ABSTRACT

Title of dissertation: ESSAYS ON MACROECONOMIC TRENDS AND CYCLES

Yiqun Chen, Doctor of Philosophy, 2016

Dissertation directed by: Professor S. Borağan Aruoba
Professor L. Luminita Stevens
Department of Economics

This dissertation describes two studies on macroeconomic trends and cycles. The first chapter studies the impact of Information Technology (IT) on the U.S. labor market. Over the past 30 years, employment and income shares of routine-intensive occupations have declined significantly relative to nonroutine occupations, and the overall U.S. labor income share has declined relative to capital. Furthermore, the decline of routine employment has been largely concentrated during recessions and ensuing recoveries. I build a model of unbalanced growth to assess the role of computerization and IT in driving these labor market trends and cycles. I augment a neoclassical growth model with exogenous IT progress as a form of Routine-Biased Technological Change (RBTC). I show analytically that RBTC causes the overall labor income share to follow a U-shaped time path, as the monotonic decline of routine labor share is increasingly offset by the monotonic rise of nonroutine labor share and the elasticity of substitution between the overall labor and capital declines under IT progress. Quantitatively, the model explains nearly all the divergence between routine and nonroutine labor in the period 1986-2014, as well as the mild decline of the overall

labor share between 1986 and the early 2000s. However, the model with IT progress alone cannot explain the accelerated decline of labor income share after the early 2000s, suggesting that other factors, such as globalization, may have played a larger role in this period. Lastly, when nonconvex labor adjustment costs are present, the model generates a stepwise decline in routine labor hours, qualitatively consistent with the data. The timing of these trend adjustments can be significantly affected by aggregate productivity shocks and concentrated in recessions.

The second chapter studies the implications of loss aversion on the business cycle dynamics of aggregate consumption and labor hours. Loss aversion refers to the fact that people are distinctively more sensitive to losses than to gains. Loss averse agents are very risk averse around the reference point and exhibit asymmetric responses to positive and negative income shocks. In an otherwise standard Real Business Cycle (RBC) model, I study loss aversion in both consumption alone and consumption-and-leisure together. My results indicate that how loss aversion affects business cycle dynamics depends critically on the nature of the reference point. If, for example, the reference point is status quo, loss aversion dramatically lowers the effective inter-temporal rate of substitution and induces excessive consumption smoothing. In contrast, if the reference point is fixed at a constant level, loss aversion generates a flat region in the decision rules and asymmetric impulse responses to technology shocks. Under a reasonable parametrization, loss aversion has the potential to generate asymmetric business cycles with deeper and more prolonged recessions.

ESSAYS ON MACROECONOMIC TRENDS AND CYCLES

by

Yiqun Chen

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
2016

Advisory Committee:
Professor S. Borağan Aruoba, Chair
Professor L. Luminita Stevens, Co-Chair
Professor John C. Haltiwanger
Professor John Shea
Professor Phillip L. Swagel

DEDICATION

To my parents, Bolin and Youqing.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my dissertation committee members at University of Maryland, in particular, my co-chairs Borağan Aruoba and Luminita Stevens, for their guidance and support in the development of this dissertation. Since the beginning of my graduate study, Borağan has taught me the most of macroeconomic models and computation techniques. His teaching has deeply inspired me and provided the foundation for me to pursue macroeconomic research. Throughout my time at Maryland, I have benefited immensely from his mentoring and constant encouragement, for which I am forever grateful. Luminita joined the Maryland faculty when I was a third-year student. I had just finished all the field courses and was fascinated by behavioral macroeconomics, but struggled deeply to find a suitable topic for a research paper. Luminita guided me through this challenging and profound process, and showed me how to transition from being a student to being a researcher and scholar. A young female economist herself, she has also done the most to encourage and support me to seize opportunities both inside and outside the Department of Economics at University of Maryland to develop my career as an economist. This dissertation would not have been possible without her support.

I have also benefited tremendously from the guidance of my other advisers at University of Maryland. I am deeply indebted to John Shea for his extensive comments on each and every draft of my papers in the past few years. His commitment to mentoring and education is legendary and a constant inspiration to students and teachers alike in the department. I have also benefited immensely from the teaching and guidance of John Haltiwanger. Much of Chapter One of this dissertation was inspired by a series of conversations with John, whose knowledge and insights into the labor market issues and macroeconomics are truly incomparable. Finally, I thank Phillip Swagel for his generosity in serving on my dissertation committee.

Many others at University of Maryland have helped me and contributed to the development of this dissertation. In particular, I thank Ethan Kaplan, Anton Korinek, Sebnem Kalemli-Ozcan, Nuno Limao and Felipe Saffie, for their useful comments on my research. I thank my wonderful fellow graduate students, Pablo Cuba-Borda, Xiaoya Dou, Jingting Fan, David Moreno, Wei Li, Lixin Tang, and Ozlem Tonguc, among others, for years of academic and emotional support. Part of this dissertation was written when I was visiting the Federal Reserve Board of Governors, and I also thank Tomaz Cajner, Charles Fleischman, Deb Lidner, Christopher Nakarda, Dave Ratner, Christopher Smith, and many others at the Board for their valuable feedback.

Last but not least, I am grateful to my family: Mimi, for her gentle and warm company in the past five years; D. C. Ruiz, for being an endless source of joy; and David, for his love, humor, and support for me in both the biggest and the smallest decisions in life. Thanks to them, the past few years as a doctoral student were both productive and filled with joy. Finally, I am eternally indebted to my parents, Bolin and Youqing. Their unconditional love and unwavering support for my education is what makes this dissertation, and everything else in my life, possible.

College Park, Maryland

April, 2016

Contents

List of Tables

List of Figures

LIST OF ABBREVIATIONS

AR	Auto-Regression
BEA	Bureau of Economic Analysis
BGP	Balanced Growth Path
BLS	Bureau of Labor Statistics
CES	Constant Elasticity of Substitution
CPS	Current Population Survey
DOT	Dictionary of Occupational Titles
DSGE	Dynamic Stochastic General Equilibrium
EOS	Elasticity of Substitution
EU	Expected Utility
FRED	Federal Reserve Economic Data
GDP	Gross Domestic Product
GMM	General Method of Moments
GPT	General Purpose Technology
HP	HodrickPrescott
IT	Information Technology
MORG	Merged Outgoing Rotation Groups
LSTC	Labor Saving Technological Change
NIPA	National Income and Product Accounts
QAR	Quadratic Auto-Regression
R&D	Research and Development
RBC	Real Business Cycle
RBTC	Routine Biased Technological Change
RTI	Routine Task Intensity
TFP	Total Factor Productivity
TOPI	Taxes on Production and Imports
WTO	World Trade Organization

Chapter 1: The Rise of the Machines and the U.S. Labor Market

1.1 Introduction

Many macroeconomic models assume that the aggregate factor shares of income are stable over the long run and that business cycles can be treated as independent of economic growth trends. Both assumptions have been challenged, however, by the structural changes in the U.S. labor market in the past 30 years.

First of all, the U.S. labor income share has declined relative to capital since the mid-1980s (Elsby, Hobijn, and Şahin, 2013; Karabarbounis and Neiman, 2014; and others). The headline measure of labor share compiled by the Bureau of Labor Statistics (BLS) historically fluctuated around a mean of 64 percent. Since the mid-1980s, however, aggregate labor share has been on a mild downward trend, and the decline accelerated notably after the 2000s to an average of around 58 percent in recent years. Alternative measures of labor share also suggest a significant, albeit smaller, decline over this period. A declining labor share has many implications, including declining government tax revenue — as labor is generally taxed at a higher rate of capital — and changes in responses to monetary policy.

Second, within labor, there has been an increasing divergence between two groups of workers: those who work in "routine-task-intensive" occupations, such as manufacturing and clerical jobs, and those who work in "nonroutine-task-intensive" occupations, such as professionals and technicians. In the past three decades, both employment and income shares of routine occupations have declined steadily and significantly relative to those of nonroutine occupations. Because nonroutine labor includes both the highest-skilled and the lowest-skilled, this divergence between nonroutine and routine occupations is also known as a "polarization" in the aggregate labor market along the worker skill-distribution, which is closely linked to the weakening of the middle class. The disappearance of "routine" jobs — particularly in the manufacturing sector — and the associated weakening of the middle class and the rise in income and wealth inequality have far-reaching implications on the shape and welfare of the macroeconomy, and have been front and center in the recent political debate.

What are the driving forces behind these structural changes in the U.S labor market? A prominent candidate is technological progress. The structural shifts from routine to nonroutine labor and from labor to capital coincide with the so-called "second machine age," as advances in computer and Information Technology (IT) have led to an increasing pace of automation, causing disruptions in the workplace (Brynjolfsson and McAfee, 2014). Economists have largely agreed that "routine-biased" technological change, such as computerization and IT progress, is one of the main driving forces behind the decline of routine occupations relative to nonroutine occupations. (Autor, Levy, and Murnane, 2003; Autor, Dorn, and Hanson, 2015; Goos, Manning, and Salomons, 2014; and others). However, it is still an open question whether technology has also contributed to the decline of the overall labor share relative to capital (Elsby, Hobijn, and Şahin, 2013; Karabarbounis and Neiman, 2014), and if so, to what extent and whether the trend will continue. In this study, I seek to answer these questions both qualitatively and quantitatively.

The second assumption that most macroeconomic models rely on is the dichotomy between long-run economic growth and short-run business cycles. However, there is strong evidence that recent business cycles in the U.S. are not independent from long-run growth trends. More specifically, the decline of routine employment since the 1980s has not been a gradual process; instead, it has taken a "stepwise" shape, with nearly all of the long-term downward adjustments occurring within a 3-year window around each recession. This concentration of trend adjustments in recessions, in turn, has caused some "anomalies" in the cyclical behavior of the labor market. For each of the three most recent U.S. recessions (1990-91, 2001, and 2007-09), employment and total hours lagged behind output for more than a year during the recovery, giving rise to unprecedented "jobless recoveries". Jaimovich and Siu (2012) suggest that the majority of the job loss in recessions and ensuing weak recoveries during this period can be traced back to the trend decline of routine labor. This is not surprising as cyclical fluctuations of business conditions and short-run market imperfections can significantly distort the timing and the magnitudes of medium-run structural changes. However, a conventional model in which economic growth trends and business cycles are independent of each other is unable to capture the structural components of labor fluctuations over business cycles.

In this study, I build a model of unbalanced growth that allows for interactions between labor market trends and business cycles. I use the model to formally assess the role of IT growth, which takes the form of Routine-Biased Technological Change (RBTC), in driving the decline of routine labor income share and employment relative to nonroutine labor, the decline of the overall labor income share relative to capital, and the rise of jobless recoveries. The model builds upon the general equilibrium framework in Krusell, Ohanian, Ríos-Rull, and Violante (2000) and incorporates the task model of Autor, Levy, and Murnane (2003). The key assumption is familiar and intuitive: that "IT capital" is highly substitutable for routine labor and highly complementary to nonroutine labor. As a result, progress in IT, or "the rise of the machines", leads to a permanent decline in the demand for routine labor and a permanent rise in the demand for nonroutine labor.

The model yields several interesting results. First, I show analytically that the overall labor income share follows a U-shaped growth path under RBTC, as opposed to a continued decline as suggested by previous studies (Morin, 2014 and Karabarbounis and Neiman, 2014) and feared by the public. This U-shaped path arises as the monotonic decline of the routine labor share is increasingly offset by the monotonic rise of the nonroutine labor share over time. More specifically, if routine labor initially constitutes a large share in employment and labor income, as is the case of the United States in the 1980s, then the substitutability between IT capital and routine labor dominates the complementarity between IT capital and nonroutine labor, and overall labor is effectively a net substitute for capital in the aggregate production function. At this stage, capital-augmenting technological change, such as IT progress, causes the overall labor's share of total income to fall. Over time, however, RBTC induces labor to flow away from the routine sector towards the nonroutine sector; as a result, overall labor and capital become increasingly complementary. After reaching an inflection point, overall labor becomes effectively a net complement for capital, and IT progress causes the labor share to rise relative to capital. In other words, RBTC implies a declining elasticity of substitution (EOS) between overall labor and capital. This result contributes to the ongoing debate on capital-labor substitutability (see Chirinko (2008) for a survey, also Karabarbounis and Neiman (2014) for a more recent discussion), as it is the first to highlight how changing composition of labor can significantly alter the relation between overall labor and capital. The declining path for the capital-labor EOS suggested by the model also provides a possible reconciliation between the recent decline of labor share and the long-run complementarity between labor and capital. Moreover, the model's prediction of a U-shaped path for the overall labor share suggests that the recent decline is only temporary; in the long run, computerization and IT progress will enhance labor's share in total income rather than depressing it.

Second, to quantify these distributional effects on RBTC, I calibrate the model to U.S. quarterly data in the period 1986-2014. Since the model features unbalanced growth paths and is therefore nonstationary, I solve it using the Extended Function Path method (Maliar, Maliar, Taylor, and Tsener, 2015). I find that, while RBTC drives large shifts in the employment and income distributions between routine and nonroutine labor, it has a relatively mild effect on the income distribution between overall labor and capital. More specifically, the calibrated model can account for nearly all of the divergence between routine and non-routine labor in employment and income shares over time. This result suggests that, on the aggregate level, RBTC is *the* driving force behind the decline of routine occupations relative to nonroutine occupations.[1] The calibrated model can also account for the mild decline of the overall labor share in the period 1986-2002, but cannot explain the acceleration of the decline after 2002, which is consistent with the timing of events: computerization and the IT revolution were major phenomena in the 1980s and 1990s, but the productivity gains from technological progress slowed down notably in the middle 2000s (Fernald, 2014).

[1] Autor, Dorn, and Hanson (2015) and Goos, Manning, and Salomons (2014) find the same result using data from local markets and across countries, respectively.

In total, my model generates a 2.3 percentage point decline of the labor income share in the period 1986-2014, which is in line with the estimates of Eden and Gaggl (2014) and Karabarbounis and Neiman (2014).

Lastly, the model generates several novel results on the interactions between trends and business cycles in the labor market. In particular, I show that when there are nonconvex routine labor adjustment costs, the declining trend of routine labor hours takes a stepwise shape with long periods of no decline and periods of large adjustments. This stepwise decline of routine hours also induces infrequent but large fluctuations in overall labor hours, which is qualitatively consistent with the patterns observed in the data. More importantly, the model demonstrates how aggregate productivity shocks can significantly affect the timing of such trend adjustments. In particular, after a long period of inaction, the economy enters a period in which the likelihood of downward trend adjustment is high. In this period, even a small negative aggregate productivity shock can trigger a large downturn in the labor market. The model identifies three such periods with heightened propensity to adjust in the past 30 years, which retrospectively coincide well with the periods of downturns in the labor market in the data. Lastly, the model suggests that the small drop of output at the beginning of 2000, which was not enough to cause a recession, triggered one of largest and prolonged labor market adjustments in recent years. This line of analysis complements Jaimovich and Siu (2012), who focus on how secular trends can generate jobless recoveries, whereas in this study, I focus on how business cycles can affect the timing and the magnitude of trend adjustments.

In addition to the main results stated above, in deriving the production structure of the routine sector, I show that the commonly used Constant Elasticity of Substitution (CES) function can be derived from a production technology that allows direct substitution between IT capital and routine labor in producing a continuous array of routine tasks. This result provides a micro-foundation for the CES production function. Moreover, in an extension of the baseline model, I introduce a second type of technological change, which I call Labor-Saving Technological Change (LSTC). LSTC augments non-IT capital and substitutes for both types of labor, and therefore can be loosely associated with the rise of globalization and offshoring. I show that the recent movements of factor shares, namely the acceleration of the decline in the share of overall labor and the slowdown of the rise in nonroutine labor's share of aggregate income after 2002, are qualitatively consistent with a second trend induced by LSTC. Quantitatively, however, the results are inconclusive.

This study is related and contributes to several large strands of literature on technological changes and labor market trends and cycles. First, it builds upon the task model of Autor, Levy, and Murnane (2003) and Acemoglu and Autor (2011), and is related to the subsequent empirical literature on the polarization of the U.S. labor market (Autor, Katz, and Kearney, 2006; Cortes, Jaimovich, Nekarda, and Siu, 2014; Lehn, 2015; and others). It is also related to studies on the recent decline of the labor income share (Bridgman, 2014; Elsby, Hobijn, and Şahin, 2013; Karabarbounis and Neiman, 2014; Rognlie, 2015). Second, this study contributes to a large literature on technological progress and unbalanced growth (Gordon, 2010; Fernald, 2014; Acemoglu, 2002; Acemoglu and Guerrieri, 2008; and others). Third, the business cycle analysis builds upon a long line of literature on labor adjustment costs (Caballero and Engel, 1993; Caballero, Engel, and Haltiwanger, 1997; Cooper, Haltiwanger, and Willis, 2015; Cooper and Willis, 2009). It contributes to a burgeoning literature on the joint dynamics of business cycles and labor market trends (Foote and Ryan, 2015; Gaggl and Kaufmann, 2014; Jaimovich and Siu, 2012; Morin, 2014). Finally, the analysis on LSTC is in line with the most recent empirical findings on the differential roles of trade and technology in the labor market (Autor, Dorn, and Hanson, 2015; Goos, Manning, and Salomons, 2014).

