against his work, the divisions felt that Kettering's actions were the result of an over-inflated ego and not in the best interests of the corporation. Leslie (1979) places much of the blame on Kettering, claiming that his insistence that only research engineers were capable of solving technical problems had alienated the production engineers. In addition, his team "grossly underestimated the difficulty of converting a prototype into a reliable, mass-produced article ready for sale to the public."

Once the extent of the problem became clear, Knudsen moved quickly, recalling all of the automobiles. Du Pont apparently took responsibility for the failure, believing he had become too personally vested in the engine's development. Although du Pont had successfully led the corporation through the post-war recession, it was time to step aside. He passed the presidency to Sloan who had consistently advised caution in putting the new engine in production cars. Sloan cancelled development of the six-cylinder copper-cooled engine for Oakland and Olds and Knudsen at Chevrolet soon cancelled production of the four-cylinder model. Without a clear market demand for the new engine and du Pont's championship of the project, the fate of the copper-cooled engine was sealed.

Disappointed in the abrupt termination of production, Kettering announced his resignation, prompting Sloan to seek a solution that would retain the inventor at GM.[60] Sloan proposed creating a new pilot operation under Kettering to develop, manufacture, and market the air-cooled automobiles. Pierre du Pont, still serving as chairman of the board, insisted the four-cylinder model had to be sold through Chevrolet. The Executive Committee negotiated a compromise, giving Knudsen and Chevrolet responsibility for development and engineering of the four-cylinder model. Kettering's team in Dayton was given complete responsibility for redesigning the six-cylinder copper cooled engine for the existing Oldsmobile chassis. If successful, a new division would be created for small scale production and sale prior to going into high volume production (Chandler and Salsbury, 1971). Kettering accepted the compromise, but not without complaint. "Accounting always kills research," he complained (Cray, 1980, p. 218).

Disappointing tests of both copper-cooled models in the spring of 1924 showed neither was ready for production. With some $3.5 million in direct costs invested in the project, du Pont and the Executive Committee were reluctant to terminate the project. A year later, the six-cylinder version still was not ready, but Chevrolet reported all major problems had been solved with the four-cylinder model. But by this time, the successful new water-cooled Chevrolet K Model was competing well with the Ford Model T. The Executive Committee did not even discuss Chevrolet's report and the copper-cooled engine project quietly faded away.[61]

---

[60] Leslie (1979) claims it is doubtful that Kettering would have followed through with his resignation. He had a large financial investment in GM and his research team was making headway on projects of great interest to him, especially the engine knock problem they would eventually solve using tetra-ethyl lead. In addition, Kettering had a history of lashing out verbally but not fulfilling his angry threats.

[61] Because of their inherent simplicity and light weight, air-cooled engines are particularly well-suited for and frequently used in motorcycles. They are also used in most aviation piston engines where weight is a

Sloan believed the failure of the copper-cooled project was in large part due to confusion of responsibilities and poor coordination. In September, 1923, he established a General Technical Committee consisting of chief divisional engineers and top research staff to coordinate the activities of the divisions and the Research Corporation. According to Sloan (1964, p. 109), the General Technical Committee "had a remarkable effect in stimulating interest and action everywhere in the corporation in matters of product appeal and product improvement, and produced a free exchange of new and progressive ideas and experience among division engineers."

The Committee was highly independent and had the authority to undertake technical studies using the facilities of the Research Corporation, the operating divisions, or outside sources. Over the course of time, the Committee studied numerous short-range engineering problems that concerned brakes, fuel consumption, lubrication, rust and oil sludge caused by the condensation of the products of combustion, and changes in steering mechanisms required for four-wheel brakes and balloon tires. The Committee even gave attention to educating dealers and sales staff on the value of engineering developments. Kettering's staff reported to the Committee on long-range studies concerning cylinder-wall temperatures, cylinder heads, sleeve-valve engines, intake manifolds, tetraethyl lead as a gasoline additive, and transmissions.

In summary, the copper-cooled engine was developed through IR&D. The project was begun by Kettering largely because of his belief that more efficient engines were important to conserving fuel and avoiding shortages or depletion. GM's executive

---

premium and the much higher airspeed facilitates cooling. In the 1930s, German engineer Ferdinand Porsche developed a simple air-cooled engine for the first Volkswagon. It was used in the VW Type 1 (Beetle) and related models, including the Porsche 911. GM abandoned the air-cooled engine until 1960 when it introduced the Chevrolet Corvair. However, none of these engine designs use copper fins.

support of the project, however, was stimulated by an external economic condition – the post-war recession – and by intra-industry market dynamics – competition with Ford's Model T in the low priced field. During development of the vehicle, economic conditions and demand for automobiles improved. In addition, market tastes were shifting – buyers were becoming interested in more style and comfort than the Model T offered. The new technology was no longer 'demanded' by the market and therefore no longer 'needed' as part of GM's market strategy.

Also, though the research lab's findings show that the technology was ready for production, it appears the product was rushed to market before design and manufacturing issues were fully worked out. Where Ford had found great success by jointly addressing issues of product design and manufacturing processes for the Model T, GM had failed by separating the two for the copper-cooled engine vehicle. Had the few cars that were sold to retail customers performed successfully, the market outcome may have been different. But with no apparent need for the new technology and no product champion at the executive level, GM chose to minimize its risk and table the project. However, the experience with the copper-cooled engine was responsible for changes in corporate policy and organization that significantly shaped the innovative trajectory of the company for the next fifty years.

5.2.4.3 Tetra-Ethyl Lead Gasoline

In 1913, Cadillac engineers discovered a violent banging in the cylinders which they blamed on Kettering's new electric self-starter and battery system. At Delco, Kettering quickly dismissed the problem after determining that it was not caused by faulty firing of the spark plugs. Kettering later became involved in a broad range of innovative efforts for three companies he founded: Delco Light Company, Dayton Metal Products, and the Dayton-Wright Airplane Company. While working on a farm lighting system that used a small internal combustion engine to generate electricity, Kettering once again encountered engine knock and found that it occurred even more violently when using kerosene. Kettering assigned the problem to his assistant, Thomas Midgely, in 1916 (Cray, 1980).

Kettering's focus shifted to aviation fuels when he began working on aircraft engines around the beginning of World War I. He began experimenting with alternatives to gasoline altogether, finding that some fuels could be used in high compression engines while others produced a significant amount of knock. Kettering and Midgely also experimented with a number of "high-percentage" fuel blends using up to 50% of an alternative fuel that produced no knocking, such as benzene from coal or olefins. Most of the additives had technical issues to overcome such as a relatively high freezing point or the tendency to 'gum.' With the end of World War I, Kettering's focus returned to the automotive industry.

As demand grew for more powerful automobiles, the industry began using higher compression engines and engineers found that the knock problem increased. Engine knock reduced the energy efficiency and therefore the power of the engine and also could crack the piston. While loss of engine power certainly bothered engineers and consumers, the efficiency issue was a more pressing concern. Fears of oil shortages are nearly as old as the automobile industry, and in 1919, geologists estimated that only 20 or 30 years-worth of oil were left in the U.S. After World War I, demand for gasoline grew so rapidly that suppliers had trouble keeping up with demand and the quality of fuel declined (Kovarik, 1994). Lower quality fuels increased the prevalence of knock and further reduced engine efficiency.

Many industry experts advocated alternative fuels, especially ethanol, as a solution to the supply problem. Ethanol had the added benefit of not producing engine knock even at high compression. However, the lower energy density of ethanol required vehicles to carry more fuel and thus more weight, which was particularly problematic for aircraft. In addition, experts concluded that there was insufficient farmland to produce both food and fuel. Although experiments were underway to convert cellulose from trees and agricultural wastes into ethanol, the conversion process was too inefficient. Prohibition also complicated the development of an ethanol fuel infrastructure and even hindered experimentation. Nonetheless, Kettering apparently was an advocate of ethanol as a long-term solution to fuel supply concerns. He believed more efficient gasoline engines would extend the life of oil supplies and could be converted later to use ethanol after the fuel production technology was more mature. Increased efficiency would then help compensate for the lower energy density of ethanol. This approach required high compression engines and a solution to the knock problem (Kovarik, 1994).

The prevailing view in the industry was that engine knock was caused by premature ignition of the fuel when the spark was too far advanced. However, Kettering and Midgely had proven that pre-ignition and knocking were two separate problems and they believed that engine knock was due to properties of the fuel itself. In addition to high percentage fuel blends, they also experimented with low concentrations of additives or so-called doping compounds. One of Kettering and Midgely's early theories was that knocking arose from the heat absorption properties of the fuel. Perhaps the lower volatility of kerosene caused it to remain in droplets in the cylinder then vaporize very suddenly after combustion had begun, leading to increased knocking. They decided to try darkening the fuel to enhance heat absorption.[62] Unable to locate any oil-soluble dye, Midgely tried adding a few drops of iodine to the carburetor and found that knock was diminished. When he later used a number of dyes, there was no effect, yet a colorless iodine compound was effective. Midgely concluded that the effectiveness of the iodine was related to its chemical properties. Since iodine was corrosive and expensive, Midgely tried similar elements: bromine, tellurium, tin and selenium. The results pointed toward heavier elements. In December, 1921, after trying perhaps thousands of compounds, he finally used tetraethyl lead (TEL) in the test engine and the knocking was silenced. In 1923, Midgely was awarded the William H. Nichols Medal from the New York section of the American Chemical Society for his discovery of the anti-knock properties of TEL.

Although TEL was the most efficient additive Midgely had tried, it was difficult to make and broke down quickly in sunlight. Adding it to gasoline resulted in lead deposits on the engine valves and spark plugs, while lead particles corroded the valve seats and exhaust system. Additional experimentation solved these problems and the additive was ready for commercialization in 1923. GM expected to capture 20% of the fuel market and estimated production and distribution costs at only one cent per gallon of treated gasoline. With a price differential of three cents per gallon, GM stood to realize an annual profit of around $24 million based on the existing fuel market.

[62] The literature provides two conflicting stories of the source of this inspiration. According to Pound (1936), Delco chemists noticed that ordinary combustion produced a bluish flame while the flame for combustion with knocking was yellow or orange. Kovarik (1994) reports that Kettering recalled that the trailing arbutus sometimes bloomed in the snow and theorized that the flower's red color might help it absorb more heat.

But by 1923, Midgely and three other lab employees were suffering from lead poisoning. Despite his illness and the warnings of several of the world's leading health experts at MIT, Harvard, and Yale, Midgely insisted that there was no risk to the public and that there would not be enough lead in automobile exhaust to cause any problems (Kovarik, 1994). Amid a growing controversy over the safety of the additive, GM and Standard Oil of New Jersey jointly founded the Ethyl Gasoline Corporation in 1924 to market gasoline doped with TEL, which they gave the brand name Ethyl. Kettering was named president and Midgely vice president.

In 1923 and 1924, three separate operations were set up to produce the lead solution. A pilot plant set up in Dayton manufactured TEL and shipped it in liter bottles to prospective customers who mixed it into gasoline at their service stations. The Dayton plant was shut down in April, 1924, after two workers on the packing line died. Midgely claimed that the workers had failed to follow safety precautions and that the liquid was safe when handled properly. The Du Pont Company was contracted to manufacture the lead solution beginning in August, 1923, and by January, 1925, Du Pont was manufacturing 1,700 gallons per day. Eight workers died at the Du Pont facility between the time it opened and the winter of 1925. However, demand for Ethyl was booming, and GM had been pushing Du Pont to speed up production. Standard Oil claimed to have a less expensive process than Du Pont's and, with Sloan's approval, set up a small plant in New Jersey in September, 1924. In late October, five workers went violently insane and within eight days had died; at least thirty others were treated for acute lead poisoning and some suffered permanent brain damage (Kovarik, 1994).

It is not entirely clear whether the New Jersey incident was solely the result of an industrial accident or explosion, as typically recounted in automotive histories, or due at least in part to the fact that Standard's production process was inherently more hazardous than Du Pont's, which had already claimed five lives before the new plant opened. Concerned over those losses, Du Pont was developing a closed system that would limit workers' exposure to the chemicals and fumes. Meanwhile, Standard's process involved higher temperatures and pressures and workers were directly exposed to concentrated fumes and process wastes. On touring the facility in September, Du Pont engineers were reportedly shocked by the dangers posed by Standard's equipment and methods as well as the inadequacy of safety precautions. When Kettering and Midgely urged Du Pont to adopt Standard's process, Du Pont's technical director refused (Kovarik, 1994).

After the October incident, Standard Oil's production plant was shut down, but the courts found the company innocent of criminal negligence in February, 1925. However, media coverage of the incident led to a public panic over acute poisoning from the use of 'loony gas' and several states moved to ban its use. GM continued to stress that, while manufacturing pure TEL was indeed dangerous, there was no risk to the public because the additive was diluted to one part in 1,300 in gasoline. In a presentation to the American Chemical Society, Midgely also asserted that no alternatives to Ethyl existed – a claim that was contrary to Midgely's own research. Meanwhile, health experts continued to weigh in on the danger of chronic exposure to lead in automobile exhaust. As the controversy grew, Kettering and Midgely were secretly removed as president and vice president of the Ethyl Gasoline Corporation (Kovarik, 1994).

In May, 1925, a conference on the safety of TEL was held by the federal Public Health Service (PHS). While the PHS had no authority to regulate chemicals, it had taken an aggressive stance against TEL. As a result of the conference, which lasted only one day, the Surgeon General was tasked with forming an expert panel to investigate the dangers of leaded gasoline. The Ethyl Corporation suspended sales of the additive pending the outcome of the investigation. The study examined the health of 252 men, comparing drivers and mechanics exposed to leaded gasoline to those who had not, as well as to workers in lead industries. Despite finding that the exposed workers' blood and stools contained high levels of lead, the panel's report in January, 1926, was inconclusive, citing the incompleteness of data regarding the effect of longer exposures on human health.

Vindicated by the report, Ethyl Corporation resumed sales of leaded gasoline in the spring of 1926 after agreeing to put warning labels on it. In 1927, Chrysler introduced the first moderately-priced automobile with a high-compression engine, made possible by higher octane fuels like Ethyl that reduced knock. By 1940, TEL was used in 90% of the gasoline in the U.S (Kovarik, 1994). The Ethyl Corporation recorded profits of $1 billion between 1924 and 1947, when its patents expired. GM's share of the profits totaled $82.6 million, while patent royalties netted $43.3 million. Du Pont realized profits of $86 million (Cray, 1980).

5.2.4.4 Duco Paint

Finishing methods for automobile bodies were inherited from the carriage industry. As many as three dozen coats of paint and varnish were required, with sanding and drying in between. Body finishing was a major production bottleneck since the total process took from two to eight weeks, depending on a number of conditions, including temperature and humidity. Dark pigments tended to dry faster, leading to the use of very dark greens and blues, and of course, black (Nieuwenhuis and Wells, 2007). The all-steel Budd bodies used on Dodge cars beginning in 1914 could be oven-baked, facilitating a shift to black enamel.

In 1921, GM appointed a Paint and Enamel committee to consider the problem of finishing car bodies. In GM's assessment of the problem, the industry absolutely needed a product that would reduce finishing time from days to hours without requiring the use of high temperatures beyond what wood could stand. The product needed to be inexpensive, capable of producing a wide range of colors, and able to last the life of the vehicle. In short, the industry needed a finish that would provide all the advantages of both varnish and enamel with none of the disadvantages of either (Pound, 1934). On the other end of the spectrum from paints and varnishes, the toy industry was using cellulose nitrate lacquers, but they dried too fast. The paint and varnish industry claimed the task was impossible.

The answer to the auto industry's finish problem came about largely by accident in a Du Pont laboratory, but traces logically from the firm's origins. In the 1830s, French chemists discovered that nitric acid mixed with starch or wood fibers produced a lightweight combustible explosive material broadly called nitrocellulose. Unfortunately, the material was unstable and therefore not practical as an explosive. Some ten years later, a German-Swiss chemist cleaned up a nitric acid spill with a cotton apron which he then hung on the stove door to dry. But once dry, the apron exploded; the chemist had serendipitously discovered what came to be known as guncotton, which was also discovered by another German chemist in the same year. Refinements in the manufacturing process improved the stability of guncotton, which was used as a propellant and a low-order explosive.

Guncotton was briefly used in artillery, but was too powerful and unstable for use in firearms. In 1884, Frenchman Paul Vieille invented a smokeless gunpowder by dissolving guncotton in ether or alcohol. The gelatinous solution was formed into thin sheets and cut into small flakes to produce a less volatile propellant that only exploded when compressed. The powder was more powerful than black powder, producing a higher muzzle velocity and therefore more accurate long-range firing with no pall of black smoke. Du Pont developed its own smokeless gunpowder in the 1890s and became a leading manufacturer the explosive for U.S. military use.

Other experiments discovered that films made of nitrocellulose combined with ether or alcohol would harden into a film, leading to the development of plastics, lacquers and photographic film.[63] The Du Pont management began a program of diversification beginning in 1902 and was particularly interested in alternative uses for nitrocellulose. In 1904, Du Pont purchased the International Smokeless Powder and Chemical Company which produced nitrocellulose lacquers. In 1910, they purchased the Fabrikoid Company, which produced artificial leather made from fabric covered with a nitrocellulose film. Du Pont improved the product and successfully marketed it to the automobile industry for tops for open cars. In 1915, Du Pont purchased the Arlington Company which manufactured pyroxylin (nitrocellulose) plastics used in combs, collars, cuffs, and automobile side curtains.

---

[63] Though not explosive, this film was still highly flammable.

In 1920, a Du Pont chemist working with nitrocellulose film noted an accidental chemical reaction that led to the creation of a thick, quick-drying lacquer capable of holding more pigment than before. Du Pont initially marketed the lacquer under the name Viscolac in 1921, but subsequently worked with General Motors Research Corporation to refine the product for the automotive industry (Du Pont, 2007). With promising progress in the labs, Pierre du Pont petitioned the Du Pont board to give GM an exclusive license to the new product, but the request was denied. The new lacquer was renamed Duco and was first used on the "True Blue" Oakland for model year 1924. Du Pont made the new finish available to the entire automobile market in 1925.

The introduction of Duco was a major step forward and was soon used by all major auto manufacturers. The quick-drying, inexpensive lacquer relieved a major bottleneck in mass producing cars. Even inexpensive cars were now available in a wide variety of rich, deep colors, opening a new era in automotive styling. The finish did not fade in sunlight or weather, nor was it damaged by the heat of the engine. And dents could be repaired and the metal repainted with matching color. However, there still were issues to be resolved with the lacquer. Duco did not produce a lustrous finish, so owners had to wax the car frequently to maintain a glossy shine. But the most significant problem was that the lacquer was not sufficiently adhesive and would sometimes strip away from metal in large sheets. The Du Pont and GM research laboratories continued to collaborate on undercoats to improve adhesion. In addition, the development of synthetic chemicals later eliminated the reliance on natural resins that were limited in quantity and varied in quality (Sloan, 1964).

5.2.4.5 Style, Annual Models, and Planning for Change

The post-war era presented the automobile industry with three challenges: 1) market saturation; 2) the growing used car problem; and 3) shifting consumer preferences for comfort and style. To stimulate sales, manufacturers relied on consumer credit and began focusing on dealer organizations, marketing, and national advertising. The issue of comfort was being partially met by the production of more closed-body cars, which rose from 17% of the market in 1920 to 72% by 1926 (Epstein, 1972). Style, however, had always been subordinate to engineering. Market saturation meant that most customers were now purchasing their second vehicle. If automakers wanted to maintain industry growth, owners would have to be enticed to trade up rather than wait until the end of the useful life of their current automobile. Manufacturers needed to engender consumer dissatisfaction.

Since the birth of the industry, performance and economics had dominated the design process and both style and comfort had been a secondary issue – style had been forced to adapt to mechanical and mass production requirements. After thirty years of development, the product and production technologies were no longer limiting factors in styling, which allowed for a new emphasis on aesthetics. At this point, body and chassis design were largely two separate processes, resulting in an awkward fit and appearance. The chassis design included the fenders, running boards, and hood, and was still based on open body styling. The separately designed closed body fit on top of the frame and was therefore narrower than the chassis and quite tall, resulting in an ungainly appearance. To make matters worse, road and engine improvements led to faster driving, and the typical automobile's high center of gravity was becoming increasingly dangerous.

GM's Alfred Sloan had long believed that a lower body was visually more appealing and, in 1926, he decided to approach automobile styling more systematically. Lawrence Fisher of the Fisher Body Company, now serving as general manager of Cadillac, shared Sloan's sensibilities on styling. Fisher visited dealers around the country and was impressed by the styling of custom bodies being produced by California dealer Don Lee for Hollywood stars. Fisher toured Lee's custom body shop where he met the shop's young director and chief designer, Harley J. Earl. Earl, who had studied at Stanford and trained in his father's carriage shop, was using novel approaches to body design. Instead of making models of various automobile components using wood and hammered metal, he designed the complete automobile as an integral whole using modeling clay. As a result, the long, low custom bodies that Earl designed blended together the elements of the body, hood, fenders, headlights, and running boards.

In 1926, Sloan and Fisher hired Earl as a consultant for the design of a new GM automobile to be sold by the Cadillac Division but at a lower price than the current Cadillac model. The new La Salle, introduced in 1927, was the first mass produced car designed by stylists. Sloan (1964, p. 269) recalled, "The La Salle looked longer and lower; the 'Flying Wing' fenders were drawn deeper than their predecessors; side windows had been reproportioned; the belt line had a new type of molding; sharp corners had been rounded off, and other design details were added giving it the unified appearance that we were looking for." Since there was no time to redesign the existing Cadillac model that year, Earl suggested using color to dress it up. Where the Cadillac had previously been offered in three colors, the 1927 model was announced with five hundred color and upholstery combinations.

Impressed with the results at Cadillac, Sloan hired Earl to direct a newly established Art and Color Section funded through the Fisher Body Division. To facilitate the operating divisions' acceptance of the new section, Sloan and Fisher leant their personal support to Earl. The first automobile completely redesigned by the Art and Color Section was the 1928 Buick. The new model was a commercial failure, in part due to style changes made for production reasons, but also because the public saw it as too radical a departure from conventional styling. Afterward, the group instituted successive style changes that allowed for an evolution along a projected line of development. The new group gradually gained acceptance as the market proved that appearance sold cars.

As of the mid-1920s, the annual model change was not a stated policy of GM or any automaker. Manufacturers had always sought to make some improvements in quality or performance every year. Major design and style changes generally were instituted as needed in response to the market and competition. Sloan and the GM division sales managers debated the idea of annual models versus continual improvements in a 1925 sales committee meeting. While no one in the industry wanted annual models, GM was finding it necessary to make more frequent changes. These changes were not regularized and the timing of them was becoming problematic from both production and marketing standpoints. In truth, because of the need to motivate owners to trade up and the new emphasis on style, GM's constant upgrading resulted in the introduction of annual models beginning in 1923. By the 1930s, GM had recognized this unavoidable fact and was planning for regularized change. The annual model change entailed a major style redesign every three years, a cycle that meshed with die life to minimizing retooling costs. Minor annual cosmetic changes were made in between. Although Sloan and Earl did not invent planned obsolescence, they were the first to institutionalize it (Sloan, 1964; Cray, 1980; Flink, 1988).

Much of GM's ability to accommodate annual model changes is attributed to William Knudsen. Knudsen was a Keim Mills employee when Ford purchased the company in 1911 and brought him to Highland Park. Later in the decade, Knudsen repeatedly tried to impress upon Henry Ford that the tired Model T had run its course (Cray, 1980). Tensions rose between the two individualists as Ford became increasingly autocratic and rigid in regard to the Model T's design. Knudsen could accept Ford's interference with his decisions, but he resented Ford's methods of countermanding his orders or telling employees to ignore them. In 1921, Knudsen resigned to avoid further confrontation with Ford, saying "I can't avoid it if I stay, and I can't stay and keep my self-respect" (Nevins and Hill, 1957, p. 168).

GM hired Knudsen in 1922 with no particular role in mind for him. Amidst problems with Kettering's new copper-cooled engine, Pierre du Pont named Knudsen vice president in charge of operations at Chevrolet. Du Pont planned for the new Chevrolet model to compete directly with Ford in the low-price car market. In 1924, Knudsen was promoted to president and general manager of the Chevrolet Division. Although the copper-cooled engine failed in the market, Knudsen, under Alfred Sloan's management, was able to turn Chevrolet around. The division lost nearly $9 million in 1921, but by 1932, Chevrolet's profits were largely responsible for keeping the corporation out of the red.

Rather than imitate the production system at Ford, Knudsen built a decentralized organization that could accommodate change and expansion. In 1924, Chevrolet produced 280,000 cars; Chevrolet models were re-styled each year and by 1928, production exceeded 1 million and forced an end to the Model T. Although Knudsen arranged machinery and operations for sequential production as at Ford, he replaced single-purpose machine tools with standard or general-purpose machines. He expanded production facilities and added new assembly plants, each run independently by a local manager. A Fisher Body plant was then attached to each assembly facility so that body production was closely coordinated with the daily output of each assembly plant. Highland Park had been extremely efficient, but also inflexible, because “every machine tool and fixture was fitted for the production of a single product whose every part had been standardized to the minutest detail” (Hounshell, 1984, p. 288). In contrast, Chevrolet production had been designed to accommodate change and marked the beginning of what Hounshell calls flexible mass production. With the addition of automated factory equipment in the 1950s, flexible mass production would also allow for variety within a product run – body style, color, trim, and powertrain options.

5.3 Birth of Chrysler, 1920-1930

Walter Chrysler left GM in 1920, unable to tolerate Durant’s “erratic decision making and arbitrary interference in Buick’s operations” (Flink, 1988, p. 68). Later that year, Chrysler was hired as executive vice-president of the Willys-Overland Company in an attempt to save its holding company, the Willys Corporation, which suffered from the post-World War I recession. While at Willys, Chrysler hired three talented engineers away from Studebaker – Fred Zeder, Owen Skelton, and Carl Breer – who brought with them twenty-eight men from Studebaker’s engineering department. The three ‘wizards’ designed an entirely new and far superior car featuring a six-cylinder in-line engine, an updraft carburetor, and semi-elliptic front and rear springs (Hyde, 2003). Chrysler planned to give the car his name and manufacture it under the Chrysler Motor Company, a newly incorporated division of the Willys Corporation. However, the parent corporation went into receivership in November, 1920, before the vehicle went into production. The Willys plant and the plans for the new car were auctioned off, purchased by the new Durant Motors, which built and sold the car as the Flint.

Chrysler then moved on to supervise the reorganization of the failing Maxwell and Chalmers Motor Companies which were in merger negotiations. Maxwell had bailed out Chalmers, a low-volume high-price producer, by leasing its plants and keeping its automobiles in production. Chrysler restored the Maxwell Company's reputation and returned the company to profitability by improving the faulty Maxwell automobile design, introducing a new axle in 1921 and aluminum pistons in 1922. Chrysler then hired the Zeder-Skelton-Breer team, which had opened a consulting firm named ZSB Engineering, to design a new six-cylinder car to be named the Chrysler Six.

The settlement of a lawsuit against Chalmers in late 1922 placed the company in receivership, allowing the Maxwell Motor Corporation to purchase the Chalmers property outright. Maxwell continued production of the Chalmers car through 1923 then retooled the plant to produce the Chrysler Six. In 1923, Chrysler signed a four-year contract to serve as president of the Maxwell Motor Corporation – the only way he could obtain the support of the board of directors for his new car. The following year, Maxwell introduced the Chrysler Six at the New York Automobile Show. The fast, powerful, and stylish automobile was an instant success at the show and in the market. Advertisements for the Chrysler focused on its technological superiority and emphasized the use of Fisher bodies as proof of its overall quality. The speed and power of the stock automobile were proven by a number of racing successes in 1924 and 1925.

Though not radical in design, the moderately priced automobile offered a host of advanced technological features as standard equipment, including a high-compression engine (4.7:1 versus the typical 4:1), hydraulic four-wheel brakes, and balloon tires. The Chrysler Six was the first moderately-priced car to offer a high compression engine, a technology that was enabled by the 1923 introduction of Ethyl (leaded) gasoline and improved crankshaft balancing. Low-pressure balloon tires were developed by Firestone in 1923 (Abernathy et al. 1983) and offered a smoother ride than older style tires.[64] They were included as standard equipment on fourteen of the models displayed at the 1924 show (Hyde, 2003). Duesenberg first introduced four-wheel mechanical brakes, a European import, in 1922 by adding standard drum brakes to the front wheels. Malcom Loughheed (later changed to Lockheed) developed a four-wheel hydraulic braking system that Duesenberg adopted on its Model A in 1923. In October of 1923, Maxwell offered hydraulic brakes as an option on its Chalmers models. However, the ZSB team found that the Lockheed brakes leaked hydraulic fluid under heavy use. Working with the Manhattan Rubber Company, ZSB designed a rubber cup to replace the rawhide one employed by Lockheed to seal in the fluid when braking. ZSB allowed Lockheed to incorporate the change into his design in exchange for free licensing of the Lockheed brake system. The 1924 Chrysler catalog referred to the system as the ‘Chrysler Lockheed hydraulic four-wheel brake’ (Hyde, 2003, p. 32).

[64] Abernathy et al. (1983), indicate that Firestone’s new low pressure balloon tires were first introduced on Ford automobiles in 1923. Although Firestone supplied Ford, Nevins and Hill (1957) state that Ford was late in adopting this innovation.

In mid-1925, the company began selling a modified Maxwell Model 25 as the Chrysler Four and discontinued the Maxwell line. That same year, Chrysler reorganized the firm as the Chrysler Corporation. Chrysler offered "new and improved" models approximately every twelve months and expanded the variety of cars it offered to include four distinct lines, with each line offering four to nine body styles. Sales tripled from 31,429 in 1924 to 106,857 in 1925 and reached 192,083 in 1927. By the late 1920s, the Chrysler Corporation had established itself as a leader in technological excellence and innovation and according to Nevins and Hill (1957, p. 444), "boasted of an engineering laboratory that would have done credit to Yale or Cornell."

