ABSTRACT

Title of Document: THE SPATIAL CONFIGURATION OF AMERICAN INEQUALITY: WEALTH AND INCOME CONCENTRATION THROUGH US HISTORY

Scott Albrecht, Doctor of Philosoph, 2014

Directed By: Roberto Patricio Korzeniewicz, PhD
Department of Sociology

Drawing on a variety of data sources—national surveys and censuses, probate and tax records, wage series and rich lists—I identify five period or regimes in US history with distinct wealth and income distributions. I argue that this periodization of inequality in the United States is a product of Arrighi’s systemic cycles of accumulation. Each cycle of accumulation is associated with a spatial configuration, a global pattern of interdependent technologies, infrastructure, institutions, networks and social relations, and ideologies, that structures the distribution and flow of wealth. Interdependence in the components of the spatial configuration means that there are periods of relative stability delineated by moments of cascading change when space is reconfigured; new patterns of wealth and income concentration emerge as a result. The principal contribution of this approach is to further our understanding
of the impact of global processes on within-country wealth and income concentration; we cannot isolate domestic market institutions and technological change from global political and economic competition.
THE SPATIAL CONFIGURATION OF AMERICAN INEQUALITY: WEALTH AND INCOME CONCENTRATION THROUGH US HISTORY

By

Scott Albrecht

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
2014

Advisory Committee:
Professor Roberto Patricio Korzeniewicz, Chair
Dr. Reeve Vanneman
Dr. Phil Cohen
Dr. Meyer Kestnbaum
Dr. Johanna Birnir, Dean’s Representative
# Table of Contents

List of Tables ............................................................................................................... iv
List of Figures ............................................................................................................... v
Introduction ................................................................................................................... 1
  Project Summary ...................................................................................................... 13

Chapter 1: A Model Of Within-Country Inequality ................................................... 19
  The Current State of US Income Inequality in Historical Context ......................... 20
  Smith and Schumpeter on Inequality ..................................................................... 24
  Institutions and Inequality ...................................................................................... 27
  Relational Inequality .............................................................................................. 31
  Modeling Spatial Inequality ................................................................................... 35
  Confined Space and Unfree Labor ......................................................................... 39
  Conceptualizing Space ........................................................................................... 42
  Innovation and Diffusion in the World-economy .................................................. 47
  From the Global to the Local ................................................................................ 59
  Systemic Causality and the Counterfactual Fallacy ............................................... 61
  Summary ................................................................................................................ 67

Chapter 2: The Colonial Inequality Regime(s) ........................................................... 69
  The Mythos of Colonial Inequality ........................................................................ 72
  A Note on Data ....................................................................................................... 77
  Colonial Crisis, Adaptation and Institutional Path Dependency ............................ 81
  The American Frontier ........................................................................................... 90
  The Global Context ................................................................................................ 93
  Interpreting Colonial Inequality ........................................................................... 102
  Conclusion ............................................................................................................ 105

Chapter 3. Commercial and Corporate Inequality Regimes ..................................... 107
  The 19th Century Transformation and Processes of Convergence ....................... 109
  Data ...................................................................................................................... 113
  The Commercial Inequality Regime .................................................................... 120
  The End of the Dutch Cycle of Accumulation and the Exploits of War .............. 121
  Commercial Wealth beyond the Dubiously Legal ............................................... 128
  The End of the Commercial Regime .................................................................... 139
  The Corporate Inequality Regime ........................................................................ 145
  The Bigger the Better ........................................................................................... 154
  America’s New Wealthy ....................................................................................... 156
  Conclusion ............................................................................................................ 157

Chapter 4. The Keynesian Inequality Regime ......................................................... 160
  The Data ............................................................................................................... 166
  Contradiction and Crisis ....................................................................................... 173
  The Intellectual Response .................................................................................... 186
List of Tables

Table 3.1 Counties, Value of Manufacturing > $100 million
(in $100 millions).................................................................148
List of Figures

Figure 0.1 Static Model of the Spatial Configuration................................5
Figure 0.2 Life Cycle of the Spatial Configuration of the World-Economy......6
Figure 0.3 Systemic Cycles of Accumulation, US Inequality Regimes,
   and Stylized Wealth and Income Concentration...............................16
Figure 1.1a. Income Share of Top 1% (1980-2008)...................................21
Figure 1.1b. Income Share of Top 1% (1913-2008)...................................22
Figure 1.2 Income Shares of Top Earners, Select Years............................23
Figure 1.3 The Lorentz Curve, Gini Coefficient, and Minimum Gini..........65
Figure 2.1. Lorentz Curve of Total Physical Wealth by Region, 1774............87
Figure 3.1 Urban Skilled to Unskilled Wage Ratio, 1817 to 1939...............118
Figure 3.2 Rates of Travel from New York City....................................136
Figure 3.3 Railroad Density..............................................................138
Figure 3.4 Domestic Manufacturing Output as a Function of the Tariff Rate....144
Figure 3.5 Number of Counties among Top 100 by Manufacturing
   Output, by Region.................................................................148
Figure 3.6 Warren and Pearson Index of Prices, 1850-1890 (1910-14=100)....149
Figure 3.7 Distribution of America’s 100 Wealthiest by Year of Birth.........157
Figure 4.1 Capital Income Ratio, United States, 1770-2010........................163
Figure 4.2 Capital Income Ratio and Top Income Share, France 1910-2010.....164
Figure 4.3 Capital Income Ratio and Top Income Share, France 1910-2010.....167
Figure 4.4 Income Share of the Top 1% (Excluding Capital Gains).............168
Figure 4.5 Income Share of the Top 1% (Excluding Capital Gains).............170
Figure 4.6 Income Share of Top 1%, France and the US..........................171
Figure 4.7 Observed And Estimated Family Income Inequality
   Gini, 1917-2008........................................................................173
Figure 4.8 Design Displacement of Capital Ships, 1905-1922.....................182
Figure 4.9 Labor Force Growth and Skilled Wage Premium......................198
Figure 4.10 Effective Tax Rate and Tax Progressivity..............................206
Figure 4.11 Indexed Minimum Wage, Constant Dollars (1938=1)..............210
Figure 5.1 Average Income by Quintile and % Change, 1973-2010............222
Figure 5.2 Employment in Manufacturing............................................225
Figure 5.3 Average Hourly Wage in Manufacturing, 1980 US$..................227
Figure 5.4 FIRE Share of Corporate Profits........................................235
Figure 5.5 Ratio of GDP of Western Europe to the United States..............240
Figure 5.6 Capacity Utilization, Total Industry, 1967 To 2011..................250
Figure 5.7 Dow Jones Industrial Average, 1980 US$.............................251
Figure 5.8 Inflation and Unemployment...............................................251
Figure 5.9 National GDP Growth (Annual %) By Income Group...............263
Figure 5.10 World Trade, 2005=100.................................................264
Figure 5.11 House Prices and Dow Jones Industrial Average, 2000 to 2013...269
Figure 5.12 Average Wage, 2000 and 2012 (2000 US$)..........................274
Figure 5.13 Global Distribution of Incomes, 1980 and 2011......................275
Figure 5.14 Global Securitization (ABS, MBS, CDO, CLO)......................278
Introduction

On the eve of the Great Depression, incomes in the United States were highly concentrated; the richest 1% of Americans earned incomes 26 times greater, on average, than the other 99%. Eighty years later, and again on the eve of financial meltdown (2007), the richest 1% again enjoyed incomes 26 times greater than the rest. In the interim, the share of income going to the top percentile fell by 60% (Piketty and Saez 2003; author’s calculations). It is the goal of this dissertation to improve our understanding of wealth and income concentration through US history: income compression and re-concentration in the 20th century, the extraordinary wealth of a group of men born in the 1840s, the first American dollar millionaires in the decade before the War of 1812, and the concentration of wealth in Charleston, South Carolina before the American Revolution.

It is the norm today to focus on a single historical moment in the study of within-country economic inequality, but a more historical perspective is not without precedent. Sixty years ago Simon Kuznets (1955: 1) addressed “the character and causes of long term changes in the personal distribution of income” in his presidential address to the American Economic Association. A quarter century later, Jeffery Williamson and Peter Lindert (1980) produced a detailed macroeconomic history of American inequality. Income inequality in the United States from the end of World War II to the Reagan Administration was historically low. Robber baron-style inequality was a condition of the past for Kuznets, Williamson and Lindert.

In the face of rising income inequality since the 1970s, researchers adopted a more immediate focus. High quality census and survey data covering the second half
of the 20th century in the United States also encouraged researchers to focus on wage
and income trends across those distributions over the last half century. That focus is
again shifting. As income concentration in the United States approaches historical
highs, there is a new appreciation for a more historical approach and a new emphasis
on top income shares. This trend is highlighted in a May 2014 special issue of
Science. The issue granted particular attention to Piketty’s (2014) Capital in the
Twenty-First Century, a historical study of top incomes. In that issue, David Autor
qualified his discussion of wage inequality in the United States by noting in the title
that it applied to the “other 99 percent” and devoted significant time and attention to
justifying research on inequalities towards the middle of the distribution.

A more historical approach to within-country inequality offers us a different
perspective on cause and effect. Technological change is the most popular
explanation in the economics literature for rising wage and income inequality in the
United States over the last four decades; other popular suspects are globalization,
falling union membership and the rise of the service or knowledge economy, less
progressive income taxes, changing family structure, and a declining minimum wage.
I offer an explanation for rising income inequality over the last four decades and other
long-term changes in the concentration of wealth and income that is unlike any other.
I begin with a common premise, that economic inequality reflects exclusive, unequal
access to wealth generating activities. Under this umbrella we could include racial
discrimination (“irrational” exclusion) or higher returns to skill (“rational” exclusion).
But instead of pursuing these well-worn paths, I point to a more esoteric variable: the
spatial configuration of the world-economy.
A spatial configuration includes economic and political relations, labor market institutions, norms, ideologies, technologies and infrastructure. One could argue that I am merely co-opting most of the more popular explanations for long-term trends in within-country inequality and placing them under a single umbrella. Within the spatial configuration, though, these variables interact, are interrelated, in meaningful ways, and we can theorize change across the spatial configuration in ways that we cannot if we approach the various processes as unique, independent explanatory variables. In other words, we can treat (more) change endogenously. Similarly, because the spatial configuration spans the world-economy, this approach can progress our understanding of the relationship between global processes and within-country inequality.

Each configuration of the modern world-economy has two key characteristics (for our purposes): first, it structures access to wealth generating activities globally. Second, it matures with time, grows old, eventually dies (kicking and screaming), and is replaced by a new configuration (Arrighi 1994/2010). As a result, we see a periodization of wealth and income concentration (I call these inequality regimes) that correlates with the duration and transition between spatial configurations. In chapters 2 through 5, I track the structure and life cycle of the spatial configurations of the modern world-economy to explain long-term trends in wealth and income concentration in the United States.

Figure 0.1 highlights some key processes (in cross-section) of the theoretical model I develop in chapter 1. At the intersection of space and institutions, the discontinuity between economic and political institutions in space transforms
economic competition into political competition and institutional change. For example, war is often motivated by groups seeking to increase their access to resources in other spaces, and domestic market institutions are often reformed politically in reaction to evolving global economic competition. At the intersection of institutions and technology, institutions influence technological change by structuring innovation and diffusion; the most obvious examples are patent and copyright law and private ownership of the means of production. At the intersection of technology and space, sites of wealth generation are determined by the co-location in space of technologies, capital (human and physical), and natural resources. Wealth and income concentration, then, is determined by access to wealth generating activities (situated in space); access to these activities is structured by their distribution in geographical space, infrastructure, and social networks, transportation technologies, and institutions (in other words, the components of the spatial configuration).
The diagram above emphasizes processes of change, but it is actually a static model of wealth and income concentration. The various components of the spatial configuration reinforce one another; there is a tendency to reinvest in existing infrastructure, technologies, institutions, and structures of power. The diagram below (Figure 0.2) highlights the key processes of change in the spatial configuration, and thus change in wealth and income concentration. First, capitalists invest in existing technologies, markets, infrastructure, etc. to expand production along the existing spatial configuration; this is the phase of material expansion. In time, accumulation of capital drives down profits, and capitalists move their capital in search of new growth opportunities; this is the phase of financial expansion. Financial expansion allows for development outside of the existing spatial configuration (i.e., new regimes of
accumulation), and contradictions between the old and the new creates crisis, conflict and instability. For example, as I discuss in chapter 4, war and the Great Depression in the first half of the twentieth century were, in part, a product of a spatial configuration built around British hegemony that was unable to accommodate Germany and the United States as emerging superpowers. Crisis and conflict ultimately destroy the foundations of the dominant spatial configuration (e.g., weaken superpowers, destroy infrastructure, create demand for new technologies, delegitimize dominant ideologies, etc.) and create space for a reconfiguration. While I develop this model to suit my own purposes, I borrow heavily from Arrighi (1994/2010) and use his periodization of the cycles of accumulation (see Figure 0.3).

Figure 0.2 Life Cycle of the Spatial Configuration of the World-Economy

- **Spatial reconfiguration** creates space for the new regime of accumulation
- **Material Expansion**: Markets (as currently organized) are saturated -> Signaling crisis -> Capital goes in search of new investment opportunities
- **Financial Expansion**: Search for new markets allows alternative regimes of accumulation to develop -> Contradictions in the spatial configuration
- **Terminal crisis (inter-state chaos)** -> Spatial reconfiguration
When we bring these two models together, the static model relating the spatial configuration to wealth and income concentration and the dynamic model of spatial configuration and reconfiguration, the net result is a thesis of inequality regimes, within which wealth and income concentrate in particular, predictable sites, delineated by moments of crisis and transformation.

What truly differentiates my project from other studies of within-country concentration is the more complete use of space and time. Consistent with a world-systems perspective, I define the unit of analysis to include all relevant processes. Most studies of within-country inequality focus on processes within the country—for example, domestic labor market institutions. Even when globalization is included as a variable researchers often stop at the border: e.g., immigration, imports/exports. But capital and labor in the United States are situated within a transnational division of labor, so the appropriate unit of analysis is not the domestic economy, but the world-economy. A key point here is the “discontinuity between economic and political institutions” (Wallerstein 1979: 35); in other words, political borders and economic markets, and the rules by which they operate, do not perfectly overlap. As a result, ‘economic’ competition across a transnational division of labor is also political, institutional, even ideological, and it can induce change that is economic, political, institutional, and ideological.

In addition to a bigger space for analysis, I make better use of that space. Space is not just a container through which objects move; it is variegated (space as place), it has distance (space as scale), and interrelationships (space as network). The creation and flow of wealth globally are structured by distance (e.g., transportation costs),
physical and institutional environments, and exchange networks and hierarchies. Demarcating space and managing movement over space are the most effective strategies for controlling access to resources today.

A spatial configuration in cross-section it is not enough to understand changes in the long-term concentration of wealth and income, but few studies analyze a sufficiently long period of time (the temporal unit of analysis) to recognize transitions between spatial configurations. In other words, these analyses again fail to include all relevant processes to understanding inequality change. My analysis covers four centuries of US (and colonial) history, through which I identify four transitions in the character of wealth and income concentration and five distinct inequality regimes (see Figure 0.3).

By defining the unit of analysis (spatially and temporally) to include the relevant processes for understanding income and wealth concentration and change, I show that change is systemic. The maturation and death of a spatial configuration are the product of processes internal to the spatial configuration. This is a major deviation from a literature that emphasizes competing bivariate relationships (i.e., the impact of globalization, technological change, taxes, the minimum wage, or unions on income inequality). Even within the world-systems community, it is common for researchers to regress a series of independent variables on the Gini coefficient across a set of purportedly independent and analytically comparable countries. I offer a more complete world-systems perspective. Instead of measuring the relative contribution of different independent variables, I emphasize the interrelationship of processes that drive wealth and income concentration. Instead of comparing experiences across a
series of countries, I locate the United States within a universe of interdependent political entities.

That said, this project is complementary, not contradictory, to the establishmentarian approach to within-country inequality. A study with a narrow unit of analysis in space and time can perform the necessary function of identifying the proximate correlates of inequality change, but it will treat change in the independent variables as the product of exogenous events – it has limited scope for understanding the processes that drive change. For example, research in skill-biased technical change often evades the issue of why change is skill-biased. By extending the unit of analysis in space and time I model these changes endogenously.

Of course, there is a cost to extending the unit of analysis in space and time. By way of a social Heisenberg Uncertainty Principle, a broader unit of analysis in time and space allows us a greater sense of momentum in change, but less precision in cross-section. This limitation is exacerbated because as we move back in time we are forced to draw on less reliable data. I am forced at times to ignore nuance and construct a narrative with broad brushstrokes. That said, it is my goal to offer a framework for reinterpreting trends over time, and I believe this project satisfies that goal.

This project hinges on the successful integration of three literatures that rarely interact, a theoretical Frankenstein’s monster. The first is the historical (period-specific) study of economic inequality in the United States. This research is generally empirically rich and theoretically specific, noting particular relationships in time but slow to generalize these results. This is the meat of the project. The skeleton, the
macro operations of the world system, is built from the integration of world-systems theory, particularly Arrighi’s (1994/2010) systemic cycles of accumulation, and trade/development economics. Finally, I use US economic history as the sinew to bind the meat to the bones.

This project represents a major deviation from most research on within-country income inequality, but I hope it is at the forefront of a shift in our understanding of that subject. I do not identify a likely causal mechanism (proximate determinant) and attempt to measure its contribution to inequality change relative to other, popular causal mechanisms. To identify proximate determinants is a worthwhile goal, but my goal is to place these proximate determinants in context. I consider the major transitions in the level of wealth and income inequality in the United States to be over-explained phenomenon in terms of the proximate determinants; it is the coincidence of so many variables that drive inequality in a coordinated direction that now warrants our attention. In other words, I do not believe that inequality change in the long-term can be calculated as the sum of a series of independent variables; the relevant variables are not independent.

I hope this project is at the forefront of a shift that emphasizes the dynamics of historical capitalism to our understanding of within-country economic inequality, a shift away from the narrow focus on bivariate relationships. The best evidence of this shift is the outrageous success of Piketty’s (2014) *Capital in the Twenty-First Century*. Instead of focusing on bivariate mechanisms, Piketty applies a logic of historical capitalism to explain long-term inequality trends. But Piketty approaches the situation from the perspective of the nation-state, and it thus limited to domestic
processes. I instead offer a discussion of *global* historical capitalism and its relationship to long-term trends in within-country inequality.

The first substantive contribution of this project is a reinterpretation of long-term trends in within-country inequality that places global processes at the center. Global structures are large and rigid; in shorter moments of time they do not produce enough variation to establish causation. This problem is exacerbated because our discussion of within-country inequality is motivated by the available data. Often that data reflects shorter periods of time, and changes in the way we collect data are correlated with bigger shifts in the global structures. For example, data collection changed dramatically during the Great Depression and after World War II. These data are valuable for understanding inequality trends since, but are less useful for tracking inequality trends as global structures were collapsing.

These challenges have led investigators to focus on domestic institutions and economic trends. This is a mistake. I highlight in this project that long-term inequality trends in the United States are better understood as a part of the *global* historical narrative. I highlight global economic forces, but also the diffusion of ideologies and institutional models internationally and the international character of war and economic crisis that drive institutional and economic change.

The second substantive contribution of this project is to offer an approach that moves beyond bivariate relationships. Unfortunately, this effort is too often interpreted as a weakness of the project, that it cannot easily be distilled into a model that is conducive to regression analysis. Instead, given the complexity of the model, I am forced to depend on historical narrative. But the literature is filled with partial
models of within-country inequality, applicable only to a specific place at a specific
time and with limited explanatory power. It would hardly be a valuable contribution
to add another approach of this ilk. Instead, I offer a model that reflects the
complexity of human social and economic interaction across a dynamic world-
economy. While it might initially appear more difficult to standardize the approach so
that it can be applied in other places and other times, my framework is, in fact,
ininitely more pliable than the standard approach that is obsessed with a single
bivariate relationship.

In short, within-country inequality change is a product of complex processes that
play out across a dynamic world-economy. This project offers the first model of long-
term trends in within-country inequality that respects this complexity.

To review, this is a study of long-term trends in wealth and income concentration
in the United States. It is unique in that I adopt a broader unit of analysis in space and
time and, from this, present a systemic model of inequality change. A motivating goal
of this project is to identify mechanisms linking within-country inequality to global
processes. To that end, I highlight the spatial configuration of the world-economy, a
self-reinforcing pattern of economic and political relations, institutions, ideologies,
technologies and infrastructure which structures the distribution of and access to
wealth generating activities in space. Domestic political and market institutions are
components of this spatial configuration, buffeted by world economic forces, so
change in the former must be understood in the context of the latter. Finally, the
spatial configuration has a natural rhythm, and changes in within-country wealth and
income concentration (the inequality regime) are synchronized with that rhythm.
Project Summary

The narrative I present on long-term changes in wealth and income concentration in the United States is, admittedly, a complex one. It is, therefore, appropriate to offer a brief roadmap before delving into the meat of the argument. I divided the project into five chapters. The first develops the theoretical model of spatial configurations and their relationship to wealth and income concentration within countries. The next four chapters deal with specific historical periods, applying the theoretical model to empirical evidence. Through the rest of this section, I offer a brief summary of the project as a whole.

I propose that large economic inequalities are the manifestation of unequal access to innovative practices. Innovation is the creation or utilization of a better, more effective, more profitable resource, technology, product or market (Schumpeter 1942/1950). Early adopters enjoy monopolistic or oligopolistic rents, returns above what would be realized in a perfectly competitive market (Sorenson 1996; 2000). In time, profitable innovations are copied (diffused), competition increases, and profits fall. A number of institutions exist to limit diffusion and protect extraordinary profits, for example, private ownership of physical and intellectual property, discrimination based on gender, race, and educational credentials, and professional associations. But the largest inequalities are spatial, and they trace the contours of space (distance, geography, infrastructure, political boundaries).

Innovation and diffusion are not random processes. First, for an innovation to be broadly adopted (diffused) it must exploit relatively abundant productive inputs and/or produce for a high-demand market (Acemoglu 1998; 2002). These vary over
space and time, so innovation and diffusion are patterned by the spatial and historical contours of the world-economy. Second, diffusion often requires the adoption of new capital and skills, exchange relationships, and the adaption of technologies and techniques to local demands – in other words, diffusion often requires a significant investment of time and resources. Third, those profiting from a new innovation can impede its diffusion. Strategies for doing so include institutional regulations (e.g., patents), the monopolization of inputs, exclusive trade relationships, and preventing transmission through political boundaries.

Self-reinforcing relationships emerge in the spatial distribution of innovation, wealth, production, and power, and patterns of diffusion and exchange are calcified by investment in physical infrastructure (Harvey’s spatial fix, 1982; 2003), institutions and networks. These spatial configurations consistently funnel excess profits to some sites and some people and away from others. By identifying sites of wealth creation globally and tracing paths of wealth diffusion, we can predict where, to whom, and to what extent wealth will accumulate. I refer to these patterns of advantage and disadvantage as inequality regimes.

A spatial configuration represents the scaffolding on which the world-economy expands, but the same structures that supported growth in the past constrain it in the future. Productive capacity increases, driving up prices for inputs and saturating markets with goods and services. Capital seeks new spaces for growth, but these developments might not be consistent with the dominant spatial configuration – existing infrastructures and technologies, international and national institutions and political and economic relations. In time, these contradictions within the spatial
The components of the configuration are interdependent, change cascades through the system (Arrighi 1994/2010: 231). Global reconfigurations can be associated with a restructuring of the international hierarchy, the rise of new economic and political ideologies, new trade patterns, institutions and associations, and investment in new means of production—the typical variables used to explain rising and falling inequality. Features of the old regime will survive—e.g., institutional expectations, relations of power, economic organization and investment. But in other respects the new regime will represent an antithesis of the old regime: an intentional rejection of old philosophies, old institutions, old hierarchies. Consequently, the level of inequality will tend to fluctuate around an equilibrium over time (Korzeniewicz and Moran 2009).

Drawing from such a perspective, I divide the history of inequality in the United States into five periods or inequality regimes (see Figure 0.3). The “colonial” regime, the first of the five, lasted from initial colonization to 1790. Wealth concentration was relatively low, but categorical inequalities (e.g., slavery, indentured servitude, coverture) were severe. High land to labor ratios encouraged an economic and political democratization for white men, and the functional distance from Europe made it difficult to tap that key site of wealth accumulation. There was significant variation across colonies, though; southern colonies that were better positioned to export to the West Indies and Europe recorded higher levels of wealth inequality across the non-slave population and larger slave populations.

Two events in 1789 set the stage for a new, “commercial” inequality regime. The US Constitution went into effect in May, and French rioters stormed the Bastille in
July. As war spread in Europe, American merchants exploited gaps in the frayed remnants of the mercantilist organization of the world-economy to enrich themselves and a few producers of uniquely American goods. In other words, the barriers that defined the colonial inequality regime began to break down. As profits fell, wealthy American merchants invested in a range of new businesses, including manufacturing built on borrowed British technology, and entrepreneurial merchants were able to push into new, western markets within the United States. The Civil War marked the end of the commercial inequality regime.

Figure 0.3 Systemic Cycles of Accumulation, US Inequality Regimes, and Stylized Wealth and Income Concentration

Notes: Dotted lines represent periods in which I emphasize wealth inequality, dashed line income inequality. CM' and MC are periods of financial and material expansion, respectively. The timing of systemic cycles of accumulation are taken from Arrighi 1994/2010.

Through the second half of the 19th century, new transportation technologies, the British emphasis on free exchange, and the growth of US domestic demand eroded
arbitrage profits in cross-Atlantic trade. The combination of intensive (more productive) and extensive (more inputs, e.g., workers) growth of the US economy has few precedents in world history, and those in a position to take an early lead in key industries are among the wealthiest men in the history of the world. More precisely, new organizational forms (i.e., the modern corporation) exploited the more rapid movement of information, goods and capital in order to control markets, secure profits, and enrich their owners against "cut-throat" competition. I refer to this as the "corporate" inequality regime.

In the first quarter of the 20th century, the Atlantic economy was torn apart by competition for scarce resources and access to new markets (both to be found, theoretically, in colonization), leading ultimately to armed competition and financial crisis. New national and global economic institutions reflected emergent economic and political ideologies that were reactions against the turbulence associated with the gold standard and British liberalism. Legislation turned against the monopolistic corporations of the last regime. Countries experimented with a variety of models, generally oriented towards managing market forces, but victory by the Allies and the rise of the United States as the new global hegemon pushed change in a direction suitable to a US-centric world-economy. Domestic markets and workers were protected incidentally by the centralization of global power, production and capital within the United States, and by Washington’s inordinate control over each. This Keynesian inequality regime lasted from the Great Depression to the 1970s.

Domestic and international institutional arrangements that defined the Keynesian inequality regime hinged on US hegemony. That hegemony, built on economic,
financial and military power, was challenged on all three fronts by the 1970s. Western European and Japanese firms began to compete with American firms, both internationally and in the United States itself; financial power shifted to New York, and the United States was not able to stay on the gold standard; the United States became embroiled in a war that it could not win in Vietnam. Rising inequality in the United States since reflects the transition to a new inequality regime, a period of financial expansion and a new economic logic of global competition. Capital from the core seeks out growth opportunities abroad, and is oriented towards reducing costs, not expanding production, when it stays home. Key financial and labor market institutions accommodated this new focus. Those in charge of managing financial flows, those in charge of managing production costs in the United States, and those in a position to profit from the international flow of goods and capital have fared well; others have not. This is the financial inequality regime.

I am faced with three challenges in the defense of my thesis. The first is to define and empirically identify distinct inequality regimes through US history, the long-term changes in wealth and income concentration. The second is to correlate transitions between these regimes with changes in the spatial configuration of the world-economy, the timing of which I borrow from Arrighi’s (1994/2010) systemic cycles of accumulation. The third is to identify the mechanisms that relate each spatial configuration with the corresponding inequality regime. In the Epilogue, I discuss the fit between model and historical narrative.
Chapter 1: A Model OF Within-Country Inequality

A study of long-term trends in wealth and income concentration is ultimately motivated by a need to understand trends today. I begin this chapter with a brief look at the current inequality trend in the United States, a look at its historical relevance, and juxtapose my approach against the perspective offered by other studies of US inequality over historical time. I follow that with a discussion of the relationship between innovation, diffusion, institutions and inequality. Innovation provides a means to secure extraordinary profits, but diffusion introduces competition, reducing profits. Institutions can influence the rate of innovation and diffusion, and thus intervene in the relationship between innovation and inequality. Institutions can make inequalities durable by naturalizing categorical distinctions. These distinctions are easiest to maintain when people in different categories do not interact (i.e., do not compete), which is most easily accomplished if they inhabit different geographical spaces. Consequently, the largest categorical inequalities globally are spatial. In fact, ‘unequal exchange’ (e.g., the exploitation of labor) always involves a manipulation of space.

Space can be manipulated because it has structure. Innovation and diffusion are constrained by physical and constructed geographies, institutional environments, the distribution of wealth, and distance. The configuration of space is durable. Resources (economic, social, political) yield greater returns when they are invested to exploit, not replace, the existing structures, and those who control the most resources are particularly invested in supporting existing structures. Within a spatial configuration, profits are consistently funneled to some and away from others. In time, though,
opportunities for growth within the spatial configuration are exhausted. Tighter competition for falling profits forces capitalists to seek out new markets, which allows for developments that contradict the existing spatial configuration.

Contradiction leads to crisis, conflict and, eventually, a reconfiguration. I use this model to link Arrighi’s (1994/2010) systemic cycles of accumulation with periodized changes in income and wealth concentration in the United States.

The Current State of US Income Inequality in Historical Context

There is nothing wrong with people succeeding and making money. But there is something wrong when the opportunity for all Americans to get ahead, to enter the middle class, and to create a better life for their children becomes more and more elusive. That is what has been happening: The ladder into the middle class and beyond has become harder and harder to climb. The American dream has slowly slipped beyond the grasp of millions (OMB 2009: 9).

In its spending request for fiscal year 2010, *A New Era of Responsibility*, the Obama administration documented growing inequality by charting top income share estimates from economists Thomas Piketty and Emmanuel Saez (2003). The President’s chart shows the share of total income going to the top 1% of earners doubling since 1980 (see Figure 1.1a). Of this chart, Daniel Henninger (2009) of the Wall Street Journal said that “Messrs. Piketty and Saez have produced the most politically potent squiggle along an axis since Arthur Laffer.”
The real value of the Piketty and Saez income share estimates is not what they tell us about the distribution of income in the United States since 1980. Several surveys offer reliable estimates and broader coverage of the distribution of income for this period.

Piketty and Saez draw on tax return data to provide consistent top income share estimates back to 1913. From this perspective, we see that income inequality in the United States today, though high, is not unprecedented (see Figure 1.1b). This is recognized widely in the literature. For example, Goldin and Margo (1992: 3) find “remarkable similarities . . . between the narrowing wage structure of the 1940s and the widening wage structure after 1970.” Harrison and Bluestone (1990: viii) prefaced their work on “the Great U-Turn” by noting “the decade of the 1980s bore a credible resemblance to the 1920s. Underneath a small sliver of society with an almost

---

1 The federal income tax was made constitutional with the ratification of the Sixteenth Amendment on Feb 3, 1913.
unfathomable wealth lay a precarious economy and an increasingly uneasy struggling middle class.”

In Figure 1.2 we can see just how remarkable these similarities are. According to the U.S. Census Bureau, income inequality in the United States hit a high point in 2006 when the Gini coefficient reached .470 (Census 2011a: Table A-2). The income shares accruing the top 10%, 5%, 1%, .5%, .1% and .01% of earners in 1928, when income concentration last peaked, are almost identical to those in 2006 despite a large drop in top income shares in the intervening years.

Looking back further, Williamson and Lindert (1980:5) wrote that “inequality among free Americans before the Revolution was not too different from that which we experience today [1980]. Yet, inequality was hardly stable for the long period in between.” Taken together, these observations suggest a cycle of rising and falling inequality: rising from the colonial period to the Great Depression, falling back to
colonial-era levels through the post-war era, and then retracing those steps since the 1980s.

![Figure 1.2 Income Shares of Top Earners, Select Years](image)

Source: Piketty and Saez (2003), updated online 2008, Table A3

Despite the empirical similarities in wealth and income inequality in different historical periods, the literature favors period-specific explanations. For example, Goldin and Margo (1992) argued that the empirical similarities between the 1940s and 1970s were largely superficial, and the politico-economic forces driving those trends were fundamentally different; “For the 1940s . . . unique historical events [and] institutional changes . . . might explain much of the narrowing” (1992: 4, italics added). Even Simon Kuznets’ (1955) explanation for long-term changes in the distribution of income emphasizes economic modernization, a one-off process.

It is my position that the changes in the proximate determinants of economic inequality—e.g., new technologies and institutions—are primarily the result of actors reacting to the evolving structure of the world-economy, and deploying a range of strategies to secure and extend their control over resources (and people). In other
words, historical events and institutional changes are unique in time, in that the strategies they adopt are influenced by local conditions. But the goals are the same. I use the rest of this chapter to discuss how strategies to control access to resources become manifest in the spatial configuration.

*Smith and Schumpeter on Inequality*

Economic inequality reflects the unequal distribution of profits. When few people are engaged in producing a good or service that is in high demand, those few can enjoy large profits. People tend to adopt those strategies—activities, techniques, locations—that generate greater profits. As they do so, existing producers are forced to compete on price to purchase inputs and attract customers, and profits fall. Eventually, the profit per person from that activity above that generated from other activities approaches zero . . . in theory, at least.

According to Adam Smith,

> The whole of the advantages and disadvantages of the different employments of labour and stock must, in the same neighbourhood, be either perfectly equal or continually tending to equality. If in the same neighbourhood, there was any employment evidently either more or less advantageous than the rest, so many people would crowd into it in the one case, and so many would desert it in the other, that its advantages would soon return to the level of other employments. (Smith 1776/2004: 85)

Smith added one condition:

> This at least would be the case in a society where things were left to follow their natural course, where there was perfect liberty, and where every man was perfectly free both to choose what occupation he thought proper, and to change it as often as he thought proper. (Smith 1776/2004: 85)

Unfortunately, as Adam Smith highlighted, this “perfect liberty” does not exist. There is a cost to acquiring the necessary tools, skills, networks, etc., to engage in a new economic activity. In some cases, market entry is entirely impossible. These
constraints allow those already enjoying extraordinary profits to continue to do so, creating durable inequalities.

Durable inequalities, though, may be less durable than their beneficiaries may hope. Even if the diffusion of profitable activities is perfectly constrained, excess profits are vulnerable to innovation—the creation of new, better, more profitable techniques in their place. Along these lines, Joseph Schumpeter argues that the traditional economic focus on price competition and advantages at the margins is myopic.

In capitalist reality as distinguished from its textbook picture, it is not that kind of competition\(^2\) which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization (the largest-scale unit of control for instance)—competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives. This kind of competition is as much more effective than the other as a bombardment is in comparison with forcing a door, and so much more important that it becomes a matter of comparative indifference whether competition in the ordinary sense functions more or less promptly; the powerful lever that in the long run expands output and brings down prices is in any case made of other stuff. (Schumpeter 1942/1950: 84-85)

While potentially disruptive to old inequalities, creative destruction entails the constant generation of new winners and losers, and can itself be the source of a “constant drive towards inequality” (Korzeniewicz and Moran 2005). Those who are on the leading edge of a successful innovation reap “spectacular rewards”, while others see their products / technologies / markets / organizations replaced and destroyed as they become less profitable

\(^2\) “competition within a rigid pattern of invariant conditions, methods of production and forms of industrial organization in particular, that practically monopolizes attention” (84).
The most popular explanations of wage inequality today point to differential returns to skill and skill-biased technological change (SBTC; Autor, Katz and Kearney 2008). Some individuals are richer, the argument goes, because they have abilities that make them more efficient and their work more profitable. New technologies and the death of old technologies can stretch that gap if the new technologies increase the demand for skilled labor. This model can then be combined with changes in household composition to explain trends in income inequality. Rarely explicated in these models are 1) the reasons technological change might be skill-biased and 2) the factors that lead to imperfect competition—that is, what prevents innovation from benefitting lower-income workers and what prevents lower-income workers from adopting higher-income strategies?

One aspect that prevents the immediate redistribution of workers from less profitable to more profitable economic activities is that with each new technology there is a learning phase, a period of time during which more workers (typically designated less-skilled) master the skills necessary to manage the new technology.

More often than not, the effects of technological change are unequalizing at first but not in the long run. For example, after the learning phase is over and workers become fully efficient in using the new technology, firms substitute relatively expensive skilled labor with more economical unskilled labor. . . . Even if the demand for skilled labor did not decline, wage inequality would fall if the supply of skilled workers caught up with demand. (López-Calva and Lustig 2010: 13-5).

Goldin and Katz (2010) describe a race between technological change and education (or innovation and diffusion) in the United States during the 20th century—inequality grows when technological change outpaces the reproduction of skilled labor.
But these approaches are, again, insufficient. First, having been derived from the human capital model of inequality they ignore the complexity of diffusion. Access to new economic activities may require a new set of skills, but it may also require access to new markets, capital and technologies. In other words, we lose sight of inequalities between workers and capitalists by focusing instead on differentiating workers. The income gap between workers and capitalists (and managers) far exceeds income gaps between skilled and less-skilled workers. This stylized fact is highlighted by rising income inequality in the United States in recent decades: a growing gap between skilled and less-skilled workers in the middle of the income distribution is dwarfed by the gap between the richest 1% of adults and everyone else (see chapter 5).

Second, they ignore the multiple dimensions of diffusion. Innovation and diffusion are social processes that are embedded in time and space (Arrighi and Drangel 1986: 20; Arrighi, Silver and Brewster 2003: 17). They respond to the spatial distribution of supply and demand as well as the social relations that constrain the movement of capital, goods, ideas, and people through time and across space. Therefore, if one is able to manipulate social institutions and relations that shape innovation and diffusion— influence the pace and location of innovation, gain access to sites of innovation, and constrain or channel its diffusion—that person or group can consistently earn profits beyond what is possible through price competition alone.

*Institutions and Inequality*

Speaking of the disruptive nature of capitalism, Polanyi (1944/2001) notes that “the rate of change is often no less important than the direction of the change itself; but while the latter frequently does not depend upon our volition, it is the rate at
which we allow change to take place which well may depend upon us” (1944/2001: 39). To Polanyi, the role of government and legislation “consists often in altering the rate of change, speeding it up or slowing it down as the case may be” (1944/2001: 39; quoted in Korzeniewicz and Moran 2009: 35). “Historically, the institutional arrangements shaping inequality have combined both sets of strategies, . . . accelerating the introduction of some innovations . . . , while simultaneously *slowing down* and restricting the rate at which [other] innovations . . . are introduced” (Korzeniewicz and Moran 2009: 35).

For this discussion, institutions encompass everything from physical and intellectual property rights (which constrain/protect access to capital, resources and profitable ideas), labor organizations, professional associations and anti-trust legislation (which constrain/protect access to employment and productive activities generally), schooling (access to skills and credentials), trade and migration regulations (access to resources and markets), financial regulations (access to capital), and institutions of political governance (which influence everything else). Polanyi argues that the market system is “an institutional structure itself [that], as we all too easily forget, has been present at no time except our own” (Polanyi 1944/2001: 40).³ In each case, the diffusion of a productive innovation (or access to the capital necessary to exploit that innovation), is sped up or slowed down by institutions—“the rules of the game” (North 1990: 3). They operate by preventing access to a market

³ A classic historical example is the English “market overt”. Only in a market, in a specific location at a specific time, could movable property (chattel) be sold without witnesses.
(goods or labor), preventing access to capital (human or physical), or preventing the intellectual or practical diffusion of the innovation.

That institutions shape economic development and the distribution of its benefits is not a new idea. New institutional economics, which seeks to explain economic growth by focusing on the institutions that mitigate transportation and information costs, redistribute risk, etc., traces its roots back to 1937 (see Coase 1937).\textsuperscript{4} Douglas North linked American economic growth to institutions and institutional change (North 1981; 1990). Recently, Sokoloff and Engerman (2000; Engerman and Sokoloff 1997; 2002) and Acemoglu, Johnson and Robinson (2001; Acemoglu and Robinson 2001) have looked at initial conditions in the New World (factor endowments, population density and geography) to explain the development of “extractive institutions” (with high inequality) in some places and “property rights institutions” (with low inequality) in others. These institutional environments are associated with different long run growth rates, linking inequality within regions to inequality between regions\textsuperscript{5} (see also North 2005; De Ferranti et al. 2003; Ravallion 1997).

Institutional environments that offer narrow access produce a distribution of power and resources in which the beneficiaries have both the incentive and capacity to reinforce that institutional environment (Glaeser, Scheinkman, and Schleifer 2003). Initial conditions favor/allow accumulation of wealth and power at the top, which is

\textsuperscript{4} Going back much further, even Adam Smith was concerned about unequal influence (Rothschild 2002)

\textsuperscript{5} Korzeniewicz and Moran (2009) critique this link by showing that the relationship between growth rates and inequality has evolved over time.
then reproduced by limiting political voice and mass education (in favor of higher education), hijacking systems for [re]distributing land, curtailing the capacity of workers to negotiate contracts collectively and to migrate, and legislating regressive taxes and public spending (Sokoloff and Engerman 2000). These can limit one’s ability to develop skills, to deploy skills, and receive a fair return for skills. These inequalities can influence the distribution of wealth and income within a population and create a group-based system of stratification.

This approach is represented empirically in work by Hernando de Soto (2002). De Soto argues that huge numbers of potential entrepreneurs (in Peru specifically) are excluded from the formal economy by bureaucratic regulations that make gaining formal property rights and licensing a business a long and expensive process. The result is a system of dysfunctional property rights that deters investment by the poor and allows the rich exclusive access to formal markets. The poor are forced to compete in the riskier informal economy.

The institutional environment also plays an important role in mediating risk. Individuals cannot control or predict economic outcomes because 1) they cannot control or predict the economic, social and political decisions of others and 2) they cannot control or predict the consequences of creative destruction (though some are more effective than others). Greater risk, by definition, produces greater inequality; the gap between winners and losers is larger. The relationship between risk and inequality is reinforced because those with fewer resources have less margin for error, so access to risky, and potentially very rewarding, ventures are open disproportionately to those with more capacity to absorb losses (Rogers 1962). Risk is
less consequential when 1) the rules of the game are clearly explicated and executed consistently (there is less uncertainty in the institutional environment itself), 2) the process of creative destruction (whereby innovation makes older technologies, goods, markets, etc. redundant) is constrained to reduce uncertainty in the speed and direction of economic change and 3) risk is pooled across many individuals.

Relational Inequality

In a review of Williamson and Lindert’s (1980) seminal study of historical inequality in the United States, Turner (1982) differentiates their model, which focuses on uneven capitalization and technological change, from more sociological models that focus on the role of power.

For most sociologists, existing inequalities create differences in power that are used to influence economic and political decisions that increase inequalities. Unless this cycle is broken by mitigating conditions, such as democratization, unionization, ideological persuasion, consolidation of power by the disadvantaged, etc., inequality tends to increase. Thus sociologists focus on those conditions that retard or accelerate capacities of various economic and political sectors to consolidate power, control key economic and political decisions, and extract economic surplus. (Turner 1982: 530)

Inequality is relational. This is obviously true in the sense that an individual cannot be unequal unto herself. Wealth and poverty do not make inequality, but their coexistence. But we can also argue that inequality is relational in the sense that “there is a causal connection between the advantage and the disadvantage of two classes” (Sorenson 2000).

We can conceptualize this “causal connection” as adopting three forms. At one extreme is market competition, the rich are rich and the poor are poor because the rich are more efficient producers. Most economists tell us that we should celebrate this “causal connection” between the wealth of some and poverty of others. At the other
extreme, the rich are rich and the poor are poor because the rich exploit the poor. I’d like to highlight a third “causal connection” which combines aspects of the two: the rich are rich and the poor are poor because the “rules of the game” favor some and disfavor others. Institutions naturalize inequalities for those that want to emphasize market competition, while critical thinkers note that the rules of the game are biased.

The privileged adopt strategies to monitor and protect their advantage while minimizing the costs of doing so.

Durable inequality among categories arises because people who control access to value producing resources solve organizational problems by means of categorical distinctions. Inadvertently or otherwise, those people set up systems of social closure, exclusion and control. Multiple parties — not all of them powerful, some of them even victims of exploitation — then acquire stakes in these solutions. (Tilly 1999: 8)

Categories are used to justify the existing inequalities when they were themselves created to justify and protect restricted access to economic opportunities.

If we ignore the relational character of inequality, we are left to focus on the unique characteristics of groups—this approach often reduces to a project of measuring skill. Skill is, generally, an ability that has been acquired by training or experience. It is economically significant because it increases the productivity of labor. But productivity is not synonymous with price or profit. For a skill to produce surplus (i.e., profits that exceed the cost of developing the skill), something must be preventing its diffusion. These barriers to diffusion are not always intentional. For example, a skill may require immense natural capacity (e.g. physical size or intelligence). But in some cases, capacities that are labeled skill are just undiffused labor techniques that have been institutionally appropriated by some group to protect
their advantage (Arrighi 2007: 46). In other words, skill is not the source of inequality but an organizational mechanism used to justify and protect an advantage.

Adam Smith highlighted this point by comparing the work performed in the town and country.

No apprenticeship has ever been thought necessary to qualify for husbandry, the great trade of the country. After what are called the fine arts, and the liberal professions, however, there is perhaps no trade which requires so great a variety of knowledge and experience. The innumerable volumes which have been written upon it in all languages, may satisfy us, that among the wisest and most learned nations, it has never been regarded as a matter very easily understood. (Smith 1776/2006: 107)

But town workers were able to protect their relatively high wages by “restraining that free competition which would most certainly occasion” a “reduction of price” through corporations and apprenticeships (Smith 1776/2006: 107).