The remainder of this chapter is organized as follows. Section 1.2 presents empirical evidence on recent structural changes and business cycle dynamics observed in the U.S. labor market and discusses the related literature. Section 1.3 presents a simple model with RBTC and shows how it can generate labor market structural changes, in particular a U-shaped growth path for the overall labor share and a declining EOS between overall labor and capital. Section 1.4 presents the full baseline model, calibrates it to U.S. quarterly data, and presents quantitative results on the goodness of fit for medium-run trends and business cycle dynamics. Section 1.5 extends the baseline model to incorporate labor adjustment costs and shows how aggregate productivity shocks may shift the timing and the magnitudes of the trend adjustments and cause them to concentrate in recessions. Section 1.6 extends the baseline model to incorporate LSTC in order to account for the acceleration of the decline of labor's share after 2002. Section 1.7 offers some concluding remarks.

1.2 Empirical Evidence and Related Literature

1.2.1 Divergence of Routine and Nonroutine Labor

The terms "routine" and "nonroutine" refer to the task content of an occupation. Conceptually, a task is routine if it can be performed by following an explicit set of rules and procedures. A task is nonroutine if it requires either abstract reasoning and creativity, or physical dexterity and direct interpersonal interaction. Generally speaking, routine tasks are inherently more suitable for automation, while nonroutine tasks must be performed by humans. It is worth noting that recent progress in machine learning and artificial intelligence has made it increasingly difficult to draw a line between routine and nonroutine tasks. Many tasks that were considered nonroutine only a few years ago, such as driving, writing, and playing chess, can now be performed by machines or computers, albeit to a limited extent. However, automation of such tasks relies heavily on the availability of big data and/or a highly controlled environment. In any case, computers, machines and robots cannot *think* or react in an uncontrolled environment like humans. Therefore, it is unlikely that the majority of the nonroutine tasks and occupations defined in this study will become routine in any foreseeable future.

I use the Current Population Survey (CPS) Merged Outgoing Rotation Groups (MORG) files available from 1979 onward as the main data source for employment, wages, hours, and occupations. I also use the task data from Autor and Dorn (2013) and the harmonized occupation classification system (*occ1990dd*) introduced by Dorn (2009) to pool occupations across different decades. I extend Dorn (2009)'s system by incorporating the most recent revisions of the Census occupation codes after 2010.

In the literature, there have been two ad-hoc approaches to matching task data to occupations. One approach, which was introduced by Autor and Dorn (2013) and used by Goos, Manning, and Salomons (2014) and Autor, Dorn, and Hanson (2015), utilizes an index of Routine Task Intensity (RTI), defined as:

$$RTI_k = \ln(T^R_{k,1980}) - \ln(T^A_{k,1980}) - \ln(T^M_{k,1980}) \tag{1.1}$$

where $T^R_{k,1980}$, $T^A_{k,1980}$, and $T^M_{k,1980}$ are measures (on a 0-10 scale) of routine, abstract, and manual task inputs of occupation k in 1980.[2] RTI measures the relative importance of routine tasks in occupations, which can also be interpreted as an occupation's potential susceptibility to displacement by automation. In their analysis of the local labor markets, Autor and Dorn (2013) label the occupations in the top 33% of the 1980-employment-weighed distribution of RTI as routine occupations.

A second approach to defining routine and nonroutine occupations is by labeling major occupation groups as one of four kinds: nonroutine-cognitive, routine-cognitive, routine-manual, and nonroutine-manual. This is the approach taken by Jaimovich and Siu (2012) and Foote and Ryan (2015), both of which study empirically the link between aggregate labor market trends and business cycles. Table 1.1 shows the classification of six major occupation groups into the four theory-based categories.[3] Using this approach, about 60% of occupations in 1980 are defined as routine, which is obviously much higher than the measure based on the RTI approach used by Autor and Dorn (2013). The discrepancy reflects the ad-hoc nature of both approaches: on the one hand, the RTI measure has a

[2]The RTI measure is based on Autor, Levy, and Murnane (2003) and the task data from the 1997 Dictionary of Occupational Titles (DOT), which assesses the occupational tasks of more than 12,000 highly detailed occupations. Autor and Dorn (2013) aggregate the detailed occupations to the three-digit occupation codes of the 1970 Census, and then map them to the harmonized *occ1990dd* occupation codes. There are five original DOT task variables. The Routine task measure, T^R_k, is an average of two DOT variables, "set limits, tolerances and standards" and "finger dexterity", which measure the intensity of routine cognitive and routine motor tasks in occupations. The Abstract task measure, T^A_k, is the average of "direction, control, and planning" and "quantitative reasoning". Lastly, the Manual task measure, T^M_k, corresponds to the DOT variable "eye-hand-foot coordination". Goos, Manning, and Salomons (2014) map the RTI measure to the European occupational classification and study job polarization in 16 Western European countries.

[3]Classifying tasks along two dimensions, routine versus nonroutine and cognitive versus manual, is, again, based on Autor, Levy, and Murnane (2003). The definition of major occupation groups varies across studies and over time. The decennial Census contains the *official* grouping, but the number and definition of the groups change from one Census to another. The six occupation groups considered in this study is a slightly aggregated and modified version of the Census groups and is discussed in Autor and Dorn (2013). Nevertheless, no matter what the definitions of the major occupation groups are, the mapping of the underlying occupations into the four theory-based categories is very stable. See Foote and Ryan (2015) for more discussion.

straightforward and appealing interpretation on the occupation level, but the cut-off value used to define routine and nonroutine — 33% in Autor and Dorn (2013) — is arbitrary. On the other hand, labeling an entire occupation group as routine or nonroutine may overlook significant variations across occupations within the group. For example, both being service jobs, motion picture projectionists have an RTI of 6.22, the second highest among all occupations, while recreation workers have an RTI of -1.89, one of the lowest.

Table 1.1
Classification of Major Occupation Groups

	Major Occupation Groups
Nonroutine Cognitive	Management, professional, technical, financial sales, public security
Routine Cognitive	Administrative support and retail sales
Routine Manual	Precision production and craft; Machine operators assemblers, and inspectors; Transportation, construction, mechanics, mining, agricultural
Nonroutine Manual	Service

In light of these discrepancies, I use a classification scheme that combines the two approaches in the existing literature to define routine and nonroutine occupations in this study. I first rank occupations by their RTI values and divide the employment-weighed distribution of RTI into three parts.[4] I consider the top 1/3 occupations in the employment-weighed RTI distribution as routine and the bottom 1/3 as nonroutine. For the middling 1/3, I consider an occupation routine if it belongs to one of the routine occupation groups

[4]The original Autor and Dorn (2013) classification uses employment in 1980 to weight the RTI distribution, but here I use the 1986 employment instead. The choice is made for two reasons. First, 1986 is the starting year of my quantitative model to which several important parameters are calibrated. Second and more importantly, under Dorn (2009)'s occupation categorization system, 68 occupations - about $1/5$ of the total - are missing in the 1980 data, while only 8 occupations are missing in the 1986 data. Nonetheless, none of the results in this study are sensitive to using employment data in 1980 to measure routineness.

listed in Table 1.1. Under this classification scheme, about 55% of occupations in 1980 are defined as routine.

Figure 1.1
Divergence of Routine and Nonroutine Occupations

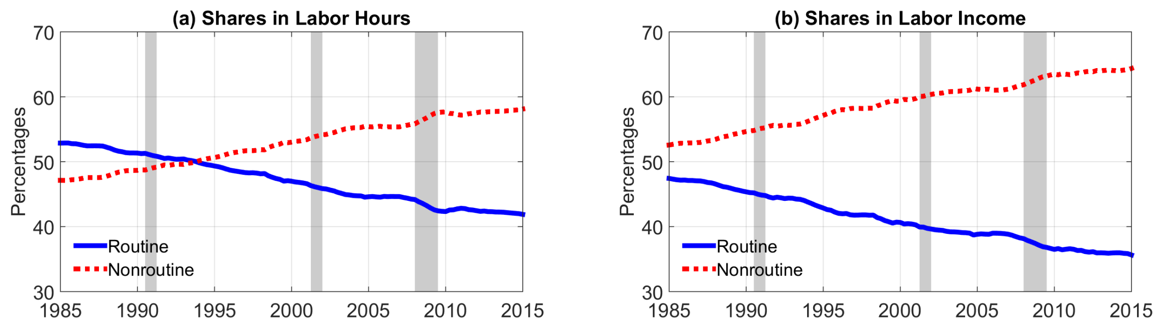


Figure 1.1 shows the divergence between routine and nonroutine occupations in terms of labor hours and income shares in the period 1985-2014. Both panels display the shares of routine and nonroutine labor within total labor input and income so that the vertical levels in each figure sum up to 1. The share of routine labor in total hours has fallen from 53 percent in 1985 to 42 percent in 2014; the share of routine income in total labor income has fallen from 47 percent to 36 percent over the same period. The decline is concentrated in the manufacturing sector, but is also widespread in retail and wholesale trade.

1.2.2 Decline of Labor Income Share

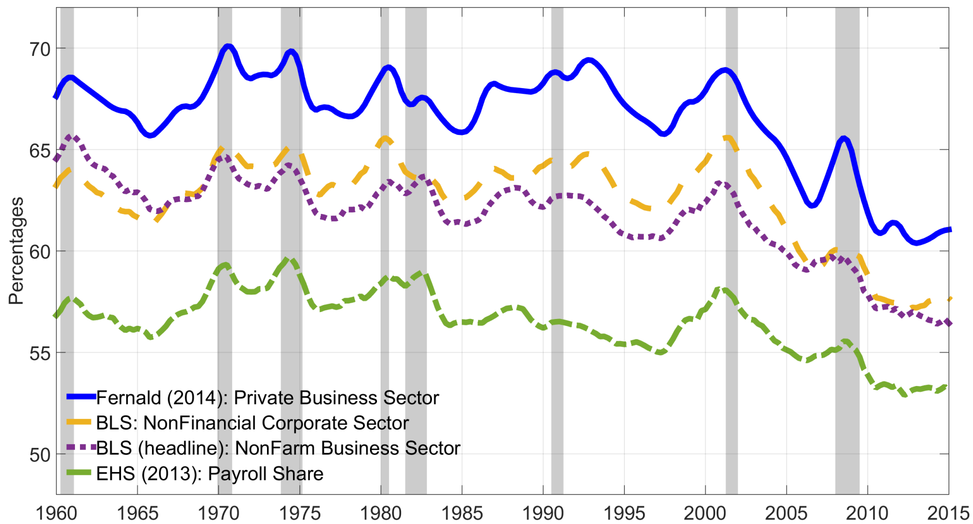
Despite its central role in macroeconomic models and policy-making, the labor share of income is difficult to measure. As a result, there has been considerable debate over whether there has been a decline of the labor share in the United States.[5]. The main is-

[5]Some of the debate focuses on the difference between the *gross* labor share and the *net* labor share. In this study, I focus on the gross labor income share as it matters for the aggregate production function and the long-run factor share stability. The net labor share, which excludes capital depreciation and taxes from total income, has shown little to no decline over time. The net income share matters for income inequality, which is beyond the scope of this study. See Rognlie (2015) and Bridgman (2014) for empirical comparisons of gross and net income shares.

sue surrounding the measurement of aggregate labor share is that the division of labor and capital income earned by the self-employed is inherently unclear, and therefore has to be imputed (Elsby, Hobijn, and Şahin, 2013; Gollin, 2002). The most commonly cited headline measure compiled by BLS suggests that the labor income share in the nonfarm business sector has declined from 64 percent in the 1980s to 58 percent in 2012-14. However, Elsby, Hobijn, and Şahin (2013) estimate that about a third of the decline of the BLS headline measure is a statistical artifact due to the underlying assumptions made to impute labor's share of proprietors' income. To be more specific, the BLS headline measure assumes that the self-employed earn the same wage as payroll employees in the same industry. Elsby, Hobijn, and Şahin (2013) show that this assumption implies a negative capital share in proprietors' income in the 1980s, suggesting that both the initial level and the subsequent decline of labor share in the headline measure have been overstated. To circumvent the measurement issue, several recent papers have focused on the payroll share in the corporate sector instead of the aggregate labor share. Using this strategy and both aggregate and industry-level data, Elsby, Hobijn, and Şahin (2013) conclude that the U.S. labor share of income declined by around 4 percent between the early 1980s and the early 2010s. Using cross-country data, Karabarbounis and Neiman (2014) document a global decline of labor share of a similar magnitude. In addition, both studies find that the majority of the decline is attributable to within-industry changes rather than to changes in industrial composition.

I construct my labor income share series for the private business sector using data from the National Income and Product Accounts (NIPA). I follow Fernald (2014) and assume that the labor share in proprietors' income is the same as that in the non-financial corpo-

Figure 1.2
Decline of Labor Income Share



rate sector.[6] Figures 1.2 displays the constructed series, as well as alternative measures for different sectors, for the period 1960-2014. As these measures show, the labor share experienced at most a mild downward trend between the 1980s and early 2000s, with substantial fluctuations. After 2002, the decline accelerated and the drop was steep. Overall, my preferred measure yields a 6 percent drop of the average labor share between the 1980s and 2010s.

It is well understood that in order to generate non-constant factor shares in the long run, two conditions must be met in the aggregate production function. First, the elasticity of substitution between capital and labor must not be 1, i.e., the aggregate production function must not be Cobb-Douglas. Otherwise, changes in factor prices in the long run will exactly offset changes in quantities, forcing constant income shares of effective capital and labor. In other words, any form of capital- or labor-augmenting technological

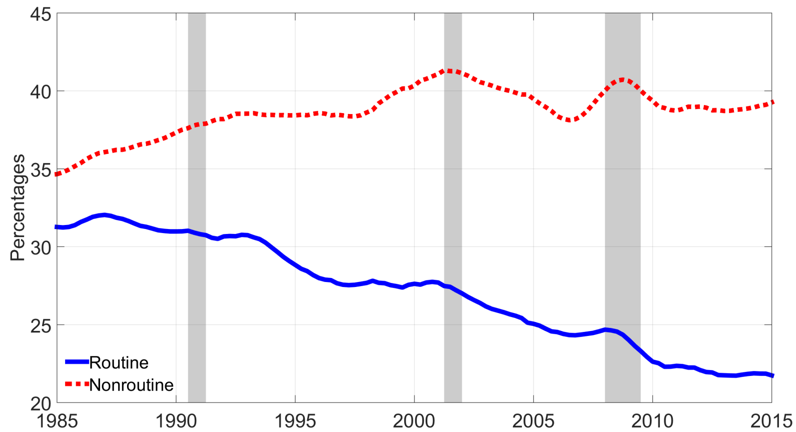
[6]The total cost equals revenue net of taxes on production and imports (TOPI), plus subsidies, plus the portion of TOPI that is properly allocated to capital (i.e. property and motor vehicle taxes).

change is effectively Hicks-neutral under the Cobb-Douglas production function. Second, the net technological change in the economy needs to be capital-augmenting. This condition stems from the simple yet fundamental asymmetry between capital and labor: capital can accumulate indefinitely, at least in theory, but per capita labor input is bounded above. When the production function is not Cobb-Douglas but the net technological change is labor-augmenting, capital accumulation (i.e., deepening) allows capital income to grow at the same pace as labor income, and thus keeps the relative factor income shares constant along a balanced growth path. The same mechanism does not exist for labor when the net technological change is capital-augmenting; as a result, factor shares can vary in the long run.

Moreover, whether a capital-augmenting technological change is *capital-biased* or *labor-biased* depends on the elasticity of substitution. When capital and labor are substitutes, capital-augmenting technological change raises the income share of capital; when capital and labor are complements, capital-augmenting technological change raises the income share of labor.

A production function with a non-unity Constant Elasticity of Substitution (CES) and capital-augmenting technological change is therefore the simplest structure for modeling changing labor shares. Karabarbounis and Neiman (2014) estimate such a production function using cross-country data on the corporate sector and find that the elasticity of substitution between capital and labor is around 1.25. They further argue that a declining price of investment goods relative to consumption goods, which is in turn driven by computerization and IT growth, explains roughly half of the overall decline in labor's share since 1975. They argue against other mechanisms including increasing markups and the changing skill

Figure 1.3
Routine and Nonroutine Labor Share in Total Income



composition of the labor force.

Elsby, Hobijn, and Şahin (2013), however, point out that the explanation for declining labor share based on capital-labor substitutability in an aggregate production function is incomplete because it is empirically inconsistent with the timing of events. Namely, the largest decline of labor share took place in the 2000s, whereas IT-driven technological progress peaked in the 1990s and possibly slowed down in the mid-2000s (see also Fernald 2014). They instead argue that globalization is more likely to be the driving force behind the recent sharp decline of labor share since the 2000s. I address this issue in Section 1.6.

Figure 1.3 applies the division of labor income between routine and nonroutine workers as in Figure 1.1b to the overall labor income share in Figure 1.2 to show the routine and nonroutine labor shares in total income. This simple decomposition reveals two important facts. First, the mild decline of the aggregate labor share is due to offsetting effects of the decline of the routine share and the rise of the nonroutine share. Second, the marked decline of the aggregate labor share after 2002 is driven primarily by a slowdown of the rise of the nonroutine share, and not by an accelerated decline of the routine share. This is consistent with the findings of Beaudry, Green, and Sand (2013), who document a decline in the demand for high-skilled workers since the early 2000s.