The biggest challenge for the company during this time was expanding production fast enough to keep pace with demand, a difficult proposition since building new facilities was both costly and time consuming. Chrysler relied on the Fisher Body Company to supply most of its car bodies, but purchased a large body plant directly across from one of its Detroit facilities in 1925. After Fisher Body became wholly owned by GM in 1926, most Chrysler bodies were supplied by the Briggs Manufacturing Company, the Edward G. Budd Manufacturing Company, and the Murray Corporation. The company expanded its existing manufacturing facilities in 1928.

However, the biggest increase in plant capacity came when Chrysler purchased Dodge Brothers in 1928. The Dodge brothers had launched their automobile manufacturing business in 1913 after ten years of manufacturing parts worth up to 60% of the value of Ford automobiles. Because of their reputation, when the first Dodge car was introduced in November, 1914, it was an instant success. Perhaps the most significant feature of the new four-cylinder Dodge was the use of an open all-steel body produced by Edward G. Budd (see section 5.4.2). The steel body gave the car superior strength and allowed the use of an oven-baked enamel finish which reduced finishing time to five days.

During the first year, Dodge only offered one model in two body styles. In 1915, it added a third 'Winter Model' which was merely the standard touring or roadster Dodge with a detachable steel top and removable glass windows. The Dodge brothers followed a strategy similar to Ford's, eschewing annual model changes and instead making minor improvements to the existing design. In 1917, the wheelbase of the Dodge automobiles was increased and a coupe and a sedan were added to the line. Two delivery trucks were introduced in 1918, and a four-door sedan, the first of its type in the industry, was added in 1920. However, there were few styling changes to the Dodge models from 1914-1920 (Hyde, 2003).

While at the New York Automobile Show in January, 1920, both of the Dodge brothers contracted influenza. The disease took the life of John Dodge on January 14. Weakened by his illness and devastated by his brother's death, Horace spent much of the year in Florida and died on December 10. The company continued to operate successfully under the leadership of Frederick J. Haynes, whom the brothers had hired in 1912 to oversee production. Following the Dodge brothers' policies, Haynes pursued engineering improvements and cost reductions rather than cosmetic changes. In the fall of 1921, the share of closed car bodies in the Dodge lineup jumped to 35%. For the 1923 model year, Haynes and Budd jointly designed two all-steel closed-body models that were introduced in the summer and fall of 1922, another industry first. Since these models used a steel frame and steel panels, they allowed Dodge to use a baked enamel finish on its closed-body cars.

In 1925, the Dodge widows sold the company to a banking house whose executives had no experience manufacturing or selling cars. The new management introduced an entirely new line of automobiles in 1927, all more expensive than the old four-cylinder Dodge models. Though solid and comfortable, they had few innovative features and failed in the market; Dodge sales and profits plummeted.[65] Though the management correctly perceived a market shift to larger engines and more luxurious cars, in going 'up-market,' Dodge had abandoned its long-time customers and entered a market segment where it had no reputation (Hyde, 2003). The company had spent nearly $15 million retooling plants to produce the new automobiles and now could not sell them. In a year when the price of automobile stock generally increased 50%, the price of Dodge Brothers stock fell, prompting the bankers to sell.

As luck would have it, Chrysler desperately needed to increase his production capacity and was as anxious to buy Dodge as the bankers were to sell it. The Chrysler line was popular and profitable and the company already had plans to introduce two new models. The new low-priced Plymouth was intended to compete with Ford and Chevrolet, while the mid-priced DeSoto was to compete with Dodge. However, competing with Ford and Chevrolet in the low-priced car market would have been extremely difficult without a large established dealer network. The acquisition of Dodge doubled Chrysler's sales outlets with the addition of a distribution and dealer network

[65] Factory sales of automobiles fell 1% (42,800) in 1926 and 20% (755,800) in 1927, then rebounded by 28% in 1928 to exceed 1925 sales. In 1926, Ford sales alone dropped nearly 450,000 from a high of nearly 2 million units. In May of 1927, Ford shut down the Rouge plant to retool for the new Model A and production ceased for almost a year. This shutdown left the low-priced field wide open for Chevrolet and the new Chrysler Plymouth.

that was generally considered one of the best in the industry (Hyde, 2003). Chrysler was also especially interested in adding casting and forging capabilities and the Dodge Brothers plants had large, modern, and efficient facilities.

According to Hyde (2003), the purchase of Dodge was not just intended to increase production capacity and capabilities, but was a strategic move to ensure the long-term survival of the company. Since Chrysler and Dodge models overlapped in terms of price categories, the purchase reduced competition in those segments while increasing Chrysler's market share within them. Most significantly, the merger would allow Chrysler to operate more like, and therefore compete with, the dominant force in the industry – General Motors. A larger size and sales volume enabled the operation of finance and export companies; greater production of parts; and economies in research, engineering, and purchasing. When the merger was complete on July, 1928, Dodge became a division of Chrysler, making it the third largest U.S. automaker behind GM and Ford. An editorial in the *Automotive Daily News* noted that the industry was now dominated by a 'big three' that accounted for nearly 75% of passenger car manufacturing (Hyde, 2003).

5.4 Enclosed Steel Bodies

The earliest automobile bodies were styled and built similarly to carriages, and even the names were borrowed from this tradition: stanhope, surrey, brougham, phaeton, tonneau, cabriolet, landaulet, and wagonette. The techniques for shaping wooden bodies – steaming and bending – resulted in simple curves with little distinguishing style characteristics from one make to another. The use of sheet metal and aluminum, beginning around 1900, allowed for greater styling differences. Columbia Electric produced the first enclosed body, constructed of a wood frame and steel panels, in 1898. Marmon introduced the first all metal (open) body using cast aluminum in 1902 (Abernathy et al., 1983). However, closed wooden bodies were expensive, costing 30-50% more than open bodies (Epstein, 1972), and did not hold up well to the vibration and shock of rough roads. The use of metal was also expensive and therefore was found only in high-priced automobiles and limousines. In 1916, 98% of new cars bodies were open (Epstein, 1972) and as late as 1920, 85% of car bodies were constructed of wood (Abernathy, 1978).[66] Motorists could protect themselves from the elements by purchasing canopy tops from automobile or accessory manufacturers. In 1905, Peerless introduced a folding top which became standard equipment by 1916 (Epstein, 1972).

Beginning in 1916, the percentage of automobiles manufactured with closed bodies rose from a mere 1.5% to 17% by 1920 and reached 72% by 1926. Epstein (1972) credits the increasing popularity of closed bodies in part to road improvements which reduced the vibrations and shocks that damaged automobile bodies. In addition, quantity production brought costs and prices down and extended the market for closed bodies beyond the wealthy. The difference in price between open and closed bodies fell to around $100 by 1926, costing about the same as optional folding canopy tops offered prior to 1916 (Epstein, 1972). The Hudson Motor Company was among the first manufacturers to produce closed-body cars in volume, introducing the affordable Essex coach at $1495 in 1921. By 1925, the price of the four-cylinder Hudson coach had

[66] The wood or wood and steel automobile body and chassis were attached to a frame consisting of two long bars. Initially the frame was constructed of wood, which produced a quiet ride. In 1903, A.O. Smith introduced a pressed steel frame (Abernathy et al., 1983). As manufacturers began instituting mass production techniques, they found suppliers could not produce wood frames quickly enough to meet the speed of moving assembly lines. As the industry shifted to pressed steel frames, A.O. Smith soon became the leader in frame production, a position it held for most of the twentieth century (Rubenstein, 2001).

dropped to $895, placing it within the low-price field (Sloan, 1964). One firm, the Fisher Body Company, is largely responsible for the development of inexpensive closed bodies.

Meanwhile, the use of all-steel bodies was also rising thanks to innovations pioneered by the Edward G. Budd Company that significantly reduced production time and costs. Manufacturing wooden bodies entailed a large amount of hand crafting that simply could not be replaced with Fordist production techniques. The adoption of all-steel bodies banished the last foothold of the craft system from automobile manufacturing. Steel also improved the strength and flexibility of bodies, making them quieter and more comfortable. At the same time, it allowed greater variety in body styling. By 1925, Budd's all-steel technology was used in 50% of U.S. body production.

5.4.1 The Fisher Body Company

The industry shift to closed bodies is attributable to Fred J. and Charles T. Fisher who formed the Fisher Body Company in Detroit in July, 1908. The brothers were third generation vehicle craftsmen with experience in carriage building and in motor vehicle body manufacturing for the C.R. Wilson Body Company, then the largest U.S. automobile body firm. The Fisher's younger brothers, six in total, joined them in the business later. While many automobile body builders, particularly older ones, employed machinery and styling inherited from the carriage business, the Fishers understood the specific needs of the automobile for strength and resilience and were therefore more progressive. Most notably, for two years they actively solicited orders for closed bodies, believing it would extend the driving season and expand the market for motor automobiles especially by appealing to women. In 1910, Cadillac ordered the first 'large' volume of closed bodies, 150 units, prompting the Fishers to organize the Fisher Closed

Body Company (Pound, 1934). The Fishers quickly gained a reputation for building high quality bodies. By 1916, after merging all their interests into a single company, the Fisher Body Company was a leader in the industry, manufacturing 370,000 car bodies a year.

As Fisher production expanded, so did concerns about capital and future demand. If the brothers issued additional stock to new holders, their control of the company could become weakened. In addition, the automobile industry was becoming increasingly concentrated and their customers could switch to a new supplier or begin manufacturing their own bodies. The Fishers' best options were to either begin producing their own automobiles or merge with an existing manufacturer. As they made tentative plans for the former option, two manufacturers initiated negotiations for mergers. William Durant had tried unsuccessfully to employ four of the Fisher brothers at GM and was alarmed by the potential loss of a source for closed bodies, which he believed might be of strategic importance. In 1919, GM outbid the competitors, purchasing 60% of the Fisher Body Company. GM signed a ten year contract to purchase, at cost plus 17.6%, as many closed car bodies from Fisher as it could supply, though Fisher was free to produce bodies for other customers. Under the terms of the merger, Fisher Body and GM each selected half the members of the board of directors for the reorganized Fisher Body. The firm's stock was placed in a five year voting trust with each side assigning two trustees. Fisher retained a majority control in a newly appointed operating committee, but GM secured majority control of the finance committee. For the next five years, Fisher essentially would be allowed to continue operating as an independent firm. However, the brothers were contractually obligated to remain with the company for those five years (Coase, 2000; Freeland, 2000).

At $26.7 million, GM's purchase price of the 60% share of Fisher Body was exceptional even in a year of frenzied acquisitions. But according to Pound (1934, p. 182), the significance of this event went far beyond the value of the physical assets obtained:

> The importance of this alliance, which was later extended to the point of complete merger, can hardly be overestimated. It gave General Motors first call on the production of the largest and best equipped body-building plants in the world. With the steady trend toward closed cars, the Fisher brothers… had pushed forward until they occupied a foremost position in their line both as to quantity and quality. The Fisher name had become known far and wide, and the presence of a Fisher body on any car recommended it to the consuming public. By this one decisive step the General Motors Corporation wrote off future body difficulties by placing that business in the hands of outstanding specialists in the field of body manufacture. With advancing complexities in design and manufacture, the wisdom of this alliance has become increasingly manifest.

Apparently, the Fisher brothers should not have been concerned about losing their customers, because they possessed a scarce resource – knowledge. Freeland (2000) argues that GM's purchase of a majority share in Fisher Body was motivated by "the desire to acquire and retain the specialized knowledge and services of the Fisher brothers. Unified ownership was thus driven much more extensively by the attempt to capture scarce knowledge than has been recognized."

While the work provided by Fisher Body continued to be efficient and of high quality, GM officials felt that the Fisher brothers were not sufficiently aware of or concerned with GM's needs, particularly in terms of expanding body production by locating new plants near GM assembly facilities. Concerned first with the bottom line for their own company, the Fishers preferred to expand existing facilities instead of building entirely new plants. In order to improve coordination between the two companies, Pierre du Pont placed Fred Fisher on GM's board of directors in 1921 then appointed him a member of the executive committee in 1922. These positions involved him in making broad decisions regarding GM's production, product design, and pricing (Chandler and Salsbury, 1971).

As the end of the five year employment contract neared in October, 1924, closed-body cars were increasingly popular and GM became concerned about retaining the Fishers. The possibility of a merger was discussed as early as 1922, but the brothers had no desire to become GM employees and were reluctant to see the family business disappear. In addition, they wished to continue receiving financial incentives comparable to the lucrative profit-sharing terms of the original contract, while GM was concerned that such an arrangement would limit Fisher Body's ability to expand with GM. A new agreement was brokered that compensated the brothers through GM's stock incentive plan rather than from Fisher's profits. To receive the full benefit, the brothers would have to stay with the company until 1929. Fred Fisher was also appointed to GM's finance committee while Lawrence Fisher joined the GM board of directors and the executive committee. This agreement retained the knowledge and skills of the brothers and had the added benefit of further aligning the interests of the two companies. From this time, it apparently was understood that the GM-Fisher relationship was permanent and the two firms would eventually merge (Coase, 2000; Freeland, 2000).

Despite the improvements in communication and coordination between the firms, availability of bodies continued to limit automobile production as late as 1925. The wooden frames in Fisher bodies limited the use of heat to speed drying and therefore created a production bottle neck. According to a GM publication, in 1923, Fisher “pioneered the use of lacquer in place of paint and varnish for bodies. This was a vital step in attaining volume production, thereby helping to bring the closed car within the reach of the average buyer” (cited by Nieuwenhuis and Wells, 2007). However, this innovation originated at Du Pont (see section 5.2.4.4) and its early adoption at Fisher was facilitated by GM’s close relationship with both firms. In fact, according to Cray (1980), the Fisher’s were reluctant to adopt the new finish because of its dull finish and difficulties getting the paint to adhere. Though GM sought an exclusive license for the new Duco lacquer, Du Pont made the paint available to the entire motor industry in 1925 (Cray, 1980; Sloan, 1964). The shift to lacquer reduced the time required to paint and trim a body from four weeks to six hours. And, as GM (cited by Nieuwenhuis and Wells, 2007) later pointed out, it “ushered in a new era of color in automotive styling.”

In 1925, the entire GM line was redesigned with an increasing emphasis on color and style and a growing number of closed bodies. Sales were booming and GM needed to expand production, which meant Fisher Body needed to expand as well. The Fisher’s were still reluctant to move to new facilities close to GM assembly plants, in part because they still provided bodies to other manufacturers. Closed-body styling was a primary means of product differentiation and the Executive Committee was concerned that control of a strategic asset would be compromised if the remaining 40% of Fisher Body’s stock were sold to outside interests, such as a speculator like Durant. Only a year after GM reached a new employment contract with the Fishers, Pierre du Pont once again suggested the companies merge. Du Pont and Fred Fisher negotiated an exchange of 1.5 shares of GM stock for each remaining share of Fisher stock, half of which was owned by the brothers. Fisher Body became a fully owned GM division. Charles Fisher joined the

Executive Committee and Sloan appointed Lawrence Fisher head of Cadillac. All but one of the six Fisher brothers eventually served as senior executives at GM. GM's complete ownership of Fisher Body cut off competitor's access to the most important supplier of closed bodies and helped propel GM to the top position in the industry (Freeland, 2000). In 1934, Fisher would be responsible for the innovation that removed the last vestige of fabric from auto bodies: the turret top.

5.4.2 Edward G. Budd Manufacturing Co. of Philadelphia

After graduating from high school, Edward G. Budd apprenticed as a machinist in a Smyrna, Delaware, iron works firm. Around 1890, Budd moved to Philadelphia where he worked for a foundry and a manufacturer of machine tools and metal presses. The art of pressing steel sheets was a relatively new idea, dating back to around 1860. In 1911, *The Horseless Age* (cited in Grayson, 1978) wrote "by the term pressed steel, engineers generally understand that the material has been worked into shape from plate or tube by means of stamping and drawing processes performed either while the material is hot or, in some cases, cold. The process is therefore one which usually changes the form of rather than the thickness of the material operated upon…"

Budd became acquainted with a Connecticut mechanic who had designed a pulley made of sheet steel which was lighter and cheaper to produce than traditional cast iron. When his friend received financing to produce the new pulley, Budd joined him as a draftsman and engineer at the American Pulley Company. The pulleys proved to be strong and flexible and were a market success. As the business expanded, American began producing stamped steel parts for Hale & Kilburn, a furniture manufacturer that specialized in seating for railways, subways and trolleys. Most Hale & Kilburn's parts were produced using cast iron, and in 1902, the company hired Budd away from American to develop lighter, cheaper pressed steel parts. Budd and co-worker Morris Lachman accomplished these goals using new techniques in oxy-acetylene welding enabled by the work of Frenchmen Henry Le Chatelier, Edmond Fouche and Charles Picard and Americans Thomas Wilson, Major J. Turner Morehead, and Eugene Bournonville.

Budd and Lachman also experimented with resistance and arc welding. The electric resistance welder was developed around 1888 by Elihu Thomson, an Englishman working in the U.S. In 1892, C.L. Coffin of Detroit received a U.S. patent for an arc welding process using a metal electrode whereby the metal from the electrode was used to fill the joint. With this new expertise in welding and pressed steel fabrication, Budd designed all-steel passenger rail cars that Hale & Kilburn produced for the Pullman Company in the early 1900s.

Budd soon began experimenting with shallow-draw sheet metal stamping, and produced a small number of sheet metal panels for the King and Paige Co.'s automobile bodies. In 1909, Budd was contacted by Emil Nelson, the chief engineer of the Hupp Motor Car Co., for help developing an all-metal body. Although advances in sheet metal manufacture made possible larger stampings of uniform thickness, the compound curves demanded by automobile bodies still required multiple stampings connected by hand welding and no Detroit firm was willing to contract to build an all metal body. Budd already believed that car bodies should not have a piece of wood "as big as a toothpick," and eagerly embraced the project (Grayson, 1978).

Since they lacked the equipment and experience necessary to draw complex shapes from a single piece of metal, Budd and Nelson designed the cowl and tonneau from formed and welded panels, and devised a system to support the parts during welding. Hale & Kilburn began supplying Hupp with pressed steel body panels and Budd began development of an all-steel body. In 1911, Budd developed and Hale & Kilburn built for Hupp the first automobile body that used steel-reinforced wood frames, called armored wood, to which sheet steel was attached. In 1912, Hupp introduced the Model 32, the first car with an all-steel body produced by Hale & Kilburn (Grayson, 1978). At this time, the painting of bodies was a production bottleneck. Because of the wood frames, the bodies could not be baked to speed drying because this could result in fires. Gluing, sanding, staining, and varnishing took weeks. By adapting stove enameling on the Model 32, Hubmobile reduced the process to one day (Grayson, 1978; Nieuwenhuis and Wells 2007).

In 1911, Hale & Kilburn was sold to J.P. Morgan and the management was replaced. The new management had no understanding of the stamped metal business and Budd had difficulty tolerating the stream of senseless orders. In the same year, Robert C. Hupp left the Hupp Motor Car Company. In 1912, Budd resigned from Hale & Kilburn and Nelson resigned from Hupp. Although the Hupp Company continued to manufacture the Model 32, it incorporated a traditional wood and steel body.

In June, 1912, Budd formed the Edward G. Budd Manufacturing Co. and hired a handful of coworkers from Hale & Kilburn.[67] Budd's son later claimed "[i]t is probably true that this small group of people represented the greatest aggregate knowledge that

[67] One of these, Joseph Ledwinka, would later be responsible for numerous lucrative patents and would be a significant factor in Budd's success.

existed in the Country at that time, of the art of making dies and using presses for the formation of light gauge steel parts" (Grayson, 1978). Budd believed that steel was the material best suited for manufacturing both automobile bodies and interior frames. Steel "gave the elasticity that was desirable so that some flexibility of the chassis was absorbed by the flexibility in the body. It gave greater strength at less weight. Our art reduced to the minimum the number of joints in the body and so kept the body water tight, wind tight, and quiet" (Grayson, 1978).

Budd continued to perfect the techniques required to manufacture all-steel bodies, eventually developing jigs and fixtures that prevented the stamped steel parts from deforming during the welding process. Soon after being appointed president of GM in 1912, Charles Nash ordered a sample metal body for Buick then ordered 2,000 metal touring (open) bodies for GM's Oakland division. Willys-Overland also ordered 2,500 metal touring bodies that year for its recently purchased Garford Truck Company. Budd purchased a facility in Detroit to paint and upholster bodies shipped from Philadelphia. Unfortunately, Garford filed for bankruptcy soon afterward and Oakland demand dropped. Budd was forced to sell the Detroit facility, but was able to keep the company afloat by manufacturing truck bodies, automobile fenders, and stamped panels and trim for railway cars.

The breakthrough came for Budd when the Dodge Brothers ordered 5,000 all-steel bodies for their new automobile in 1914. With its instant success, Dodge ordered 50,000 all-steel bodies in 1915. In June of 1915, Budd was granted a patent on the design of a welded all-steel touring car body. Demand for the new Dodge exceeded Budd's production capabilities, so the C.R. Wilson Body Co. supplied Dodge with bodies using wooden frames (Grayson, 1978; Nieuwenhuis and Wells, 2007). In addition to complete bodies, Budd supplied stamped parts for Buick, Reo, Willys-Overland, and later, Chevrolet.

In 1916, Budd introduced a hard-top with optional side curtains that could be attached to the Dodge chassis during inclement weather and removed during fair weather, thus creating the Dodge 'Winter Model'. In 1917, Budd developed an all-steel sedan for Dodge, but dedication of Budd facilities to the war effort delayed production. In 1922, Budd and Dodge's Frederick Haynes developed an all-steel coupe body for Dodge that was introduced in the 1923 model year. It was the industry's first all-steel closed-body car. Due to limitations in sheet metal width and concerns over noise problems, the roof section still used a wood and wire frame over which fabric was stretched.

As closed bodies gained in popularity during the 1920s, Budd expanded production rapidly. Significant economies of scale realized from high production volumes combined with improved welding and engraving machinery allowed Budd to cut the average body price by 40% in 1922. By 1925, Budd's all-steel technology was used in 50% of U.S. body production (Nieuwenhuis and Wells, 2007). Ledwinka continued to improve manufacturing techniques, developing a huge welding machine capable of producing steel sheets 140 inches long – longer than those rolled by steel mills. This allowed Budd to stamp entire side panels from a single sheet (Grayson, 1978).

Although Budd did not invent the steel body or any of the core technologies used – production of sheet steel, forming by hydraulic presses, resistance welding, and the use of jigs and fixtures to prevent distortion – he was the first to see the possibilities for these processes to revolutionize body manufacturing. His synthesis and adaptation of existing tools and processes eliminated the remaining need for craft techniques in body production, beginning with the elimination of handcrafted wood components and ending with the replacement of hand bolting of steel panels by automated welding in 1925 (Abernathy et al., 1983). Nieuwenhuis and Wells (2007) assert that researchers have under-valued Budd's contributions to the displacement of craft manufacturing by mass production. "[T]he early Fordist revolution in mass manufacturing of cars was incomplete, and required the introduction of Budd's all-steel body technology to resolve key bottlenecks in the manufacture and painting of vehicle bodies and thereby enable a transformation of the industry." Without a doubt, the closed steel body marked a major change in the functional specifications and design attributes of the automobile. According to Abernathy (1978, p. 18), "The very concept of the automobile was changed for the consumer by an early technological advance in body design. The introduction of closed steel bodies during the 1920s raised a whole new set of criteria for automotive design – passenger comfort, room, heating and ventilation, and quietness of ride."

5.5 Time Sales and Financing

Even before 1910, consumers had become willing to forego savings and investments or go into debt to own an automobile. Banks generally disapproved of financing the purchase of luxury items like "pleasure cars," a perspective which persisted until sometime in the 1920s (Sloan, 1964). As a result, Americans began mortgaging their homes to obtain the necessary cash, leading sometimes to foreclosure (Flink, 1970). However, the automobile became more utilitarian and more affordable just as the wealthy market became saturated. Dealers soon found that some promising sales prospects required extension of credit. In 1911, the Studebaker Corporation announced it would accept dealer-endorsed notes from farmers and other responsible buyers (Smith, 1968; Flink, 1988). In the following years, dealers began arranging for customers of financial standing in the community to make a down payment of 66-75% with the balance due in one or two payments within one to three months (Epstein, 1972).

Success with these arrangements emboldened some dealers to develop installment plans under which they held title to the car until it was completely paid for. This technique required that dealers make arrangements with local banks and finance companies. While banks still were not willing to handle large quantities of loans for automobiles, they were willing to lend money to specialty finance companies. Experience soon showed that these installment plans were a good low-risk investment when the automobile was well insured and the payment plan took into account the depreciation of the car, which was used as security against the loan. By assessing relatively high fees for these low-risk loans, the finance companies prospered (Pound, 1934).

As the scale of the automobile business grew, dealers shouldered ever-increasing costs of new car inventories and the facilities required to display and stock them. In addition, they needed to provide service facilities and space for stocking and selling used cars. As a result, the specialty finance companies began financing dealer stock as well as consumer purchases (Pound, 1934). Under increasing pressure from their dealers, manufacturers of moderately priced cars began to view credit sales as an alternative to Ford's strategy of progressively lowering prices (Flink, 1988). In 1915, John Willys created the industry's first captive lending agency to finance time sales of Overland and Willys-Knight cars (Sloan, 1964; Cray, 1980). In 1916, the company was reorganized to finance twenty-one makes of cars, including all GM makes, Dodge, Ford, Hudson, Maxwell, Reo and Studebaker (Flink, 1988).

GM became the first manufacturer to directly finance consumer purchases of new and used cars when Durant and Raskob created the General Motors Acceptance Corporation (GMAC), a GM owned subsidiary, in 1919. GMAC also financed wholesale purchases by GM dealers. According to Sloan (1964, p. 306), GMAC was established to prove the validity of consumer financing of automobiles and to "crusade for reasonable rates." Over the next few years, GMAC formulated and refined policies regarding loan terms and structure in order to minimize the risks associated with default. GM also commissioned a multi-year study by economist E. R. A. Seligman who reported that installment credit strengthened consumers' willingness and ability to save, increased purchasing power, and stabilized and increased production. Published in 1927, the study "had a strong influence… in bringing about an acceptance of installment selling among bankers, businessmen, and the public" (Sloan, 1964, p. 306).

The use of installment sales became an important strategy in reducing swollen inventories during the recession that followed World War I, particularly the stock of used cars traded in as a down payment on a new automobile. As the automobile market saturated in the 1920s and manufacturers began pursuing annual model changes to stimulate new car sales, disposing of used cars became a major issue and consumer credit became essential to the industry. Because an inexpensive used vehicle offered more options and conferred more status than a new Model T, the sale of these cars on easy credit also cut into Ford's market. This effect was reinforced by the fact that Henry Ford refused to engage the Ford Motor Company in consumer lending and did not endorse installment sales by Ford dealers.

By early 1921, there were over 110 automobile finance companies in existence and by 1925, 75% of new and used automobile sales were made on time-payment plans (Flink, 1988; Epstein, 1972). Prior to the 1920s, merchants arranged time-payment purchasing for furniture, sewing machines and pianos, and mortgages were commonplace for buying homes. But now the concept was increasingly applied to all durable consumer goods and was rapidly becoming a mainstay of the U.S. economy (Rae, 1965; Flink, 1988). One dealer in 1928 went so far as to assert that it was practically a patriotic duty to buy and sell on credit, since keeping Americans wanting more than the bare necessities kept Americans working and thus kept the economy growing (Cray, 1980). Eventually, even the Ford Motor Company was forced to followed suit. Ford established a credit corporation to fund Ford dealers and customers in 1928, concurrent with the changeover from the Model T to the Model A. Hounshell (1984, p. 293) deemed the new Universal Credit Corporation "perhaps the most revolutionary change (considering Ford's detestation of credit buying) wrought by the changeover."

5.6 Technical and Institutional Context

5.6.1 World War I, Recession, and Economic Recovery

From the beginning of World War I, the European automobile industry played a key role in what would be the world's first mechanized war. Rapid conversion to the war effort led to a phenomenal expansion of capacity and employment, mainly for the production of trucks, tractors, and tanks, but also for diversification into munitions, aircraft engines, airframes, and other items useful to the military. The war effort also led to the modernization of French, British, and Italian factories.

With the possibility of U.S. involvement in the war looming, the U.S. government called on the domestic automotive industry to help with American preparedness. Hudson vice-president Howard Coffin was tapped to serve on the Navy Department Advisory Committee in August, 1915, then was appointed chair of the Committee on Industrial Relations, an agency of the Council of National Defense, in 1916.

Even after the nation entered the war on April 6, 1917, U.S. automobile production continued unabated until well into 1918, making 1917 a banner year for the industry. Hundreds of automobile industry executives volunteered for the war effort, though most of the industry's support came from the smaller manufacturers whose market position had been deteriorating despite the boom conditions. Because of their smaller investments in specialized plants and equipment, these firms were able to convert to wartime production more easily and at lower cost than Ford and GM. In addition, the luxury automobile chassis of Locomobile, Packard, Peerless, and White were easily adapted to support two- and three-ton truck bodies for the military. By 1916, Packard was already making more trucks than cars. White had received orders to supply trucks to the Allies at the onset of hostilities and had found truck production so profitable that it discontinued making cars in 1917. The Nash Motor Company, formed in 1916 from the ashes of the Thomas B. Jefferey Company, grew to be the leading truck manufacturer by the end of the war (Rae, 1959; Flink, 1988).

Henry Ford had taken a pacifist stance against the war in late 1915, but revised his position and vowed to stand behind the president when diplomatic ties with Germany were severed in early 1917. Ford Motor Company engineers were responsible for process innovations that allowed inexpensive mass production of Liberty engine cylinders, resulting in a contract for Ford to produce all the cylinders for Liberty engines produced in the U.S. and a contract to make 5,000 complete engines. Ford produced a wide variety of other items for the war effort, including armor plate, caissons, shells, steel helmets, submarine detectors, and torpedo tubes. Ford also built 60 Eagle Boats and completed two tank prototypes too late to be used in the conflict.[68] The company's most significant contribution, however, was 39,000 Model T cars, ambulances, and trucks which supplied both American and Allied forces (Nevins and Hill, 1957).