A new technique which has not yet diffused (or, if diffused, its practice is in some other way constrained6), is more profitable, and that reward is often interpreted as a return to skill and a prime mover of rising inequality. As in Smith’s example, when a group is able to slow or control diffusion, they can hoard opportunity and retain excess profits in the name of skill. The concept of human capital conceptualizes skill as a property that can be rented for a wage, but we can also consider the privilege to acquire or deploy a skill to be a property.

Historically, those activities in which children could not or were not usefully employed held a natural advantage in slowing diffusion. With Adam Smith’s husbandry, for example, the employment of children not only increased the labor

6 To produce for the market, producers need access to any material inputs, tools, labor (skill and time), and access to the market. Producers can control competition by constraining access to any of these.
supply, tightening competition, but through their employment the required skills were diffused broadly and cheaply. But for the first European migrants to what is now the United States, the knowledge and capacity to produce food was the most valuable skill and one that was too often in short supply.

Over the last century, formal education, especially public education, and the demands of industrial employment revolutionized the dynamics of skill diffusion. In the United States, the high school movement between 1910 and 1940 and the GI Bill in 1944 allowed populations to access training (and credentials) that they could not receive at home and from which they would have otherwise been excluded (Goldin and Margo 1992; Goldin and Katz 2010). One result is that education is increasingly used as a proxy for skill or human capital more generally (historically, occupation and literacy have been the dominant proxies).

In the last two decades of the 20th century, inequality within groups with similar educational experiences (residual inequality) began to increase in the United States (Lemieux 2004). Formal education is becoming relatively less important in managing the diffusion and deployment of skill and is, therefore, a less efficient measure of skill. As the demands on the labor force evolve or groups struggle to control diffusion, institutions that channel diffusion are created, recreated and destroyed, but within the context of the existing institutional environment. Guilds are replaced by professional associations and trade unions, and formal institutions of public education replace home and religious training.

In short, economic inequality reflects the operation of social relations in time and space. These social relations embed themselves in economic institutional
arrangements and categories. These arrangements can be sustained over time, despite the constant turbulence of creative destruction because the benefactors reinvest profits in new profit-generating activities. These relations, though, would be impossible to maintain over time without the role space plays in locating innovation and channeling diffusion.

To review, inequality is the product of unequal access to wealth-generating activities. First, innovation creates new wealth-generating activities (and destroys old ones). Second, the diffusion of that innovation is constrained in some way. All diffusion is constrained in time, but institutions also affect the rate and direction of diffusion. Inequalities, are structured by socially-constructed institutions, but the role of institutions is hidden, and inequalities naturalized, by categorical distinctions. In the next section I introduce the role of space. Space exacerbates each of these processes. Natural, built and institutional environments interact in space to structure innovation and diffusion.

*Modeling Spatial Inequality*

Distance can separate workers from work and producers from markets. Barriers between places—e.g., political borders, linguistic/cultural differences—further exclude workers from capital, more productive technologies and more profitable markets. On the other hand, workers in the right place at the right time benefit from limited competition. Today, the location of workers has more to say about their income than their education, occupation, experience, gender, age, effort and capacity combined (Korzeniewicz and Albrecht 2012).
It is difficult to empirically measure the impact of differential market access across space on inequality. If we assume that the primary principle separating workers from markets is national borders and the physical distance between countries, the impact of space is titanic. The single most important variable influencing an individual’s income in the world today is country of residence. The inequality between nations explains about ⅞ of total global inequality (i.e. the unequal distribution of income across all individuals worldwide; Korzeniewicz and Moran 1997; Milanovic 2005; Firebaugh 2003). These gaps have proven durable over the last century (Milanovic 2005; Bourguignon and Morrisson 2002).

Admittedly, these inequalities reflect more than differences in market access. One alternative explanation for international inequalities are differences in human capital accumulation between countries and regions. But differential market access is, in part, responsible for human capital inequalities. Workers in poorer regions are less productive in part because they are excluded from working with the most productive technologies and in the most profitable fields, and therefore have less incentive and fewer surplus resources to invest in training and education. They might also be spatially excluded from training and educational opportunities (which are made available through markets as well) even if they wished to pursue them.

The second problem with the human capital explanation is that even after controlling for the education, occupation and experience of workers, the gaps are still massive (Lucas 1990; Krueger 1968). For example, the economic gains of becoming a building laborer in New York versus Mumbai are much larger than the gains of
becoming an engineer in Mumbai versus a building laborer in that same city (Korzeniewicz and Albrecht 2012).

Measures of between-country inequality ignore the contribution of spatial exclusion within nations. Intra-national inequalities are relatively small compared to the international gaps, but we are missing an important piece of the story if we ignore intra-national spatial lumpiness. These range from North/South regional inequalities (Williamson 1965) to more granular systems of racial segregation (e.g. South African apartheid or urban residential segregation). In the United States, Massey (2007; Massey and Denton 1993) and others have noted the impact of residential segregation. “Spatial segregation renders stratification easy, convenient, and efficient because simply by investing or disinvesting in a place, one can invest or disinvest in a whole set of people” (Massey 2007: 19; see also Massey and Denton 1993). If our scope is sufficiently granulated, there is little empirical difference between within and between-space inequality; just the mechanisms that spatially separate workers from work are different at dissimilar spatial scales. In other words, there is not one appropriate unit of analysis in the study of inequality, and our interpretation of inequalities should vary with the unit of analysis.

The impact of differential market access across space is not limited to what we typically think of as between-space (e.g., international) inequality. Korzeniewicz and Moran (2009) point to the spatial character of exclusion for differentiating between high and low inequality institutional arrangements. High inequality is the product of exclusion within a region. Low inequality also depends on exclusion, but the excluded are outsiders. “The institutional arrangements characteristic of within-
country [low inequality] do restrict access to opportunity for large sectors of the population, except that excluded populations now are located primarily outside national borders” (78). The total distribution of resources is similar, but one region enjoys low inequality by creating a more unequal distribution between regions. For example, southern and eastern European immigration at the end of the 19th century undoubtedly increased inequality in the United States during that period but helped these poorer regions close the gap with the richer United States.

Researchers adopting a narrow unit of analysis (processes within a rich country over a short period of time) conclude that domestic labor market institutions are important and international competition less so in driving international inequalities. That is only true because domestic labor market institutions limit competition from abroad. Capital and labor in rich countries are cognizant of the costs and benefits of those mechanisms (e.g., citizenship requirements, tariffs, price floors and other subsidies), and that calculus plays a role in domestic institutional and political negotiations. For example, there are many examples of workers seeking protection from competition from immigrants, but we should also recognize that employers seek compensation if immigration restrictions limit their pool of employable workers. Changes in global labor market conditions reverberate through domestic labor market institutions.

In short, international inequalities are larger than within country inequalities, in part because rich countries protect workers against international competition and hoard access to scarce resources. These mechanisms do not undermine the concept of a transnational division of labor, but reflect the “discontinuity between economic and
political institutions” (Williamson 1979:35). Markets are socially constructed institutions, so those who do not benefit from naked economic competition often seek to reconstruct them. Labor and capital renegotiate their relationship with an awareness of the potential impact on international competition.

Confined Space and Unfree Labor

The unfettered operation of supply and demand for allocating labor is undervalued as an equalizing force. Obviously, certain individuals are more productive than others, by innate capacity, training or choice, and therefore garner more from market exchange. And the creative destruction of economic progress displaces some workers while rewarding others. But wage labor, in which workers are able to negotiate the price of their labor with potential employers without constraint, has a better track record of fairly compensating workers than other historical alternatives. Unconstrained wage labor is not equality, but the invisible hand can be more progressive than visible ones.

Wallerstein (1974) went so far as to say, “When labor is everywhere free, we shall have socialism” (pg. 127). On the other hand, it is “the combination of free and ‘unfree’ labor and land that in fact characterizes the capitalist world-economy” (pg. 149). Through slavery, tenancy, sharecropping and other coerced forms of non-wage labor, capitalists extract profits. “Somewhere in a remote village at this moment a non-wage worker is producing a surplus in which, via multiple intermediaries, each one of us is partaking, if to different degrees” (pg. 127).

Free labor is mobile, and because no labor is perfectly mobile no labor is perfectly free. Labor is ultimately unfree when workers are stripped of their right to sell their
labor—i.e., when they are constrained to a single work site and ‘employer’. Slaves are forced to ‘sell’ their labor in a specific work site. Workers have a stronger negotiating position, and thus receive higher wages, as they gain access (including spatial access) to new potential work sites, just as a merchant that has access to more markets will tend to find a higher price for her wares. There are many barriers that can separate workers from work, but none is more important than physically/spatially preventing movement between work sites.

The same is true of the sale of goods; goods are ultimately consumed in space, and physically limiting access to those sites of consumption reduces the sale value of those goods. Buts goods are fundamentally more mobile than labor, so those who profit from their exchange have an inherent economic advantage over those who depend on selling their labor. Financial capital is even more mobile, and today the flow of financial capital is limited only by its liquidity.

For Wallerstein (1979), a necessary characteristic of the capitalist world-economy is the multiplicity of states in a single economy. Within this setting, actors must pursue profits. “The attempts of these actors to use non-market devices to ensure short-run profits makes them turn to the political entities which have in fact power to affect the market – the nation-states” (1979: 17). More powerful states impose themselves on weaker territories, and capitalists in core regions reap the benefits of ‘unequal exchange’. “If states did not exist, in short, capitalism would have had to invent them” (Harvey 2006: 105).

Ignoring “non-market devices”, we can apply the Heckscher-Ohlin model to explain any relationship between globalization and inequality (Williamson 2006):
regions will export those local resources that are relatively underutilized. Demand for labor will increase relative to capital where demand for labor was low, and vice versa, so that inequality will fall where demand for labor was relatively low (and inequality high) and will rise where demand for labor was relatively high (and inequality low), so that international inequalities will converge. But we cannot ignore non-market devices; considering that humans have bought and sold other humans in the market, the very concept of non-market devices as a distinct category is a farce.

‘Unequal exchange’ (e.g., the exploitation of labor) always involves the manipulation of space. Even the most coercive and violent systems for organizing labor have to account for space in their operations; for example, it is not slavery if slaves can easily leave one place and access markets in another (and, thus, the importance of the Fugitive Slave Act of 1850 in the run up to the American Civil War). But when spatial control over slave labor is imposed, slave owners are able to convert their spatial dominance of slaves into a price advantage over competing producers. Some of those profits are then reinvested to bolster their spatial power.

Power over space can be used in other ways to restrict market access and create opportunities for excessive profits. Market exclusion creates opportunities for arbitrage. The ‘law of one price’ suggests that the same good in different markets, after accounting for transportation costs, should have the same price. If not, actors can purchase goods at the cheaper market and sell them at the more expensive market to earn a risk-free profit. If only one actor is able to move goods between markets, that actor can extract profits with little risk and relatively little investment.
Control over market access also allows insiders to hoard opportunities (Tilly 1999). Returning to *The Wealth of Nations*, Adam Smith notes that higher wages in towns should attract workers from the country.

In every country of Europe we find, at least, a hundred people who have acquired great fortunes from small beginnings by trade and manufacturing, the industry which properly belongs to towns, for one who has done so by that which properly belongs to the country . . . [S]tock and labor naturally seek the most advantageous employment. They naturally, therefore, resort as much as they can to the town, and desert the country. (Smith 1776/2004: 109)

Towns developed a system of “voluntary associations and agreements, to prevent that free competition which they cannot prohibit by bye-laws” (Smith 1776/2004: 109). In other words, town residents hoarded the more profitable employment by restricting access to it (Tilly 1999; Korzeniewicz and Moran 2009). Going further, town employment is also more profitable in part because country employment is not hoarded. Because labor in the country is cheaper than labor in towns, town workers are able to import cheap foodstuffs and other raw materials through which they are able to add value and maintain their higher wages—again emphasizing the relational character of inequality. Smith did note that, in time, even if labor was not free to find employment in the towns, opportunities for growth in towns would become scarce and investment would flow to the country, reducing wage gaps.

*Conceptualizing Space*

The manipulation of space (e.g., barriers to migration) is critical to our understanding of economic inequalities between and within countries. But I have focused to this point on the political manipulation of space. A better understanding of the interaction between space and inequality requires a richer conceptualization of space. Space is not just a geometric area, “a simple and immutable container within
which social processes occur” (Harvey 2006: 77). Social processes occur through space and time, and the operation of these processes is influenced by the distribution of power.

The ability to control the timing and spacing of human activities . . . reflects the distribution of power and the control of resources. Relations of power, structures of inequality, and practices of domination and subordination are embedded in spatial design and relations. Thus spatial arrangements are both products and sources of other forms of inequality. (Tickamyer 2000: 806)

To understand the relationship between space and diffusion, and, therefore, between space and inequality, we need to draw on all three components of space: as place, as scale and as network (Tickamyer 2000; Lobao 1996). Unequal exchange occurs at the intersection of all three.

Space as Place. In the sociology of spatial inequality, place is “the particular locale or setting” (Tickamyer 2000: 806). It includes the “production of nature” through the interaction of productive technologies and the local environment (Smith 1990). “Regional consciousness and identities, and even affective loyalties” may build within a region from shared consumption patterns and culture that can cut across class lines (Harvey 2006: 102). In short, a place is both a location with certain material characteristics, but also a geographic region with an organizing principle and institutional environment—e.g. the sovereign state or municipality.

Place finds expression in economic models that include local ecological or institutional characteristics as variables (see Sachs and Gallup 1998). Adam Smith pointed to differences in place to explain the advantage of town over country: “The inhabitants of a town, being collected into one place, can easily combine together. . .
The inhabitants of the country, dispersed in distant places, cannot easily combine together” (Smith 1776/2004: 109).

Too often, explanations of inequality within and between places reduce to detailing characteristics of place. For example, Firebaugh (2003) links recent developments in global inequality to the “deepening industrialization of poor nations” (Firebaugh 2003: 23). Industrialization of poor nations, it is argued, has produced international convergence but new intra-national inequalities.

Any approach that focuses primarily on developments within countries is incomplete because of Wallerstein’s discontinuity between political and economic institutions. The global economy is charged by a single division of labor—the temptation to treat some sovereign states as isolated economic systems is only further evidence of the impact of exclusion (see space as network) on the practical functioning of the world-economy. The state is the focal point of power when the place is a country, but power is also organized and barriers erected at other scales.

Space as Scale. Scale is the most underappreciated component of space in the study of inequality. Scale has been invoked by the inherently more spatial subdisciplines in the social sciences (Tickamyer 2000), often for its role in impeding diffusion. For example, John Snow mapped cholera outbreak clusters in London in 1854. Princeton’s European Fertility Project in the 1970s found evidence of diffusion in demographic transitions (Coale and Watkins 1986). Sociologists have also used scale

---

7 He additionally notes economic integration, technological change, demographic windfalls, growth of the service sector, and the fall of communism
to map residential segregation and the geographic distribution of poverty, and how the two may be related (Massey 2007; Iceland, Weinburg and Steinmetz 2002).

In economics, scale has two theoretical manifestations: economies of scale and transportation costs. Early growth models assumed constant returns to scale; productivity was independent of the spatial proximity of other productive activities. Constant economies of scale made growth models insensitive to space (Pritchett 2006). Paul Krugman and other trade economists in the 1980s began to emphasize agglomeration as an outcome of economies of scale, factor price differences, and transportation costs (Krugman 1991; Helpman and Krugman 1985; Fujita, Krugman and Venables 2000). In short, this means that “new production tends to be drawn to existing production locations” (Harvey 2006: 98).

Space as Network. Inequality is relational—one cannot be unequal in isolation. For example, to Marx, the condition, or even the existence, of the proletariat does not make sense in isolation, but only in relation to the bourgeoisie, and vice versa. Tilly (1999) bridged Marx’s exploitation and Weber’s social closure to develop a model of durable inequalities. “I claim that an account of how transactions clump into social ties, social ties concatenate into networks, and existing networks constrain solutions of organizational problems clarifies the creation, maintenance, and change of categorical inequality” (1999: 21).

These unequal relations are embedded in and across space. Relations between places can facilitate or hamper interactions between segments of populations in each place differently, and are therefore an important source of inequality within and between places. Dependency and world-systems approaches have drawn on this
relational character of inequality to theorize the impact of unequal relations between larger social units (Wallerstein 1979; Rubinson 1976; Cardoso and Faletto 1979). “The strength of the state machinery in core states is a function of the weakness of other state machineries” (Wallerstein 1979: 21).

Embedded spatial relations operate within and between regions. For example, in the world-economy today weaker states tend to be associated with higher levels of inequality because they lack the capacity and incentive to respond to the poor (Rubinson 1976). On the other hand, stronger states are built on broader compromises, so advantages won by the state are more broadly enjoyed within the country (Phillips 1993; Arrighi 1994/2010: 325).

Alternatively, we can conceptualize actors within regions as members with limited participation rights in the world-economy. Actors with greater participation rights (more extensive market access both within and between regions) are in a stronger position, and therefore will tend to see higher profits. Within a region, these advantaged actors can reinvest excessive profits to reinforce their advantaged position—e.g., slave owners pushing pro-slavery legislation. In the modern world-economy, elites within a country can manage exports and imports to a country, and therefore monopolize on the profits from the local comparative advantage. The sites of production, exchange, and consumption, and power, are distributed unevenly across the spatial plane, and an individual’s economic opportunities are dependent on their position relative to the sites of profit and the barriers erected to restrict movement to them.
Innovation and Diffusion in the World-economy

At face value, this next section is a review of literature on international inequalities, but the goal is not to explain international inequalities. Instead, a discussion of durable international inequalities highlights the scaffolding of the world-economy that structures the production and accumulation of wealth in space. The configuration of space and the inequalities that configuration produces are intimately linked; the configuration consistently funnels wealth to some and away from others, and the beneficiaries of the configuration are, in turn, literally invested in sustaining the status quo. In other words, I show how a contoured space (composed of place, scale and networks) allows for spatial structures that constrain innovation and diffusion. Drawing on trade economics and development literatures, I argue that these spatial structures are self-reinforcing and form durable configurations that constrain innovation and diffusion.

Adam Smith sought to explain “the wealth of nations,” and countless social scientists have investigated uneven national economic development since. I next explore a portion of this literature that deals specifically with “spatial inertia”, or why the rich tend to stay rich and the poor tend to stay poor. This spatial inertia is a symptom of structured distributions of innovation, production, exchange, migration, and wealth, in other words, a spatial configuration of the world-economy. An analysis of the formation (and collapse) of spatial configurations is important for our purposes because they are responsible for consistently funneling profits to some and away from others.
David Harvey (2006) organizes arguments of uneven geographic development in four categories: *environmentalist, constructivist, geopolitical*, and *historical/diffusionist*. First, environmentalist arguments look at “the ways in which human adaptations to variegated environmental possibilities underlie territorial specializations, divisions of labor, and the creation of distinctive regional ways of life” (Harvey 2006: 73). For example, Diamond (1997) has argued that Eurasian dominance is a product of ecological advantages, including climate, high-yield grains and domesticable animals.

Ecological models can also be relational. Pomeranz (2000) argues that ecological, cultural and political differences between pre-industrial England and China have been overstated. The real advantage that led to England’s industrialization was a coincidental co-location of coal and population centers in England and the country’s access to the natural resources of the New World. New World resources allowed England to break from a cycle of adopting more intensive agricultural techniques to feed a growing population on a constrained supply of arable land.

Environmental approaches are incomplete if they do not consider the interaction between natural endowments, productive technologies, labor, competition, and consumption. For example, the profitability of mineral deposits depends on the demand for that resource in other economic activities and the technology for economical extraction [profitability is dependent on our capacity for material appropriation]. When we make this model dynamic by adding technological change and local resource depletion, unequal development is not static, but fluctuates with technology regimes and time.
Harvey’s next two sets of arguments on uneven development are *geopolitical* and *constructivist* (Harvey 2006). Both point to the role of state capacity, but the two approaches differ on the important mechanisms that link state capacity to development. Geopolitical arguments point to an institutional survival of the fittest; certain arrangements—state structures, political blocs—privilege some populations over others (for example, see Birdsall 2002; Sachs 2005; North 1990). Constructivist arguments emphasize the effect of stronger states on weaker states (see Arrighi 1991; Wallerstein 1983/1996).

Harvey’s final set of arguments to explain uneven development are *historicist/diffusionist*. These arguments generally adopt models that focus on technological innovation and adaption; as it pertains to industrialization, the West developed a new set of technologies that allowed it to produce more efficiently. Those regions that adopt this new set of technologies share in these advantages. Other regions fail to jump on board and remain backwards.

Alone, historicist/diffusionist models predict a temporary spike in interregional inequality followed by convergence through diffusion. Firebaugh (2003), for example, expects convergence to come through the diffusion of industrial technologies and techniques. But models that focus on diffusion in time only fail to account for an interdependent global economy populated by variegated natural, built and institutional environments, and distances that unevenly distribute comparative advantage across space.

It is important to recognize that the global economy evolves with time. “The fundamental impulse that sets and keeps the capitalist engine in motion comes from
the new consumers’ goods, the new methods of productions or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates” (Schumpeter 1942/1950: 83), so developmental models that were successful in the past are unlikely to be successful in the future, in part because they were already exploited successfully. Arrighi, Silver and Brewster (2003), for example, argue that industrialization is no longer the path to convergence between countries. Industrial capacities have diffused, so the profit margins enjoyed by the early industrialists have shrunk. American producers have shifted away from manufacturing (employment in manufacturing has dropped to levels last seen in the 1940s despite heavy growth in the size of the labor force (BLS 2010)), not because it is less productive, but because it is less profitable and, as a result, market share has shifted elsewhere.

Second, successful innovations—those innovations that generate profits and are diffused—exploit a resource base, an institutional environment, a labor supply, and profitable markets. In other words, innovations are located in space, and its diffusion requires adapting that innovation to a new environment. One implication of spatially embedded innovation is that creative destruction also adopts a spatial component: innovation in one place, suited only to that place, can make previously profitable activities particularly suited to another place redundant. More frequently, innovation accumulates; it exploits economies of scale and wealth generated by prior innovation.

Before Schumpeter, Marx pointed to processes of creative destruction in capitalism. Capitalism is inherently vulnerable to overaccumulation crises; productive capacity exceeds demand, prices fall, and profits are erased. “And how does the bourgeoisie get over these crises? On the one hand by enforced destruction of a mass
of productive forces; on the other, by the conquest of new markets, and by the more thorough exploitation of old ones” (Marx and Engels 1848).

This concept was extended by Arrighi (1994/2010) who noted cycles of material and financial expansion, with each cycle ending in interstate chaos and the emergence of a new hegemon—Holland from Genoa, to Britain, to the United States. Brenner (2006:4, 8) argues that we are now in the midst of “a long-term and system-wide economic downturn” from “the over-capacity and over-production that result from intensified horizontal, intercapitalist competition.” Larry Summers (2013) borrowed a term from the 1930s, secular stagnation, to describe a situation of savings exceeding investment, and investment rates already too low to further induce borrowing. Harvey (1982) argued that uneven development means that overaccumulation crises are not always synchronized across the system, but can be regionally specific. “The timing of upturns and downturns in the accumulation cycle can then vary from one region to another . . . . different regional rhythms . . . can just as easily compensate each other as build into some vast global crash” (Harvey 1982: 427-28).

Innovation creates opportunities for profit and therefore attracts capital. In time, though, that innovation becomes old, diffused, and supplanted by something newer and better. But the capital invested in the worn-out innovation is already sunk (Brenner 2006), and a system of institutions and networks are structured around it. In other words, the “destruction” of creative destruction lags behind and fails to create space locally for more creation. Profits dry up, and capitalists adopt one of Marx’s solutions—the destruction of productive forces (war, economic depression) and/or the conquest of new markets (colonization, financialization of the economy). When one
industry is supplanted or relocated in a place, it will also affect neighboring industries that in some way profited from their co-location (Pritchett 1997), and when the push for change gains enough momentum, the system organizing production and exchange, the spatial configuration, undergoes a revolution.

In practice, inequalities between some regions are deeper, and inequalities across other regions flatter, and these patterns are more stable than we might expect from the processes of creative destruction described above. This spatial inertia is the product of place-independent production, economies of scale and other institutional and non-institutional constraints on innovation and diffusion that often overpower those forces that encourage convergence between regions.

A key principle of developmental spatial inertia is that some production is not strictly rooted in space. “A merchant . . . is not necessarily the citizen of any particular country. It is in a great measure indifferent to him from what place he carries on his trade; and a very trifling disgust will make him remove his capital, and together with it all the industry which it supports, from one country to another” (Smith 1776/2006: 394; quoted in Arrighi 2007: 63). In other words, some producers are (somewhat) free to move production to that location which offers them the greatest profits.

The meaning of place-independent production has changed since Adam Smith. At the time of the first national census of 1790, when 90% of the US population worked in agriculture, work sites were deeply rooted in space; farmland cannot be shifted to another location to take advantage of cheaper labor or higher prices. Early manufacturing was also tied to natural resources—running water for power and
mineral deposits. Space was revolutionized in the 19th century, largely by the steam engine. Not only did the steam engine dramatically reduce transportation costs (on land and water), but it also offered a source of power that was more reliable and powerful than animal power and more mobile than water power. Production is never completely place-independent (a lesson which has been reinforced as world cities have survived the theorized onslaught of the cyber age), but through time, people have shifted to work that is more dependent on the institutional and built environments and labor supply than the natural environment and resource base.

The dominant convergent force between regions is that capital can find greater profits in less developed regions precisely because they are less developed. In economics, this is known as the catch-up effect:

Consider two countries producing the same good with the same constant returns to scale production function, relating output to homogeneous capital and labor inputs. If production per worker differs between these two countries, it must be because they have different levels of capital per worker: I have just ruled everything else out! Then the Law of Diminishing Returns implies that the marginal product of capital is higher in the less productive (i.e., in the poorer) economy. If so, then if trade in capital good is free and competitive, new investment will occur only in the poorer economy, and this will continue to be true until capital-labor ratios, and hence wages and capital returns, are equalized. (Lucas 1990: 92)

Returning once more to Smith’s discussion of town and country, he noted that in time profits in the town would be reduced enough by competition that town residents would find it necessary to invest in the country, and the two would converge.

The Marxian version also emphasizes diminishing returns. Competition for profits between capitalists drives down the returns on capital to zero. Like Smith, Marx noted that the law of falling profits “is only a tendency” and “it is only under certain circumstances and only after long periods that its effects become strikingly

But it is more than the will of the rich that prevents capital from pursuing cheap labor. Looking back at Lucas’ description of the catch-up effect, the logic appears sound, but the assumptions are not robust. Lucas (1990) notes that three assumptions behind his quote above are violated: that labor across countries is homogenous, that capital markets are perfect, and that there are no external benefits of human capital.

Lucas’ first violated assumption—that labor supplies are homogenous—has some effect on aggregate income levels, but wage gaps between similarly skilled individuals across countries are still huge (Krueger 1968; Korzeniewicz and Albrecht 2012). In other words, regions do not fail to converge because of differences in the base levels of human capital.

Second, because of poor institutional development or active intervention, capital fails to move efficiently. Poor institutions are costly, and can create inequalities, because they prevent the efficient allocation of resources. An individual or group can be excluded from an opportunity because they are not able to access the investment capital necessary to purchase the raw materials, tools, training or transportation to produce and distribute a finished product, even when the investment would be profitable. They are not directly excluded from competing economically by market failure.

More important is Lucas’ third violated assumption, that there are no external benefits of human capital or, more generally, constant returns to scale. For a number
of reasons, there are benefits to setting up shop closer to other shops, none more important than the co-location of experienced labor. As a result, productive activities tend to cluster; the role of agglomeration economies has received increasing attention, especially in trade economics (Krugman 1991).

Harvey (1982, 2003) discusses agglomeration as the outcome of a ‘spatial fix’. Faced with an accumulation crisis, capitalists invest in construction and infrastructure, which can absorb capital in the moment and make a region more productive and efficient in the future. Larry Summers recommends that the government begin a large investment in infrastructure to overcome secular stagnation. The fix (solution) is temporary, but the fix (spatially immobile) has a longer timeframe. Large scale construction projects root production in place. “The vast quantities of capital fixed in place act as a drag upon the capacity to realize a spatial fix elsewhere” (Harvey 2003: 116).

Capitalists and workers resist capital’s spatial relocation—a ‘switching crisis’. “Territorial alliances, which often became increasingly powerful and more deeply entrenched, arise to protect and enhance the value of capital already committed within the region” (Harvey 1982: 428). Capitalists are locked into these alliances to protect spatially fixed investments and because they are dependent on the extra-market interventions of the state, which require cross-class compromise.

Once fixed physically and institutionally, producers in these locations have both the incentive (to reduce costs) and resources to drive innovation (Arrighi, Silver and Brewster 2003), and their innovations tend to reflect the needs of that place. Convergent forces tend to dominate across territories united by a “regional
consciousness” (e.g. within a country, though some regional gaps can be still be surprisingly durable (Williamson 1965)), while divergent forces tend to dominate elsewhere. The result is the formation of relatively stable zones of predominating prosperity and zones of predominating depression (Arrighi and Drangel 1986; Arrighi, Silver and Brewster 2003: 17).

Part of “conserving privileges already won” is restricting access to these fixed productive technologies.

[A] social relationship may provide the parties to it with opportunities for the satisfaction of spiritual or material interests. If the participants expect that the admission of others will lead to an improvement of their situation, an improvement in degree, in kind, in the security or the value of the satisfaction, their interest will be in keeping the relationship open. If, on the other hand, their expectations are of improving their position by monopolistic tactics, their interest is in a closed relationship. (Weber 1968: 1:43; quoted in Tilly 1999: 7)

A territorial alliance, especially one built on notions of nationhood, will often seek to limit immigration. On the other hand, if they truly have captured a valuable advantage, outsiders will be looking to gain spatial access.

One result that may arise from such a scenario is the contemporary condition of ‘everything but labor globalization’—a world-economy organized around mobile capital and spatially excluded labor. With the tendency for production to be centralized, labor mobility is the most powerful force towards convergence. Convergence is not about the relative productive capacity of a region, but the relative productive capacity per worker. Workers want to maximize wages which, in the right circumstances, translates to migration. In fact, given the tendencies towards perverse capital flows, from poor to rich, labor mobility may be the best option available.
Immigrants from poor countries who find work in rich countries might earn as much as ten times more than they did before migration (Word Bank 2002).

That migration can positively affect the migrating individual is straightforward, but the economic theory of interregional migration and interregional convergence is complex and inconclusive (Lucas 2003), but a growing body of empirical work supports the connection. Puga (1998) links greater levels of regional inequality in Europe compared to the United States to lower mobility between regions (nations in Europe and states in the United States) despite greater centralization of productive activities in the United States. Pritchett (2006) provides evidence of a similar phenomenon when comparing the United States and the Atlantic economy historically with less-developed countries today. Other researchers have demonstrated wage convergences that coincide with variable population growth rates (Blanchard and Katz 1992; Glaeser and Gyourko 2003).

In short, if the sites of productive innovation are less geographically rooted than in the past, but have become physically rooted, territorial alliances and economies of scale will encourage the agglomeration of production. Labor mobility becomes the primary vehicle for regional convergence (short of a spatially reorienting overaccumulation crisis as noted by Smith and Marx). Labor mobility, in this context, is diffusion by infusion; the most profitable techniques or capital do not diffuse spatially to workers, so workers move to the most profitable sites.

In summary, durable regional inequalities are a product of innovation (creation) across an ecologically, infrastructurally, demographically, and institutionally variegated space, and the uneven distribution of power to manipulate that space.
Innovations are localized in space to exploit unique environments. The concentration of wealth, infrastructure, complementary industries, exclusive institutions and knowledge mean that profits tend to be reinvested locally. Space becomes configured as the incentives to invest and trade are structured by the distribution of infrastructure, labor, capital, institutions and technologies. As wealth accumulates in some regions, costs rise. Productive capacity exceeds local demand (overaccumulation) in wealthy regions, and financial capital from those regions finds higher returns in non-traditional markets.

Over time, many of the pillars that supported the existing spatial configuration are weakened. Physical capital in the core becomes less profitable, but also makes it more difficult to adopt newer, potentially more profitable technologies. Institutions that structured relations domestically and internationally are undermined by changing economic demands. Capitalists turn to the state to protect their interests in markets around the world, and emerging markets, bolstered by capital investments from rich countries, look to bring their international status in line with their growing economic clout. These contradictions lead to interstate chaos, or a reshuffling of the global hierarchy. War, civil and international, can destroy physical infrastructure, and paves the way for a new spatial configuration (although many of the pieces end up falling close to where they had been laid previously).

Arrighi (1994/2010) describes this process as systemic cycles of accumulation. Each cycle has two phases, material expansion and financial expansion. Material expansion (a rapid investment in productive capacity) leads to overaccumulation. The transition to financial expansion is marked by a switching crisis, after which capital is
increasingly shifted out of production and trade and into finance. This transition only
forestalls the eventual terminal crisis, a series of events that marks the end of the
“dominant regime of accumulation.”

The spatial configuration consistently funnels wealth to certain activities and
places and away from others, creating stable patterns or regimes of inequality.
Therefore, it follows that a change in the configuration of space should introduce a
new pattern or inequality regime. In other words, we should expect a synchronization
between spatial reconfigurations and inequality regime changes, or in the context of
Arrighi’s systemic cycles of accumulation, the synchronization of crises (be they
signal or terminal) and inequality regime change.

*From the Global to the Local*

Direction of Change. Spatial reconfigurations provide the conditions for change, but
say nothing about the nature of that change. Crisis creates a state of disequilibrium, a
period during which social action has greater leverage than when the system is in a
stable equilibrium. But pre-existing power structures, while vulnerable, continue to
exist. Institutions continue to define the “rules of the game” even if those rules are
more likely to be broken. Existing infrastructure, natural and constructed
environments, and agglomeration points of economic activity are more likely to recur
in the new configuration, but some are made redundant and others are physically
destroyed in the transition. In short, the new configuration and new regime will
reflect a transformation of the old, not the creation of a new. Change is path-
dependent (Acemoglu and Robinson 2012).
On the other hand, the emergent system will reflect a conscious reaction against the old system precisely because the old institutions and hierarchies will define the points around which action will be organized. Social action will focus on key pivot points of the old system, but the balance of power shifts because the crisis reflects a failing of the old arrangements. Those parties advantaged by the old arrangements will be weakened because 1) the configuration that structured their advantage is weakened and 2) the logic or ideology that justified the old arrangement will lose authority. As a result, change in the institutions governing a particular set of arrangements, e.g., the relationship between workers and managers, will tend to fluctuate cyclically. The net result is that, if inequality rose with the last transition, there is a greater than random chance that inequality will fall with the next regime change.

Finally, transition generates room for political, institutional, and technological innovation. These innovations diffuse globally, so that institutional, political and technological forms can spread rapidly around the world. The emerging global hegemon emphasizes innovation (and change in general) that is conducive to a new spatial configuration centered on it – e.g., the United States quickly took up the mantle from the United Kingdom as World War II was coming to a close to cajole other states and the international system to adopt institutional and political forms that were consistent with its vision of the emerging spatial configuration.

The result is that the new regime will tend to adopt three characteristics: 1) it will share many features of the old regime; 2) change will tend to be cyclical; and 3) new
“solutions” to the crisis of the world-economy will tend to diffuse through the world-economy, often disseminating from the inside out (or core to periphery).

I propose the following hypotheses: 1) We can predict points of wealth accumulation (the character of an inequality regime) from the configuration of the world economy. 2) Inequality regime change will be synchronized with spatial reconfigurations. 3) The new regime will share many features with the old regime, but will tend to produce a reversal in the degree of wealth and income concentration, and some characteristics of the new regime will have been disseminated from the most powerful state(s) at the time.

_Systemic Causality and the Counterfactual Fallacy_

In this project I employ a world-systems perspective. This means that I accept as a matter of principle that the key processes driving income inequality in the United States are not all contained within its political boundaries: e.g., global flows of goods, capital, and labor/technology, ideologies, and policies; unequal relations and access to markets; global economic dynamics (systemic cycles of accumulation); domestic investment patterns (economic specialization and access to capital). The historical example of slavery is particularly useful. Slavery was the principle source of inequality in the colonies, both in the unequal relations between slaves and slave owners and in the unequal relations between slave owners and free households without slaves. The institution of slavery was adapted from slave models in colonies to the south. The supply of slaves was made possible by unequal relations between Europe and Africa. And a relatively large share of goods produced with slave labor
were exported to Europe and the Caribbean. Slavery, and its impact on inequality, within the United States makes no sense without a global perspective.

This perspective has a few important methodological implications. First, it is not appropriate to treat countries as equal and independent observations. They are not independent. Country-level outcomes in the United States depend on processes within countries around the world and the networks between these countries. And they are not equal. Concepts of Westphalian sovereignty aside, Luxembourg and China are not analytically comparable. They should not be treated as standardized units of observation.

Second (and building on the first), generalization, and therefore significance testing, is devoid of meaning because there is only one world system. A study of income inequality in the United States can apply theory to define mechanisms, develop hypotheses, and identify correlations. But that single observation, with a sample size of one, represents the entire universe of the study.

A popular approach to the study of within-country inequality is to look at inequality change within a country over time and test for significant covariance with explanatory variables. The potential for omitted variable bias using this approach is overwhelming. Not only do we need to recognize the potential impact of reflexivity (past outcomes can influence future outcomes), but the entire global-historical context changes with time. These limitations ultimately mean that it is impossible to establish an appropriate counterfactual. Any quantitative measure of a variable’s impact requires a control – a situation in which the variable is not allowed to vary. But in the case of the United States there is no control. For the reasons stated above, the United
States of the past nor Germany in the present is an appropriate control for the United States in the present.

For example, what would income inequality in the United States look like today if the minimum wage had kept pace with inflation? Many researchers have approached this problem with some form of multivariate analysis, comparing observations across countries or over time. These all assume some scenario of *ceteris paribus* – what would inequality look like if all else was equal. But all else *cannot* be equal; in fact, all else must be different; it is a fundamentally different world in which the minimum wage in the United States keeps pace with inflation. Minimum wage does not change (or not *not* change) in a vacuum, and it defies rationality to propose a scenario in which it does. Therefore, it is fundamentally impossible to uniquely measure the impact of the minimum wage (or any other explanatory variable) on inequality in the United States.

The emphasis on measuring the relative contribution of a host of potential explanatory variables on inequality is motivated by a search for policy implications. If we could isolate the impact of the minimum wage on income inequality, for example, minimum wage policy would be better informed. This is a valid goal. But this project has a different goal. I seek to explain the long-term changes in wealth and income concentration. Projects that adopt a narrower unit of analysis seek to explain fluctuations in the size distribution of income (i.e., inequality) over a short period of time. By contrast, I seek to explain the larger historical transitions. The difference between the two is notable in Figure 0.1b. My focus is on the major swings between
1920 and 1940, and again after 1970, not the substantial but impermanent fluctuations we see at shorter time scales.

In this sense, this project is not meant to challenge the traditional literature on within-country inequality, and in fact I draw heavily on that research to identify the key mechanisms linking spatial configurations to the concentration of wealth and income in the United States. The difference is the interpretation of causality. Change is not the cumulative outcome of a host of independent variables; it is systemic. I do not treat minimum wage policy as an independent variable, but as an institutional solution, one of many, designed to achieve a particular outcome. My goal, then, is to identify and explain a constellation of changes (economic, institutional, technological) that, in turn, create a new economic elite and increase or decrease the concentration of wealth and income.

The Dependent Variable. Throughout the project I refer to within-country inequality as the phenomenon I am explaining, but I do so only as a form of shorthand. My real focus is wealth and income concentration, or the share of wealth or income that is controlled by those at the top of the distribution. In other words, I limit my attention to one point in the distribution instead of the entire distribution. I do so for four reasons. First, changes in the distribution of income or wealth across the entire distribution is excessively complex; a mechanism that impacts the distribution of income at the top of the distribution may not impact those at the bottom of the distribution, and vice versa.

Second, the distribution of incomes across the entire distribution is often driven by developments at the top of the distribution. Abstractly, this is true because the
richest tend to control such a large share of total wealth and income. This can be demonstrated geometrically using the Gini coefficient, the most popular point estimate of inequality in a population. The Gini coefficient is equal to twice the area between the Lorentz curve (which traces cumulative income or wealth by cumulative population) and the 45 degree line. In the chart below (Figure 1.3), the Gini coefficient would be equal to twice the entire shaded area (light and dark).

Using this metric, if we are given a top income share—for example, the top 10% of earners receive 33% of total income—the geometry of the Gini coefficient gives us an upper and lower bound of possible values, a maximum and minimum Gini coefficient for the entire population. On the chart above, the lightly shaded area would be the minimum Gini—if income were to be distributed as evenly as possible across the population given that the top 10% of earners receive 33% of total income, this would be our income distribution. This area is equal to \((y-x)/2\), where \(y\) is the income share and \(x\) is the proportion of the total population, and the minimum Gini is equal to \(y-x\), or .230. We can also calculate the maximum Gini as \(1-x/y\), or .700. In this case, the true Gini is .451. The minimum Gini is equal to 51.4% of the true Gini. Assuming a log normal distribution of income (and large population distributions rarely deviate enough from a log normal distribution to invalidate this assumption), the minimum Gini based on the income share of the top 10% of earners generally represents about 50% of the true Gini, the top 5% about 33% of the true Gini, and the top 1% represents about 13%. 
Empirically, the importance of the top shares may be greater than this abstract discussion suggests. For example, according to the Census Bureau (2014), the US Gini for household income inequality increased from .403 to .468 between 1980 and 2009. During that time, the income share of the top 5% of households increased from 16.5% to 21.7%, lifting the minimum Gini during that period from .115 to .167. This rise in minimum Gini (of .052) is equal to 80% of the total rise in the true Gini over that period (and 94% of the rise between 1980 and 2000).

In addition to its direct impact on inequality, top income shares can inform us about the distribution of incomes below. Income shares at other levels are correlated because income and wealth distributions tend to follow predictable patterns. Gains by the top 10% of earners are usually mirrored by smaller gains for the next 10% and losses for the bottom 40%. For example, Piketty and Saez’s (2003) top income shares explain about 97% of the variance in US household income inequality even though
they only provide data for the top 10% of earners (see chapter 4 for a more detailed discussion).

Third, the abstract and empirical importance of top income shares is also theoretically important because it forces us to emphasize explanations for rising or falling inequality that focus on the richest members of a society. I would argue that it is more sociologically relevant if the top 1% of earners capture an additional 5% of total income than if the top 40% of earners capture an additional 10%. Most research on rising income inequality in the United States, though, has focused on the latter.

Finally, historical data is often collected and/or reported only for the richest individuals or households, e.g., the percent of total wealth or income controlled by the richest 1% or 10% of households. Additionally, before 1913 there is little reliable information on the distribution of incomes, particularly at the top of the distribution. Therefore, from 1600 to 1913, I focus primarily on trends in the concentration of wealth; after 1913 I look also at the concentration of income.

**Summary**

In this project I offer an explanation for the long-term changes in the concentration of wealth and income in the United States. Against a literature that emphasizes multivariate models of change, with researchers promoting competing explanatory variables, I offer a systemic model. Specifically, the spatial distribution and flow of wealth globally is structured by an integrated configuration of international and domestic institutions, physical infrastructure, economic and political ideologies, and informal networks. These conditions determine to whom and to what extent wealth and income should become concentrated.
Change is the result of contradictions within these spatial configurations. As the configuration ages, profits fall in older economic activities and new competitors—new technologies, new markets, rising national powers—are constrained by the institutional arrangements, relations and infrastructure of the spatial configuration. These contradictions lead ultimately to crisis, chaos and a reconfiguration. The character of wealth and income concentration change as a result.

Each of the following chapters deal with a specific historical period: the colonial inequality regime from 1600 to 1790, the commercial regime from 1790 to 1860, the corporate regime from 1865 to 1929, the Keynesian regime from 1929 to 1973, and the financial regime from 1973 to the present. I also use chapter 5 to speculate on the next inequality regime. Each chapter describes the crisis that drove inequality regime change, the global context of crisis, and situates the emergent inequality regime within the new spatial configuration.
Chapter 2: The Colonial Inequality Regime(s)

The story of wealth and income concentration in the United States is composed principally of two narratives. The first tracks the systemic cycles of accumulation of the world-economy. Early colonization correlated with the financial expansion near the end of the Genoese cycle. The end of the colonial inequality regime (and the end of British control of the territory) can be causally linked to conflict at the end of the Dutch cycle (see chapter 3). As such, wealth concentration in the colonial inequality regime reflects the configuration of space during the Dutch cycle (see Figure 0.3).

The second narrative is the secular convergence of the United States with the centers of wealth and power in the world-economy. In the early colonial period, the territory that would become the United States was a second-class colony. By the mid-20th century, the United States was the center of economic, financial, and military power in the world-economy. Economic inequality is ultimately a product of the exclusive access to wealth generating activities situated in space, the intersection of functional distance and institutional constraints on access. The colonial inequality regime is a useful foil for understanding the effect of convergence on wealth and income concentration in the 19th and 20th centuries.

In this chapter I discuss the spatial configuration of the Dutch cycle, the peripheral position of the colonies in that configuration, and the resulting degree and character of wealth concentration in the colonies. Our interpretation of economic inequality in the territory that would become the original United States during the colonial period depends on where we choose to look. The experience of puritanical Massachusetts bears little resemblance to that of South Carolina. Inequality in South
Carolina was high if we consider only free households, and it defies modern comparison if we include that half of the population living in bondage (Jones 1977; 1980). Relatively egalitarian Pennsylvania was more dependent on indentured servitude than any other colony (Herrick 1969). Inequality rises if we count British proprietors and married women whose legal rights, including ownership, passed to their husbands (Shanahan and Correll 1997). In other words, our interpretation of colonial economic inequality will vary depending on geography and the unit of observation.

That being said, wealth concentration across free, resident colonial households was lower than in contemporary England, and lower than in the United States towards the end of the 19th century. Wealth and status were principally defined by land ownership, and colonial administrators were giving land away. With land in abundance, labor was scarce and relatively pricey. Because land was relatively cheap (Christensen 1981), the very wealthiest often distinguished themselves by accumulating other, more portable assets; slaves and ships were particularly important.