Figure 1.4
IT and Non-IT Capital Share in Total Income

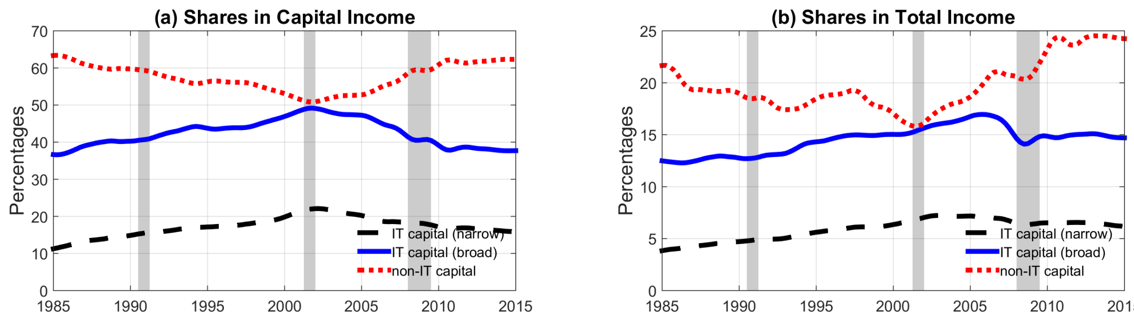
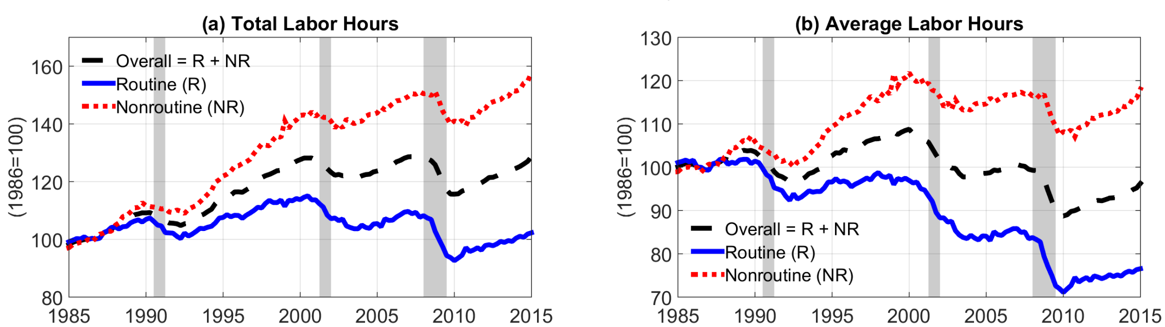


Figure 1.4 shows the decomposition of capital income over the same time period. IT capital can be narrowly defined as information processing equipment and software, or broadly defined as equipment and software. Non-IT capital includes nonresidential structures. This figure reveals that the mild decline of the labor income share is driven by the rise of IT capital share prior to the 2000s, but the sharp decline of labor share after the 2000s was driven by the increase of income accrued to structures. Moreover, panel (a) shows that the crowd-out of IT-capital income by non-IT capital started around 2002, and panel (b) shows that the share of IT-capital income within total income started to decline around 2005. This is consistent with Fernald (2014), who documents a slowdown of IT productivity growth prior to the onset of Great Recession.

1.2.3 Labor Hours: Trend vs. Business Cycles

Figure 1.5 shows the total and average hours of routine, nonroutine, and overall labor for the period of 1985-2014. Data on overall labor hours is for the U.S. nonfarm business sector as published by the BLS Productivities and Costs release. Average labor hours is defined as total hours divided by the trend labor force.[7] This measure removes the effect of population growth and most demographic changes on the evolution of labor inputs. Since in this study I focus on structural changes due to demand-side channels and economic growth due to technological changes, I use average labor hours as the main measure for labor input. All series in Figure 1.5 are normalized to 100 in 1986Q1 to allow direct comparisons of their paths.

Figure 1.5
Labor Hours over Business Cycles (1986=100)



Visual inspection yields several observations. First, after controlling for population growth and demographic changes, average routine labor hours display a clear and significant downward trend throughout the period, while average nonroutine hours grow strongly before the 2000s but flatten afterwards. This pattern of divergence between routine and

[7] I use the trend labor force measure constructed by Aaronson, Cajner, Fallick, Galbis-Reig, Smith, and Wascher (2014).

nonroutine hours echoes the divergence of their respective income shares discussed in the previous section. Second, the decline of average routine hours is not a gradual process; instead, it has followed a roughly stepwise path since the mid 1980s and was largely concentrated in the short periods surrounding the three recessions. Third, all three series of average hours are procyclical. Note that the procyclicality of average nonroutine hours does not contradict the absence of decline in the *total* nonroutine hours during recessions documented by Jaimovich and Siu (2012). Instead, it shows that the absence of decline in the total nonroutine labor hours is largely driven by population growth. Fourth, routine and overall labor hours display significant delays in recovery after the end of each recession, causing the so-called "jobless recoveries." Nonroutine hours, however, show no such delay in recovery.

1.3 A Simple Model of Unbalanced Growth

In this section, I construct a simple model of unbalanced growth to study analytically the key mechanism through which RBTC drives structural changes in the labor market. The simple model has two sectors, a "routine" sector and a "nonroutine" sector. I first derive the production function of the routine sector with direct substitution between IT capital and routine labor at producing a continuous array of tasks. I show that the commonly used CES production structure is a special case of this general form of technology. I complete the final goods production function by introducing complementarity between IT capital and nonroutine labor. I then conduct comparative statics analysis to show how IT growth drives divergence between routine and nonroutine labor hours and income shares, as well as a U-shaped growth path of the overall labor income share.

1.3.1 Routine Task Production

First, I describe the routine sector and the routine-labor-replacing aspect of RBTC in detail. The setup follows Acemoglu (2010) and Hawkins, Michaels, and Oh (2013). Throughout the paper, I use "machines", "computers", and "IT capital" interchangeably.

The intermediate product of the routine sector is a CES aggregate of an array of tasks:

$$Y_r = \left[\int_0^1 y(i)^{\frac{\phi-1}{\phi}} \, di \right]^{\frac{\phi}{\phi-1}}$$

where $\phi > 1$ is the elasticity of substitution between tasks. Any task can be performed by either a machine or routine labor:

$$y(i) = \begin{cases} m(i)/\kappa(i) & \text{if task i is performed by machine} \\ l_r(i)/\iota(i) & \text{if task i is performed by routine labor} \end{cases}$$

Here $m(i)$ and $l_r(i)$ denote the quantities of computer capital and routine labor used in completing task i. One unit of $y(i)$ is produced with either $\kappa(i)$ units of machinery or $\iota(i)$ units of routine labor. I assume that the array of tasks $i \in [0, 1]$ are ordered based on the relative productivity of machines and labor in completing them. Without loss of generality, I assume that tasks at which machines are relatively productive are ordered first. That is, $i = 0$ is the most machine-productive task and $i = 1$ is the most labor-productive. I assume, additionally, that the relative productivity of capital, $\Phi(i) = \frac{\partial y(i)/\partial m(i)}{\partial y(i)/\partial l_r(i)} = \frac{1/\kappa(i)}{1/\iota(i)} = \frac{\iota(i)}{\kappa(i)}$, is strictly decreasing in i. The monotonicity of $\Phi(i)$ and the uniform distribution of i indicate a single crossing and the existence of a critical task $i = \theta_m$, such that:

$$y(i) = \begin{cases} m(i)/\kappa(i) & \text{if} \quad i < \theta_m \\ l_r(i)/\iota(i) & \text{if} \quad i > \theta_m \end{cases}$$

At $i = \theta_m$, the firm is indifferent between using machines or labor. Competitive factor markets guarantee that the marginal cost of $y(i)$ must be the same regardless of the means of production at the critical task θ_m. That is,

$$\Phi(\theta_m) = \frac{\iota(\theta_m)}{\kappa(\theta_m)} = \frac{R_m}{w_r} \tag{1.2}$$

where w_r is the wage for routine labor and R_m is the rental rate for machines, taken as given by the firm. Everything else equal, the higher the θ_m, the larger the machines' share in the production of the routine sector. If the productivity of machines grows relative to that of routine labor over time, then in equilibrium more and more tasks will be carried out by machines.

Within the routine sector, the firm's optimization problem can be solved in two steps. First, given the total quantity of computer capital *M* and routine labor L_r, the firm chooses which tasks to perform with machines and which to perform with labor, and decides the quantities of each factor to allocate to their respective tasks. This is a static problem that yields a revenue function of the firm in terms of M and L_r. The second step is to choose M and L_r in each period to maximize profits in a dynamic setting.

I assume perfectly competitive markets. The first step of the problem is given by:

$$\max_{\{y(i)\}_{i=0}^{1},\theta_m} \quad \left[\int_0^1 y(i)^{\frac{\phi-1}{\phi}} di\right]^{\frac{\phi}{\phi-1}} \tag{1.3}$$

$$\begin{aligned} \text{s.t.} \quad M &= \int_0^1 m(i)di = \int_0^{\theta_m} \kappa(i)y(i)di \\ L_r &= \int_0^1 l_r(i)di = \int_{\theta_m}^1 \iota(i)y(i)di \end{aligned}$$

The optimality conditions of the static problem yield the optimized production function for the routine sector:

$$Y_r^* = \left[A(\theta_m)^{\frac{1}{\phi}} M^{\frac{\phi-1}{\phi}} + B(\theta_m)^{\frac{1}{\phi}} L_r^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}} \tag{1.4}$$

where:

$$A(\theta_m) = \int_0^{\theta_m} \kappa(i)^{1-\phi} di \quad \text{and} \quad B(\theta_m) = \int_{\theta_m}^1 \iota(i)^{1-\phi} di \tag{1.5}$$

To understand these terms, recall that $1/\kappa(i)$ is output per unit of capital in task i and $\phi > 1$. $A(\theta)$ is thus an aggregate over the average product of machines for the tasks they perform, and may be interpreted as an index of machine productivity. The same idea applies to $B(\theta_m)$ and labor productivity.

The critical task level θ_m is determined by the capital-labor ratio and relative productivity:

$$\frac{M}{L_r} = \frac{A(\theta_m)}{B(\theta_m)} \left[\frac{\kappa(\theta_m)}{\iota(\theta_m)}\right]^{\phi} = \frac{A(\theta_m)}{B(\theta_m)} \Phi(\theta_m)^{-\phi} \tag{1.6}$$

We can also verify that the production function satisfies constant returns to scale.

The relative productivity of machines grows exogenously due to IT progress. I assume that for any task i,

$$\Phi_t(i) = X_{m,t}\bar{\Phi}(i) \tag{1.7}$$

where $\bar{\Phi}(i)$ is a time-invariant component of the relative productivity and $X_{m,t}$ is the growth in IT. Implicit in this assumption is that IT is a "general purpose technology" (GPT), as growing computing power raises the productivity of *all* computer capital relative to routine labor.[8]

Special Case: CES between M and L_r

The task model described above gives us the micro-foundation of the aggregate production function of the routine sector. It also gives us an interpretation of the key variable θ_m as the critical task. However, due to the presence of $A(\theta_m)$ and $B(\theta_m)$, the general form of the production function (1.4) is technically difficult to work with. Here I show that CES can be embedded with certain specifications of $\bar{\Phi}(i)$. Incorporating the growth trend in computer productivity, I assume that the productivity of machines satisfies $\kappa(i,t) = (X_{m,t}\alpha_m)^{-1}\, i^{\frac{1-\epsilon}{\epsilon}}$ for task i, while the productivity of routine labor satisfies $\iota(i,t) = \iota(i) = \alpha_r^{-1}(1-i)^{\frac{1-\epsilon}{\epsilon}}$, with $0 < \epsilon < 1$ to ensure that machines and routine labor are substitutes. The relative

[8]It is also common in the literature to model the technological growth as an exogenous decline of the price of IT capital (see Autor, Levy, and Murnane (2003), Morin (2014), and Lehn (2015) for examples). With perfect competition, it can be shown that these models are isomorphic to my model and yield essentially the same evolution and dynamics of key variables.

productivity of machines is given by:

$$\Phi(i) = X_{m,t} \left(\frac{\alpha_m}{\alpha_r} \right) \left(\frac{1-i}{i} \right)^{\frac{1-\epsilon}{\epsilon}} \in (0, \infty) \quad \text{for} \quad i \in (0,1) \tag{1.8}$$

It turns out that under this specification, the optimized production function of Y_r^* in (1.4) collapses into the CES form (after scaling by a constant term):

$$Y_r^* = \left[\lambda_m \left(X_{m,t} M \right)^{\epsilon} + (1-\lambda_m) L_r^{\epsilon} \right]^{\frac{1}{\epsilon}} \tag{1.9}$$

where $\lambda_m = \frac{\alpha_m^\epsilon}{\alpha_m^\epsilon + \alpha_r^\epsilon}$. The critical task level θ_m is given by:

$$\theta_m = \frac{\lambda_m \left(X_{m,t} M \right)^{\epsilon}}{\lambda_m \left(X_{m,t} M \right)^{\epsilon} + (1-\lambda_m) L_r^{\epsilon}} \tag{1.10}$$

This CES version of the production function has been the most commonly used specification in the literature (See Autor, Levy, and Murnane (2003), Autor and Dorn (2013), and Morin (2014) for examples).

It is worth noting that under this special structure of $\Phi(i)$, the optimized production function for the routine sector no longer depends on the elasticity of substitution between tasks, ϕ; instead, it depends on the elasticity of substitution between computer capital and routine labor, which is assumed to be constant.

1.3.2 Final Goods Production

A unique final good is produced competitively by combining inputs produced by the routine and nonroutine sectors in a nested CES form:

$$Y_t = \left[\lambda_n Y_{n,t}^{\sigma} + (1-\lambda_n) Y_{r,t}^{\sigma}\right]^{\frac{1}{\sigma}} \tag{1.11}$$

$$Y_{n,t} = L_{n,t} \tag{1.12}$$

$$Y_{r,t} = \left[\lambda_m (X_{m,t} M_t)^{\epsilon} + (1-\lambda_m) L_{r,t}^{\epsilon}\right]^{\frac{1}{\epsilon}} \tag{1.13}$$

where $L_{n,t}$ is nonroutine labor input and λ_n is a scaling parameter. The key assumptions about the production structure are: (1) the elasticity of substitution between M_t and $L_{r,t}$ is $\frac{1}{1-\epsilon} > 1$, and (2) the elasticity of substitution between $L_{n,t}$ and $Y_{r,t}$ is $\frac{1}{1-\sigma} < 1$. That is, IT capital is a net substitute for routine labor and a net complement for nonroutine labor.

The model economy admits a representative household with the simple lifetime utility function

$$\sum_{t=0}^{\infty} \beta^t \ln(C_t) \tag{1.14}$$

where C_t is consumption at time t and β is the time discount factor. For simplicity, labor, L_t, is supplied inelastically and there is no population growth or other form of technological change besides $X_{m,t}$. Labor market clearing requires:

$$L_{n,t} + L_{r,t} = L_t \equiv 1 \tag{1.15}$$

and the resource constraint is given by:

$$Y_t \geq C_t + M_{t+1} - (1 - \delta_m)M_t \tag{1.16}$$

where δ_m is the depreciation rate of IT capital.

Competitive Equilibrium

I denote the prices of the $Y_{n,t}$ and $Y_{r,t}$ goods as $p_{n,t}$ and $p_{r,t}$. I normalize the price of the final good, P_t, to 1 at all times so that

$$1 \equiv P_t = \left[\lambda_n^{\frac{1}{1-\sigma}} p_{n,t}^{\frac{-\sigma}{1-\sigma}} + (1 - \lambda_n)^{\frac{1}{1-\sigma}} p_{r,t}^{\frac{-\sigma}{1-\sigma}}\right]^{\frac{1-\sigma}{-\sigma}} \tag{1.17}$$

Since labor is supplied inelastically, routine and nonroutine labor receive the same wage, w_t. A competitive equilibrium is defined in the usual way as paths for factor and intermediate goods prices $[w_t, R_{m,t}, p_{n,t}, p_{r,t}]_{t\geq 0}$, labor and capital allocations $[L_{n,t}, L_{r,t}, M_t]_{t\geq 0}$, and consumption and savings decision $[C_t, M_{t+1}]_{t\geq 0}$ such that firms maximize profits, households maximize utility, and markets clear.

Since markets are complete and competitive, we can appeal to the second welfare theorem and characterize the competitive equilibrium by solving the social planner's problem of maximizing the utility of the representative household subject to exogenous technological progress $X_{m,t}$, labor market clearing condition (1.15), and resource constraint (1.16), together with the initial condition that all real variables start with positive values. The objective function in this program is continuous and strictly concave, and the constraint set forms a convex-valued continuous correspondence. Thus the social planner's problem has a unique solution, which corresponds to the unique competitive equilibrium.

There are two steps to deriving the social planner's solution. First, given the capital and total labor level in each period, M_t and L_t, there is a static optimal division of $L_{r,t}$ and $L_{n,t}$ that maximizes Y_t so as to achieve the largest possible set of allocations that satisfy the constraint set. The second step is the dynamic choices of M_{t+1} and C_t, which depend on the intertemporal rate of substitution.