William Durant of GM had also opposed war production, refusing Henry Leland's request to covert Cadillac facilities to the production of Liberty engines. Leland and his son immediately resigned and formed the Lincoln Motor Company to manufacture the engines. Shortly thereafter, under the mounting pressure of public opinion, Durant began token production of Liberty engines at both Cadillac and Buick (Flink, 1988). The Liberty engine itself had been designed for mass production by Packard engineers as a replacement for the Allied engine designs which required hand machining (Rae, 1965).

Supervision of the automobile industry had been assigned to the Automobile Products Section of the Council of National Defense and the War Industries Board. To coordinate government requirements and industry facilities, automobile manufacturers established a committee with representatives of the National Automobile Chamber of Commerce, the Ford Motor Company, and the parts makers. In reality, the committee largely functioned to prevent government interference with industry operations. In

[68] Eagle Boats were small (200-foot) steel submarine chasers intended for use in the battle against German U-boats. Only seven were delivered before the signing of the Armistice and none saw service in WWI.

response to an apparent steel shortage in late 1917, the War Industries Board proposed curtailing the industry's supply of alloy steel for manufacturing passenger cars. Though spokesmen for the industry denied the existence of an actual shortage, the industry was divided on the matter. While large manufacturers lobbied for the leeway to independently solve the supply issue, small producers whose markets were already disappearing preferred to suspend all passenger car production and relinquish their facilities to the government for the duration of the war. The debate continued until the reality of the shortage became unavoidable by mid-1918. The War Industries Board cut the steel allocation to the industry to allow for passenger car production in the last six months of the year at a level of only half of that achieved in the last six months of 1917. In truth, the diversion of plant facilities to war production had already substantially reduced production for civilian use such that total production for 1918 was only 53% of the previous year's (Rae, 1959).

U.S. automakers generally believed that the diversion of plant facilities to the war would lead to unfulfilled demand and a booming market after the conflict ended. With the advent of peace on November 11, 1918, manufacturers rode a wave of optimism, scrambling to reconvert facilities and expand capacity to meet the backlog in demand that by mid-1919 had reached an estimated one million vehicles. Worried about speculation on the stock market and the growing use of installment sales plans for new cars, the Federal Reserve Board had raised the discount rate in November, 1919, and speculators began liquidating over-inflated stock. Meanwhile, rising commodity prices had increased the cost of living. With the decrease in purchasing power, consumer spending slowed. Veterans returning from the war began swelling the unemployment ranks and the oversupply of labor pulled down wages, leading to labor unrest. Although automotive employees generally were not involved, strikes in the coal and steel industries affected supply lines. By the summer of 1920, the country was slipping into an economic recession. To make matters worse, foreign markets for farm commodities collapsed and the American farmer – the foundation of the automobile market – returned to hard times. The 1919 boom in automobile demand had been fleeting, based on inflated wartime prices, easy credit, and over-blown expectations. Now the boom had collapsed. Automobile sales plummeted and trainloads of cars stood undelivered.

Although the recession would prove to be brief – just an economic readjustment to peace-time production and prices – it hit the automobile industry hard. Many manufacturers, especially the largest ones, were in the midst of optimistic expansions and were over-extended. Henry Ford was developing the huge River Rouge complex and deeply in debt from his buyout of the minority stockholders. GM's expansion plans and Durant's loose management, not to mention his questionable financial dealings, left GM in an equally precarious position.

Ford was the first to respond to the slumping demand for cars, cutting prices in September, 1920, as much as 30%, thereby cutting further into its competitors' sales. Soon, twenty-three manufacturers followed suit; GM was not among them. Price cuts failed to sufficiently stimulate industry sales, and by December, production was shut down at Buick, Dodge, Ford, Maxwell-Chalmers, Nash, Packard, Reo, Studebaker, and Willys-Overland. Pierre du Pont had taken over leadership at GM and continued to resist price cuts but was forced to slow production. Nonetheless, the industry recorded profits for the year, largely based on sales in the first half of the year (Rae, 1959; Rae, 1965; Flink, 1988).

Ford's sales began to pick up in early 1921 and production was gradually resumed. Sales of other manufacturer's cars remained slow, and the last remaining holdouts began to cut prices. Pierre du Pont finally entered the price war in May and two rounds of price cuts eventually left a gap of $110 between the Chevrolet and the Model T. However, the more expensive Chevrolet came with features that were only offered at additional cost on the Model T and Ford cut the price one last time. Both vehicles were losing money at the final prices, but the cuts enabled Ford to win 62% of the passenger car market in 1921. Sales of all GM cars only amounted to 14% of the market that year, yet the corporation was able to survive the recession (Cray, 1980).

Automobile sales rebounded in 1922 and reached a new high of 3.6 million in 1923. GM emerged from the recession a stronger corporation, having undergone a major reorganization and begun a complete redesign of its vehicles. Other companies were unable to meet their obligations and went into receivership. Among these less fortunate firms were Maxwell-Chalmers and the Willys Corporation, the remains of which would eventually give birth to the Chrysler Corporation. Meanwhile, firms that had refrained from ambitious expansion plans were better positioned to weather the downturn. These companies emerged relatively unscathed and included Dodge, Hudson, Nash, Packard and Studebaker (Rae, 1959). After a brief economic slowdown in 1924, the country entered a period of rising prosperity that brought steady growth in automobile sales.

5.6.2 Road Improvements, 1917-1930

With the U.S. turning out motor vehicles and supplies for the war effort, the rail lines to eastern ports became clogged in 1917. Highway Transport Committee Chairman Roy Chapin, who was also president of the Hudson Motor Car Company, was tasked with identifying road routes for trucks to be driven from assembly plants in the Midwest to the docks in the east. Loaded with other freight, these caravans demonstrated the viability of long-distance trucking as an alternative to shipping freight by rail and simultaneously drew attention to the need for an interconnected highway system. In addition, Chapin arranged for local authorities to keep the truck routes open in winter, demonstrating the feasibility of year-round use of major highways (Rae, 1959; Rae, 1965; Flink, 1988).

Between 1916 and 1921, Congress had distributed a total of $75 million to states to improve postal roads. Federal support for building roads was substantially increased with the passage of the Federal Highway Act of 1921, which appropriated $75 million for matching grants in 1922 alone. Initially, state expenditures were funded through licensing fees and property taxes or general funds. However, sparsely populated states found these revenues insufficient. To pay for highway construction, Oregon, New Mexico, and Colorado instituted the first gasoline taxes in 1919. State and federal revenues from registration fees and taxes on gasoline, vehicles, and parts reached nearly $189 million in 1923. Public highway expenditures that year totaled $1.5 billion, while the amount spent on investment in and upkeep of automobiles and highways combined reached $6 billion. By 1929, all states had instituted gasoline taxes, raising $432 million that year and entirely replacing the use of property taxes and general funds for main roads (Nevins and Hill, 1954; Cray, 1980; Flink, 1988).

A 1933 report by the President's Research Committee on Social Trends concluded that the automobile was the largest single contributor to the expansion of local taxes between 1913 and 1930. Highway expenditures, funded largely by use taxes, were only part of this burden. The automobile had also required expenditures to solve traffic congestion, increased crime, and traffic control. In addition, increasing suburbanization, aided by automobility, required the provision of facilities and public services for suburban communities (Flink, 1988). However, it is important to note that the desire for suburbanization predated the automobile. The rise of factory production late in the 19th century, combined with swelling immigration, led to polluted conditions and an increasing urban concentration of poorly paid workers and their families. The automobile enabled the growing middle class to fulfill a pre-existing desire to escape the city while simultaneously strengthening this demand This led eventually to socio-economic and racial segregation and the impoverishment of urban centers. Thus, while the automobile and improved roads got "rural America out of the mud," it also caused a host of problems for urban planning.[69]

Early in the development of the automobile, critics had warned that the innovation would not find market success until the roads were improved. However, construction of roads passable by motor vehicles lagged well behind the development and diffusion of the automobile itself. The capability to navigate rough roads was therefore a main design criterion for those vehicles, like the Model T, that were ultimately successful in the first

[69] Although the motor vehicle conferred great initial economic and social benefits to rural communities, the shift from horses to tractors had unexpected negative feedbacks. High equipment costs made it more profitable to farm larger tracts and often forced farmers to take out mortgages. Operating costs were further increased by the need to replace manure with chemical fertilizers. As farm productivity rose, prices fell. This shift ultimately led to the displacement of farm laborers, over-production of staple crops, and the erosion of agricultural profits. Thus, impoverished rural populations began migrating to the cities in the 1920s (Cray, 1980; Flink, 1988).

decade of the 20th century. This resulted in cars with high road clearance and rugged suspension, as well as a preference for open bodies since the structural integrity of closed wooden bodies rapidly deteriorated from vibration and shock. With road improvements by the 1920s, these design attributes were unnecessary. They were also undesirable and unsafe, since they resulted in a rough, uncomfortable ride and a high, precarious center of gravity. Thus, road improvements were a factor in the shift in market preferences toward comfort and luxury in the mid-1920s.

5.6.3 Refueling Infrastructure

Around the turn of the century, gasoline was widely available at country stores for use in lighting and stationary engines. At that time, kerosene dominated the petroleum product markets and gasoline was a by-product for which supply outstripped demand. Around 1905, the largest demand for gasoline was for use as a solvent or in chemical and industrial plants; motor vehicles consumed only 600,000 of about 7 million barrels produced annually. However, in 1916, automobiles consumed approximately 1.2 billion gallons of gasoline, or more than 28 million barrels (Melaina, 2007).

Gasoline was distributed using the system established for kerosene consisting of around 100 refineries and a large network of small, dispersed, bulk storage facilities. Fuel was transported from the refinery to bulk stations in barrels or tank wagons and then distributed to retailers by horse-drawn wagons. A variety of dispersed methods were used to deliver gasoline to consumers. Retailers stocked cans of gasoline, up to five gallons each, on store shelves. General stores and garages that sold or used a higher volume of fuel stored gasoline in barrels. The gasoline was ladled it into a pitcher, carried to the vehicle, and poured through a chamois-lined funnel into the tank.

Consumers could also purchase home refueling tanks and pumps, which were popular with wealthy motorists. Many wealthy urban motorists used their vehicles for country touring on weekends and stored them in downtown parking facilities, many of which dispensed gasoline from handcarts with tanks and pumps. These handcarts were also used as mobile refueling centers, as were dispensing tank wagons.

Curb pumps, the precursor to modern gas stations, were introduced in 1907 and consisted of an underground storage tank, a gasoline pump located at the curb, a flow meter, and a dispensing hose. Curb pumps were refilled by tank wagons and were the dominant form of refueling in the late 1910s and early 1920s. Most were operated as a sideline and were owned by and located in front of businesses that either catered to motorists or wished to attract more customers to their primary business. Curb pumps were a common street-side fixture and were installed in an estimated 140,000 locations by 1927. In fact, they were so prevalent that lines of waiting vehicles became an urban nuisance and 14 major cities had banned their installation by 1923.

Dedicated gasoline stations also appeared around 1907 and were distinguished from dispersed methods by the use of land and buildings dedicated to refueling and related services. They also sold larger volumes of fuel and had higher capital and operating costs. The dedicated station rose in popularity after 1915 as motorists became increasingly dissatisfied with long waiting lines at curb pumps and on-the-spot price increases and unreliable fuel quality experienced with many dispersed refueling methods. Dedicated stations were often owned by oil companies, which projected an image of reliability and legitimacy. In addition, these stations began emphasizing service and amenities and "the service mentality was firmly established across the gasoline marketing industry" by the end of the 1920s (Melaina, 2007). These services included free air and water; tires, batteries, and accessories; oil and battery checks; free maps; clean restrooms; and helpful attendants. Early records on the number of refueling stations are inconsistent but suggest that there were around 15,000 in 1920. The first reliable statistics became available with the 1924 census which reported 46,904 dedicated refueling stations[70]; by 1929 the number had risen to 121,513 (Melaina, 2007). Melaina (2007) concludes that "the takeoff period for gasoline stations occurred between 1915 and 1925… so the rise of gasoline filling stations followed rather than preceded the rise of gasoline vehicles."

5.6.4 Market Saturation and the Used Car Problem

Before 1920, the typical new car customer was buying his first automobile. But by 1923, the market for new cars was reaching saturation. Although the total number of personal vehicles in use was still rising, sales of new cars had leveled off as shown in Figure 5-1. According to Sloan (1964, p. 163), after 1923 "the role of the new car was to cover scrappage and growth in car ownership." To a large extent, market saturation was an inevitable result of the phenomenal success of the industry's efforts to decrease production costs. According to Hounshell (1984, p. 186), "Ford pursued specialization in production to its logical conclusion – gross overproduction… Until, Ford, complete saturation of the world market by a single product had remained only a theoretical concept."

[70] The U.S. Census figures include establishments that derived more that half their revenue from the sale of motor fuels.

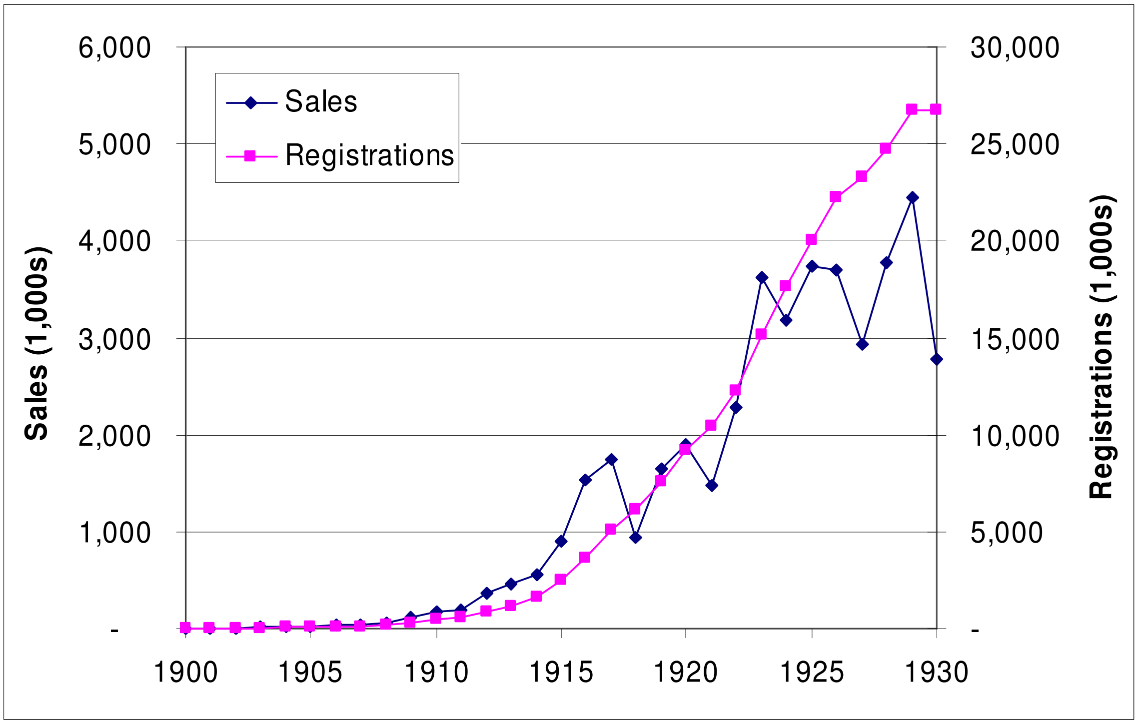


Figure 5-1: Automobile Sales and Registrations
Source: Cain (2006), see Appendix A.

As replacement sales came to represent a larger fraction of the market, the used car was increasingly traded in as a down payment on a new car. Initially, manufacturers considered the disposal of these used cars to be solely the dealers' responsibility. Though there are no reliable statistics on trade-ins, the used car market first became an issue for manufacturers during the post-war recession beginning in the fall of 1920. According to Cray (1980, p. 187) "[f]or the first time, dealers found themselves unable to deal off the used cars taken as down payments on new-car sales. Used car prices fell 25 percent, and still the vehicles sat there. A 'used-car problem,' as it would become known, had surfaced." After the economy recovered, new car sales in 1923 reached a new record when consumers purchased 3.6 million new automobiles but they also traded in 2.8 million used cars (Cray, 1980).

For the first 25 years of the industry, manufacturers had focused on expanding production quickly enough to satisfy market demand. Suddenly, they were competing with each other for fewer and fewer first time-buyers. Saturation presented two new challenges for the industry: disposing of used cars and stimulating demand for new cars. The industry already had a mechanism in place for assisting with the first challenge: installment sales. But the availability of used cars on easy credit meant even more competition for the manufacturers of inexpensive cars. The automakers needed to convince consumers that new vehicles offered something their current cars and used cars did not. To meet this need, manufacturers pursued two strategies: national advertising and the annual model change.

Industry emphasis shifted from production to marketing and sales. Between 1914 and 1924, GM's advertising budget per vehicle increased tenfold. The industry became one of the largest users of magazine advertising, with expenditures rising from $3.5 million in 1921 to $6.2 million in 1923 and to 9.3 million in 1927 (Cray, 1980; Flink, 1988). But the automakers were not alone in their increasing use of national advertising. Because of improved transportation, distribution of manufactured products was no longer confined to a local sales territory around the factory. By the 1920s, national advertising of brand-name products like Pepsodent toothpaste, Coca-Cola, Welch's grape juice, Listerine, Campbell's soup, RCA radios, and now automobiles, was common (Cray, 1980). Automobile advertisements emphasized utility and the freedom to travel. In addition, automakers began to target women in their advertising, emphasizing style, comfort, refinement, social status, and even safety.

By 1910, automakers were well aware of the disadvantages of frequent model changes and most opted to make continual improvements to existing models. New models were introduced as needed in response to perceived market conditions. Although no manufacturers had policies regularizing change, they continually tracked and responded to market trends, introducing new or completely restyled models when they perceived it to be necessary. But by the 1920s, many found they were making increasingly frequent style changes in an attempt to entice buyers to trade up and to win market share from competitors through product differentiation. While GM managers argued against annual model changes, the corporation in effect introduced a new model every year beginning in 1923. Despite manufacturers' resistance, the annual model change was emerging from evolutionary forces as an unavoidable feature of the industry. GM incorporated style as a major consideration in automobile design with the establishment of the Art and Color Section in 1927 and in that same year introduced the first mass produced stylists' car, the La Salle. In 1930, the corporation acknowledged that the annual model change had in fact already arrived and led the industry in planning for regularized change.

5.6.5 Role of Women

A number of exogenous factors increased women's influence in automobile purchases. Educational opportunities for women had been steadily improving. Women were empowered with the achievement of universal suffrage in 1920 and the widespread dissemination of birth-control information. Labor-saving devices freed more of the home-maker's time, but women were increasingly likely to work outside the home as well. Their participation in the workforce during World War I began to alter public views on women's role in society. Prior to the war, when a woman worked outside the home it was regarded with shame, but during the war, it became a patriotic and heroic act. Although most of these women left the fields and factories after the war, their expectations were permanently effected. In addition to gaining experience in the workforce, married women whose husbands served in the war also found themselves making decisions as heads of household. While war widows were forced to maintain these responsibilities, others gladly relinquished them to their returning husbands. But many found their old subordinate roles in the family less satisfying.

Changing economic conditions were also creating new employment opportunities for women. With the rise of industrialism and scientific management, men were promoted to an entirely new class of jobs in middle management and clerical jobs became available to women. A new consumer culture gave rise to chain department stores and new opportunities as clerks. While the percentage of women in the labor force did not grow appreciably during the 1920s, between 1920 and 1930, the fraction of employed women who were married grew from 19% to 25% in professional jobs, and from 18% to 32% in clerical jobs (Nottingham, 1947). With the possibility of paid employment, women began to view marriage as less a requirement for financial stability and more a partnership and source of personal gratification.

The young generation was particularly influenced by World War I and the age of prosperity that followed. A large number of young men did not return from the war; most servicemen did return, but their outlook was permanently altered. Facing a high probability of death in the trenches, these young men adopted a live-for-the-moment attitude, looking for and finding extreme experiences. They had broken with society's structure, and returning to their prior lifestyle proved difficult. Meanwhile, young ladies had also broken with that structure in supporting the war effort at home. These women adopted a similar, live-for-fun attitude and experimented with new sexual freedoms. The 'New Woman' cut her hair short, smoked, drank, and wore makeup. She shortened her skirts and loosened her clothing so that she could dance to jazz music. The automobile suited the flapper's gregarious lifestyle perfectly, affording freedom and thrill.

The safety bicycle and the automobile also played a role in emancipating women, giving them new freedom of mobility. Wealthy urban women had first adopted electric cars, but with electric self-starters and improved transmissions, gasoline cars became accessible to women. With the availability of closed-body cars which protected them from wind, weather, mud, and engine grease, rural and urban women alike took to the automobile. Women wanted cars that were lower to the ground because they were easier to board. They wanted roomier, comfortable interiors with plenty of headroom for hats. For women, the mechanical reliability of the automobile was presupposed and high priority was given to elegance of style, imaginative use of color, luxurious comfort, and ease of operation. Married women may have left decisions about purchasing a horse and carriage to their husbands, but they asserted their preferences when it came to automobiles. Meanwhile, prosperous single women were also becoming a factor in the market.

5.7 Analysis

The narrative history presented in this chapter begins in 1916, well into the first specific phase. This phase lasted until approximately 1918, after which the industry entered a second transitional phase. By 1926, a new dominant design had emerged and the industry entered a second specific phase. In this section, I use the framework presented in section 2.6 to analyze the characteristics of innovation and continued development of institutions during the first specific and second transitional phase. I examine the mutual adaptation of the technology, institutions, and physical infrastructure and discuss the role of alignment and embedding, as well as unexpected events, in the stability of the developing automobile regime. Many of the issues discussed here have roots in events that occurred prior to 1916, but come to fruition between 1916 and the late 1920s. Therefore, this analysis also builds on the narrative in chapter 4.

Section 5.7.1 analyzes the industry trend toward vertical integration and its impact on innovation during the second transitional phase. I identify three factors that contributed to vertical integration in the automobile industry: 1) a trend in the TIEC toward large industrial combinations that could influence the price and supply of raw materials; 2) process innovations that required continuous plant operation and carefully timed delivery of parts; and 3) tacit knowledge that existed only in human capital. I conclude that the influence of vertical integration on innovation depended on the specific strategies pursued and the resulting rules and capabilities that were embodied in corporate structure, operating procedures, facilities, and human capital.

Section 5.7.2 identifies five factors that triggered a second transitional phase: 1) World War I and the post war recession; 2) market saturation; 3) the availability of inexpensive all-steel closed bodies; 4) and improved roads; and 5) the increasing market influence of women. These changes in the TIEC resulted from exogenous events, innovations, and issues associated with the very success of the technology and the development of supporting infrastructure. While the transitional phase did not present serious challenges for the industry until the middle of 1920, the stage had been set by the middle of 1918. Therefore, I mark the beginning of this phase in 1918.

Section 5.7.3 summarizes firm strategies in responding to the changing environment of the transitional phase. I find that firm histories, the personal beliefs and perspectives of firm leaders, and corporate structures and culture – *i.e.*, firm-specific rules – shaped individual firm strategies for innovation and their adaptive capabilities. In addition, I find that unexpected events presented insurmountable difficulties for some firms, while other entrepreneurs were able to exploit the opportunities presented by those firms' failures.

Section 5.7.4 reviews the emergence of a new dominant product design that arose from co-evolutionary forces. I find that production processes and management systems, *i.e.* corporate rules, must also be aligned to rules in the TIEC and to the technology. Therefore, I conclude that the specific phase also is characterized by a particular dominant business model that is aligned with the TIEC of that era. Thus, a transitional phase may involve adaptations and innovations in the business model as much as (or rather than) the product design and the dominant business model of a specific phase may represent a recent innovation for that time.

Section 5.7.5 reviews the characteristics of innovation during the first specific phase and the second transitional phase. Contrary to the description of the industry life-cycle model, I conclude that there was no clear industry-wide trend toward process innovations and away from product innovations during the first specific phase, nor was there a reverse trend during the second transitional phase. I also find no clear trend in the shifting importance of the mechanisms of learning: learning-by-interacting with suppliers remained the most prominent mechanism of innovation during both phases. Innovation also remained synthetic, involving collaborations by teams of entrepreneurs and between manufacturers and suppliers. Finally, I find that innovations in business operations played a critical role and should not be over-looked in studies of technological progress and transitions.

Section 5.7.6 discusses the process of embedding, alignment and stability during the first specific phase and second transitional phase. I find that the development of 'hard' supporting infrastructure lagged behind the expansion of markets for the automobile and did not become a significant factor in stability until the second transitional phase. Beginning in the latter part of the first specific phase, manufacturers also began using consumer financing to alter the selection environment, an institution which aligned with and reinforced trends in the TIEC. The stability of the first specific phase was undermined by the post-war recession and evolving consumer preferences that became misaligned with the existing dominant product design. This development was marked by feedbacks and co-evolutionary effects. The emergence of a new dominant product design brought the product and rules (consumer preferences) into realignment. Finally, the introduction of TEL was challenged by emerging trends in the TIEC, but the innovation was successful due to the immaturity of institutions governing public health and GM's activities to shape the selection environment.

Finally, section 5.7.7 summarizes the events during the time period covered by this chapter and the findings of this analysis.

5.7.1 Vertical Integration

An industry trend toward vertical integration began even prior to the onset of the first specific phase. Henry Ford began this process in 1905 when he established the Ford Manufacturing Company to manufacture parts for the Model N (see section 4.1.6). William Durant pursued vertical integration beginning in 1907 at Buick with the acquisition of the Weston-Mott Axle Company and continuing in 1908 at GM with an aggressive acquisition strategy that resulted in partial or complete ownership of a multitude of suppliers (see sections 4.2.1 and 4.2.2). This process continued throughout the specific phase and, by the onset of the second transitional phase in 1918, both Ford and GM were large, vertically integrated companies. In chapter 4, I identified learning-by-interacting with suppliers as a critical mechanism of innovation in the transitional and early specific phases. Therefore, vertical integration and the resulting corporate structure had significant implications for the ability of these industry leaders to innovate and adapt during the second transitional phase.

This section examines the factors contributing to the trend toward vertical integration and discusses the impact of vertical integration on the innovative capabilities of Ford and GM. I find that three factors contributed to vertical integration in the automobile industry: 1) a trend in the TIEC toward large industrial combinations that could influence the price and supply of raw materials; 2) process innovations that required continuous plant operation and carefully timed delivery of parts; and 3) tacit knowledge that existed only in human capital. I also conclude that vertical integration in itself did not necessarily stifle innovation, but that the influence of integration on innovation depended on the specific strategies pursued and the resulting rules and capabilities that were embodied in corporate structure, operating procedures, facilities, and human capital.

The traditional explanation of the trend toward vertical integration asserts that it was more economical for manufacturers to make their own parts. According to Smith (1968, p. 110), by 1929 "it was still possible to become an automobile manufacturer by buying components and assembling them, but it was impossible to cut much of a figure this way or to survive in competition… With competition at fever heat, the manufacturer could ill afford to share any profits with suppliers. Companies were growing more vertical as well as horizontal. The rich were getting richer; the poor were getting poorer." However, parts suppliers generally could achieve the same low costs that manufacturers could through economies of scale, especially if guaranteed volume sales through long term contracts. Meanwhile, purchasing parts obviated the need for investments in plants and equipment, an important consideration in a rapidly growing industry. On the other hand, the price of raw materials – coal, iron ore, and steel – was a concern since suppliers were large industrial combinations whose control of a large portion of the market allowed for price fixing. Thus, the wide-spread trend toward combinations and price fixing was an important contextual factor for the industry. But availability was a bigger concern since supplies were restricted several times due to labor strikes in the railroad and mining industries.

According to Epstein (1972), vertical integration in the early automotive industry became attractive because it ensured an adequate and timely supply of materials and parts, and therefore facilitated continuous plant operation. By the early 1900s, demand for automobiles was so great that manufacturers were not concerned with market share, but rather with adequate production volume and prompt delivery of parts. Olds, the first mass producer of automobiles, built 3,299 cars in 1902, but could have sold 4,000 if suppliers had been able to fill his orders. A few years later, Buick found that "[t]he public would take Buicks as fast as they could be turned out; delay in delivery of even a minor part might cost a tremendous sum" (Pound, 1934, p. 87). Suppliers of parts and raw materials were often the cause of plant downtime at Ford, and this became more problematic once continuously moving assembly lines were fully implemented and timing became crucial. As late as 1934, Pound (p. 87) reported that "in this industry utter dependence on certain forms of goods was so essential that practically all the survivors in the stern battle for existence waged during the past thirty years are those who have been working toward self-determination, seeking positions where their operations could not be shut off by shrinkage of those essential supplies."

Another factor in the move toward vertical integration was the appropriation of knowledge. Suppliers of key components possessed singular expertise that, for the most part, resided in human capital. Through partial ownership in the Ford Motor Company, Ford established a relationship with the Dodge brothers to provide nearly complete vehicles, excluding only bodies, wheels, and tires. This move was critical in Ford's early success. The Dodge brothers brought necessary start-up capital, but as manufacturers of engines for the only mass produced vehicle to date, they also brought scarce knowledge. After Henry Ford made clear his intention to bring this expertise in-house, the Dodges successfully launched their own manufacturing firm. Similarly, GM's partial, and later complete, ownership of the Fisher Body Company was a strategic move to secure exclusive access to the Fisher brothers' unique skill and knowledge in closed body production.

Ford and General Motors pursued significantly different strategies for vertical integration, resulting in different corporate structures and different emphases for innovation. Henry Ford brought nearly all production, including processing of raw materials, in-house at the Rouge complex. He established complete, centralized, financial and managerial control. Without the tempering effects of a board of directors or executive committee, strategic decisions were made entirely according to Ford's personal goals, beliefs, values, and perceptions. The rules governing interactions at Ford included those established for Fordist production that were embodied in standard operating procedures, routines, and capital equipment, and one additional overarching rule: defer to Henry Ford. Employees unable to comply were forced out or left in frustration. As a result, innovation at Ford had stagnated by the late 1910s, just as a new transitional phase was beginning.