Early colonial wealth accumulation was limited by market size and reach. In 1700, the purchasing power of Europe (gross domestic product adjusted for the relative cost of goods locally) was 20 times that of what would become the United States, and England’s purchasing power was more than 7.5 times greater than that in the colonies (Maddison 2008; author’s calculations). The ability to accumulate wealth in the colonies was tied to one’s capacity to exploit American resources and export goods to Europe or the sugar colonies; the colonial economy was not large enough,
and too dependent on subsistence farming, to produce huge profits for a large-scale producer. But the North American colonies in the early colonial period were relatively devoid of high-demand exports (i.e., goods that Europeans were willing to pay a high premium to obtain) and profits were constrained by high transportation costs and mercantilist trade policies.

In short, the principle source of wealth (i.e., land), was broadly distributed and wages were relatively high; the American economy was too small to support a local, large-scale producer, so little wealth could be accumulated by producing for the domestic market; Americans struggled to find exports (like sugar and silver) that were valuable enough in Europe to overcome the political, technological and organizational constraints to penetrating European markets. As a result, wealth inequality across free households during this period was relatively low and fluctuating with the price of exports.

I begin this chapter with a brief discussion of contemporary perspectives and interpretations of inequality in the colonies and early United States before turning to the empirical evidence on wealth inequality in the colonies. I then discuss the transformative effect of the unique physical environment of the Americas on imported European institutions and traditions; because the functional distance from Europe created a sense of isolation, the process of institutional adaptation is often treated locally, but I emphasize that functional distance is a component of the spatial configuration. Institutional change occurred at the intersection of space, broadly conceived, and the existing institutional arrangements. I then consider the impact of the various components of the colonial space on wealth and income concentration:
imported institutions, the frontier, mercantilism, and the intersection of technology, natural resources and the functional distance to markets. Ultimately, the inequality experience of the colonies, from trade to slavery, even their relative isolation, was a global phenomenon.

*The Mythos of Colonial Inequality*

Economic inequality in the region was a matter of great philosophical, moral, and political consequence in the decades after the Revolution. Political and economic democratization were theoretically linked. For example, James Madison in a letter to Thomas Jefferson drew a causal arrow from political freedom to economic egalitarianism, “I have no doubt but that the misery of the lower classes will be found to abate wherever the Government assumes a freer aspect, & the laws favor a subdivision of property” (1787, quoted in Soltow 1989: 12).

In that same year, John Adams also linked democratization with lower levels of inequality, but detailed a more pernicious mechanism:

Perhaps, at first, prejudice, habit, shame or fear, principle or religion, would restrain the poor from attacking the rich, and the idle from usurping on the industrious, but the time would not be long before courage and enterprise would come, and pretexts be invented by degrees, to countenance the majority in dividing all the property among them, or at least, in sharing it equally with its present possessors. Debts would be abolished first; taxes laid heavy on the rich, and not at all on the others; and at last a downright equal division of every thing be demanded, and voted. (John Adams 1787, quoted in Soltow 1989: 19).

Like Adams, and now writing after the French Revolution, Alexis de Tocqueville feared ‘equality in servitude’: "But one also finds in the human heart a depraved taste for equality, which impels the weak to want to bring the strong down to their level, and which reduces men to preferring equality in servitude to inequality in freedom”
(Tocqueville 1835: ch. 3). But he also saw the arrow of causation running in the other direction,

Among the new objects that attracted my attention during my stay in the United States, none struck my eye more vividly than the equality of conditions. I discovered without difficulty the enormous influence that this primary fact exerts on the course of society; it gives a certain direction to public spirit, a certain turn to the laws, new maxims to those who govern, and particular habits to the governed. (Tocqueville 1835: preface).

In the Federalist Papers, James Madison emphasized the risk presented by class conflict to political stability.

The most common and durable source of factions has been the various and unequal distribution of property. Those who hold and those who are without property have ever formed distinct interests in society. Those who are creditors, and those who are debtors, fall under a like discrimination. A landed interest, a manufacturing interest, a mercantile interest, a moneyed interest, with many lesser interests, grow up of necessity in civilized nations, and divide them into different classes, actuated by different sentiments and views. The regulation of these various and interfering interests forms the principal task of modern legislation, and involves the spirit of party and faction in the necessary and ordinary operations of the government. (Madison 1787: No. 10).

This was not an abstract interest in economic inequality for its own sake. The American Revolution was a project in bringing political governance in line with the existing material conditions of social life. Before the political revolution there was an economic revolution that transformed economic institutions and social relations and would eventually make British rule untenable. In this sense, inherently at issue was a comparison of economic inequality in the colonies to that in Europe, as it was intimately linked to independence and a new model of governance. Contemporary observers, both American and European, often noted greater equality among the colonists and early Americans than in Europe.
Lord Adam Gordon, 1764: The levelling principle here, everywhere operates strongly and takes the lead, and everybody has property here, and everybody knows it (quoted in Mereness 1916: 449, Williamson and Lindert 1980: 9).

Benjamin Franklin, 1772: I have lately made a Tour thro’ Ireland and Scotland. In those Countries a small Part of Society are Landlords, great Noblemen, and Gentlemen, extremely opulent, living in the highest Affluence and Magnificence: The Bulk of the People Tenants, extremly poor, living in the most sordid Wretchedness, in dirty Hovels of Mud and Straw, and cloathed only in rags. I thought often of the Happiness of New England, where every Man is Freeholder, has a Vote in publick Affairs, lives in a tidy, warm House (quoted in Soltow 1989: 10).

While not fundamentally disagreeing with the observation of general equality, the reaction of La Rochefoucauld to his trip in 1797 is far less romantic, “The inhabitant here is proprietor and cultivator; that he lives as he pleases, must be admitted; but in the most remote and uninhabited parts of America that I have visited, I have never seen a greater proportion of wretched habitations” (quoted in Soltow 1989: 13).

The early United States had a high rate of landownership, but land did not correlate with wealth as it did in Europe. For example, it required about 1/50th as much labor time to buy a parcel of land in America as it would to buy that same land in England in 1790 (Christensen 1981). Land was cheap on the frontier but it was also unimproved and distant from markets, so it did not generate the same revenues as land in Europe. As a result, despite high rates of landownership, life on the frontier, was decidedly nasty, brutish and short. “Travelers into the frontier regions were often appalled at the meager diet, squalid housing, and near-absence of creature comforts to which homesteaders subjected themselves while they sank most of their time and resources . . . in to the business of farm development” (Vickers 1996: 218). Conditions were less stark in older settlements, but the North American colonies were notably short on wealth at the top of the distribution.
Observers of American wealth and poverty were also selective (Soltow 1989). Jefferson admitted that beggars do present themselves in larger towns, but “these are usually foreigners” (Jefferson 1787: Query 14). Property was held widely among free men, but 17% of the US population was enslaved in 1800. More than 25% of South Carolina, Virginia, Georgia, Maryland and North Carolina (representing four of the country’s seven most populous states) were enslaved in that year (1800 Census, Historical Census Browser 2004). British coverture meant that married women’s legal rights (including property rights) were subsumed under those of her husband. In light of these categorical inequalities, to claim that “everybody has property here, and everybody knows it” reflects a myopic perspective on US wealth inequality.

At the time, the leading thesis for American exceptionalism was the transformative effect of the American frontier. Land was relatively cheap, as noted above, and many colonies employed headright grants to attract settlers, whereby land was given to those willing to cross the Atlantic. James Madison (1786) admitted that superior governance might not be the only factor in the abatement of the misery of the lower classes:

Our limited population has probably as large a share in producing this effect [comparative comfort of the mass of people in the United States] as the political advantages which distinguish us. A certain degree of misery seems inseparable from a high degree of populousness. If the lands in Europe which are now dedicated to the amusement of the idle rich, were parceled out among the idle poor, I readily conceive the happy revolution which would be experienced by a certain proportion of the latter. But still would there not remain a great proportion unrelieved? (Madison 1786)

A century later, Frederick Jackson Turner (1893/1996) formalized and extended the logic relating the American frontier to the ‘American character’ – democratic, aggressive, innovative, and egalitarian. The harsh, wild environment forced settlers to
explore new social forms and practices. To quote Turner’s own gratuitously romantic language:

In the settlement of America we have to observe how European life entered the continent, and how America modified and developed that life and reacted on Europe. . . The wilderness masters the colonist. It finds him a European in dress, industries, tools, modes of travel, and thought. It takes him from the railroad car and puts him in the birch canoe. It strips off the garments of civilization and arrays him in the hunting shirt and the moccasin. It puts him in the log cabin of the Cherokee and Iroquois and runs an Indian palisade around him. Before long he has gone to planting Indian corn and plowing with a sharp stick, he shouts the war cry and takes the scalp in orthodox Indian fashion. In short, at the frontier the environment is at first too strong for the man. He must accept the conditions which it furnishes, or perish, and so he fits himself into the Indian clearings and follows the Indian trails. Little by little he transforms the wilderness, but the outcome is not the old Europe . . . . The fact is, that here is a new product that is American. (Turner 1893/1996: ch. 1).

Drawing on the economic theories of Malthus and Ricardo, Kearl, Pope and Wimmer (1980) extend the Turner thesis to hypothesize that inequality was lower on the frontier, and that inequality increased with population density.

In this line of thought, the colonial historian Jackson T. Main (1976) argues that European class structure, with highly concentrated levels of wealth, was exported to the colonies, but these institutions proved impossible to maintain in the colonies.

Likewise, institutional economists Daron Acemoglu, Simon Johnson and James Robinson (2001; Acemoglu and Robinson 2012) differentiate the institutional models of the North American colonies from those in South America, and associate these differences with native population density and political structure and sources of exportable wealth (see also Engerman and Sokoloff 1997; 2002; Sokoloff and Engerman 2000).

Williamson and Lindert (1980) began their discussion of colonial wealth by claiming “comparative levels of European and American inequality have never been
seriously debated” (p. 10). But Lee Soltow’s (1989) study of the First Direct Tax of 1798 suggested that wealth inequality at the turn of the 19th century was as high as it was on the eve of the Civil War. As Peter Lindert (2000) notes, this would suggest that early American wealth inequality was also as high as in England at the same time. Others have noted that estimates of colonial wealth inequality, particularly from Jones (1977; 1980), severely underestimate experienced inequality by ignoring coverture, indentured servitude and slavery, and large estates owned by British nonresidents (Shammas 1993; Shanahan and Corell 1997).

The reality is that there was no single inequality regime through the colonial and Revolutionary periods. Inequality is best understood at the intersection of discontinuous political and economic institutions; as such, a study of economic inequality across the North American British colonies and United States under the Articles of Confederation should focus on the intersection of regional variations (e.g. diverse geographies) and shared experiences (e.g., British institutional histories) as they interact over time (e.g., westward expansion). In other words, economic inequality in the colonial inequality regimes was low in some places and high in others, was steady and rose over time, was defined by slavery in some regions and by the absence of slavery in others. The aggregate is theoretically and experientially meaningless even if it was mythologically relevant.

A Note on Data

Wealth data for this chapter are drawn almost exclusively from probate records, inventories of the property of the recently deceased. Through the 1960s and 1970s, researchers developed methods to adjust these values for the age distribution of the
living population (the estate multiplier; Lampman 1962; Jones 1977) and adjusting for underrepresented populations (Smith 1975; G Main 1977; 1974; JT Main 1965; 1976; Jones 1977). Audits of the estate multiplier method have “established unambiguously that adjusting for age distribution affects only the levels and not the trends in wealth inequality” (Williamson and Lindert 1980: 12). Comparisons of probates against local censuses have found that about 60% of potential wealth holders\(^8\) were probated, and coverage was positively correlated with age and higher for whites (G Main 1976; Smith 1975). But researchers have also found that both very rich and very poor populations are surprisingly well represented in probate (Williamson and Lindert 1980).

Unfortunately, there is no equivalent source of data on incomes. Williamson and Lindert (1980) argue that income and wealth inequality will be monotonically related given a few “innocuous” assumptions. Soltow (1989) derives an estimate of income based on home prices. But divergent trends in wealth and income inequality in the United States since 1980 (with the former stable for several decades while the latter drove upwards) suggest a simplifying assumption about the relationship between wealth and income inequality may not be in order.

The relationship between wealth and income inequality suffers two complications. The first is fairly straightforward: high incomes do not exclusively target the wealthiest. If we assume wealth reflects the distribution of past incomes (which

\(^8\) Children, married women, slaves and servants are generally treated as non-potential wealth holders. There are two methods for dealing with these populations. The first is to focus only on free adult men. The second is to assume zero wealth for non-wealth holders and to include them in the wealth distribution.
requires a constant savings rate and a very loose definition of income, see below), we may deduce that there is a lagged, linear relationship between wealth and income inequality. But those activities and actors generating high incomes change over time. Wealth and income inequality can move in diametrically opposite directions if non-wealthy individuals begin to earn very high incomes.

The second complication hinges on the definition of income. We can include up to four components in a definition of income: cash incomes (which may or may not include government transfers; e.g., wages, EITC), non-cash benefits (e.g., subsidized health insurance, WIC), more-liquid asset appreciation (e.g., rising stock prices), and less-liquid asset appreciation (e.g., rising home prices). The definition of income a researcher uses often depends on the source data. Research from surveys often define income in terms of cash income, those drawing on tax data can include capital gains. Only the more enterprising include the value of benefits or non-cash government transfers as these can be difficult to appraise (see CBO 2011).

Typically, only ideologically motivated research includes all components. For example, Richard Burkhauser, Cornell economist and adjunct scholar of the neoconservative American Enterprise Institute, has challenged the measure of rising income inequality in recent decades in the United States from standard sources, and recently used the rising costs of homes and medical insurance to argue that “comprehensive” income inequality is not rising (Armour, Burkhauser and Larrimore 2013). Including non-liquid asset appreciation as income is problematic because it does not track the capacity to consume. For example, a rising home price does not increase the utility of that home to the residents. Assuming home prices rose
proportionately across the region, the residents cannot sell the home to convert that new wealth into consumption because they will need to reinvest those gains in a now more expensive new home. The same is true of medical insurance, the cost of which has grown faster than the marginal benefits of that insurance on health. To cash in on these new sources of “income,” and individual would need to sacrifice housing and health. Modern finance offers some options for leveraging non-liquid assets, but colonial households would have found it nearly impossible to convert non-liquid asset appreciation into consumption.

So we are faced with two options. The first is to ignore convention and define income comprehensively. The second is to admit that we should not use wealth and wealth inequality as an indicator of incomes. Instead of depending on flimsy wage data and unreliable assumptions about the relationship between wealth and income, I choose instead to focus exclusively on wealth and wealth inequality through the colonial period. More wage data is available in the 19th century, and I will draw on that data in chapter 3, but the unfortunate reality is that quality income data is not available until 1913.

Probate records allow for a relatively reliable measure of wealth inequality within a region over time, but complications arise in cross-colony comparisons. First, counties and colonies had different rules for which assets were inventoried, with more gaps in coverage as you move south (Williamson and Lindert 1980). Second, coverage by researchers, to extract and analyze probate archives, follows the same trend, with greater coverage in the north, less moving south.
Against this provincial scattering of coverage, only Alice Hanson Jones (1977; 1980) offers a cross-colony sample. This 1774 benchmark is critical for cross-colony comparison, but also for “national” comparisons to the censuses of the mid-19th century. On the other hand, we must recognize that we are dealing with a sample of 919 individuals drawn from sources with inconsistent definitions of wealth, and we should not draw conclusions that extrapolate beyond the data. For the most part, even those critical of the 1774 benchmark do not attack Jones’ methodology, but how those results have been interpreted.

*Colonial Crisis, Adaptation and Institutional Path Dependency*

It is the general argument of this dissertation that interstate crisis associated with systemic cycles of accumulation creates space for spatial reconfiguration – economic and institutional change – with important consequences for within-country inequality. This leads to a periodization of inequality change. That being said, the initial regime change in the American colonies was not systemic but geographic. Regime change was not the product of a global spatial reconfiguration, but an adaptation of European institutions to a very non-European environment.

While the initial regime-defining crises of the colonial period were not synchronized with global systemic cycles of accumulation, global processes played a crucial role in defining trends through the 17th and 18th centuries. Most important was slavery, an essentially global institution. Slaves were kidnapped from one part of the world by men from another part of the world exercising superior technology and economic resources. They were transported to the North American colonies to produce goods that only made sense as exports either to the sugar colonies or Europe.
The experience of each colony varied substantially by geography, the composition and intentions of the colonizers, and timing. But they also shared some key features—e.g., isolation from the principal sites of wealth generation and accumulation, European institutional and political traditions, and low population density. Divergence between colonies was relatively small compared to the divergence between the North American colonies and those further south. Those societies organized around sugar or the exploitation of native populations experienced higher levels of inequality (Engerman and Sokoloff 2002).

Early colonial history is rife with examples of reformation in response to contradictions between the institutional models introduced in the colonies and the physical realities facing colonists. Economic struggles and political strife in response to weakly embedded institutions created space for change. The result was a new set of institutional arrangements that varied some from colony to colony and dramatically from what the original governing bodies had intended.

Virginia. The Virginia Company of London established the Jamestown Settlement in 1607. The early colonists were poorly prepared for the challenges they faced. The Jamestown site was selected because it was uninhabited; it was uninhabited because it was a malaria-ridden swamp. Despite that, natives attacked the company within two weeks of reaching Jamestown (then James Fort). They arrived too late in the season to plant. The first colonists lacked the skills to produce their own food. Instead, they depended on strong-arming nearby tribes. In John Smith’s famous ‘Rude Letter,’ he emphasized the need for appropriately skilled workers:

When you send againe I intreat you rather send but thirty Carpenters, husbandmen, gardiners, fisher men, blacksmiths, masons, and diggers up of
trees, roots, well provided; then a thousand of such as we have: for except wee be able both to lodge them, and feed them, the most will consume with want of necessaries before they can be made good for any thing. (Smith 1608).

Investors had unreasonable expectations of economic returns in the colonies based on others’ experiences in South America: “There was not talke, no hope, no worke, but dig gold, refine gold, load gold” (Anas Todkill, quoted in Acemoglu and Robinson 2012: 22). With the second shipment of supplies, the Company demanded that the colonists send commodities sufficient to pay for the voyage, a lump of gold, and evidence of a route to the South Sea. The colonists failed to meet this “good faith” request.

The functional and normative collapse of Jamestown came in the winter of 1609/1610. Known as the Starving Time, 80% of the colonists perished. Historical accounts and recent archeological evidence suggest that some turned to cannibalism:

. . . driven through unsufferable hunger unnaturallie to eat those thinges which nature most abhorred, the flesh and excrements of man, as well of our owne nation as of an Indian, digged by some out of his grave after he had laien buried three daies & wholly devoured him; others, envying the better state of bodie of any whome hunger had not yet so much wasted as there owne, lay weight and threatened to kill and eat them; one amonge the rest slue his wife as she slept in his bosome, cut her in peeces, powdered her & fedd uppon her till he had clean devoured all partes saveinge her heade . . . (Journals of the House of Burgesses of Virginia, Papers of the General Assembly 1623/24, McIlwain 1915: 29).

In 1613, communal farming was abandoned and plots of land were granted to settlers. Economic circumstances immediately began to improve. In 1619 the first representative assembly met, forerunner to the Virginia General Assembly. But the most important development for the future of Virginia came in the form of a plant.

John Rolfe and his family were heading to Jamestown aboard the Sea Venture, a purpose built ship to bring supplies to the colonists, in 1609 when the ship was caught
in a hurricane and was deliberately driven into Bermuda. While in Bermuda, Rolfe collected seeds of *Nicotiana tabacum*. He began exporting tobacco in 1612, Jamestown’s first successful export.

The introduction of tobacco as a cash crop in the region would dovetail with another development of the 1610s. The headright system, implemented to attract settlers to Virginia, granted a plot of land to anyone that paid their own way, or that of another, across the Atlantic. This encouraged well-to-do settlers to invest in indentured laborers, who would provide labor for a period of time and entitle the master to another plot of land. African workers were first imported as indentured labor in 1619. The legal rights of African indentured laborers were trimmed over the next decades; in 1661 slavery as a lifelong, hereditary legal status made its first appearance in Virginia statute (Hashaw 2007). Just under 40% of the residents of Virginia in 1790 were slaves (Historical Census Browser 2004).

Maryland. Maryland likewise adopted headright and indenture servitude to attract new settlers, but Leonard and Cecil Calvert (Lord Baltimore) hoped to govern those new settlers under feudalistic precepts. In 1638, four years after landing in Maryland, and in response to popular discontent, Leonard Calvert was forced to abandon this experiment for British common law, and the republican Assembly took over the right to initiate legislation.

Carolina. After an initial bid to establish a Carolina colony in the early 17th century failed, title was granted to eight Lords Proprietors in 1663. The most active of the eight, Anthony Ashley-Cooper (Lord Shaftesbury), with his secretary, John Locke, drafted the “Fundamental Constitutions of Carolina”. The Fundamental Constitutions
contained progressive interpretations of representative government and religious
tolerance. But it also sought to establish a tightly controlled, feudalistic society,
employing serfs and slaves, controlled by a landed, titled aristocracy. The
representative parliament was only allowed to debate those subjects pre-approved by
the Lords Proprietors.

The Lords Proprietors hoped to profit from land sales, but were forced to lower
the price when early sales were anemic. Northern Carolina was settled by migrants
from inland Virginia, who planted tobacco and used Virginian ports to export their
crops. Settlers in southern Carolina relied more heavily on slave labor to cultivate rice
and indigo, and exported their crops through Charleston (then Charles Town). The
Fundamental Constitutions proved unpopular, especially among those in the southern
half of Carolina, and was never ratified. Political unrest led ultimately to the split
between North Carolina and South Carolina and, in 1729, the two colonies were
reorganized as royal colonies. But the institutional impact of the Fundamental
Constitutions would endure. On one hand, South Carolina would continue to be a
highly structured, manorial society; in 1790, 43% of the population of South Carolina
was enslaved (more than half were enslaved in 1820) and more than a third of white
households owned slaves (Historical Census Browser 2004). On the other, there was a
tradition of relative religious freedom and male suffrage with low property
requirements.

Georgia. In addition to other lesser factors, the experience of Georgia was different
from Carolina because 1) it came into being much later (1733) and 2) the British
claim was much less secure; it bordered Spanish Florida, and the French did not cede its claim on Georgia for another 30 years.

James Oglethorpe, trustee of the Province of Georgia, followed in the same intellectual tradition as Shaftesbury and Locke. But he created the ‘Oglethorpe Plan’ 70 years after the Fundamental Constitutions of Carolina. While Locke sought a fair and sustainable distribution of land across a slave-owning aristocracy, Oglethorpe envisioned a slave-free, yeomen society built on a fair and sustainable distribution of land. The plan included a detailed script for settling Savannah and the distribution of towns across the Province. The Oglethorpe Plan was short lived. The Spanish invaded in 1742, and the Trustees turned control of the colony over to the crown in 1752. But the impact on Georgia, from the layout of Savannah to differences between it and South Carolina, endured.

The differences between mid-Atlantic and southern colonies are small in comparison with differences between them and colonies further north. Not only are Pennsylvania and Massachusetts geographically different than Virginia and South Carolina, the intentions of its settlers were different (Vickers 1996). While Oglethorpe sought to establish Georgia as a society of yeoman farmers, colonies further north came closer to that ideal; only South Carolina and Virginia had more slaves per capita than Georgia in 1790 (Historical Census Browser 2004).

Regional wealth inequalities in 1774 track more or less as we would expect (Jones 1984). Figure 2.1 tracks the distribution of wealth in each of three regions, the South, the middle colonies, and New England, using a Lorentz curve: individual wealth is ranked and then charted with cumulative population (from poorest to richest) on the
x-axis and cumulative wealth on the y-axis. The gap between this curve and the diagonal, denoting equality, is the Gini coefficient, a standard measure of inequality. Wealth inequality in the South was greatest, even before accounting for the slave population. Wealthy merchants in Salem and Boston drove up inequality in New England relative to the middle colonies, where wealth inequality was the lowest.

Figure 2.1. Lorentz Curve of Total Physical Wealth by Region, 1774

Institutional Path Dependency. To this point I have highlighted differences between colonies, but more important in the long run of US history is the average difference between these North American British colonies and other New World colonies. Some
key building blocks of an emerging US institutional environment were 1) the social structure and population density of the native population (Engerman and Sokoloff 2002), 2) the high land/labor ratio, 3) the influence of English institutional traditions (North 1990; Coatsworth 1993), 4) the composition of the immigrant populations (Vickers 1996) and 5) the economic viability of natural regional products in European and West Indian markets.

The native population in North America was not as dense or hierarchically structured as in Central and South America, and more easily removed by force or disease. British North America, therefore, was not built on top of an existing social structure, and was not built with heavy state military intervention. There were some early efforts to meet high labor demands with semi-free and slave native labor, but these were largely unsuccessful (Engerman and Sokoloff 2002).

Following the elimination of the native population, colonized North America was heavy in resources but light in labor. Given the distance between labor supplies and labor demand, mechanisms were created to bring the two together. Among the solutions used were indentured servitude (indentured in Europe or redemptioners), prisoners, and slavery. For those not categorically exploited, the high land to labor ratio was empowering, economically and politically. For example, Lord Baltimore’s failure to transplant the manorial system from England was typical of the North American colonies (Galenson 1996).

Aspects of a more democratic legal, political and economic environment were transplanted from England, for example, fee simple land tenure and representative
political bodies. Often, these forms re-emerged against the wishes of early proprietors and crown officials.

The institutional environment, which varied regionally, was heavily influenced by the joint outcome of the composition of the immigrant population and the economic viability of local exports. For example, New England attracted family groups and produced goods similar to those produced in England. Consequently, there was an emphasis on self-sufficient homesteading. On the other hand, the Carolinas were colonized by elites from Barbados. The cultivation of rice on slave plantations was encouraged both by the physical and institutional environment. But conditions elsewhere were more extreme: 43% of those in South Carolina were slaves in 1790 versus 94% in Haiti in 1789 (Historical Census Browser 2004). The net result was relatively low levels of wealth inequality in the North American colonies versus colonies elsewhere (Engerman and Sokoloff 2002; Acemoglu, Johnson and Robinson 2001).

The bifurcated legal and economic structures of the North and South was a defining principle of US history through the 1960s. Income inequality is still notably higher on average in Southern counties relatively to the rest of the United States, and intergenerational mobility significantly lower in those same countries (Chetty et al 2014). These differences can be traced back to the introduction of chattel slavery, which can again be traced back to the different economic opportunities available in the North and South. But it is myopic to then conclude that the different inequality regimes of the North and South were simply a product of local economics. African slavery in North American was an innovative response (with very perverse
consequences) to global social hierarchies, networks, and ideologies, geographies, technologies and infrastructure, and global supply and demand.

The American Frontier

The frontier thesis has two components. First, after rejecting native claims on land and removing native populations through disease and violence, European settlers had more land than people to work it. If we recombine the modern-day states of Virginia and West Virginia, this territory is three quarters the size of Great Britain; Carolina (North Carolina plus South Carolina) is 94% the size of Great Britain; the territory of the 13 original colonies is 61% that of Great Britain, France, Spain, The Netherlands and Denmark combined, but in 1790 the booming population of the United States was about half that of Great Britain (8 million to 4 million). The high land to labor ratio increased the value of labor relative to land, and the relative scarcity of labor reduced the marginal benefit of additional land beyond what the owner could work.

The second component, emphasized by Turner (1893/1996), is that living conditions on the American frontier (which stretched to the Atlantic Coast in the 17th century) forced social/cultural/political adaptations. “[F]reshness, and confidence, and scorn of older society, impatience of its restraints and its ideas, and indifference to its lessons, have accompanied the frontier” (Turner 1893/1996: ch. 1).

The impact of these two components is evident in the first decades of each new colony. From John Smith requesting more “diggers up of trees” to Shaftesbury and Locke being forced to reduce the price of land in Carolina to attract settlers, the relative value of labor (peasantry) to land was apparent. And Lords Baltimore and Shaftesbury sought to reestablish a landed aristocracy in the New World but failed...
(though this probably had less to do with men abandoning the “garments of civilization” for moccasins than Turner claims).

From this, many have theorized that the American frontier reduced economic inequalities (e.g., Turner 1893/1996; Main 1976; Kearl, Pope and Wimmer 1980). Specifically, newer settlements had lower levels of inequality as cheap land was distributed fairly evenly across the first entrants. Inequality then rose as later arrivals took ownership of smaller plots of less productive land or were landless. The shifting frontier also offered a safety valve for older settlements, whereby poorer colonists moved west, levelling wealth distributions in the east while also populating those newer, lower inequality settlements further west (Kearl, Pope and Wimmer 1980).

The evidence supports this hypothesis only after admitting strong qualifications (Curti 1959; Kearl, Pope and Wimmer 1980; Steckel 1990). Kearl, Pope and Wimmer (1980) found that time of arrival was the best indicator of wealth in Utah between 1850 and 1870 (Schaefer 1987 produced similar results for Arkansas and Texas). The first settlers grab the best land and entrench themselves economically, politically and socially in the region. Inequality rises over time as immigration drives up land prices (George 1881). But Curti (1959) found similar wealth distributions in Trempealeau County, Wisconsin as in comparable towns in Vermont in that same period; Trempealeau County first appears in the US Census in 1860 with 2,560 inhabitants (Historical Census Browser 2004). Bolton (1984) found surprisingly high levels of wealth inequality in Arkansas in 1840; Arkansas had only 14,000 inhabitants in 1820 and just under 100,000 in 1840 (Bolton 1984; Historical Census Browser 2004).
Inequality in newer settlements rose to match that in older settlements, and consistently did so within a decade or two.

The first issue with the frontier as safety valve hypothesis is that successive waves of settlers find fewer economic opportunities (Menard and Carr 1979; Kearl, Pope and Wimmer 1980; Menard, Harris and Carr 1974). For example, Menard (1973) found that 90% of indentured servants that arrived in Maryland in its first decade of settlement and stayed for at least 10 years became landowners, and a few of these former indentured servants were in time among the wealthiest planters in the colony. “One needed only a few simple tools, a few head cattle, and about fifty acres of land to set up as an independent planter [in the Chesapeake], all within reach of newly freed servants or free immigrants of modest means” (Menard 1996). Just a quarter of a century later, though, less than half would become landowners and none would become wealthy (Menard 1973; Carr and Menard 1979; Galenson 1996).

Second, the accessibility of the frontier is also overplayed. The ‘frontier as safety valve’ thesis depends on constant westward expansion, but westward expansion was not cheap or easy; it required a sizable capital investment for supplies, skills, and an enormous investment of labor (Vickers 1996). As settlers moved inland, especially into Appalachia, they were partially cut off from larger eastern cities and ports, which were necessary to get a return on their investment.

On the other hand, we can argue that standard measures of wealth inequality on the frontier are misleading. New settlements were populated by individuals and families that had just invested (or left behind) much of their existing wealth to migrate to the newer settlement (Steckel 1990; Schaefer 1987). These poor migrants
pushed up wealth inequality. But in their migration they purchased a valuable asset—
a new *location*. They were often rewarded by rapid upward mobility, which is
captured statistically by low rates of landlessness and high rates of land accumulation
by the landless in early American history (Steckel 1990; Soltow 1989).

On the whole, I argue that neither the environmental characteristics of the
American frontier nor the high resource to labor ratio in the colonies, had a
substantive effect on wealth distributions on the whole. The opportunities that
motivated homesteaders to push back the frontier also attracted the relatively young,
relatively poor migrants from Europe that drove up wealth inequality in colonial cities
(Williamson and Lindert 1980).

**The Global Context**

What if the New World (or North America specifically), had appeared off the
Iberian coast near the end of the 15th century? Migrants from Europe to the New
World could travel by foot. The relatively high land to labor ratio would be quickly
expunged. Slavery in North America would be much less appealing to landowners.
Commodity prices would converge, and colonists would have less incentive to focus
on value dense goods for export. European elites would more easily maintain political
control over new settlements through better communication and the more rapid
deployment of force. The politically transformative effect of the frontier would have
played out very differently. In short, while the geography of the American frontier
may have played a role, its impact on inequality can only be understood within the
global geography of the Atlantic economy.
The cost and danger of crossing the Atlantic kept the population density of Europeans in North America low, and a small population meant that little wealth could be amassed within the North American colonies alone. The United Kingdom and Europe had 20 and 150 times, respectively, the purchasing power of the territory that would become the United States in 1700 (those figures would fall to 1.2 and 4 by 1860; Maddison 2008). Real wealth, and real wealth inequality, was derived from trade with Europe.

But there were three general requirements for a colonist to profit from trade with Europe. First, they needed to produce a good that could not be more cheaply supplied in Europe; second, the good had to be value-dense; third, they could not be at the wrong end of monopolized trade. Unlike the silver mines and sugar plantations further south, the early North American colonies were relatively devoid of value-dense minerals and commodity crops. The few exceptions, rice from South Carolina, tobacco, furs, whale oil, etc., highlight the potential economic impact of those crops (more on this below). Conversely, John Law’s inability to extract profits from Louisiana highlights the challenges faced by export-oriented producers.

John Law and the Compagnie d'Occident. The disconnect between the territory that would become the United States and the contemporary centers of global wealth accumulation is highlighted in the brief history of the Compagnie d'Occident. The company and its Chief Director, John Law, are better known as a historical case study of financial misallocation from “irrational exuberance.” But juxtaposed against the Dutch East India Company (VOC), the rise and fall of the “Company of the West”
demonstrates the challenge of extracting profits from Louisiana before the rise of cotton.

The VOC was established in 1602 with a 21-year monopoly over Dutch colonial activities in Asia. The functional goal of the company was to merge existing efforts and raise new capital through stocks to form an enterprise large enough to manage internally the capital and labor to make long-term investments (i.e., enjoy economies of scale). It’s ‘initial public offering’ raised 6½ million guilders, almost eight times more than the British East India Company (Frentrop 2003).

Consistent with the logic guiding the formation of a large, monopolistic trading company with powers to wage war and negotiate treaties, establish colonies and prosecute, even execute, convicts, the early days of the VOC were oriented more towards establishing a larger foothold and greater influence in the region than generating profits through trade. VOC officer Jans Pieterszoon Coen was particularly infamous for his violent pursuit of market power: “We cannot make war without trade, nor trade without war” (quoted in Findlay and O’Rourke 2007: 208).

In time these efforts paid off financially.

The commercial profits of this aggressive strategy were substantial. By the 1650s the VOC had established an effective and highly lucrative monopoly on the export of cloves, mace and nutmeg . . . and was becoming a major conduit for India textile exports. . . [B]y the 1690s the number of ships [returning laden with Asian goods] was 156. . . Between 1700 and 1750 the tonnage of Dutch shipping sailing back around the Cape doubled. As late as 1760 it was still roughly three times the amount British shipping (Ferguson 2009: 136-7).

VOC shareholders were the principal beneficiaries. By 1650 the initial investors had received dividends worth 8 times their original investment, an annual rate of
return of 27%, and VOC stock increased in value 700% between 1602 and 1733 (De Vries and Van der Woude 1997; Ferguson 2009).

John Law was inspired by the innovation and financial success of the VOC. A predecessor to fellow Scot, Adam Smith, Law believed that money was only a means of exchange, and that it was the exchange itself that constituted real national wealth. He believed that the VOC’s and Amsterdam Exchange Bank’s monetary conservatism were limiting (Ferguson 2009). He believed he could maximize profits by combining regional monopoly powers with liberal stock issues and the power to print money.

In France he was given the opportunity to test his system. His Company of the West was granted a monopoly over commerce with Louisiana in 1717, as well as control over internal colonial affairs, for 25 years, and he sold shares in the company to raise 100 million livres of initial capital. In subsequent years Law extended his financial clout: the royal mint and tax collection, the China and East India companies, control over tobacco revenues. He financed these moves with ever higher-priced share offerings (Ferguson 2009).

Ostensibly, the displacement that justified higher share prices was the promise of future profits from Louisiana. That was why Law devoted so much effort to conjuring up rosy visions of the colony as a veritable Garden of Eden, inhabited by friendly savages, eager to furnish a cornucopia of exotic goods for shipment to France... In reality, the share price was supported by lending printed money to buyers (Ferguson 2009: 145).

In other words, Law was managing one of history’s largest pyramid schemes.

Within 4 years, John Law was persona non grata in France. The value of paper money from the royal mint collapsed along with the share price of the Mississippi Company (successor to the Company of the West), and food prices soared. Louisiana
failed to produce the profits necessary to sustain the pyramid, and not because Law
did not try:

[A] grand new city was established at the mouth of the Mississippi: New
Orleans... [A] few thousand impoverished Germans from the Rhineland,
Switzerland and Alsace were recruited to act as colonists. But what the
unfortunate immigrants encountered when they reached Louisiana was a
sweltering, insect-infested swamp. Within a year 80 per cent of them had died
of starvation or tropical diseases like yellow fever (Ferguson 2009: 145)

New Orleans in 1721 was described as “a place of one hundred wretched hovels in a
malarious wet thicket of willows and dwarf palmettos, infested by serpents and
alligators” (Pierre François Xavier de Charlevoix, quoted in Seidenberg and
Weissman 2012: 439). A year after this flattering description, a hurricane flattened
the city.

In reality, Law was ahead of his time: “Such visions, as we know, were not
wholly without foundation, but their realization lay far in the future” (Ferguson 2009:
145). The sugar industry first took root in the region at the end of the 18th century.
The Mississippi River became a major commercial route through the last two decades
of that century. Then the population of New Orleans doubled in the 1830s, foreign
exports doubled between 1831 and 1833, and in 1840 New Orleans was the wealthiest
and third-most populous city in the United States (Lewis 2003). The Louisiana case
was extreme but not unique. The United States would become the largest, richest
economy in the world, but it lacked the instant moneymakers of the Spice Islands,
Cerro de Potosí or Madeira.

That being said, wealth inequality in the colonies was consistently linked with the
relative success of exports. After the American Revolution, John Jacob Astor would
amass an incredible fortune beginning with the fur trade, and whaling would become
a lucrative, export-oriented market centered in Nantucket. Sugar in Louisiana and cotton across the South would also fund large estates in the first half of the 19th century. Before the Revolution, there were fewer successful exports. Nine of the 10 richest colonists in Jones’ (1980) sample of 919 probates in 1774 lived in Charleston, South Carolina, and Charleston is by far the wealthiest (and most unequal) “district” in her sample (which included Boston and Philadelphia). Wealth in the region was tied to exporting rice, which reached 81.5 million pounds in 1773 (Dethloff 1982).

Likewise, wealth inequality in Maryland fluctuated with tobacco prices, the region’s major export (Williamson and Lindert 1980). A levelling of the wealth distribution between 1640 and 1670 (Menard, Harris and Carr 1974) correlates with tobacco prices (Williamson and Lindert 1980). The pattern repeats itself at the beginning of the 18th century; “So capital gains and losses from changes in export demand seem to account for Maryland’s colonial wealth instability” (Williamson and Lindert 1980: 20). And rising wealth inequality in Boston between 1700 and 1730 is tied to asset appreciation in portable personal property (“slaves, servants, currency, bonds, mortgages, book debt, stock in trade, and ships”), not physical asset accumulation (Nash 1976, Williamson and Lindert 1980: 19). “Thus the ‘cycles’ in wealth concentration can be readily associated with Boston’s trade conditions” (Williamson and Lindert 1980: 19).

For colonists, the challenge of accumulating European-style wealth was further exacerbated by trade policies explicitly designed to funnel wealth, through a positive balance of trade, back to Europe.
Mercantilism. Mercantilism emerged in 16th century Europe as a tool to protect the interests of the state (and raise revenue for state functions) by controlling (and taxing) foreign trade to raise revenue and ensure a positive balance of trade. Mercantilist policies regulating trade were also used to create a class of very wealthy merchants (with privileged access to restricted trade) to whom the state could turn for financial assistance during crises (McCusker 1996; colonial charters were often used as compensation for backing the Crown as well). Within this paradigm, more extensive, regulated trade generates more wealth for the state and its merchant elite, feeding the drive for colonial expansion. North American colonization was largely bankrolled by merchants that would never travel to the colonies, but were looking to control a piece of the new trade.

For the colonists, mercantilism limited where to and from whom they could export or import certain goods. These policies were explicitly enumerated in the Navigation Acts in the 1660s. The 1661 amendment to the original Act of 1660 added a list of enumerated articles (sugar, tobacco, indigo, and cotton; rice joined the list in 1704) that must be exported directly to England from the colonies. The Act of 1663 required all goods heading for the colonies first pass through England. (Dethloff 1982). Measuring the impact of the Navigation Acts on economic development in the colonies, let alone wealth inequality, is challenging. The most cited estimate of the impact of the Acts on income, from Thomas (1965), estimates lost income from exports at about 1% of per capita income.

The acts gave the colonists favorable access to English markets, stimulated the rapid expansion of colonial marine industries, contributed to an increase in the supply of English manufactured goods and services, established a supportive legal framework which eased credit for colonial trade, and provided that trade enjoyed the protection of the Royal Navy. The shipping and shipbuilding industries particularly prospered. (Dethloff 1982)

Second, through a policy now known as “salutary neglect”, the most notorious features of the Navigation Acts were rarely enforced. Fearing war with France from the north, the British did not want to irk colonists (at least until the end of the French and Indian War). For example, South Carolinians shipped rice directly to Portugal and the West Indies even though this was explicitly prohibited (though 80% did go directly to England; Dethloff 1982).

In this environment, colonial merchants and supporting industries fared well in many cases. For example, increasing or higher-than-expected inequality in economies as diverse as Maryland (Land 1965; G. Main 1977), Philadelphia (Smith 1984), and New England (Jones 1984) can be attributed specifically to the emergence of a wealthy commercial elite (Main 1965). As far back as 1687, the wealthiest quarter of taxable Bostonians, with about two-thirds of the cities taxable wealth, was populated by merchants with assets throughout the British Empire (Henretta 1965).

In reference to rising wealth inequality in Maryland after 1710, Main (1977: 570) noted that a solution to the paradox of rising inequality in the face of declining proportions of immigrants and freed servants in the white population lies in the growth of great fortunes during these years despite the secular depression in the tobacco
market. Valuable clues to the origin of this wealth lie in the inventoried assets themselves, because one characteristic common to the majority of personal estates worth £1,000 sterling or more is the involvement of their owners in mercantile and financial activities.

Likewise, “[t]he increase in inequality in Philadelphia, the area of pre-Revolutionary America for which historians have measured the greatest widening of the gap between rich and poor, resulted from economic rather than demographic change, as the rewards of commercial development accrued primarily to the affluent groups that controlled the economic system” (Smith 1984: 645). The Cabots and Crowninshields amassed fortunes and were influential in the international trade of rum, slaves, opium, tea, Madeira wine, Valencia oranges and Málaga grapes working out of Salem, Massachusetts in the 18th century.

These examples, though, do not rule out the potential for a more unequal outcome in a free-trade counterfactual. For example, a focus on merchant wealth north of Virginia ignores the fact that the burden of mercantilist policies was greatest in the South (Ransom 1968). Mercantilism was designed to funnel profits to the crown and domestic (European) merchants. American merchants were prevented from exploiting their local comparative advantage and, like the early settlers on the frontier, British merchants were able to enter the new market first and had advantages—experience, networks, financing, and political support—that made it difficult for colonial merchants to compete directly.

In summary, we can draw a few general conclusions about wealth inequality in the colonies. First, the frontier was a source of mobility, but its impact on overall wealth inequality is probably overstated. Wealth, and wealth inequality, is generally linked to exports, and fluctuating inequality is often correlated with price cycles in
those export markets. But the bigger conclusion is that there was not just one
inequality regime, level or trend:

If one were to take 1690 or 1700 as a base, the wealth inequality series
reported . . . would suggest mixed trends, but, on average, a drift toward
greater wealth concentration for the seven or eight decades prior to the
Revolution. This characterization holds for rural Connecticut (but not for
Hartford County), for rural Massachusetts (but not for rural Suffolk County),
for Boston as well as Portsmouth (New Hampshire), and for Philadelphia as
well as nearby Chester County. It does not hold for Maryland, however, which
exhibits stability from the 1690s onward [but gently rising inequality from
1710]. New York City is another exception, since it had a stable wealth
distribution between 1695 and 1789. (Williamson and Lindert 1980: 16-17).

Interpreting Colonial Inequality

There are three key theses on wealth inequality and inequality trends in the
colonial period (Williamson and Lindert 1980).

-European class structure made its way across the Atlantic, with a highly
concentrated wealth distribution, but this model was unsustainable against the
American frontier, and the American Revolution destroyed any remains of the
high inequality European class structure (JT Main 1965; 1976; Turner
1893/1996).

The unviability of the European class structure in the colonies was evident. But by no
means were the arrangements that replaced it overwhelmingly egalitarian: slavery in
the south, indentured servitude in the middle Atlantic colonies. Inequality on the
frontier quickly rose to match that in eastern settlements. Williamson and Lindert
(1980) dismissed the notion that American inequality approached the levels seen in
Europe (see also Kulikoff 1971; Lindert 2000), and while that may be true (among
free households), there is little evidence of an egalitarian trend over this period.

Low inequality was the initial condition, but the supply of labor increased
relative to land, increasing the value of the latter relative to the former.
Formidable barriers to migration west limited the effectiveness of the frontier
‘safety valve’. The result was a tendency for convergence with European
There is ample evidence that mobility was less pronounced among later arrivals, but there is little evidence for a long-run trend towards rising regional inequality across the colonies. Instead, we see regional fluctuations with little secular, let alone universal, tendency (Williamson and Lindert 1980). One solution to this apparent paradox – closing opportunity but stable inequality – is that a burgeoning middle class and new demand for labor accompanied the maturation process of a settlement; stretching at the tails was offset by growth in the middle of the distribution.