The optimal division of $L_{r,t}$ and $L_{n,t}$ is given by equalizing their marginal products, which in turn equals the competitive wage:

$$w_t = \lambda_n \left(\frac{Y_t}{Y_{n,t}}\right)^{\frac{1}{\sigma}} \left(\frac{Y_{n,t}}{L_{n,t}}\right) = (1-\lambda_n)(1-\lambda_m) \left(\frac{Y_t}{Y_{r,t}}\right)^{\frac{1}{\sigma}} \left(\frac{Y_{r,t}}{L_{r,t}}\right)^{\frac{1}{\epsilon}} \tag{1.18}$$

It is useful to define $\Theta_{n,t}$ and $\Theta_{m,t}$ as the share of nonroutine labor in total income and the share of IT capital in routine sector income respectively:

$$\Theta_{n,t} = \frac{\lambda_n L_{n,t}^{\sigma}}{\lambda_n L_{n,t}^{\sigma} + (1-\lambda_n) Y_{r,t}^{\sigma}} \quad \in (0,1) \tag{1.19}$$

$$\Theta_{m,t} = \frac{\lambda_m (X_t M_t)^{\epsilon}}{\lambda_m (X_{m,t} M_t)^{\epsilon} + (1-\lambda_m) L_{r,t}^{\epsilon}} \quad \in (0,1) \tag{1.20}$$

We can compute the overall income share of labor in the economy as

$$s_{L,t} = \Theta_{n,t} + (1-\Theta_{n,t})(1-\Theta_{m,t}) \tag{1.21}$$

with the income shares of nonroutine labor and routine labor each given by

$$s_{L_n,t} = \Theta_{n,t} \tag{1.22}$$

$$s_{L_r,t} = (1-\Theta_{n,t})(1-\Theta_{m,t}) \tag{1.23}$$

The income share of capital is therefore

$$s_{M,t} = (1-\Theta_{n,t})\Theta_{m,t} \tag{1.24}$$

Finally, we can define η_t to be the share of nonroutine labor in total labor hours:

$$\eta_t = \frac{L_{n,t}}{L_t} \tag{1.25}$$

and rewrite equation (1.18) as

$$\frac{\eta_t}{1-\eta_t} = \frac{\Theta_{n,t}}{(1-\Theta_{n,t})(1-\Theta_{m,t})} \tag{1.26}$$

The left side of equation (1.26) is the ratio of hours of nonroutine and routine labor; the right side is the ratio of income. Equation (1.26) shows that the optimal division of $L_{r,t}$ and $L_{n,t}$ occurs when these two ratios are equalized. This result depends on several simplifying assumptions, including complete and competitive markets and inelastic labor supply. Introducing wedges on the labor supply side can lead to departures from (1.26), including preference of one type of labor over another and frictions in labor adjustments. While equation (1.26) helps simplify the analysis, the results discussed below do not hinge on this exact equation.

1.3.3 Labor Market Structural Changes

In this section, I derive several analytical results regarding the growth path of the routine and nonroutine shares in hours and income, as well as the income shares of overall labor and IT capital.

Proposition 1.3.1. *In the competitive equilibrium,*

$$\frac{d\ln(\eta_t)}{d\ln(X_{m,t})} = \left[\frac{\eta_t}{1-\eta_t} \cdot \left(\frac{1-\sigma}{\epsilon-\sigma} \cdot \frac{1}{\Theta_{m,t}} - 1\right) + \frac{1-\sigma}{\epsilon-\sigma} \cdot \frac{1}{\Theta_{m,t}}\right]^{-1} > 0 \tag{1.27}$$

for $\sigma < 0$ *and* $0 < \epsilon < 1$.

Proof. Using $L_{n,t} = \eta_t L_t$ and $L_{r,t} = (1-\eta_t)L_t$, equation (1.18) can be rewritten as

$$\frac{\eta_t}{(1-\eta_t)} \cdot (1-\eta_t)^{\epsilon} L_t^{\epsilon} = \frac{\lambda_n(\eta_t L_t)^{\sigma}}{(1-\lambda_n)(1-\lambda_m)} \cdot [\lambda_m(X_{m,t}M_t)^{\epsilon} + (1-\lambda_m)(1-\eta_t)^{\epsilon} L_t^{\epsilon}]^{\frac{\epsilon-\sigma}{\epsilon}} \tag{1.28}$$

This further yields

$$X_{m,t} = \left[(1-\eta_t)\left(\frac{L_t}{M_t}\right)\lambda_m^{-\frac{1}{\epsilon}}\right]\left\{\left[\frac{\lambda_n}{(1-\lambda_n)(1-\lambda_m)}\right]^{\frac{\epsilon}{\sigma-\epsilon}}\left(\frac{\eta_t}{1-\eta_t}\right)^{\frac{\epsilon(1-\sigma)}{\epsilon-\sigma}} - (1-\lambda_m)\right\}^{\frac{1}{\epsilon}}$$

Taking derivatives

$$
\begin{aligned}
\frac{d\ln(X_{m,t})}{d\ln(\eta_t)} &= \frac{dX_{m,t}/X_{m,t}}{d\eta_t/\eta_t} \\
&= \left(\frac{1-\sigma}{\epsilon-\sigma}\right)\cdot\frac{\left[\frac{\lambda_n}{(1-\lambda_n)(1-\lambda_m)}\right]^{\frac{\epsilon}{\sigma-\epsilon}}\left(\frac{\eta_t}{1-\eta_t}\right)^{\frac{\epsilon(1-\sigma)}{\epsilon-\sigma}}}{\left[\frac{\lambda_n}{(1-\lambda_n)(1-\lambda_m)}\right]^{\frac{\epsilon}{\sigma-\epsilon}}\left(\frac{\eta_t}{1-\eta_t}\right)^{\frac{\epsilon(1-\sigma)}{\epsilon-\sigma}}-(1-\lambda_m)}\cdot\left(\frac{1}{1-\eta_t}\right)-\left(\frac{\eta_t}{1-\eta_t}\right) \\
&= \left(\frac{1-\sigma}{\epsilon-\sigma}\right)\cdot\frac{(1-\lambda_m)/(1-\Theta_{m,t})}{(1-\lambda_m)/(1-\Theta_{m,t})-(1-\lambda_m)}\cdot\left(\frac{1}{1-\eta_t}\right)-\left(\frac{\eta_t}{1-\eta_t}\right) \\
&= \left(\frac{1-\sigma}{\epsilon-\sigma}\right)\cdot\frac{1}{\Theta_{m,t}}\cdot\left(\frac{1}{1-\eta_t}\right)-\left(\frac{\eta_t}{1-\eta_t}\right) \\
&= \frac{\eta_t}{1-\eta_t}\cdot\left(\frac{1-\sigma}{\epsilon-\sigma}\cdot\frac{1}{\Theta_{m,t}}-1\right)+\frac{1-\sigma}{\epsilon-\sigma}\cdot\frac{1}{\Theta_{m,t}}
\end{aligned}
$$

Finally,

$$
\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}=\left[\frac{\eta_t}{1-\eta_t}\cdot\left(\frac{1-\sigma}{\epsilon-\sigma}\cdot\frac{1}{\Theta_{m,t}}-1\right)+\frac{1-\sigma}{\epsilon-\sigma}\cdot\frac{1}{\Theta_{m,t}}\right]^{-1}>0 \tag{1.29}
$$

as $\frac{1-\sigma}{\epsilon-\sigma}>1$ and $\frac{1}{\Theta_m}>1$, for $\sigma<0$, $0<\epsilon<1$, and $0<\Theta_{m,t}<1$. *Q.E.D.*

□

Proposition 1.3.1 states that the IT capital-augmenting technological progress $X_{m,t}$ induces labor inputs (hours) to flow from the routine sector to the nonroutine sector. The direction of the flow results directly from the assumed range of parameters, σ and ϵ, which govern the relative elasticities of substitution between IT capital and the two types of labor. In addition, the model predicts that in the long run, continued RBTC will eventually render routine labor obsolete, as all routine tasks will be performed by computers and machines, while all human labor inputs will be nonroutine in nature.

The next set of theorems consider the impact of IT progress on income shares:

Lemma 1.3.1. *In the competitive equilibrium,*

$$\frac{d\ln(\Theta_{m,t})}{d\ln(X_{m,t})} = \epsilon(1-\Theta_{m,t})\left[1+\frac{\eta_t}{1-\eta_t}\cdot\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}\right] > 0 \quad (1.30)$$

$$\frac{d\ln(s_{L_n,t})}{d\ln(X_{m,t})} = \frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} = \sigma\left(\frac{\epsilon-1}{\epsilon-\sigma}\right)\cdot\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}\cdot\frac{1}{1-\eta_t} > 0 \quad (1.31)$$

for $\sigma < 0$ *and* $0 < \epsilon < 1$.

Proof. First I show that $\frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} > 0$. Using the definition of $\Theta_{n,t}$ in equation (1.19) and taking the derivative with respect to $X_{m,t}$, we get:

$$\frac{d\Theta_{n,t}}{dX_{m,t}} = \sigma\Theta_{n,t}(1-\Theta_{n,t})\left[\frac{1}{\eta_t}\frac{d\eta_t}{dX_{m,t}} - \frac{\Theta_{m,t}}{X_{m,t}} + \frac{(1-\Theta_{m,t})}{1-\eta_t}\frac{d\eta_t}{dX_{m,t}}\right]$$

which implies that

$$\begin{aligned}\frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} &= \sigma(1-\Theta_{n,t})\left[\frac{d\ln(\eta_t)}{d\ln(X_{m,t})} - \Theta_{m,t} + (1-\Theta_{m,t})\frac{\eta_t}{1-\eta_t}\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}\right]\\ &= \sigma\left[\frac{d\ln(\eta_t)}{d\ln(X_{m,t})} - \Theta_{m,t}(1-\Theta_{n,t})\right]\end{aligned}$$

To show $\frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} > 0$ it suffices to show

$$\frac{d\ln(\eta_t)}{d\ln(X_{m,t})} - \Theta_{m,t}(1-\Theta_{n,t}) < 0 \quad (1.32)$$

as $\sigma < 0$, or equivalently

$$\frac{d\ln(X_{m,t})}{d\ln(\eta_t)} - \frac{1}{\Theta_{m,t}(1-\Theta_{n,t})} > 0 \quad (1.33)$$

One can show the following:

$$
\begin{aligned}
\frac{d\ln(X_{m,t})}{d\ln(\eta_t)} &- \frac{1}{\Theta_{m,t}(1-\Theta_{n,t})} = \left(\frac{1-\sigma}{\epsilon-\sigma}\right)\cdot\frac{1}{\Theta_{m,t}}\cdot\left(\frac{1}{1-\eta_t}\right) - \left(\frac{\eta_t}{1-\eta_t}\right) - \frac{1}{\Theta_{m,t}(1-\Theta_{n,t})} \\
&= \left(\frac{1-\sigma}{\epsilon-\sigma}\right)\cdot\frac{1}{\Theta_{m,t}}\cdot\frac{(1-\Theta_{n,t})(1-\Theta_{m,t})+\Theta_{n,t}}{(1-\Theta_{n,t})(1-\Theta_{m,t})} - \frac{1}{(1-\Theta_{n,t})}\left[\frac{\Theta_{n,t}}{(1-\Theta_{m,t})}+\frac{1}{\Theta_{m,t}}\right] \\
&= \left(\frac{1-\epsilon}{\epsilon-\sigma}\right)\cdot\frac{1}{\Theta_{m,t}}\cdot\frac{(1-\Theta_{n,t})(1-\Theta_{m,t})+\Theta_{n,t}}{(1-\Theta_{n,t})(1-\Theta_{m,t})} \\
&= \left(\frac{1-\epsilon}{\epsilon-\sigma}\right)\cdot\frac{1}{\Theta_{m,t}}\cdot\frac{1}{1-\eta_t} > 0, \quad \text{as} \quad 0<\epsilon<1, \sigma<0
\end{aligned}
$$

We can further show that

$$
\begin{aligned}
\frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} &= -\sigma\left[\frac{d\ln(X_{m,t})}{d\ln(\eta_t)} - \frac{1}{\Theta_{m,t}(1-\Theta_{n,t})}\right]\cdot\left[\Theta_{m,t}(1-\Theta_{n,t})\cdot\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}\right] \\
&= \sigma\cdot\left(\frac{\epsilon-1}{\epsilon-\sigma}\right)\cdot\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}\cdot\frac{1}{1-\eta_t}
\end{aligned}
$$

which proves the first half of Lemma 1.3.1.

I now show that $\frac{d\ln(\Theta_{m,t})}{d\ln(X_{m,t})} > 0$. Again, using the definition of $\Theta_{m,t}$ in equation (1.19) and taking the derivative with respect to $X_{m,t}$, we get:

$$
\frac{d\Theta_{m,t}}{dX_{m,t}} = \epsilon\Theta_{m,t}(1-\Theta_{m,t})\left[\frac{1}{X_{m,t}} + \frac{1}{1-\eta_t}\frac{d\eta_t}{dX_{m,t}}\right]
$$

which implies that

$$
\frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} = \epsilon(1-\Theta_{m,t})\left[1+\frac{\eta_t}{1-\eta_t}\cdot\frac{d\ln(\eta_t)}{d\ln(X_{m,t})}\right] > 0 \tag{1.34}
$$

using the result from Proposition 1.3.1 that $\frac{d\ln(\eta_t)}{d\ln(X_{m,t})} > 0$. This completes the proof of

Lemma 1.3.1.

Q.E.D. □

Lemma 1.3.1 states that IT progress raises the income share of nonroutine labor and the income share of IT capital *within* the routine sector. Using this result, we can prove the following:

Proposition 1.3.2. *In the competitive equilibrium,*

$$\frac{d\ln(s_{L_r,t})}{d\ln(X_{m,t})} = -\left[\frac{\Theta_{m,t}}{1-\Theta_{m,t}} \cdot \frac{d\ln(\Theta_{m,t})}{d\ln(X_{m,t})} + \frac{\Theta_{n,t}}{1-\Theta_{n,t}} \cdot \frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})}\right] < 0 \tag{1.35}$$

for $\sigma < 0$ *and* $0 < \epsilon < 1$.

Proof. Follows directly the definition of routine and nonroutine labor income shares in equation (1.22) and the result from Lemma 1.3.1. *Q.E.D.* □

Proposition 1.3.2 states that monotonic growth in IT productivity $X_{m,t}$ drives a monotonic decline in the routine labor share in total income. Since both types of labor receive the same wage in equilibrium, Proposition 1.3.1 already implies that the income of nonroutine labor would rise relative to routine labor due to growth in $X_{m,t}$. What Proposition 1.3.2 shows is that the routine labor share of *total* income also falls. Again, this is due to the complementarity (substitutability) of nonroutine (routine) labor and IT capital in the production function.

Proposition 1.3.3 (U-Shaped Time Path of Overall Labor Share). *In the competitive* *equilibrium,*

$$\frac{d\ln(s_{L,t})}{d\ln(X_{m,t})} = \Theta_{m,t}(1-\Theta_{n,t})\left[\frac{-\sigma(1-\epsilon)\Theta_{n,t}\Theta_{m,t}-\epsilon(1-\sigma)(1-\Theta_{m,t})}{(1-\epsilon)\Theta_{n,t}\Theta_{m,t}+(1-\sigma)(1-\Theta_{m,t})}\right] \tag{1.36}$$

and

$$\begin{aligned}
\frac{d\ln(s_{L,t})}{d\ln(X_{M,t})} &< 0, \quad \textit{when} \quad \frac{\Theta_{m,t}\Theta_{n,t}}{1-\Theta_{m,t}} = \frac{s_{L_n,t}s_{M,t}}{s_{L_r,t}} < \frac{\epsilon(1-\sigma)}{\sigma(\epsilon-1)} \\
\frac{d\ln(s_{L,t})}{d\ln(X_{M,t})} &> 0, \quad \textit{when} \quad \frac{\Theta_{m,t}\Theta_{n,t}}{1-\Theta_{m,t}} = \frac{s_{L_n,t}s_{M,t}}{s_{L_r,t}} > \frac{\epsilon(1-\sigma)}{\sigma(\epsilon-1)}
\end{aligned}$$

for $\sigma < 0$ *and* $0 < \epsilon < 1$.

Proof. Using the definition of the total labor income share in equation (1.21) and the results from Lemma 1.3.1 and Proposition 1.3.1, we have

$$\begin{aligned}
\frac{d\ln(s_{L,t})}{d\ln(X_{m,t})} &= \frac{1}{s_{L,t}}\left[\Theta_{n,t}\Theta_{m,t}\frac{d\ln(\Theta_{n,t})}{d\ln(X_{m,t})} - (1-\Theta_{n,t})\Theta_{m,t}\frac{d\ln(\Theta_{m,t})}{d\ln(X_{m,t})}\right] \\
&= \Theta_{m,t}(1-\Theta_{n,t})\left[\frac{-\sigma(1-\epsilon)\Theta_{n,t}\Theta_{m,t}-\epsilon(1-\sigma)(1-\Theta_{m,t})}{(1-\epsilon)\Theta_{n,t}\Theta_{m,t}+(1-\sigma)(1-\Theta_{m,t})}\right]
\end{aligned}$$

Note that $\Theta_{m,t}(1-\Theta_{n,t}) > 0$ and $[(1-\epsilon)\Theta_{n,t}\Theta_{m,t}+(1-\sigma)(1-\Theta_{m,t})] > 0$ for any concave production function, and the sign of $\frac{d\ln(s_{L,t})}{d\ln(X_{m,t})}$ depends on the sign of the numerator inside the bracket on the right side. Since both $\Theta_{n,t}$ and $\Theta_{m,t}$ are continuous, monotonically increasing functions in time (by Lemma 1.3.1) and have a range of $(0,1)$, $\frac{\Theta_{n,t}\Theta_{m,t}}{(1-\Theta_{m,t})} \in (0,\infty)$ is also a monotonically increasing function in time. Continuity and monotonicity in the range of $(0,\infty)$ imply a single crossing of $\frac{\Theta_{n,t}\Theta_{m,t}}{(1-\Theta_{m,t})}$ at $\frac{\epsilon(1-\sigma)}{\sigma(\epsilon-1)} > 0$. The rest of Proposition

1.3.3 then follows.