Meanwhile, GM purchased all or part ownership of suppliers and initially allowed them to continue operating as independent firms. During this time, innovation continued on whatever level set by individual company policies. There was no single identifiable set of rules for the corporation as a whole, except perhaps for the understanding that, during the years of his tenure, William Durant could interfere at will. GM's strategy of acquisition, along with Durant's ultimately destructive actions in the stock market, led to corporate associations and the appropriation of human capital that defined GM's innovation capabilities and strategies for the next fifty years. The most significant of these associations were those with the E.I. Du Pont de Nemours and Fisher Body

Companies and with Pierre du Pont, Alfred Sloan, and Charles Kettering. Collectively, the decisions of Durant, du Pont, and Sloan were of critical importance in shaping GM's strategies and success in the second transitional phase.

The financial and managerial reforms achieved between 1920 and 1925 finally established a coherent set of rules for GM that was embodied in corporate structure and policies regarding operations, communication, and financial reporting. After du Pont and Sloan began coordinating divisional operations through corporate policy, their establishment of decentralized control allowed for a greater level of innovation than what was realized at Ford. This innovation was largely realized through IR&D and was institutionalized within the Research Corporation. However, the establishment of financial controls to coordinate management of the divisions, the separation of R&D from operations, and the establishment of corporate policy on product development likely inhibited the rate of innovation realized in production automobiles. Dodge, and later Chrysler, would emerge as industry leaders in incorporating product innovations. Neither firm pursued vertical integration to the extent that Ford and GM did.

5.7.2 Initiation of a Second Era of Ferment

Beginning with the onset of conflict in Europe, four factors created misalignments that undermined the stability of the specific phase. These factors resulted in a shift in consumer preferences that triggered a new transitional phase: 1) World War I and the post war recession; 2) market saturation; 3) the availability of inexpensive all-steel closed bodies; 4) improved roads; and 5) the increasing market influence of women. These factors include exogenous events, innovations, and issues associated with the very success of the technology and the development of supporting infrastructure. These initiating events were marked by pervasive co-evolutionary effects.

In general, the industry first felt the challenges of the transitional phase in the middle of 1920 when the post-war recession sent sales plummeting. However, the stage had been set by the middle of 1918 when diversion of facilities to the war and a steel shortage limited production to 53% of the level achieved in 1917. In fact, despite banner sales in 1920, average sales for the years 1918-1920 were still lower than total sales in 1915. Therefore, I will mark the beginning of the transitional phase in 1918.

While the war and ensuing recession were exogenous events, the response of the industry during and after the war created feedbacks that magnified their impact on the industry. First, using trucks to deliver war supplies to the ports demonstrated the feasibility of long-haul trucking and year-round driving. This provided the impetus for increased federal funding for improved roads, a process already underway. Second, the industry (correctly) perceived that decreased production during the war had resulted in pent up demand for automobiles. When the post-war recession hit, companies that had invested in aggressive expansion plans had an especially difficult time coping with the sudden drop in sales. Many firms did not survive the crisis.

Meanwhile, market saturation was a result of the industry's past success. The industry had to either scale down production to serve a stagnant market consisting of replacement sales or find an innovation that would stimulate accelerated replacement or create demand in a new market. In choosing to innovate, a firm would be taking on the nexus role, actively using variation to alter the selection environment in terms of consumer preferences, and thereby altering the functional requirements of the automobile. However, consumer demands were already shifting due to other factors. Improved roads and affordable closed bodies were increasing the feasibility of year-round driving, and all-steel closed bodies were more durable and comfortable. Women were driving more and influencing more sales thanks to closed bodies, electric self-starters, and shifting cultural patterns that were accelerated by the war. The industry's response to these shifts – an increasing emphasis on style and comfort and frequent style change – would be self-reinforcing.

5.7.3 Firm Strategies for Adaptation

This section summarizes firm strategies in responding to the changing environment of the transitional phase. The resources available to firms depended on institutional linkages, corporate policy, and access to human as well as financial capital. Therefore, I find that firm histories, the personal beliefs and perspectives of firm leaders, and corporate structures and culture – *i.e.*, firm-specific rules – shaped individual firm strategies for innovation and their adaptive capabilities. Because of industry concentration, firm-specific rules had a large and rapid impact on the development of the technology and the industry as a whole. This supports the use of a behavioral approach with a micro-focus in the analysis of innovation and socio-technical change. In addition, I find that unexpected events presented insurmountable difficulties for some firms, while other entrepreneurs were able to exploit the opportunities presented by those firms' failures.

GM was one of the firms caught in midst of aggressive expansion when the recession struck. Durant's market activities only made matters worse. But because of GM's prior acquisition strategy and Durant's past blunders, the corporation was also lucky to have managerial and technical talent in Alfred Sloan, Pierre du Pont, and Charles Kettering. These activities also resulted in ties to the E.I. Du Pont de Nemours Company that gave GM access to that company's war surplus. With these resources, the corporation was able to respond aggressively to the challenges before it. Durant was ousted, the corporation was reorganized, and the new management instituted financial controls. With the financial crisis partially under control, GM began to address market saturation, which meant fierce competition with other manufacturers in a market of relatively fixed size. GM rationalized the product line and attempted to penetrate the low-price market through product innovation (the copper-cooled engine).

When the new technology failed, the economy was already recovering, so GM instead redesigned the conventional Chevrolet. GM was successful in challenging Ford, not because of a new innovation, but because of the company's understanding of the importance of an existing one in a new era. According to Sloan (1964, p. 160), "The last decisive element in this competition, I believe, was the closed body, which itself was by far the largest single leap forward in the history of the automobile since the basic car had been made mechanically reliable."

This understanding of changing conditions, enabled in part by systematic study of the market, led to GM's next three-part strategy: pursuing national advertising, emphasizing style, and making frequent model changes. Although the corporation did not formulate style and model change policies during the transitional period, the management structure that Sloan instituted was inherently responsive to market conditions. Thus, by the time designing for style and planning for regular change were institutionalized, the annual model change was already a *de facto* policy.

GM's final and arguably most significant strategic move during the transition was born from the effort to rationalize the product line and fill an existing gap. First, by offering "a car for every purse and purpose," GM essentially re-introduced product differentiation and turned a single mass market, where competition was based on price, quality, and options, into a group of niche markets defined by socio-economic status as well as functional requirements. Within each niche, competition was based on price and quality, and GM sought to place itself near the top in both. Then, when adding the Pontiac in 1926, GM decided to share chassis and body parts across divisions and price categories to provide product variety while preserving economies of scale. GM now had a competitive alternative to Fordism.

GM had turned the crises of the post-war period into entrepreneurial opportunity. In 1927, Moody's Investors Service praised the results: GM had "reorganized its system of buying raw materials, harmonized its factories, reformed its sales methods, improved it advertising, organized a department to study public demand, and developed the greatest automobile research organization in the world. In so doing it had approximately doubled its rate of profit" (Nevins and Hill, 1957, p. 403).

When the recession hit, Ford faced the added challenge of being in debt from the stock buyout and court ordered dividend. Ford suspended production at Highland Park and the Rouge. During the shutdown, Ford streamlined the company's organization and improved coordination of supplies, but he also ruthlessly slashed the payroll. While GM was departing from past practice and building an effective management structure, Ford was holding the line and eliminating vital business functions. Ford responded to the post-war crisis by clinging to the Model T, continuing to reduce prices, and pursuing ever deeper vertical integration. His pursuit of vertical integration was in part due to price fluctuations encountered leading up to and during the war, though such fluctuations had always been a feature of the industry. But by staying with previously successful strategies and taking them to an even greater extreme, Ford illustrated the folly of what Gharjedaghi (2006) calls inertia and sub-optimization. The strategy did nothing to address saturation, shifting preferences, or other manufacturers' introduction of frequently restyled closed-body cars.

While Flink (1988, p. 240) asserts that "[t]he annual model change and diversity of product were incompatible with Fordist production methods," the problem went far deeper. The annual model change and diversity of product were incompatible with Henry Ford's values. In 1907, he had been able to perceive the market potential for the Model T because he not only knew the farmer's needs, but because the farmer's values were part of his psyche: hard work, thrift, and resourcefulness. Ford had always known that the Model T would someday need to be replaced and he was aware of market trends. But the market shift in the 1920s was near anathema. Ford's beliefs and perspectives, his personal rules, had served him well in the emergent and specific phase, but now were misaligned with changes in the definition of the technology and other developments in the TIEC. Because of his intolerance for frills and conspicuous consumption, he continued to believe that automobiles were and *should* be service vehicles. Unlike GM's Sloan, Ford was unwilling to *accept* that the functional definition of the automobile had been altered from a technology for basic personal transportation to one that provided transportation service, demonstrated social status, and provided satisfaction from personal ownership.

However, the misalignment of Ford's strategies with the trend toward market saturation is harder to explain. Ford's stated intention was to build a car that, for all practical purposes, would last forever (Ford and Crowther, 1922). Such a strategy would lead inevitably to a saturated market and stabilized sales. Ford could only have hoped to maintain market share and therefore stable production levels. However, the industry response of the annual model change, led by GM, not only accelerated replacement sales to maintain market growth, but also eroded Ford's market position.

The reality of the market shift was inescapable, especially after other manufacturers began to feed demand for style with frequent model changes. In order to accommodate a complete product change to the Model A, Ford had to institute a complete shutdown, scrap most of the company's specialized machine tools, and develop and purchase new process equipment. Inertia and suboptimization cost Ford dearly in terms of lost production, capital outlays, and the loss of critical human capital that could have increased the firm's adaptability.

While many smaller firms were unable to survive the financial crisis of the post-war recession, Walter Chrysler was able to exploit the opportunity this presented. Chrysler was able to convert the Maxwell and Chalmers companies from failing enterprises in 1920 into the third largest manufacturer by 1928. In doing so, he also created an opportunity out of unexpected events that were disastrous for the Dodge brothers. The influenza epidemic in 1920 took the lives of the company's founders, leading eventually to the sale of the Dodge Brothers in 1928. In buying Dodge, Chrysler increased its volume and sales force enough to operate much like the industry leader, GM.

Walter Chrysler was successful mainly because of his strategy to maintain a position of technological leadership, an approach that also had served Dodge Brothers well.[71] Though Chrysler did not necessarily originate many major innovations, the company was the first to incorporate a number of advanced technologies in a moderately-priced car. Much like GM, Chrysler collaborated with suppliers to improve technologies, established a well-funded research laboratory, offered improved models annually, developed several distinct lines of vehicles, and established dealer and consumer financing.

5.7.4 Dominant Design and Business Model

This section reviews the emergence of a new dominant product design that arose from co-evolutionary forces. However, I find that production processes and management system, *i.e.* corporate rules, must also be aligned to rules in the TIEC and to the technology. Therefore, I conclude that the specific phase also is characterized by a particular dominant business model that is aligned with the TIEC of that era. Thus, a transitional phase may involve adaptations and innovations in the business model as much as (or rather than) the product design and the dominant business model of a specific phase may represent a recent innovation for that time.

The "dominant design" is not a static concept, but rather a temporarily stable description of the majority of the market. Even during a specific phase, it co-evolves with TIEC in a gradual, continuous fashion. During a stable specific phase of sufficient duration, a product's functional description and attributes may evolve considerably,

[71] Interestingly, after leaving Ford, C. Harold Wills attempted this same strategy in designing the Wills-Saint Claire introduced in 1921. The car was an excellent work of engineering, but was years ahead of its time. The engine was so complicated that no ordinary mechanic could repair it and the vehicle was a market failure. C.H. Wills & Company went into receivership in 1922 (Sorensen, 1956; Rae, 1959).

appearing radically changed after decades without any disruptive changes in the TIEC or the industry. During the first specific phase, from roughly 1910 to 1918, the functional definition of the automobile was stable but the product characteristics used to fulfill them evolved. The most significant change was in engine technology. In 1912, 91% of new automobiles sold included four-cylinder engines while only 8% were six-cylinders; by 1920, the fractions were 73% and 25% and the trend toward six-cylinders continued through the next decade.

As discussed in section 5.7.2, a new transitional phase was initiated around 1918 by WWI and other changes in the TIEC. It is difficult to pinpoint exactly when a new dominant design emerged and marked the onset of a new specific phase. However, it is clear that before Ford discontinued the Model T in 1927 the automobile had been functionally redefined. I therefore place the emergence of a dominant design around 1926. By that time, providing basic reliable transportation to the masses was no longer enough. In the earlier days, when the expense of motor vehicles restricted ownership to the wealthy, owning *any* automobile conveyed social prestige. But with ownership attainable by the masses, it now mattered *which* vehicle a person drove. To all who saw it, the automobile suggested that its owner had attained a certain level of income and social status. And much like a work of art, the automobile designed for style had an aesthetic value that yielded personal satisfaction. According to Epstein (1972, p. 58), "[l]ike a handsome piece of furniture, a picture, or an artistic piece of jewelry, it not only is used but it reaches a deep-seated sense of ownership."

The automobile in 1926 was functionally defined as a technology to provide year-round, comfortable, rapid, personal transportation while displaying social status and imparting the pleasure of personal ownership. The product attributes that fulfilled the requirements of this definition were an all-steel, closed-body car, powered by a four- or six-cylinder gasoline internal combustion engine. However, the trend toward larger engines continued in the second specific phase with eight-cylinder engines eventually becoming dominant.

The product strategy that GM adopted in 1923, "a car for every purse and purpose," was the ultimate expression of this new definition. When a buyer traded in a Chevrolet for an Oakland (later Pontiac), he was climbing the economic and social ladder. Owning a Cadillac said that you had 'arrived.' The institutionalization in 1927 of designing for style completed GM's transition to the new era. While Ford and others criticized GM for artificially creating demand through planned obsolescence, they overlooked the fact that GM did not invent the idea of the car as a style product. No manufacturer's actions caused the market change; consumer preferences and automobiles co-evolved with the TIEC.

The onset of a second transitional phase taught the industry hard lessons in planning for and accommodating change. Not only did the new definition of the automobile dictate continual change, but in a co-evolutionary system, change is the rule and adaptation is necessary for survival. According to Hounshell (1984, p. 264), "The Ford Model T dictum of maximum production at minimum cost gave way to planning for change." This new dictum meant new product strategies, new business organization, and new systems of production. 'Rigid' mass production gave way to 'flexible' mass production.

Thus, a second but related distinguishing characteristic of the specific phase becomes apparent from this research: the existence of a dominant business model. Rycroft and Kash (1999) assert that as industries mature, production processes evolve from craft to mass to synthetic production. However, this research indicates that what I have called a specific phase is characterized by a particular business model that includes the production process and organizational structure. Thus, a transitional phase may involve adaptations and innovations in the business model as much as (or rather than) the product design. The production methods seen during the specific phases reviewed in this research each were new innovations at the time, as was the organizational structure originated at GM in the 1920s. During the first specific phase, the dominant business model entailed "Fordist production" (volume production; standardized fully-interchangeable parts; special or single purpose machinery; and continuously moving assembly), and centralized control. The dominant business model of the second specific phase entailed flexible mass production, staff and line organization, and coordinated de-centralized control using financial controls. There is good reason to expect that in other eras, specific phases will entail yet other business models with production systems and organizational structures that are recent innovations for their time, for example lean production and the Toyota Production System.

5.7.5 Characteristics of Innovation

This section reviews the characteristics of innovation during the first specific phase and the second transitional phase. Contrary to the description of the industry life-cycle model, I conclude that there was no clear industry-wide trend toward process innovations and away from product innovations during the first specific phase, nor was there a reverse trend during the second transitional phase. I also find no clear trend in the shifting importance of the mechanisms of learning. Learning-by-interacting with suppliers was the most prominent mechanism of innovation during both phases, while spillovers were and learning-by doing were also important and R&D became a major factor for the first time during the transitional phase. Innovation remained synthetic, involving collaborations by teams of entrepreneurs and between manufacturers and suppliers. Finally, I find that innovations in business operations deserves mention and should not be over-looked in studies of technological progress and transitions.

The product and industry life-cycle model predicts that innovation during the specific phase, which for the automobile first occurred between roughly 1910 and 1918, will shift emphasis from products to processes. This was certainly true at Ford, which constituted only 10% of the market in 1910 but roughly 50% between 1914 and 1919. It should be recalled, however, that Ford's major process innovations, which actually began prior to the onset of the specific phase, involved a great deal of concurrent product innovations. Further, a trend toward process innovation was not characteristic of the industry as a whole and a number of significant product innovations were introduced during this timeframe, including the large volume production of closed body cars (1910), the all-steel body (1912), the electric self-starter (1911), and mass production of the all-steel body (1914).

This conclusion is supported by Klepper and Simons (1997) who analyze the list of automotive innovations from Abernathy, *et al.* (1983). Using a five year moving average (shown in Figure 5-2), the authors find no clear trend in product innovation from

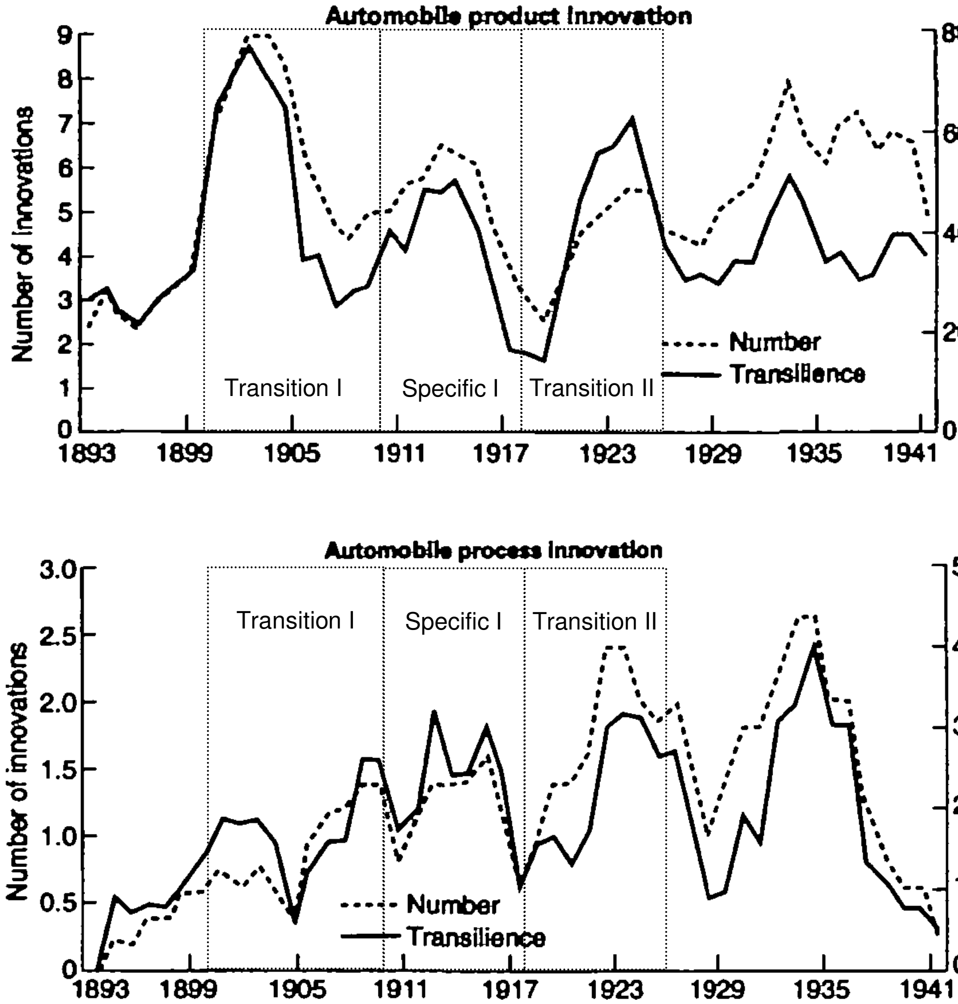


Figure 5-2 Automotive Product and Process Innovation
Source: Reproduced from Klepper and Simons (1997) with phase depiction by author.

1893 through 1940, both in terms of the number of innovations and their transilience index – a subjective measure of the innovation's impact on production systems and market linkages.[72] However, the authors do find a generally increasing trend in the

[72] Abernathy et al. (1983) define transilience as a measure of the overall competitive impact of an innovation, including disruption to production systems and market linkages. The authors use a seven point scale to assign transilience scores, based on their judgment, to each innovation in an extensive list covering the period from 1893 to 1981. A value of 7 represents innovations that were "very disruptive for products or processes." See Abernathy et al (1983) p. 109-110 and p. 153. Klepper and Simons (1997) calculate the sum of the squared transilience scores for all product or process innovations in each year then graph the five-year moving average to smooth out year-to-year variations.

number and transilience index of process innovations during that timeframe. Interestingly, product innovations generally occurred more frequently (averaging between 2-9 per year versus 0-2.5) and were more disruptive to production (transilience roughly 16-80 versus 0-40).A few points are worth noting on examination of Figure 5-2, which shows the phases identified in this research overlaid on Klepper and Simons' graph. First, midway through the first specific phase (1914), product innovation was indeed waning and was then reinvigorated by the end of the post-war recession (1921). This is consistent with the life-cycle model and the dates I have drawn to delineate the phases, but the specific phase lasted only another four years, which provides very few data points from which to draw conclusions. If the specific phase had lasted longer before being interrupted by the war, the decline might have proved to be merely a brief slump.

In addition, WWI diverted productive and innovative resources as automobile manufacturers were involved in both the development and manufacturing of military equipment. For example, Packard engineers were largely responsible for the design of Liberty aircraft engines and Ford developed innovative manufacturing processes for the engine cylinders. This diversion of resources would have delayed the introduction of new automotive innovations even if events and changes in the TIEC had not triggered a new transitional phase. Thus, the war in itself could explain the pattern of product innovations shown in Figure 5-2 between 1917 and 1919. If, for these reasons, the years during the war and the ensuing recession (1918-1921) are excluded, the transilience index of process innovations between the first specific and second transitional phases is roughly constant, though there is an overall increase in their number. Therefore, while the work of Klepper and Simon supports the assertion that product innovation stagnates during the specific phase and is re-invigorated with a new transitional phase, it is not conclusive. Further, it fails to provide evidence of a trend toward greater emphasis on process innovation during the specific phase.

As we have seen, Ford's focus on process innovations after 1908 was the result of Henry Ford's personal rules (goals, values and perceptions) and not necessarily an inherent feature of industry or firm maturity. Ford pursued both vertical integration and centralized control and after 1913, product innovation was stifled in preference for control of supply lines and a singular focus on production costs. Meanwhile, innovation was institutionalized at the corporate level in GM and later strengthened considerably with the purchase of Kettering's enterprises in 1918.

Much of the significant product innovation during the specific phase between 1910 and 1918 originated with suppliers, including: large-scale production of closed bodies by Fisher for Cadillac, the all-steel body by Hale & Kilburn for Hupmobile, and the electric starter by Delco for Cadillac. Dodge Brothers was incorporated during this timeframe and quickly established itself as an innovative leader by incorporating Budd's all-steel body. Budd also contributed a significant process innovation – electric spot welding – that served as an enabling technology for all-steel bodies. Thus, interactions with suppliers continued to play a critical role in development, as did spillovers. Budd's innovations were highly synthetic, representing the culmination of years of teamwork beginning at Hale & Kilburn and in collaboration with Emil Nelson, the chief engineer of the Hupp Motor Car Co.

During the second transitional phase (1918-1926), most of the significant product innovations – four-wheel hydraulic brakes, balloon tires, and all-steel closed bodies – again originated with suppliers. Duco lacquer, primarily a process innovation for the automobile industry, also originated with a supplier and, though it was a product of a guided research, the particular application largely resulted from serendipity. Another significant innovation during the second transitional phase was the development of leaded gasoline (TEL) at GM's Research Corporation. Although part of a guided research program, the breakthrough was largely one of trial and error. While it did not represent an innovation in automobile technology, it solved a major industry problem and enabled the use of engines with higher compression ratios. It is also the first instance where environmental factors became an issue for the industry.

Perhaps most interesting during this phase is the importance of innovations in business operations which are generally overlooked in studies of technological progress. GM's "car for every purse and purpose" market segmentation was a brilliant solution to market saturation and an emerging functional definition (see section 5.7.4). The corporation's later decision to standardize parts across divisions, an innovation in both product and process, enabled this market segmentation and was equally revolutionary for the industry and market. Other significant business innovations include planning for annual model changes, design for style, and the use of time sales.[73] While planning for change may be considered an innovation in business operations, it also required important changes to production processes, specifically in capital equipment, which can be attributed to William Knudsen.

[73] Time sales were introduced during the first specific phase but became a dominant feature of the industry in the second transitional phase.

Market segmentation and annual model changes designed for style were innovations based on a fundamental understanding of emerging user preferences, which GM tracked with rather sophisticated data analysis. At the same time, the appearance of stylish new vehicles every year, marketed as symbols of status, reinforced and fed these trends in preferences. These innovations, along with the use of consumer financing can be considered as efforts to alter the selection environment by accelerating replacement sales and extending the market to more consumers. Thus, the new functional definition of the automobile, the new design attributes that fulfilled them, consumer preferences, and financing institutions arose in a co-evolutionary fashion.

Finally and perhaps most significantly, GM's innovations in corporate structure and operations established a new standard business model for industry in general (see section 5.7.4). These innovations resulted from the collective and synthetic efforts of Pierre du Pont, Alfred Sloan, Donaldson Brown and Albert Bradley. While many of these innovations resulted from in-house efforts, largely learning-by-doing, the entrepreneurs themselves were associated with GM due to past supplier relationships and financial difficulties and brought with them experience gained in other industries.

5.7.6 Alignment and Stability

Chapter 4 discussed how the automobile, scientific management and Fordism were aligned with evolving higher-level rules within the TIEC. This section summarizes the the process of embedding which continued through the first specific phase and became more pronounced during the second transitional phase. I find that the development of 'hard' supporting infrastructure lagged behind the expansion of a mass market for the automobile and did not become a significant factor in stability until the second transitional phase. Beginning in the latter part of the first specific phase, manufacturers also began using consumer financing to alter the selection environment, an institution which aligned with and reinforced trends in the TIEC. The stability of the first specific phase was undermined by the post-war recession and evolving consumer preferences that became misaligned with the existing dominant product design. This misalignment was due to: 1) an exogenous factor – the increasing influence of women in the automobile market; 2) two factors related to the technology's success – market saturation and road improvements; and 3) an innovation – the all-steel closed body. Changing preferences were also reinforced by the industry's response to them as well as to the war and the challenge posed by market saturation. The emergence of a new dominant product design and business model brought the product and rules (consumer preferences) into realignment. Finally, the introduction of TEL was challenged by emerging trends in the TIEC, but the innovation was successful due to the immaturity of institutions governing public health and GM's activities to shape the selection environment.

Physical infrastructures for roads and refueling began periods of significant growth late in the first specific phase. Federal funding for roads began in 1916 and was significantly increased in 1921. This development was supported by the initiation of gasoline taxes in 1919, which became the dominant source of state funding for main roads by 1929. The availability of gasoline from dispersed and non-dedicated sources prior to the turn of the century was a factor in the success of gasoline-powered vehicles in early markets. A physical infrastructure for refueling gasoline motor vehicles began development with the introduction of curb pumps and dedicated stations in 1907.

Between 1915 and the early 1920s, the curb pump was the dominant form of refueling, but growth in dedicated stations began around this time. By 1930, dedicated service stations were overtaking the curb pump both in terms of number and sales volume. Thus, the development of roads and filling stations, the 'hard' infrastructure that still exists today, lagged behind the expansion of a mass market for the automobile. In fact, both entered a period of substantial growth just as the automobile industry was entering the second transitional phase. Therefore, physical infrastructure did not become a significant factor in stability until the second transitional phase.

In 1910, bankers disapproved of loans for automobiles, which they still considered a luxury. Beginning in 1916, captive lending agencies began financing time sales of automobiles and this became an important strategy in adapting to the challenges presented by the second transitional phase. The implementation of consumer financing by dealers and manufacturers can be considered efforts to alter the selection environment by extending the market for new vehicles to more consumers and by improving the desirability of used cars relative to inexpensive new cars, particularly the Model T. Such financing had been limited to furniture, sewing machines, pianos, and homes, but the concept became a mainstay of the U.S. economy in the late 1920s. Thus, the use of time sales for automobiles aligned with and reinforced trends in the TIEC.

The stability of the first specific phase was undermined by an exogenous event – the post-war recession – and by evolving consumer preferences that became misaligned with the existing dominant product design. The increasing influence of women in the automobile market represented an exogenous factor in this misalignment. Two other factors were due to the technology's success – market saturation and road improvements.

The fourth factor was an innovation – the all-steel enclosed body. While continuing road improvements during the transitional phase should have served to support the industry, they were destabilizing as long as the dominant design remained consistent with the poor roads of the earlier era. The Model T design in particular lagged behind the evolution the supporting infrastructure. Road improvements and the trend toward year-round use, both of which were reinforced by events during the war, along with saturation and the increasing role of women dictated new design requirements. Once the automobile had been redefined and a new dominant design had emerged, the product and rules (consumer preferences) became realigned.

During the transitional phase, one additional misalignment arose. The introduction of tetra-ethyl lead (TEL) represents the first time that the environmental health effects of automobile use became a prominent issue. However, in 1925, institutions to protect public and worker health were still in their infancy despite the work of Progressives (see section 3.4.1), whose influence waned in the post-war era. While university experts and the federal Public Health Service (PHS) raised the TEL issue to national attention, the federal government had no authority to regulate chemicals. Meanwhile, Thomas Midgely and GM actively worked to refute the claims of health experts and calm the public, despite suffering from lead poisoning himself. Thus, while the use of TEL was misaligned with a rising national consciousness about worker and public health, no formal legislative rules were in place to block the use of the additive. The PHS study was unable to make conclusive statements regarding the health effects of long-term exposure, and the courts found Standard Oil innocent of criminal negligence in the TEL plant accident. Therefore, GM's efforts at altering the selection environment were successful as the controversy faded. However, the movement in the TIEC toward greater concern with public and worker environmental health would later create more serious misalignments with automobile and fuel use.

5.7.7 Summary

The first specific phase of the automobile industry extended from roughly 1910 to 1918. Though product attributes evolved slowly, the functional definition of the automobile remained stable during this phase and was aligned with rules within the TIEC. The specific phase was characterized by a dominant business model that was also aligned with the TIEC. Firms whose operations or products were less aligned struggled financially. Although this alignment supported the stability of the specific phase, physical infrastructure lagged behind the expansion of the market and became more of a factor during the second transitional phase.