Wealth inequality in the United States did not rise between the end of the 18th and end of the 19th centuries, which means that colonial wealth inequality was already high (Soltow 1989; Shammas 1993; Shanahan and Corell 1997; Henretta 1965; Lemon and Nash 1968).

This position is significant because it is well established that wealth inequality near the end of the 19th century in the United States was as high as any we have ever measured in American history. In other words, this third group argues that wealth inequality at the time of the American Revolution was already on par with levels in Europe.

The empirical support for this approach comes from two sources. The first is a study of the First Direct Tax survey of 1798 by Lee Soltow (1989). Soltow finds that wealth inequality in this survey, an assessment of the taxable assets across the United States, was as high as that measured in the 1860 census and far higher than measured by Jones (1980) for 1774 from probate records. The second source are post-hoc adjustments to existing probate sources to include excluded populations (Shanahan and Corell 1997).

Belief in a preindustrial period of equality in America rests largely on an accounting that ignores the wealth of colonial officials, British residents, and the legal institutions of coverture, indentured servitude, and slavery. When
these elements are taken into account, the wealth of the adult population is more concentrated in the top quintile of the population in colonial times than in the nineteenth and twentieth centuries. The top quintile held 95 percent of the wealth, while the other 80 percent of the population held only 5 percent (Shammas 1993: 427)

The first point is that a survey of holdings can yield very different results than a survey of households, because many of the wealthiest proprietors were actually living in England. As a result, A. H. Jones underestimated the wealth share of the richest 1% (Shammas 1993). On the other hand, British proprietors helped compress inequality among resident households as they claimed rents that would have otherwise accumulated in the colonies.

Second, “the history of wealth inequality over the entire period changes considerably if one measures inequality on the basis of the adult population rather than by household” Shammas 1993: 415). To demonstrate this, Shammas (1993) compared the distribution of wealth across the adult population in 1774 to that in 1870 (Soltow 1975). Shammas emphasizes the greater concentration of wealth in the top quintile in 1774 versus 1870 to argue that belief in an egalitarian pre-industrial America rests on ignoring British proprietors, slavery, indenture and coverture. The more reasonable interpretation is two part: 1) wealth was more concentrated in the top quintile in 1774 (95% to 92%) because the bottom three quintiles were legally prevented from owning property; 2) wealth was more concentrated at the very top of the distribution in 1870 than 1774, the wealth share of the top 1% increased 32% from 1774 to 1870, from 28% to 37%. (Shanahan and Correll 2000 employ a different technique for adjusting probates from 1774 and the wealth census of 1860, but
produce similar results – wealth became more concentrated at the top, but the share of wealth going to the bottom 60% increased from 1774 to 1860.)

Conclusion

And thus our interpretation of inequality in the colonial period depends on where we look to find our representative sample. Looking across free adult males, wealth was less concentrated at the very top of the distribution in 1774 than in the 19th or 20th centuries, but wealth was more concentrated in the top half or quintile of the distribution in 1774 than the second half of the 20th century (Williamson and Lindert 1980). If we look across all adults, categorical inequalities in 1774 mean that wealth was much more concentrated in the top quintile in 1774 than in the 20th century and concentration at the very top of the distribution in 1774 and the 1950s is on par, but wealth is less concentrated at the very top in both periods than in the 19th century (Shammas 1993).

And our interpretation depends on where we look geographically. The wealthiest 10% in Delaware, Pennsylvania, New Jersey and New York controlled less than 40% of the regions wealth, less than 50% in Connecticut, Massachusetts, Rhode Island, and New Hampshire, and just under 70% in Maryland, Virginia, North Carolina, South Carolina, and Georgia (Jones 1980). Likewise, there was variation in regional trends up to 1774, inequality rising in some places but not others.

These variations are more than empirical artifact. Instead, they reflect the unique intersection of institutional and physical environment. Where the United States today has regions marked by higher levels of inequality, and others by greater equality (and some of the variation between them can be traced back to the colonial period), it
makes empirical and theoretical sense to also treat the United States today as having a
single distribution. To theorize an inequality state and trend for the entire region in
the colonial period, a region that changes in size and composition over time, is a
meaningless exercise in aggregation. The impact, even presence, of slavery,
indentured servitude, coverture, mercantilist trade policies, transportation
technologies, export-oriented production, the frontier or headright land grants on
wealth inequality varied from place to place.
Chapter 3. Commercial and Corporate Inequality Regimes

Wage and wealth inequality rose through the 19th century and into the 20th. Through the first half of the century, American merchants exploited gaps in the frayed remnants of the mercantilist organization of the world-economy and profited from huge price gaps between the United States and Europe. I refer to this as the commercial inequality regime. Through the second half of the century, new transportation technologies and the British emphasis on free exchange eroded arbitrage profits in cross-Atlantic trade. But the combination of intensive (more efficient) and extensive (more inputs, e.g., land and workers) growth of the US economy allowed those in a position to take an early lead in key industries to amass extraordinary amounts of wealth. More precisely, the modern corporation exploited the more rapid movement of information, goods and capital to control markets, secure profits, and enrich their owners. I refer to this as the corporate inequality regime. The focus of this chapter is to describe how these developments are correlated with and causally linked to the British cycle of accumulation (see Figure 0.3).

I begin the chapter with a discussion of a few trends and developments (convergence, rising wage inequality, slavery and emancipation, and the military subordination of Native Americans) that underlie the economic space of the commercial and corporate inequality regimes: westward expansion, economic development, particularly in the North, and better, faster, more reliable movement between the American frontier and Europe. I then discuss the available data, and limitations of those data sources, on wealth concentration and wage inequality through the 19th century.
We then take a step back to the 18th century and the Dutch cycle of accumulation to the put the American Revolution and Napoleonic Wars, and their impact on wealth inequality in the United States, in context; American merchants were well-positioned to exploit war in Europe that marked the beginning of the British cycle of accumulation. Beyond the immediate impact of war, new transportation technologies, internal improvements in the United States, and the death of mercantilism in the Atlantic brought American merchants unprecedented access to both European markets and the American frontier. Wealth initially accumulated in the hands of those best positioned to move goods between previously isolated (or mostly isolated) markets, and wealth inequality increased.

I then turn our attention to the end of the commercial inequality regime and the beginning of the corporate inequality regime. Freer trade initially produced extraordinary profits for merchants, but competition increased and profit margins fell. Specifically, prices in Europe and the United States converged, so that the profit potential of buying goods in one market and selling in another collapsed. But as prices converged, the size of the US economy boomed, so that domestic capitalists no longer needed to export to Europe to become wealthy. Instead, the key to wealth accumulation in the second half of the 19th century was controlling access to markets. The modern corporation integrated new transportation and communication technologies with a new organizational form and mechanisms for accumulating capital that allowed a smaller number of individuals to grab ever larger shares of key (and exponentially expanding) industries. Wealth accumulated in these firms, and
specifically in the hands of their largest shareholders. Wealth concentration in the United States again increased.

*The 19th Century Transformation and Processes of Convergence*

The United States’ 19th century is a dynamic bridge between the isolated colonial period and the hegemonic 20th century. I note in chapter 2 that a key variable in the character of the (multiple) inequality regimes of the North American colonies was their functional distance from the core of the world-economy. On the other hand, the current cycle of accumulation centers on the United States (see chapters 4 and 5 and Figure 0.3 above). It was during the British cycle of accumulation that the US economy, and its position in the world-economy, was transformed. I posit this transformation is best understood in terms of convergence.

Convergence took two forms. First, the multiple, localized inequality regimes of the colonial period converged into a single, national US inequality regime. There continue to be important institutional variations across the United States, but in the 19th century the impact of slavery and emancipation on wealth inequality merits special attention.

The slave population in the United States increased from 700,000 to 4,000,000 between 1790 and 1860, slaves made up more than half of the populations of Mississippi and South Carolina in 1860 and almost half of the populations of Texas, Florida, Georgia, and Alabama, and the share of the South that was enslaved increased between 1790 and 1860 (Historical Census Browser 2004). The production of cotton (in bales) increased 1,300 times, from 3 to 3,841 over that same period (Historical Statistics 2006), and the share of US exports in cotton increased from
below 40% before 1830 to about 60% in the 1850s. In the last decade before the Civil War land, slave, and wealth ownership were relatively concentrated in the South: farmers with more than 1,000 acres owned 21% of the total improved acreage in the Cotton South (versus less than 1% through most of the Midwest), only 25% of Southern families owned slaves (Wright 1970), and 90-95% of Southern wealth was owned by slaveowners (Soltow 1989). All signs point to a large, growing, and highly concentrated economic sector, conditions that should produce concentrated wealth; conversely, emancipation should have resulted in a massive redistribution of wealth.

But the narrative of economic development in the North trumped that of the South. Despite huge gains in cotton production, the slave population as a share of the total US population fell from 18% to 13% between 1790 and 1860 (Historical Census Browser 2004). Cotton represented a greater share of exports in 1860 than in the past, but the ratio of exports to total GDP fell by two-thirds, from more than 15% in 1800 to about 5% in 1860. The South lost its colonial income-per-capita advantage over the North during the post-Revolution recession, and had not yet rebounded to pre-Revolution incomes as late as 1840; in the meantime, Northerners doubled their incomes in real terms during that period (Lindert and Williamson 2012). In short, wealth was still concentrated in the South (Williamson and Lindert 1980), but the relative size of the pie was smaller. Ironically, Southern cotton exports provided the foreign reserves the North needed to import capital goods from England (North 1961).

At the other end, emancipation could have reduced wealth inequalities across the United States by allowing former slaves to own wealth when they had been
categorically denied that opportunity in the past. Unfortunately, the material benefit of emancipation to slaves was less than one might have hoped for, and the redistribution of wealth shares from South to North through the Civil War far outweighed any effect of emancipation (Williamson and Lindert 1980). The impact of emancipation on wealth concentration across the United States was limited, in short, because Southern wealth was less important than it had been in the past. Even in 1860, the national size distribution of wealth is largely unaffected by how slaves are counted methodologically (Gallman 1969). On the other hand, the trend of concentrating wealth between 1774 and 1860 is more steep if we exclude slaves (as property and as population) because their impact on the economy as a whole was less substantial in 1860 than 1774 (Williamson and Lindert 1980).

Convergence is also manifest in two processes that shrunk the functional distance between the United States and the core of the world-economy. First, new technologies and trade policies removed trade barriers between the United States and Europe, so that by the end of the century, Americans and Europeans were effectively competing in a single market. The best evidence of this is price convergence; for example, prices in Chicago and Liverpool were very different at the beginning of the 19th century and similar at the beginning of the 20th (North 1955; Persson 2004). If a good was substantially more expensive in one location versus another, an entrepreneurial merchant would buy goods from the latter and sell them in the former.

The second process bringing the United States closer to the core of the world-economy was that the core of the world-economy moved in that direction. The purchasing power of Europe was 13 times greater than that of the United States in
1820, four times greater in 1870, and 1¼ times greater in 1913. The US economy would have been tied (with Spain) for the fifth largest economy of Europe in 1820, the largest by the end of the 1870s, and many times larger than any other by 1913 (Maddison 2008). Some of this growth was due to increased productivity – per capita purchasing power in the United State quadruped between 1820 and 1913 – but more important was extensive growth – the population of the United States increased almost 25 times over that period (Maddison 2008; Historical Census Browser 2004). I will argue in the conclusion that because convergence reset the rules of access to the sites of wealth accumulation, it was associated with wealth concentration.

Critical to the story of US economic development and convergence with Europe is westward expansion and the military subordination of Native Americans. The British opposition to westward expansion was a factor in the American Revolution, the politics of westward expansion (free or slave) was a critical source of tension between North and South before the Civil War, and Hitler expressly cited US Lebensraum as a precedent that he pursued in World War II (Baranowski 2011). The key process, then, of 19th century US economic and political development was westward expansion and development on territory that was involuntarily relinquished by its original inhabitants. It is beyond the scope of this project to attempt an empirical assessment of the impact of this process on the distribution of wealth across the entire continent, but it is important to note that westward expansion and the extensive development of the US economy was built on a massive, brutal transfer of wealth.
Data

Wealth and wage inequality in the United States increased through the 19th century, but the timing of that growth is hard to pin down due to imperfect comparability across data sources. In this next section I discuss the available sources, some of the particular complications in those sources, and offer an interpretation of inequality trends through the century. Specifically, wealth became more concentrated and the distribution of wages more unequal between 1790 and 1860, and again between 1860 and World War I.

Wealth. The complexity of the American economy increased through the 19th century, but data on the distribution of wealth in the 19th century are little better than what is available for the colonial period. After the samples from 1774 (Jones 1980) and 1798 (Soltow 1989), the next reliable benchmarks on the distribution of wealth are the censuses of 1850 (real estate only), 1860, and 1870 (Soltow 1975; Steckel 1990). King (1927), Gallman (1969) and Kopczuk and Saez 2004 have estimated the distribution of wealth in the United States from 1916 using tax records. Unfortunately, because all four major benchmarks—1774, 1798, mid-19th century and 1916—draw on different data sources—probates, survey, census, and tax returns—reliability and comparability between them have been a source of debate.

There are a number of regional samples, but they lack the necessary coverage to draw conclusions for the nation as a whole: Kulikoff 1971 (Boston, 1790 to 1845, tax records), Bolton 1982, 1998 (Arkansas 1800 to 1860), Sarson 2000 (Maryland 1800 and 1820), Smith 1973 (Hingham, Massachusetts, 1765 to 1880), Main 1976 (Massachusetts, 1829 to 1891); Newell 1980 (Butler Co., Ohio, 1803-1865), Soltow
1971 (Wisconsin, 1860 forward), and Steckel and Moehling 2001 (Massachusetts, 1820 to 1910). There is also an extensive literature on height and standard of living with implications for inequality between regions and classes (Fogel 1986; Steckel 1979; 1986; 1995; Komlos 1987; 1998; Hiermeyer 2010).

I note at the end of chapter 2 that wealth in the United States probably became both more concentrated and more broadly diffused through the first half of the 19th century: existing wealth trickled down to populations that had been excluded previously (e.g. indentured servants), but new wealth went disproportionately to those at the very top of the distribution. One important caveat here is the timing of rising wealth inequality. Williamson and Lindert (1980) argue that wealth inequality rose between 1820 and 1860 (based on a few regional samples from disproportionately urban populations). This position, though, contradicts Soltow’s (1989) 1798 results, which show relatively high inequality at the beginning of the 19th century. These latter results are frequently used to argue that colonial wealth inequality was already high, but from this grand historical perspective we often forget that much can, and did, change in the 25-year period between 1774 (the first “national” benchmark) and 1798. Specifically, I will argue that wealth became substantially more concentrated in the 1790s with the onset of the commercial inequality regime.

Williamson and Lindert (1980: 47) argue that wealth inequality after the Civil War remained at a high, uneven plateau, but admit that “the half-century between the 1870 census and the onset of modern estate tax returns [1919] . . . is an empirical Dark Age.” Using the 1890 census, Gallman (1969: 13) estimated that the (gross) wealth share of the wealthiest .031% doubled or tripled from 1850 to 1890, from
around 7.2-7.6% to 14.3-19.1%. Drawing on the same census, Holmes (1893; Williamson and Lindert 1980) estimated that wealthiest 1% controlled 25.8% of net worth. Williamson and Lindert (1980) note that this is only marginally below the figure produced by Lampman (1962) for 1922, but this conclusion may be misleading for two reasons. First, wealth inequality in 1922, after World War I, was significantly lower than a decade earlier; the income share of the richest 1% fell 20% between 1914 and 1920 (Piketty and Saez 2003), and King (1927) and Williamson and Lindert (1980) estimate that the wealth share of the wealthiest 1% also fell about 20% between 1912 and 1923. Second, Kopczuk and Saez (2004) put the wealth share of the top 1% in 1922 at 36%, 33% higher than Lampman (1962). While issues with data comparability (asset and sample coverage, gross versus net wealth, etc.) mean that the results are far from conclusive, there is plenty of room for wealth inequality to have risen throughout the period from the Civil War to 1913 (and again through the 1920s).

Wages. The transformation of the US economy entailed a redefinition of labor. The emancipation of slaves, the end of indentured servitude, the forced expulsion of native populations and mass immigration of Europeans changed the way labor was organized. The expansion of semi- and fully proletarianized labor powered industrialization in the United States, and also increased the importance of both wage inequalities and inequalities between labor and capital. The emphasis of this dissertation is on income and wealth concentration at the top of the distribution. Because wage and income data before 1913 do not capture trends at the top of the distribution, my empirical and historical analysis focuses on trends in wealth
inequality alone until the 20th century. That being said, a quick note on trends in skill premia through the 19th century is warranted, as this helps set the stage for an analysis on income inequality in the 20th century.

The story of wage inequality through the 19th century is complicated by limited data. We have well-established bookends: US wage inequality was low by European standards at the beginning of the 19th century, but high at the beginning of the 20th century (Phelps-Brown 1968; Habakkuk 1962; Rosenberg 1967; Williamson and Lindert 1980). Before the American Revolution, colonists frequently complained about British regulations against immigration—more immigrants meant more labor in a labor starved nation, and greater regional influence. Habakkuk (1962) noted the relative abundance of skilled labor in 1820. “English visitors a century earlier characterized America as a nation endowed with cheap skills and expensive ‘raw’ labor” (Williamson and Lindert 1980: 67). A British machinist was paid 105% more than a common laborer in 1825; an American machinist received only 50% more (Rosenberg 1967).

A century later conditions had reversed; “shortly before World War I, the premium on skilled labor was extraordinarily high in America” (Williamson and Lindert 1980: 67). Taussig (1927: 58-60; quoted in Atack and Passell 1994) found the “comparatively low rate of pay for the unskilled” prior to World War I “markedly peculiar.” In building trades, the skilled to unskilled wage ratio was 2.17 in the United States versus 1.54 in the United Kingdom in 1909 (Phelps-Brown 1968). And the relative dearth of skilled labor in the United States influenced technological change in that country versus Britain: “If skilled labor is a good substitute for
machinery in production, then cheap skilled labor in Britain may have made it unprofitable to adopt capital-intensive methods. In America, by contrast, the relative scarcity of skilled workers may have required the adoption of machines that could be operated by the relatively abundant unskilled workers. This would also be perfectly consistent with the existence of difference in productivity between the two countries” (Atack and Passell 1994: 205).

The timing of rising wage inequality is more contentious in the literature. Estimates on wage inequality trends in this period are informed by a number of wage series; researchers have collected advertised wages for jobs in a particular place and requiring a particular set of skills over time. The most important sources before 1890 are the *Report of Persons and Articles Hired* (Margo 1999; 1820 to 1860), and the *Aldrich and Weeks Reports* (Burgess 1920, Coehlo and Shepherd 1976; these series have been indexed by Williamson 1975, see figure 3.1). Unfortunately, these wage series hardly represent the full distribution of material well-being—e.g., rising urban skilled to unskilled wage ratios in the antebellum period may have signaled a significant wage trend, but antebellum skilled workers were hardly the period’s economic elite and urban residents were still a demographic minority, so the impact of this trend on the overall distribution of incomes or wealth is difficult to assess. Other series include Adams (1968, Philadelphia laborers), Rothenberg (1992, agricultural day laborers in Massachusetts), Layer (1955, textile workers), Zabler (1972; iron workers in Eastern Pennsylvania), and Lebergott (1964, various sources and occupations). From 1890 to the 1940 Census (the first to collect data on wage and salary income), researchers collected more representative wage data (e.g., Burgess
1920, Douglas 1926, Long 1960; these and more are included/discussed in Williamson and Lindert 1980, Goldin and Margo 1992, and Margo 1999), which helps to compensate for a dearth of other data sources.

As with wealth, Williamson and Lindert (1980) argued that rising wage inequality was concentrated in an antebellum surge (see Figure 3.1), but this position has been widely criticized on empirical and theoretical grounds (Grosse 1982; Margo and Villaflor 1987; Atack and Passell 1994: 540). In reality, while we can identify trends for specific worker groups—e.g., clerks in Massachusetts—there is insufficient coverage to draw meaningful national conclusions.

Figure 3.1 Urban Skilled to Unskilled Wage Ratio, 1817 to 1939

Kuznets (1955) directed attention to industrialization as a source of rising inequality during this period. According to Kuznets, mechanization of production in the United States created tiered wages across the United States, higher in manufacturing, lower in agriculture. When workers were concentrated in one sector or the other, inequality was relatively low; when workers were dispersed between the
two, inequality was high. The evidence suggests a more complex story involving immigration, capital intensity and firm size.

First, immigration tended to increase skill premia. Beginning with a wave of Irish immigrants after 1845, followed by a larger wave of southern and eastern Europeans, American laborers found that they were increasingly forced to compete not only with European goods but also Europeans themselves for work. Consistent with the Hecksher-Ohlin trade model, Williamson and Lindert (1980) found a strong empirical link between a growing labor force (natural growth and from immigration) and wage inequality (O’Rourke and Williamson 1999).

Also, workers that used new technologies (or the white collar workers that managed their use) benefitted from an institutional environment that lacked the tools to broadly distribute the literacy, numeracy and other skills required. Unlike husbandry, industrial skills are not learned by children through home-based production, and the existing institutions for distributing skills—e.g., guilds and trade unions—were ill-equipped for training an industrial labor force. Workers employed in the largest firms with the greatest capitalization also earned the highest wages (Margo 1999). It was not until the high school movement early in the 20th century that these skills were broadly diffused, and wage gaps immediately began to compress despite increased automation of production through World War II. The universal importance of skill-acquisition through education in this period is demonstrated by the large returns to education enjoyed by Iowa farmers (Goldin and Katz 1999). In short, the relationship is much more complex than the two sector model proposed by Kuznets;
new skills, new technologies, and larger firms worked their way into all industries, benefitting some workers at the expense of others.

I focus the rest of the chapter on wealth concentration. Empirically, wealth became more concentrated from the last decades of the 18th century through the first decades of the 20th century, but the mechanisms of accumulating wealth evolved through the century. I turn our attention next to the commercial inequality regime and wealth concentration through the first half of the 19th century.

**The Commercial Inequality Regime**

The empirical evidence suggests that wealth concentration in the United States rose between 1774 and 1860, and perhaps considerably between 1774 and 1798. This concentration marks the end of the colonial inequality regime and the beginning of the commercial inequality regime, and it is correlated with the end of the Dutch cycle of accumulation and the beginning of the British cycle.

The shift from Dutch cycle to British cycle of accumulation can be mechanically tied to wealth concentration in the United States through two related processes. First, American political independence and war in Europe at the end of the 18th century granted American merchants special access to European markets that had been closed previously; among these pioneers of the commercial inequality regime are the United States’ first dollar millionaires. The second process involved a fundamental shift in the relationship between the United States and the rest of the world-economy. I highlight three developments: the end of mercantilism, better transportation technologies, and US internal improvements. These developments allowed American
merchants to connect an expanding US economy with resource hungry European markets, and get very rich in the process.

In the next section, I detail the narrative that links the end of the Dutch cycle of accumulation to American political independence and the Napoleonic wars and, in turn, the first American dollar millionaires and the beginning of the commercial inequality regime. In the subsequent section we look beyond the epiphenomenon of re-exporters and smugglers to the roots of the commercial inequality regime in the British cycle of accumulation.

*The End of the Dutch Cycle of Accumulation and the Exploits of War*

In his 1728 *A Plan of the English Commerce*, Daniel Defoe argued that “the Dutch must be understood as they really are, the *Carriers of the World*, the middle Persons in Trade, the Factors and Brokers of *Europe*: That, as is said above, they *buy* to *sell* again, *take* in to *send* out; and the greatest Part of their vast Commerce consists in being supply’d from all Parts of the World, that they may supply all the World again” (Defoe 1728: 192, emphasis and spelling in original). Defoe’s economic insight must be critically received as he would die three years later while hiding from creditors, but he did offer an apt, often quoted, description of the Dutch mercantile empire that stood at the center of a world-economic material expansion.

But Defoe’s description was more accurate of his past than his present. The Dutch world-trading system depended on market access, and that access was being materially threatened by the broader adoption of mercantilist policies. “Down to 1720 countries such as Prussia, Russia, Sweden, and Denmark-Norway had lacked the means . . . to emulate the aggressive mercantilism of England and France. But in the
years around 1720 a heightened sense of competition among the northern powers, combined with the diffusion of new technology and skills . . . led to a dramatic change. Within the next two decades most of northern Europe was incorporated into a framework of systemic industrial mercantilist policy” (Israel 1989: 383-4; quoted in Arrighi 1994/2010: 144-5).

But as one door closed another opened. By restricting the exchange of goods, the northern European powers created trade imbalances and investment opportunities in newly isolated markets, and a greater incentive to fight for market access. Dutch money could reach markets even when Dutch goods were blocked. Before mid-century, leading Dutch business entities began moving capital instead of goods through their international networks (Arrighi 1994/2010: 146). “By the 1760s, all the states of Europe were queuing up in the offices of the Dutch money-lenders: the emperor, the elector of Saxony, the elector of Bavaria, the insistent king of Denmark, the king of Sweden, Catherine II of Russia, the king of France and even the city of Hamburg (although it was Amsterdam’s successful rival) and lastly, the American rebels” (Braudel 1984: 246-7; quoted in Arrighi 1994/2010: 146).

The transition from Dutch material to financial expansion is important not only as a precursor to the emergence of British hegemony, but Dutch finance also played an important role in the series of wars that drove the transition from the colonial inequality regimes to the commercial inequality regime in the United States. The first of these is the Seven Years War from 1754 to 1763. Great Britain sought to extend control over colonies and trade at the expense of France and Spain; Prussia sought to extend its influence in continental Europe. Great Britain and Prussia allied against a
French-Austrian-Spanish front. British success in the war set the stage for it to emerge as the new global hegemon over the next quarter century. Great Britain gained colonial territory in Africa and the West Indies in addition to Spanish Florida and the bulk of New France in North America. France lost colonial territory (e.g., New France and French Louisiana) and gained war debt.

Critical for our purposes is the link between the Seven Years War and the American Revolution. “The year 1763 marked a turning point in the British-colonial relationship”:

Britain had emerged victorious from the long war with France. But the war had left Britain with an enormous public debt—all the larger because Parliament had reimbursed over 40 percent (1.069 million “pounds”) of colonial government contributions to the war effort—and a growing conviction that the colonies must bear a greater share of the cost of maintaining the empire. . . . Therefore, the Crown imposed a series of new taxes and reformed colonial administrative practices to enforce new and existing taxes better in order to generate additional revenues (Atack and Passell 1994: 67-68).

The British had hesitated to impose strict rule on the colonies in the past (e.g., salutary neglect) because it feared French interference from the north. With the French effectively removed from North America, and war debts weighing on the treasury, the British sought to collect dues from the colonists. The colonists violently objected.

When war in the colonies did break out, the colonists turned immediately to the French for assistance. The French spent 1.8 billion livres in the Seven Years War, inflating the national debt to a value seven times greater than revenues. Confronted with these obligations, the French spent another 1.3 billion livres in support of the American colonists in their war for independence (Conway 1995). The fiscal burden
culminated in the financial crisis of 1787. The French Revolution and European wars that followed in the 1790s and early 19th century were crucial to wealth concentration in the United States at the beginning of the commercial inequality regime (a point I return to below).

The immediate impact of American independence on wealth inequality was more a matter of international politics than domestic institutional change. The colonies had long enjoyed relative self-rule by popular assembly under the British (an attempt to impose minimal taxes led to war). The Articles of Confederation recognized state-based sovereignty built around existing political bodies, but established a weak national alliance. The individual states largely lacked the status to pursue meaningful foreign, monetary or fiscal policy.

But after the Revolution the former colonies operated outside of the British Empire. American merchants were no longer legally constrained by British mercantilism, but some trade routes within the Empire that had been open to the colonists were now closed. And even where British and Americans might have competed on equal footing legally, British merchants benefitted from a substantial competitive advantage. The British were more experienced and drew on established trade and financial networks, had access to London insurance markets, enjoyed the protection of the British navy, and shipped goods in superior ships. By trouncing the Dutch in the Fourth Anglo-Dutch War (1780-1784), British naval power was dominant in the Atlantic.

And the British sought to extend their advantage over the United States. Lord Sheffield argued in 1783 that England’s “great national object is to raise as many
sailors and as much shipping as possible” and capture all Anglo-American trade in
British ships, and for some time these efforts were moderately successful (Sheffield
1783; quoted in Bruchey 1990: 114). That Sheffield’s pamphlet was deemed worthy
of response by Thomas Paine (1783) and Tench Coxe (1791), among others, no doubt
reflects its impact on American commerce: “Since the publication of this pamphlet in
England, the commerce of the United States to the West Indies, in American vessels,
has been prohibited; and all intercourse, except in British bottoms, the property of and
navigated by British subjects, cut off” (Paine 1783).

The impact on American commerce was tangible. “Trade appeared to decline in
the post-Revolutionary years, and shipping interests did complain bitterly about their
plight” (Atack and Passell 1994: 73). Official exports fell by half between the early
1770s and 1793 (Bruchey 1990); actual exports (including smuggled goods) grew
over that period but failed to keep pace with population growth (Atack and Passell
1994); and per capita income fell about 10% between 1774 and 1805 (Atack and
Passell 1994). Given that concentrated wealth in the colonial period was associated
with mercantile interests in the North and export-oriented plantations in the South, the
post-Revolutionary suppression of American exports undoubtedly compressed wealth
distributions through the 1780s.

Sheffield also noted that American shipping suffered because the individual states
lacked leverage in commercial negotiations with foreign powers (Sheffield 1783;
Bruchey 1990). Under these conditions, Alexander Hamilton (1787) noted that “there
is scarcely anything that can wound the pride or degrade the character of an
independent nation which we do not experience.” He highlighted constraints on trade
and international influence: “Are we entitled by nature and compact to a free participation in the navigation of the Mississippi? Spain excludes us from it. . . Is commerce of importance to national wealth? Ours is at the lowest point of declension. Is respectability in the eyes of foreign powers a safeguard against foreign encroachments? The imbecility of our government even forbids them to treat with us. Our ambassadors abroad are the mere pageants of mimic sovereignty.” (Hamilton 1787). On these grounds, Charles Beard argued in 1913 that federalism in the United States was driven by “merchants, money lenders, security holders, manufacturers and shippers, capitalists, and financiers” who sought to benefit from consistent, national institutions and better terms for trade and credit (Beard 1913: 17; quoted in Atack and Passell 1994: 75). McGuire and Ohsfeldt (1984) analyzed voting patterns and concluded that delegates with merchantile interests were more likely to vote for ratification than other delegates.

The new Constitution went into effect in 1789 and foreign trade was already accelerating by 1790. But growth in trade was more a result of war in Europe than the new legal framework (Atack and Passell 1994; Bruchey 1990). “By virtue of American neutrality during the European wars, American ships could trade with both sides. The demand for U.S. products increased substantially, and the demand for American shipping services increased yet more dramatically” (Atack and Passell 1994: 77). The French opened their ports to American ships in 1793. Foreign exports doubled from $19 million to $40 million between 1790 and 1800 (Atack and Passell 1994). This new independence was interrupted in 1807 and by the War of 1812, but
(relatively) free trade effectively replaced mercantilism by the end of the Napoleonic Wars (McCusker 1996).

More important than traditional exports, though, was the new business of re-exports. In an ironic reversal of pre-Revolution mercantile roles, Americans were able to exploit a legal loophole, a war-induced salutary neglect: British West Indian products were shipped to the United States, relabeled as American, and then shipped to Britain (or elsewhere in Europe) without being unloaded or facing much risk of confiscation by the French or British Navies (Bruchey 1990). The dollar value of re-exports grew from nothing in 1790 to $45 million in 1800, more than ordinary exports (Atack and Passell 1994).

The power of this embryonic commercialization to create wealth (and inequality) is best captured by the fact that America’s first (literal) millionaires were merchants at the end of the 18th century, and much of their wealth came from re-exports (Bruchey 1990). In fact, the predominant feature of the who’s-who of America’s wealthy at the turn of the century was participation in trading practices of dubious legality. Before the Revolution, the Cabots made their fortune operating a fleet of privateers running opium, rum and slaves. Elias Hasket Derby profited immensely from the American Revolution as he equipped or owned shares in dozens of privateering ships. He grew his wealth in the 1790s and was probably the wealthiest American in 1799. Stephen Girard is also reputed to have profited from the American Revolution by trading with the British, but it was through neutral shipping in the midst of war in Europe that Girard became one of the wealthiest individuals, in relative terms, in US history (Bruchey 1990). This rapid accumulation of new wealth
at the end of the post-independence recession may help to explain the incongruous results from Soltow’s (1989) First Direct Tax of 1798 to A.H. Jones’ (1980) sample from 1774.

**Commercial Wealth beyond the Dubiously Legal**

The re-exporters and smugglers demonstrate the importance of access to consumers, especially when that access is limited. But just as important through the rest of the first half of the 18th century was access to raw materials within the United States. Those who entered those markets first, often employing shady business strategies to do so, and connected them with distant buyers often became very wealthy.

John Jacob Astor, then, is emblematic of the commercial inequality regime. Astor did not shy from smuggling opium, but his major industry was furs. Astor exploited the Jay Treaty in 1794, which opened trade with Canada, to build a fur empire. By 1800, Astor was worth about $250,000, well short of Girard and Derby, but still a wealthy man for the period. Where Astor differed from Girard and Derby, though, is that his business was as much about expanding west as accessing markets to the east. In addition to a global fleet that traded regularly with China, Fort Astoria in 1811 became the first US community on the Pacific Coast, and Astor built a trading empire in the Great Lakes region. Astor was the wealthiest antebellum American; Forbes estimated that his peak net worth was equal to about $110 billion in 2006 dollars, far more than any American today (*Forbes* 2007).

I highlight three developments that defined the commercial inequality regime. The first, the end of mercantilism, requires that we again take a step back in time to
1776 and Adam Smith's *Wealth of Nations*. American merchants rushed to exploit the gaps opening in old mercantilist networks. To do so, they transported goods using new technologies and new infrastructure, the second and third developments of the commercial inequality regime.

Adam Smith and the End of Mercantilism. After mercantilism helped expedite the end of the Dutch cycle of accumulation, the British cycle of accumulation is best defined by a rapid push towards free trade imperialism. This evolution has its roots in wars the Dutch funded during their financial expansion in the second half of the 18th century. In the shadow of a generation of costly wars to defend mercantilist trade interests, and the relationship between these mounting costs and political instability in France, the economic and political foundations of mercantilism as a philosophy of political economy were shaken. Adam Smith provided the theoretical sledge hammer that was used to destroy the remaining supports.

Adam Smith opened the *Wealth of Nations* by noting “the nation will be better or worse supplied with all the necessaries and conveniences for which it has occasion” as production (GDP in modern parlance) “bears a greater or smaller proportion to the number of those who are to consume it” (Smith 1776/2004: XV). This premise is precisely in line with mercantilist logic, but he quickly strayed from the party line. Where mercantilists sought to bloat this proportion by discouraging the consumption of imports and otherwise manipulating the terms of trade, Smith emphasized instead “the skill, dexterity, and judgment with which its labour is generally applied” (Smith 1776/2004: XV).
How, then, does one increase the “skill, dexterity and judgment” with which labor is applied? “The greatest improvement in the productive power of labour, and the greater part of the skill, dexterity, and judgment with which it is any where directed, or applied, seem to have been the effects of the division of labour” (Smith 1776/2004: 3). In other words, productivity is increased through specialization. Specialization depends on exchange; without exchange we must all be autarkic and utterly without specialization.

Because exchange encourages specialization, specialization increases productivity, and productivity generates wealth, wealth is maximized through exchange: “As it is the power of exchanging that gives occasion to the division of labour, so the extent of this division must always be limited by the extent of that power, or, in other words, by the extent of the market” (Smith 1776/2004: 15). Smith’s conclusion was revolutionary—national wealth was maximized by extending markets, not manipulating them.

That Adam Smith began with a mercantilist premise but a classical (economic) conclusion reflects more than original thought. Mercantilism sought to maximize the specie wealth of the crown and its loyal supporters with a positive balance of trade. This wealth could then be deployed to extend royal influence, soft and hard power. Smith’s logic reflected an economic democratization, the modernization of the economy and warfare (and thus classical economics stands as the first modern school of economic thought).
In a classic dialectical turn, this transition finds its roots in mercantilism itself. Mercantilism required an internalization of war- and state-making capacity in order to profit from the territorial expansion mercantilism encouraged.

Partly through commands to state bureaucracies and partly through incentives to private enterprise, the rulers of France and of the United Kingdom internalized within their domains as many of the growing number of activities that, directly or indirectly, entered as inputs in the war-making and state-making process as was feasible. In this way they managed to turn into tax revenues a much larger share of protections costs . . . By spending these enhanced tax revenues within their domestic economies, they created new incentives and opportunities to establish ever new linkages between activities and thus make wars pay for themselves more and more. . . . [A]n increasing number of civilians were mobilized to sustain indirectly, and often unknowingly, the war-making and state-making efforts of rulers. War-making and state-making were becoming an increasingly roundabout business which involved an ever-growing number, range, and variety of seemingly unrelated activities. (Arrighi 1994/2010: 51)

Adam Smith, then, anticipated the “wave of rebelliousness,” beginning with the American Revolution, that brought about “a thorough transformation of ruler-subject relations” and “the establishment of an entirely new kind of world hegemony (British free-trade imperialism)” (Arrighi 1994/2010: 53). In short, the end of mercantilism was important not only for what it meant in terms of foreign access to domestic ports, but it marked a revolution in the relationship between the state and the economy; we will see this pattern replicated at the end of the British cycle of accumulation and the beginning of the US cycle.

The geography of classical economic thought is also no accident. The British rode mercantilist logic to a position of preeminence in the Atlantic, but from that position imperialism made sense only as a means to open closed markets, not to close open markets. Thus, the Scottish Adam Smith argued that political economy should be oriented towards expanding markets. And the British sat at the top of the economic
food chain, and so it was apparent to the British economist David Ricardo that it should specialize and exploit its comparative advantage – thus relegating the periphery to providing British with the raw materials to feed industry.

Adam Smith was not the first to critique mercantilist thought, and the change in policy towards the colonies after the Seven Years War reflected a material change in attitudes in the British Empire regarding its relationship with its colonies. The British began a project of unilateral trade liberalization in the 19th century (McCusker 1996), and the repeal of the Navigation Acts and Corn Laws in the late 1840s marked the de jure end of mercantilism.

Transportation Technologies. After 1790, Britain increased public expenditures six fold, from 22 million pounds to 123 million pounds, to meet the demands of war. This spending boom fueled a rapid expansion of the capital goods industry, particularly iron. Two decades later war ended but the productive capacity remained; repurposing these capital goods fueled a transportation revolution: iron ships and iron train cars on iron railways powered by steam engines (Arrighi 1994/2010; McNeill 1984).

“Combined with the contemporaneous spread of mechanization within the textile industry, these innovations transformed the British capital goods industry into an autonomous and powerful engine of capitalist expansion” (Arrighi 1994/2010: 165). There are few better examples of how crisis gives birth to a new spatial configuration and material expansion of the world-economy.

The transportation revolution made British free-trade imperialism logistically possible. In order to pursue their Ricardian comparative advantage, the British needed to reduce the price of food to encourage a shift in employment from agriculture to
industry and reduce the relative cost of British industrial exports. Repealing the Corn Laws at mid-century was an important legislative step, but cheaper ocean transportation was also essential.

The transportation revolution was fundamental to the commercial inequality regime, but it was also a major player in the end of the commercial inequality regime. For example, clippers were competitive with steam-powered ships until mid-century (Atack and Passell 1994), but in the second half of the 19th century transportation costs collapsed. For example, Liverpool paid a 60% premium on grain relative to Chicago, 93% on meat and animal fats, and 80% on iron in 1870. In 1912, the price gap shrunk to 14% on grain, 18% on meat and animal fats, and 20% for iron (O’Rourke and Williamson 1992). Merchants depend on arbitrage for profits - buying a good at a lower price in one market and selling it at a higher price in another. Shrinking price gaps are antithetical to mercantile profits.

And the new steamships brought more than goods across the ocean. For example, 20% of New Yorkers were foreign immigrants in 1820, but 50% of the population and 80% of the wage-earning workforce were born outside the United States in 1850; 400,000 immigrants entered the city in the 1840s alone (Licht 1995). Immigration transformed labor in the United States and provided the muscle behind the American industrial revolution. Intensive and extensive growth of the US economy meant that capitalists in the second half of the 19th century could focus their attention on producing for a domestic market and depend less on exports. I will revisit these points when we turn our attention to the end of the commercial inequality regime.
Internal Improvements. The ocean-going transportation revolution required little in the way of infrastructural investment, but it also found an entrenched and stubborn technology in the form of the wind-powered clipper (Atack and Passell 1994). Land-based transportation, on the other hand, required a more substantial initial investment, but the potential gains, even in the first half of the 19th century, were immense.

Through the colonial period, wealth generation in the colonies was largely limited to coastal regions. As noted in chapter 2, though theoretically open to homesteaders, migration inland was difficult, dangerous, and expensive. “Penniless immigrants, even those from European farms, went mainly to the cities to do wage labor by necessity. Successful agriculture pioneering required extensive capital, including working family members” (Hughes and Cain 2002: 92). Massive transportation costs consumed any profits that could be earned through exchange, so farmers that were not effectively close to markets were limited to subsistence farming, which did not justify the investment to move west (Atack and Passell 1994).

Three variables, then, influenced westward expansion: the cost of migrating (including land purchases), the price of exports, and the cost of transporting those exports. For example, higher grain prices following a series of disappointing harvests in the 1830s induced a wave of settlers to go west through that decade (land entries exceeded 20 million acres in 1835 versus a previous annual peak of about 6 million in 1819; Robbins 1942; Atack and Passell 1994).

Lower transportation costs were brought about by a wave of internal improvements. Investment in canals took off after 1825, dipped a decade later, and then exploded from 1835 until 1842 (Goodrich 1961; Atack and Passell 1994).
Erie Canal opened in 1825 and was a huge success; by 1860 it was carrying almost as much trade as the Mississippi and Ohio Rivers (Atack and Passell 1994). Investment in canals waned as the popularity of railroads grew. There were 3,000 miles of rail in the United States in 1830, 9,000 miles in 1850 and 30,000 miles in 1860 (Historical Statistics 2006).

The result was a major reduction in transportation costs over land. Charles Paullin (1932) estimated travel times from New York City to points across the United States (see Figure 3.2). In 1800 (Paullin 1932: Plate 138A), Paullin estimated that it would take more than four weeks to reach New Orleans and six weeks to reach Chicago. Those times would be cut in half by 1830, three weeks to Chicago and two weeks to New Orleans (Plate 138B). In 1857 (Plate 138C), the travel time to New Orleans would be cut in half again to about six days, but the time to Chicago would be slashed to less than two days. By 1930 (Plate 138D), a traveler from New York could reach any point in the United States in four days or less by rail.

The case of New Orleans is illustrative of the commercial inequality regime. The Louisiana territory, once at the center of Law’s failed Mississippi Company, passed from French to Spanish and back to French hands before being sold to the United States when Napoleon abandoned plans for an American empire. Unfettered access to the Mississippi River (no longer would Hamilton need to worry about Spanish interference), combined with steamboats, gave Americans a direct route from frontier to ocean. New Orleans blossomed in subsequent decades, and was among the richest cities in the world in the 1830s.
Figure 3.2 Rates of Travel from New York City
Where New Orleans is illustrative of the commercial inequality regime, it is in the juxtaposition of Chicago and New Orleans that we can see the relationship between internal improvements and the transition to the corporate inequality regime in the United States. New Orleans was a commercial center, one of the largest and probably the richest city in the United States after 1830. But Chicago was the hub of a growing internal economy. Paullin (1932) produced another series of maps showing the extension of rail between 1840 and 1870 in the United States (Plates 138L and 140A; Figure 3.3). Rail connected much of the northern coast in 1840. In 1870, Chicago was at the center of a massive rail network that connected the Midwest and Northeast and, through the transcontinental railroad completed in 1869, a single route to the west coast. New Orleans and much of the South, by comparison, was relatively isolated.
After the Civil War, wealth in the United States was increasingly associated with internal, industrial growth, not raw exports (e.g., sugar and cotton) to Europe.

Figure 3.3 Railroad Density

Source: Paullin 1932

Returning to the impact of the commercial inequality regime on wealth concentration in the United States, expanding market access to the east and west allowed merchants to profit extraordinarily from arbitrage, purchasing goods in one market at a lower price in one place and selling them at a higher price elsewhere. Wealth also accumulated in commercial nodes (e.g., New York and New Orleans). Again, John Jacob Astor is a useful example; Astor was also a major investor in New York real estate.

Unfortunately, we cannot definitively compare these predictions against the empirical record. Wealth was more concentrated in 1860 than 1774, and more concentrated in 1798 than 1774. Soltow (1989) argues that inequality in the
distribution of real estate in 1860 was only marginally higher than in 1798, and this
gap can be explained by measurement error. Taken together, we can conclude that
wealth concentration either increased rapidly to 1798 from 1774 (and most likely
from 1790) and then plateaued until the Civil War, or wealth concentration increased
to 1798, and then portable wealth became significantly more concentrated to 1860. I
cannot differentiate empirically between the two scenarios, but the above narrative is
generally consistent with either, more so with the latter.

The End of the Commercial Regime

As is the case with each inequality regime, the conditions that defined the
previous regime would, in time, drive the transition to the next. Immigration, internal
improvements and westward expansion grew the US economy and reduced American
dependence on European markets. Falling transportation costs (both in terms of
time/risk and the initial capital investment) and greater access to international ports
increased competition and reduced the potential for arbitrage profits. And the
technology behind the British industrial revolution that absorbed raw materials from
around the world diffused. In other words, those conditions that produced America’s
first dollar millionaires were the same that eventually ate into their profit margins.