Q.E.D. $\square$

Proposition 1.3.3 is one of the most important results of this study. It states that if the economy starts in a state in which routine labor has a sufficiently large share in production and total income, i.e., $\frac{s_{L_n,t}s_{M,t}}{s_{L_r,t}} < \frac{\epsilon(1-\sigma)}{\sigma(\epsilon-1)}$, as is the case of the U.S. economy in the 1980s, then the total labor income share will first fall and then rise again in response to ongoing growth in IT productivity, following a U-shaped path. To understand this result, it is useful to note the following relation:

$$\frac{d\ln(s_{L,t})}{d\ln(X_{M,t})} = \underbrace{\frac{s_{L_r,t}}{s_{L,t}} \cdot \frac{d\ln(s_{L_r,t})}{d\ln(X_{M,t})}}_{<0} + \underbrace{\frac{s_{L_n,t}}{s_{L,t}} \cdot \frac{d\ln(s_{L_n,t})}{d\ln(X_{M,t})}}_{>0} \tag{1.37}$$

That is, the elasticity of the overall labor income share with respect to RBTC is the sum of the elasticities of routine and nonroutine labor income shares with respect to RBTC, weighted by their respective shares within overall labor income. Lemma 1.3.1 and Proposition 1.3.2 have established that IT progress causes a monotonic decrease in the routine income share and a monotonic increase in the nonroutine labor income share. If routine labor initially accounts for a sufficiently large share of the economy, then its fall will first outweigh the rise of nonroutine income, causing the total labor income share to fall. Over time and as labor continues to flow from the routine sector to the nonroutine sector (Proposition 1.3.1), there will be a turning point after which the rise of nonroutine labor will outweigh the fall of routine labor, and the total labor income share will start to rise again. In other words, the U-shaped path stems from the monotonic decline of the routine labor share being increasingly offset by the monotonic rise of the nonroutine labor share. This result suggests that the recent observed decline of the labor income share will reverse in the future. Moreover, if IT productivity continues to grow indefinitely, the total labor income share will continue to rise and approach 1 in the long run. Proposition 1.3.3 also naturally implies that the IT capital income share, $s_{M,t} = 1 - s_{L,t}$, will first rise and then fall. Furthermore, Proposition 1.3.3 states that the timing of the turning point depends on the relative magnitudes of the complementarity and the substitutability between IT capital and the two types of labor.

Last but not least, the U-shaped path for the overall labor income share implies a declining elasticity of substitution between overall labor and capital in this model:

Proposition 1.3.4 (Declining Capital-Labor EOS). *The Elasticity of Substitution between IT capital and overall labor is given by*

$$EOS(M_t, L_t) = \left\{1 + \left[\frac{-\sigma(1-\epsilon)\Theta_{n,t}\Theta_{m,t} - \epsilon(1-\sigma)(1-\Theta_{m,t})}{(1-\epsilon)\Theta_{n,t}\Theta_{m,t} + (1-\sigma)(1-\Theta_{m,t})}\right]\right\}^{-1} \tag{1.38}$$

and

$$\begin{aligned}\frac{1}{1-\epsilon} &\geq EOS(M_t, L_t) > 1, \quad \textit{when} \quad \frac{\Theta_{m,t}\Theta_{n,t}}{1-\Theta_{m,t}} = \frac{s_{L_n,t}s_{M,t}}{s_{L_r,t}} < \frac{\epsilon(1-\sigma)}{\sigma(\epsilon-1)} \\ \frac{1}{1-\sigma} &\leq EOS(M_t, L_t) < 1, \quad \textit{when} \quad \frac{\Theta_{m,t}\Theta_{n,t}}{1-\Theta_{m,t}} = \frac{s_{L_n,t}s_{M,t}}{s_{L_r,t}} > \frac{\epsilon(1-\sigma)}{\sigma(\epsilon-1)}\end{aligned}$$

for $\sigma < 0$ *and* $0 < \epsilon < 1$.

Proof. The elasticity of substitution between capital and labor takes a general form:

$$EOS(M_t, L_t) = \frac{d\ln(M_t/L_t)}{d\ln(w_t/r_t)} = \left[1 + \frac{1}{1-s_{L,t}}\frac{d\ln(s_{L,t})}{d\ln(X_{m,t})}\right]^{-1} \tag{1.39}$$

Using the result from 1.3.3, we get the expression in 1.3.4. Q.E.D. □

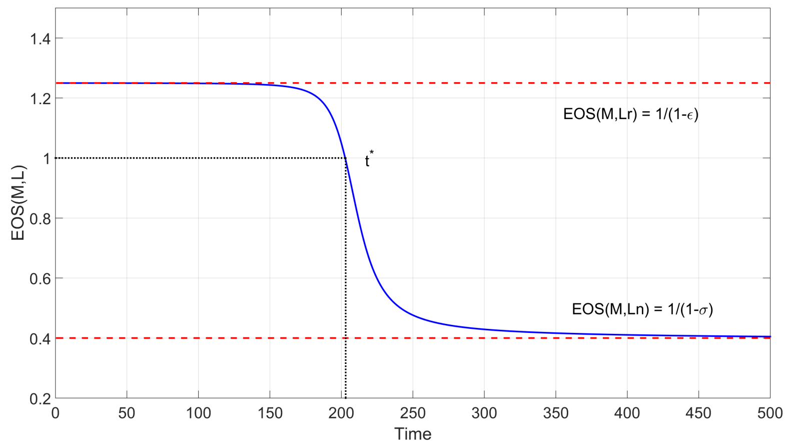
Proposition 1.3.4 provides an alternative perspective on the time-varying nature of the relationship between overall labor and capital that is central to this model of unbalanced growth. As the composition of labor shifts from routine-intensive to nonroutine-intensive tasks over time, overall labor and capital become increasingly complementary. Given that RBTC is capital-augmenting, the overall labor share declines initially when capital and labor are net substitutes and rises later when capital and labor are net complements. The economy reaches the inflection point when $EOS(M_t, L_t)$ is exactly 1. In addition, the substitutability between overall labor and capital is bounded between the substitutability between capital and routine labor and the substitutability between capital and nonroutine labor. As a result, the model yields a declining path for the elasticity of substitution between the overall labor and capital as depicted in Figure 1.6.

It is also easy to see that

$$\lim_{\theta_{n,t}\to 0} EOS(M_t, L_t) = \frac{1}{1-\epsilon} \quad \text{and} \quad \lim_{\theta_{n,t}\to 1} EOS(M_t, L_t) = \frac{1}{1-\sigma} \tag{1.40}$$

There has been a wide range of estimates of the elasticity of substitution between capital and labor in the literature (see Chirinko 2008 for a survey). Most estimates of the EOS between overall labor and capital have been less than 1, while a recent influential paper by

Figure 1.6
Declining Capital-Labor Elasticity of Substitution



Karabarbounis and Neiman (2014) documents a global decline of labor share and estimates the EOS to be around 1.25. My results provide a possible reconciliation of these two observations by emphasizing the time-varying nature of the EOS under RBTC.

1.3.4 Asymptotic State of Model Economy

The shifts of factor shares naturally imply that there is no Balanced Growth Path (BGP) in the model for $X_{M,t} < \infty$. The factor shares in the asymptotic state of the economy are described as in Proposition 1.3.5:

Proposition 1.3.5. *In the long run,*

$$\begin{aligned}
\lim_{X_{M,t}\to\infty} \eta_t &= 1 \\
\lim_{X_{m,t}\to\infty} s_{L_r,t} &= 0 \qquad \textit{and} \qquad \lim_{X_{m,t}\to\infty} s_{L_n,t} = 1 \\
\lim_{X_{M,t}\to\infty} s_{L,t} &= 1 \qquad \textit{and} \qquad \lim_{X_{M,t}\to\infty} s_{m,t} = 0
\end{aligned}$$

and production function:

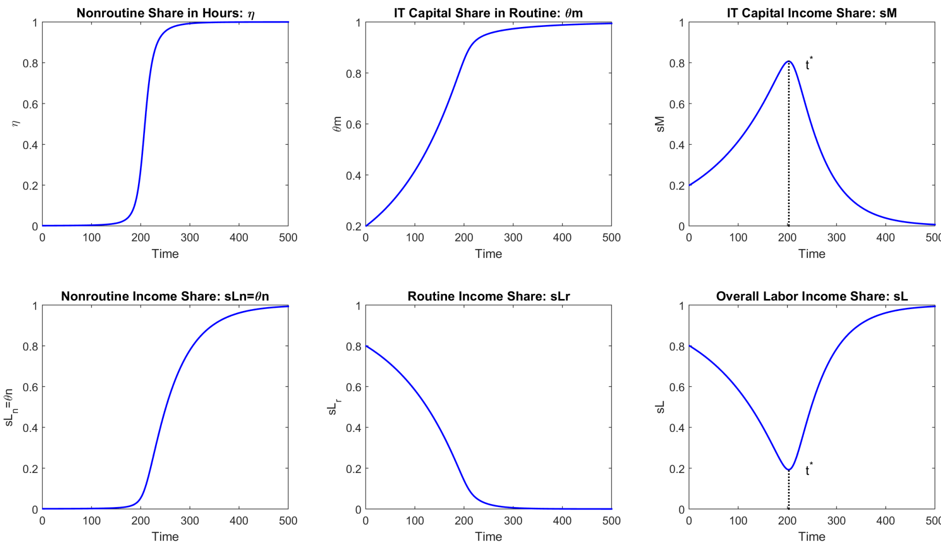
$$\lim_{X_{M,t}\to\infty} Y_t = \left[\lambda_n L_{n,t}^{\sigma} + (1-\lambda_n)(X_{M,t}M_t)^{\sigma}\right]^{\frac{1}{\sigma}}, \quad \sigma < 0$$

Proof. Both $L_{r,t}$ and $L_{n,t}$ are bounded above at 1. By Equation 1.19, $\lim_{X_{M,t}\to\infty} \Theta_{m,t} = 1$ and $\lim_{X_{M,t}\to\infty} \Theta_{n,t} = 1$. The rest of the results follow by definitions. *Q.E.D.* □

Proposition 1.3.5 states that in the long run, continued growth in RBTC will eventually render routine labor obsolete, as all routine tasks will be performed by computers and machines, while all human labor inputs will be nonroutine in nature. Moreover, because nonroutine human inputs are a complement for the routine tasks performed by computers, technology is labor-biased in nature and labor share of income will rise and eventually approach one in the long run.

The long-run properties of the model economy presented in Proposition 1.3.5 are driven by the exogenous nature of the technological progress in the simple model. In reality, however, the IT capital income share is unlikely to decline indefinitely, because technologies are created by firms choosing to invest in R&D, and the endogenous production of technology will slow down eventually as the IT capital income share declines. We conjecture that adding a "self-correcting" mechanism between endogenous technological progress and capital income to the model would create another inflection point on the right side of the U-shaped curve, after which the rise of the labor share would slow down and approach a level less than one asymptotically. Understanding fully the feedback between technological change and the transitional dynamics would require an endogenous growth model (see Acemoglu and Restrepo, 2015 for example), which is beyond the scope of this study.

Figure 1.7
Simple Model Simulation: Time Paths of Key Structural Variables



In my quantitative exercise below, I take into account a potential eventual slowdown of technology by fitting the data on recent IT progress to a logistic function with a long-run limit.

1.3.5 Simple Model Simulation

To help readers visualize the analytical results presented above, I simulate the simple model and show the dynamic time paths of the key structural variables (Figure 1.7) as well as the capital-labor elasticity of substitution (Figure 1.6). In this exercise, I use reasonable parameter values whenever possible, but set the initial factor shares to extreme values in order to show the entire range of values for the model variables. In particular, I solve the model using the following parameterization: $\beta = 0.985$, $\delta_m = 0.04$, $\epsilon = 0.2$, $\sigma = -1.5$, $\lambda_m = 0.1548$, $\lambda_n = 2.9292e-08$, and $X_{m,t} = 1.03X_{m,t-1}$ with $X_{m,0} = 1$. Parameters $\beta, \delta_m, \epsilon$ and σ are set to the same values as in the full quantitative model discussed in detail in the next section. Parameters λ_m and λ_n are calibrated so that the initial factor shares are $s_{M,0} = 0.2$, $s_{L_r,0} = 0.799$, and $s_{L_n,0} = 0.001$. Under this parameterization, nearly all labor in the model economy is routine initially. I simulate the model for 600 periods until it approaches its asymptotic state, in which nearly all labor has become nonroutine and the overall labor income share approaches 1.

1.4 Full Quantitative Model: Frictionless Baseline

In this section, I describe the full quantitative model, in which a few extra ingredients are introduced in order to take the model to the data. First, I add a second type of capital, "non-IT capital," to the final goods production function, as it accounts for more than 20 percent of national income in the data. Second, I assume that the household labor supply is elastic, which allows me to study fluctuations of overall labor hours. Third, I introduce aggregate Total Factor Productivity (TFP) shocks to generate business cycles. I continue to assume that markets are complete and competitive and thus I only describe the social planner's problem.

1.4.1 Setup

Final goods production takes four inputs: IT capital, M; non-IT capital, K; routine labor, L_r; and nonroutine labor, L_n. Production is also subject to an aggregate productivity shock, z_t, in each period. The production function of the final good is:

$$Y_t = z_t K_t^{\alpha} \left\{ \lambda_n L_{n,t}^{\sigma} + (1 - \lambda_n) \left[\lambda_m \left(X_{m,t} M_t \right)^{\epsilon} + (1 - \lambda_m) L_{r,t}^{\epsilon} \right]^{\frac{\sigma}{\epsilon}} \right\}^{\frac{1-\alpha}{\sigma}} \tag{1.41}$$

By adding non-IT capital as a Cobb-Douglas term in the production function, all the analytical results derived above continue to hold. It is easy to see that adding K does not affect the mechanism through which RBTC drives the divergence between L_r and L_n. We can further show that the U-shaped growth path for the overall labor income share is also preserved if the elasticity of substitution between K and overall labor is always 1. Recall that the labor income share follows a U-shaped path because the elasticity of substitution between IT capital and overall labor falls over time, from more than 1 to less than 1. Since the elasticity of substitution between K and labor is always equal to 1 under Cobb-Douglas, the elasticity of substitution between *overall* capital - including both M and K - and overall labor in the full model follows a similar path, yielding a U-shaped path for the labor income share in the long run.

I assume that the exogenous RBTC process $X_{m,t}$ follows a logistic growth path:

$$X_{m,t} = \frac{\bar{X}_m}{1 + e^{-\mu(t-\tau)}} \tag{1.42}$$

where $X_{m,0} = 1$. Under a logistic growth path, IT productivity grows approximately exponentially at the initial stage. After reaching an inflection point at time τ, the growth rate starts to decline and $X_{m,t}$ eventually stabilizes and approaches $\bar{X}_m$ asymptotically.

The TFP shock z_t follows an AR(1) process:

$$\log(z_t) = \rho_z \log(z_{t-1}) + u_t \tag{1.43}$$

with $0 < \rho_z < 1$ and $u_t \sim N(0, \sigma_z^2)$. I abstract from TFP growth and other forms of technological change in the baseline model for the sake of a clean exposition of the effects of RBTC.