The stability of the first specific phase was undermined by factors that can be placed in three categories:

1) exogenous factors: WWI, the post-war recession, and the increasing influence of women in the automobile market;
2) the technology's success: market saturation and road improvements; and
3) innovations: the inexpensive all-steel closed body.

These factors influenced the industry through the actions of government agents, economics (the price and availability of raw materials, capital, and labor), and changing consumer preferences. As a result, the functional definition and attributes of the dominant design from the first specific phase were no longer aligned with consumer preferences.

Up through the onset of the second transitional phase, three factors contributed to increasing vertical integration in the automobile industry:

1) a trend in the TIEC toward large industrial combinations that could influence the price and supply of raw materials;
2) process innovations that required continuous plant operation and carefully timed delivery of parts; and
3) tacit knowledge that existed only in human capital.

Vertical integration had a great influence on each company's innovative focus and capabilities, but this influence depended on the specific strategies pursued to achieve integration and the resulting rules and capabilities that were embodied in corporate structure, operating procedures, facilities, and human capital. Therefore, during the second transitional phase, I find that firm histories, the personal beliefs and perspectives of firm leaders, and corporate structures and culture – *i.e.*, firm-specific rules – shaped individual firm strategies for innovation and their adaptive capabilities. In addition, I find that unexpected events presented insurmountable difficulties for some firms, while other entrepreneurs were able to exploit the opportunities presented by those firms' failures.

There was no clear industry trend toward process innovation versus product innovation during the specific phase, and no reversal of that trend during the transitional phase that began in 1918. Rather, individual firms pursued specific strategies for innovation. For example, Ford focused on reducing costs and controlling supplies, consistent with company strategy and the dictates of Fordism. I also find no clear trend in the shifting importance of the mechanisms of learning. product innovation occurred largely through interaction with suppliers who generally initiated these novelties. GM's partial ownership of its suppliers created a partner relationship and fostered a higher rate of product innovation than was found at Ford. However, newcomer Dodge Brothers quickly assumed the role of industry leader in product innovation with the incorporation of all-steel bodies. This role was taken on by Chrysler in the late 1920s. Innovation also remained synthetic, involving collaborations by teams of entrepreneurs and between manufacturers and suppliers. In addition, innovations in business operations played a critical role in the second transitional phase and should not be over-looked in studies of technological progress and transitions.

The new dominant product design of 1926 arose from co-evolutionary forces, with industry responses to the challenges of the transition and new product offerings reinforcing trends in consumer preferences. The new product definition that emerged by 1926 was misaligned with Fordism and dictated a new dominant business model – "Sloanism" and flexible mass production. Thus, the specific phase is characterized by a particular dominant business model that is aligned with the TIEC of that era and a transitional phase may involve adaptations and innovations in the business model as much as (or rather than) the product.

By the late 1920s, the technology and dominant business model were aligned consumer preferences. With increasingly aligned institutions and the addition of a dedicated physical infrastructure, the gasoline automobile had become part of a stable socio-technical regime.

6.0 Results and Conclusions

…the car is not one invention but a mechanical creation composed of hundreds, if not thousands, of inventions. In truth, we are still inventing the car, for the car is an ever-changing assembly of ideas, systems and parts.

-- J. Barach, 2007

6.1 Introduction

As discussed in chapter 2, the quasi-evolutionary (QE) and multi-level perspective (MLP) theories assert that niches play a fundamental role in socio-technological transitions. Niches are understood to be distinct application domains, small in scale and scope, that are characterized by specific functional requirements. Because of the special functional requirements of the niche, users are willing to accept higher cost or lower performance on other non-essential functionalities compared to those attained with existing products available in mass markets. Thus, products tailored to niches are not competitive in other niches or mass markets because the functional requirements are not comparable across markets. For a product to compete in a mass market, which is large in scale, it must provide a functionality with wide appeal and compete with alternative products based on price, quality, and performance, which can be objectively compared.

The QE and MLP theories describe transitions as a process of niche accumulation, niche branching, or niche proliferation. The development or penetration of new niches allows a technology to expand the overall scale of its markets. Eventually, the technology achieves cost reductions through economies of scale and learning effects that allow it to conquer a mass market. This phenomenon relies on learning processes, primarily learning-by-doing and learning-by-interacting with users, within niche markets that alter the technology and the systems in which it is produced and used. As a result, a novel technology becomes embedded in a new rule set through the mutual adaptation of the technology and institutions, where institutions are understood to include organizations and rules. Rules may be cognitive, normative, or legislative, and are defined as the complex set of constraints that shape interactions among agents and between agents and the physical elements in the system. The co-evolutionary process of embedding results in the alignment of the technology, institutions, and the physical artifacts of the system, forming a stable system or regime. The embedding process arises from the interaction of agents through institutional linkages in what QE theory calls the technological nexus. Agents in the nexus role link selection and variation by translating information from both realms; shaping interactions; and harmonizing social and market needs with the results of scientific and technological research.

Chapter 1 posed a set of research questions related to the assumptions underlying this niche hypothesis; the role of the technological nexus in niche dynamics and the diffusion to wider markets; and regime dynamics and stability. For reference, these questions are repeated in section 6.2. The following sections directly address each research question and discuss related findings of this research.

Section 6.3 summarizes the role of niche markets in both the emergence of the motor vehicle and its diffusion to mass markets. I find that niche markets did indeed play an important role in the transition to motor vehicles but that this role was not entirely consistent with the assertions of the niche hypothesis. I also find that the mechanisms of learning involved in the improvement of motor vehicle performance and reductions in cost were more complex than the descriptions found in QE and MLP theory. Section 6.4 reviews the activities aimed at altering variation and selection, as well as other forces that produced mutual adaptations in the technology and institutions. I find that the technological nexus concept is useful in understanding the process of embedding. However, I find that this process occurred simultaneously with or lagged behind the diffusion of the automobile into the mass market rather than within the context of niche markets. I also find that the technological nexus concept is not sufficient for understanding all the important forces behind the co-evolution of technologies and institutions. Section 6.5 discusses other factors related to niche-regime interaction that were important in facilitating the transition to motor vehicles, including the role of compatibility with existing regimes, the stability of existing and emerging regimes, and the role of exogenous events. I find that the diffusion of the motor vehicle is more consistent with a description of gradual unfolding and adaptation similar to the concept of 'hybridization' (Geels, 2002) than with the description of niche proliferation and rapid regime renewal. Section 6.6 summarizes these findings and discusses the implications for theory and policy. Last, section 6.7 presents an agenda for future research.

6.2 Research Questions

- Is the role of niche markets as described in the literature historically accurate for the transition from horse-drawn to motorized vehicles?
    - Were learning-by-interacting with users and learning-by-doing the primary processes involved in improving the performance and cost of early motor vehicles?
    - Did growth in niche markets facilitate economies of scale?
    - Is niche branching, niche proliferation or niche market growth an accurate description of the diffusion of the gasoline motor-vehicle into mass markets?

- Is the linkage between selection and variation described by the technological nexus an accurate account of the interactions that produced the technological and institutional adaptations that facilitated wider diffusion of the automobile?
  - Who were the agents involved and from what regimes were they drawn?
  - How did elements of the existing regime(s) influence their identities and roles? How were these identities and roles altered?
  - What formal institutions were developed and what contributed to their stability and longevity?
  - How were the variation and selection environments altered? Did this co-evolution occur within the context of niches?
  - Was the selection-variation linkage the most significant network contributing to technological and institutional adaptations?
- Were other factors related to niche-regime interaction important in facilitating the transition to motor vehicles?
  - In what ways were early motor vehicles compatible with the existing regime(s)? How did this compatibility evolve and eventually diverge?
  - Did regime instability contribute to the transition? If so, what were the causes of this instability?
  - What role did unexpected exogenous events play in existing and emerging regime instability and the adaptation of institutions?

6.3 Role of Niche Markets

This research shows that niche markets did indeed play an important role in the transition to motor vehicles. However, that role is not entirely consistent with the assertions of the niche hypothesis. This section summarizes the role of niche markets in both the emergence of the technology and its diffusion to mass markets. I draw out several consistencies and discrepancies between the history and the assertions of MLP and QE theories. These findings are also summarized in section 6.6.

6.3.1 Emergent Phase, Through 1900

The first experiments with motorized transport, which occurred in the late 18th and early 19th centuries, applied steam engine technology to transporting heavy equipment for military applications and rail freight. By the late 19th century, motorized rail transport had become the primary mode of commercial long-distance transportation for freight and passengers in the U.S. Although steam-powered cable systems initially served urban passenger transportation, electric rail cars, introduced in 1888, became the dominant technology for this application by 1902. Meanwhile, the bicycle emerged in the 1880s as a form of personal transportation. This technology first found application in a niche market providing sport and entertainment for daring young men. By the late 1880s, technological improvements had expanded the bicycle market to include men, women, and children of all physical abilities. The bicycle provided exercise and recreation as well as practical, short-distance, personal travel.

Thus, by the time motor power vehicles found application in personal transport, motorized rail systems had developed into two regimes for commercial transport. Practical personal travel was served by an existing regime based on draft animals while personal transport for sport and recreation was served by an emerging bicycle regime. Although the rail and draft animal regimes were fairly stable, problems had appeared by the turn of the century. The railroads had come under scrutiny for abuses of monopoly power and price discrimination. Thus, business practices were becoming misaligned with a growing cultural trend known as Progressivism. Horse-drawn public transport in urban areas had difficulty negotiating steep inclines and was contributing to problems with wastes, odor, flies, and illness. The success and growth of this regime was causing difficulties with continuing on the current trajectory. Meanwhile, success of the bicycle was altering consumer perceptions and creating a demand for performance characteristics (greater speed and distance) that the technology was unable to fulfill. The emerging bicycle regime was facing technological limits.

During the emergent phase, personal motor vehicles first found commercial success in two niche markets serving very wealthy consumers. These niches were defined by the function they served. Electric automobiles were used for short trips within cities and were the technology of choice for wealthy women. They provided practical urban transport, conveyed social prestige, and were attractive as a solution to rising problems with the existing animal-powered transport regime in this function. Steam-powered and gasoline internal combustion vehicles were used for touring the countryside. They provided sport and recreation and also conveyed social prestige.

Based on the niche hypothesis, we might consider that the application of steam and electric power to personal vehicles was representative of niche branching – the use of a radical innovation in subsequent application domains or market niches – where the original application was in rail transport. However, engines used in commercial transport were too heavy and bulky for personal transport and developing a small engine or electric motor and battery with sufficient performance represented a significant innovative challenge. Therefore, the technologies for steam and gasoline power used in early motor vehicles derived more from engines for stationary uses than from other motorized transport applications. Meanwhile, electric vehicle technology relied on battery technology and not primary electricity generation and transmission technologies used for electric rail systems. The *inspiration* to apply motor power to personal vehicles did not arise from rail transport either. Instead, it arose from the bicycle regime. Therefore, niche branching from the rail regime is not an accurate description of the emergence of niche markets for motor vehicles. Nonetheless, much of the knowledge base relevant to motor vehicles was shared with rail transport.[74]

The MLP asserts that problems in an unstable regime present opportunities for a new technology to serve as a solution, effectively creating new niches (Geels, 2004). Thus, one could argue that the early success of electric automobiles in the urban market niche was facilitated by problems using horse-drawn vehicles in cities. However, electric rail systems were already established as an effective solution and bicycles also could provide clean individualized transportation in town. Again, electric motor vehicles were conceived as a technology for urban applications only after the bicycle altered perceptions about personal transport. This change provided the inspiration for motor power to replace horse or human power for personal, rather than commercial, transport. Electric vehicles could provide a combination of the advantages of electric rail systems (cleanliness) and personal carriages (individual flexibility) without the disadvantages of the bicycle (physical exertion and limited cargo capacity). Thus, it was emerging misalignments in *multiple* regimes that provided an opportunity for electric cars to enter the urban niche, where they competed with carriages, electric rail cars, and bicycles.

[74] We do see a number of automobile entrepreneurs with experience manufacturing rail cars, notably Walter Chrysler and Edward Budd. Henry For also worked for a manufacturer of rail cars, but only for six days. His main experience was with stationary engines.

The MLP also asserts that the articulation of new functionalities plays a role in niche development. In a stable regime, a new functionality creates niche applications and allows a radical technology to develop relatively unnoticed by the existing regime. In an unstable regime, entrepreneurs may promote the new functionality of their innovation as a solution to regime problems. No new functions were articulated for electric vehicles in urban transport – they provided practical transport and social prestige and in this manner served as a direct replacement for fine carriages for the wealthy. They also provided clean intra-city transport, which was a direct replacement for electric rail cars or bicycles. However, they were able to provide these functions simultaneously, which represents a form of hybridization, as discussed below. The use of steam- and gasoline-powered motor vehicles for touring did represent a new function by combining the flexibility of horse-drawn vehicles and bicycles with the speed and long-distance capabilities of rail transport. This function – rapid, cross country, personal transport – provided entertainment, sport, and thrill, as well as social prestige. As the MLP predicts, motor vehicles in this application did not compete with the existing technologies for practical long distance transport (rail and carriage). However, to some extent, it did compete with bicycles as wealthy urbanites began driving motor vehicles in city parks for recreation.

In both of the niche markets for automobiles, the combination of functional features from multiple regimes represents a sort of hybridization. Hybridization here is interpreted as the merging of functions, performance features, and technologies, both old and new, from multiple regimes to solve particular problems or provide new functionality through symbiosis. The MLP describes hybridization as a mechanism in niche proliferation, but identifies only the combination of new and existing technologies to solve particular problems while causing a minimum of regime disruption.

At this point, the automobile case history supports the assertion that radical technologies emerge in niche markets. The history is also generally consistent with the pathways of niche development described by the MLP. However, it suggests two areas where the MLP lacks sufficient description. First, the emergence of the motor vehicle in early niche markets can only be understood through consideration of multiple regimes. In MLP language, the analysis must include more than one meso-level system. This deficiency was raised by Raven (2005) and accounted for in the framework developed in section 2.6 by considering direct linkages to other industries as well as indirect linkages through the technical, institutional, and ecological complex (TIEC). This is also consistent with the assertion of Kemp et al. (1998) that "[t]he success of niche formation is… linked to structural problems, shifts and changes within the existing regime(s). The ultimate fate of processes of niche formation depends as much on successful processes within the niche as on changes outside the niche: it is the coincidence of both developments that gives rise to niche development patterns." However, this research suggests that much more emphasis should be given to this underdeveloped aspect of niche formation.

Second, a more complete conception of hybridization would include the merging of functionality and performance features in addition to technological artifacts. Further, this combination can yield synergies that result in new functions or applications in addition to novel solutions.

6.3.2 Transitional Phase I, 1900-1910

We now turn to consideration of the transitional phase where the technology moves from niche to wider markets. This phase began around 1900 for the automobile. Did processes within the niche facilitate this diffusion, and if so, what were these processes? The MLP and QE theories assert that, within the niche, manufacturers learn about users' needs and users learn about product performance through learning-by-interacting. Manufacturers then improve the technology in terms of those needs and achieve lower costs through learning-by-doing. Consequently, growth of niche markets, niche branching, or niche proliferation allows manufacturers to reduce costs through economies of scale.

Only the gasoline-powered motor vehicle was able to break out of its niche market and conquer wider markets. The early niche for gasoline vehicles was the relatively small, wealthy touring market. While other uses for the automobile and geographic features likely defined many other niche markets, as of the turn of the century, the automobile was widely regarded as a toy for the rich. But the market for automobiles quickly shifted from the very wealthy to the lower upper class, primarily merchants and physicians, for whom the automobile provided utility as well as sport. The Olds curved dash, which sold for $650, accounted for one third of the 4,100 motor vehicles sold in 1900, the year it was first introduced. Many small Midwestern manufacturers were also selling light, inexpensive surrey-type vehicles like the curved dash that appealed to wealthy farmers. Although these vehicles were well suited to navigating rural roads, they tended to rattle apart quickly on them. The key to penetrating the rural market was producing an inexpensive vehicle that could navigate rough roads yet was rugged and reliable. With this goal in mind, Ford introduced three successive inexpensive models during the transitional phase: the Model A in 1903 with prices starting at $750, the Model N in 1906 starting at $500, and the Model T in 1908 starting at $825.

Geels (2005b) asserts that growth in the touring niche in the first years of the 20th century was responsible for gasoline cars gaining a lead over electric- and steam-powered automobiles. However, this growth had stagnated by 1905 as the market for expensive vehicles saturated fairly early. As shown in Figure 4-2, the market for inexpensive vehicles was already well established in 1903 and sales had already surpassed that of both the medium- and the high-price category. By 1910, the market for new automobiles priced above $2,775 had stabilized and represented only 10% of the market. Demand for moderately priced vehicles continued to grow, but at an ever decreasing rate. With a midpoint around $2,000, the cost of these vehicles still constrained this market to wealthy manufacturers and capitalists. Meanwhile, the least expensive vehicles appealed to two new market segments: the upper-lower class for whom the automobile fulfilled utilitarian purposes as well as conveying social prestige and a larger market consisting of affluent farmers for whom the automobile was primarily utilitarian.

It is difficult to argue that processes within the early niche market were responsible for a reduction in price that allowed motor vehicles to branch into the inexpensive market. As discussed in section 4.5.4, the trend within the high-price category between 1903 and 1908 was, if anything, toward *higher* price, not lower. And while the performance of these vehicles was being steadily improved, it was being improved relative to the functional requirements of the niche market: speed, power, size, and weight were all increasing. While niche branching was likely occurring between 1900 and 1910 within the slowly growing medium-price class of vehicles, these markets still consisted of a limited number of very wealthy consumers. The trend within this category was toward every higher quality but not necessarily lower price. Performance and reliability in both categories increased as well, but the remaining core requirements for the mass market were not being addressed: affordability, ruggedness, ability to navigate rough roads, and ease and affordability of repair.

Meanwhile, the lowest price category steadily gained market share after 1907, and the trend within that category was clearly toward ever lower price. The decrease in production costs and price for inexpensive vehicles was not the result of a widening of the touring market, but rather of Henry Ford's business strategy to consciously target the rural market and continuously reduce prices. Cost reductions were realized from learning-by-doing and scale economies, among other mechanisms, but they were realized *simultaneously* with a product redefinition that moved the automobile out of the wealthy niches into the mass market. Product improvements for the mass market were also accomplished *in* the mass market. The functional requirements of the mass market were decidedly different from those of the wealthy touring market: practicality, ruggedness, and affordability of ownership replaced speed, power, and display of wealth. This difference in functional requirements resulted in a product design with decidedly different attributes. The Model T was distinctly a "no frills" utilitarian vehicle.

6.3.3 Specific Phase I, 1910-1918

Much of the downward trend in prices within the lowest price category after 1907 was due to sales of the Model T, which by 1915 was priced from $390 and constituted around 70% of sales of vehicles priced under $1,375. As the industry approached the specific phase, Ford actively sought to drive down the cost of manufacturing motor vehicles specifically to make them affordable to the masses. In this pursuit, he was fairly unique. However, Ford's greatest accomplishment was the manufacture of a product of both high quality and low cost. While one or the other had been achieved within the industry, it was the combination that allowed Ford to conquer the large rural market. This combination was achieved by jointly addressing issues of product design and manufacturing processes. Ford designed the Model T with the requirements of mass production in mind, jointly evolving the product design, machine tools, and production process.

The volume production of automobiles, perfected by Ford around 1913, was enabled by four facets of the TIEC: 1) a broad-based acceptance of the ideals of individualism as expressed in capitalism; 2) the evolution of the American system of manufactures begun in the arms-manufacturing industry and perfected through the sewing machine and bicycle industries; 3) the disintegration of traditional craft society; and 4) mechanistic scientific approaches based in methodological reductionism that led to the development of scientific management.

Ford's production team was uniquely able to synthesize and exploit the opportunities provided by this social and institutional context. The key aspects of mass production were simplification of design, standardization of parts, precision machining, continuous motion, carefully timed speed, and the use of labor-saving mechanisms. Ford did not originate any of the component process innovations involved in this system: standardized parts all built to common gauge, precision manufacturing, high volume production, specialized workers performing increasingly minute tasks, control and planning by supervisors and engineers, hierarchical management, specialized machine tools, and continuously moving assembly. Ford was not even the first to apply any one of them, but was the first company to fully integrate them. In doing so, the Ford team relied on knowledge spillovers, learning-by-doing, and a great deal of inspiration. Clearly, for the automobile, the QE and MLP theories' singular focus on learning-by-doing is inconsistent with this development.

As manufacturing costs were reduced, Ford continuously lowered prices, despite the fact that the market during the first specific phase was still clamoring for more production than the industry could provide. Ford therefore set the stage for competition based on price for a given level of quality. For example, Durant explicitly aimed to compete with Ford in developing the 1915 Chevrolet Model 490, named for the price of the 1914 Model T. The automobile had clearly conquered a mass market by 1916: total sales reached over 1.5 million units and 91% sold for under $1,375; 51% of sales were priced under $675 and the Model T had captured half of the total market.

6.3.4 Mechanisms of Innovation

While learning-by-interacting with users and learning-by-doing were important mechanisms in the innovations that improved the ability of motor vehicle to meet consumers' needs and reduce costs, they were not the only significant processes. Table 6-1 summarizes the significant innovations discussed in chapters 3 through 6 and places them within the industry phase structure. This section discusses the role of each mechanism of learning in turn then provides a summary of conclusions.

Clearly, knowledge spillovers were of primary importance in the emergent and first transitional phase, ending around 1910. This knowledge resided in human capital and initially was embodied in plants and equipment in the bicycle, carriage, and machine tool industries. In general, firms in these industries became suppliers or evolved into automobile manufacturing firms. Suppliers thus served as the link for transmitting inter-industry and international knowledge spillovers. Spillovers continued to be extremely important during the first specific phase from 1910 to 1918 as the performance of motor vehicles continued to be substantially improved and costs were further reduced. Spillovers figured prominently in four of the most significant innovations of this phase: Budd's all-steel body (begun while Budd was at Hale & Kilburn), Delco's electric self-starter, the incorporation of the moving assembly line in Ford's production system, and sales on credit.

What is not reflected in Table 6-1, but is apparent from the historical narrative, is that user interactions were important in both product refinements and in firms' decisions to adopt technologies already in the market. Although Ford did not need to interact with users to discover the needs of the rural market, he did rely on feedback from dealers and buyers to identify defects as he focused on product innovations to improve reliability. Meanwhile, supplier interactions provided the technological means to improve the product.

In fact, learning-by-interacting with suppliers figured prominently in all phases. These interactions often included collaborations to solve problems with the existing component or to design entirely new ones. For example, Ford and Holley collaborated on early carburetor development during the first transitional phase.[75] During the first specific phase, Hupp and Edward Budd collaborated to develop all-steel bodies while Budd was employed at Hale & Kilburn. During the second transitional phase, Chrysler and Lockheed worked together to improve four-wheel hydraulic brakes. Suppliers often initiated contact with manufacturers and petitioned for the adoption of their technologies. This relationship is demonstrated in Ford's use of stamped steel and vanadium steel, Cadillac's incorporation of the self-starter developed by Kettering at Delco, and GM's adoption of closed bodies developed by Fisher. The unique relationship between GM and supplier Du Pont played a significant role in the collaborative development of Duco pyroxylin lacquer and the commercialization of tetra-ethyl lead (TEL). At Ford, where extensive vertical integration was realized through in-house production of nearly all parts and components, supplier interactions were not significant after about 1916.

The role of R&D within automotive manufacturing firms in the timeframe of this analysis is decidedly less than would be expected from contemporary industry. This is to some extent a reflection of the fact that the idea of corporate research laboratories was still relatively new.[76] In addition, at least through 1910, firms were operating on slim cash reserves and were primarily interested in re-investing profits in capacity expansion. However, by 1910, automobile manufacturers were using materials testing laboratories to improve product design. This is consistent with the assertions of the industry-phase model. However, R&D did not appear to be a significant factor until the second transitional phase. By 1920, a broad range of R&D was institutionalized at GM with the

[75] This collaboration is not reflected in Table 6-1 because it did not represent a new innovation. It does, however, illustrate the importance of supplier interactions in improving early product performance.

[76] Thomas Edison built the first U.S. industrial research laboratory in Menlo Park in 1876.

General Motors Research laboratories by incorporating the knowledge and talents of Charles Kettering and his staff at Delco. Through vertical integration, GM had brought the innovative efforts of the corporation's suppliers in-house. Kettering's and Delco's innovative efforts were then subject to the constraints imposed by GM's market strategies. This was demonstrated by the termination of the copper-cooled engine program and the selection of TEL over fuel blends in addressing the problem of engine knock. By late 1920, Chrysler had a research laboratory worthy of comparison to those of Ivy League universities.

Perhaps most interesting is the emergence during the second transitional phase of innovations in business operations. GM's market segmentation and decision to standardize parts across divisions, an innovation in both product and process, combined to enable both variety in product and economies of scale. Other significant business innovations include planning for annual model changes, design for style, and the use of time sales. Time sales were introduced during the first specific phase but became a dominant feature of the industry in the second transitional phase. While planning for change may be considered an innovation in business operations, it also required important changes to production processes, specifically in capital equipment. Market segmentation and annual model changes designed for style were innovations based on a fundamental understanding of emerging user preferences, which GM tracked with rather sophisticated data analysis. Finally and perhaps most significantly, GM's innovations in corporate structure and operations established a new standard business model for industry in the second specific phase. While these innovations resulted from in-house efforts, largely learning-by-doing, the entrepreneurs themselves relied on experience gained in other industries and were associated with GM due to past supplier relationships and financial difficulties.

In conclusion, while learning from users and learning-by-doing were important mechanisms in improving the performance and cost of early motor vehicles, they were not the only significant ones. Table 6-1 shows a striking contribution from spillovers and learning-by-interacting with suppliers during all phases. In addition, this table and the analysis provided in section 5.7.5 show that, contrary to the description of the industry life-cycle model, there was no clear industry-wide trend toward process innovations and away from product innovations during the first specific phase, nor was there a reverse trend during the second transitional phase. I also find no clear trend in the shifting importance of the mechanisms of learning as the industry progressed. Finally, I find that innovations in business operations deserves mention and should not be over-looked in studies of technological progress and transitions.

Table 6-1 Sources of Selected Innovations and Industry Phase

	Phase	Technology/Innovation Description	Type[b]	Source: Spillover	R&D	Learning by Doing	Learning by Interacting[a]: Users	Suppliers
	Changes in TIEC: problems with expanding scale of current technology (stabling space and cost, waste, odor, and flies); demand for rapid personal transport stimulated but not met by bicycle; advances in machine tool technology; changing perceptions of efficiency and human labor							
< 1900	Emergent	Motive power (batteries, steam and gasoline engines)	1	X				
		Product configuration	1	X carriages, bicycles				
1900-1910	Transitional	Manufacturing techniques (standard parts, volume production, single purpose machines)	2	X arms, sewing machines, bicycles		X		X
1903		Pressed steel frame (A.O. Smith)	1	X bicycles				X
1907		Integrally cast cylinder block and crankcase (Ford)	1,2[c]			X		
1908		Vanadium steel (Ford)	1[d]	X Europe	X	X		X

[a] The distinction of user versus supplier interaction is taken from the perspective of the automobile manufacturer. An innovation that originates with the supplier at its initiative is considered to come to the manufacturer through interactions with suppliers.

[b] Type denotes (1) product, (2) process, (3) marketing or (4) other innovation.

[c] Abernathy et al. (1983) classify these as process innovations. I give a joint classification to product design changes made for incorporation of innovative manufacturing techniques.

[d] Abernathy et al. (1983) classify this (and material changes in general) as a process change. Given our discussion of the impact of the use of vanadium steel on the design of the Model T and its resulting superior performance, I classify this as a product change.

	Phase	Technology/Innovation Description	Type[b]	Source: Spillover	R&D	Learning by Doing	Learning by Interacting[a]: Users	Suppliers
1908		Detachable cylinder head (Ford)	1,2[b]			X		
1908		Magneto integrated in flywheel (Ford)	1,2[b]			X		
1909		Branch assembly (Ford)	2			X		
1910-1920	Specific	Functional definition: basic, reliable, inexpensive, personal transportation capable of traveling long distances and negotiating rough roads Dominant design: water-cooled, four-stroke, four-cylinder, gasoline internal combustion engine mounted vertically under a front hood; shaft drive; magneto; steering wheel Dominant business model: capitally intensive volume production using standard parts (Fordist production)						
1910		Closed body (Fisher on Cadillac)[e]	1				X	X
1912		Electric spot welding (Hale & Kilburn; enabling tech. for all-steel body)	2	X rail cars				X
1912		All-steel body (Hale & Kilburn on Hupmobile)[f]	1			X		X
1912		Electric starter (Delco on Cadillac)	1	X electric cash registers				X

[e] This marks the first mass production (150 units) of closed bodies. The first closed body, made of a wood frame and steel panels, was introduced in 1898 by Columbia Electric.

[f] The first all-metal body automobile was a 1902 Marmon with a cast aluminum body. The all-steel Hupmobile body was developed by Edward G. Budd at Hale & Kilburn, and was only available for one year. The first mass produced (5,000 units) all-steel bodies were built for Dodge by the Edward G. Budd Manufacturing Co. in 1914.