In this next section I identify the origins of the corporate inequality regime.
Specifically, Americans borrowed and adapted British technologies and employed
semi-proletarianized labor to launch their own industrial revolution. The politically
and economically more powerful North became more interested in protecting infant
industries than profiting from raw exports. New productive, transportation and
communication technologies encouraged the growth of large scale enterprise, and
growing global competition for raw materials prompted vertical integration in those enterprises to protect access to raw materials and other markets. Early entrepreneurs were able to capture industries within innovative organizational forms and ride US economic growth to accumulate massive amounts of wealth and power until the US government intervened in the 20th century to encourage competition between firms and manage competition between labor and capital. Wealth concentration was higher in the decades before the Great Depression than at any other point in US history.

Bringing British Technology to the United States. To protect their technological advantage in textile manufacturing, the British banned the export of that technology. And when the technology was stored in the heads of trained artisans, they forbade the embarkation of those individuals. To sneak past custom officials, Samuel Slater disguised himself as a farm laborer when emigrating from England for the United States (Licht 1995). The technology had to be adapted to the unique material conditions of the United States; for example, Samuel Slater contracted entire families to work in his factory because he was unable to find wage laborers, and early manufacturing labor was provided primarily by young women and children (Licht 1995). But in time Slater and others helped initiate a technological revolution in the United States.

More important for our purposes is where Samuel Slater turned to finance the implementation of the templates in his head. “In 1790 the textile mill of Almy and Brown opened in Pawtucket, Rhode Island, using British technology pirated by Samuel Slater. This venture was financed out of the merchant capital accumulated by the partners. . . . Most early manufacturing followed this pattern. . . . The Lowells of
Boston, for example, used their extensive mercantile fortune to establish the nation’s first industrial city on the falls of the Merrimack River—Lowell, Massachusetts.” (Atack and Passell 1994: 178). In fact, the primary backer of Lowell, Massachusetts was Samuel Cabot Lowell, a node joining the Cabot and Lowell families, among the two most important merchant families in Boston at the time (Atack and Passell 1994).

John Jacob Astor built his fortune on the fur trade, but he also profited immensely by investing that wealth in New York real estate. John Avery Parker used his shipbuilding resources to become a whaling kingpin and then gain a controlling interest in an iron mill (whaling provided the oils that lubricated the industrial revolution before petroleum; Atack and Passell 1994). The embargo of 1807 and increased competition in subsequent decades were particularly important in motivating merchants to diversify (Atack and Passell 1994). “The great mercantile families that had accumulated wealth in the colonial commerce placed their surplus capital in large-scale ventures: in Lowell mills, but also in further trade, banking, canal and railroad construction, and mining development” (Licht 1995: 35). As in the case of Samuel Slater, those to benefit from new American industry were often not members of the merchant elite, but they did often turn to that merchant elite, directly or indirectly, for financing.

The huge upfront costs on canals and railroads required new financial mechanisms and institutions (Atack and Passell 1994), and “the expansion of trade and the growth of business enterprise led (indeed compelled) some merchants to specialize in finance” (Klein 2007: 11). The New York Stock Exchange was organized in 1792 to meet new demands for capital investment, but financing was
largely available only from private investors until the end of the 19th century, and the banking system was far from stable during this period.

Manufacturing output increased at a rapid clip in the decades before the Civil War, faster in percentage terms than after the war. When there was an intersection of a growing immigrant, urban, less-skilled labor force, mercantile capital, and access to Midwestern raw materials, manufacturing expanded. “In 1820 agricultural workers accounted for 58 percent of the Massachusetts labor force. By 1840, when absolute agricultural employment in Massachusetts was at its peak, the share of labor in agriculture had fallen to 40 percent. In the next decade farm employment fell precipitously from 87,500 to 55,700, or only 15 percent of the labor force. This precipitous decline coincided with the opening of rail links to the West.” (Atack and Passell 1994: 178).

By combining the extensive growth of the US economy (from migration and westward expansion) with the intensive growth offered by industrialization, the US economy was transformed. Before the Civil War, real wealth in America was tied to demand in and access to European markets, because the potential purchasing power in Europe was much greater than in the United States. In 1820, France and the United Kingdom each produced three times more than the United States, and Germany produced twice as much. The United States passed France and Germany in terms of gross domestic product between 1850 and 1860, and it passed the United Kingdom by 1870. In 1870, the United States produced 40% as much as those three countries combined; that figure jumped to 85% in 1910. By 1920 US GDP was greater than that in the UK, Germany and France combined (Maddison 2008).
developmental strategy. The export-oriented South—slave plantations were a corporate adaptation to maximize profits from the European trade—preferred an export-oriented economic policy; cotton made up more than half of total exports between 1815 and 1860 (Klein 2007). The industrializing North preferred protection from British imports. Cotton textiles were particularly vulnerable to British competition (Atack and Passell 1994).

The trade economics of the Civil War can be succinctly summarized in Figure 3.4 below. James (1981) estimates the effect of a tariff on real incomes for various groups in 1859. James estimates that a tariff up to 40% of the price of imports would be positively correlated with incomes for laborers and capitalists, but negatively correlated with incomes for land and slaveowners. Others have produced different income elasticities, but the general conclusion stands: Northern manufacturers stood to benefit from protectionist policies while southern and western landowners did not.
Victory by the North accelerated the transformation of the US economy. Many Southerners believed “King Cotton” would force the British to come to their aid in the case of open conflict with the North, but their confidence backfired. Instead, the British intensified production in India. In the years before the war, two organizations formed in England "to encourage the increased cultivation of cotton in every part of the world suited for its growth" (Fifth Annual Report of the Cotton Supply Association, quoted in Logan 1958: 472). They identified India as a suitable home for investment, and the Indian share of British cotton imports grew from 31% in 1861 to 90% in 1862. India increased exports by more than 40% during the Civil War (Logan
Combined with slowing growth in the demand for cotton (Wright 1974), cotton after the Civil War was no longer king.

The North exploited its grip on federal power during and after the Civil War. In addition to abolishing slavery, a Northern Congress passed the Morrill Tariff and Revenue Act (the first attempt at a national income tax) and the Legal Tender Act (authorizing the Treasury to issue $150 million in greenbacks) in 1861, the Homestead Act (opening up more of the West to settlers), the Morrill Land Grant Act (creating agricultural colleges) and the Pacific Railway Act (providing federal funds for the transcontinental railroad) in 1862. The demands of war drove policy during this period, but another theme with a longer time frame emerged: The United States was adopting more aggressive, expansive federal fiscal and monetary policies. The net result was a fundamental transformation, economic and political, of the United States.

**The Corporate Inequality Regime**

At first blush, the story of the corporate inequality regime is simple: a small set of massive corporations rose up in the last half of the 19th century, and those at the top of the ladder became excessively rich. We are familiar with those at the apex—Carnegie (US Steel), Rockefeller (Standard Oil), Henry Ford (Ford Motor), Jay Gould (the archetypal ‘robber baron’) —but further down the ladder compensation was also higher for those in larger corporations (Margo 1999). To place the rise of the corporate inequality regime in context, I highlight four developments: the growth of the domestic market, infrastructural integration of the Union, cut-throat competition through Europe’s Long Depression, and the innovation of the corporate form.
Growth of the Domestic Market and Infrastructural Integration. The United States surpassed the United Kingdom in the early 1870s to become the largest economy in the world (Maddison 2008), but the maturation of the US economy involved more than extensive growth. According to Maddison, the per capita purchasing power of the United States fluctuated around 80% of that in the United Kingdom from 1820 until the end of the 1870s, but would match that of the United Kingdom in 1901 and was significantly higher (about 7%) by World War I (Maddison 2008; author’s calculations). The US economy was more productive and capital intensive (more efficient and using more machines) than its European peers by 1900 (Atack and Passell 1994). For example, as late as 1870 blast furnaces in the United States needed about 67% more fuel and 40% more workers to match Belgian furnaces in pig iron production. US furnaces would match those in Belgium in both respects in 1890, and were more efficient in 1900 (Allen 1977).

The extensive and intensive growth of the economy was facilitated by upgraded infrastructure. Canals and rail moved goods more quickly and cheaply than roads and unimproved rivers. But the revolution in information was more dramatic. Until the 1840s, news travelled along the same routes as everything else. By 1852, though, the United States had 23,000 miles of telegraph wire, 23,000 more than in 1844, and information could be moved almost instantaneously between major cities (Atack and Passell 1994). Markets expanded in the first half of the 19th century as barriers to exchange were lowered and transportation costs fell. But the combination of transportation and information technologies allowed for a different kind of expansion.

At the intersection of industrialization and improved infrastructure, manufacturing output increased and shifted westward. Table 3.1 lists counties by decade that produced over $100 million in manufacturing (in current dollars; prices were largely stagnant over the period in question; Historical Census Browser 2004; author’s calculations). New York and Philadelphia first make the list in 1860. Twenty years later in 1880, New York and Philadelphia still top the list, and Massachusetts had more representatives than any other state, but Chicago (Cook Co.) surged into third, and Cincinnati (Hamilton), Pittsburgh (Alleghany), and St. Louis also topped $100 million. In 1900, New York City continues to dominate US manufacturing – Manhattan (New York Co.) and Brooklyn (Kings Co.) were both in the top five, and Hudson Co., New Jersey is also on the list. But Chicago (Cook), Pittsburgh (Allegheny), and St. Louis are also among the top six. Of the counties producing more than $100 million in manufacturing in 1900, 48% were more than 300 miles from the East Coast. Consequently, in 1850, 49 of the top 100 counties by manufacturing output were in the Northeast, but that number fell over the next 50 years as the Midwest (1860 to 1880) and then the rest of the country (1880 to 1900) increased their share (see Figure 3.5).
### Table 3.1 Counties, Value of Manufacturing > $100 million (in $100 millions)

<table>
<thead>
<tr>
<th></th>
<th>1850</th>
<th>1890</th>
<th>1900</th>
</tr>
</thead>
<tbody>
<tr>
<td>(None)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1860</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW YORK (NY)</td>
<td>1.6</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>PHILADELPHIA (PA)</td>
<td>1.4</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>1870</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW YORK (NY)</td>
<td>3.3</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>PHILADELPHIA (PA)</td>
<td>3.2</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>ST LOUIS (MO)</td>
<td>1.6</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>MIDDLESEX (MA)</td>
<td>1.1</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>SUFFOLK (MA)</td>
<td>1.1</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>1880</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW YORK (NY)</td>
<td>4.7</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>PHILADELPHIA (PA)</td>
<td>3.2</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>COOK (IL)</td>
<td>2.5</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>KINGS (NY)</td>
<td>1.8</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>SUFFOLK (MA)</td>
<td>1.3</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>MIDDLESEX (MA)</td>
<td>1.3</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>ST LOUIS CITY (MO)</td>
<td>1.1</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>HAMILTON (OH)</td>
<td>1.1</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>ALLEGHENY (PA)</td>
<td>1.1</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>ESSEX (MA)</td>
<td>1.0</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

Source: Historical Census Browser 2004

### Figure 3.5 Number of Counties among Top 100 by Manufacturing Output, by Region

![Graph showing the number of counties among Top 100 by manufacturing output, by region.](image)

Source: Historical Census Browser
Europe’s Long Depression. Stagnant prices through the last quarter of the 19th century were a matter of great political concern. Europe in particular was beset by two decades of price deflation and low growth during the Long Depression (1873-1896). Popular opinion about the severity of the crisis did not match reality: "A depression of prices, a depression of interest, and a depression of profits; there is that undoubtedly. I cannot see any reason for believing that there is any considerable depression in any other respect" (Alfred Marshall, quoted in Musson 1959). This was particularly true of the United States, where strong economic growth resumed after 1879. But it was not immune to the deflationary pull exerted by Europe—price convergence was particularly strong during this period. The Civil War dramatizes the price effects, but prices for farm products, textiles, and metals were higher in the mid-1850s than in 1890 (Figure 3.6).

Figure 3.6 Warren and Pearson Index of Prices, 1850-1890 (1910-14=100)

![Graph showing the index of prices from 1830 to 1900 with different categories: Farm products, Textile products, Metals and metal product. The source is Warren and Pearson 1933.]
The Long Depression was central to Kondratieff’s (1925) model of major economic cycles, and Arrighi (1994/2010) associates the Long Depression in Europe with the collapse of the British-centered material expansion.

Generally speaking, our analysis of systemic cycles of accumulation has shown that every material expansion of the capitalist world-economy has been based on a particular organizational structure, the vitality of which was progressively undermined by the expansion. . . . as the mass of capital that sought reinvestment in trade increased under the impact of rising or high returns, a growing proportion of the economic space needed to keep returns rising or high was being used up (Arrighi 1994/2010: 232)

From this position, the Long Depression was a depression not in the sense that output flagged, because quite the opposite was the case, but in the sense that production expanded too fast to maintain profits. Specifically, rapid industrialization in Germany and the United States was leaving Britain behind.

There is other evidence of stagnation and inefficiency in several of the major industries, such as iron and steel, coal, and cotton. In iron and steel, for example, Britain was losing her technological lead; she was failing to modernize her plant, to develop new processes (like the Thomas basic process, for example), or to modify her industrial structure with the same rapidity as Germany and the United States—owing to conservatism, the heavy cost of replacing old plant, and deficiencies in technical education. . . . The cotton industry also experienced a declining growth both of production and productivity, and though Lancashire goods still dominated world markets, there was evidence of declining efficiency and of conservatism, as shown for example by the slow adoption of ring spinning and automatic looms in this country and by the very slight fall in real costs during this period as compared with the United States. Similarly, in the development of new industries, such as chemicals and electrical engineering, Britain was lagging behind Germany and the United States (Musson 1959: 206-7).

We can attribute British “conservatism,” their dependence on older, less efficient technologies, with the strictures of the existing spatial configuration. From the mid-1860s, British capital flowed internationally in hopes of finding higher returns (Williamson 1964; Arrighi 1994/2010). It did so because the United States
and Germany were in a better position to adopt newer, more efficient technologies (Atack and Passell 1994). The result, as we will see 100 years later in the United States, was the financialization of the British economy (Musson 1959).

As prices and profits in Europe fell, the economic fruits of exporting to Europe from the United States soured. “The discipline of the market posed a serious threat to the growing investment in increasingly specific capital goods and human capital. As a result, firms sought to maintain or increase profits and reduce risk by controlling price and output—that is, through monopolization. Increasingly, competition was viewed as ‘ruinous’ or ‘cutthroat.’” (Atack and Passell 1994: 481). American firms were in a unique position because the United States had the scale, both geographically and demographically, to specialize internally. Firms in the United States sought to exploit this advantage by internalizing and integrating production and distribution. When American firms did turn outward, it was to secure access to cheaper inputs abroad: direct investment from large US companies reached 7% of GNP in 1914 (Atack and Passell 1994).

The Corporate Form. The mechanization of the US economy exaggerated the impact of business cycles because the cost of fixed capital was less variable than labor; i.e., the firm could not layoff or cut wages to their machines (Prechel 2000). As such, there was little incentive to cut production when prices fell, and in some cases production would increase in response to lower prices in order to generate the same level of profits against slimmer profit margins.

Trade associations began to organize pooling agreements in the 1870s to establish minimum prices and divide markets (e.g., create regional monopolies), but these were
only artificial solutions. Instead of segmenting markets, firms sought to strengthen
ties to “centralize control over independent firms competing for the same market”
(Prechel 2000: 28). Firms turned to mergers and holding companies, but legal and
organizational complications with these arrangements led ultimately the creation of
trusts; small owners relinquished control of business activities in exchange for trust
certificates that guaranteed a share of profits (Prechel 2000).

The final step in the horizontal and vertical integration of production in the
modern corporation depended on the unique political geography of the United States.
First, in a capitalist world-economy defined by the discontinuity of economic and
political institutions (Wallerstein 1979), the United States enjoyed the advantage of
being sufficiently large in size, geographically and demographically, to internalize a
division of labor. In this environment, the United States was a leader in protectionist
economics and isolationist politics at a time when the economic orthodoxy of the
hegemon pushed regional specialization and exchange. Second, the US economy was
partially segmented between states. For example, firms were defined as foreign, with
no legal rights, outside their state of incorporation (Prechel 2000). But competition
between states for tax revenue, with New Jersey at the cutting edge, forced states to
liberalize incorporation laws in the 1890s.

The result was a proliferation of large, integrated trusts and holding companies.
“Almost nonexistent at the end of the 1870s, these integrated enterprises came to
dominate many of the most vital industries within less than three decades” (Chandler
1977: 285). Through horizontal integration (i.e., merging firms producing similar
goods), firms were able to control access to end-markets. Through vertical integration
(i.e., integrating supply and distribution chains), they secured access to inputs. In so doing, they also became themselves the leading sites of innovation having captured access to the abundant resources, sources of profit and the capital to fund development and implementation. Where the independent American inventor was a major player decades before, innovation became the realm of large research departments within or funded by the largest firms (Bruchey 1990).

These arrangements, in turn, were dependent on the faster transfer of goods, capital and information made possible by internal infrastructural improvements.

The integration of mass production with mass distribution allowed manufacturers to lower costs and increase productivity largely through the more effective management and coordination of production and distribution. The result was a new kind of firm—the modern business enterprise—characterized by many distinct operating units and managed by a hierarchy of salaried employees. In these firms the level of economic activity was such that administration coordination within the firm was more efficient and profitable than coordination through the market. The visible hand of the manager had replaced the invisible hand of the market. These firms are those whose names are familiar to this day: Pillsbury, Singer, International Harvester, Armour, Swift, Standard Oil, Remington, American Tobacco, and Diamond Match, to name but a very few. Big business had arrived. It brought with it lower costs, quality control, and dependability. Unfortunately, it was often the brainchild of “robber barons” who sought to reap the profits from monopoly. (Atack and Passell 1994: 472-73)

While the primary beneficiaries were those that owned the capital and managed the exchange networks, workers in large firms also enjoyed profits above those available to workers that were exposed to the full force of global competition.

This new organizational form gave American capitalists an advantage domestically and internationally, and played a critical role both in their ability to accumulate huge amounts of wealth and in the United States emerging as the global economic hegemon in subsequent decades.
The most imposing barrier to entry in these industries was the organization the pioneers had built to market and distribute their newly mass-produced products. A competitor who acquired the technology had to create a national and often global organization of managers, buyers, and salesman if he was to get the business away from the one or two enterprises that already stood astride the major marketing channels (Chandler 1977: 299, quoted in Arrighi 1994/2010)

The Bigger the Better

Bringing together the growing size and integration of the US economy, stiffening competition from Europe, and new organizational forms (and the legal and financial institutions to support them), firms in the United States had every incentive to grow large. This was not the case in the very recent past. “Most American industry in 1860 . . . was widely dispersed, rural, small-scale, and simple” (Atack and Passell 1994: 458). The reason for this is twofold. First, transportation costs made it impossible for larger firms to exploit economies of scale (centralizing production) since the raw inputs first had to be shipped in and then the output shipped out. “A number of authors have tried to measure the importance of scale economies using a variety of different production function forms. Most have concluded that scale economies, if and when they existed, were small and soon exhausted. Firms did not have to be large to be competitive in the first half of the 19th century” (Atack and Passell 1994: 193). Returns on investment were lower in larger firms (Atack and Passell 1994: 206). Falling transportation costs killed small firms that depended on distance from larger, more efficient competitors to survive. Through the second half of the 19th century, the theoretical benefits of economies of scale grew, so that James (1983) estimated that the ideal capacity of a flour mill in 1890 was infinite, i.e. a monopoly.
Second, even where the incentives were right, firm size was dictated by access to capital, and organizing large amounts of capital for a centralized production facility required a financial sophistication that did not yet exist. The evidence is in the exceptions. New York and Philadelphia were manufacturing hotbeds in the 1870s, but manufacturing establishments in Rhode Island and Massachusetts had twice as many workers on average as establishments in New York and Pennsylvania in 1870 (Historical Census Browser 2004; author’s calculations). The largest establishments in the Northeast were funded by wealthy New England merchant families (e.g., the Lowells). The New York Stock Exchange (1817), Chicago Board of Trade (1848), and more informal networks built around canals and railroads facilitated the financialization of the economy. Early holding companies and trusts were motivated by the need to accumulate capital (Prechel 2000).

As incentives changed, firm size grew. The average value produced annually by manufacturing establishments in the United States doubled between 1840 and 1890 (Historical Census Browser 2004; author’s calculations). The change was more dramatic in certain industries. Between 1869 and 1919, wage earners per establishment tripled in textiles, paper, and lumber, quintupled in furniture, stone, clay and glass, primary and fabricated metals, and machinery, and increased almost tenfold in petroleum and coal (O’Brien 1988). This last case is particularly relevant as Rockefeller’s Standard Oil controlled 90% of the oil industry in 1880 and two-thirds in 1911.
America's New Wealthy

Wealth concentration in the corporate inequality regime was historically high as it created a new class of super-wealthy Americans. At the beginning of the 19th century there were a handful of millionaires across the United States. Soltow’s (1989) survey identified 130 men among 1.7 million sampled with wealth exceeding $50,000 in 1798. The New York Sun counted 25 millionaires in 1845 (though this was most likely an undercount; Klein 1995). That number exploded to 4,047 in 1892 according to a survey by the New York Tribune. Of those, 82% did not inherit their wealth; Chicago was second only to New York City in the number of millionaires, and only 2% of those derived their wealth from an inheritance (New York Tribune 1892; Klein 1995). Then in 1918, at which point price inflation does become a factor, Forbes identified more than 200 Americans with an annual income over $1,000,000 (Forbes 1918).

The growth of millionaires and multimillionaires is not, alone, evidence of wealth concentration because there was more wealth to go around by the end of the century than at the beginning. But there is still evidence of wealth concentration when we control for the overall size of the economy. Klepper and Gunther (1996) produced a list of the 100 wealthiest Americans as a ratio of the national GNP at the time. When we chart these 100 individuals by birth year, the results are overwhelming (see figure 3.7). Of the 100, 71 were born in the 19th century and 22 between 1831 and 1841. Refining the filter a bit more, five of the 11 wealthiest Americans of all time (inflation-adjusted dollars) were born between 1834 and 1839 (John Rockefeller, Andrew Carnegie, Frederick Weyerhaeuser, Jay Gould, and Marshall Field; New York
These men began their business careers during the Civil War (Licht 1995) and were at the cutting edge of new wealth creating opportunities.

Figure 3.7 Distribution of America’s 100 Wealthiest by Year of Birth

![Chart showing distribution of America’s 100 wealthiest by year of birth from 1700 to 1950.]

Source: Klepper and Gunther 1996; Wealth as a ratio of national GNP

**Conclusion**

I included the commercial and corporate inequality regimes in a single chapter because there is value in their juxtaposition. First, the transformation from colonial outpost to emerging global power is critical to understanding the commercial and corporate inequality regimes, and also sets the stage for inequality trends in the 20th century. Rising wealth inequality in the first half of the 19th century reflected the disparities, in size and technological sophistication, between European and American markets. Merchants accumulated unprecedented levels of wealth by tapping European markets, exporting raw materials to feed European industry. Wealth concentration in
the second half of the century was built on domestic industrial empires, tapping into
the largest, most efficient economy in the world.

The inequality experience of the United States through the 19th century
complicates efforts to theorize a direct relationship between “globalization” and
within-country inequality. Globalization is a multifaceted process, and its impact on
the distribution of wages, incomes or wealth will vary across countries and across the
distribution within countries. For example, the variable linking liberalization of the
Atlantic economy in the first half of the 19th century with rising wealth concentration
in the United State was the wealth disparity between the United States and Europe.
But progressive trade liberalization through the century encouraged competition and
eliminated arbitrage profits. Then, in the second half of the century, the variable
linking globalization and wealth concentration was international economic
convergence (and the related institutional adaptations to deal with falling prices and
profits). At first glance we might note a positive correlation between trade
liberalization and rising inequality in both periods and draw some meaning from that,
but the relationship is obviously much more complex as the principle mechanism
(i.e., the value of European markets relative to US markets) linking trade and
inequality switched signs.

Instead of theorizing a direct relationship between globalization and inequality,
we need to identify the sites of wealth generation and accumulation, and constrained
access to them. The inequality regime of the first part of the 19th century in the United
States was defined by constrained access, through trade, to global sites of wealth
accumulation. It was the global constraints on exchange – technology, international
politics and networks, large capital investment – that drove up wealth inequality in the United States, not international trade \textit{per se}. Then, in the second half of the 19th century, a wave of globalization and “free trade” was marked by rising wealth concentration in the United States, but the relationship was again more complex. The United States was now a key site of wealth accumulation globally, and the modern corporation offered restricted access to that wealth through the vertical and horizontal integration of business activities.

In other words, the process of globalization is important to inequality and concentration trends because it redistributes market access. For some, globalization means greater access to new markets; for others, this means new competition in old markets. The Heckscher-Ohlin model tells us that inequality should be negatively correlated with trade where labor is relatively abundant, as those regions will specialize in labor-intensive exports, increase the demand for labor, and close the gap between labor and capital. The opposite is true where capital is abundant. But the story is more complicated because the capital, networks, and institutional support to engage in trade are not evenly distributed across a population; market access is also a property and it can be monopolized. For example, cotton was a labor-intensive activity, so its export to new markets should have increased the demand for, and compensation to, labor. Demand increased, but because the labor was enslaved, compensation did not.
Chapter 4. The Keynesian Inequality Regime

Income, wealth and wage inequality in the United States, and in other countries around the world, dropped markedly through the second quarter of the 20th century. The more important outcome is that income and wealth were still more broadly distributed two decades later. I argue that the Great Depression and World War II, terminal crises of the British cycle of accumulation, drove a spatial reconfiguration with an accompanying revolution in institutions, ideologies, technologies, and power structures, the backbone of the Keynesian inequality regime (see Figure 0.3). As a result, the benefits of economic growth after World War II were more broadly distributed than before the war.

Income and wealth inequality fell dramatically between the onset of the Great Depression and the end of World War II. The income share of the top 1% of earners in the United States fell from 19.6% to 10.5% between 1928 and 1944 (Piketty and Saez 2003). The share of total wealth owned by the wealthiest 1% fell from 36.5% to 25.5% over that same period (Kopczuk and Saez 2004; see also Kuznets 1953 and Lampman 1962).

Income and wealth compression between 1928 and 1944, in isolation, can be chalked up to current events, namely a financial crisis followed by an industrial, labor-intensive world war. The former devalues financial assets owned disproportionately by the rich, and the latter can destroy financial and physical assets and increases demand for labor. But two decades later, in the mid-1960s, income and wealth inequality remained at or below levels reported in the mid-1940s: “The more challenging part needing explanation is the nonrecovery of top capital incomes”

Trends in wage inequality suggest that the shock of the 1930s and 1940s masked a subtler transition that began earlier. After rising monotonically (though inconsistently) through the 19th century, wage gaps between more-skilled and less-skilled workers began to shrink around 1910 (Williamson 1975; Golden and Katz 1999). Growth in the supply of labor slowed and access to education spread (Williamson and Lindert 1980; Goldin and Katz 1999).

Presented with these results for the United States, Germany, and England, Simon Kuznets (1955) hypothesized that industrialization was at its core. He explicitly couched within-country inequality within a modernization framework by asking, “Does inequality in the distribution of income increase or decrease in the course of a country's economic growth?” (Kuznets 1955:1). He argued that inequality rose and then fell in England, Germany and the United States as these countries industrialized, and that this inverted-U of rising and falling inequality could be a universal tendency of industrializing economies.

What Kuznets did not discuss is that the experience of England, Germany and the United States was not unique. In the face of global economic and political crises, income concentrations fell in other core countries and their colonial offshoots between the end of the 1920s and the mid-1940s. In a sample of eleven countries, the income share of the top 1% fell by an average of 35% (Alvaredo, Atkinson, Piketty and Saez 2012). After the war, incomes did not re-concentrate in these countries.
As the ‘Great Compression’ was global, it needs a global explanation; a model that looks only at processes within countries (Kuznets inverted-U curve); even an explanation that allows for networked flows of goods, people and capital (e.g., the Heckscher-Ohlin trade model, see Lindert and Williamson 2001), cannot capture the coordinated compression of income distributions.

Piketty (2014) offers an alternative explanation for income compression in countries around the world in the 1940s. Piketty associates the size distribution of incomes with the ratio of the value of capital to the total national income in a country. Because capital tends to be more concentrated than labor incomes, as the ratio of capital to income in a country increases, the share of all incomes (including rents and capital gains) going to those at the top of the distribution increases. With the Great Depression and World War II, capital was physically destroyed by war, devalued by financial crisis, or liquidated by its owners to maintain their standard of living during a prolonged recession. Inequality stayed low over the next two decades as capital/income ratios were slow to recover against rapid labor productivity gains and high top marginal tax rates. The Great Depression and World War II are exceptional and external to Piketty’s central narrative of capital accumulation.

Developed around the experiences of Great Britain and France, this narrative has little explanatory power when applied to the United States. The capital/income ratio in the United States did fall only 28% between 1930 and 1950 (see Figure 4.1), while the income share of the richest 1% fell by more than half during that period. Assuming capital generated returns of 5% (borrowing here from Piketty’s own argument), the income share of returns to capital would have fallen from 26% to 18%
during the period. Even if all capital in the United States were owned by the richest 1%, the effect of a falling capital/income ratio on top income shares would fall well short of the figures Piketty himself helped produce (Piketty and Saez 2003). Looking ahead, a steadily increasing capital/income ratio in the second half of the twentieth century is also not enough to explain observed trends. This is especially true if we remove housing, the ownership of which is more broadly distributed than other forms of capital.

Figure 4.1 Capital Income Ratio, United States, 1770-2010

When we overlay the capital/income ratio in France on the income share of the top 1% of earners in that country, the fit is imperfect (see Figure 4.2). Specifically, the capital/income ratio collapsed between 1910 and 1920, but was fairly stable thereafter; on the other hand, the income share of the richest 1% in France fell by half between 1920 and 1950. The experience in Great Britain was similar.
The capital/income ratio is not enough to explain the compression of incomes in the United States, nor is it impressively correlated with top income shares in France. But even without a convincing bivariate relationship with income concentration, the capital/income ratio can still be a piece of the narrative. It is consistent with Arrighi’s cycles of accumulation and my proposed relationship with the concentration of incomes and wealth within countries that World War I, the Great Depression and World War II re-prioritized some capital forms (increasing the value of some while devaluing others) and caused capital to change hands within and between countries.

Where we fundamentally disagree, though, is that this was an exceptional event. It is my contention that this global great compression reflects a reconfiguration of the world-economy – a reorganization of the physical spaces and technologies of production, of inter- and intranational economic and political relations, and an ideological revolution that rationalized these new arrangements.
In the chaos and competition of the period, the entire continuum of theorized social arrangements came into play, and a wide variety were realized, from fascism to communism. In the aftermath of war, the victors were able to encourage the adoption of a new social contract (of which the welfare state was an important component) that had evolved in their countries. States implemented policies that emphasized domestic demand, financial insurance and regulation, and boosting state capacity to manage these new responsibilities. States invested directly in human capital development through education, and encouraged domestic investment in physical capital. Combined with the economic leveling of financial collapse and war, these new arrangements empowered labor, raised wages, and established a new inequality status quo in the global core.

The transition was, in part, an ideological reaction against the failed arrangements of the 19th century, but it also reflected changes in the material conditions of production and the global organization of economic life. As noted by Karl Marx in the mid-19th century and Simon Kuznets one hundred years later, industrialization and urbanization fundamentally changed the organization of economic, social and political life. For example, a new policy emphasis on full employment was discussed in terms of economic growth (i.e., maximizing production and demand), but there was also a growing awareness of the political threat represented by an underemployed proletariat.

Globally, the United States exploited its unique geopolitical position to emerge from war as the new global and financial hegemon. As such it exerted its ‘moral leadership’ to restructure the world-economy in its interests. This project included
hosting talks to create new rules for international finance and trade, but Washington also imposed its productive, financial and military clout to influence domestic institutions elsewhere. Innovative ‘New Deal’-style institutional arrangements replaced the gold standard, which had been the requirement for full participation in the British-centered world-economy before World War I.

I began the chapter by examining the evidence of a national and global great compression. In the next section I take a closer look at the contradiction and crisis that brought an end to the British systemic cycle of accumulation; specifically, a spatial configuration organized around British hegemony, free trade and a rigid gold standard could not accommodate emerging superpowers in the United States and Germany. I then turn to the ideological response to the Great Depression, which included a new appreciation for the interrelationship of political and economic stability and the state’s role in maximizing production and demand. Finally, I detail many of the mechanisms linking the Great Depression and World War II to lower income inequality in the United States.

The Data

Bourguignon and Morrisson (2000) offer the most complete empirical estimates of global inequality between and within countries historically. They estimate income shares for nine income decile shares and two vintiles for 33 country groups over 11 benchmark years from 1820 to 1992. The authors conclude that there was a “substantial decline in within-country inequality between 1910 and 1950” (pg. 734).

Figure 4.3 charts global inequality and the Theil decomposition of inequality into between- and within-country inequality based on their estimates. Between 1910 and
1950 within-country inequality falls almost 40%. Outside of this one historical moment, the authors estimate that global within-country inequality has been markedly stable over the last two centuries (see Korzeniewicz and Moran 2009 for a theoretical treatment of this point).

Figure 4.3 Decomposition of Global Inequality Between and Within Countries, 1820 to 1992

Source: Bourguignon and Morisson (2000)

When we shift our focus to top income shares, inequality falls faster and further. Top income shares are inherently more volatile: they reflect a smaller population and capital gains make up a larger share of top incomes and are also more volatile than wages. But because top income shares represent a unique population and have a different composition than other incomes, the rise and fall of top income shares can be theoretically distinct from other measures of inequality that consider the entire distribution of incomes (e.g., Gini coefficient).

Over the last decade, Thomas Piketty, Emmanuel Saez and colleagues have collected a databank of top income shares from tax returns for dozens of countries
(Alvaredo, Atkinson, Piketty and Saez 2012; see Atkinson, Piketty and Saez 2011 for a discussion of data reliability). Figure 4.4 charts the average income share of the top 1% of earners (excluding capital gains) for the 11 countries with at least one data point for 1926-1928 and 1944-1946. In all eleven cases, the income share of the top 1% fell. On average, the income share of this group fell 35%. While the fall was greatest in France and Japan, countries that bore the brunt of World War II, countries as diverse as Canada, India, and New Zealand also saw the income share of their richest 1% fall.

Figure 4.4 Income Share of the Top 1% (Excluding Capital Gains)

Source: Alvaredo, Atkinson, Piketty and Saez; Data for Sweden from 1930.

In the United States, the income share of the top 10% of earners fell from about 1/2 to 1/3 of total income during the period; the income share of the top 1% fell from just under 1/4 in 1928 to half that by 1944 and less than 10% in 1953 (Piketty and Saez 2003).
To put this in perspective, we can translate these points into a minimum Gini coefficient. If we are given only one data point, we can estimate a minimum Gini coefficient as $S - P$, where $S$ is the income share of the group and $P$ is the percentage of the population represented from the top. For example, the average income share of the top 1% in the United States between 1926 and 1928 was 18.8%. If this 18.8% was evenly distributed among that group and the remaining 81.2% was evenly distributed among the remaining 99%, the Gini coefficient (the minimum Gini) would be $0.188 - 0.01 = 0.178$. By 1944-1946 the top share fell to 11.1% for a minimum Gini of 0.101. The difference, $0.178 - 0.101 = 0.077$, is roughly equal to the increase in the Gini coefficient in the United States from 1967 to 2011 (Bee 2012), but in less than half as much time.

More revolutionary than the collapse of top income shares was the lack of rebound after the war (Piketty and Saez 2006). Figure 4.5 shows that the income share of the top 1% in nine of the 11 sampled countries was lower two decades later than it had been at the end of World War II. Only in Finland and France were the richest 1% grabbing a larger share of total income, and these shares were still substantially lower than in the late 1920s.
A closer look at the United States establishes two distinct points during which top income shares dropped sharply (see Figure 4.6). First, the stock market crash cut the income share of the top 1% of earners from 19.6% to 15.3% between 1928 and 1931. Between 1931 and 1940 this value fluctuated some but was fairly stable. Then between 1940 and 1944 the income share of the richest 1% fell from 15.7% to 10.5%.

Highlighting the global nature of the Great Compression, Figure 4.6 compares the de-concentration of incomes in France and the United States. The experience is similar in the two countries in terms of extent and timing; for example, if we fix the mean and standard deviation of year-on-year change for France and the United States but allow for the timing of change to vary, the probability that the correlation between the two lines (top income shares in the United States and France between 1925 and 1955) would exceed the measured value (.953) is 1.63%. This probability is small enough to reject the null hypothesis that the timing of change in the two countries is random.
relative to one another. In short, top income shares fell in both countries by the same amount and at the same time.

Figure 4.6 Income Share of Top 1%, France and the US

Source: Alvaredo, Atkinson, Piketty and Saez 2012

Using estate tax returns, Kopczuk and Saez (2004) estimate that the share of wealth owned by the wealthiest 1% fell from 40.3% to 28.4% between 1930 and 1932 and then steadily to 22.6% in 1949, just over half of what it had been two decades earlier.

The picture of economic inequality through the rest of the distribution is less certain. The 1940 census collected data on wages, and the Gini coefficient for family incomes can be estimated back to 1947 using the Current Population Survey. Researchers have collected wage series to track absolute and relative wage trends for various occupations (e.g., Burgess 1920, Douglas 1926, Long 1960; Margo 1999). Drawing on these sources, beginning from 1910, there is clear evidence in the United States that wages of less-skilled workers increased relative to gross domestic product during this period (Goldin and Margo 1992) and that wages of less-skilled workers
caught up with those of more skilled workers (Williamson and Lindert 1980). Goldin and Margo (1992) found no pre-1940 compression of white collar wages with unskilled labor.

Wage compression during World War II was more dramatic. “Education and skill premiums plummeted, geographical differences in wages were reduced, and the variance of wages within these groups fell substantially” (Goldin and Margo 1992: 2). The range between the logged wage at the 10th and 90th percentile in 1940 was 1.41; it fell to 1.06 in 1950 (Goldin and Margo 1992).

In an effort to estimate a more complete distribution of incomes, I combined top income share estimates with the unemployment rate. Using the income share of the top .01%, .1%, 1%, 5% and 10% and the unemployment rate, I estimated the Gini coefficient in family incomes reported by the Census Bureau (2011) from 1947 to 2008. These six variables explained more than 95% of the variance in the Gini coefficient between 1947 and 2008.

I then projected these results back to 1917 (see Figure 4.7). Projected family income inequality fell some in 1929 before unemployment overpowered the impact of the stock market losses. Inequality fell some through the 1930s as unemployment settled down, and then plummeted after 1940 as the war fueled full employment. Of course, the validity of this projection requires that the relationship between top income shares, unemployment and aggregate inequality holds over time, and it extrapolates beyond the data as unemployment at the height of the Great Depression approached 25%. That being said, the de-concentration of top incomes and
reemployment of almost a quarter of the labor force between 1932 and 1944 made for the most significant inequality event in the history of the United States.

Contradiction and Crisis

As with the previous chapters, I hypothesize that contradictions in the spatial configuration of the world-economy (principally defined here by British liberalism and the gold standard) led ultimately to interstate chaos, the destruction of physical and financial assets, a relocation of centers of power and production, and, as a result, space and momentum for institutional change. The three decades from 1914 to 1945 fit the bill perfectly: two world wars and a global Great Depression, mass destruction of the means of production (primarily in continental Europe), a relocation of the center of power from England to the United States, the emergence of new political and economic ideologies, and accelerated political and institutional change. A new world order emerged that favored capital, trade, and immigration controls, fiscal
policies to buoy demand and manage unemployment, more investment in human
capital, greater government intervention in the relationship between capital and labor,
and new progressive taxes to fund a more active state. As a result, within-country
inequality fell in countries around the world.

Unresolved Crises. In the United States between 1929 and 1933, GDP per capita fell
47.0% in nominal terms. Double digit deflation exaggerates the collapse; in constant
dollars, GDP per capita fell 28.6% (Williamson 2013; BLS 2013, CPI; author’s
calculations). The value of industrial products fell more than 50% between September
1928 and June 1932 (FRED 2013, INDPRO), as did prices on farm products (BLS
2013, WPU01). Unemployment peaked around 25% and the Dow Jones Industrial
Average fell more than 85% (FRED 2012, DJIA).

The impact of the Great Depression was not limited to the United States. Between
1929 and 1932 industrial production fell more than 20% in Great Britain and France
and more than 40% in Germany; foreign trade fell more than 60% in all three
countries (Blum, Cameron and Barnes 1970). Of the 16 major economies of Western
Europe (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland,
Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United
Kingdom), real GDP per capita (Geary-Khamis dollars) fell in 12 between 1929 and
1932, and rose only 2% over three years in the other four (Portugal, Ireland, Norway,
and Denmark). GDP per capita fell more than 10% across Western Europe and more
than 25% in the western offshoots (Australia, New Zealand, Canada and the United
States; Maddison 2008).
Recession between 1929 and 1933 was only the first of two in the decade. Though less severe and shorter lived, recession in 1937 and 1938 further highlighted the fundamental weakness of the economy. The recession of 1937 is often attributed to Roosevelt’s first steps to pull back the reflationary policies instituted during his first term, and the Banking Act of 1935, which reigned in liquidity (Eggertsson and Pugsley 2006); that is to say, an effort to retract Depression-era, emergency economic policies pushed the US economy back into recession.

Also, US recovery before 1937 was driven largely by monetary reflation (Friedman and Schwartz 1963; Bernanke 2000), which was made possible in part to the deteriorating political situation in Europe. The economy showed little sign of self-correction (Romer 1992). In other words, the same fundamental misallocations that led to the Great Depression were not corrected; the economic contradictions of the past were not resolved.

Likewise, a decade before the Great Depression, Keynes (1919) feared that the Treaty of Versailles did nothing to resolve the problems that led to and were exacerbated by World War I

The Treaty includes no provisions for the economic rehabilitation of Europe,—nothing to make the defeated Central Empires into good neighbors, nothing to stabilize the new States of Europe, nothing to reclaim Russia; nor does it promote in any way a compact of economic solidarity amongst the Allies themselves; no arrangement was reached at Paris for restoring the disordered finances of France and Italy, or to adjust the systems of the Old World and the New. (ch. 6)

War had disrupted the organization of economic and political life through much of Europe, and the resolution of war did not provide any solutions.

This population secured for itself a livelihood before the war, without much margin of surplus, by means of a delicate and immensely complicated
organization, of which the foundations were supported by coal, iron, transport, and an unbroken supply of imported food and raw materials from other continents. By the destruction of this organization and the interruption of the stream of supplies, a part of this population is deprived of its means of livelihood. (ch. 6)

Keynes ended *The Economic Consequences of the Peace* prophetically: “But who can say how much is endurable, or in what direction men will seek at last to escape from their misfortunes?” (ch. 6).

To capture the absurdity of the 1920s global economic system (and, specifically, America’s isolationist and protectionist policies therein), Roosevelt (quoted in Frieden 2007: 145-6) situated the following fictional dialogue in Lewis Carrol’s *Alice in Wonderland*:

A puzzled, somewhat skeptical Alice asked the Republican leadership some simple questions:

“Will not the printing and selling of more stocks and bonds, the building of new plants and the increase of efficiency produce more goods than we can buy?”

“No,” shouted Humpty Dumpty. “The more we produce the more we can buy.”

“What if we produce a surplus?”

“Oh, we can sell it to foreign consumers.”

“How can the foreigners pay for it?”

“Why, we will lend them the money.”

“I see,” said little Alice, “they will buy our surplus with our money. Of course these foreigners will pay us back by selling us their goods?”

“Oh, not at all,” said Humpty Dumpty. “We set up a high wall called the tariff.”

“And,” said Alice at last, “how will the foreigners pay off these loans?”
“That is easy,” said Humpty Dumpty. “Did you ever hear of a moratorium?”

And so, at last, my friends, we have reached the heart of the magic formula of 1928.

Many Europeans were dependent on loans from the United States, but excess capital dried up in 1928. A speculative boom in the United States absorbed excess liquidity, and then the crash of 1929 resulted in a liquidity crunch (Landes 1966; Arrighi 1994/2010).

US economic and financial policy in the interwar period, and the inadequacies of the Treaty of Versailles, highlight the fundamental contradictions of the 19th century British-centered free trade system: a system that depended on international cooperation was increasingly defined by interstate competition and uncertainty because economic and political power shifted from those who were the primary beneficiaries of a financially organized system of free trade to those whose participation was less lucrative.

Crisis and War. Vladimir Lenin (1917) famously attributed the outbreak of World War I to competing imperialisms: “the war of 1914-18 was imperialistic (that is, an annexationist, predatory, plunderous war) on the part of both sides; it was a war for the division of the world, for the partition and repartition of colonies, ‘spheres of influence’ of finance capital, etc.” In the most immediate sense, war in the 1910s and 1940s can be attributed to German efforts to match its growing military-industrial power with control over world resources, “this obsession [with Lebensraum] drove German rulers to try first to follow in the British [external], and then in the US [internal] path of territorial expansion. However, their attempts triggered a sudden escalation of interstate conflicts” (Arrighi 1994/2010:63).
Ironically, while post-hoc observers have documented a list of causal conditions before the Great War, contemporaries discounted the possibility of war among the major powers (Ferguson 2008). In 1914, the socialist journalist Henry Noel Brailsford argued that “In Europe the epoch of conquest is over and . . . it is as certain as anything in politics that the frontiers of our national states are finally drawn. My own belief is that there will be no more wars among the six great powers” (Brailsford 1914, quoted in Ferguson 2008: 299). Financial markets did not bat an eye at the assassination of the Archduke Ferdinand as traders were convinced of the security of the European political and financial system (Ferguson 2009).

This confidence emanated from international economic and financial interdependence. Much like Friedman’s Golden Arches (2000) and Dell (2005) theories of conflict prevention, contemporaries believed that interlocking supply chains and investments offered sufficiently strong incentives to prevent the escalation of conflict to war (Rowe 2005). European and American capitalists had a shared interest in a system that was making some of them exceptionally wealthy: “The belle époque of the Edwardian era marked the high point of Britain’s free-trade imperialism. The wealth and power of the propertied classes, not just in Britain but of the entire Western world, had attained unprecedented heights” (Arrighi 1994/2010: 277).