The representative household's lifetime utility function is given by:

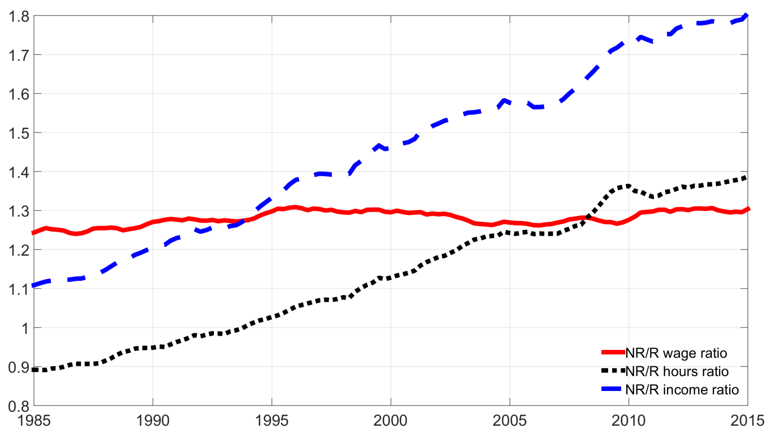
$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \ln(C_t) - \chi(L_{r,t} + \rho L_{n,t}) \right\} \tag{1.44}$$

where parameter χ governs the relative utility of consumption and leisure, and ρ governs the relative disutility from supplying routine and nonroutine labor. The functional form of the utility function has several important implications regarding equilibrium wages and labor supply functions. First, the utility function assumes a constant ratio between disutilities from supplying routine and nonroutine labor, which in turn implies a constant ratio between routine and nonroutine wages in equilibrium in the baseline frictionless model. This simplifying assumption is motivated by the relative constancy of relative wages in the data (Figure 1.8).[9] Second, assuming that in the long-run $X_{M,t}$ stabilizes, the utility function implies that the substitution effect of the permanent real wage growth on the supply of labor bundle $(L_r + \rho L_n)$ exactly offsets the income effect. As a result, $(L_r + \rho L_n)$ will stabilize in the long run. This means that if and only if $\rho = 1$, then the overall labor $L = (L_r + L_n)$ is stable along the transitional path. In the empirically relevant case that $\rho > 1$, the overall labor $L = (L_r + L_n)$ will be decreasing along the transition path and then

[9] The constancy of the relative wage between routine and nonroutine labor masks substantial movements of the relative wage between the high-skilled nonroutine-cognitive workers and the middle-skilled routine workers, as well as the relative wage between the routine workers and low-skilled nonroutine-manual workers, in different sub-periods. Moreover, it is likely to be driven by the way topcoded income is treated in the CPS. Following the labor literature, I multiplied topcoded income — mostly nonroutine — in the CPS by a factor of 1.5. However, since the topcoding only changes infrequently, the procedure does not completely alleviate the bias of topcoding over time.

stabilize at a level permanently lower than its initial value as the long-run $X_{M,t}$ stabilizes.[10] Lastly, note that this decline of long-run labor hours is not due to the linear nature of the labor disutility in (1.44); instead, it results from the wedge between L_n and L_r and the changing composition of L. In other words, using logarithm or power disutility functions for labor in (1.44) will yield the same long-run results.

Figure 1.8
Relative Wages of Routine and Nonroutine Labor



Finally, the resource constraint of the economy is given by:

$$\begin{aligned} C_t &+ K_{t+1} - (1-\delta_k)K_t + M_t - (1-\delta_m)M_t \\ &= z_t K_t^{\alpha} \left\{ \lambda_n L_{n,t}^{\sigma} + (1-\lambda_n) \left[\lambda_m \left(X_{m,t} M_t \right)^{\epsilon} + (1-\lambda_m) L_{r,t}^{\epsilon} \right]^{\frac{\sigma}{\epsilon}} \right\}^{\frac{1-\alpha}{\sigma}} \end{aligned}$$

where δ_m and δ_k are the depreciation rates of the two types of capital.

[10] To see this, suppose that the utility function (1.44) is maximized subject to a static budget constraint $w_r L_r + w_n L_n = C$. The first order conditions with respect to the intra-temporal routine/nonroutine labor-consumption choices are then given by $w_r = \chi C$ and $w_n = \rho\chi C$. Putting these three equations together, we get $L_r + \rho L_n = \frac{1}{\chi}$ and $\Delta L_r = -\rho \Delta L_n$. If and only if $\rho = 1$ does the increase in L_n exactly offset the decrease in L_r. If $\rho > 1$ and $\Delta L_r < 0$, then $\Delta L = \Delta L_r + \Delta L_n < 0$.

1.4.2 Calibration

I calibrate the model to U.S. quarterly data in the period 1986-2014. I choose 1986 as the starting point because the output gap was low at that time, so the economy could be thought of as approaching steady state. The complete list of calibrated parameters is in Table 1.2.

Table 1.2
Benchmark Calibration of Model with RBTC and Unbalanced Growth

Parameters	Values	Calibration Targets or Sources
β: Households' discount factor	0.985	Average annual capital return = 4%
α: Non-IT capital income share	0.20	Average non-IT capital income share
δ_k: Non-IT capital depreciation	0.015	Average non-IT capital depreciation
δ_m: IT capital depreciation	0.045	Average IT capital depreciation
χ: Relative weight of labor disutility	2.273	Labor hours = 0.33 in 1986Q1
ρ: Relative disutility from L_n, L_r	1.251	NR/R wage ratio in 1986Q1
λ_m: Scaling parameter	0.223	Routine income share = 0.33 in 1986Q1
λ_n: Scaling parameter	0.254	Nonroutine income share = 0.34 in 1986Q1
ϵ: Substitution between M and L_r	0.20	EOS=1.25, Karabarbounis & Neiman (2014)
σ: Substitution between M and L_n	-1.50	EOS=0.40, model fit of η_t from 1986-1990
μ: Growth rate of RBTC	0.014	Logistic function fitting of DiCecio (2009)
τ: Inflection point of RBTC	270	Logistic function fitting of DiCecio (2009)
$\bar{X}_m$: Asymptotic level of RBTC	29.48	Logistic function fitting of DiCecio (2009)
ρ_z: Persistence of TFP shocks	0.90	Estimation using output data, 1986-2014
σ_z: Variance of TFP shocks	0.045	Estimation using output data, 1986-2014

There are a total of 15 parameters in the baseline model, and I divide them into four groups. The first group of parameters, including β, α, δ_k, and δ_m, are calibrated by matching the average value of the associated targets over the entire 1986-2014 period. First, I set the household discount factor $\beta = 0.985$ to match a long-run capital return of 4 percent per year. Next, I calibrate the parameters related to the two types of capital. I use data from the detailed Fixed Assets Accounts and National Income and Product Accounts (NIPA) published by the Bureau of Economic Analysis (BEA) to construct empirical counterparts of IT and non-IT capital. Since I focus on technological changes that increase the substitutability of IT capital and routine labor, the main criterion I use in distinguishing M and K in the data is whether a type of fixed asset performs tasks that can potentially be done by labor. By this criterion, M should include not only high-tech capital such as information processing equipment and software, but also various types of industrial equipment that have been increasingly computerized and are central to automation. Therefore, in the baseline calibration, I include both nonresidential equipment and software as M and nonresidential structures as K.[11] Under this categorization, the non-IT capital income share $\alpha = 0.2$. The average annual depreciation rate is approximately 18 percent for IT capital and 6 percent for non-IT capital over the period 1986-2014. Thus, I calibrate the quarterly depreciation rates $\delta_m = 0.045$ and $\delta_k = 0.015$.

The second group of parameters, including χ, λ_m, λ_n, and ρ are calibrated by matching the value of the associated targets at the start of 1986. As the model is time-dependent, this allows me to match the initial points of the time paths for labor hours and income shares to the data and assess the model's ability to match the subsequent growth paths. In particular, I calibrate $\chi = 2.273$ to set initial overall labor hours to $L = 0.33$.[12] I set the two scaling parameters $\lambda_m = 0.223$ and $\lambda_n = 0.254$ to match the initial shares of routine and

[11] All the results discussed below are robust to using narrower definitions of M.

[12] In theory, since the model is non-stationary and solved using backward induction, one cannot calibrate parameter values to target objects in the first period without solving the entire time paths, which is very costly. Therefore, in practice, I calibrate these parameters by constructing a pseudo "steady state," shutting down both RBTC growth and TFP shocks and computing the "steady state" of the economy as if in a stationary environment in period 1. The bias from this approximation is negligible and has near zero impact on the results.

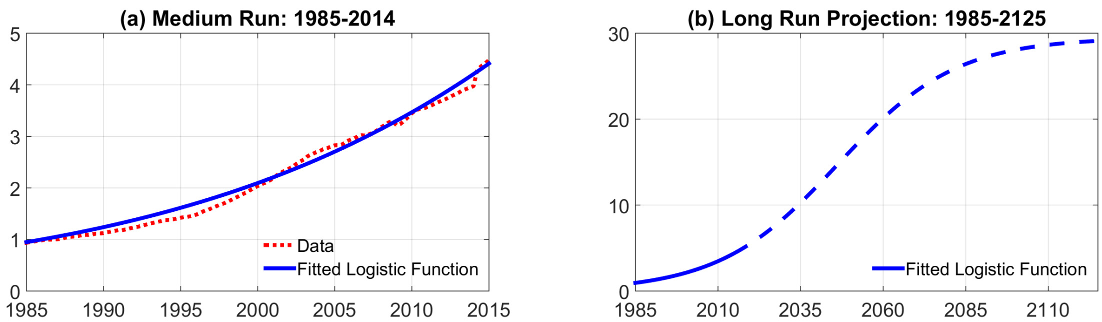
nonroutine labor in total income. I set parameter $\rho = 1.25$ to match the initial nonroutine and routine wage ratio as shown in Figure 1.8.

The third group of parameters includes ϵ and σ, which govern the elasticities of substitution between IT capital and the two types of labor, and the three parameters that govern the exogenous RBTC progress. These parameters are of central importance to the dynamics of the model and have been discussed at the end of Section 1.3.3. When taking into consideration the distinction between IT and non-IT capital (or between equipment and structures), estimates of the elasticity between IT capital and labor are consistently between 1 and 2 (Tevlin and Whelan (2003), for example). In the baseline calibration, I set the EOS between IT capital and routine labor to $\frac{1}{1-\epsilon} = 1.25$, which is the estimate of Karabarbounis and Neiman (2014) for the EOS between overall capital and labor and therefore serves as a conservative value for the EOS between M and L_r.[13] To calibrate the elasticity of substitution between M and L_n, I follow the strategy of Lehn (2015) and set it to match the growth path of η_t between 1986 and 1990. I then assess the model's ability to match labor market structural shifts after 1990. The calibrated elasticity of substitution between M and L_n is $\frac{1}{1-\sigma} = 0.4$. Both elasticities are well within range of the estimates in the literature.

I use the inverse of the relative price of equipment measured by DiCecio (2009) to proxy for the exogenous $X_{m,t}$ process. I fit the series to a logistic function as in (1.42) and obtain values for parameters μ, τ and $\bar{X}_m$. The fitted logistic growth path is shown in Figure 1.9a. I then extrapolate the series to 600 periods (1.9b) to study long-run projections of the model.

[13] A high value in the literature is 1.67, estimated by Krusell, Ohanian, Ríos-Rull, and Violante (2000) as the EOS between equipment and unskilled labor.

Figure 1.9
Exogenous RBTC Process



The last group of parameters, ρ_z and σ_z, govern the AR(1) process of TFP shocks. To set their values, I first solve the model and use the cyclical components of output from the data to back out the z_t series.[14] This way, the cyclical components of output generated by the model matches those in the data, which allow me to focus on the trend movements later. I then use the extracted z_t series and estimate that $\rho_z = 0.9$ and $\sigma_z = 0.045$.

1.4.3 Computation Method

Since the model does not have a balanced growth path and therefore cannot be "de-trended" by $X_{m,t}$, I solve the model over a finite horizon with backward induction. The algorithm is akin to the Extended Function Path method studied in Maliar, Maliar, Taylor, and Tsener (2015). In period 0, agents learn that $X_{m,t}$ will start growing and stop after T periods. T is chosen to be large enough such that $X_{m,t}$ would have approached its long-run level. I first solve for the value function in the last period and then iterate backwards. The solution to the model is the entire set of value functions and decision rules along the time path. In addition, I adapt the Envelope Condition Method of Arellano, Maliar, Maliar, and Tsyrennikov (2013) into the backward induction routine to solve the frictionless baseline

[14]The cyclical components of output are obtained by applying an HP filter with $\lambda_{hp} = 1600$.

model. I use value function iteration to solve the model with nonconvex labor adjustment cost. Lastly, I use the Smolyak Method of Judd, Maliar, Maliar, and Valero (2014) to deal with the moderately high dimensions.

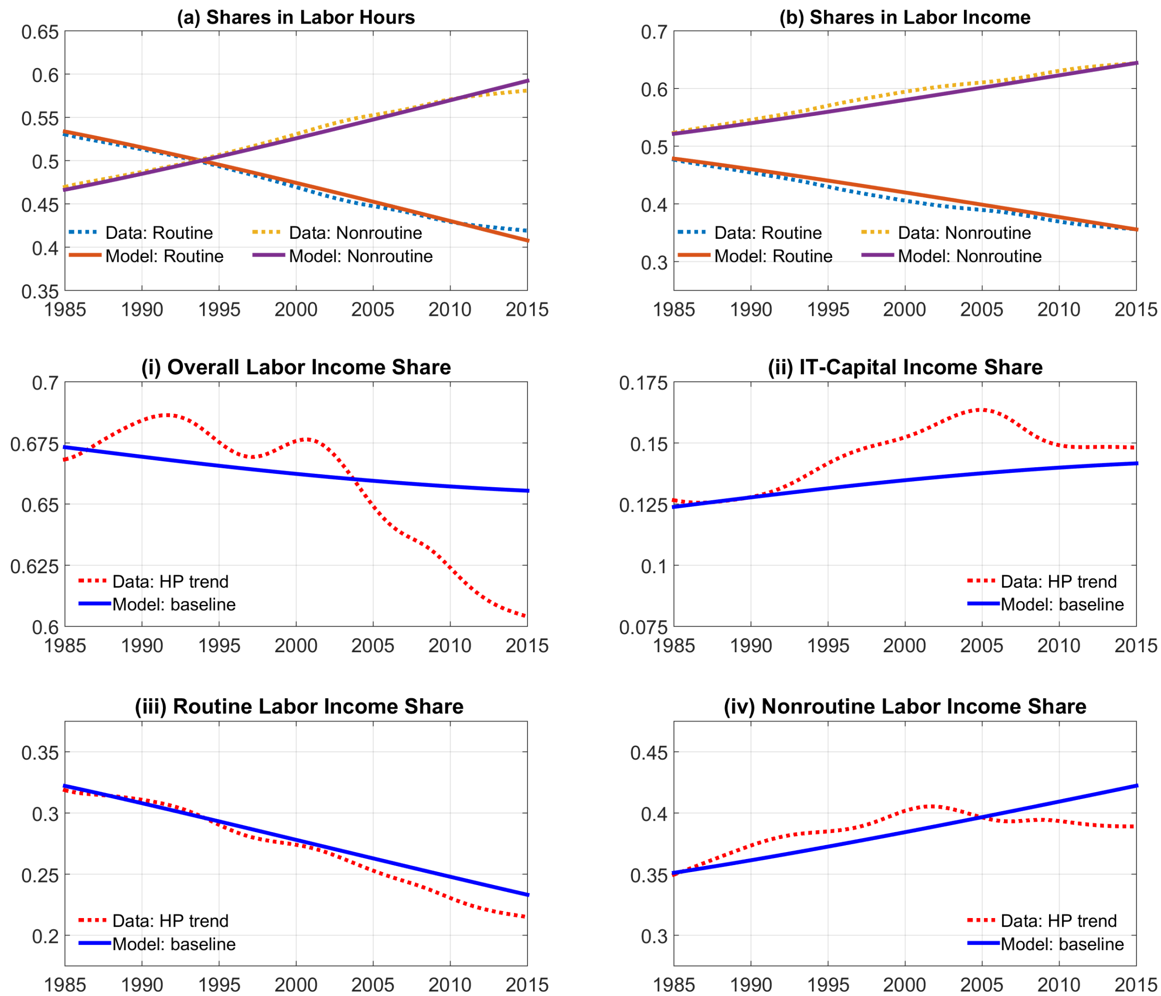
1.4.4 Results

I now present results from the frictionless baseline model and assess the model's ability to match the medium-run labor market trends in the data. I then discuss the model's long run predictions, extrapolating the IT growth process into the future.

Figure 1.10 describes the model's quantitative performance at matching the observed U.S. labor market trends in the period 1986-2014. First and foremost, the model does a good job at matching the divergence between routine and nonroutine shares in labor hours and labor income, which provides strong evidence that IT growth or RBTC is the main driving force behind the divergence of labor market outcomes between the two groups of workers.

Second, the model generates a mild decline of the overall labor income share, consistent with the trend in the data observed from the mid-1980s to the early 2000s. It cannot, however, account for the sharp decline of labor's share observed after 2002. In addition, because the income share of non-IT capital is assumed to be constant, the model generates a corresponding, mildly increasing trend in the IT capital income share, which is qualitatively consistent with the data. Note that the problem with labor share cannot be fixed by simply allowing a non-constant share of non-IT capital in this model. If we assume (realistically) that non-IT capital is a relative complement for IT capital and both types of labor, RBTC would cause its income share to rise over time. However, in the data, the non-IT capital

Figure 1.10
Model Performance: Medium-Run Structural Changes (1986-2014)



income share fell mildly from the mid-1980s through the early 2000s. After the early 2000s, it rose sharply, as the income share of structures soared during the housing boom. See Rognlie (2015) for details on this issue.

Third, the model does a good job at matching the decline of the routine labor share of total labor income and the corresponding rise of the nonroutine income share, in particular before the early 2000s. The model does not capture the accelerated decline of the share of routine labor in total income or the slowdown and mild fall of the nonroutine labor share of total income after the early 2000s. Together with the fact that RBTC explains most of the divergence between routine and nonroutine labor throughout the sample period, one can infer that another force is behind the accelerated fall of routine labor's share and the marked slowdown of nonroutine labor's share after the early 2000s. Put differently, a model with one smooth monotonic trend in technology cannot possibly match the time-varying medium-run trends in factor shares observed in data.