	Phase	Technology/Innovation Description	Type[b]	Source				
				Spillover	R&D	Learning by Doing	Learning by Interacting[a] Users	Suppliers
1913		Moving assembly line (Ford)	2	X Flour milling, meat packing, *etc*		X		
1914		Large-scale production V-8 engine (Cadillac)	1		X			
1914		Mass production of all-steel body (Budd on Dodge)	1					X
1916		Sales on credit (Willys)	3	X sewing machines, furniture			X dealers	
1918	Economic shock: post-war recession, overcapacity, and market saturation; precipitates high failure rate and further firm consolidation, and thus ever-increasing market concentration							
	Changes in TIEC: new consumer preferences for comfort, style, and performance, due to improved roads, inexpensive closed bodies, market saturation, increasing prosperity, increasing role of women in purchases							
1918-1926	Transitional							
1921		Market segmentation by economic class (GM)	1,3			X	X	
1921		Four-wheel mechanical brakes (Duesenberg)	1	X Europe				
1922		Four-wheel hydraulic brakes (Lockheed on Duesenberg)	1		X	X		X
1923		Balloon tires (Firestone)	2		X	X		X
1923		All-steel closed body (Budd on Dodge)	1		X			X
1923		Tetra-ethyl lead gasoline additive (GM)	4		X			

	Phase	Technology/Innovation Description	Type[b]	Source: Spillover	R&D	Learning by Doing	Learning by Interacting[a]: Users	Suppliers
1923		Annual model change (GM)	1,2, 3	X fashion		X	X	
1924		Duco lacquer finish (Du Pont on Oakland)	2	X gunpowder, toys	X			X
1926		Standardization of parts across product line (GM Oakland Pontiac)	1,2			X		
1926-?	Specific	Functional definition: stylish, comfortable, personal transportation that conveys social prestige and the pleasure of personal ownership. Dominant design: electric starter, closed body, competition based on styling, increasing body and engine sizes culminating in V-8 engine Dominant business model: Large, vertically integrated corporation; "flexible mass production"; staff and line de-centralized organization with financial controls ("Sloanist" management).						

6.3.5 Summary

Growth in niche markets did not provide the economies of scale that reduced the costs of motor vehicles and facilitated diffusion to mass markets. Further, the needs of the mass market were not discovered within the niche market, nor were product improvements relative to these needs achieved in the niche market. And because the vast rural market can not be accurately described as a market niche, niche branching, niche proliferation, or niche market growth are not accurate descriptions of the diffusion of the gasoline motor-vehicle into mass markets. In fact, niche branching occurred *after* diffusion to the mass market. The gasoline motor vehicle conquered the urban niche occupied by electric vehicles after being vastly improved for use in the mass market.

Therefore, the role of niche markets in the transitional phase and early specific phase of the automobile is not consistent with the characterization found in the MLP and QE theories of transitions. This, however, is not to say that niche markets were not important or possibly even critical. Competing technological configurations first found applications in niche markets which served to demonstrate technical feasibility relative to niche market needs. These markets also demonstrated the relative advantages, as well as the limitations, of each alternative. Within the niches, firms developed technical capabilities, established reputations, and formed networks for financing, manufacturing, and sales. Racing was extremely important in working out technological issues and demonstrating to the public the capabilities and the performance of automobiles generally and of technological options and manufacturers specifically. This role will be discussed further in section 6.3.5.

Finally, while learning from users and learning-by-doing were important mechanisms in improving the performance and cost of early motor vehicles, they were not the only significant ones. Table 6-1 shows a striking contribution from spillovers and learning by interacting with suppliers. This finding supports the assertion that the MLP and QE theories would benefit from a more complete description of the learning mechanisms involved in the co-evolutionary technological and institutional adaptations that facilitate diffusion. It also has implications for the definition of the technological nexus, as discussed in section 6.4.

6.4 Mutual Adaptations and the Technological Nexus

In QE theory, variation and selection are coupled through institutional linkages that constitute the 'technological nexus.' Actors serving in the nexus role translate the requirements of the selection environment into objectives for technological development and also alter the selection environment, in part by imposing the requirements of technology on the selection environment. These activities produce co-evolutionary technological and institutional adaptations that facilitate wider diffusion and result in the new technology becoming embedded in the wider socio-technical complex.[83]

This section reviews the activities aimed at altering variation and selection, as well as other forces that produce mutual adaptations in the technology and institutions. I find that the technological nexus concept is useful in understanding the process of

[83] Recall that in the framework presented in section 2.6 institutions are defined to include both organizations and rules (cognitive, normative, and regulative). They include any form of social construct or constraint on human interactions, including organizational constructs, such as governmental bodies and research and industry organizations; formal constraints, such as legislation, economic rules, and legal contracts; and tacit constraints, such as shared perceptions, beliefs, social customs, and moral codes. For example, races are a social custom with shared cultural meaning and expectations and therefore are an institution that shapes interactions.

embedding. However, I find that this process occurred simultaneously with or lagged behind the diffusion of the automobile into the mass market rather than within the context of niche markets. I also find that the technological nexus concept is not sufficient for understanding all the important forces behind the co-evolution of technologies and institutions. I identify four deficiencies in the nexus description and suggest some ideas that would complement the nexus concept.

Enthusiast publications and the popular press were the first institutions to emerge in the nexus role. They facilitated communication among engineers and altered selection by influencing public opinion and expectations of motor vehicles, the technological options and specific manufacturers. In this role, they were not necessarily a source of unbiased information. Manufacturers also made extensive use of the press, beginning with Pope's initiation of the press interview as a means to introduce new models. Through the Association of Licensed Automobile Manufacturers (ALAM) and its counterpart, the American Motor Car Manufacturers Association (AMCMA), manufacturers used the press extensively to sway public opinion and reassure consumers during the Selden patent dispute.

The press initiated another institution, racing, that shaped both variation and selection. Racing was established first in Europe and evolved from the custom for bicycles. Speed and reliability races demonstrated the capabilities of the technical options and specific models, thus shaping public perceptions and also stimulating the interest of investors. Thus, by directing funding to particular technologies and manufacturers, races and demonstrations influenced the direction of variation.

However, Ford argued that races also limited the industry's early development (Ford and Crowther, 1922). The technology itself – its very existence, its specific design, and its use – alters the selection environment. The first functional definition of the automobile and its use in that function initially served to inhibit diffusion into wider markets. As conspicuous spectacles, races reinforced public opinion that motor vehicles were no more than toys for the very wealthy. In addition, the use of automobiles by wealthy motorists in elaborate festivities and reckless touring aroused indignation and anger in those with lower financial status. However, this negative opinion faded after the introduction of inexpensive motor vehicles that provided a valuable and affordable function for these populations. Ransom Olds served in the nexus role by interpreting initial public skepticism into design requirements for the curved dash – simplicity of use and repair. Experience with using this vehicle then served to overcome the skepticism that necessitated its design specifications, thereby easing the 'harsh' selection environment. Ford's interpretation of market requirements and the resulting vehicle designs continued this mutual adaptation.

Other activities of manufacturers through emerging industry trade organizations served the nexus role. The ALAM and the AMCMA in particular were conceived as organizations to connect manufacturers to each other and to customers by arranging for public exhibition new vehicles (trade shows); promoting races; promoting the sale of members' cars; encouraging public interest in automobiles; promoting good roads; and through the exchange of technical information. In enforcing the Selden patent, the ALAM claimed to be protecting the industry from disreputable manufacturers that would have created public distrust and a negative opinion of the automobile. By restricting licenses and controlling production output, the ALAM also sought to maintain high prices, thus affecting variation. The issue was settled by the legal system, and these higher level rules ultimately supported the independent manufacturers of the AMCMA, though this group had already disbanded. After the legal decision in 1911, the ALAM's boundary spanning role was transferred to the Automobile Board of Trade and then, in 1914, to the National Automobile Chamber of Commerce (NACC). The ALAM's technical work was transferred to another prominent industry organization, the Society of Automobile Engineers (SAE), which served as a linkage among manufacturers and suppliers. While the SAE did not connect manufacturers to customers, it did influence both variation and selection. As a technical society, the SAE represented the shared knowledge base which served as a cognitive constraint on innovation by shaping the perceptions of engineers. But more significantly, the confusion caused by the huge number of parts used by the industry constituted an adverse selection environment which the SAE attempted to alleviate through parts standardization. While the organization succeeded at easing the parts burden and improving the quality of materials, it was unable to make the selection environment significantly more benign for small manufacturers.

Other than the Selden patent dispute, the automobile industry was fairly free to develop without governmental restrictions during the emergent and first transitional phase. However, formal constraints on automobile use were emerging in the form of local and state speed and registration laws. Initially, these laws were misaligned with the capabilities of the automobile and served to limit its usefulness and therefore its diffusion. Through motoring clubs, enthusiasts joined forces with manufacturers to remove this element from the selection environment, eventually shifting efforts from opposition to coordination. In this process, the purpose of motoring clubs shifted from organizing social activities for the wealthy elite to activism for all motorists. By 1910, registration laws were more coherent and actually served to create a more benign selection environment, since fees were used for road improvements.

The first instance of government action related to the environmental health effects of automobile use arose during the second transitional phase after the introduction of tetra-ethyl lead (TEL). An industrial incident and the statements of university experts and the federal Public Health Service (PHS) raised the TEL issue to national attention. However, institutions to protect public and worker health were still in their infancy and the federal government had no authority to regulate chemicals in 1925. Thus, although the use of TEL was misaligned with a rising national consciousness about worker and public health, no formal legislative rules were in place to block its use. GM actively worked to refute the claims of these experts and the PHS was unable to make conclusive statements regarding the health effects of long-term exposure. In addition, the courts found Standard Oil innocent of criminal negligence in the illness and deaths related to the TEL plant accident. Therefore, GM's efforts at altering the selection environment were successful and the controversy faded.

The poor condition of roads initially had served to limit the usefulness of the automobile and was a major design consideration for inexpensive vehicles. Once again, motorists, through clubs, joined forces with manufacturers to promote road improvements. Many of these agents were continuing efforts they had undertaken as bicycle enthusiasts and manufacturers. Because motor vehicles offered an alternative to railroads for transporting agricultural goods, farmers joined these efforts around 1909 through the National Grange. Farmers were motivated by economics and by the perceived abuses of monopoly power and price discrimination by the railroad, an issue brought to national attention by Progressives. The Good Roads movement produced modest results, but improvements were overtaken by the increasing demands of motor vehicle use until the appropriation of federal funding for road improvements beginning in 1916. But the major breakthrough in road construction resulted from the industry's response to events that occurred during WWI which demonstrated the feasibility of both long-distance trucking to transport freight and the year-round use of highways.

While motorists' and manufacturers' efforts to coordinate registration laws and improve roads did alter the selection environment, they did not simultaneously influence technology development. Therefore, they cannot be designated as nexus activities as described by QE theory. However, these actions *indirectly* influenced the technology and resulted in mutual adaptations. This feedback came through consumer expectations as roads were dramatically improved after the war and the idea of year-round use became common. This change was one of several factors that initiated a second transitional phase, or era of innovative ferment, between 1918 and 1926. As a result, the automobile's functional requirements and attributes were significantly altered to align with the much improved condition of roads. Two other related factors contributed to the second transitional phase: market saturation and the used car problem. Manufacturers' innovative responses to these challenges – time sales and annual model changes – were designed to alter the selection environment.

By the beginning of the specific phase (1910), mutual adaptation had resulted in a dominant design that was aligned with existing physical infrastructure (roads and a dispersed refueling system) and supporting organizations were actively working to improve that infrastructure as well as legislative rules that created an adverse selection environment. Thus, the process of embedding was well underway and many of the activities responsible for the co-evolution of technologies and institutions were consistent with QE theory's technological nexus description. However, most of these adaptations occurred during the transitional phase and were simultaneous with or lagged behind the diffusion of the automobile into the mass market. In addition, they continued and were reinvigorated with the onset of the second transitional phase.

As an agent associated with niche markets, the ALAM worked to protect these niches and restrict market diffusion.[84] Motoring clubs also represented a niche market (wealthy touring) but changed roles during the process to represent the needs of the mass market. One prominent nexus agent, Albert Pope, was drawn from a competing regime (bicycles) and was initially associated with a niche market (electric vehicles). Pope apparently viewed motor vehicles as an opportunity – a solution to falling bicycle sales – rather than a threat and adapted his role to support rather than stifle the emerging industry. He also eventually adapted from manufacturing electric vehicles to gasoline vehicles. Pope's activities demonstrate the potential significance of existing networks and agents to emerging industries.

This research provides evidence that the technological nexus concept is useful in understanding the process of embedding but suggests that it is not sufficient. I identify four deficiencies in the nexus description of the interactions that produce technological and institutional adaptations. In general, these deficiencies stem from the singular focus

[84] The ALAM attempted to maintain the motor vehicle's standing as an expensive product for the wealthy. In addition, it was organized to enforce the Selden patent that was owned by the Electric Vehicle Corporation.

on agent activities that directly link variation and selection and result in simultaneous adaptations.

First, some of the important linkages between variation and selection are not mediated by agents, but arise from system properties. For example, saturation of the wealthy touring market prior to 1910 served to down-select specific firms and thus influenced variation. Similarly, activities that alter the selection environment may not directly alter variation (or vice versa), but the system structure and properties may create indirect feedbacks that do. This is illustrated by the chain from activities of motoring clubs to improved funding for roads to altered consumer preferences to the re-specification of vehicle design.

Third, the technological nexus does not account for all the linkages among agents which are important to bringing about technological and institutional adaptations. Throughout the history reviewed here, suppliers were extremely important as originators of or collaborators in product innovations. This suggests that two networks that are not included in the technological nexus are crucial to understanding technological adaptations to selection pressures. The first is the manufacturer-supplier relationship. Second and less obvious is the relationship between suppliers and consumers of the end-product. This raises an interesting question for further study, especially in relation to innovation in the automobile industry today: how do suppliers collect and interpret information from a selection environment to which they are not directly connected?

Finally, events that arise outside the system, like WWI or the Panic of 1907, may directly alter the selection environment, stimulate agent responses that directly or indirectly influence it, or alter the resources available to certain firms for innovation. Such events can be critical in enabling, altering, or halting the process of embedding.

In addition, this research suggests some ideas that would complement the nexus concept. First, particular individuals with high visibility, resources and influence might serve as 'nexus champions,' promoting institutional and technological co-adaptations in much the same spirit as 'product champions' promote specific innovations. Examples from this historic review include the role of Albert Pope in championing the Good Roads movement and Henry Ford's role in championing the rights of the independent manufacturers during the Selden patent dispute. Second, the combination of the boundary spanning nexus role, exogenous events, and the effects of system structure and properties yields synergies that produce greater impact on variation and selection than the sum of the impact of each process occurring in isolation. This typical feature of complex system dynamics is illustrated by the multiple factors and processes that initiated the second transitional phase from 1918-1926.

6.5 Other Factors

This section 6.5discusses factors related to niche-regime interaction that were important in facilitating the transition to motor vehicles, including the role of the compatibility of the automobile with existing regimes, the stability of existing and emerging regimes, and the role of exogenous events. I find that the diffusion of the motor vehicle is more consistent with a fit-stretch pattern of gradual unfolding and adaptation than one of niche proliferation and rapid regime renewal. I also find that the combination of instabilities in multiple existing regimes provided opportunities for the emerging automobile industry. While a process of alignment was increasing the stability of the automobile industry during the first specific phase, I find support for the argument that a stable regime did not coalesce until the second specific phase when the technology abandoned the last vestiges of carriage styling and automobile production techniques became fully aligned with new, higher level rules in the TIEC.

6.5.1 Compatibility and Fit-Stretch

Section 2.4.2 described a diffusion pattern of hybridization that Geels (2002) asserts may occur in unstable regimes. In this process, new technologies are merged with existing ones to solve particular problems in the existing regime with a minimum of disruption to that regime. Continuous development eventually leads to the articulation of new functionalities related to the novelty's capabilities and new technical designs are tailored to these functions. Over time, the hybrid technology's form and function co-evolve through a fit-stretch pattern as illustrated in Figure 2-2.

Although problems were arising with the expanding scale of the carriage-based transportation regime in urban use, the regime was still relatively stable, particularly for serving rural transport. However, the development and diffusion of motor vehicles shows many parallels with the hybridization pathway and fit-stretch pattern.

The first gasoline-powered vehicles can easily be described as a combination of a new technology with elements of the existing regime. Some of the vehicles built during the emergent phase consisted of a small gasoline internal combustion engine attached under a carriage's floorboards and connected to the wheels using a chain drive adapted from bicycles. Even after commercial production began in the transitional phase, many automobiles were constructed using bicycle wheels attached to bodies produced by carriage manufacturers. Indeed, most of the inexpensive vehicles that first penetrated the mass market were still merely motorized horse buggies (Flink, 1988). The design gradually evolved, departing from the restrictions dictated by horse power and taking greater advantage of the options and capabilities afforded by the gasoline engine. By 1910, when the first dominant design emerged, the engine had been moved to the front and enclosed under a hood; a steering wheel had replaced the tiller; and a rotating shaft had replaced the chain drive. The last vestiges of carriage styling disappeared with the emergence of the second dominant design incorporating the all-steel closed body in introduced in 1923 and the integrated design of the stylists' car in 1927.

With the cumulative effects of a century of innovation in the firearms, sewing machine, and bicycle industries, the manufacture of automobiles also required only that skills and knowledge be "transferred from existing uses to new ones" (Rosenberg, 1963). These manufacturing processes then evolved to meet the specific needs of the technology and were soon providing improvements to machine tools that diffused to other industries. Eventually, an entirely new production system was developed by the industry that was suited to the requirements of high-volume, high-quality, low-cost production. Again, the last vestige of manufacturing techniques inherited from the carriage industry disappeared with adoption of all-steel bodies.

Vehicle design was also tailored to the physical infrastructure of existing regimes. Practical automobiles were capable of negotiating rough roads and could be repaired by mechanics experienced with stationary engines and farm machinery. Gasoline vehicles gained a market advantage over steam and electricity in part because fuel was widely available through a dispersed system established for the production and distribution of kerosene for lighting and heating. Beginning during the first specific phase, the requirements of the automobile were then imposed on the infrastructure. Roads were improved and gasoline of predictable quality became available at dedicated stations. In 1923, a fuel additive was developed to enable higher compression ratios. By the late 1920s, road improvements made faster, year-round driving possible, and these capabilities were reflected in a new dominant design that was lower to the ground and incorporated all-steel closed bodies

Combining this description with the findings of section 6.3, it appears that the diffusion of the automobile is more characteristic of Geels' (2002) hybridization mechanism described in section 2.4.2 and the general fit-stretch pattern shown in Figure 2-2 than niche proliferation. Indeed, one could argue that the gradual unfolding of the automobile regime represented an internal transition (see section 2.6) for the carriage and bicycle industries. Indeed, the bicycle industry clearly went into rapid senescence with the emergence of motor vehicles and many manufacturers adapted to automobile manufacturing. However, supporting the argument for internal transition would require a more thorough study of the disruption to the carriage industry, specifically an analysis of the number of firms that adapted to automobile manufacture, the number that failed, and the number of new firms that arose from outside either the carriage or bicycle industries.

6.5.2 Regime Stability

As discussed in section 6.3, problems had emerged in rail and draft animal regimes by 1900. Business practices of the railroads were becoming misaligned with the trend toward Progressivism, while the success and growth of the carriage regime combined with increasing population densities was causing difficulties continuing on that trajectory. Meanwhile, the emerging bicycle regime was facing technological limits just as its success was altering consumer perceptions and creating demand for greater performance. These instabilities did provide opportunities for the emerging automobile industry, but mainly through their combination. The emerging misalignments in multiple regimes provided an opportunity for electric cars to enter the urban niche. Meanwhile, the combination of misalignments in multiple regimes *and* the articulation of a new functionality opened a new niche for touring and later contributed to wider diffusion. In addition, the automobile itself was aligned with increasing individualism; this was particularly true for gasoline cars since touring conveyed a sense of freedom never before experienced. Again, the *coincidence* of developments and processes in *multiple* regimes *and* within niches is critical to understanding the emergence of the motor vehicle.

As the automobile industry developed, the success of the Ford Motor Company can be partially attributed to the alignment of Fordism with evolving higher-level rules within the TIEC regarding capitalism, mechanistic reductionism, and the value of efficiency. This success in turn reinforced those trends. A process of embedding, as discussed in section 6.3.5, began in the transitional phase and continued through the specific phase, bringing the technology, industry, institutions and infrastructure into further alignment. Development of aligned physical infrastructures for roads and refueling began periods of significant growth late in the first specific phase and became a major factor in the second transitional phase.

Although many institutions were still in their infancy and dedicated physical infrastructure lagged behind diffusion of the automobile, the trend toward alignment gave the industry some stability in the first specific phase. However, this stability was undermined by an exogenous event – the post-war recession – and by evolving consumer preferences that became misaligned with the existing dominant product design. Four factors contributed to this misalignment: 1) the increasing influence of women in the automobile market; 2) market saturation; 3) road improvements; and 4) the introduction of mass produced all-steel closed bodies. The first of these is an exogenous development in the TIEC. The second and third factors can be attributed to the success of the automobile. The fourth is a breakthrough innovation. Once the automobile had been redefined and a new dominant design had emerged, the product and consumer preferences were realigned, marking the start of the second specific phase.

In chapter 2, I asserted that a socio-technical regime is understood to exist when a technology and sector have achieved the specific phase and when the stability of that specific phase is supported by the alignment of institutions and existing physical infrastructure with the dominant product design, production processes, and organizational forms. Although supporting rules, organizations, and physical infrastructures were certainly co-evolving with the automobile during the first specific phase, one could argue that a regime did not coalesce until the second specific phase. This is consistent with Nieuwenhuis and Wells' (2007) argument that true mass production did not occur until the development of the all-steel closed-body car in 1923. At that point, the production process was completely aligned with the 'new' rules regarding capitalism, mechanistic reductionism, and efficiency. This argument is also consistent with the gradual unfolding of the regime as presented in section 6.5.1, which describes the complete break of the technology design from the old regime with the introduction of the stylists' car, the GM La Salle, in 1927. This marks the completion of the fit-stretch pattern as shown in Figure 2-2. In this interpretation, the industry might be considered to have experienced a 26-

year transition, with the period from 1910-1918 representing a lull in the era of ferment. However, without a measurable definition of 'alignment' and 'stability,' the determination of whether an industry is in the specific phase or has achieved the status of a regime remains a subjective one.

6.5.3 Role of Exogenous Events

In this research, exogenous events or developments are understood as external forces that arise with no direct causal linkages to the regime or technology of interest. This research identified four notable exogenous developments during the period of this study.

First, the Panic of 1907 occurred just as the high-priced touring market was saturating and likely contributed to the drop in sales in this price bracket in 1908. Therefore, it also contributed to the shakeout phenomenon by contributing to the failure of firms manufacturing these vehicles in 1908, when the highest number of firms exited the industry. Thus, the interaction of a random event and developments in the selection environment influenced variation and further reinforced the trend toward less expensive vehicles and the emerging dominant design. In addition, by reducing sales of high-priced cars and likely influencing the financial position of Alexander Malcomson, the Panic of 1907 also serendipitously supported Henry Ford's efforts to gain financial control of the Ford Motor Company. Finally, funding for the ALAM's Mechanical Branch was cut as a result of the panic and its work was transferred to the SAE in 1910.

Second, the influenza epidemic of 1920 took the lives of John and Horace Dodge. The Dodge widows eventually sold the firm to bankers who were unable to manage it successfully. In 1928, this presented an opportunity for Chrysler to expand its distribution and sales network, add production capabilities, achieve economies of scale in production and research, and reduce competition. With the purchase of Dodge, Chrysler became the third largest U.S. automaker.

Third, WWI and the post-war recession had both immediate and co-evolutionary impacts on the industry. Many small firms that were struggling, particularly those manufacturing larger vehicles, converted to war-time production of trucks and continued in this market after the war. The industry's response to the nation's call for help in the war effort also demonstrated the feasibility of both year-round road travel and long-haul trucking for freight transport. This led to increased federal funding for road construction and shifted consumer perceptions of motor vehicle use. Then, many firms did not survive the post-war recession. Those that did were fundamentally altered as this event triggered a transitional phase.

Finally, the increasing role of women in motor vehicle market decisions also factored into the initiation of a transitional phase. This role reinforced the shift in consumer preferences toward closed bodies, lower clearance, and design for style.

6.6 Conclusions

6.6.1 Research Questions

This research supports the assertion that niche markets play a significant role in the emergence of a new technology. For the automobile industry, competing technological configurations first found applications in niche markets. These niches served to demonstrate the feasibility of the technology relative to niche market needs and the technology's limitations. They also served to demonstrate the relative advantages of alternative technological approaches compared to each other and to the existing regime.

Firm capabilities and reputations were established in the niche and networks were formed for financing, manufacturing, and sales. The technology was initially extremely immature, and user feedback assisted manufacturers in improving the product's performance relative to users' needs.

However, this study also demonstrated that the role of niche markets was not consistent with the assertions of the niche hypothesis and that the process of niche branching, niche accumulation or niche proliferation is not an accurate description of the transition to motor vehicles. The product was improved within niche markets, but only as defined by niche market requirements; the core requirements for the mass market, particularly the need for a combination of high quality and low cost, were not being addressed. Further, the initial niche market for gasoline motor vehicles, the wealthy touring market, did not widen; by 1907 it was actually saturating. In addition, prices were not reduced in this niche market; they were in fact rising. Meanwhile, the market for mid-priced cars, which represented more functional uses but still consisted of wealthy consumers, was growing slowly. However, product and process improvements were again geared toward the needs of this market for higher quality and power.

Instead, cost reductions from learning-by-doing and scale economies were realized *simultaneously* with a product redefinition that moved the automobile out of the wealthy niches into the mass market. Subsequently, product improvements for the mass market were accomplished in the mass market. While learning from users and learning-by-doing were important in this innovative process, it also involved a wider set of learning mechanisms, with spillovers and learning-by-interacting with suppliers figuring prominently. For the gasoline motor vehicle, niche branching into the urban niche occurred *after* diffusion to the mass market. Finally, the transition to automobiles is more appropriately described as a process of unfolding consistent with the mechanism of hybridization, where elements of old regimes are combined with new technologies.

This research finds that that the technological nexus concept is useful in understanding the co-evolution of technology and institutions, but also finds that it does not provide a complete description of this process. Important factors not included in the technological nexus include supplier networks, exogenous events, and feedbacks arising from system structure and properties. Due to system dynamics, the combination of these factors yields synergies that produce greater impact on variation and selection than the sum of the impact of each process occurring in isolation.

In conclusion, the product requirements for niche markets are, by definition, distinctly different from those of mass markets. Therefore, learning-by-interacting with users in niche markets is inadequate for understanding the needs of the mass market. Instead, new technologies diffuse to mass markets simultaneously with a redefinition of the product that specifies appropriate functional requirements for the mass market. For example, the computer was originally conceived as a technology for scientific applications, but moved into the mass market after being redefined to provide functions like word processing that were valuable to a wider market. The underlying process of redefinition is likely to be unique for each technology, though general patterns may emerge. For the automobile, this redefinition was the result of the personal beliefs, values, experience, and inspiration of the entrepreneurs. Finally, supplier networks and linkages to other industries and regimes are critical to understanding the innovations that improve performance relative to mass market needs and facilitate diffusion.

6.6.2 Implications for Theory

The goal of this dissertation was to answer specific policy-relevant questions regarding the role of niche markets in the transition to new socio-technical regimes. This research did not attempt to develop a new model of innovation or a complete theory of technological transitions, but rather used concepts and insights from several bodies of theory to investigate the research questions. However, this research has important implications for the theory of socio-technical transitions.

First, this research found that the QE and MLP theories of socio-technical transitions would benefit from a richer description of innovation as provided by the mechanisms of learning described in section 2.5.2.

Second, evolutionary theories generally assert that innovation is triggered by new scientific or technical breakthroughs and relegate social and institutional issues to the role of 'focusing devices.' In QE theory, social and institutional issues are actively addressed by 'heterogeneous engineers' to realize successful designs. In the MLP, these issues provide opportunities for new technologies to serve as a solution, thereby creating a niche. Thus, transitions occur when processes on different levels link up and reinforce one another. This research has shown that social and institutional issues can trigger an era of ferment and even serve as the *basis* of an innovation, such as scientific management or Fordism. This finding is consistent with the description of innovation presented in section 2.6, where of the role of the entrepreneur is to create value by exploiting *any* change in the TIEC – technical, institutional (economic, social, etc.), or ecological. Thus, the role of social and institutional factors in innovation could be even more central than the description found in the QE and MLP theories.

Third, the QE and MLP theories tend to focus on a single existing regime, though the theories themselves do not preclude the influence of multiple regimes. Meanwhile, this research found that the combination of misalignments in multiple regimes and other developments in the TIEC were critical to understanding why the automobile industry emerged when it did and in the particular niche markets that it did. Therefore, these theories would benefit from a more explicit accounting of the interactions among multiple regimes and niches.

Fourth, in chapter 5, I found that firm histories, the personal beliefs and perspectives of firm leaders, and corporate structures and culture – *i.e.*, firm-specific rules – shaped individual firm strategies for innovation and their adaptive capabilities. These firm-specific rules had a large impact on the development of the technology and the industry as a whole. This finding supports the use of a behavioral approach with a micro-focus in the analysis of innovation and socio-technical change. Thus, theories of socio-technical transition would benefit from the incorporation of concepts from micro-level theories, such as research in the sociology of entrepreneurship, institutional logic, and institutional entrepreneurship, among others.

6.6.3 Implications for Policy

Complex co-evolutionary systems present difficult challenges for policy-makers. Agents are widely dispersed throughout a network with diffuse connections that are often obscure. Meanwhile, the elements of the TIEC and the attributes of the product and production system are continually changing and adapting. These features combine to complicate the identification of targets for regulations, incentives, and sanctions.

Influencing such systems requires complex decision-making and complex, adaptive policy strategies (Rycroft and Kash, 1999).

In addition to local problems, the expansion of the current transportation regime based on personal gasoline ICE vehicles is posing large-scale problems, including dependence on a finite fossil resource that is concentrated in regions outside our control and contributions to global climate change through the emission of greenhouse gases. Alternative technologies currently available to solve these problems all fall short of the existing regime on one or a number of functional requirements such as cost, performance, reliability, fuel availability, convenience, and social acceptability. Current research and policy aim to alleviate these deficiencies.

Though many efforts are underway to improve the tools available for analysis, policies generally fail to account for the difficulties presented by co-evolutionary systems. For example, the existence or lack of opportunities, technological or institutional, arising outside the regime of interest typically is not represented in policy analyses. Further, models used for these analyses typically do not consider the potential for conditions such as consumer and societal valuations to change or for new opportunities to arise in the TIEC. Because of these facts, niche management and other policies may be slow or insufficient to bring alternative technologies into widespread use and may yield unexpected and suboptimal results.