But this security was a façade. The “twenty years . . . of great splendor” followed a reflation of prices that allowed capitalists to escape the stagnation of the 1870s and 1880s, but the new financial liquidity was due in part to an escalation of the arms race (Arrighi 1994/2010: 277). Peace, then, was not a product of interests shared broadly,
but a delicate balance of power organized as economic interdependence communicated conditions of international assistance (Rowe 2005; Arrighi 1994/2010).

The fundamental source of conflict was that the global economy grew beyond the capacity of its primary beneficiaries to manage. The argument for free trade is that it encourages specialization based on comparative advantage. This organization of production benefits most those small (but relatively rich) countries that cannot specialize internally; the most dogmatic free traders at the turn of the 19th century were England and the Low Countries (Frieden 2007). A global economy centered on these countries was stable when they held the world’s workshops, but over the 40 years from 1870, “modern manufacturing spread from its limited base in Britain and northwestern Europe” (Frieden 2007: 59).

In 1850, the gross domestic product of the United Kingdom was 70% that of the United States and Germany combined. That figure fell below 60% around 1870, below 50% around 1880 and below 40% at the turn of the century. In the last few years before World War I, the GDP of the United Kingdom fell below 30% of the United States and Germany combined, less than that of Germany alone, and less than half the GDP of the United States. In 1890, the combined economies of Australia, New Zealand, Canada, and Argentina were 15% the size of the United Kingdom. In 1910, that figure topped 40% (Maddison 2008, author’s calculations; Frieden 2007). In 1870, Britain, Belgium and France together produced nearly half of the world’s industrial output. Industrial output in Germany and the United States was on par with and double that of Britian by 1913, respectively, and industrial output of Britain,
Belgium and France fell to a fifth of the global total (Frieden 2007). Economic power was shifting away from northwest Europe.

Ironically, though predictably, the late-19th century spatial configuration contained the seeds of its own demise. For example, facilitated by relatively free trade and the gold standard, Britain exported more than half of its capital in the years before World War I, much of that to the United States, and earned significantly higher returns on foreign investments than the domestic variety. Ten percent of British national income came from foreign investments, shipping, insurance and other international services (Frieden 2007). Successful investment encouraged rapid economic growth elsewhere, shifting power to countries whose interests were not in line with the existing configuration.

Rapid industrialization, particularly in Germany and the United States, increased competition for world resources, but where the United States and United Kingdom were able to control access through their internal and external empires, respectively, Germany was forced to pay an indirect tribute to each, the UK as the center of world commerce and the US as a recipient of German labor and capital (Arrighi 1994/2010: 62). Economic competitiveness depended on access to scarce resources that, particularly for European states, depended on access to colonies.

This “discontinuity between economic and political institutions” (Wallerstein 1979: 35) seeded the transition from inter-enterprise to inter-state competition and an arms race. Before the nations of Europe battled with bullets they found other ways to compete. Tit-for-tat trade tariffs rocked the global economy after the stock market crash of 1929. Even before it was signed in 1930, the historically aggressive Smoot-
Hawley tariff in the United States was met with retaliatory measures from Europe and Canada. The impact was substantial.

U.S. imports from Europe declined from a 1929 high of $1,334 million to just $390 million in 1932, while U.S. exports to Europe fell from $2,341 million in 1929 to $784 million in 1932. Overall, world trade declined by some 66% between 1929 and 1934. More generally, Smoot-Hawley did nothing to foster trust and cooperation among nations in either the political or economic realm during a perilous era in international relations. (US Department of State 2003)

Madsen (2001) estimates that economic and legislative barriers caused trade to contract 33% between 1929 and 1932.

Another tool was currency manipulation. By purchasing foreign currency, a country could (and can) make its currency relatively cheap (expanding the supply) while making the foreign currency relatively dear (decreasing the supply). For example, in the early 1930s, New Zealand devalued its currency relative to the pound sterling. This gave them a relative, and significant, advantage over Denmark as the two were the principal suppliers of butter to Great Britain. This launched a series of retaliatory devaluations between the two countries over the next three years. “By the end of 1933 the two currencies were back to roughly where they had started against each other, but ... competitive devaluations had heightened political tensions and protectionist pressures” (Frieden 2007: 185).

The Anglo-German confrontation was the most notorious arms race, but much of Europe and even South America got involved. The race was principally naval as the British threatened and devised plans to blockade German ports, and the Germans sought to build a fleet capable of preventing such a blockade. In 1889 the Royal Navy officially adopted the two-power standard, to field a navy equal to the combined prowess of the second and third most powerful navies (e.g., the German and
American navies). Despite aggressive investments to this end, the two-power standard eventually proved impossible due to the rapid expansion of the German and American navies. The world’s oceans were being filled with ever larger, more powerful, more destructive ships (see Figure 4.8).

Figure 4.8 Design Displacement of Capital Ships, 1905-1922

![Graph showing design displacement of capital ships from 1904 to 1922.](image)

Note: Displacement in tons of all capital ship classes
Source: Breyer 1973

Intensifying competition and political tension, manifested in trade and currency manipulation and an arms race, raised the costs while reducing the benefits of international economic participation. To place war in Europe in the second decade of the twentieth century at the feet of over-accumulation, during a period of financial expansion, would have seemed implausible to pre-war observers, but the international
ties that were supposed to prevent war were instead the mechanisms that dragged the world into a provincial conflict (Rowe 2005).

World War I revolutionized the financial world in two critical ways. First, Britain borrowed heavily from the United States to manage “the necessities of life and warfare” and liquidated assets in the New York Stock Exchange at heavily discounted prices (R.H. Brand, quoted in Milward 1970: 46; Arrighi 1994/2010: 279). As a result, liquidity began to flow to the United States, and less flowed back to Britain (Williamson 1964). This had two important (and related) impacts: first, it was a key first step towards the United States emerging as the new global financial hegemon after World War II – the United States no longer depended on British capital; second, the liquidity crunch in the United States a decade later would be all the more contagious internationally as the United States had become a major source of liquidity in international markets.

The second revolution of the financial world from World War I was that global finance collapsed almost completely between July and November 1914 (Ferguson 2009). John Maynard Keynes, the principal economic ideologue of the post-war era, made a name for himself digging up cash for the British war effort while working at the Treasury. After the war, an international consensus emphasized restoring currency stability and convertibility, but individual nations deployed controls on capital, trade and migration to secure stability: “While the intent was the freeing of trade, the effect was its strangulation” (Polanyi 1957: 27, quoted in Arrighi 1994/2010: 282).

The mishmash of prewar institutions (e.g., the gold standard) and postwar material conditions was combustible.
The gold standard is the key to understanding the Depression. The gold standard of the 1920s set the stage for the Depression of the 1930s by heightening the fragility of the international financial system. The gold standard was the mechanism transmitting the destabilizing impulse from the United States to the rest of the world. The gold standard magnified that initial destabilizing shock. It was the principal obstacle to offsetting action. It was the binding constraint preventing policymakers from averting the failure of banks and containing the spread of financial panic. For all these reasons, the international gold standard was a central factor in the worldwide Depression. Recovery proved possible, for these same reasons, only after abandoning the gold standard (Eichengreen 1992: xi).

As noted earlier, the economic absurdities of the 1920s fed directly into the economic catastrophe of the 1930s.

In turn, economic crisis, both of the immediate recession and the fundamental contradictions in the structure of the world-economy, ultimately offered an answer to Keynes (1919) rhetorical question about what men can endure and where they will seek escape. In this, there is a clear causal connection between the inevitable economic collapse, extremist politics and war near the end of the 1930s (James 1990). For example, the Weimar Republic depended on American loans, but American banks closed the spigot at the onset of the Great Depression. In the face of economic recession and consistent with the financial theory of the day, Chancellor Bruning implemented a strict policy of austerity and deflation, and unemployment rose above 30% before his ouster in 1932. (Fisher’s theory of debt deflation, arguing that the Great Depression was caused by an accumulated consumer debt followed by a period of massive deflation, was published later that year.) Hitler was sworn in as Chancellor in January 1933, and in 1939 he invaded Poland.

More than 100 million people served in military units during World War II, and it resulted in an estimated 50 to 85 million deaths (Somerville 2008). Focusing on the
United States, industrial production trebled between May 1938 and February 1944 (FRED 2013, INDPRO), and food prices doubled between 1939 and 1943 (BLS 2013, WPU01). Unemployment essentially disappeared, falling from 14.6% to 1.2% between 1940 and 1942 (BLS 2011) and output per capita increased 82% between 1939 and 1944 (Williamson 2013). Because the other belligerent countries were actively engaged in destroying each other’s productive capacity, production in the United States soared relative to Europe. Before the end of the world, US production was on par with the rest of the world combined (Bolt and van Zanden 2013; see chapter 5). As New Deal policies were introduced to manage an economy in crisis, a series of measures (e.g., price and wage controls) were implemented to manage an economy in total war; we will return to these measures below.

The net result was that the economic globalization that defined the British-centered spatial configuration was dead. The gold standard, a central feature of the system, was abandoned. Trade and migration networks collapsed. Though victorious in war twice over, British hegemony ended, and the United States rose up to take that crown. After the war, US production equaled all of Western Europe combined, and the US controlled 50% of the world’s gold reserves and two-thirds of global monetary reserves (Eichengreen 1996). After World War II, the sun set on the American “empire”, but it did not set on its armed forces.

Bringing the narrative back to the focal point of this study, emerging contradictions in the late-19th century global spatial configuration led to economic depression and war. In addition to destroying assets owned disproportionately by the rich, they created space and a momentum for institutional and ideological...
experimentation that led ultimately to lower levels of economic inequality in the United States and elsewhere. In the next section I take a closer look at the evolving intellectual environment through the Great Depression to World War II.

The Intellectual Response

Economics as a discipline tends to emphasize its evolutionary progressivity as a science, but it is also historically produced knowledge (Cohen and Emmett 2012). A quick study of the history of economic thought highlights the intimate relationship between system-wide economic failure and theoretical ‘paradigm shifts’. The 1930s was no different.

Economic orthodoxy in the 19th century achieved a theoretical purity that has not and never again can be matched in the social sciences. Classical economics rested on a single, empirically established relationship, Adam Smith’s observation that specialization increased productivity. From this point, Smith, Ricardo and their ilk deduced that the degree of specialization (and thus productivity) was inextricably linked to the scale of autarky, the unit within which all the necessities of life could be produced. Larger markets, linked through exchange, encouraged greater specialization, productivity, and wealth. Any barriers to that exchange effectively shrank the size of the market, and therefore reduced specialization, productivity and wealth. The timing of the development and application of this logic in the world-economy is not a coincidence; transportation technologies exploded the functional size of markets in the first half of the 19th century (see chapter 3).

The theoretical principles supporting the gold standard were fuzzier; many economists regarded the commitment to the gold standard as a symptom of a precious
metal fetish (Frieden 2007). Instead, we can interpret it as a political stipulation of laissez-faire economics across political boundaries.

Most orthodox supporters of the system argued that substantial state intervention in the market would interfere with the natural operation of the gold standard. They believed that unemployment compensation, aid to troubled farmers, and extensive social programs for the poor would impede adjustments required by the gold standard; such programs would keep wages and prices from falling as necessary to keep economies in balance. (Frieden 2007: 30).

In other words, the gold standard hamstrung government intervention through fiscal and monetary policy that, in turn, was conducive to the operation of the free market.

Economic orthodoxy was not consensus. Williams Jennings Bryan famously centered his bid for the presidency of the United States in the 1890s on decrying the gold standard: “you shall not crucify mankind upon a cross of gold.” Participation in the global economy, though, required that the state and its treasury abide by the rules of the game under British hegemony (Frieden 2007).

The intellectual paradigm shift of the 1930s was fundamentally different from other transitory periods in one critical sense: “the Depression [gave] birth to macroeconomics as a distinct field of study” (Bernanke 1995: 1). Where microeconomics continues to be the domain of neoclassical economics, the crisis of the 1930s turned intellectuals and policy makers away from the constraints of the gold standard and in pursuit of tools for managing economies nationally and internationally. More than a new field of study, the emergence of macroeconomics as a field of study marked a fundamental shift in the relationship between the state and the economy. There are myriad examples of this new relationship, from new regulations, social insurance schemes, and public work projects, but none are more
fundamental in the United States than the Current Population Survey. The CPS was initiated in 1940 by the Work Projects Administration under the name Monthly Report of Unemployment (BLS 2013a); it was a new tool, introduced by a new agency, to closely monitor an economic indicator that had only recently been technically defined.

As had been the case in the past and would again be the case in the future (see chapter 5), the social and economic theories of the Edwardian era were completely unprepared for the chaos of the subsequent decades; decades of relative stability and security (low interest rates) and strong economic growth (especially for those controlling mobile capital) had offered the contemporary political and economic elite a sense of arrival (Ferguson 2009). As that illusion shattered, existing paradigms were quickly reimagined. I highlight three thinkers that offered insights into the crisis but also captured the revolution in economic thought: Karl Polanyi, Joseph Schumpeter and John Maynard Keynes.

Polanyi 1944 - *The Great Transformation*. Polanyi argued that markets played a small role in human societies in the past, and that the apotheosis of the self-regulating market was an intentional project of the modern nation-state. The system required that all components of production, including land, labor and money, be subjected to the market principle; workers were intentionally proletarianized: “under such a system we can not exist unless we buy commodities on the market with the help of incomes which we derive from selling other commodities on the market” (p.97).

Fundamentally, the transition dis-embedded economic activity from social relations. “The true criticism of market society is not that it was based on
economics—in a sense, every and any society must be based on it—but that its economy was based on self-interest. Such an organization of economic life is entirely unnatural” (p.249).

Such an institution could not exist for any length of time without annihilating the human and natural substance of society; it would have physically destroyed man and transformed his surroundings into wilderness. Inevitably, society took measures to protect itself, but whatever measures it took impaired the self-regulation of the market, disorganized industrial life, and thus endangered society in yet another way. It was this dilemma which forced the development of the market system into a definite groove and finally disrupted the social organization based upon it (p.3-4).

The result was political chaos, national and international, as groups violently pursued a broad range of measures to secure “professional status, safety, and security, the form of a man’s life, the breadth and depth of his existence, the stability of his environment” (p.153). In reference to the chaos of the 1930s, Polanyi argued that the “only alternative . . . was the establishment of an international order endowed with organized power which would transcend national sovereignty” (pg. 22).

Schumpeter 1942 – Capitalism, Socialism and Democracy. While the mechanisms involved are very different, Schumpeter also argued that there is a fundamental tendency in capitalism to destroy the institutions organizing economic and social life: “its very success undermines the social institutions which protect it, and ‘inevitably’ creates conditions in which it will not be able to live” (p.61).

The contradiction is rooted in the duplicitous character of capitalism itself—creative destruction. In this sense, Schumpeter was responding to institutionalism, another stream of economic thought that gained popularity in the 1930s and emphasized the costs of monopoly and imperfect competition (for example, Robinson
1933; Chamberlin 1933). He argued that the success of capitalism was rooted in innovation, not perfect competition (see chapter 1).

Innovation is both creative (the introduction of a new technology or good) and destructive (the old technology or good become redundant), and so the success of capitalism amounted to a constant churning of technologies, skills, firms, markets, institutions, etc. “Capitalism, then, is by nature a form or method of economic change and not only never is but never can be stationary” (p.84). As such, capitalism is constantly producing both winners and losers, and “[s]ecular improvement that is taken for granted and coupled with individual insecurity that is acutely resented is of course the best recipe for breeding social unrest” (p.160), and so it is undermined by its own success.

Schumpeter’s unique perspective on capitalist progress offered a unique remedy. Because progress is fundamentally rooted in innovation and not price competition, barriers to perfect price competition are not contrary to the capitalist project: “There is no more of a paradox in this than there is in saying that motorcars are travelling faster than they otherwise would because they are provided with brakes” (p.88). Regulated, corporate capitalism could speed along faster than an economy defined by perfect competition because it faced less risk of crashing.

Keynes 1936 – The General Theory. The fundamental argument of Keynes’ General Theory of Employment, Interest and Money is that fiscal policy works (see chapter 5 for the monetary rebuttal). Economic policy should be oriented towards achieving full employment (the exclusion of deficient-demand unemployment). Employment is a product of demand, and demand can be stimulated by spending, so governments can
pursue their economic goal of full employment by spending to eliminate deficient-demand unemployment.

If the Treasury were to fill old bottles with banknotes, bury them at suitable depths in disused coalmines which are then filled up to the surface with town rubbish, and leave it to private enterprise on well-tried principles of laissez-faire to dig the notes up again (the right to do so being obtained, of course, by tendering for leases of the note-bearing territory), there need be no more unemployment and, with the help of the repercussions, the real income of the community, and its capital wealth also, would probably become a good deal greater than it actually is. It would, indeed, be more sensible to build houses and the like; but if there are political and practical difficulties in the way of this, the above would be better than nothing (p. 129).

State intervention is essential because economic recessions have an irrational component. Demand depends on consumer confidence, which is in turn undermined by economic recession. Government spending is necessary to “kick start” private spending.

Keynes extended this logic to emphasize reducing interest rates and international monetary reform to encourage spending – consumption and investment – in the private sector (Tily 2007; Davidson 2002). The logic relating interest rates and unemployment would be formalized in the Phillips curve, and, combined with counter-cyclical spending, would become the most important economic levers of macroeconomic policy.

The need for domestic and international oversight of the economy can also be related to a field that was first formally defined by John Von Neumann in a paper in 1928 and then in book form in 1944. Game theory emphasizes interactive decision-making. Both domestically and internationally, the outcome of an individual strategy depends on the strategy adopted by other agents. For example, in a recession, an investment might only be successful if consumers are confident that others will also
invest and are therefore willing to spend. As Keynes emphasized, government spending can substitute for consumer confidence until consumer confidence is restored.

Though their logics varied substantially, Polanyi, Schumpeter and Keynes agreed that markets should not be left to their own devices. Taken together, they offered a damning critique of classical, laissez-faire economics:

1) “Self-regulating” markets take a huge toll on human relationships and the environment, are socially disruptive and self-destructive

2) Government intervention is essential to prevent market failure and limit the human and environmental cost.

3) National state intervention is not enough; stable economic and financial systems require international cooperation and regulation.

The intellectual and institutional response to evolving material conditions began before intellectuals formalized the logic. For example, the Works Progress Administration was active in 1935, before Keynes recommended the Treasury bury bottles of cash; the United States was organizing international talks in Bretton Woods, New Hampshire before Polanyi argued that the alternative to chaos was an “international order endowed with organized power”; Josiah Wedgwood added a preface in 1939 to his book, *The Economics of Inheritance*, “political democracies that do not democratize their economic systems are inherently unstable” (quoted in Piketty 2014: 520). But Keynes’ work, especially, became the ideological core of the new inequality regime, and framed public economic debate for decades afterwards. It is within this new intellectual framework that the ‘Great Compression’ persisted.

Korzeniewicz and Moran (2009) argue that low inequality in high income countries is achieved by spatially excluding the economically disadvantaged; there is
a tradeoff between within and between country inequalities. This requires that the state has both the incentive and capacity to perform these functions. The concurrence of rising inequality between countries and falling inequality within many countries in the first half of the 20th century suggest that capacity and incentive converged between 1913 and 1950. State capacity (revenue and regulatory powers) increased with the threat and then execution of war. Immigration and trade policy and the gold standard were modified to give states more leeway in fiscal policy and to control the movement of capital, goods and people over their borders. The revolution in economic thought targeted this new capacity on maximizing demand, stabilizing the financial system, and regulating trade and immigration in a sustainable way.

*The Mechanisms of the Great Compression*

Research into the causes of inequality change, particularly in the United States over the last four decades, favors analyses that decompose inequality change into component parts. As I noted in chapter 1 and reiterate in chapter 5, this approach has limits. A decomposition of change requires some assumption of the counterfactual – inequality change while holding the single variable constant. The assumption of a counterfactual is complicated by 1) omitted variable bias and 2) dependence between the variables – the change in variable x and its impact on inequality are both dependent on variable y. The attempt to measure effect size while accounting for these complications, especially when looking only at a single case (e.g., the United States), is more than ambitious. General equilibrium analysis claims to overcome these limitations, but only by binding human agency and reflexivity; it is a fundamental premise of my analysis that equilibria are social constructs.
Fortunately, it is not my goal to weight and rank the specific mechanisms driving inequality change. Instead, in the following section I relate the crisis and the subsequent intellectual response described above to mechanisms that have a clear relationship to falling levels of wealth, wage and income inequality. These mechanisms can be organized into three categories: 1) supply and demand of less-skilled labor; 2) destruction of and constraints on financial incomes and wealth; 3) long-term consequences of emergency institutions.

Supply and demand of less-skilled labor. The most famous statement on falling inequality in the first half of the 20th century comes from another intellectual, Simon Kuznets (1955: 18): “One might thus assume a long swing in the inequality characterizing the secular income structure: widening in the early phases of economic growth when the transition from the pre-industrial to the industrial civilization was most rapid; becoming stabilized for a while; and then narrowing in the later phases.”

For Kuznets, the principal mechanism driving inequality through this process is sector dualism. Productivity is higher in industry than agriculture, so wages are higher as well. Because of this wage gap, wage inequality is higher if workers are spread between the two, industrial and pre-industrial, than if they are clustered in one or the other. The effect is that inequality tracks an inverted U-curve as the majority of labor shifts from agriculture to manufacturing, or traditional to modern employment.

But as we noted in chapter 3, rising inequality in the last half of the 19th century involved much more than a bifurcation of labor. While there was a wage gap between agriculture and manufacturing, “American inequality trends appear within sectors and regions. They are not just an aggregate artifact of migration between sectors”
(Williamson and Lindert 1980: 142). Sector dualism has little to say about the concentration of incomes at the very top of the distribution.

The evidence outside of the United States for an inverted U-curve is far from overwhelming; to borrow the oft-cited line from Ravallion (1995:415), “The rejection of the inverted U hypothesis could hardly be more convincing.” While additional research has corroborated the cross-national inverted-U relationship between development and inequality, there is no relationship between growth and inequality within a country over time (Ravallion 1995; Deininger and Squire 1998; Schultz 1998), and Fields and Jakubson (1994) even found a significant U pattern (not inverted) between growth and inequality. Korzeniewicz and Moran (2005) argue that the static inverted-U reflects the hierarchy of the world-economy, so that middle-income countries tend to have higher levels of inequality even as global wealth, and thus the definition of middle income, rises. If economic development does not drive within-country inequality the modernization thesis has no paddle.

But technological change and industrialization played a key role in the inequality narrative for the United States. A change in the mode of production can fundamentally change the relationship between the state, business and labor. For example, the capital to labor ratio is significantly higher in manufacturing than agriculture in the early 20\textsuperscript{th} century (Williamson and Lindert 1980). Industrialization in the United States meant profits were increasingly extracted from machines instead of workers. Investment patterns in physical and human capital vary across sectors, and, as the composition of the US economy by sector evolved, investment patterns on
the aggregate evolved as well (Williamson and Lindert 1980; Lichtenstein 1995; Levy and Temin 2007).

An anecdotal example is Ford’s $5 wage. In 1914, Ford instituted a $5/day wage for many workers. The public relations claim was that the company wanted workers to be able to buy their own Ford, but the policy had the more critical goal of reducing attrition. Assembly-line work was miserable, especially for a labor force unfamiliar with industrial discipline. Given the high costs of turnover and mechanization of automobile production, Ford was willing to invest more in wages to reduce costs in recruitment and training. In the coming decades, employers would accept other, more formal, compromises with labor, occasionally negotiated by the state.

Unfortunately, the literature on the labor-saving and labor-augmenting technological change and labor demand is painfully confused by limited and inconsistent definitions of skill. Skill is variably defined in terms of education and occupation, and the relationship between these indicators and skill are tautologically confirmed by wage hierarchies. The categorization of skilled and unskilled by education varies over time; a high school educated worker can be skilled in the late 19th century and unskilled in the mid-20th. When, exactly, did an artisan with little classroom education but a lengthy apprenticeship stop being highly skilled?

A sectoral shift to manufacturing and a surge in manufacturing productivity in the early 20th century is causally linked in the literature to rising inequality in this period (Morishima and Saito 1968; Keller 1973; Williamson 1976). But demand for workers in manufacturing is associated with a massive wage compression and low levels of inequality from the 1940s (Goldin and Margo 1992; see chapter 5); in the 1910s,
manufacturing was “skill-intensive” relative to agriculture (Williamson and Lindert 1980), but manufacturing employment was often a substitute, not a compliment, to high school enrollment (Goldin and Katz 1997). An attempt to reduce any development during this period as a trend in the relationship in wages between “skilled” and “unskilled” workers is folly.

Despite these complications, some relationships are fairly easy to establish. Immigration into the United States reduced wages (see chapter 3; Goldin 1994; O’Rourke and Williamson 1999). It would then follow that falling immigration rates would have the obverse effect. Immigrants as a share of the US population began to fall after 1913 with the outbreak of World War I and legislation restricting immigration in 1921 and 1924; only after 1970 did that figure begin to rebound (Gibson and Lennon 1999). Along those lines, and related to Kuznet’s model of economic modernization and demographic transition theory, the US fertility rate fell through the first half of the 20th century to 1946. The correlation between the growth rate of the labor force and the wage premium for ‘skilled’ labor is non-negligible (see Figure 4.9; Williamson and Lindert 1980).

The implication of the figure below goes beyond the mere relationship between wages and supply of labor; the measure opposite labor force growth is the skilled wage premium (skill being a loosely defined concept aggregated from various wage series). Labor force growth through US history generally reflects the addition of a disproportionate share of less-skilled workers, either because they are young, from fertility, or, in the case of immigrants, because they are less-skilled or skills do not translate internationally.
Notably, the period from 1929 to 1948 falls well below the linear fit between labor force growth and change in the wage premium. The implication is that, for this period in particular, the wage premium fell more than would be explained by slowing labor force growth (the opposite is true for 1820 to 1840). The simple solution is that there was a change not just in labor force size, but composition.

The High School Movement beginning in 1910 served the function of changing the composition of the US labor force. The percent of US youths with a high school diploma exploded from 9% to 50% between 1910 and 1940 (Goldin and Katz 1997). In 1940, 73% of 15 to 18 year olds were enrolled in high school (Goldin and Katz 2007). The Movement also entailed a transformation in the function of secondary
education, from an emphasis on training for college to an emphasis on training for life (Goldin and Katz 2007), though greater access to public universities in the 1940s would cause that emphasis to again reverse. The US labor force was fundamentally transformed—the typical American was better educated, and women’s access to education, and then employment, also surged (Goldin 2006). This undoubtedly played a role in compressing wages (Goldin and Katz 1997). Access to tertiary education would expand rapidly after World War II.

The narrative becomes more complicated when we also consider the other side of the coin, labor demand. On one hand, changes in demand through this period can be explained by historical events: “Most of the narrowing in wage differentials, for example, took place in the 1910s and the 1940s, periods close to or coinciding with the two world wars” (Goldin and Katz 2007: 29). Likewise, Chiswick and Mincer (1972) show that falling unemployment during World War II was an important factor in falling inequality during that period, and my analysis above concurs. But the world wars ended, and the impact on wage differentials did not.

Goldin and Katz (2010) estimate changes in skill premia (college to high school, high school to high school dropout) since 1915 using changes in relative supply and various time dummies to reflect changes in relative demand. The sparseness of data for the earlier period limits the precision with which they can identify changes in relative demand, but consistent with results reported above they find that the relative demand for college educated workers fell from 1915 to 1949.

The economic solution points to technological change. Technological change is not measured directly, but instead is operationalized by the capital to labor ratio; an
increase in the ratio of capital to labor indicates that new technology increased the
returns of capital relative to labor. A number of studies indicate that technological
change had a labor-saving bias from the beginning of the 19th century and again
through the 1920s, but reversed course in the 1910s and after the 1920s (see
Williamson and Lindert 1980, Morishima and Saito 1968). More specifically, output
shifted to less-skilled labor-intensive sectors during periods of labor-augmenting
technological change (Williamson and Lindert 1980).

The potential contribution of technological drift on wage differentials is limited.
For example, wage differentials trended up and down through the 1910s, 1920s and
1930s, but they collapsed in the first half of the 1940s just as industrial output tripled
(Goldin and Margo 1992). Demand for industrial products to feed the war effort was
met by exploiting the growing class of high school educated workers. This was not
exogenous technological change, but a planned and structured reorganization of
economic activity.

Destruction of and constraints on financial incomes and wealth. As noted earlier,
financial systems shuddered in the first year of World War I. The income share of the
top 1% of earners in the United States fell from 18.2% to 14.5% between 1914 and
1920, only to rebound soon after. A decade later, the Dow Jones Industrial Average
lost 80% of its value from 1929 to 1932 (FRED 2012). Because only 5% of families
were actively associated with the stock market in 1929 (Galbraith 1954), the top 1%
of earners in the United States were disproportionately affected, and their share of all
incomes fell from 23.9% to 15.5% between 1928 and 1931 (Piketty and Saez 2003).
In a preliminary study, Atkinson and Morelli (2011) find that top income shares can move up or down with a financial crisis, and that “there is more evidence that financial crises are followed by rising inequality” (pg. 49). For example, top income shares excluding capital gains were higher in 2010 than in 2007 (Alvaredo, Atkinson, Piketty and Saez 2012). There is no doubt that top income shares were adversely affected by the stock market crash of 1929, but these shares could have rebounded. But the crash of 1929 was followed by a significant restructuring of American and global finance with the goal of preventing another collapse.

The Securities and Exchange Commission was created in 1934 to regulate the securities industry and stock and option exchanges. The 1933 Banking Act introduced the Federal Deposit Insurance Corporation to insure commercial banking deposits. Four provisions of the 1933 Banking Act, generally known as Glass-Steagall, limited the involvement of commercial banks in securities and separated commercial and investment banks. The act also introduced other forms of regulation: it prohibited interest payments on demand deposits (e.g., checking accounts); it allowed the Federal Reserve Board to limit interest rates on time deposits; it sought to limit “speculative” uses of bank credit through oversight from the Federal Reserve Banks.

Given the financial chaos of the period, it is difficult to isolate the impact of these measures from the noise, but the Dow Jones Industrial Average grew at a significantly faster rate between 1921 and 1929 in constant terms than at any other point in history (FRED 2012, author’s calculations), and wages in finance relative to other sectors fell substantially through the 1930s (Philippon and Reshef 2009). As we
will see in more detail in chapter 5, financial deregulation is associated with a concentration of income and wealth in the financial services sector.

Long-term consequences of emergency institutions. As noted above, the global economic system prior to World War I was not conducive to government intervention through fiscal or monetary policy. But the political and economic chaos in the last half of the 1910s created space and demand for government social services.

The spread of social insurance and welfare statism both in terms of countries covered and concerning the scope of provisions and its beneficiaries increasingly became part of social and economic policy in the aftermath of World War I. While previous decades had seen rather cautious experimentation with various schemes of social insurance, social security principles were now gaining ground, developing rapidly in most European countries and European settler nations and also spreading to the European periphery and beyond Europe. (Kuhnle and Sander 2010:75)

The transformation of the relationship between the state and its citizens (an important distinction from residents) over the next quarter century is anecdotally captured by the publication of the Beveridge Report in 1942. The British economist William Beveridge chaired the inter-departmental group tasked with performing a survey of existing social services in the UK and to make recommendations. Beveridge did not shy from the opportunity; he argued that this “revolutionary moment in the world’s history is a time for revolutions, not for patching” (Beveridge 1942: 6).

The recommendations of the report traced the new logic of reformed capitalism. Beveridge argued that social security would be achieved through cooperation between the state and individual, but that “the State in organising security should not stifle incentive, opportunity, responsibility; in establishing a national minimum, it should leave room and encouragement for voluntary action by each individual to provide
more than that minimum for himself and his family” (Beveridge 1942: 6-7). The report became the explicit reference point for the foundation of the post-war welfare state in the UK and elsewhere. While not without its critics, the report was well received and explicitly referenced in policy discussions around the world (Barnett 2001; Ferguson 2009).

“The arguments for state insurance extended beyond mere social equity” (Ferguson 2009: 33). Through two world wars and the Great Depression, attention turned to uncertainty. War and economic chaos highlighted the inadequacies of individual or private institutional approaches to insurance. The lesson first introduced in World War I was further impressed during the Great Depression – the liquidity crunch affected actors well outside the financial services industry and around the world – and World War II: “In Japan, as in most combatant countries, the lesson was clear: the world was just too dangerous a place for private insurance markets to cope with. . . [I]ndividuals could not be expected to insure themselves against the US Air Force” (Ferguson 2009: 207).

Greater intervention by the state in economic affairs, particularly where that intervention was oriented towards security, insurance and protection, was accepted by capital and labor alike as a necessary response to the chaos of previous decades. Business interests were not particularly excited about the full range of changes being implemented, especially since some programs redistributed incomes through progressive taxes and means-tested transfers. But the state could no longer operate on the grounds that the “business of government is business.” “Unrestrained competition sets [capitalists] to struggling with each other and soon arouses resistance . . . among
the proletariat that they are exploiting. . . . [S]tates, perceiving that it is impossible to leave employers and employees to contend in anarchy, elaborate a social legislation” (Pirenne 1953: 516, quoted in Arrighi 1994/2010). Before World War II, conflict between capital and labor fed political instability. During World War II, the mechanization of war obliterated any remaining division between industrial production and state power. After the war, the threat retained physical form in the Soviet Union.

The concentration of workers in factories and densely populated urban neighborhoods influenced the outcome of state intervention.

[I]n democratic societies the growing political power of the urban lower-income groups led to a variety of protective and supporting legislation, much of it aimed to counteract the worst effects of rapid industrialization and urbanization and to support the claims of the broad masses for more adequate shares of the growing income of the country (Kuznets 1955: 17).

In other words, economic modernization is not merely an economic process, but is also political and institutional.

For the United States, four institutional developments are particularly important in the history of wage and income inequality: progressive federal income tax, government transfers, collective bargaining rights, and wage regulation.

The 16th amendment was ratified in 1913, granting the federal government the right to directly tax individual incomes. The amendment sat astride two eras. On one hand, the income tax was argued to be a more efficient means of generating revenue than tariffs; this appeased free traders. On the other hand, Theodore Roosevelt dragged Republicans behind the amendment to raise funds to compete in the naval
arms race and generate funds to pursue the Progressive Republican agenda (Buenker 1981).

Top marginal tax rates, and tax rates in general, followed a consistent pattern through the first half of the 20th century. During each crisis – World War I, the Great Depression, World War II – tax rates rose. After the crisis, rates would fall but to a higher level than before the crisis. For example, an individual with an income of $500,000 (1982 dollars) in 1916 paid 2.2% of that in federal income taxes. That figured jumped to 18.0% in 1918, and then fell to 9.9% in 1924. That same income would be taxed at 7.8% in 1931, 12.2% in 1932, and 14.0% in 1935. Before tax rates had an opportunity to fall, World War II raised taxes on that $500,000 of income to 35.0% in 1941 and 52.0% in 1945. The rate fell to 41.2% in 1949 (Tax Foundation 2013, author’s calculations).

Because income taxes tend to be progressive, they can have a significant impact on the distribution of incomes. The relationship between inequality and taxes post-tax is mechanical, but high taxes can also effect incomes pre-tax as firms adjust their investment decisions based on expectations of returns (Piketty and Saez 2003; Piketty 2014); ceteris paribus, higher taxes on higher incomes reduces the return on that investment, so firms direct funds elsewhere.

By this logic, a focus on top marginal tax rates is insufficient; we must also consider the progressivity of taxes. To operationalize tax progressivity (see Figure 4.10), I used a uniform distribution of incomes from 0 to $1,000,000 (1982 US$) and estimated the total taxes paid by individuals across that distribution. I subtracted the Gini coefficient post-tax from that before (Gini=.333). The progressivity of taxes is
largely a product of top tax rates, as the tax rate on an income $500,000 and the progressivity measure track very closely. They do diverge some in the 1940s when tax rates on lower incomes rose significantly.

Figure 4.10 Effective Tax Rate and Tax Progressivity

Isolating the impact of income taxes on the distribution of incomes during this period is made functionally impossible because of the complexity of causal relationships between income tax rates, crisis, and other economic and institutional changes (many of which depended on new federal revenues to cover costs), but the correlation between top marginal tax rates and top income shares through the 20th century is substantial (Piketty and Saez 2003). As I discuss in chapter 5, empirical research has linked tax rates and inequality (Saez, Slemrod, and Giertz 2012; Altig and Carlstrom 1999), and the 1980s offered a wonderful natural experiment: Reagan twice substantially reduced tax rates and top income shares jumped in response.

The impact on new government spending on inequality is obvious in some cases, but more obscure in others. An obvious case is Social Security. It has played a major
role in reducing poverty among the elderly. Less obvious is the impact of unemployment insurance and make-work programs. They act as competitors to private employers. A worker should be less likely to accept a pay cut during a recession and more selective when seeking employment if there is a safety net in place. More broadly, life course research has demonstrated that the impact of negative events – unemployment, injury, divorce – accumulate over time as they increase the probability of other negative events. This can drive up inequality within a cohort over time (Albrecht 2007). Public insurance schemes can break that cycle.

The relationship between labor-related legislation in the 1930s and union membership is undeniable. In 1932, Norris-La Guardia declared that work contracts disallowing union membership as a condition of employment were unenforceable by federal law, and it prevented federal courts from issuing injunctions against non-violent labor disputes. More important, the Wagner Act in 1935 guaranteed private sector workers the right to organize, bargain and act collectively. It established the National Labor Relations Board to monitor union elections.

Union membership among nonagricultural workers (Wagner does not cover agricultural workers) was at 11.3% in 1933. That figure rose to 13.2% in 1935, 27.5% in 1938 and 35.5% in 1945 (Hirsch and Macpherson 2003). Of course, legislation on paper was not enough to ensure the growth of trade unions in the United States; the last half of the 1930s was rife with political, legal and street battles. Union membership growth halted after 1946; Taft-Hartley in 1947 repealed some protections granted in Wagner.
Again, the correlation for this period between income inequality and union power is non-negligible. Union membership fell between 1920 and 1929; the income share of the richest 1% rose from 14.5% to 18.4% in that period (Williamson and Lindert 1980; Piketty and Saez 2003; see also Baran and Sweezy 1966 and Keller 1973). That figure fell to 14.7% in 1938 and 10.5% in 1944 as union membership rose.

Economic historians are often critical of a hypothesized relationship between union membership and inequality. The correlation might be spurious: “unionization and union strength may be only a manifestation of strong market demands for labor in general and unskilled labor in particular—demands which themselves must be explained” (Williamson and Lindert 1980: 140). Even at their height trade unions represented only a minority of US workers, and they may have been responsible for driving down wages for non-union workers – although union demands were more in line with the interests of “low-skill” workers after the 1930s than before (Williamson and Lindert 1980).

Unions influence wage inequality directly and indirectly. Directly, unions compress wages among union workers by 1) promoting equal pay for equal work and 2) trimming the gap between high and low wage workers (Freeman 1980; Card, Lemieux and Riddell 2003). Indirectly, trade unions institutionalize “norms of equity”. Wage inequality tends to be lower when wage rates are institutionalized. Western and Rosenfeld (2011) find that inequality in wages among non-union workers is lower in areas with high union membership. Reuther’s Treaty of Detroit, a contract between UAW and General Motors in 1950, became the model for a new
social contract by which wages kept pace with rising labor productivity for the next few decades (Levy and Temin 2007; Lichtenstein 1995).

Organized labor more effectively supports pro-labor politicians than unorganized labor. For example, Republicans gained 72 seats in the House of Representatives in 1938. The recession of 1937 was the key factor, but a bitter split between the AFL and fledgling CIO also played a role. A Republican congress would later repeal some New Deal legislation and pass Taft-Hartley over Truman’s veto.

In addition to organic changes in supply and demand, higher taxes, and stronger unions, wage compression in the United States was also a direct act of Congress. The Federal Labor Standards Act established a minimum wage of $.25 an hour (about $3.90 in 2010 dollars, 54% of the actual 2010 value; see Figure 4.9; DOL 2013; BLS 2013, CPI; author’s calculations). Then, in an effort to manage an economy overheating in World War II, the National War Labor Board (NWLB) was created in 1942 to regulate wages.

Wage compression was not within the mandate of the NWLB, but many policies had that effect. For example, standardized wages reduced within-firm and within-industry variance in wages; wages could be increased to $.40 an hour without NWLB approval (33% greater than the minimum wage). The net effect, because the NWLB offered more exceptions for “substandard” wages, was wage compression by fiat (Goldin and Margo 1992).

The real value of the minimum wage doubled in the two decades after FLSA, and coverage expanded in the 1960s (see Figure 4.11). The federal government continues to enforce a national minimum wage, but the declining real value of that minimum
Wage in recent decades is one potential source of rising inequality in recent decades (DiNardo, Fortin and Lemieux 1996; Lee 1999; Teulings 2000; 2003). Conversely, it seems plausible that introducing and raising the minimum wage in the three decades after 1938 would compress wages. But in this case the minimum wage was moving in step with rising living standards generally, so its impact was limited.

Figure 4.11 Indexed Minimum Wage, Constant Dollars (1938=1)

The NWLB was dissolved in 1945, but Thurow (1975) argues that legislated wage compression established new normative expectations about wage differentials that were influential long after the immediate justification for controls vanished. Another possible explanation for the stickiness of command economy wage compression is that firms adjusted their practices – capital investments, employee training – and organization to rationalize the war-era wage structure.
A New Social Contract

The United States emerged from World War II in a new, lower inequality regime; as I have detailed above, it reflected changes in the underlying material, normative, institutional, political, and ideological conditions. While these changes produced lower levels of inequality, it is important to note that egalitarianism was rarely a guiding principle. Beveridge (1942) listed want, disease, ignorance, idleness and squalor, not inequality, as the obstacles to reconstruction. The new social contract hinged on two principles: 1) economic security requires strong domestic demand from a secure, productive, high-wage labor force; 2) economic security and international peace require nationally and internationally regulated finance and trade (Eichengreen 1996).

Domestic demand. Henry Ford’s explicit intention with the $5 wage, that workers would be able to buy their own Ford, is relevant. A firm cannot survive selling its products only to its employees, but a country can. The first half of the 20th century highlighted the costs and risks of depending on global trade and finance.

Specialization may be associated with greater productivity, but it also creates an economic dependence on political entities with conflicting interests. The political consequences of economic recession are compounded because 1) the state has little power to mitigate its impact and 2) this economic impotence is put on display.

Keynesian economic thought offered an alternative. Fiscal and monetary policy can effectively manage economic downturns by manipulating domestic demand. Through the 1930s it became economic orthodoxy to employ counter-cyclical spending to mitigate the impact of economic recession, and in subsequent decades to
manipulate interest rates to encourage investment and spending when appropriate. While counter-cyclical spending makes sense to boost demand in the short run, in the long run greater demand must be matched with increased productivity. Otherwise domestic demand will target foreign imports with debt (see chapter 6).

Make-work projects, unemployment and old-age insurance schemes, expanded public education, collective bargaining rights and negotiated arrangements with capital to avoid work stoppages, and public investments in infrastructure had the net effect of raising and stabilizing demand while also encouraging productivity gains to match. Within the new social contract, the state was responsible for encouraging job creation and investment, education, and eliminating or managing economic recessions. This was a dramatic change from earlier decades, during which corporations secured markets through mergers, horizontal and vertical integration to protect against "cut-throat" competition. Legislation (antitrust laws) and academic opinion (e.g., Robinson 1933 and Chamberlin 1933) turned against monopolistic corporations. In the Keynesian regime, "cut-throat" competition was to be avoided by continuously expanding the pie, not by allowing a single corporation to control more of the same pie.

As such, after the compression of wages, incomes and wealth through the Great Depression and World War II, economic gains in the post-war era were not equally distributed; those at the top of the distribution grabbed a larger share of profits. But gains were distributed proportionately; income shares were fairly stable for two decades, and the capacity of workers to purchase grew in hand with their capacity to produce.
Regulated Finance. As war raged in Europe, there was quasi-consensus that the international financial system had failed its functionaries, and that a post-war arrangement would restrict capital flows.

It is now highly responsible doctrine, in academic and banking circles alike, that a substantial measure of direct control over private capital movements, especially the so-called hot money varieties, will be desirable for most countries not only in the immediately ahead but also in the long run as well. . . . This doctrinal volte-face represents a widespread disillusionment resulting from the destructive behavior of these movements in the interwar years. (Arthur Bloomfield, 1946, quoted in Abdelal 2007: 45)

Keynes agreed that “control of capital movements . . . [would be] a permanent feature of the post-war system” (quoted in Abdelal 2007: 46). These controls were a central part of the post-war social contract (Eichengreen 1996).

But it was also apparent that control over financial flows would require explicit international coordination. “The absence of a high degree of economic collaboration among the leading nations will . . . inevitably result in economic warfare that will be but the prelude and instigator of military warfare on an even vaster scale” (Harry Dexter White, quoted in Pollard 1985). To this end, Franklin Roosevelt organized a conference of 44 allied nations in Bretton Woods, New Hampshire in 1944.

In that conference, the delegates agreed on a system for organizing exchange rates, created the International Monetary Fund to encourage financial cooperation and temporarily offset payment imbalances, and the International Bank for Reconstruction and Development which would expand later into the World Bank Group. (See chapter 5 for a discussion of the Bretton Woods Agreement in operation and its eventual demise.)
In addition to international cooperation in finances, the beggar-thy-neighbor response to the Great Depression highlighted the importance of international cooperation in trade as well. The charter for the International Trade Organization was negotiated in 1948, but it never received approval from the US Congress. The less ambitious General Agreement on Tariffs and Trade (GATT) was signed in 1947. In addition to a series of trade rounds to negotiate multilateral reductions in trade tariffs, GATT offered mechanisms to negotiate and arbitrate trade disputes.