Figure 1.11
Model Performance: Long-Run Projection of Labor Share

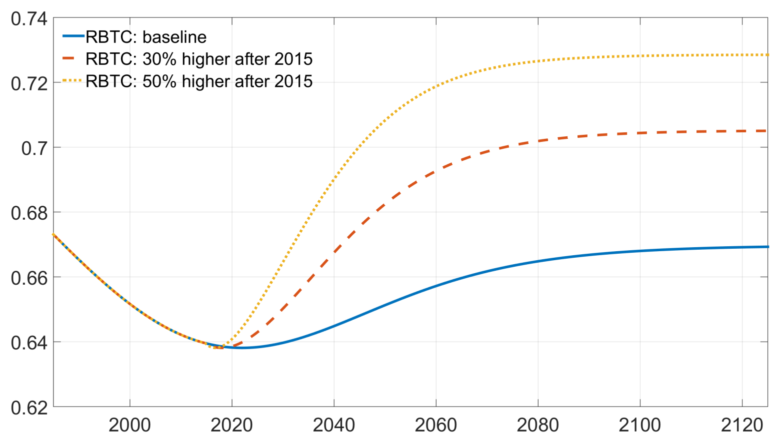
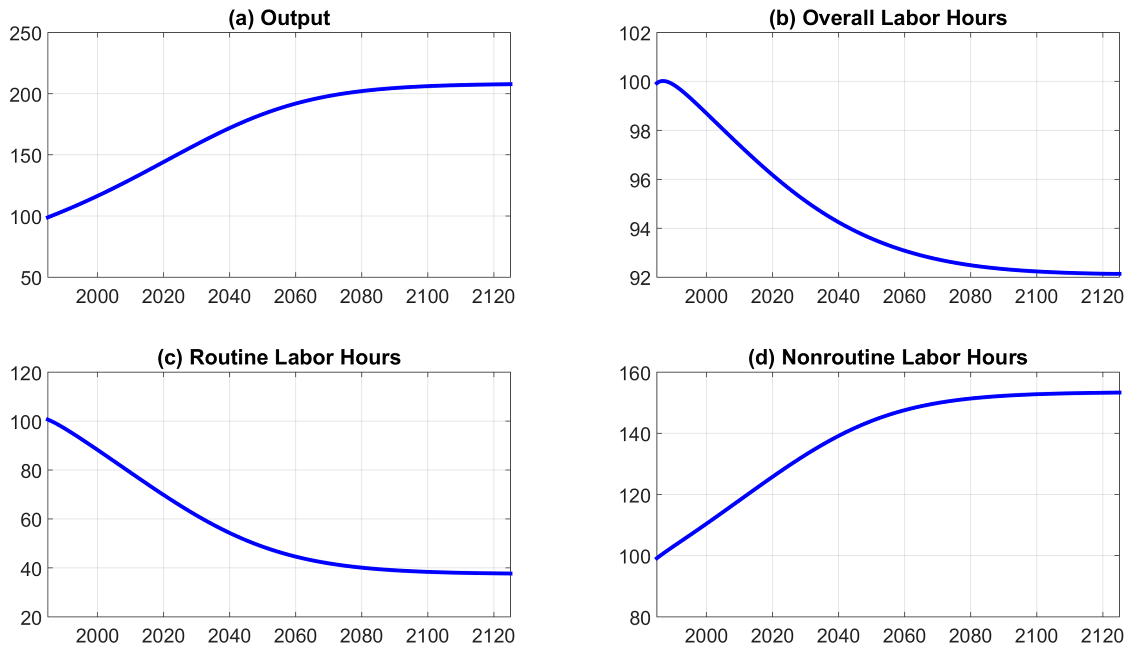


Figure 1.11 shows the long-run projections of overall labor shares under the extrapolated logistic growth path of RBTC depicted in Figure 1.9 ("baseline") as well as two alternative paths with higher $X_{m,t}$ growth rates after 2015. Several observations can be made. First, the U-shaped growth pattern is evident in all three series. Second, *ceteris paribus*, the faster the growth rate of $X_{m,t}$, the sooner the inflection point in the labor share path, and the higher the asymptotic level of $X_{m,t}$, the higher the asymptotic level of labor's share. In the baseline projection in which the $X_{m,t}$ path is extrapolated from data, labor's share stops falling around 2030. The caveat is, of course, that the projection does not take

into account the fact that labor's share fell much more in the data than in the model during the 2000s due to reasons that RBTC cannot explain.

Figure 1.12
Model Performance: Long-Run Projections of Output and Labor Hours



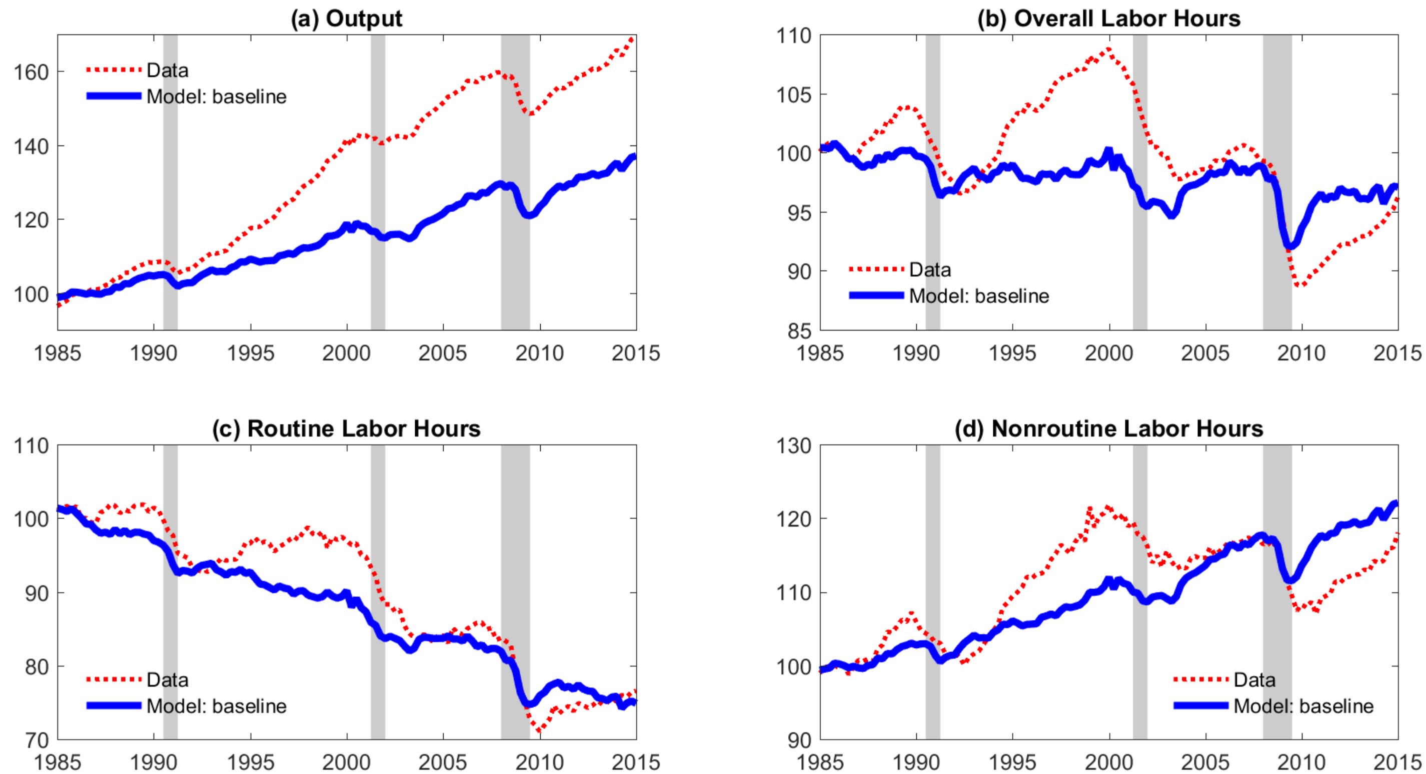
Lastly, Figure 1.12 shows the model's long-run projections of output growth and labor hours. Note that RBTC can only account for about 60% of output growth between 1986-2014 (as shown in Figure 1.13a). Thus, Figure 1.12(a) only depicts the portion of output growth that is driven by exogenous IT progress, which is assumed to continue grow logistically. The model also predicts that routine (nonroutine) labor hours decline (rise) monotonically before the exogenous RBTC process slows down. More importantly, the calibrated model predicts that overall labor hours, which is the sum of routine and nonroutine hours, declines in the long run. This is in contrast with the predicted evolution of the overall labor income share, which follows a U-shaped path. The difference is that the income shares are completely pinned down by the production side if factors are paid their marginal products, but the equilibrium levels of labor hours are determined by labor demand and supply jointly. As discussed before, the utility function allows the income effect of permanent wage growth on labor supply to outweigh substitution effect along transition path, resulting in a permanently lower level of overall labor hours.

1.5 Interactions Between Trends and Cycles: The Role of Labor Adjustment Costs

In this section, I explore the interactions between the secular trend and business cycles. Traditionally, economic growth and business cycles are often treated as separate subjects and studied in different models. The underlying assumption is that factors that drive booms and busts in economic activities are independent of the forces that drive long-run growth. I show that this is not the case for IT progress. In the presence of short-run frictions, such as nonconvex labor adjustment costs, the long-run decline (rise) of routine (nonroutine) labor may give rise to endogenous fluctuations.

Figure 1.13 shows the time paths of output and labor hours generated by the frictionless baseline model presented in the previous section. The TFP shocks used in the simulation are backed out by matching the cyclical components of output to the data. The frictionless baseline model does not generate interesting interactions between the secular trend and business cycles. This is not surprising, as without frictions, output and labor hours fluctuate around their respective long-run trends, with the cyclical components identical to those from a model without exogenous RBTC growth (not shown here). Since the inevitable trend adjustments are costless in the short run, they are instant. As a result, the frictionless baseline model does not capture the observed stepwise pattern of the routine labor decline

Figure 1.13
Output and Labor Hours Over Business Cycles: Frictionless Baseline Model



(Figure 1.13c), which is not surprising. In addition, the underlying trend components for routine and nonroutine labor hours offset each other, so the frictionless model does not generate the large swings of overall labor hours observed in the data (Figure 1.13b).

Short-run frictions are necessary to induce interesting interactions between the trend and the cycle. To study frictions, I introduce a simple form of nonconvex labor adjustment cost and assess its impact on the model's performance. I assume that the adjustment cost takes the form of Organizational Disruption as follows:

$$AC(L_{r,t-1}, L_{r,t}) = \begin{cases} 0 & \text{if} \quad L_{r,t} = L_{r,t-1} \\ \omega Y_t & \text{if} \quad L_{r,t} \neq L_{r,t-1} \end{cases} \tag{1.45}$$

That is, adjustment of routine labor disrupts the production process and incurs a loss of output of ωY_t.[15] Organizational Disruption is a form of nonconvex adjustment cost, which induces inaction regions in the labor hours decision rule. Moreover, since the size of the adjustment cost depends on the output level, it is more costly to adjust labor during booms than during recessions. Cooper, Haltiwanger, and Willis (2015) and Cooper and Willis (2009) find that a model with disruption costs provides a better match of employment dynamics in plant-level data than a model containing fixed or quadratic adjustment costs.

The resource constraint of the economy is given by:

$$\begin{aligned} C_t \quad &+ \quad K_{t+1} - (1-\delta_k)K_t + M_t - (1-\delta_m)M_t + AC(L_{r,t-1}, L_{r,t}) \\ &= \quad z_t K^{\alpha} \left\{ \lambda_n L_{n,t}^{\sigma} + (1-\lambda_n)\left[\lambda_m \left(X_{m,t} M_t\right)^{\epsilon} + (1-\lambda_m) L_{r,t}^{\epsilon}\right]^{\frac{\sigma}{\epsilon}} \right\}^{\frac{1-\alpha}{\sigma}} \end{aligned}$$

[15] For simplicity, I have abstracted from other types of adjustment cost, including both nonroutine labor adjustment cost and capital adjustment cost.

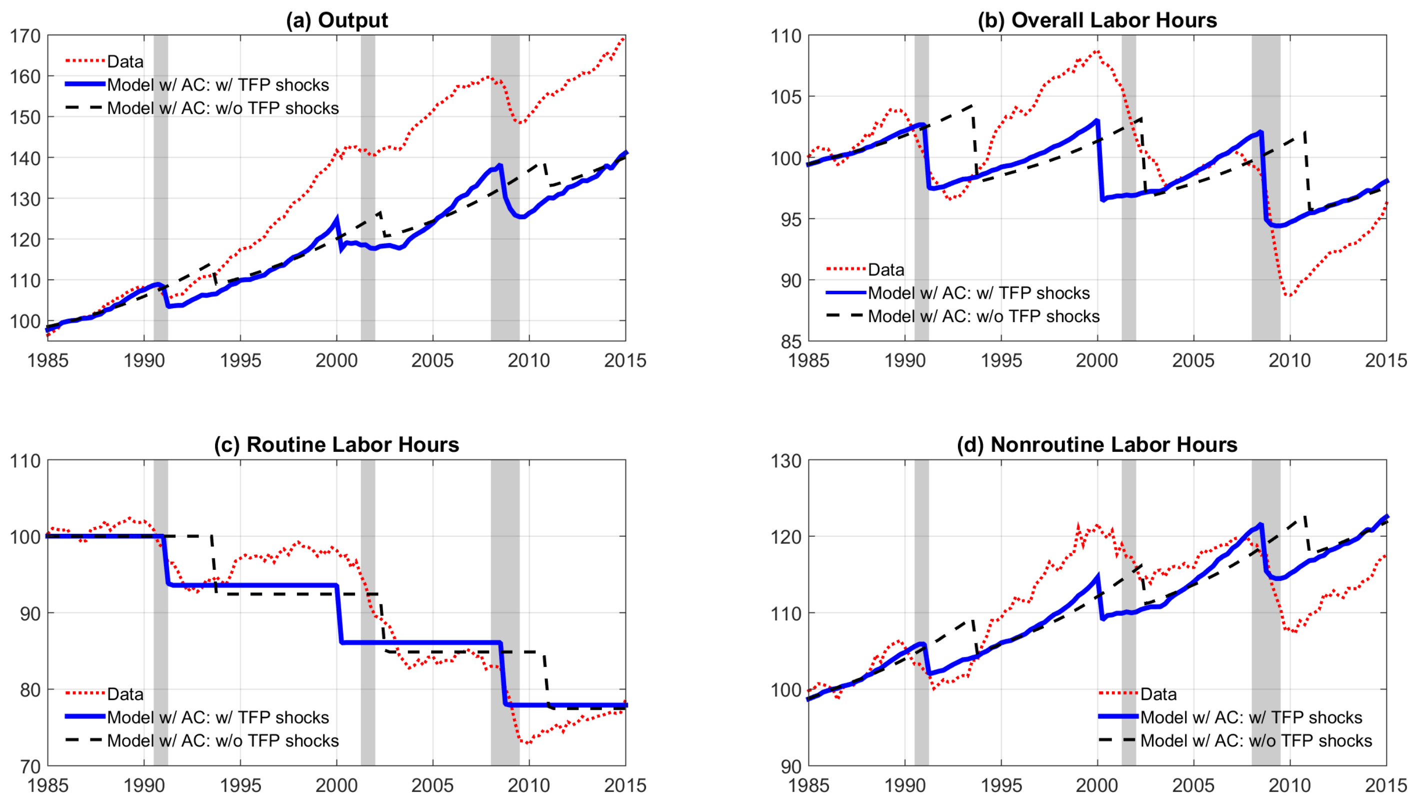
To calibrate the parameter ω that governs the size of the nonconvex labor adjustment cost, I refer to Cooper, Haltiwanger, and Willis (2015), who find that a model with a disruption cost of around 0.8 percent of output at the plant level can match the observed moments in the aggregate employment and hours data. Since in this model I assume an adjustment cost for routine labor only, I scale their estimate by the routine labor income share in 1986 and obtain $\omega = 0.00275$.

Results

Figure 1.14 shows sample paths generated by models with routine labor disruption costs, both with (blue line) and without (black dashed line) aggregate TFP shocks. In both cases, introducing a nonconvex routine labor adjustment cost allows the model to generate a stepwise decline in routine hours (Figure 1.14c) as well as large swings in overall labor hours (Figure 1.14b). To see this, recall that nonconvex disruption costs generate inaction zones in the routine hours decision rule, which then lead to long periods of inaction interrupted with periodic and large downward adjustments that bring routine labor stocks close to their long-run trend levels. Moreover, since nonroutine labor is a complement for routine labor in final goods production and since the adjustment of nonroutine labor is costless in the model, nonroutine hours also fall when routine hours adjust. Overall labor hours, as a result, display infrequent but large swings in the model with long-run trends and labor adjustment costs.