Below I use two cases to illustrate how co-evolutionary issues not considered in typical policy-making can yield unexpected and unsatisfactory results. First, the failure of ZEV mandates to bring electric vehicles into widespread use illustrates the shortcomings of niche management when new opportunities arise within the TIEC due to co-evolutionary effects of the regulation itself. Second, the outcome of CAFE legislation is used to illustrate how failure to account for changing consumer preferences due to co-evolutionary effects can erode the benefits of policy. Both policies altered the selection environment by specifying functional requirements for new vehicles. This discussion is not intended to provide a thorough policy analysis, but rather to illustrate the policy implications of the co-evolutionary perspective.

6.6.3.1 California's ZEV Mandate

In 1990, the California Air Resources Board (CARB) realized it would have serious difficulty meeting federally-mandated air quality goals under the Clean Air Act. In response, CARB passed low-emission vehicle (LEV) regulations establishing state-level exhaust emission standards that were more stringent than those established at the national level. The regulations also established the Zero Emission Vehicle (ZEV) program, which required that 2% of new vehicles sold in the state beginning in 1998 have zero emissions. The fraction beginning in 2003 was set at 10%. In 1990, the battery-powered electric vehicle was the only vehicle capable of meeting the ZEV criteria of zero tailpipe and zero evaporative emissions.

Because auto manufacturers claimed they could not meet the 1998 deadline, CARB delayed implementation until 2003, but the 10% rule for that year remained. After the original LEV regulations were drafted, Honda and Toyota offered hybrid electric vehicles for sale in the U.S. In part due to this new technological opportunity, CARB revised the ZEV rules in 2001, creating two new categories of standards for partial ZEVs (PZEVs) and advanced technology partial ZEVs (AT-PZEVs). To be certified as a PZEV, a vehicle must meet the most stringent LEV emission standard (super ultra-low emission or SULEV[85]) and have no evaporative emissions. Provisions were made for manufacturers to claim partial credit toward the ZEV mandate for the sale of both types of PZEVs, but a fraction of the ZEV mandate had to be met with AT-PZEVs.

In 2002, manufacturers challenged the legality of the revised mandate and implementation was delayed by the courts pending the outcome of the suit. California further revised the ZEV rules, hoping to put an end to the legal battle and restore the program by 2005. Negotiated with the auto manufacturers, the new rules allow more credits from PZEVs but require that manufacturers produce a minimum number of fuel cell and electric vehicles – when they become economically feasible.

The goal of the LEV regulation was to bring the state into compliance with the Clean Air Act. Meanwhile, the official goal of the ZEV mandate was to promote the development of advanced technology for zero emissions. In total, the policy may be seen as two-pronged. The LEV standards served a short-term goal of doing better with existing technology until an alternative was commercially available; functional requirements (emission standards) were established that were attainable in the near-term. The ZEV mandate was intended to serve the longer term goal of developing the alternative technology; zero emissions represented a long-term goal for the functional requirements of automobiles.

In general, setting targets in terms of functional requirements allows industry to determine the most cost effective technology available to meet policy goals. However,

[85] As modified in 2004, the LEV regulations contain standards for certifying vehicles as low emission (LEV), ultra-low emission (ULEV), and super ultra-low emission (SULEV) vehicles. For passenger cars, ULEV tailpipe emissions are reduced by 50-84% compared to federal Tier 1 emissions while SULEV tailpipe emissions are reduced by 76%-97%.

the ZEV mandate was widely regarded as a policy to promote one particular technology – electric vehicles – and this interpretation is supported by the most recent policy revision that specifies requirements for electric and fuel cell vehicles. Essentially, CARB officials attempted to use the opportunity presented by short-term local problems with moving forward in the current regime to create a niche for electric vehicles, which could serve as a solution to a long-term global problem. Proponents of the original rules were optimistic that a niche market would be created, eventually leading to commercially viable models. But after fifteen years, the rules had not been implemented and the market was still non-existent. As a means for advancing electric vehicle technology, the regulation had failed.

Using the lessons and language from this dissertation, we can propose a plausible explanation of what happened and why. First, automakers actively sought to alter the selection environment by lobbying for a delay in implementation and then challenging the legality of the regulations under federal law. Lobbying efforts claimed that the state-level rules were misaligned with the technology's and industry's capabilities. The legal challenge claimed that state-level rules were misaligned with federal-level rules.

Second, automakers responded to the altered selection environment by developing innovations within the existing socio-technological paradigm that required no adaptation on the part of users. As a result, before the 2003 effective date of the ZEV mandate, many conventional gasoline vehicles were capable of meeting the stringent SULEV emission standard established in the LEV regulations. Incremental improvements in electronic engine and catalytic control technologies were able to meet the overarching goals of the ZEV program – reducing automotive emissions. In fact, exhaust from a certified SULEV may be cleaner than the ambient air in major cities like Los Angeles.

Therefore, as an air quality policy, the LEV regulation was extremely successful in meeting its short-term goal. But if policy-makers hoped that the ZEV provision of the regulation would serve the longer term goal of facilitating a transition to a new technological paradigm, then this part of the policy has been a failure.

While electric or fuel cell vehicles may appear to be superior technologies for solving long-term problems as the transportation system continues to expand in scale, incremental improvements yielded products that were “good enough” at solving the short-term problem of meeting State Implementation Plans for the Clean Air Act. In addition, this solution was compatible (aligned with) with existing use systems (infrastructure and rules). Because these incremental improvements realigned the technology with new rules, both legislative and normative, regarding ‘clean’ vehicles, this success at least temporarily eliminated an opportunity for niche creation and eroded political support for longer-term goals.

At this point, it remains to be seen whether hybrid electric technology will achieve high volume production through niche accumulation or other mechanisms, thereby yielding technological improvements and cost reductions in battery and component technologies that can be incorporated in full electric vehicles. Such a pathway would be consistent again with Geels (2002) hybridization mechanism and the fit-stretch pattern.

6.6.3.2 CAFE

In 1973, OPEC production cuts and the Arab oil embargo resulted in a dramatic increase in petroleum fuel prices as well as shortages at gasoline stations. In an effort to decrease U.S. dependence on foreign oil, Congress passed the Energy Policy and

Conservation Act (EPAct) of 1975, which included regulations known as the Corporate Average Fuel Economy (CAFE) standards for light vehicles. Under CAFE, the sales-weighted fuel economy of each manufacturer's fleet of new vehicles must meet or exceed the CAFE standard, or the manufacturer must pay fines. When the standards were established, pick-up trucks were used primarily as work vehicles. Policy-makers set lower targets for light trucks in order to maintain the features and capabilities necessary to perform in that role. Figure 6-1 shows that, following passage of CAFE, the new car fleet average fuel economy rose from 15.8 mpg in 1978 to 28.6 mpg by 1988. It hovered at or below that value until 2002. New light truck average fuel economy increased from 13.7 mpg in 1975 to 21.6 mpg in 1987, but it declined slowly through 2002.

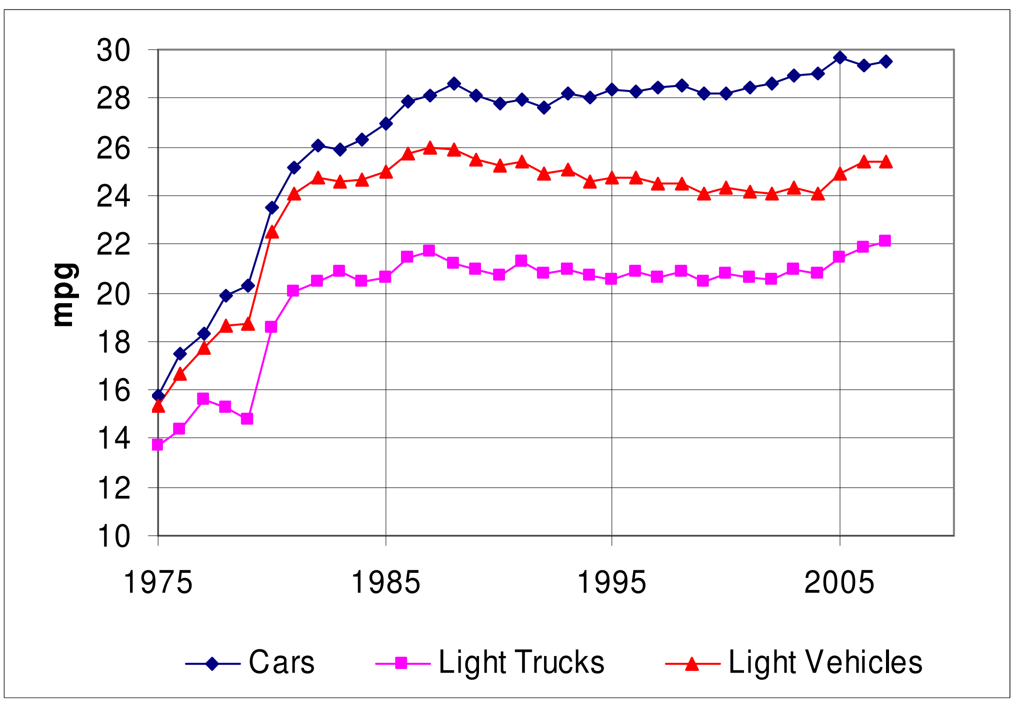


Figure 6-1 Average Fuel Economy of New Light Vehicles
Source: Sales weighted laboratory composite values calculated from U.S. EPA (2007b) Appendix E

Since the establishment of CAFE, the composition of the light vehicle market has changed. Figure 6-2 shows that while the market share of pick-up trucks has remained stable at about 15%, the market share of cars has decreased from about 80% to 50% while vans and sport utility vehicles have increased from 6% to 32%. Initially, most of the growth in this last category was in medium sized vans and SUVS. Since 1990, nearly all growth has been in medium and large SUVs.

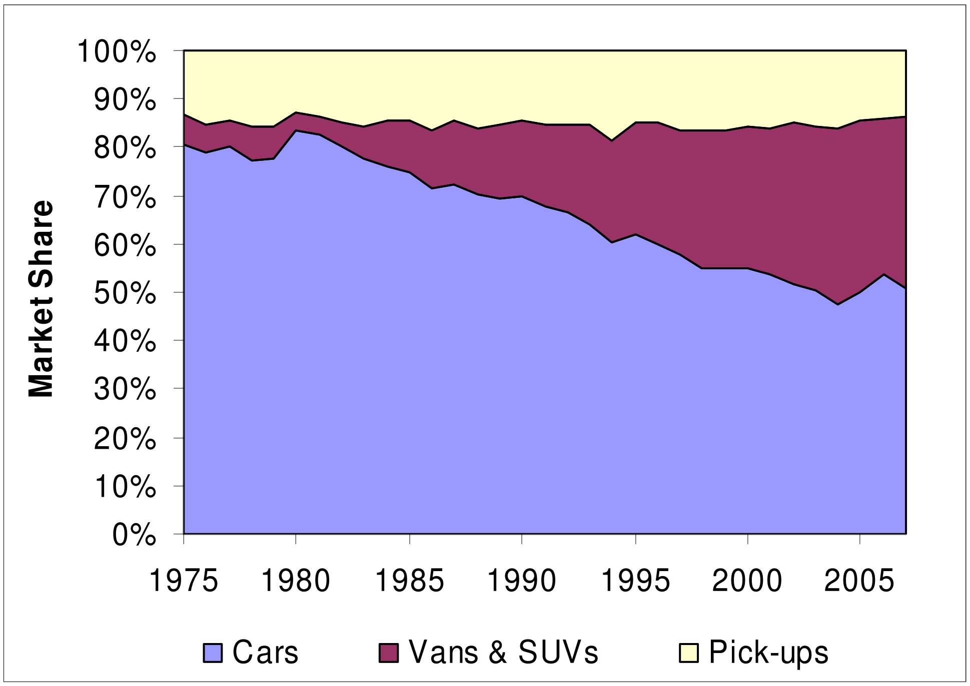


Figure 6-2 Market Share of Cars, SUVs, and Pick-ups
Source: U.S. EPA (2007b) Appendix E

Apparently, market preferences have shifted since the institution of CAFE standards. To try to understand why, consider the physical attributes of these vehicles. Figure 6-3, Figure 6-4, and Figure 6-5 show the EPA rated fuel economy, the sales-weighted average weight and horsepower-to-weight ratio for the same three classes of vehicles as well as the fleet average. The horsepower-to-weight ratio, or power density, generally indicates a vehicle's ability to accelerate and is therefore a good measure of performance or functional capability. In 1975, cars and pick-ups weighed essentially the same and had nearly equal performance. The industry's first response to meeting the new standards was to switch to smaller engines and lighten the vehicles by decreasing their overall size; decreasing the thickness, and thus the strength, of structural components; decreasing insulation; and lightening other components where possible. The resulting cars and trucks were lighter and more fuel efficient, but they were also less powerful as witnessed by the drop in power density, as well as noisier, less comfortable, and, many argue, less safe. However, once the CAFE standards were met, technical progress turned toward improving these attributes within the constraints of the regulation. Changes in frame and body design have increased strength, while improvements in engine design have produced significantly more power, both without compromising fuel economy.

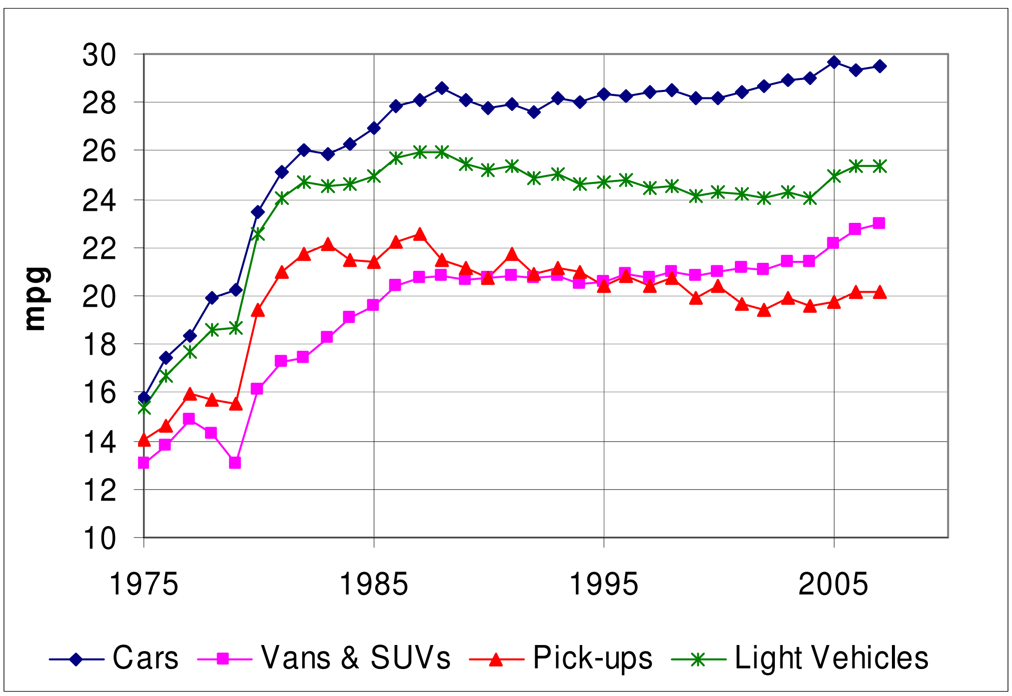


Figure 6-3 Average Fuel Economy of New Light Vehicles
Source: Sales weighted laboratory composite values calculated from U.S. EPA (2007b) Appendix E

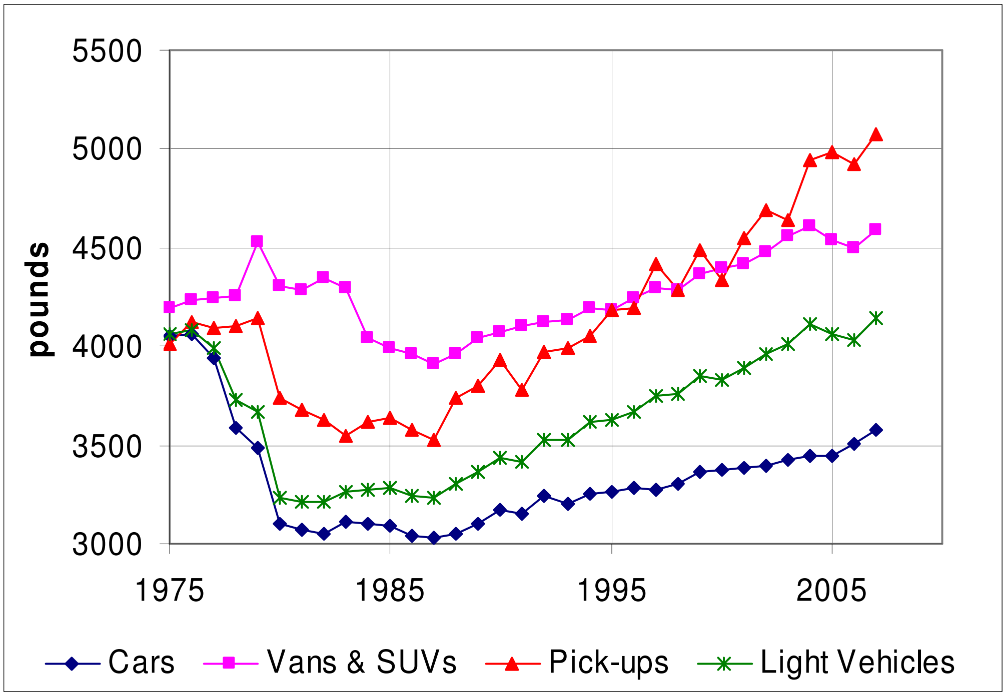


Figure 6-4 Average Weight of New Light Vehicles
Source: U.S. EPA (2007b) Appendix E

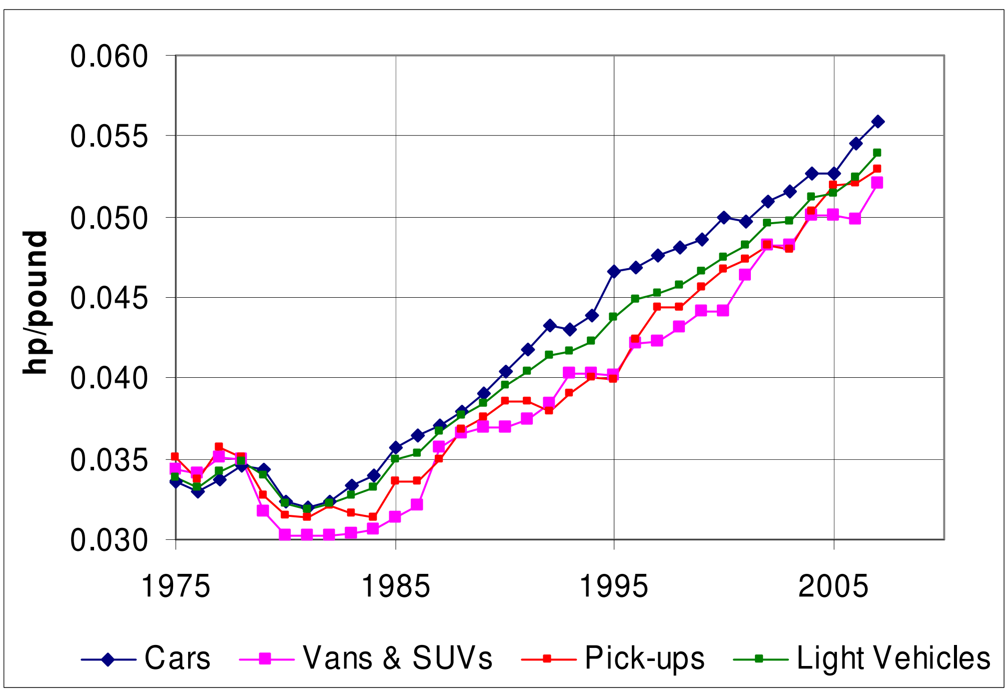


Figure 6-5 Average Power Density of New Light Vehicles
Source: U.S. EPA (2007b) Appendix E

What is more significant, however, is that the weight and performance of cars has diverged from that of light trucks. Despite weight increases since 1987, cars have remained substantially lighter than they were in 1975. Vans and SUVs are about the same weight as they were in 1980, and pick-ups are considerably heavier than they were. Most interestingly, the average light vehicle has returned to about the same weight as it was in 1975. In 1975, the performance of both truck classes was slightly better than that of cars, but cars now show a considerably higher power density than either truck class.

Based on what we have learned about functional definitions and co-evolutionary systems, we can propose a tentative theory for what has happened. In setting different standards for cars and light trucks, the CAFE policy actually created a functional distinction that did not exist and essentially defined a new class of vehicles. Cars were restricted to vehicles capable of achieving a certain fuel economy given 1970s technology. The attributes that made it difficult for light trucks to meet the more stringent car standard fulfilled a functionality that also was available in cars prior to CAFE, for example towing boats, campers, and trailers. To fulfill this function, consumers began showing increased preferences for features on vans and SUVs, of which there were very few models in 1975. Manufacturers responded by tailoring this class of vehicles into a set of products better suited to fulfill the functionalities no longer available in cars as well as those not available on pickup trucks. Eventually, this vehicle attained a set of attributes consistent with a new functional definition and likely incorporated some new functions as well. Meanwhile, the functional requirements and attributes of the car changed as well, becoming perhaps more divergent from both classes of light trucks. While the state of technology today might allow cars to meet both the CAFE standard for cars and the functions initially shifted to vans and SUVs, a sufficient solution already exists and may appeal to a new set of preferences that have evolved with the product.

Although CAFE successfully raised the average fuel economy of both cars and light trucks, the results in improving light vehicle fleet fuel economy were less than would be expected given market conditions at the time of promulgation. The policy did induce innovation, but it also altered the TIEC, resulting in the co-evolution of consumer preferences, firm strategies, and the technology. Meanwhile, the policy remained static except for minor adjustments in the actual value of the standard. As a result, the positive effects of the policy continued to be eroded until about 2004. While it is only conjecture, the reversal of this trend is likely due to high fuel prices brought on by several factors, including hurricane Katrina, difficulties at refineries in meeting new fuel formulation standards, the war in Iraq, and market speculation in oil futures. Further, the CAFE standard for light trucks was revised upward in 2006.

Again, this explanation of the outcomes of CAFE does not represent a thorough analysis and would obviously benefit from a careful analysis of the data and other evidence. But it does serve to illustrate unintended consequences of policy intervention in co-evolutionary systems.

6.6.3.3 Summary

In light of this research, effective policy must consider the functional requirements of the technology as defined by consumer demands and preferences; potential strategies of firms in responding to policy and the market; and the potential for new innovations and developments in the TIEC to alter either consumer preferences or firms' entrepreneurial opportunities, or both. In addition, policy should provide for feedback and adaptation to changing realities. This last consideration is extremely tricky since leaving policies open for adaptation could mean allowing the potential for revisions that run contrary to the policy's original goals, especially if political influences change.

In addition, this chapter has repeatedly highlighted the role of multiple developments and processes occurring within niches, within multiple regimes, and within the wider TIEC in the emergence of the motor vehicle, its diffusion into wider markets, and the initiation of new transitional phases or eras of innovative ferment. These developments include:

1) new scientific or technological breakthroughs;
2) difficulties moving forward within the existing regime(s) due to:
    a. problems arising from the expanding scale and scope of the existing regime(s);
    b. approaching technological limits, real or perceived, of the existing paradigm;
3) shifting consumer preferences;
4) misalignments between wider developments in the TIEC (e.g. cultural values, world views, etc.) and the existing regime(s); and
5) unexpected events.

This result can be interpreted from two perspectives, one passive and one active. In the passive interpretation, new technologies emerge and transitions occur when multiple developments and misalignments result in a more favorable selection environment. In the active interpretation, transitions occur when entrepreneurs *exploit* the opportunities presented by multiple developments and misalignments. While the difference is subtle, it leads to two different and complementary strategies. First, policymakers need to develop adaptive policies as discussed above. This requires that they gather sufficient information to sense and respond to the challenges these developments pose to existing policy regimes. Second, policymakers should be taking the active role of entrepreneur, searching for, linking, and exploiting opportunities as they arise.

6.7 Future Research

This research examined the mechanisms of innovation and processes of co-evolution in the transition to motor vehicles. The findings presented in this chapter raised several questions for additional research that would add depth to the current study. Meanwhile, the goal of this research is to inform policy aimed at influencing such a transition from the current regime. Therefore, another natural extension of this research would add breadth. This leads to two agendas for future research regarding innovation and transitions in the automobile industry.

6.7.1 Agenda for Adding Depth

This research found support for the argument that the gradual unfolding of the automobile regime represented an internal transition for the carriage and bicycle industries that evolved into a distinct regime in 1926. Two follow-on studies are suggested to further explore this argument. The first would apply the framework used in this study to examine the response of the carriage industry during the same era to determine how disruptive this transition was to the existing regime. Specifically, did the regime adapt or did it enter senescence? The second study would develop quantifiable definitions and appropriate measures of 'alignment,' and 'stability' and a definition of 'regime' based on these measures. This definition would be developed using a set of industries that have experienced misalignments, instabilities, and transitions. It would then be applied to the automobile history to determine when that industry achieved the status of a regime.

A third study that would add depth to the current research would examine the questions raised regarding producer-supplier relationships. Suppliers were identified as a major source of innovation between 1900 and 1927. This study would focus on suppliers that provide parts and components directly to manufacturers (tier one) in the current era. Specifically, this study would address whether suppliers continue to be major originators or collaborators in innovation; how suppliers stay in touch with the needs of end-consumers; and how the producer-supplier relationship influences innovative behavior.

6.7.2 Agenda for Adding Breadth

This agenda consists of three studies. The first would apply the analytic framework used in the current study to a more contemporary era for the automobile industry beginning in the 1970s. This study would examine: 1) sources of instability and the onset of transitional and specific phase(s); 2) changes in consumer preferences and the product functional definition; 3) the emergence of a dominant design; and 4) processes of innovation and adaptation.

The second study would make a more rigorous examination of adaptive consumer preferences in the same timeframe. This study would take advantage of the wealth of data available and augment the framework used in the current research with statistical analysis. This research would examine how consumers make purchasing decisions, how consumer preferences change in this co-evolutionary context, and how these preferences and processes are manifested in the market.

The third and final study would apply the lessons learned in the current research and the two more contemporary studies outlined in this section to analyze recent policy measures, including the Corporate Average Fuel Economy (CAFE) standards and the California Zero Emission Vehicle (ZEV) mandates.

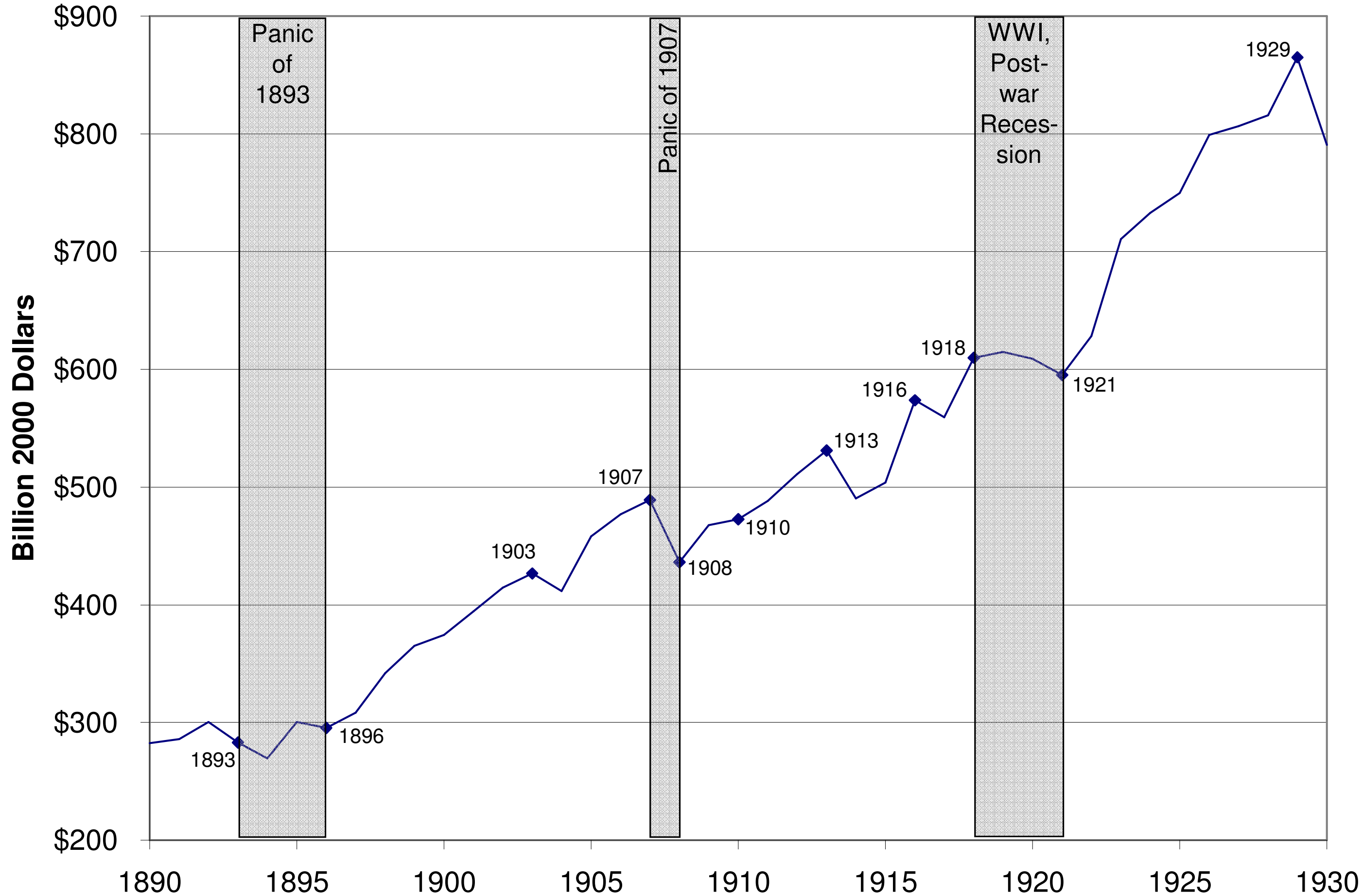
Appendix A: Vehicle Sales and Registrations, 1900-1930

	Automobiles (1,000s)		Trucks and Buses (1,000s)		Ford (1,000s)
	Sales	Registrations	Sales	Registrations	Sales*
1900	4.1	8.0			
1901	7.0	14.8			
1902	9.0	23.0			
1903	11.2	32.9			
1904	22.1	54.6	0.7	0.7	1.7
1905	24.2	77.4	7.0	1.4	1.7
1906	33.2	105.9	0.8	2.2	1.6
1907	43.0	140.3	1.0	2.9	8.4
1908	63.5	194.4	1.5	4.0	6.4
1909	123.9	306.0	3.2	6.1	10.6
1910	181.0	458.4	6.0	10.1	18.7
1911	199.3	618.7	10.6	20.8	34.5
1912	356.0	901.6	22.0	42.4	78.4
1913	461.5	1190.4	23.5	67.7	168.3
1914	548.1	1664	24.9	99.0	248.3
1915	895.9	2490.9	74.0	158.5	472.4
1916	1525.5	3617.9	92.1	250.0	730.0
1917	1745.7	5118.5	128.1	391.1	656.2
1918	943.4	6160.4	227.2	605.5	498.3
1919	1651.6	7576.9	224.7	897.8	941.0
1920	1905.5	9239.2	321.7	1107.6	463.5
1921	1468.0	10493.7	148.0	1281.5	971.6
1922	2274.1	12273.6	269.9	1569.5	1306.6
1923	3624.7	15102.1	409.2	1849.1	2019.0
1924	3185.8	17612.9	416.6	2176.8	1929.1
1925	3735.1	20068.5	530.6	2587.5	1920.1
1926	3692.3	22200.2	608.6	2932.2	1565.3
1927	2936.5	23303.5	464.7	3110.2	424.4
1928	3775.4	24688.6	583.3	3326.4	750.3
1929	4455.1	26704.8	881.9	3583.9	1870.3
1930	2787.4	26749.9	575.3	3715.1	1451.6

* Ford data for 1904-1917 is for cars only for the model year ending in July of the year given. Data for 1918-1930 is for the calendar year and includes trucks.