*The US-Centered Spatial Configuration*

I noted at the beginning of this chapter that top incomes became less concentrated in countries around the world, not just in the United States. But no two countries took the same path to lower inequality – some of the discussion above is irrelevant to or contradicts the experience of the ten other countries in Figure 4.2. It is beyond the scope of this project to consider these national experiences independently, but some experiences are shared and highlight one way the global can influence the local.

First, the political response to interstate chaos varied across countries, but the stimulus was broadly the same. Wealth and income distributions were compressed by the destruction of rich-owned assets and strong labor demand. Second, faced with the bedlam of war and depression, “the answer adopted more or less everywhere was for government to take over, in effect to nationalize risk” (Ferguson 2009: 207). Third, there was some institutional convergence after the war as the institutional innovations of the victors were adopted elsewhere. For example, the Japanese Advisory Council for Social Security in 1949 acknowledged the influence of the British welfare state, and the impetus was described by an observer as “Beveridge for the Japanese”
Likewise, the Holding Company Liquidation Commission had the express purpose of liquidating the largest holding companies in Japan, much as the Sherman and Clayton antitrust acts had sought to do in the United States.

Finally, institutional diffusion was not left to chance. Through December 7, 1941 the physical distance between the United States and the chaos enveloping Europe was a major geopolitical advantage that it exploited wholeheartedly. The United States had long been a leader in protectionist economic policies; the Smoot-Hawley Tariff was particularly egregious. Though its own political leadership played a lead role in the organization of the League of Nations, the United States never ratified the covenant, primarily because it included an obligation to assist members beset by external aggression. The United States refused to formally engage in either world war until there was an attack on its citizens. And the United Kingdom was servicing war debt to the United States until 2006 despite fighting on the same side.

During World War II, though, the United States stepped into its new hegemonic role (“the power associated with dominance expanded by the exercise of ‘intellectual and moral leadership’” [Arrighi 1994/2010: 29]) and began to actively structure the post-war global economy. In addition to various international institutions, the United States took an active role in domestic politics. For example, one of the objectives of the post-war American occupation of Japan was “to replace a feudal economy by a welfare economy” (Ball 1948: 15, quoted in Ferguson 2009). As with the British-centered system, participation in the postwar global economy required that one play by the rules.
The Marshall Plan served a dual function. On one hand, it was a tool to encourage reconstruction in Europe to prevent the spread of communism. On the other, the asymmetry between the wealth and cohesiveness of the United States and fragmentation and poverty of foreign markets created the “roots of the impasse which after the Second World War prevented the recycling liquidity back into the expansion of world trade and production. Eventually, the impasse was broken by the ‘invention’ of the Cold War” (Arrighi 1994/2010: 304).

**Conclusion**

The 20th century, industrial global economy was centered on a 19th century colonial empire. The mismatch between industrial capacity and inputs, despite the theoretical emphasis on free trade, was exacerbated by space and politics. Powerful people campaigned to open borders to trade and keep the world on gold (Frieden 2007), but World War I and then the Great Depression made the purist neoclassical argument untenable.

The 1930s saw a fit of experimentation of national political and economic models but with some consistent themes, e.g., increased state capacity and economic intervention. This Cambrian explosion was followed by a Darwinian paring in the 1940s. The United States emerged from war as the center of global economic activity by any conceivable measure (see chapter 5) and, with its allies, sought to construct a new world economic order, a moderated, negotiated and arbitrated version of British liberalism based on the “neoclassical synthesis”: Keynesian macroeconomics and neoclassical microeconomics. Neoclassical free markets were the ideal, but state
moderation was essential when local markets were scaled up, especially internationally.

Income inequality fell in countries around the world, but the new economic ideology rarely recognized egalitarianism as a driving principle. War and the Great Depression destroyed assets owned disproportionately by the rich, while war drove up the demand for labor, especially less-skilled labor, but the resulting compression of income and wealth would have been temporary if the economic fundamentals had not changed.

The new inequality status quo, to some degree, reflected the rejection of Say’s law – that supply creates its own demand. Escaping the Great Depression required more than reducing wages to again spur production. Instead, growth was driven by demand, not supply. Because the state could not secure foreign demand, it focused on the domestic variety. Transfers and make-work projects funded demand during a recession. Higher, secure wages, matched by investments in human and physical capital and infrastructure to increase productivity, were key to a strong, stable economy.

Lower inequality, though, cannot be explained completely by the economic paradigm shift. Labor was empowered by war and the political instability of the 1930s, and slower growth in the supply of labor related to industrialization and World War I. This translated into a political and legal environment that was more conducive to collective action (with an implicit, and in the late 1940s explicit, condition that organized labor exercise a certain degree of patriotism). Organized labor was
fundamental in securing wage growth tied to gains in productivity, and the gains of economic growth were shared proportionately through the 1950s and 1960s.

The Keynesian inequality regime had a limited shelf life. The levers that Washington used to execute Keynesian adjustments broke in the 1970s. The material and legal foundations of organized labor began to erode almost immediately. And the new, post-war world order collapsed as the world-economy evolved. The financial inequality regime took form in the 1980s during the Reagan Administration.
Chapter 5: The Financial Inequality Regime

Income inequality in the United States was relatively low for more than two decades after World War II. By the 1970s, though, the institutional arrangements that had defined the post-war low inequality regime became untenable. No longer could the United States profitably be the world’s workshop. Attention shifted from maximizing productivity, employment and demand, and financial stability, to freeing financial capital flows, reducing costs, and monetary economic policy. As a result of these changes, the financial services industry was able to extend its influence in the economy, alter investment strategies, and grab a higher share of business profits. Corporate management profited from weaker labor laws, financial regulation and corporate governance. The net impact is that incomes and wages for many Americans have been stagnant for decades while those at the top of the distribution have enjoyed rising profits and compensation. This shift from low inequality to rising inequality in the United States, the beginning of the financial inequality regime, correlates with the transition from material to financial expansion in Arrighi’s most recent cycle of accumulation (see Figure 0.3).

The standard approach for explaining rising income inequality in the United States since the 1970s is to identify a covariate of inequality change – e.g., economic restructuring, technical change, financialization – and use some form of decomposition or correlation analysis to estimate its relative contribution to rising inequality.

My approach is fundamentally different. I focus on the interrelationships between these developments and argue that they are components of a single transition, a
transition with economic, political, even ideological ramifications. After World War II, domestic and international institutional arrangements reflected a global economy with the United States firmly in the center. These arrangements broke down as competition from Europe and Japan intensified and more peripheral countries pushed back against American hegemony. In the 1970s, the United States was beset by high inflation and high unemployment. Employment and wages in manufacturing began to decline. Economic uncertainty spawned a neoliberal revolution that Ronald Reagan rode to the White House in 1980 (by promising that his office was not the solution to what ailed the nation). Finance (among other industries) was deregulated, federal income taxes slashed, and union membership fell dramatically. Economic restructuring, technical change, deregulation and lower top marginal tax rates are among the interrelated pieces of the new equilibrium born in the 1980s that has allowed the size distribution of incomes to grow. These should not be treated, ultimately, as independent explanatory variables because they are each key pieces of a broader narrative.

This argument is similar to that offered by Hacker and Pierson (2010): those at the top of the income distribution exercised political power to deregulate finance, slash taxes, and constrain corporate governance and labor unions. The results are similar. But while Hacker and Pierson discount the role of “globalization” in favor of politics, it is my position that the proposed political shift is embedded in the global-historical transition of the 1970s.

In the next section, “the data”, I take a closer look at the empirical evidence that income inequality is rising in the United States. In “the mechanisms” I discuss the
familiar proximate determinants of rising income inequality in the United States and their interrelationships. In the next three sections, “transformation of American hegemony”, “the crisis”, and “the response”, I place those proximate determinants within a broader global-historical context. I use the Bretton Woods Agreement to represent a spatial configuration of the world-economy constructed in the aftermath of World War II, and discuss its inability to cope when economic and political power shifted outside the United States (or at least beyond the control of the US government). I conclude the chapter with some speculative comments on the end of the financial inequality regime and the American cycle of accumulation, and future directions.

The Data

I noted in the introduction that income inequality in the United States began to rise at some point in the 1970s. The Gini coefficient for household income inequality rose 26% from .326 to .411 between 1974 and 2006 (DeNavas-Walt, et al 2012). To put that in perspective, a Gini of .326 would put the United States today in the same neighborhood as Canada and below Belgium and Switzerland, but instead the US is actually in the vicinity of Georgia, Qatar and Turkey (World Bank 2013).

Incomes have become more concentrated at the top of the distribution. While average incomes for the richest 20% grew by $51,409 (2010 $US) between 1973 and 2010, incomes for the bottom quintile fell (see Figure 5.1). According to estimates from the Congressional Budget Office (2011), which account for occupational benefits (e.g., health insurance) and government transfers, incomes of the richest 1%
grew 275% between 1979 and 2007 versus just 18% for the poorest fifth of American households.

Two mechanisms drove income concentration (CBO 2011). First, all forms of market income (e.g., labor, capital gains, business income) became more concentrated at the top of the distribution. Second, those sources of incomes that were already most concentrated, capital gains and business income, grew faster than other sources during this period.

Using the top income share estimates from Piketty and Saez (2003, updated to 2011), the 95% poorest tax units saw their incomes fall between 1973 and 2008 in constant dollars, and the richest 1% captured 72% of income growth over that period. Estimates from the Census Bureau (2010) are more conservative, but still the bottom 60% of households captured only 10% of income growth between 1973 and 2008.

The income share of the richest 1% of Americans (including capital gains) more than
doubled between 1980 and 2007, from 10.0% to 23.5%, and the experience of the top 10% and top .01% are similar (Piketty and Saez 2003, updated to 2011).

Growing income gaps are not exclusive to the very top of the distribution:

The slowing of the growth of overall wage inequality in the 1990s hides a divergence in the paths of upper-tail (90/50) inequality—which has increased steadily since 1980, even adjusting for changes in labor force composition—and lower-tail (50/10) inequality, which rose sharply in the first half of the 1980s and plateaued or contracted thereafter (Autor, Katz and Kearney 2008: 300).

Inequality in hourly wages increased over 40% between 1973 and 2007 (Western and Rosenfeld 2011). In addition to a significant educational income gap, residual inequality (inequality within groups) grew with wage instability – inequality is higher when earnings fluctuate more wildly – and now represents a significant portion of aggregate wage and earnings inequality (Gottschalk and Moffitt 2009; Shin and Solon 2008). In short, the gap between different kinds of workers is growing, and the gap between similar workers is also growing; about half of this growth is due not to persistent gaps between workers but greater wage uncertainty over an individual’s lifetime.

While inequality in the distribution of incomes and wages stretched at almost every possible point, inequality in the distribution of wealth has been a different story. Wolff (2012) reports that the wealth share of the wealthiest 1% increased marginally from 34.4% to 34.6% between 1969 and 2007, but also noted a more substantial increase in overall wealth inequality as the Gini coefficient rose from .811 to .870 during that period. Kopczuk and Saez (2004) reported lower top wealth shares in 2000 than in 1969. The stability in wealth inequality led Armour, Burkhauser and Larrimore (2013) to argue that if you consider asset appreciation (e.g., rising house
prices) as income, as well as benefits and transfers, income inequality did not increase between 1979 and 2007.

The narrative of stable wealth inequality in the United States became much more complicated with the collapse of the housing bubble in 2007. Non-high net worth (NHNW) Americans are more heavily invested in real estate than their wealthy counterparts. For example, principal residence represents 9.4% of total wealth for the wealthiest 1% and 66.6% for the bottom 60%; financial assets make up 75.7% and 12.0%, respectively (Wolff 2012). Therefore, the “creation” of new wealth in the housing bubble disproportionately buoyed the total net worth of NHNW Americans. But median household wealth fell 47% between 2007 and 2010 and reached levels last seen in 1969 (Wolff 2012). The effect of falling house prices on wealth inequality was initially offset by losses in equity markets, but since 2009 stock prices have recovered while housing prices fell further before stabilizing (Federal Housing Financing Agency 2013; Federal Reserve Bank of St. Louis 2013).

The Mechanisms

Beginning in the 1980s, rising income inequality in the United States has inspired an academic industry searching for an explanation. Instead of evaluating the relative merits of the proposed explanations, I highlight a few key variables, discuss their relationships with income inequality and, more important, discuss their relationships to one another. It is my argument that these are variables are not competing explanations, but principal characters in a larger narrative of global economic change.

Economic Restructuring/Deunionization/SBTC. Employment in manufacturing in the United States increased 85% between the beginning of 1939 and the end of 1943, and
then fell 28% to 1946 as the nation disarmed. Manufacturing employment would again approach its war-era peak in 1953 as the US military headed to Korea, but would not surpass the 16.6 million employed in 1943 until 1965. It reached a new peak in 1969, fluctuated through the 1970s, and then began a fairly consistent descent in the 1980s. In 2009, employment in manufacturing would decline to levels last seen in the early 1940s before the war machine reached its peak capacity (BLS 2013; see Figure 5.2).

There are two, seemingly contradictory explanations for falling employment in manufacturing. First, workers in manufacturing today are more productive, so fewer are needed to match past production. Over the last three decades, productivity growth in manufacturing has outpaced gains in the economy at large, so even as employment in manufacturing as a share of total employment has fallen, total production in manufacturing as a share of GDP has been roughly constant over the last half of the
20th century (BLS 2013, productivity in non-farm business sector; Mishel 1989; Morris and Western 1999).

Demand for manufactured goods has declined relative to services in the United States and other rich countries, but has grown globally (Brauer 2004). American economic domination at the end of World War II translated into high demand for American manufactured exports (Chevan and Stokes 2000), but increased productivity since has not equated with a growing demand for US manufacturing exports – the share of US manufacturing of the global total has fallen from around a third in the early 1980s to a fifth in 2010 (Levinson 2014), and China has now surpassed the United States in value added in manufacturing (World Bank 2013). In short, global demand for manufactured goods has increased, productivity in manufacturing has increased in the United States, but employment in manufacturing has declined along with US manufacturing as a share of the global total.

Meanwhile, wages in manufacturing have stalled. Between the end of 1945 and the beginning of 1973, average constant wages (adjusted for inflation) in manufacturing almost doubled ($4.67 to $9.11 in 1980 US dollars). If we recognize that an average of 30% of workers between 1945 and 1970 were employed in manufacturing, wage gains in the sector represented a major source of income growth for the middle deciles of the US income distribution. But as of September 2013 average wages have fallen for four decades in constant terms (BLS 2013, author’s calculations) despite substantial productivity gains during that period (see Figure 5.3).
The logic correlating manufacturing employment with income inequality traces back to Simon Kuznets (1955). As we noted in chapter 4, Kuznets linked inequality trends in the United States, Germany and the United Kingdom to industrialization and urbanization. In turn, when inequality began to rise, Harrison and Bluestone (1990) pointed to deindustrialization and the rise of service employment.

A relationship between economic structure – the distribution of workers, capital and investment by industry – and inequality rests on variations across industries.

Economic restructuring is important because agricultural and natural resource jobs are fundamentally different from manufacturing jobs, which in turn are fundamentally different from service jobs. Different industries have different wage structures and different work schedules for their employees, require different levels and types of education, differ in the types of relationships that exist between owners and workers, and vary in the proportion of the workforce that is either male or female (Albrecht and Albrecht 2009: 520).

Unlike the Kuznets’ model, which emphasizes sector dualism, the economic restructuring argument focuses on higher levels of inequality in services than in
manufacturing. Manufacturing employment was a source of middle-class incomes for a mass of high school educated workers, but in a post-industrial economy, these workers are forced into low-wage service jobs (Morris and Western 1999).

Research empirically connecting rising income inequality to economic restructuring – which includes the rising employment in the service sector with deindustrialization – has returned mixed results (Raffalovich 1990; Tilly et al 1986). Consequently, economic restructuring as an explanation for rising inequality has gone out of favor in some academic circles. I propose that the empirical link is weak in some studies for two reasons. First, to truly measure the impact of economic restructuring we must compare real inequality change against a counterfactual history sans economic restructuring. But there are no empirical and very little theoretical grounds on which we can base this counterfactual scenario. For example, if we allow for employment in manufacturing to continue expanding, is that because manufacturing firms are investing in labor and not capital, reducing productivity gains? Or is the United States able to grab a larger share of the global market for manufactured goods? If so, how? The problem with comparing against a historical counterfactual is that, because it did not occur, it is inherently illogical.

We can create meaningful comparison groups if we divide the United States into regions. This approach has identified large effects of deindustrialization on inequality (Albrecht and Albrecht 2009). Unfortunately, we cannot scale up these results to draw conclusions for the United States as a whole.

Second, and related to the first, economic restructuring is a part of a broader constellation of changes; it is nonsensical to treat it as a truly independent variable.
Instead of seeking to divvy out responsibility for rising income inequality to sets of phenomena – e.g., economic restructuring, skill-biased technical change and the skill/unskilled wage gap, financial deregulation and the financialization of the economy, tax policy – I will emphasize the interrelationships between these phenomena and their net impact on the distribution of incomes across the United States. For example, Western and Rosenfeld (2011) empirically link de-unionization with rising wage inequality between 1973 and 2007. As I discussed in chapter 4, unions can impact wages for member and nonunion workers by influencing policy, investment schedules, wage norms and pushing employers to use higher wages to bribe workers away from organizing. The decline of organized labor also reduced the influence of labor on management (Western and Rosenfeld 2011). With that constraint removed, changes in executive regulation and increased instability from financial speculation pushed up CEO compensation (Hoskisson, Castleton and Withers 2009).

Though union membership began to decline as a share of total employment earlier, researchers often point to 1981, and specifically Reagan’s victory over PATCO (Professional Air Traffic Controllers Organization), as a major turning point (Sherman and Voss 2000; Levy and Temin 2007). The assault on organized labor was an intentional policy response to the post-Bretton Woods crisis; employers used corporate donations to influence legislators against pro-labor legislation (Hacker and Pierson 2010). But it also reflected the associated decline of manufacturing. Union coverage in private sector manufacturing is consistently almost twice that in the private sector generally (Hirsch and Macpherson 2003). Competition from Europe
and Japan was a direct assault on the stronghold of organized labor in the United States.

Causality moved in both directions – declining profits in manufacturing weakened US labor unions, but the political attack on unions reduced union coverage in manufacturing and, ultimately, wages and employment in that sector. By way of counter-example, a more militant labor movement in continental Europe in the late 1960s and 1970s influenced policy that, in turn, led to different patterns of investment (Acemoglu 1998). In short, the development and adoption of new technologies is not an external development, it is institutionally constrained.

In other words, the direction of technical change is institutionally constrained. New technologies complement some workers, so they become more productive/efficient, while making other workers/tasks redundant. In other words, technical change can be biased. Economists, in particular, associate rising inequality, especially in the 1980s, with skill-biased technical change (SBTC). Over the last half of the 20th century, technical change in the United States was skill biased, but over the first half, as noted in chapter 4, it was not (Williamson and Lindert 1980; Autor, Katz and Kearney 2008).

The bias of technical change is generally treated as an exogenous condition of economic life, but research in directed technical change emphasizes that technical change and innovation/diffusion more broadly, like all other economic activities, respond to conditions of supply and demand (as discussed in chapter 1). Specifically, Acemoglu (1998) argues that the shift to SBTC in the 1970s in the United States reflected the growing supply of relatively cheap (both in large supply and
young/relatively inexperienced) college-educated workers. Specifically, and consistent with Goldin and Katz (2010), he argues that technical change became skill-biased in the decades before, but only in the 1980s, when the growth in the supply of college educated workers began to slow did the wage gap between more and less-skilled workers expand.

But Acemoglu (1998) also highlights that supply and demand are not just about raw numbers. As the case of continental Europe highlights, labor market rigidities, unbalanced tax policies and other regulations can influence research, development and capital investment, and can thus influence the direction of technical change. Though continental Europe generally had access to the same technologies driving SBTC in the United States and United Kingdom, inequality stayed low in these countries for several more decades (for the most part).

Economic restructuring can also be tied to changes in family structure and the feminization of poverty (Wilson 1987):

The decline of manufacturing, the loss of jobs to the suburbs, and the rise of low-wage service-sector employment dramatically reduce the number of inner-city jobs that pay wages sufficient to support a family. This situation leads, in turn, to high rates of unemployment and underemployment, and to shrinkage in the pool of male household heads financially able to support a family. Marriage thus becomes less attractive and less available to poor women, unwed childbearing increases, and female-headed families proliferate. (Albrecht et al 2000)

Eggers and Massey (1992) and Albrecht et al (2000) found strong support for an association between structural transformation and increased levels of poverty, with family structure as a mediating variable, for black urban populations and nonurban populations, respectively.
Global competition created a sense of economic crisis, particularly in the power centers of organized labor. But as noted by Bluestone and Harrison (1984), the critical factor that would lead ultimately to rising inequality was the intentional response of firms. For decades, manufacturing had been a source of rising wages for millions of American workers, and those benefits spilled over to workers in other sectors. But in the post-Bretton Woods world, this arrangement was perceived by employers as excessively rigid and costly. They pulled investment out of permanent labor – wages in manufacturing, especially, stagnated as a result – and pushed for greater flexibility in labor and financial markets. We now turn our attention to those financial markets.

Deregulation and Financialization. Another approach to explaining growing inequality in the United States, which has gained momentum since the credit crunch in 2007 and 2008, is to point to the financialization of the American economy (Tomaskovic-Devey and Lin 2011; Hacker and Pierson 2010; Kenworthy 2010; Philippon and Reshef 2009; Rauh and Kaplan 2010; Sum et al 2008; Crotty 2009). Financialization of the economy has two components: the growth of the financial services sector and, relatedly, the financialization of non-financial firms (Tomaskovic-Devey and Lin 2011).

Arrighi (1994/2010) and others with a more global perspective often highlight a financialization of the global economy beginning in the dusk of the 1960s, with the move to floating exchange rates, the explosion of Eurocurrency markets and currency exchanges, and growing sovereign debt. For the United States alone, while there are obvious signs of financialization through the entire post-war period, the rate of
financialization jumped in the 1980s. Financialization accelerated as firms sought external investment opportunities (financial expansion) and the state lost, intentionally or unintentionally, the capacity to regulate financial flows and products; that is to say, financial markets became effectively deregulated.

In the United States, financial regulation at its height in the 1960s had three major components. The first were state laws that prevented intrastate banking. Only 12 states allowed unrestricted statewide branching in 1970. The other 38 were deregulated by 1994 (Strahan 2002). The second were national regulations that prohibited the integration of particular activities within certain firms – e.g., commercial and investment banking (Glass-Steagall, 1933), and commercial banking and insurance (Bank Holding Act, 1956). The project to remove these restrictions began in earnest in 1987 and was complete in 1999 (Financial Services Modernization Act; Philippon and Reshef 2009). The third major form of regulation was interest rate ceilings, introduced with the Banking Reform Act of 1933 and removed between 1980 and 1984 (Philippon and Reshef 2009; Strahan 2002).

The latter two were introduced in response to the Great Depression and were oriented towards greater financial stability by 1) limiting contagion and 2) controlling capital flows. State banking restrictions had their own, provincial logic, but also contributed in limiting the concentration of financial power. Combined with the stabilizing role of the Federal Deposit Insurance Company (FDIC), these regulations had an impressive track record.

In turn, deregulation allowed financial firms to expand and diversify, and for financial power to concentrate in a few firms (Davis 2009). Given these potential
Deregulation of financial activities combined with new sources of liquid financial capital. Federal Reserve chairman Paul Volcker raised interest rates as the 1970s came to a close, which stemmed inflation but also attracted capital flows to US banks. Reagan launched a wave of debt spending, somewhat reminiscent of the arms race launched a century earlier that ended the Long Depression, funded by high interest treasury bonds. And the migration of the center of global financial power from Washington to New York was complete.

The impact of this financialization of the US economy on income (and wealth) inequality was twofold, direct and indirect. Directly, financial profits and compensation in financial services ballooned in the 1980s with the help of financial deregulation and the global financial surge (Philippon and Reshef 2009; Hacker and Pierson 2010). For example, the number of US billionaires in finance increased from 8 to 104 between 1987 and 2012, or 8 to 52 in real terms (1987 US$). The percent of US billionaires in finance almost doubled from 13 to 25% over that period (Forbes
Lin and Tomaskovic-Devey (2013) estimate that between $5.8 to $6.6 trillion have been transferred to the financial sector since 1980. Profits in FIRE (finance, insurance and real estate) surged in the 1980s and 1990s (see Figure 5.4).

Figure 5.4 FIRE Share of Corporate Profits

Source: BLS 2013; Notes: Finance, Insurance and Real Estate (FIRE); corporate profits in FIRE as a percent of all corporate profits

Financial expansions have historically been associated with rising inequality in the global center of finance (Phillips 1993). One reason is that finance is scalable; while a manufacturing firm must employ more workers to expand production, a single financial trader can execute a deal of any size. While concentrating capital in the United States was a competitive advantage during the material expansion, during a financial expansion debt spending can force out domestic production:
“Financialization appears to have crowded out capital investment in real productive assets” (Tomaskovic-Devey and Lin 2011: 546).

The indirect relationship between financialization and inequality is more complex, but fundamentally, financialization increases risk by shortening evaluation periods. The financial sector is in the unique position of being able to profit from uncertainty and instability, and it therefore has few incentives to pursue stability (Arrighi 1994/2010; Lewis 1990; this problem has been perversely multiplied by the moral hazard of financial firms that are “too-big-to-fail”). Corporate raiders and stock market speculation added risk and instability to the economy (Crotty 2009) by shortening the time frame within which producers, especially the publicly traded variety, operated. Publicly traded companies became less interested in the long-term, invested relationships with labor that had prompted the Treaty of Detroit (Acemoglu 2002). With less than half of corporate free cash going to reinvestment since 1980 (Davis 2009), investors are not invested (in the non-financial sense of the term) in firms.

This transition is often associated with the shareholder value model. Shareholders are the owners of the firm. The goal of the firm is to generate profits to enrich its owners – the shareholders. But the relationship between shareholder and firm, especially minority shareholders, is fundamentally different for a publicly traded company; they tend to respond to shorter time frames.

In response to different incentives, management employs a different set of strategies: “Shareholder value strategies, such as mergers, layoffs, and investments in labor saving technology, led to reduced employment, particularly for unionized
workers, but not to increased profitability” (Tomaskovic-Devey and Lin 2011: 546; Fligstein and Shin 2007). In other words, the relationships between financialization, economic restructuring, and deunionization are reciprocal, and rising inequality is a product of their net effect on the distribution of wages and incomes. On one hand, wages for male workers have been stagnant for four decades. On the other, compensation in finance and to non-financial management has exploded - most of the gains have come in the form of stock and stock options (Frydman and Jentor 2010), a choice influenced, in part, by the tax code.

Tax Policy. One half of the Reagan administration’s neoliberal assault was on government regulation – setting markets free. The other half was an assault on government fiscal policy. From this perspective, reducing spending was a winner on both ends. On one hand, government spending was worse than ineffective: Milton Friedman (1975) once said, “I say thank God for government waste. If government is doing bad things, it's only the waste that prevents the harm from being greater.” On the other, lower spending, and smaller government generally, meant that there was less need to collect market distorting taxes.

Top marginal tax rates topped 90% in the 1940s and 1950s. Between 1963 and 1965 the top rate fell to 70%. Then in 1981, only seven months after taking office, Reagan signed the Economic Recovery Tax Act into law. The law cut individual top marginal tax rates from 70% to 50% and reduced estate and corporate taxes. The Tax Reform Act of 1986 again reduced top marginal tax rates from 50% to 28%.

While illuminating, a focus on top marginal tax rates is incomplete. They affect only a small handful of individuals and only a percentage of their incomes – and both
of those values changed over time as well. Beginning only in the 1980s were tax brackets adjusted for inflation, so the scale would slide over time. And top marginal rates do not account for deductions and exclusions. To account for these limitations, Saez, Slemrod, and Giertz (2012) report marginal tax rates by income group. For example, the tax rate for the richest 1% of Americans hovered around 50% between 1960 and 1980, but then fell from 54% to 28% between 1980 and 1988. The correlations with top income shares is notable. The income share of this group (the richest 1%) held steady around 8% for two decades before the 1980s, then jumped from 8% to 13% as top tax rates fell (Piketty and Saez 2003).

In response to tax cuts in the 1980s, a large literature in economics sought to estimate the economic impact of these cuts, and taxes in general, on federal revenues (a relationship generally described in terms of the Laffer curve from chapter 1) and incomes generally. A related literature looked at the relationship between tax rates and market incomes (incomes before taxes and transfers). If incomes at the top of the distribution are taxed at a very high rate, firms get more return on their investment on incomes further down the distribution. During the 1980s, the correlation between top marginal tax rates and income concentration is striking, and more sophisticated approaches also find a significant relationship between top tax rates, income concentration (Saez, Slemrod, and Giertz 2012) and distributional inequality (e.g., Gini coefficient; Altig and Carlstrom 1999): “Concretely, the two phenomena are perfectly correlated: the countries with the largest decreases in their top tax rates are also the countries where the top earners’ share of national income has increased the
most (especially when it comes to the remuneration of executives of large firms)” (Piketty 2014: 521).

In summary, a host of related changes – deunionization, technological change, economic financialization, declining top marginal tax rates, etc. – allowed incomes to concentrate at the top of the distribution. This changes, though, did not occur in isolation. In the next section of this chapter, I detail the contradictions in the post-war spatial configuration that led ultimately to crisis, space for institutional reform, a reconfiguration of the world-economy, and rising inequality in the United States.

**Transformation of American Hegemony**

US global hegemony after World War II was built on three interdependent pillars – military, economic and financial power. The United States had the world’s most destructive military force – especially after the Korean war-inspired re-mobilization (OMB 2014: 185), and also had a physical presence around the world: “no state had previously based its own troops on the sovereign territory of other states in such extensive numbers for so long a peacetime period” (Krasner 1988: 21).

The gross domestic product of the United States in Geary-Khamis dollars (adjusted for purchasing power) peaked over 40% of the global total in 1945 (Bolt and van Zanden 2013; author’s calculations; see Figure 5.2). Assuming a higher cost of commodities in the United States than the global average, US real GDP probably topped 50% of the global total in that year. If we combine the 12 countries of Western Europe with consistent coverage in the Maddison data (Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, and the UK), their combined GDP was 3.4 times greater than in the United States in 1870, but
the United States would surpass that consortium in total production during World War II by about 60% and, after some convergence in the 1950s, would settle at around par in the 1960s (Figure 5.5).

Finally, the United States, through World Wars I and II, assumed the role of the world’s banker from the United Kingdom. It printed the de facto global currency (through the Bretton Woods Agreement), it controlled more than half of the world’s gold reserves after World War II, and, in 1950, two-thirds of global monetary reserves (Eichengreen 1996: 114).

The three pillars supported one another and also conferred a degree of moral and intellectual authority to the United States. Popular examples of their interdependence are the military industrial complex – the Korean and Cold War buildups helped resolve a liquidity issue from the demobilization after World War II (Arrighi 1994/2010) – interstate loans that carried political conditions, and economic sanctions.

Figure 5.5 Ratio of GDP of Western Europe to the United States

Source: Bolt and van Zanden 2013; Notes: "Western European" includes Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, and the UK
for bad behavior. I discussed in chapter 4 how the United States exploited its military, financial and economic advantage to shape the emerging spatial configuration.

But less than three decades later, in the 1970s, American hegemony was challenged on all three fronts. The French forced Nixon to scrap the Bretton Woods Agreement, and financial power shifted from Washington to New York. War in Vietnam critically wounded the perception of US military might, both in terms of capacity and as a moral force, domestically and internationally. US economic hegemony was fundamentally challenged with the reemergence of Western Europe and Japan as economic powers par excellence. Through the 1970s inflation, unemployment and energy prices spiked while the Dow Jones Industrial Average fell in real terms.

Despite these challenges, by the 1990s the United States was again firmly entrenched as the global hegemon, perhaps more so than ever with the fall of the Soviet Union. New York was the center of global financial activity. The United States still had the world’s largest economy, almost by an order of magnitude (in the mid-1990s, only Japan had a GDP more than 50% of the US, only Japan’s and Germany’s surpassed 25%, and only five other countries had a GDP greater than 10% of that in the US [World Bank 2013]), and it seemed to be in very good shape by the most popular indicators: unemployment and inflation fell back to healthy levels and the Dow Jones Industrial Average surged. And in a military exercise aimed in part at keeping a lid on energy prices, the US demonstrated in the Persian Gulf War that, when need be, it could still depend on raw physical force.
In this context, the political scientist Francis Fukuyama famously argued in 1992, in stark contrast to my thesis, that we had reached the “end point of mankind’s ideological evolution”:

a remarkable consensus concerning the legitimacy of liberal democracy as a system of government had emerged throughout the world over the past few years, as it conquered rival ideologies like hereditary monarchy, fascism, and most recently communism. . . . That is, while earlier forms of government were characterised by grave defects and irrationalities that led to their eventual collapse, liberal democracy was arguably free from such fundamental internal contradictions (Fukuyama 1992: Introduction).

The perception of arrival is the natural product of a newly emerging configuration, a confluence of political, economic and institutional forms that seem free of the “irrationalities” that are exposed through time.

But the structures of the world-economy and of the United States were fundamentally transformed between the 1960s and the 1990s, and the new “financial” regime that emerged through the 1980s reflected both an adaptation to and an active creation of these new realities. In the next section I explore the contradictions in the Keynesian Regime that were exposed between the 1960s and 1970s. I focus on the structure of US economic and financial hegemony in a world no longer defined by the scars of world war and economic crisis in the first half of the 20th century. I then turn to the political, economic and institutional adaptations to crisis. Finally, I explore the relationship between these adaptations and economic inequality in the United States – rising income inequality, a growing wage gap between more and less-educated workers, but stable wealth inequality.
The Crisis

Though perhaps more symbolic than mechanical, the collapse of the Bretton Woods Agreement between 1971 and 1973 succinctly captures the structural transformations of the global economy and ensuing crisis that would force a shift from the Keynesian to the financial inequality regime. Through the Agreement established in 1944, foreign currencies were pegged to the US dollar, and dollars were redeemable for gold at $35 per ounce. This system was different from the gold standard of the past in three critical ways. First, pegged exchange rates were adjustable in certain, specific situations. Second, states retained some control of international capital flows. Third, the International Monetary Fund (IMF) was created as a form of international FDIC; it could extend financing to countries at the wrong end of a balance-of-payments imbalance (Eichengreen 1996).

This system depended fundamentally on the United States as a financial anchor – the strength of the dollar and the ability of the Treasury to redeem dollars for gold. The latter seemed a safe bet given the size of its reserves (Eichengreen 1996). This meant that, barring a major and sustained turn in its balance of payments, the United States would be able to keep the value of the dollar fixed relative to gold, and a major and sustained turn was unlikely in the foreseeable future given the manufacturing prowess of the United States. In short, by building on the unique economic and financial position of the United States, the Bretton Woods System established a less rigid, more robust gold standard that positioned the US dollar as the de facto global currency.
The challenge in the late 1940s was to bring Europe back up to speed so they could re-establish currency convertibility without absorbing a stability-threatening current account deficit. “Postwar Europe had immense unsatisfied demands for foodstuffs, capital goods, and other merchandise produced in the United States and only limited capacity to produce goods for export; its consolidated trade deficit with the rest of the world rose to $5.8 billion in 1946 and $7.5 billion in 1947” (Eichengreen 1996: 98).

Buying goods from the United States that Europeans could not produce domestically required US dollars. Dollars could be obtained by selling local currency or gold. If trade was balanced, that local currency and gold would be returned through the sale of exports. But in order to maintain a modest standard of living (and social harmony), European currencies were overvalued. As a result, imports were attractive and exports were uncompetitive.

Near the end of 1949 twenty-four countries devalued their currencies, and in 1950 war broke out in Korea. The former made European goods more attractive to international buyers, particularly the United States, and the latter increased US demand for goods both domestic and international. Over the next year the US current-account surplus was cut in half (Eichengreen 1996). It took almost another decade, but on the last day of 1958, member countries of the European Payment Union restored currency convertibility and, finally, the Bretton Woods System swung into full effect. In other words, while the United States was still the global economic hegemon, it was no longer the overwhelming economic bully of the late 1940s.
As noted in the last chapter, countries around the world built reserves with both gold and US dollars under the fundamental logic that the latter was convertible into the former at a fixed rate. And the system was stable as long as there was no question that the US gold reserves were sufficient to match foreign monetary liabilities – the value in gold of US dollars in foreign currency reserves. The Belgian economist Robert Triffin noted that this arrangement had the potential for a liquidity crisis that very much resembled a bank run: fear that the US could not meet its liabilities would incite a scramble to convert dollars into gold. He theorized that the United States would restrict the supply of dollars to inflate its value and encourage central banks to hold their dollars, but that other countries would respond in kind, pursuing deflationary policies. The result would be a global liquidity crisis.

Triffin was wrong about the sequence of events that would lead to the collapse of the Bretton Woods System, an issue we will return to shortly, but the fundamental logic was correct: US foreign liabilities surged while a negatively trending current account put downward pressure on the value of the dollar (Arrighi 1994/2010). In short, the conditions that justified the the Bretton Woods System – US economic hegemony and massive gold and foreign currency reserves – were undermined. The US share of total global production fell to around 22% by 1970 (Bolt and van Zanden 2013, author’s calculations) and US gold reserves fell from $20 billion at the end of the 1950s to half that by 1970 (Walter 1991).

Through the 1960s the United States and other members of the Bretton Woods System launched a series of efforts to save gold convertibility. Domestically, these included capital controls, increasing commercial staffs in US embassies to boost
exports, even rules preventing US citizens from collecting gold coins. “The array of devices to which the Kennedy and Johnson administration resorted became positively embarrassing. . . . Dealing with the causes required reforming the international system in a way that diminished the dollar’s reserve-currency role, something the United States was still unwilling to contemplate” (Eicengreen 1996: 129).

The United States was not alone in its desires to save the existing system. In addition to a number of bilateral arrangements to buy foreign currency to maintain workable exchange rates, the London Gold Pool aggregated gold reserves from eight central banks in an effort to defend the price of gold relative to the US dollar through interventions in the London gold market. Through these and other efforts, the system staggered along through the end of the 1960s before finally collapsing in 1971 when Nixon halted the conversion of US dollars for gold. (In 1980, the price of gold hit $615, 18 times the pegged value under the Bretton Woods Agreement [Green 2007].)

It is not my argument that the failing of the Bretton Woods System caused income inequality in the United States to surge for the next four decades. Instead, it is my contention that the failing of the Bretton Woods reflects the end of a post-war material expansion that had sustained the previous low inequality regime in the United States; the Bretton Woods Agreement was a key component of that regime.

So, why did the Bretton Woods System fall apart? The short answer is that it reflected the logic of a postwar global economy. In this geo-historical context, the United States was the economic anchor that could stabilize global finances – it was, after all, almost as large as the rest of the global economy combined. But this arrangement was a historical blip. Western Europe and Japan had the human capital
base, technology, and access to financial capital to rebound quickly. The US
economic advantage, in reality, was organizational:

As long as trade and production in Western European states and in their
former colonies were organized by the mixture of familial and state capitalism
which had emerged out of the disintegration of the nineteenth-century world
market economy, US corporate capital had a decisive competitive advantage
in conquering markets for final outputs and sources of primary inputs through
direct investment and the vertical integration of the intervening sub-processes
of production and exchange. But as an increasing proportion of European and
former colonial trade and production was so conquered and reorganized, the
further expansion of US corporations came to be constrained ever more tightly
by the imposition of organizational barriers to entry that they created for one

The United States both underestimated the extent of the damage in the immediate
aftermath of war and underestimated how quickly these economies could transition
from dependents to competitors.

This transition is made manifest materially in the current account balance. In less
than two decades, the United States went from major net exporter to net importer.

Nevertheless, the phase of material expansion of the 1950s and 1960s
resembled all others in one key respect: its very unfolding resulted in a major
intensification of competitive pressures on each and every governmental and
business organization of the capitalist world-economy and in a consequent
massive withdrawal of money capital from trade and production (Arrighi

Costs of production – rising prices for inputs – increased faster than new sources
of demand were created, and the impact was felt in the 1970s in almost all core
countries (Brenner 2006). Multinational firms and finance sought growth
opportunities outside the United States, so that by 1966 US direct investment totaled
7% of GNP, which ironically led Europeans to fret about an “American challenge”
(Arrighi 1994; Wilkins 1970).
This latter development reflected another critical change in the structure of US global hegemony. In the turn of the decade between the 1940s and 1950s, Washington deployed finance as a political tool, targeting allies and the well behaved. Strict capital controls were imposed in Europe to jimmy rig the Bretton Woods System when fixed exchange rates were nonsensical economically.

Then over the next two decades central banks and politicians in the United States and Europe lost control over financial flows. Finance as a share of US GDP grew. Debt spending in the United States infused the global financial system with more dollars. This combined with petrodollars, institutional investors in the United States and Europe (e.g., pension funds), and revenue from core firms that could not be profitably reinvested (Davis 2009; Tomaskovic-Devey and Lin 2011). Despite explicit efforts in Washington to maintain control, finance was globalized and began to operate outside of political borders. This process was only further accelerated with the death of Bretton Woods, and now with flexible exchange rates, governments outside the United States found themselves free to spend (Arrighi 1994/2010).

This transition is best represented by the expansion of the Eurodollar market. Eurodollars are simply deposits denominated in US dollars but held in banks outside the United States. Until the 1950s, US dollars that made their way outside the United States were generally repatriated either to buy US exports or invested in US money markets. But as the quantity of US dollars internationally grew so did the size of the Eurodollar market. US multinationals, particularly, exploited Euromarkets to acquire US dollars for European (and other international) subsidiaries without the regulations
and transaction costs of domestic sources as Washington sought to crack down on financial flows.

Foreign dollar loans that had previously come under the regulatory guidelines of examination of US government agencies simply moved out of their jurisdictional reach. The result has been the amassing of an immense volume of liquid funds and markets – the world of Eurodollar finance – outside the regulatory authority of any country or agency (Eugene Birnbaum of Chase Manhattan Bank, quoted in Frieden 1987: 85).

The result was a global financial revolution:

What is most striking about the last few decades is the liberalization of capital flows between the major countries and the incredible growth of the Euro-markets, which has averaged about 30 per cent per year since the 1960s. This has so far outstripped the growth of global trade and output that financial flows now utterly dominate real flows between countries in quantitative terms (Walter 1991: 2000, quoted in Arrighi 1994/2010: 208).

Despite these flesh wounds, the Bretton Woods System survived until member states lost faith. The back-breaking straw came from France, who demanded $191 billion in gold in August 1971 (Frum 2000). This development reflects two interrelated structural changes. First, as already noted, the system failed to accommodate changing economic realities. Second, the United States no longer had the financial, economic or, arguably, military strength to impose its will on the economic system. France, and De Gaulle in particular, was never happy with the privileged position the system granted to the United States, but were particularly perturbed when the United States was able to exploit the system to help fund a war in Vietnam when the system had constrained France’s abilities to finance its own military efforts in Indochina (Eichengreen 1996).

Overaccumulation in the core came to a head in the 1970s. On one hand, economic growth in the core was still relatively strong through the decade.
constant dollars, US GDP grew at an annualized rate of 3.2% and GDP in the Euro area at 3.5% (World Bank 2013, author’s calculations). Both figures are higher or on par with those in any decade since, and much of the slowdown from the 1960s can be attributed to the oil embargo in 1973 and 1974.

On the other hand, there are clear signs that the 1970s marked the end of a material expansion. For example, capacity utilization, the ratio of industrial production to potential production, began to fall in the 1970s and continues to fall (see Figure 5.6). After accounting for business cycles, we can identify a clear (and statistically significant) downward trend in the percent of US productive capacity that can be profitably deployed. The Dow Jones Industrial Average was lower in 1982 than it had been in 1966 in nominal terms, and was significantly lower in real terms (see Figure 5.7). Inflation spiked, as did unemployment. More fundamentally, there was a growing sense of uncertainty about future growth prospects.

Figure 5.6 Capacity Utilization, Total Industry, 1967 To 2011

Source: Fed 2013; Trough, peak and total are linear trends. Shaded areas identify periods of economic recession as defined by the National Bureau of Economic Research 2013.
High inflation with high unemployment was particularly troublesome. In the strain of Keynesian economic thought popular in the day, governments could fight high unemployment with spending (which would increase the money supply and cause inflation), such that there was a negative relationship between the two. The portmanteau “stagflation” was coined to describe the scenario in the 1970s in which persistent high unemployment and inflation resided together (see Figure 5.8).
High unemployment had two fundamental causes in the 1970s. The first was a recession brought on by OPEC’s oil embargo. High energy prices increased the cost of doing business. The second was that firms responded to narrowing profit margins, evidenced in the Dow Jones Industrial Average. Firms sought to reduce costs, but wage levels were relatively rigid, so they reduced the cost of labor by employing fewer laborers. Manufacturing alone dropped 3.6 million workers between October 1974 and February 1975 (BLS 2013).

High inflation reflected a few developments. As economic growth slowed, the growth of the money supply outpaced it and costs rose. The money supply was fed by government spending domestically and internationally (Vietnam). In the Bretton Woods world, excess dollars had to be absorbed in foreign reserves, giving the United States more fiscal flexibility. That was no longer the case.

Finally, the OPEC oil embargo was not a random, external event. In a direct challenge to US hegemony, the oil exporting nations called out the US and Europe for supporting Israel. Much of the foreign policy of the United States since has been oriented towards remedying this uncertain dependence on foreign energy. It also highlighted one of the consequences of material expansion and causes of spatial reconfiguration – resource depletion: the United States had once been an oil exporter itself.

These conditions created a sense of crisis in the United States that extended beyond the economic challenges the country was facing. Domestically, the government’s key levers of economic manipulation seemed to break. Internationally, US hegemony was challenged on all three fronts: military prowess, economic power,
and financial control. The result was a new ideological debate, a new set of rules governing economic activity, a new organization of production, and a shift in the distribution of economic activity in the United States by sector. In short, we changed what we make, how we make it, why we make it, and, ultimately, who benefits from it all.

The Response

“The recent stagnation of American incomes and the rise of inequality have their origins in the growth of global competition and specifically in a distinctive array of business strategies adopted by American corporate managers to cope” (Bluestone and Harrison 1984: xii, emphasis added). Economic competition is not an explanatory variable. New competition can prime change, but the exact response is made within the context of the existing economic, political and institutional environment: “the particular content and form of these changes are affected by the historically specific institutional arrangements that define the parameters and options available to social actors” (Prechel 2000: 2).

Broadly, we can understand the response to crisis in the 1970s as a choice between domestic demand and comparative advantage. I noted in chapter 4 that transition to the Keynesian Regime can be summarized as a new emphasis on domestic demand and investment as the means for stable economic growth. Many of the institutional changes of the period were oriented towards securing higher wages for workers, while also paying for those wages by increasing productivity – physical and human capital investments.
On the other hand, as I noted at the beginning of this chapter, the United States emerged from World War II as the global economic hegemon. It exported foodstuffs, manufactured durables and capital around the world, particularly to Europe. Through the Bretton Woods System it was able to maintain a global economic and military presence, at a reduced cost for the latter and at a profit for the former. In other words, it was able to play domestic demand at home while exploiting comparative advantage internationally.

But strong domestic demand was not limited to demanding domestically produced goods. The United States was the largest commercial market in the world, and in the 1960s it became a major net importer (Arrighi 1994/2010). Strong domestic demand came into conflict with an international competitive advantage. The United States was able to spend its way to strong growth through most of the 1960s, but US firms increasingly opted to invest internationally, and spending failed to reign in unemployment in the 1970s. The US seemed to be faced with a choice: pursue a more autarkic, labor-oriented model of growth or reduce costs domestically – both wages and government spending – to retain an international economic and financial competitive advantage and global military presence.

And this brings us to Milton Friedman. Upon his death in 2006, his obituary in *The Economist* proposed that “he was the most influential economist of the second half of the 20th century (Keynes died in 1946), possibly of all of it” (*The Economist* 2006). And the juxtaposition with Keynes is no accident. While Keynes embodied the economic ideology of the postwar material expansion, Friedman headlined the
monetarist movement that became the principal ideological alternative post-Bretton Woods.

The fundamental monetarist argument is that the money supply plays a major role in determining economic output/investment and consumption. The logic stands firmly in contrast to Keynesian-style fiscal policies. While Keynes’ *General Theory* favored government spending in a recession to boost confidence and demand, Friedman and his ilk argued that government spending merely increased the money supply and disrupted the equilibrium in the supply and demand for money (i.e., price inflation).

For decades economists had noted an empirical tradeoff between inflation and unemployment. Government spending increased the supply of money, driving up prices, but was also correlated with falling unemployment rates (known as the Phillips curve). Friedman argued that this association was an illusion. Workers were fooled in the short term into believing that rising wages meant more purchasing power, so they returned to work for lower real wages. Unemployment then rebounded when workers adjusted their expectations and demanded higher wages. On top of that, inflation (and deflation) adds uncertainty and cost to termed contracts because agents must also consider the future value of money as well as the current value of goods/capital.

Extending this concept more broadly, Friedman argued that government management of the economy could only produce illusory benefits. He argued in *Capitalism and Freedom* (1962) that the scope of government action should be limited to enforcing contracts, promoting competition, minimally managing the money supply, and protecting those incapable of doing so themselves (“madman or
child”). Government intervention was not only a threat to economic efficiency but individual freedom broadly defined. "Concentrated power is not rendered harmless by the good intentions of those who create it" (Friedman 1962).

This logic fundamentally contradicted the postwar social contract adopted throughout the West. Government intervention in the postwar era was oriented towards promoting domestic demand by securing higher wages for workers and benefits to those temporarily unemployed or unable to participate in the labor force. But Friedman argued that the government could not succeed. Instead, intervention only sustained economic inefficiencies and led to price inflation, both of which had real economic costs long term. On these grounds, Friedman won the Noble prize for economics in 1976 and Ronald Reagan argued in his first inaugural address in 1981 that “government is not the solution to our problem; government is the problem”

The paradigm shift in economic thought in the 1970s and 1980s was tangible (Tomaskovic-Devey and Lin 2011). Of his own experience, Friedman noted in 1998 that he was now part of “the mainstream of thought, not, as we were 50 years ago, a derided minority” (Friedman and Friedman 1998). That is not to say that Friedman single-handedly redirected American economic thought. Instead, he was the very capable messenger of a transition that reflected a change in the material conditions of economic life.

The regime change of the 1970s and 1980s, in this light, was a product of institutional changes – e.g., deregulation of finance – that were the manifestation of subtle and explicit propaganda/lobbying from economic thinkers, like Friedman, and business interests – e.g., investment banks – but also an economic reorganization
motivated by new forms of competition and guided by the evolving institutional environment. Business interests pointed to rising competition and falling profits and argued that labor costs and regulation were the cause. Guided by the economic philosophies of Friedman and his ilk, they stressed deregulation, tax cuts, declawing labor, and smaller government generally, and in the 1970s and 1980s their pleas were heeded (Harvey 2005).

**Reviewing the Narrative of 20th Century Inequality**

In the 1930s and 1940s, economic and political crisis destroyed financial and physical capital, infrastructure, social networks and organizations, and delegitimized institutions, polities and social ideologies. This laid the groundwork for a spatial reconfiguration and a new phase of material expansion. But the grounds for material expansion were finite. Only in the 1950s did the new configuration really take root, and in the 1960s evidence of the inherent contradictions in the system were becoming obvious.

As competition intensified, financial capital was pulled out of productive enterprises in search of new investment opportunities. The mobility of capital, especially with the dawning of the digital age, meant that financial flows soon dwarfed trade and other processes of globalization. Perhaps more important, these flows were largely beyond the purview of geographically-defined states. Financial agents were able to exploit the inefficiencies and irrationalities of the US-centered Bretton Woods system, and the system became untenable.

In the United States, the failure of the Bretton Woods system and the conceptualization of the world-economic structure it reflected forced a series of tough
choices. Economic uncertainty in the 1970s was a product of the intersection of overaccumulation in the core and attacks on US hegemony more broadly. In this environment a new economic logic, represented best by Friedman’s monetarism, took root that explicitly contradicted the social contract of the last quarter century and offered a solution to the economic malaise.

The new regime that emerged through the 1980s represented a thorough reversal of the Keynesian regime. Taxes on top incomes, labor unions, financial and other corporate regulations were scaled back or scrapped altogether. Weaker labor and shorter investment windows redirected investment strategies and, in turn, technological change; management was incentivized to maximize short-term profits, not long-term productivity.

And the major economic indicators seemed to justify these changes, at least temporarily and for some. The Dow Jones Industrial Average surged through the 1990s, and the United States (or at least those in a position to exploit global economic and financial growth) seemed as on top of the world as ever before. But this model required a sacrifice of “legitimate” demand (funded by income) in the name of lowering costs and controlling prices; domestic consumption was maintained through debt spending financed by the liberated financial services industry.

The 21st Century Narrative

The economic uncertainty of the 1970s was a distant memory to many as the United States reached “the end of history” in the 1990s. In constant terms (adjusted for inflation), the Dow Jones Industrial Average [DJIA] trebled between 1982 and 1995, and then increased 2.7 times to the end of 2000 (Fed 2013, author’s
calculations). GDP per capita in the United States grew at 2.6% per year between 1991 and 2000 in constant terms. The United States even ran a budget surplus between 1998 and 2001; the last budget surplus was 1969, and in only seven other years did government revenues surpass outlays since the Great Depression (OMB 2014: Table 1.1).

But more important than the economic trends was the sense that we had discovered a perfect concoction of economic, political and social institutions in Western liberal democracy (Fukuyama 1992). The term “Washington Consensus” was coined in 1989 to describe a specific set of economic policies to be exported globally, and has been used more broadly in reference to the neoliberal ideal: free markets, monetarist, supply-side economics, and a minimalist state. “During [the 1990s], the world has been under the impression that there was a clear and robust consensus about what a poor country should do to become more prosperous” (Naim 1999).

Of this period, Arrighi (1994/2010) wrote,

the bourgeoisie of the West came to enjoy a belle époque in many ways reminiscent of the “wonderful moment” of the European bourgeoisie eighty years earlier. The most striking similarity between the two belles époques has been the almost complete lack of realization on the part of their beneficiaries that the sudden and unprecedented prosperity that they had come to enjoy did not rest on a resolution of the crisis of accumulation that had preceded the beautiful times. On the contrary, the newly found prosperity rested on a shift of the crisis from one set of relations to another set of relations. It was only a matter of time before the crisis would re-emerge in more troublesome forms. (pg. 334-5)

The euphoria of the 1990s soured in the next decade. The psychological landscape was altered by the terrorist attacks on September 11, 2001. The DJIA lost 33% of its value between the first half of 2001 and September 2002. Over the course of the
The DJIA only briefly topped its 2001 peak for about three cumulative months in 2007 in constant terms; in other words, after almost two decades of astronomical growth, stock market gains were negligible between 2001 and the end of 2013 (Fed 2013). GDP per capita grew at 0.7% per year between 2001 and 2009. After eight total months of economic contraction between 1983 and 2000, the United States experienced 26 months of falling production in the first decade of the 21st century (World Bank 2013; NBER 2013).

An unscientific study of book titles in the current affairs section of a bookstore suggests that the United States and the rest of the West, and their associated economic models, are in perpetual decline, or require dramatic overhauls to repair (following the authors’ diagnoses and prescription, of course). Economic stagnation, uncertainty, and debt, combined with rapid economic growth in China and other large but relatively poor economies, suggest that the current structure of the world-economy is being transformed. In this context we are faced with two pressing questions:

1) Where does the financial crisis of 2007 fit?

The financial crisis was deep, global, centered in the core, and threatened essential market processes. The crisis and ensuing recession inspired a new round of financial regulations, social protest against a system that seemed to enrich a few (e.g., the 1%) at the expense of the rest (the 99%), and measurably lowered income inequality. But income inequality in the United States quickly rebounded, and wealth inequality is probably higher than before.

2) When is the next regime change and where will it take us?
The financial crisis highlighted the fragility of a debt-funded financial expansion. Neoliberal economic orthodoxy stumbled, and the financial services industry is now subject to tighter regulation. There has been a significant shift in the sites of wealth generation and accumulation; China, in particular, has emerged as the global economic dynamo. As with previous transitions, national and international institutional arrangements and relationships are becoming archaic as they struggle to govern a global economy that is being materially reconstructed. The when and where are yet to be determined, but there are plenty of clues that the next inequality regime change is nearing.

If inequality regime change in the United States is truly synchronized with Arrighi’s systemic cycles of accumulation, we should now be looking for a terminal crisis to mark the end of the US cycle of accumulation and the financial inequality regime. In the next section I take a closer look at the financial crisis of 2007: can we interpret a financial crunch with an impact on economic growth and political stability globally as the terminal crisis? While the crisis shared many characteristics with the Great Depression, it more appropriately understood as foreshadow to a terminal crisis. In the final section of this chapter, I note some indicators of a spatial configuration in crisis and speculate on future directions.

The Financial Crisis of 2007-2009. The financial crisis of 2007-2009 was arguably the worst of its kind since the 1930s, and even shared many characteristics with the cataclysmic collapse that launched the Great Depression. Steven Gjerstad and Nobel Laureate Vernon Smith (2009) argued that

[t]he events of the past 10 years have an eerie similarity to the period leading up to the Great Depression. . . .It appears that both the Great Depression and
the current crisis had their origins in excessive consumer debt -- especially mortgage debt -- that was transmitted into the financial sector during a sharp downturn.

For example, housing prices tripled between the end of the 1990s and the peak in 2006, and the ratio of residential mortgage debt to household wealth almost tripled between 1920 and 1929 (.102 to .272; Gjerstad and Smith 2009).

Other similarities may have been precluded by state intervention, motivated by research on the Great Depression. For example, Irving Fischer (1933) argued that the root of the Great Depression was over-indebtedness followed by deflation (making it more difficult for debtors to meet obligations). Friedman and Schwartz (1963) highlighted the constriction of the money supply. Then Federal Reserve Chair Ben Bernanke is a noted scholar of the Great Depression, and the eponymous “Bernanke Doctrine” is oriented towards combatting deflation. As a result, the Federal Reserve devised policy with an eye towards preventing Depression-style deflation.

Fundamentally, though, the Great Depression and financial crisis of the last decade are similar in that both were big and both were global.

In the United States, the financial crisis was intimately linked to falling house prices, which peaked in 2006. Then, between 2007 and 2012 median household wealth – about two thirds of which is composed of the value of the principal residence – fell 47% (Wolff 2012). The recession spread to other industries as those exposed to losses in the mortgage market were unable to meet their other obligations. Credit risk, measured as the gap between the interest rate on interbank loans (three month LIBOR) and short-term US government debt (three month T-Bill), tripled between July and August 2007 and then doubled again to September 2008 (Fed
Firms dependent on short term loans (commercial paper) to cover expenses were forced to scale back or halt operations. Between mid-2007 and February 2009, the value of the Dow Jones Industrial fell more than 50% (by way of comparison, the DJIA fell about 75% between the second half of 1929 and the end of 1931; Fed 2013).

The impact of the financial crisis and ensuing recession was not isolated to the United States. Global market capitalization (the total value of issued shares of publicly traded companies) fell more than 50% between October 2007 and February 2009, losing $39.4 trillion dollars over that period (WFE 2013). GDP growth in high income countries, as defined by the World Bank, ground to a halt in 2008 and fell 3.7% in 2009. On the other hand, GDP growth in middle and low-income countries continued, virtually unabated in the case of low-income countries (see Figure 5.9; World Bank 2013).

Figure 5.9 National GDP Growth (Annual %) By Income Group.

Source: World Bank 2013
In the face of credit constraints, global trade fell 20% by volume between April 2008 and May 2009, and 16% between October 2008 and January 2009 (CPB 2014). The value of global trade fell 23% between July 2008 and February 2009. Even more dramatic was the collapse in the trade of primary commodities. The real value of fuels fell 67% in the second half of 2008, and the real value of other primary commodities fell 40% (see Figure 5.10). In short, the crisis impacted finance, production and trade, rich and poor, but the impact was greater for the former than the latter.

Figure 5.10 World Trade, 2005=100

![Graph showing world trade from 2003 to 2013](image)

Source: CPB 2014

In an effort to unfreeze financial markets in the United States, Congress authorized the expenditure of $700 billion through the Troubled Asset Relief Program (TARP) in October 2008. That number was later reduced to $475 billion. In the wake of the financial meltdown and expensive bailouts, public ire targeted the financial services industry and insufficient federal regulation of finance: the crisis originated in mortgages under conditions that would have been prohibited in the past; it spread to other firms through derivatives that were not regulated and were given top credit ratings despite their shaky foundations; Washington was forced to bail out banks that,
in the new regulatory environment, became “too big to fail”. More generally, leverage (assets purchased with borrowed funds) had ballooned across firms and households in recent decades, multiplying gains in good times and losses in bad.

The most visible manifestation of discontent was the beginning of the Occupy movement in 2011. Particular motives varied from site to site and occupant to occupant, but the movement represented a protest against the concentration of economic and political power in large corporations and the global financial system, with an emphasis on a perceived distinction between an empowered 1% and everyone else (the 99%).

But even before the Occupy movement began in earnest, a sense of class conflict was growing publicly. In a survey conducted by the Pew Research Center, 47% of respondents in 2009 reported seeing strong or very strong conflicts between rich and poor in the United States. That figure jumped to 66% in 2011 (Morin and Motel 2013). In another survey of marketing and public relations executives in the financial services industry, 81% said they were worried about negative public reaction to executive compensation in the financial industry, and more than half said that Occupy Wall Street had a real impact on the business (Benoit 2012). The International Monetary Fund feared “an explosion of social unrest” in the United States and Europe in response to the worst jobs crisis since the 1930s (Evans-Pritchard 2010).

The most substantial policy response to these conditions was the Dodd-Frank Wall Street Reform and Consumer Protection Act. The Obama Administration sought a “sweeping overhaul of the financial regulatory system, a transformation on a scale not seen since the reforms that followed the Great Depression” (Obama 2009). Dodd-
Frank was signed into law in July 2010. On July 16 of that year, the headline in the Wall Street Journal read, “Law Remakes U.S. Financial Landscape,” and the authors claimed that the law represented a “rewrite of rules touching every corner of finance” (Paletta and Lucchetti 2010).

Some major goals of the law were bringing transparency to derivatives (shifting many over-the-counter products to public exchanges), a procedure for winding down failing financial firms, and better regulation of credit rating agencies. Perhaps the most significant change was the “Volker Rule;” the fundamental logic of this rule, formalized only at the end of 2013, is to prohibit deposit-taking firms (commercial banks) from engaging in proprietary trading (investment banking). It is, essentially, a weakened version of Glass-Steagall.

The ideological impact of the recession was more fundamental than angry protestors at the wrong end of a business cycle.

For economists 2008 was a nightmare. The people who teach and research the discipline mocked by Thomas Carlyle, a 19th-century polemicist, as “the dismal science”, not only failed to spot the precipice, many forecast exactly the opposite—a tranquil stability they called the “great moderation” [a phrase attributed particularly to Ben Bernanke]. While the global economy is slowly healing, the subject is still in a state of flux, with students eager to learn what went wrong, but frustrated by what they are taught. (The Economist 2013: “Keynes New Heirs”)

The Economist (2013) explicitly compared the state of 21st century, post-crisis Economics to the 1930s.

In the early 1930s economics was in a terrible state. The global economy was stuck in a rut, and economists could not explain why. Two Britons changed things. In 1933, John Maynard Keynes, an economist at Cambridge University, supplied the raw ingredients: a new theory that explained how deficient demand could lead to persistent recessions and long-term unemployment. The ideas were radical but technical. They really took off when John Hicks, then also at Cambridge, distilled Keynes’s ideas into a
simple model which quickly became the backbone of undergraduate teaching. *(The Economist 2013: “Keynes New Heirs”)*

The ideological evolution at the International Monetary Fund [IMF] has been particularly interesting. Historically, the IMF has been the purveyor of the financial logic of the Washington Consensus to the world. In its mission to promote financial security, the IMF makes assistance to member countries contingent on meeting nominal goals in fiscal austerity; in response to the Asian financial crisis, the IMF also set targets for structural reforms. But in the onset of the financial crisis, then-managing director Dominique Strauss-Kahn (2008) laid out the case for a “Targeted Fiscal Boost.” He argued that monetary policy is still the first line of defense, “But monetary policy may not be enough”. He offered two explanations for this. First, monetary policy may fail because banks “have suffered substantial capital losses and thus want to consolidate their balance sheets and avoid taking on additional risk”. Second, in language that Keynes might have used himself “there is a risk that if a slowdown really takes hold, it will be hard to shake off.”

The pro-fiscal turn received another boost in 2013 when the IMF’s chief economist, Olivier Blanchard, released a working paper that suggested that the IMF has underestimated the fiscal multiplier (Blanchard and Leigh 2013); in other words, government spending can produce gains in the economy that exceed the original outlay. The IMF in its World Economic Outlook lowered some growth projections based on overzealous austerity and suggested that fiscal adjustment in some cases should be delayed (IMF 2012).

It is in this environment that the Obama administration, in its 2009 spending request, and as quoted at the beginning of this dissertation, identified lower income
inequality as a policy aim (OMB 2009). In the years since, income inequality has remained on the political agenda.

In short, trillions of dollars of assets were destroyed (at least on paper), and trade networks were disrupted by the liquidity crunch. The government response was substantial: spending in the hundreds of billions and an overhaul of the financial regulatory framework. Public opinion shifted, academic orthodoxy has been critiqued, and income inequality became a legitimate target of policy. The financial crisis of 2007 potentially recreated the conditions for a spatial reconfiguration and inequality regime change in the United States.

And the immediate impact of the recession on the distribution of incomes across the United States was non-negligible. According to the Census Bureau, the Gini coefficient of household incomes fell .007 between 2006 and 2007, the largest single year drop since the 1960s (Census 2012: Table A-3). The income share of the top 1% (excluding capital gains) fell from 18.3% to 16.7% between 2007 and 2009, and, including capital gains, fell from 21.5% to 17.5% during that period (Piketty and Saez 2003; updated to 2012). In other words, incomes became less concentrated at the top of the distribution and more broadly distributed throughout.

But the impact of the recession on the distribution of incomes was temporary. While the richest 1% absorbed 49% of losses between 2007 and 2009, that same group captured 95% of the gains during the recovery between 2009 and 2012. Their share of total income, including capital gains, rebounded to 21.5% in 2012 and surpassed the 2007 peak in 2012 when you exclude capital gains (Piketty and Saez 2003). The Gini coefficient for household incomes, which fell .0072 between 2006
and 2007, recovered that value twice over by 2011, growing from .463 to .477 (Census 2014: Table A-3).

Top income shares recovered, and the conditions sustaining relatively low wealth inequality faltered. I noted earlier that the wealthiest are invested disproportionately in financial assets and the rest in their principal residence (i.e., their homes; Wolff 2012). Housing and stock prices both began to fall in 2007, but the Dow Jones Industrial Average plummeted in 2008 while housing prices continued a slower descent (see Figure 5.11). Given the unequal composition of household wealth, this turn allowed non-wealthy households to grab a greater share of total wealth. But beginning in 2009, equity prices rebounded quickly while housing prices continued to fall, stabilizing only in 2011. The result is that median household wealth in the United States fell 47%, but the 100 wealthiest Americans were $279 billion wealthier in 2011 than 2009 (in nominal terms; Wolff 2012; Forbes 2012, author’s calculations).

Figure 5.11 House Prices and Dow Jones Industrial Average, 2000 to 2013

Why was the crisis’ impact on inequality not more longstanding? The first part of the answer is that economic crises produce only temporary income compressions, as noted in chapter 4 (Atkinson and Morelli 2011). Asset prices can fluctuate more wildly than wages. Prices fall quickly during a crisis, but then rebound quickly afterwards. On the other hand, high unemployment tends to have a longer shelf life. Consequently, inequality tends to fall at the onset of the crisis, but then rise again after; this was the case with the Great Depression before the onset of World War II. Enduring inequality compression requires a more substantial material and institutional transformation.

We have not yet seen that substantial transformation of the material and institutional organization of economic life, but this transformation may still be underway. For example, Dodd-Frank, though more than three years old, is largely incomplete. The law allows existing regulating bodies to propose specifics in how the law is executed, but regulators consistently miss deadlines (Villarreal 2013). A final version of the Volker Rule was recently published, and Treasury Secretary Jacob Lew allegedly warned bankers that the rule could be more stringent than they were expecting (Hopkins and Hamilton 2013). Working from an early draft proposal, Standard and Poor’s estimated that the rule “could reduce combined pretax earnings for the eight largest U.S. banks by up to $10 billion annually” (Harper 2012).

Globally, the impact of new regulations is already palpable. Deutsche Bank estimates that employment in the ten largest financial firms will be 20% below its 2010 peak in 2014, and average pay at Goldman Sachs and J.P. Morgan fell 5% in the first 9 months of 2013. Investment banking revenues were down about 5% in 2013.
and 25% from 2009 (The Economist 2014, “The Law of Small(er) numbers). J.P. Morgan claims that one in seven employees works in compliance (personal correspondence). “A failure to cut costs fast enough means that the industry’s profitability has been ruined. Average returns on equity for the biggest investment banks slumped to about 8% last year, according to McKinsey. Without deep cost cuts it reckons this figure will fall to 4% by 2019” (The Economist 2014, “The Law”).

A second part of the answer as to why the impact of the crisis was short-lived is that the financial crisis was quickly followed by a second, related crisis – the sovereign debt crisis. In the United States, debt crisis was experienced primarily as sympathy pains from Europe (except where self-inflicted by Congressional gridlock) as interest rates on government debt have been at record lows. But anger over large bailouts to financial firms combined with the specter of a Southern European-style sovereign debt crisis inspired a more aggressive version of Friedman’s monetarism.

Where the Occupy movement was the visual manifestation of anger over unrestrained corporate and financial power and the resulting economic inequalities and instability, the Tea Party was the manifestation of an ideological movement against government spending and regulation of economic life. The Tea Party echoes positions that have long been part of the Republican platform, but tend to be ideologically extreme. Tea Partyers do not actively support higher levels of income inequality in general, but by opposing progressive taxation (e.g., high top marginal tax rates) and progressive government spending, they oppose the most powerful levers for effecting a compressed distribution of incomes.
In summary, the financial crisis temporarily reduced the size distribution of incomes, and it may have been a precursor to future changes, but the backlash against government intervention also highlights the indeterminate direction in which that change will ultimately flow. For example, Madestam, et al (2013) show that Tea Party organization affected political outcomes more than can be explained by latent attitudes alone. In other words, the direction of future change will be shaped by the efforts of social actors.

Crisis and Synthesis. While the proximate cause of the crisis was unregulated mortgage backed securities and collateralized debt obligations, the fundamental origin was the mismatch between profit incentives in a new global economy and 20th century institutions and logics for organizing and regulating that activity. In the United States, wages have been stagnant for male workers for almost four decades (DeNavas-Walt, Proctor and Smith 2012). As a result, household income growth through the bottom half of the distribution has been anemic. Productive capacity and output continued to grow, both in the United States and elsewhere in the global economy, but capacity utilization (the ratio of actual to potential production based on the existing capital stock) in the United States slid steadily (Fed 2013). US firms and the global economy depended on demand from the world’s largest consumer market, the United States, but US household incomes were not keeping pace.

The stopgap solution came in the form of debt spending, sovereign and consumer. Economic stability, low interest rates, and financial deregulation encouraged leveraged spending in the United States to compensate for stagnant wages. The housing bubble and sub-prime mortgage lending were exemplary cases of Harvey’s

Since the crisis, asset prices have surged in the United States, but full economic recovery in the core may be an impossible utopia. US unemployment has gradually fallen from a peak of 10% in 2009, but at 6.7% in December 2013 is still higher than at any point in the previous two decades (BLS 2013). Asset prices in the United States through January 2014 surged beyond what is justified by economic fundamentals (The Economist 2013, “More Bricks, Fewer Bubbles”). Of this condition, Larry Summers suggests that the rich world may be trapped in “secular stagnation,” in which saving outpaces investment (Summers 2013), a defining quality of a configuration in the final stages of a financial expansion.

Economic instability in the core is not the only sign of the next spatial reconfiguration. Globally, sites of wealth generation and accumulation have been shifting. The obvious case is China. China averaged more than 10% growth between 1990 and 2007, and in 2007 was producing 5.5 times more in constant terms than in 1990. The Gross Domestic Product grew by $1.5 trillion between 2000 and 2007 (in 2005 dollars), equal to about 3% of the global production in that year and 15% of global growth between 2000 and 2007 (World Bank 2013, author’s calculations).

Aggregate measures of economic growth for China underestimate China’s economic transformation and, in turn, its transformative impact on the global economy. For example, 58 million Chinese live in provinces that would be classified by the World Bank as high income by GNI (Gross National Income) per capita
(Tianjin, Beijing and Shanghai), and another 158 million will cross that threshold by 2015 (Jiangsu, Inner Mongolia, Zhejian). The number of dollar billionaires in China increased from 8 to 95 between 2006 and 2012; by way of comparison, the number of billionaires in the United Kingdom, Germany and France, combined, increased from 93 to 105 (Forbes 2012). And average (constant) wages across a sample of 14 occupations in Shanghai almost tripled between 2000 and 2012 (UBS 2012; Korzeniewicz and Albrecht 2012; see Figure 5.12). In other words, wealth is quickly accumulating, and wage costs rising, in those parts of China that the rich world has increasingly become dependent on as a source of cheap labor.

Figure 5.12 Average Wage, 2000 and 2012 (2000 US$)

Perhaps more important that surging production in China will be China as an increasingly important destination for consumer goods. “In the 1950s and 1960s the world-economy was transformed by the emergence of the American consumer. Now China seems poised to become the next consumption superpower. In all likelihood, it
has just overtaken Japan to become the world’s second-biggest consumer economy. Its roughly $3.3 trillion in private consumption is about 8% of the world total, and it has only just begun” (The Economist 2014, “Chinese Consumers”). Growing supply and demand from China will further transform global financial flows and sites of wealth accumulation.

While China is the extreme example, both in terms of population size and the pace of growth, the global distribution of incomes more generally has been transformed in recent decades. Using GDP per capita and the within-country distribution of incomes from national surveys, I estimate the global distribution of incomes for 1980 and 2011. During that period, the distribution was reshaped from bimodal in 1980, clearly identifying a rich and poor world, to unimodal in 2011, reflecting a fundamental transformation in where wealth is produced and how it is then distributed (see Figure 5.13).

Figure 5.13 Global Distribution of Incomes, 1980 and 2011
The parallels with the world-economy a century earlier are significant. Germany and the United States, industrial latecomers, used British capital, the newest technologies, and relatively large populations to transform the global economy beyond what the arrangements of British liberalism could sustain. The result was interstate tension (an arms race and war), financial instability, chaos and inequality regime change (in the United States and elsewhere). If these similarities can hold up over time, we should expect a reconfiguration of the world-economy and a related inequality regime change in the United States.

And in what direction will that take us? The challenge in prognosticating the direction of inequality change is that the period leading to a regime change creates space for social action but does not determine the outcome. But if we consider the principle dimensions of inequality change in the 20th century, some predictions are plausible.

Rising wages in China have led some to predict a return of manufacturing to the United States. The Boston Consulting Group (Sirkin, Zinser and Rose 2013) explicitly laid out these expectations:

Export manufacturing has recently become the unsung hero of the U.S. economy. . . . [T]he country’s exports have been growing more than seven times faster than GDP since 2005. As a share of the U.S. economy, in fact, exports are at their highest point in 50 years. . . . But this is likely to be just the beginning. We project that the U.S., as a result of its increasing competitiveness in manufacturing, will capture $70 billion to $115 billion in annual exports from other nations by the end of the decade. About two-thirds of these export gains could come from production shifts to the U.S. from leading European nations and Japan. By 2020, higher U.S. exports, combined with production work that will likely be “reshored” from China, could create 2.5 million to 5 million American factory and service jobs associated with increased manufacturing.
Rising prices in China combined with falling energy prices in the United States have increased the country’s competitiveness in this sector.

In addition to creating new jobs, the re-emergence of manufacturing in the US economy could empower labor. At first blush, manufacturing returning to the United States seems like a harbinger of good things to come for the labor movement. But the political environment is hardly nurturing; the victory against public-sector labor unions in Wisconsin, for example, highlighted this. And rich world manufacturing is about specialization, adaptability, speed and personalization. It is an economic activity completely distinct from the Fordist manufacturing of the past.

New regulations and declining profits in the financial services industry could signal a definancialization of the US and global economies. That clearly seemed to be the case in 2008 as securitization (the process of pooling assets to sell to investors) and IPOs (initial public offerings) were well below high marks reached earlier in the decade. Both are important in the process of financialization as they add a degree of separation between asset creation (e.g., forming a new company) and ownership. As of 2013, IPOs and securitization were still well below their pre-crisis peaks but trending upwards (The Economist 2014, “Back from the Dead”; see Figure 5.14).
A key covariate of inequality change in the United States in the 20th century was the progressivity of federal income taxes. The focus since 2007 on sovereign debt would suggest that raising top tax rates would be a focal point on political agendas. But any effort to do so elicits cries of class warfare (the alternative to class warfare, we must presume, is acknowledging that the wealthy have already won). Top marginal tax rates will stay closer to 30% than 90%, barring a major new demand for revenues by the federal government.

And what could drive that new demand? The most significant candidate is the same event that was a major impetus for the introduction of a federal income tax a century ago – total war. It is well beyond the scope of this paper to predict the outbreak of war on a global scale, but it is not unthinkable. The Economist (2013, “Look Back with Angst”) led its 2013 year-end issue by arguing that a “century on, there are uncomfortable parallels with the era that led to the outbreak of the first world war”. And The Economist does not delve into Arrighian systemic cycles of
accumulation, which predict that growing contradictions in the spatial configuration of global production would lead to interstate chaos . . . again.

It is this final variable that makes forecasting regime change difficult. No one in 1913 or even 1929 could have predicted the coming transformation of the US and global economies. Income inequality in the United States today could plummet if interstate chaos paves the way for more financial regulation, more progressive income and estate taxes and a new era of material expansion, with manufacturing returning to the United States. On the other hand, change could reflect the emergence of the corporate inequality regime a century and a half ago: the remaining barriers to global horizontal and vertical integration of production could choke competition and funnel profits to those at the top. Does the rise of China and other emerging market economies signal the shift of economic power out of the United States, or will rising costs and consumption in those countries allow for a new material expansion in the United States? Will protests against the power and profits of large corporate and financial firms result in meaningful regulation, or will concerns about government regulation and capital flight win out? The answers to these questions will shape the next American inequality regime.
Epilogue: Assessing the Argument

I argued in chapter 1 that long-term trends in wealth and income concentration reflect changes in the spatial configuration of the world-economy. People, polities, institutions, networks, geographies, infrastructure, technologies, capital, etc. are located in space. Some components of the configuration are resistant to change (e.g., it is difficult to relocate most infrastructure), and the components of the configuration are interdependent (e.g., the implementation of a technology requires the co-location of people with the right skills, the right capital and infrastructure, the right geography in some cases, and the technology must diffuse through networks and over political and institutional boundaries). As a result of this resistance to change and interdependence of its components, the spatial configuration is durable over time.

When change does occur it cascades through the entire system: new economic and political philosophies, new institutions to implement those philosophies, new productive and transportation technologies, and new infrastructure to handle the new technologies. Domestic and international hierarchies become unstable. The world-economy is reconfigured.

The spatial configuration structures the flow of wealth through the world-economy; access in and across space to wealth-generating activities is regulated by the technologies, infrastructure, geography, institutions, polities, and networks of the existing spatial configuration. Wealth and income concentrates in the hands of those that are able to exploit wealth-generating activities and tap those markets where wealth has accumulated. The result is a relatively stable regime of wealth and income concentration. Conversely, when the configuration of the world-economy is
disrupted, there can be relatively rapid change in where and to whom wealth accumulates.

How well does this model fit long-term trends in wealth and income concentration through US history? I break this discussion into three parts. First, what needs to be explained? In other words, what are the historical trends in wealth and income concentration in the United States? Second, is there a correlation between the dependent and independent variables, or do transitions between American inequality regimes correlate with the crises and reconfigurations of Arrighi’s systemic cycles of accumulation? Finally, can we establish causality from the independent to the dependent variable, or can we identify mechanisms linking spatial configurations to American inequality regimes?

Empirically, the pattern of wealth and income concentration through the 20th century is well documented: income concentration fell markedly around World War II and began to rise again in the last quarter of the 20th century. The historical record on wealth inequality before the Civil War is a matter of debate, but both sides of that debate produce estimates that show wealth becoming more concentrated from the colonial period through the Civil War and peaking around World War I.

It is also important to identify the principal beneficiaries in each period. The historical record is less clear on this count, but I identify patterns in the characteristics of those at the top of wealth and income distributions over time: Southern plantation owners through the colonial period; Northern merchants after the barriers of colonization and mercantilism began to break down, especially during the Napoleonic Wars; industrialists in the second half of the 19th century; managers (e.g., CEOs) and
the financial services industry since 1970. Labor, though not at the top of the income distribution, was a principal beneficiary of arrangements in the post-war period.

Combining the trends in wealth and income concentration with the characteristics of those at the top of the distribution, we can identify five distinct periods: low inequality across free resident households, but large categorical inequalities that benefitted Southern slave owners, through the colonial period; a commercial expansion, made possible by war in Europe, that produced America’s first dollar millionaires and a concentration of wealth between Independence and the Civil War; an explosion of huge corporate conglomerates after the Civil War, and the greater concentration of wealth and income in the hands of their owners and managers; a levelling of the distributions of wealth and income between the Great Depression and the 1970s; rising income inequality since the 1970s as incomes surged in management and finance.

Do these five inequality regimes correlate with Arrighi’s systemic cycles of accumulation. The short answer is yes, as demonstrated in Figure 0.3 reproduced above. Synchronization in the 20th century is definitive: the Keynesian and financial inequality regimes on one hand, and the US-centered material and financial expansions on the other. Synchronization between inequality regime change and systemic cycles of accumulation is less obvious in the 18th and 19th centuries, but my historical account links the beginning of the British systemic cycle of accumulation with the transition to the commercial inequality regime in the last quarter of the 18th century, and the shift to British-centered financial expansion in the 19th century with the transition to the corporate inequality regime in the United States.
Can we make the leap from correlation to causation? Though not definitive, the narratives of chapters 2 through 5 are suggestive. I mechanically link patterns in wealth and income concentration to the spatial configuration by 1) identifying global sites of wealth accumulation and 2) identifying the relevant barriers/facilitators to accessing those sites across different industries. For example, Europe was the key site of wealth accumulation through the colonial period, but technological limitations and institutional boundaries limited access to European markets; some plantation owners were able to accumulate large estates by skirting mercantilist restrictions and exploiting high prices for their exports.

American merchants were the primary beneficiaries at the end of the 18th century when war in Europe gave them unique access to new markets. Early industrial titans were able to reap huge profits by riding rapid US economic growth on top of new organizational forms designed to limit competition. Global competition in the 1970s allowed managers and the financial services industry to rewrite New Deal legislation, extract investment out of labor, and move capital around to find new growth opportunities globally.

Having established a plausible causal relationship between the spatial configuration and long term trends in wealth and income concentration in the United States, the next step is to rule out alternative explanations. But few truly alternative explanations exist. My approach does not contradict the existing literature on within-country inequality; in fact, I draw heavily on those results (e.g., skill-biased technical change) to identify key mechanisms of inequality change in the various historical periods. The difference is by extending the unit of analysis in space and time I am
able to treat these processes endogenously. Consequently, my challenge is not to rule out alternative explanations, but to establish the interdependence of these alternative explanations within the spatial configuration of the world-economy. It is left to the reader to judge the extent to which I succeeded in doing this.

Three other efforts do warrant special attention. First is Kuznets’ inverted-U curve hypothesis. Technically, Kuznets does not offer an alternative hypothesis to my own; the focus on sector dualism highlights wage gaps through the middle of the distribution, not income concentration at the top of the distribution. But I would also argue that Kuznets’ particular specification of the relationship between economic development and inequality is less important than the more general assumption that there is a relationship between development and inequality.

I note in chapter 5 that productive arrangements can have distinctive inequality personalities; e.g., large-factory manufacturing can facilitate the organization of labor and increase productivity for moderately skilled workers, and thus be associated with lower levels of inequality. Linking development to within-country inequality, then, requires only that we assume a consistent pattern in the developmental pattern across countries – e.g., agriculture to manufacturing to service. Taken to its extreme, this logic suggests that global economic processes matter only in so far as they facilitate (e.g., diffusion of industrial technologies) or hinder (e.g. colonialism) national economic development.

I forego a full examination of the empirical merits of the modernizationist approach to within-country inequality, and go only so far as to say that researchers have challenged the stages of development model (Arrighi, Silver and Brewer 2003)
and the correlation between the level of development and within-country inequality (Korzeniewicz and Moran 2005). It better serves my purposes to note that the history of economic development in the United States does not make sense in a strictly modernizationist framework. The industrialization of the American economy in the 19th century was spurred by British technology and capital, and it was built on the labor of European immigrants; war in Europe at the beginning of the 20th century allowed the United States to become the global manufacturing and financial leader; global competition in the second half of the 20th century pushed the United States to develop a service-oriented economy; international trade policy played a role in the American Revolution and the Civil War, and belligerent trade policy tore apart the world-economy in the decade before World War II. The main characters in US economic history are both foreign and domestic; it does not seem prudent to extract this experience from its global-historical context and impose it elsewhere.

The second effort is Goldin and Katz’s (2010) race between education and technology. Fundamentally, the argument is that technology and skill tend to be complementary, so new technology increases the demand for skill while education increases the supply of skilled workers. When technological change outpaces education, inequality increases with the skill premium. Again, this approach is not directly in contrast with my own as the race between education and technology will be made manifest more in the middle of the distribution. Also, innovation (technology and new techniques) and diffusion (education) are key components of my model. But Goldin and Katz generally assume education and technology are exogenous.
The third effort only recently made its splash in the within-country inequality literature. Piketty (2014) argues that within-country inequality is a product of capital accumulation. In short, ownership of capital is more concentrated than incomes, so within-country inequality is driven by the ratio of returns to capital to national income (or the relative incomes of capitalists and workers), and exacerbated by the ratio of capital to national income (because it reflects the value of physical capital relative to human capital). The capital to income ratio tends to rise over time, returns to capital are fairly constant over time, so within-country inequality will rise over time unless capital is destroyed (reducing the capital to income ratio) or national income growth is extraordinary (greater than 5%). In Piketty’s account, within-country inequality in rich countries increased over the last few centuries with the exception of one historical moment (the Great Depression and World War II) that destroyed capital and produced extraordinary income growth rates. (I noted in chapter 4 the empirical problems with applying Piketty’s model to the Great Compression.)

Reviewers have noted that constant returns to capital, if capital is allowed to accumulate relative to national income, seems to contradict the basic economic assumption of price reflecting the intersection of supply and demand—supply increases, demand does not, so price should fall (Milanovic 2013). Piketty himself posits that it is this assumption (that the accumulation of capital would lead to falling profits) that led Marx astray, “a historical prediction that turned out to be quite wrong” (Piketty 2014: 59). But Piketty’s position is not heretical economically, nor does it contradict the logic of overaccumulation; it only seems that way because he presents his model of within-country inequality in the context of the nation-state.
The logic of overaccumulation does not presuppose a falling return to capital overall, but falling profits in those markets that had received the heaviest investment (the dominant regime of accumulation). This results in a “negative judgment on the possibility of continuing to profit from the reinvestment of surplus capital in the material expansion of the world-economy” as it is currently configured (Arrighi 1994/2010: 220). Capitalists in the center of the world-economy strategically specialize in high finance, and by finding new avenues for investment (and driving speculative bubbles) can forestall crisis and even enjoy a “wonderful moment;” Fukuyama (1992) is not alone in confusing the peak of a financial expansion for the resolution of the human dialectic. If we focus only on rich countries and treat them independently, as Piketty does, we can conclude that returns on investment are fairly constant and that the process of capital accumulation can continue indefinitely. But if we recognize that Britain before World War I managed high returns to capital by shifting capital flows to the United States and Germany, and that this ultimately destabilized the structure of the world-economy, we can see that the Great Depression and World War II were not exceptional historical moments, but products of systemic cycles of accumulation in the world-economy.

In short, as I argued in the introduction, the quality that sets this project apart from other efforts of its kind is the more complete use of time and space. The United States does not develop economically, adopt new technologies, expand access to education, accumulate capital or find returns to that capital in a vacuum.

Two arguments justify a focus on ‘domestic’ (instead of global) processes in the study of within-country inequality (particularly within the United States): 1) domestic
labor market institutions are important and 2) international trade and immigration are of insufficient magnitude to drive current trends. But when we extend the unit of analysis to include all relevant processes, we see that domestic labor market institutions are not independent of global processes and the magnitude of international trade and migration are not independent of domestic labor market institutions. To quote Wallerstein (1979: 35) one last time, we must recognize the importance of the “discontinuity between economic and political institutions.” In these overlapping spaces, economic actors deploy a range of strategies, not just price competition, to gain an advantage. In other words, competition is economic, political, ideological, and its impact is economic, political and ideological. These are not independent spheres of analysis, but interrelated components within the system of the world-economy.

From this summation I offer two take-away messages. First, within-country inequality is a product of global processes in two key ways: 1) wealth accumulates in global space and 2) domestic economic and political ideologies, political and market institutions, and patterns of investment in infrastructure and technology are shaped by global competition. Second, within-country inequality is a conscious human objective. Economic actors intentionally deploy a set of strategies to grab a larger share of incomes. The success of any such strategy depends on the strength of the opposition and the structural conditions (emphasized in this dissertation) that can facilitate or deny exclusive access to wealth-generating activities. Institutional and technological change are choices, the products of conflict and negotiation, and justified by ideological narratives; they are not exogenous. Income inequality in the
United States today is rising because enough people wanted it to do so, and they deployed the right set of strategies to achieve that goal. The current trend will be reversed when enough people decide it is worth disrupting the *status quo* to redistribute access to wealth-generating activities.
References


Albrecht, Scott. 2007. “Age Structure of US Income Inequality.” Master’s Thesis, University of Maryland, Department of Sociology.


Coxe, Tench. 1791. *A Brief Examination of Lord Sheffield’s Observations on the Commerce of the United States.* from the press of M. Carey: Philadelphia


Defoe, Daniel. 1728. *A Plan of the English Commerce: Being a Compleat Prospect of the Trade of this Nation, as well as the Home Trade as the Foreign.* London: Charles Rivington, at the *Bible and Crown* in St. Paul’s Church-Yard.


Kondratieff, Nikolai. 1925. The Major Economic Cycles.


Smith, John. 1608. “Letter sent to the Treasurer and Councell of Virginia from Captaine Smith, then President in Virginia.” Retrieved from


US Census Bureau (Census). 2014. “Selected Measures of Income Dispersion,
1967-2010.” Historical Income Tables: Income Inequality, available at

from http://future.state.gov/when/timeline/1921_timeline/smoot_tariff.html

Vickers, Daniel. 1996. “The Northern Colonies: Economy and Society, 1600-

Villarreal, Alexandra. 2013. “Regulators miss 70.1 percent of 398 Dodd-Frank
rules.” Bank Credit News June 7, 2013. Retrieved from
http://bankcreditnews.com/news/regulators-miss-70-1-percent-of-398-dodd-
frank-rules/10769/.

and the Origins of the European World-Economy in the Sixteenth Century.

Press: Boston.


Walter, Andrew. 1991. World Power and World Money: The Role of Hegemony


York.


Wilkins, Mira. 1970. The Emergence of Multinational Enterprise. Cambridge
University Press: Cambridge.