Sources: Industry – L.P. Cain (2006) Tables Df343-346 and Df339-342; Ford – Nevins and Hill (1954 and 1957)

Appendix B: Real GDP, 1890-1930



Source: Johnston and Williamson, 2007.

Glossary

Emergent phase – the emergent phase of a technology begins with the first practical application of a new idea (an innovation). This phase is characterized by radical innovation, flexible production capacity, niche markets, and the emergence of new firms. See Table 2-1.

Exogenous factor - exogenous factors (events or developments) are external forces that arise with no identifiable causal linkages to the regime or technology of interest.

Experimental niche –non-commercial exploration of a technology in a temporary societal experiment where it is protected from market selection pressures and from the engineering and use practices established by the existing socio-technical regime.

Fordism –the production system developed at the Ford Motor Company between 1908 and 1913 entailing large volume production of a single product with centralized control. Elements include highly standardized parts, extensive use of specially-designed single-purpose machinery, minute subdivision of labor, and the automated flow of materials and work (continuous motion).

Hybridization – the merging of functions, features, and technologies, both old and new, from multiple regimes to solve particular problems or provide new functions through symbiosis.

Industrial combination – an aggregation of corporations under a central or controlling corporation. Use of the term 'combination' around 1900 was preferred among businessmen over the term 'trust,' which implied the existence and exercise of monopolistic powers.

Institutions – any form of social construct or constraint on the interaction between human agents and between human agents and the physical artifacts in a system. Institutions include organizations and rules.

Mass markets – application domains that are large in scale and scope and that are characterized by functionalities with wide appeal. Demand is sufficient to allow high volume production, resulting in economies of scale. Competition is based mainly on price and quality.

Niche branching – the use of a radical innovation in subsequent application domains or market niches.

Niche hypothesis – The assertions of the quasi-evolutionary and multi-level perspective literature regarding the fundamental role of niches in technological transitions.

Niche markets – distinct application domains that are small in scale and scope and that are characterized by specific functional requirements. Because of the special functional requirements of the niche, users are willing to accept higher cost or lower performance on other non-essential functionalities compared to those attained with any existing products available in established (mass) markets.

Regime optimization – incremental improvement within an existing regime.

Regime renewal – radical innovations that depart from the existing trajectory and require a change in the architecture and rules of the existing regime and therefore in the behavior of various actors.

Regime instability – an existing technology looses viability due to emerging misalignments between the elements of the system, including the technology, rules and other institutions, and infrastructure. Regime instability may also involve exogenous factors.

Retention – the embedding of technologies in engineering practices, production systems, organizational routines, consumption and usage patterns, and supporting and complimentary systems that result in the stability of search activities and the patterns of technological change.

Rules – cognitive, normative, or legislative constraints on human interactions and the interactions between human agents and the physical elements of a system. Rules may be formal or informal and include legislation, economic rules, legal contracts, social conventions, moral codes, world views, and shared knowledge and beliefs.

Sectoral system – a set of new and established products for specific uses and the set of agents carrying out market and non-market interactions for the creation, production, sale and use of those products.

Selection - "[a]ny framework in which agents interact in order to choose between competing patterns of behaviour" (Metcalfe, 1995). Selection involves social, political, and market mechanisms and is shaped by institutions and infrastructure.

Senescent phase – the phase of a product's lifecycle when a new innovation begins to dominate that product's market. The old technology finds application only in niche markets where it holds a performance advantage or where heritage design is required. Innovation slows or stops and is tailored to the needs of niche users. Eventually, firms are unable to survive and the industry ceases to exist. See Table 2-1.

Socio-technical regime – a socio-technical regime is understood to exist when a technology and sector have achieved the specific phase and when the stability of that specific phase is supported by the alignment of institutions and existing physical infrastructure with the dominant product design, production processes, and business model.

Specific phase – the phase of a product's lifecycle when products proliferate around a dominant design and ever more emphasis is given to process innovation. There is little product differentiation and competition in mass markets is based primarily on cost and quality. See Table 2-1.

Spillover – an innovation that originates outside the sectoral system – in other industries, institutions or other countries. Spillovers may involve either a transfer of knowledge (knowledge spillovers) or economic benefits (rent spillovers). This research is concerned with knowledge spillovers that arise outside the motor vehicle sector.

Strategic niche management (SNM) – the creation, development, and controlled phase-out of protected spaces for the development and use of promising technologies by means of experimentation. SNM is a policy recommendation that arises from the quasi-evolutionary literature.

Synthetic innovation – "creating new and enhanced technological products and processes with previously unattained performance by combining components, knowledge, and capabilities in ways that deliver synergism" (Rycroft and Kash, 1999).

Technological nexus – the set of institutional linkages between selection and variation. Agents acting in the technological nexus role link selection and variation by translating information from both realms; shaping interactions; and harmonizing social and market needs with the results of scientific and technological research. These actions result in the mutual adaptation of the technology and institutions.

TIEC – technical, institutional, and ecological complex – the larger system within which a sectoral system is embedded and which serves as the sectoral system's environment. It includes physical artifacts, agents, and institutions; elements include meso- and macro-level constructs including physical infrastructures, regimes, regionally or nationally established rules, cultural values and culturally learned patterns of behavior, *etc*. See section 2.6 and Figure 2-3.

Transitional phase – the phase of a product's lifecycle when a dominant design begins to coalesce as users' needs, product features, design, and components become standardized. The product diffuses into mass markets. Innovation slows, becomes incremental, and is focused on cost reduction and quality enhancements. See Table 2-1.

Variation – the purposive and guided, yet inherently stochastic, generation of diversity or novelty through human creativity and of the discovery of new possibilities for action.

References

Abernathy, W.J. (1978) *The Productivity Dilemma: Roadblock to Innovation in the Automobile Industry.* Baltimore, MD: Johns Hopkins University Press.

Abernathy, W.J., K.B Clark and A.M Kantrow (1983) *Industrial Renaissance: Producing a Competitive Future for America.* New York, NY: Basic Books, Inc.

Abernathy, W.J., and J.M. Utterback (1978) Patterns of industrial innovation, *Tech. Rev.* Jun/Jul, 40-47.

Afuah, A.N. and J.M. Utterback (1997) Responding to structural industry changes: A technological evolution perspective, *Industrial and Corporate Change* 6 (1), 183-202.

Alchian, A. (1963) Reliability of progress curves in airframe production, *Econometrica* 31 (4), 679-693.

Ahmad, S. (1966) On the theory of induced invention, *Economic Journal* 76 (302), 344-357.

Anderson, P., and M.L. Tushman (1990) Technological discontinuities and dominant designs: a cyclic model of technological change, *Administrative Science Quarterly* 35 (4), 604-633.

Argote, L. and D. Epple (1990) Learning curves in manufacturing, *Science* 247 (4945) 920-924.

ASME (2008) http://anniversary.asme.org/history.shtml, accessed March 25, 2008.

Baloff, N. (1966) The learning curve – some controversial issues, *J. Industrial Econ.* 14 (3) 275-282.

Barach, J. (2007) *Automobile History*. http://www.motorera.com/history/history.htm, accessed June 28, 2007.

Barnes, J.K. (1921) The men who “standardized” automobile parts, *The World’s Work* 42 (2), 204-208.

Cain, L.P. (2006) Motor-vehicle factory sales, by vehicle type: 1900–1996, and Motor vehicle registrations, by vehicle type: 1900–1995, in S.B. Carter, S.S. Gartner, M.R. Haines, A.L. Olmstead, R. Sutch, and G. Wright, *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/ISBN-9780511132971.Df184-57710.1017/ISBN-9780511132971.Df184-577

Chandler, A.D. and S. Salsbury (1971) *Pierre S. du Pont and the Making of the Modern Corporation*. New York, NY: Harper & Row, Pub., Inc.

Christensen, C.M. (1997) *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston, MA: Harvard Business School Press.

Clarke, L., J. Weyant, and A. Birky (2006) On the sources of technological change: Assessing the evidence, *Energy Economics* 28, 579-595.

Coase, R.H. (2000) The acquisition of Fisher Body by General Motors, *J. Law and Econ.* 43 (1), 15-31.

Cohen, W.M. and D.A. Levinthal (1989) Innovation and learning: The two faces of R&D, *Econ. J.* 99 (397), 569-596.

Cohen, W.M. and D.A. Levinthal (1990) Absorptive capacity: A new perspective on learning and innovation, *Admin. Science Quarterly* 35 (1), 128-152.

Conceiçãoa, P, M.V. Heitora, F.Veloso (2003) Infrastructures, incentives, and institutions: Fostering distributed knowledge bases for the learning society, *Tech. Forecasting & Social Change* 70 (7) 583–617

Cowan, R. and S. Hultén (1996) Escaping lock-in: The case of the electric vehicle, *Tech. Forecasting Soc. Change* 53 (1), 61-80.

Cray, E. (1980) *Chrome Colossus.* New York, NY: McGraw-Hill.

Day, G.S. and D.B. Montgomery (1983) Diagnosing the experience curve, *J. Marketing* 47, 44-58.

Dijkema, G.P.J., P. Ferrãob, P.M. Herdera, and M. Heitor (2006) Trends and opportunities framing innovation for sustainability in the learning society, *Tech. Forecasting and Social Change* 73, (3), 215-227.

Dosi, G. (1982) Technological paradigms and technological trajectories, *Research Policy* 11, 147-162.

Dosi, G. (1988) Sources, procedures, and microeconomic effects of innovation, *J. Econ. Lit* 26 (3), 1120-1171.

Dosi, G. (1997) Opportunities, incentives and the collective patterns of technological change, *Econ. J.* 107, 1530-1547.

Du Pont (2007) http://heritage.dupont.com, accessed September 7, 2007.

Dutton, J.M. and A. Thomas (1984) Treating progress functions as a managerial opportunity, *Acad. Mgt. Rev.* 9 (2), 235-247.

EIA (2007) *Annual Energy Review 2006*, Report No. DOE/EIA-0384(2006).

Elzen, B., R. Hoogma, and R. Kemp (2003) Managing the transition to sustainable transport through strategic niche management, in *Transportation, Energy, and*

Environmental Policy: Managing Transitions, VIII Biennial Asilomar Conference, September, 2001, 175-203. Washington, DC: National Academies Transportation Research Board.

Encyclopedia Britannica (1926) Mass production, in *The Encyclopedia Britannica*, 13th ed., vol. 30, 821-23. New York, NY: The Encyclopædia Britannica Co., Ltd.

Epstein, R.C. (1972) *The Automobile Industry: Its Economic and Commercial Development.* New York, NY: Arno Press. Reprint of the 1928 edition.

Flink, J.J. (1970) *America Adopts the Automobile, 1895-1910*. Cambridge, MA: MIT Press.

Flink, J.J. (1988) *The Automobile Age*. Cambridge, MA: MIT Press.

Ford, H. and S. Crowther (1922) *My Life and Work*. Garden City, NY: Garden City Publishing Co.

Freedman, D.H. (1992) Is management still a science? *Harv. Bus. Rev.* Nov./Dec., 26-38.

Freeland, R.F. (2000) Creating holdup through vertical integration: Fisher Body revisited, *J. Law and Economics* 43 (1), 33-66.

Freeman, C. (1982) *The Economics of Industrial Innovation*. Cambridge, MA: MIT Press.

Freeman, C. (1994) The economics of technical change, *Cambridge J. Econ.* 18, 463-514.

Freeman, C. (1996) The greening of technology and models of innovation, *Technological Forecasting and Social Change* 53, 27-39.

Geels, F.W. (2002) Technological transition as evolutionary reconfiguration processes: a multi-level perspective and a case-study, *Research Policy* 31, 1257-1274.

Geels, F.W. (2004) From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory, *Research Policy* 33, 897-920.

Geels, F.W. (2005a) Processes and patterns in transitions and system innovations: Refining the co-evolutionary multi-level perspective, *Tech. Forecasting & Social Change* 72, 681-696.

Geels, F.W. (2005b) The dynamics of transitions in socio-technical systems: A multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860-1930), *Technology Analysis & Strategic Management* 17 (4), 445-476.

Gharajedaghi, J. (2006) *Systems Thinking: Managing Chaos and Compleixity*. Boston, MA: Butterworth-Heinemann.

Ghemawat, P. (1985) Building strategy on the experience curve, *Harv. Bus. Rev.* March/April, 143-149.

Godin, B. (2005) The linear model of innovation: the historical construction of an analytical framework. Project on the History and Sociology of S&T Statistics, Working Paper No. 30. Montreal, Quebec: Canadian Science and Innovation Indicators Consortium. Available at http://www.csiic.ca/PDF/Godin_36.pdf accessed July 20, 2006.

Grayson, S. (1978) The all-steel world of Edward Budd, *Automobile Quarterly* 16 (4), 352-367.

Griliches, Zvi (1957) Hybrid corn: an exploration in the economics of technical change, *Econometrica* 25, 501-522.

Grübler, A. (1998) *Technology and Global Change*. Cambridge, UK: Cambridge University Press.

Grübler, A., N. Nakićenović, and D.G. Victor (1999a) Dynamics of energy technologies and global change, *Energy Policy* 27, 247-280.

Grübler, A., N. Nakićenović, and D.B. Victor (1999b) Modeling technological change: Implications for the global environment, *Annu. Rev. Energy Environ* 24, 545-569.

Hayes, R.H. and Abernathy, W.J. (1980) Managing our way to economic decline, *Harvard Business Review* July/Aug., 67-77.

Hilton, G.W. (1969) Transport technology and the urban pattern, *J. Contemporary History* 4 (3), 123-135.

Hirsch, W.Z. (1952) Manufacturing progress functions, *Rev. Econ. Stat.* 34 (2) 143-155.

Hirsch, W.W. (1956) Firm progress ratios, *Econometric* 24 (2), 136-143.

Hodbay, M. (1998) Product complexity, innovation and industrial organisation, *Research Policy* 26, 689-710.

Holdren, J.P. (2006) The energy innovation imperative: Addressing oil dependence, climate change, and other 21st century energy challenges, *Innovations* 1 (2), 3-23.

Holling, C.S. (1994) Simplifying the complex: the paradigms of ecological function and structure, *Futures* 26 (6), 598-609.

Hoogma, R., R. Kemp, J. Schot, and B. Truffer (2002) *Experimenting for Sustainable Transport: The Approach of Strategic Niche Management*, New York, NY: Spon Press.

Hounshell, D.A. (1984) *From the American System to Mass Production, 1800-1932*. Baltimore, MD: Johns Hopkins University Press.

Hugill, P.J. (1982) Good roads and the automobile in the United States 1880-1929, *Geographical Review* 72 (3), 327-349.

Hyde, C.K. (2003) *Riding the Roller Coaster: A History of the Chrysler Corporation*. Detroit, MI: Wayne State University Press.

Jarvis, G.K. (1972) The diffusion of the automobile in the United States: 1895-1969. PhD dissertation, University of Michigan.

Johnson, L., D. Greene and A. Birky (2003) Is the barrel half full or half empty? Implications of transitioning to a new transportation energy future, in *Transportation, Energy, and Environmental Policy: Managing Transitions,* report of the VIII Biennial Asilomar Conference, September, 2001, 104-129. Washington, DC: National Academies Transportation Research Board.

Johnston. L.D. and S.H. Williamson (2007), What Was the U.S. GDP Then? http://www.measuringworth.com/usgdp/, accessed March 26, 2008.

Joskow, P.L. and N.L. Rose (1985) The effects of technological change, experience, and environmental regulation on the construction cost of coal-burning generating units, *RAND J. Econ.* 16 (1), 1-27.

Kanigel, R. (1996) Frederick Taylor's apprenticeship, *Wilson Quarterly* 20 (3), 44-51.

Kanigel, R. (1997) *The One Best Way: Frederick Winslow Taylor and the Enigma of Efficiency*. New York, NY: Viking.

Kantrow, A.M. (1980) The strategy-technology connection, *Harvard Bus. Rev.* 58 (4), 6-21.

Kemp, R. (1994) Technology and the transition to environmental sustainability, *Futures* 26 (10), 1023-1046.

Kemp, R., J. Schot and R. Hoogma (1998) Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management, *Technology Analysis & Strategic Management* 10 (2), 175-195.

Kirsch, D.A. (2000) *The Electric Vehicle and the Burden of History.* New Brunswick, NJ: Rutgers University Press.

Klepper, S. and K.L. Simons (1997) Technological extinctions of industrial firms: an inquiry into their nature and causes, *Industrial and Corporate Change* 6 (2), 379-460.

Klepper, S. and K.L. Simons (2005) Industry shakeouts and technological change, *International J. Industrial Organization* 23, 23-43.

Kovarik, W.J. (1994) Charles F. Kettering and the 1921 discovery of tetraethyl lead in the context of technological alternatives. Presentation to the Society of Automotive Engineers Fuels & Lubricants Conference, Baltimore, MD. http://www.runet.edu/~wkovarik/papers/kettering.html, accessed October 12, 2007.

Langlois, R.N. (2002) Schumpeter and the Obsolescence of the Entrepreneur. UConn Department of Economics Working Paper No. 2002-19. Available at SSRN: http://ssrn.com/abstract=353280

Leslie, S.W. (1979) Charles Kettering and the copper-cooled engine, *Technology and Culture* 20 (4), 752-776.

Lieberman, M.B. (1984) The learning curve and pricing in the chemical processing industries, *RAND J. Econ.* 15 (2), 213-228.

Malerba, F. (1992) Learning by firms and incremental technical change, *Econ. J.* 102 (413), 845-859.

Malerba, F. (2002) Sectoral systems of innovation and production, *Research Policy* 31, 247-264.

May, G. S., ed. (1990) *The Automobile Industry, 1896-1920.* New York, NY: Facts on File.

Melaina, M. W. (2007) Turn of the century refueling: A review of innovations in early gasoline refueling methods and analogies for hydrogen, *Energy Policy* (35), 4919-4934.

Metcalfe, J.S. (1994) Evolutionary economics and technology policy, *Econ. J.* 104 (425), 931-944.

Metcalfe, J.S. (1995) Technology systems and technology policy in an evolutionary framework, *Cambridge J. Econ.* 19, 25-46.

Montgomery, D.B. and G.S. Day (1985) Experience curves: evidence, empirical issues, and applications, in H. Thomas and D. Gardner, eds., *Strategic Marketing and Management*. New York, NY: John Wiley & Sons Ltd.

Mowery, D.C., and Rosenberg, N. (1979) The influence of market demand upon innovation: A critical review of some recent empirical studies, *Research Policy* 8, 102-153.

Mowery, D.C. and Rosenberg, N. (1998) *Paths of Innovation: Technological Change in 20th Century America*. New York, NY: Cambridge University Press.

Nakićenović, N. (1986) The automobile road to technological change: diffusion of the automobile as a process of technological substitution, *Technological Forecasting and Social Change* 29, 309-340.

Nelson, R.R. and R.N. Langlois (1983) Industrial innovation policy: lessons from American history, *Science* 219 (4586), 814-818.

Nelson, R.R. and S.G. Winter (1977) In search of useful theory of innovation, *Research Policy* 6, 36-76.

Nevins, A. and F.E. Hill (1954) *Ford: The Times, the Man, the Company*. New York, NY: Charles Scribner's Sons.

Nevins, A. and F.E. Hill (1957) *Ford: Expansion and Challenge 1915-1933*. New York, NY: Charles Scribner's Sons.

Nieuwenhuis, P. and P. Wells (2007) The all-steel body as a cornerstone to the foundations of the mass production car industry, *Industrial and Corporate Change* 16 (2), 183-211.

Nordhaus, W. D. (1973) Some skeptical thoughts on the theory of induced innovation, *Quarterly Journal of Economics* 87 (2), 208-219.

Nottingham, E.K. (1947) Toward an analysis of the effects of two World Wars on the role and status of middle-class women in the English-speaking world, *Amer. Soc. Rev.* 12 (6), 666-675.

O'Connor, M. (1994) Complexity and coevolution: Methodology for a positive treatment of indeterminacy, *Futures* 26 (6), 610-615.

Pavitt, K. (1984) Sectoral patterns of technical change, *Research Policy* 13, 343-373.

Pound, A. (1934) *The Turning Wheel: The Story of General Motors Through Twenty-five Years 1908-1933*. New York, NY: Doubleday, Doran & Co., Inc.

Price, B. (1990) Frank and Lillian Gilbreth and the motion study controversy, 1907-1930, in D. Nelson, ed., *A Mental Revolution: Scientific Management since Taylor.* Columbus, OH: Ohio State University Press.

Rae, J.B. (1959) *American Automobile Manufacturers: The First Forty Years.* Philadelphia, PA: Chilton Co.

Rae, J.B. (1965) *The American Automobile: A Brief History.* Chicago, IL: Univ. of Chicago Press.

Raff, D.M.G. (1991) Making cars and making money in the interwar automobile industry: economies of scale and scop and the manufacturing behind the marketing, *Business History Review* 65 (4), 721-753.

Raff, D.M.G. and M. Trajtenberg (1997) Quality-adjusted prices for the American automobile industry: 1906-1940, in T.F. Bresnahan and R.J. Gordon, eds., *The Economics of New Goods*. Chicago, IL: University of Chicago.

Rapping, L. (1965) Learning and World War II production functions, *Rev. Econ. Statistics* 47 (1), 81-86.

Raven, R. (2005) Strategic niche management for biomass: A comparative study on the experimental introduction of bioenergy technologies in the Netherlands and Denmark. PhD dissertation, Technische Universiteit Eindhoven.

Revelle, R., and H.E. Suess (1957) Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO2 during the past decades, *Tellus* 9 (1), 18-27.

Rip, A. (1995) Introduction of new technology: Making use of recent insights from sociology and economics of technology, *Technology Analysis & Strategic Management* 7 (4), 417-431.

Rip, A. and R. Kemp (1998) Technological Change, in Rayner, S. and E. L. Malone, eds., *Human Choice and Climate Change, Vol. 2: resources and Technology*, 327-399.

Rosenberg, N. (1963) Technological change in the machine tool industry, 1840-1910, *J. Econ. History* 23 (4), 414-443.

Rosenberg, N. (1976) *Perspectives on Technology.* Cambridge, UK: Cambridge University Press.

Rosenberg, N. (1994) *Exploring the Black Box: Technology, Economics, and Theory.* Cambridge, UK: Cambridge University Press.

Rosenberg, N. (1996) Uncertainty and technological change, in R. Landau, T. Taylor, and G. Wright, eds., *The Mosaic of Economic Growth.* Stanford, CA: Stanford University Press.

Rubenstein, J.M. (2001) *Making and Selling Cars: Innovation and Change in the U.S. Automotive Industry.* Baltimore, MD: Johns Hopkins University Press.

Ruttan, V.W. (1997) Induced innovation, evolutionary theory and path dependence: sources of technical change, *Econ. J.* 107, 1520-1529.

Rycroft, R.W. and D.E. Kash (1999) *The Complexity Challenge: Technological Innovation for the 21st Century.* London, UK: Pinter Publishers.

Rycroft, R.W. and D.E. Kash (2002) Path dependence in the innovation of complex technologies, *Technology Analysis & Strategic Management* 14 (1), 21-35.

SAE (2007) An abridged history of the SAE, http://www.sae.org/about/general/history/, accessed September 28, 2007.

Sahal, D. (1981) Alternative conceptions of technology, *Research Policy* 10, 2-24.

Schon, D. (1967) *Technology and Change*. New York, NY: Delacorte.

Schmookler, J. (1962) Determinants of industrial invention, in R.R. Nelson, ed., *The Rate and Direction of Inventive Activity: Economic and Social Factors.* Princeton, NJ: Princeton University Press.

Schmookler, J. (1966) *Invention and Economic Growth*. Cambridge, MA: Harvard University Press.

Schumpeter, J.A. (1939) *Business cycles: a theoretical, historical and statistical analysis of the capitalist process*. New York, NY: McGraw Hill.

Schot, J.W. (1992) Constructive technology assessment and technology dynamics: The case of clean technologies, *Science, Technology, & Human Values* 17 (1), 36-56.

Schot, J., R. Hoogma, and R. Elzen (1994) Strategies for shifting technological systems, *Futures* 26 (10), 1060-1076.

Schumpeter, J.A. (1942) *Capitalism, socialism and democracy*. New York, NY: Harper.

Simons, K.L. (1995) Shakeouts: Firm Survival and Technological Change in New Manufacturing Industries. PhD dissertation, Carnegie Mellon University.

Sloan, A.P. (1964) *My Years with General Motors*. Garden City, NY: Doubleday & Co., Inc.

Smith, P.H. (1968) *Wheels within Wheels*. New York, NY: Funk & Wagnalls.

Smith, K. (2000) Innovation as a systemic phenomenon: rethinking the role of policy, *Enterprise & Innov. Management Studies*, 1 (1), 73-102.

Smits, R. (2002) Innovation studies in the 21st century: Questions from a user's perspective, *Tech. Forecasting and Social Change* 69 (9), 861-883.

Soete, L. and A. Arundel (1993) *An Integrated Approach to European Innovation and Technology Diffusion Policy: A Maastricht Memorandum.* Commission of the European Communities, SPRINT Programme, Luxembourg.

Solow, R.M. (1957) Technical change and the aggregate production function, *Rev. Econ. Statistics* 39 (3), 312-320.

Sorensen, C.E. (1956) *My Forty Years with Ford.* New York, NY: W.W. Norton & Company, Inc.

Stigler, G.J. (1951) The division of labor is limited by the extent of the market, *J. Political Econ.* 59 (3), 185-193.

Tallman, E.W. and J.R. Moen (1990) Lessons from the panic of 1907, *Economic Review – Federal Reserve Bank of Atlanta* (75) May/Jun 1990, 2-12.

Taylor, F.W. (1895) A piece rate system, *Trans. Amer. Society Mech. Engineers* XVI, 856-883.

Taylor, F.W. (1911) *Shop Management*. New York, NY: Harper and Brothers.

Tushman, M.L. and L. Rosenkopf (1992) Organizational determinants of technological change: Toward a sociology of technological evolution, *Research in Organizational Behavior* 14, 311-347.

Unruh, G.C. (2000) Understanding carbon lock-in, *Energy Policy* 28, 817-830.

U.S. EPA (2007a) *Inventory of U.S. Greenhouse Gas Emissions And Sinks: 1990-2005,* Report No. EPA 430-R-07-002, April. Washington, DC: USEPA.

U.S. EPA (2007b) *Light-Duty Automotive Technology and Fuel Economy Trends: 1975 Through 2007*, Report No. EPA420-R-07-008, September. Washington, DC: USEPA.

U.S. Patent Office (2008) U.S. Patent No. 549,160 dated November 5, 1895. U.S. Patent and Trademark Office website, http://www.uspto.gov/patft/index.html, accessed March 13, 2008.

Utterback, J.M. and W.J. Abernathy (1975) A dynamic model of product and process innovation, *Omega* 3 (6).

Weisberger, B.A. (1979) *The Dream Maker: William C. Durant, Founder of General Motors.* Boston, MA: Little, Brown and Co.

Whipp, R. and P. Clark (1986) *Innovation and the Auto Industry*. New York, NY: St. Martin's Press.

Wik, R.J. (1951) Steam power on the American farm, *Agricultural History* 25 (4), 181-186.

Witt, U. (1992) Evolutionary concepts in economics, *Eastern Econ. J.* 18 (4), 405-419.

Womack, J.P, D.T. Jones, and D. Roos (1990) *The Machine that Changed the World: The Story of Lean Production*. New York, NY: Macmillan Publishing Co.

Wright, T.P. (1936) Factors affecting the cost of airplanes, *J. Aeronaut. Sci.* 3, 122-128.

Yelle, L.E. (1979) The learning curve: historical review and comprehensive survey, *Decision Sciences* 10, 302-328.

Zimmerman, M.B. (1982) Learning effects and the commercialization of new energy technologies: the case of nuclear power, *Bell J. Econ* 13 (2), 297-310.