More importantly, aggregate productivity shocks can trigger these trend adjustments, giving rise to concentrated declines of labor hours in recessions (compare the black dashed lines and the blue solid lines in Figure 1.14). As the blue solid lines show, the model with

Figure 1.14
Output and Labor Hours Over Business Cycles (1986-2014): Model with Labor Adjustment Cost



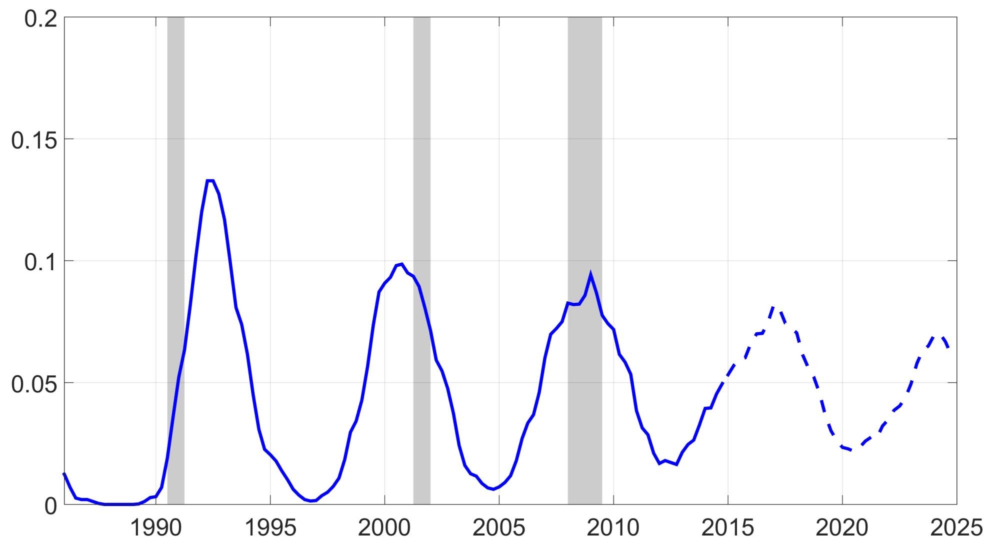
nonconvex adjustment costs and TFP shocks is able to capture the timing of the last three labor market downturns.

Furthermore, when there is a long-run reallocation trend, the sensitivity with which routine labor decisions adjust to productivity shocks changes over time; more specifically, it increases as the gap between the existing stock of labor and the frictionless "target" grows over time during the period of inaction, and declines each time after an adjustment.[16] In other words, given an initial state of the economy, a long-run reallocation trend and nonconvex adjustment costs together lead to periods with high likelihood of labor adjustment and periods with low likelihood of labor adjustment. Figure 1.15 plots the estimated probability of adjustment for the period 1986-2014 and identifies three such cycles in the past 30 years, which have largely coincided with the business cycle.[17] When the propensity to adjust is high, even a small negative TFP shock can trigger a large labor market downturn. In this vein, the model explains why the small drop of output at the beginning of 2000, which was not enough to cause a large recession, nevertheless triggered a large and prolonged decline of labor hours that lasted several years, giving rise to the most severe example of a jobless recovery.

[16]The "gap approach" is common in the labor adjustment cost literature (Caballero and Engel, 1993; Cooper, Haltiwanger, and Willis, 2015). The gap approach studies the dynamics of labor adjustment as a function of the gaps between firms' labor choices under adjustment costs and the counterfactual frictionless targets. The firm adjusts labor when the gap is large and remains in inaction when the gap is small. In a stationary environment, the frictionless target fluctuates around a constant mean, so the sensitivity of the adjustment decisions is constant.

[17]I simulate the model with routine labor disruption costs using 1000 randomly generated TFP series (AR(1)). The simulation starts in 1983 as the stock of routine labor at that time is close to trend after having adjusted during the 1982 recession. Each TFP series generates a stepwise declining path of routine hours, with different timing of adjustments. Figure 1.15 plots the percentage of adjustments occurring in each period.

Figure 1.15
Probability of Adjustment



1.6 The Decline of Labor in the 2000s

A competing explanation for the divergence between routine and nonroutine occupations and the decline of the overall labor share is globalization. Globalization, and offshoring in particular, is potentially associated with the divergence between routine and nonroutine occupations because there is a moderate correlation between the "codifiability" and the "offshorability" of a job. The idea is that routine-intensive tasks are well suited for automation because they can be computerized and are also likely to be suitable for offshoring because they can be performed at a distance without substantial loss of quality. However, an important difference remains, as codifiability is about whether a task can be broken down into a set of pre-fixed procedures and produced cost-effectively by a computer or a machine, while offshorability is about whether physical proximity is essential in performing a sequence of tasks. For example, software engineering is a nonroutine but highly offshorable job, while being an inspector on an assembly line is routine but not offshorable. The correlation between a standard measure of offshorability and RTI is around 0.4 (Goos, Manning, and Salomons, 2014).[18] In addition, offshoring is associated with a declining labor share, as it induces a shift of capital intensity in domestic production: as U.S. firms move their more labor-intensive production stages to China, the remaining domestic production in the U.S. on average becomes more capital-intensive.

Some studies have argued that RBTC and trade play overlapping roles in driving the two discussed labor market structural shifts. These two forces may also be mutually reinforcing.[19] However, a few recent papers have suggested that the effects of RBTC and trade on the decline of labor's share and the divergence between routine and nonroutine occupations have been quite different: RBTC is much more important than offshoring in explaining the divergence between routine and nonroutine occupations, while import competition and offshoring are more important than RBTC in driving the decline of overall employment and labor's share, especially after 2002 (see Autor, Dorn, and Hanson (2015), Goos, Manning, and Salomons (2014), Michaels, Natraj, and Van Reenen (2014), and Elsby, Hobijn, and Şahin (2013) for details).[20]

Building on these insights, I extend the baseline model with RBTC to incorporate a notion of offshoring without engaging in a full-blown trade model. Specifically, I introduce a second form of technological change that augments a third type of capital, which I denote as IM_t. The key assumption is that IM_t is more substitutable with both types of labor than

[18]For details on the construction of the measure for offshorability, see Blinder and Krueger (2013).

[19]RBTC and trade are mutually reinforcing, as the rise of offshoring owes largely to declining cost of transportation and (especially) communication technologies, which makes it difficult to completely separate the two processes. However, there is reason to believe that these interactions are of at most second-order importance relative to their respective effect on the labor market (Acemoglu, Gancia, and Zilibotti, 2015).

[20]Goos, Manning, and Salomons (2014) also show that there is a shift away from routine occupations toward nonroutine occupations within industries, but that RBTC also leads to significant between-industry shifts in the structure of employment in 16 West European countries.

standard non-IT capital in the aggregate production function; therefore I call this technology Labor-Saving Technological Change (LSTC). [21] The augmented production function is given by:

$$Y_t = z_t K_t^{\alpha} \left[\lambda_{im} \left(X_{im,t} IM_t \right)^{\gamma} + (1-\lambda_{im}) \left\{ \lambda_n L_{n,t}^{\sigma} + (1-\lambda_n) \left[\lambda_m (X_{m,t} M_t)^{\epsilon} + (1-\lambda_m) L_{r,t}^{\epsilon} \right]^{\frac{\sigma}{\epsilon}} \right\}^{\frac{\gamma}{\sigma}} \right]^{\frac{1-\alpha}{\gamma}} \tag{1.46}$$

in which $X_{im,t}$ represents LSTC. The nested CES structure of the aggregate production function (1.46) has several implications. First, it implies that IM capital is more substitutable with labor ($0 < \gamma < 1$) than with non-IT capital, K_t, which is key to the role of LSTC in depressing labor's share. Second, since IM_t can substitute for the entire task bundle produced by IT capital and labor, LSTC does not induce divergence between routine and nonroutine labor, as both types are equally substitutable by IM_t. Last but not least, LSTC decreases the share of IT capital in total income.

It is useful to denote the task bundle produced by IT capital and the two types of labor as $Y_{ML,t}$:

$$Y_{ML,t} = \left\{ \lambda_n L_{n,t}^{\sigma} + (1-\lambda_n) \left[\lambda_m (X_{m,t} M_t)^{\epsilon} + (1-\lambda_m) L_{r,t}^{\epsilon} \right]^{\frac{\sigma}{\epsilon}} \right\}^{\frac{1}{\sigma}} \tag{1.47}$$

and I define $\Theta_{im,t}$ as follows:

$$\Theta_{im,t} = \frac{\lambda_{im} \left(X_{im,t} IM_t \right)^{\gamma}}{\lambda_{im} \left(X_{im,t} IM_t \right)^{\gamma} + (1-\lambda_{im}) Y_{ML,t}^{\gamma}} \tag{1.48}$$

The income share of IM capital is given by: $s_{IM,t} = (1-\alpha)\Theta_{im,t}$, and the income shares

[21] In the context of offshoring, IM_t can be thought of as imported intermediate goods. Another possible empirical counterpart is intangible assets.

of IT capital and the two types of labor are all scaled by $(1 - \Theta_{im,t})$.

I assume that the law of motion for IM capital is given by:

$$IM_{t+1} = g(I_{im,t}) = I_{im,t}^{\phi} \tag{1.49}$$

where $I_{im,t}$ stands for investment in IM_{t+1}. Parameter $\phi < 1$ characterizes the production function of new IM capital with diminishing marginal returns. Also, I assume that the existing stock of IM capital depreciates fully after production, which is a harmless simplifying assumption.[22] The resource constraint in the economy is given by:

$$Y_t = C_t + I_{k,t} + I_{m,t} + I_{im,t} \tag{1.50}$$

where investment $I_{k,t}$ and $I_{m,t}$ are defined in the same way as in the main model. Everything else not described here is the same as in the main model.

Analytical Results

Here I conduct comparative static analysis for the model with both LSTC and RBTC. As in Section 1.3.3, the following propositions are derived from a simplified model without non-IT capital, as non-IT capital accounts for a constant share, α, in the total income and therefore does not affect the properties of the structural changes induced by exogenous technological changes.

[22] The specification of equation (1.49) allows me to relate the model to offshoring more easily, as IM_{t+1} can be thought of as imported intermediate goods being produced by combining offshoring investment $I_{im,t}$ and foreign labor (not explicitly accounted for in this model) with a production technology like (1.49). However, a standard law of motion for capital, such as $IM_{t+1} = I_{im,t} + (1 - \delta_{im})IM_t$, delivers essentially the same results in the model.

Proposition 1.6.1. *In the competitive equilibrium,*

$$\frac{d\ln(\eta_t)}{d\ln(X_{im,t})} = 0 \quad (1.51)$$

$$\frac{d\ln(s_{IM,t})}{d\ln(X_{im,t})} = \frac{d\ln(\Theta_{im,t})}{d\ln(X_{im,t})} = \gamma(1-\Theta_{im,t}) > 0 \quad (1.52)$$

$$\frac{d\ln(s_{L,t})}{d\ln(X_{im,t})} = \frac{d\ln(s_{M,t})}{d\ln(X_{im,t})} = \frac{d\ln(s_{L_n,t})}{d\ln(X_{im,t})} = \frac{d\ln(s_{L_r,t})}{d\ln(X_{im,t})} = -\gamma\cdot\Theta_{im,t} < 0 \quad (1.53)$$

for $0 < \epsilon < 1, \sigma < 0$, *and* $0 < \gamma < 1$.

Proof. See online Appendix. □

Proposition 1.6.1 states that LSTC does not induce labor to flow between routine and nonroutine sectors; it increases the share of the IM capital in total income, and decreases the shares of labor – routine, nonroutine, and overall – and IT capital in total income. This result qualitatively establishes that LSTC can generate the additional decline of the overall labor income share and the slowdown of the nonroutine labor income growth observed in the data after 2002.

Proposition 1.6.2. *In the competitive equilibrium,*

$$\frac{d\ln(s_{L,t})}{d\ln(X_{M,t})} < 0, \quad \textit{when} \quad \frac{\Theta_{m,t}\Theta_{n,t}}{1-\Theta_{m,t}} < \frac{(\epsilon-\gamma\Theta_{IM,t})}{(\gamma\Theta_{IM,t}-\sigma)}\cdot\frac{(1-\sigma)}{(1-\epsilon)} \quad (1.54)$$

$$\frac{d\ln(s_{L,t})}{d\ln(X_{M,t})} > 0, \quad \textit{when} \quad \frac{\Theta_{m,t}\Theta_{n,t}}{1-\Theta_{m,t}} > \frac{(\epsilon-\gamma\Theta_{IM,t})}{(\gamma\Theta_{IM,t}-\sigma)}\cdot\frac{(1-\sigma)}{(1-\epsilon)} \quad (1.55)$$

for $0 < \epsilon < 1, \sigma < 0$, *and* $0 < \gamma < 1$.

Proof. See online Appendix. □

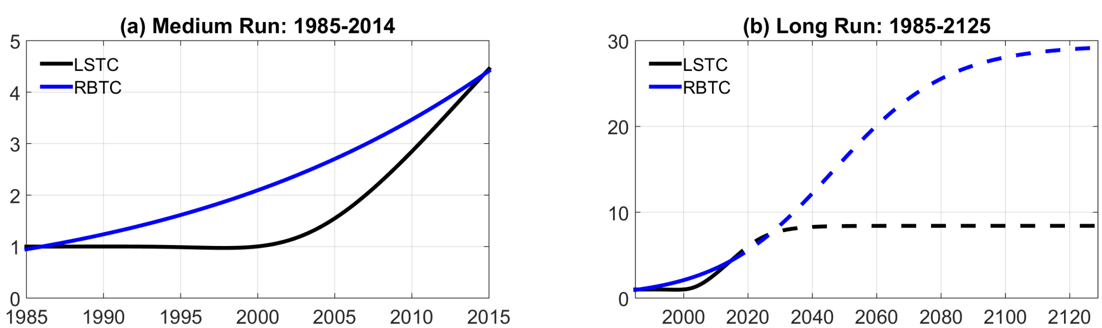
Proposition 1.6.2 asserts that even when IM capital is present, RBTC continues to induce a U-shaped path for the overall labor income share, as long as the initial IM capital share is relatively low $(s_{IM,0} < \epsilon/\gamma)$. This mild condition stems from the dual effects of RBTC in this model: within the task bundle $Y_{ML,t}$ produced by IT capital and labor, RBTC induces divergence between routine and nonroutine labor and raises the income shares of nonroutine labor and IT capital relative to routine labor. Because $Y_{ML,t}$ is a substitute for IM capital, RBTC also raises the overall income share of $Y_{ML,t}$, which in turn raises the income shares of IT capital and both types of labor, relative to IM capital. For nonroutine labor, the two effects work in the same direction and RBTC unambiguously increases the share of nonroutine labor in total income. However, for routine labor, the first effect of RBTC is negative while the second effect is positive; as a result, RBTC has an ambiguous effect on the share of routine labor in total income when IM capital is present. Only when the second effect is relatively small, e.g., when the IM capital share is low, does RBTC depress the routine labor income share and consequently induce a U-shaped pattern for the overall labor share. This result implies that countries with different IM capital shares may see different effects of IT growth/RBTC on the labor income share.

Propositions 1.6.1 and 1.6.2 establish that LSTC and RBTC have different implications for long-run movements of factor shares. While RBTC induces a U-shaped path of labor's share, LSTC generates a path that is monotonically declining. The ultimate path of $s_{L,t}$ therefore depends on the relative sizes of these two exogenous forces. Under an empirically reasonable assumption that LSTC will abate much sooner than RBTC, the overall labor share will follow a U-shaped path with a delayed turnaround.

Numerical Experiment

Propositions 1.6.1 and 1.6.2 have established analytically that the model with LSTC and RBTC can generate factor share movements that are qualitatively consistent with the data. In this section, I assess the model's quantitative performance. I find that when calibrated using trade data, the model delivers very little additional decline of the overall labor share or slowdown of the nonroutine labor share. This provides tentative evidence suggesting that trade alone cannot quantitatively explain the accelerated decline of labor after 2002, which is consistent with the findings by Eden and Gaggl (2014). However, more rich open-economy models with amplification mechanisms for LSTC may have different quantitative implications.The closed-economy setup is subject to a second caveat. In the closed-economy model, the income of IM_t is treated as part of total income. In reality, the value added from imported intermediates is not part of domestic income, which means that the decline of the shares of labor and IT capital under LSTC within the domestic economy is further dampened in the model by national accounting. I leave possible interpretations and extensions to future research.

Figure 1.16
Calibrated LSTC Process



Here I proceed as follows. To illustrate the working of the model, in the following numerical exercise, I simply pick the part of the parameter space that allows me to obtain a good match for the data, and I choose reasonable parameter values whenever possible. In particular, I set the parameter $\gamma = 0.75$, which implies an elasticity of substitution of 4 between IM capital and the task bundle produced by IT capital and labor, which is in the range of estimated import elasticities of substitution. I choose $\phi = 0.6$, which is roughly equal to the capital income share in China. I choose $\lambda_{im} = 0.133$ to set the initial share of IM to be 0.03. I then feed the model with the exogenous RBTC process used in the main model and the LSTC process depicted in Figure 1.16. The range of the LSTC growth is approximately consistent with the import penetration ratio of China, but with a much sharper kink at 2002. In fact, in this numerical exercise, I assume that RBTC starts at the beginning of the simulation, while LSTC does not start until 2002, which is around the time that China joined the World Trade Organization (WTO).[23] This assumption roughly corresponds to the timing of important events: computerization started in the late 1970s and was a major phenomenon in the 1980s-90s, while the rise of globalization and offshoring did not become a significant event until the 2000s. The remaining parameter values are calibrated using the same targets as in the benchmark model.

Figure 1.17 shows that under this parameterization, the model with RBTC and LSTC can match the entire path of the overall labor share, as well as the divergence between routine and nonroutine income shares. In particular, as the decomposition shows, RBTC drives the divergence between routine and nonroutine labor shares and hours, while LSTC generates equal percentage declines in both types of labor inputs and income shares. This simultaneous dampening on routine and nonroutine labor leads to a decline of overall labor

[23] China joined WTO in December 2001.

Figure 1.17
Medium-Run Factor Income Shares: RBTC + LSTC

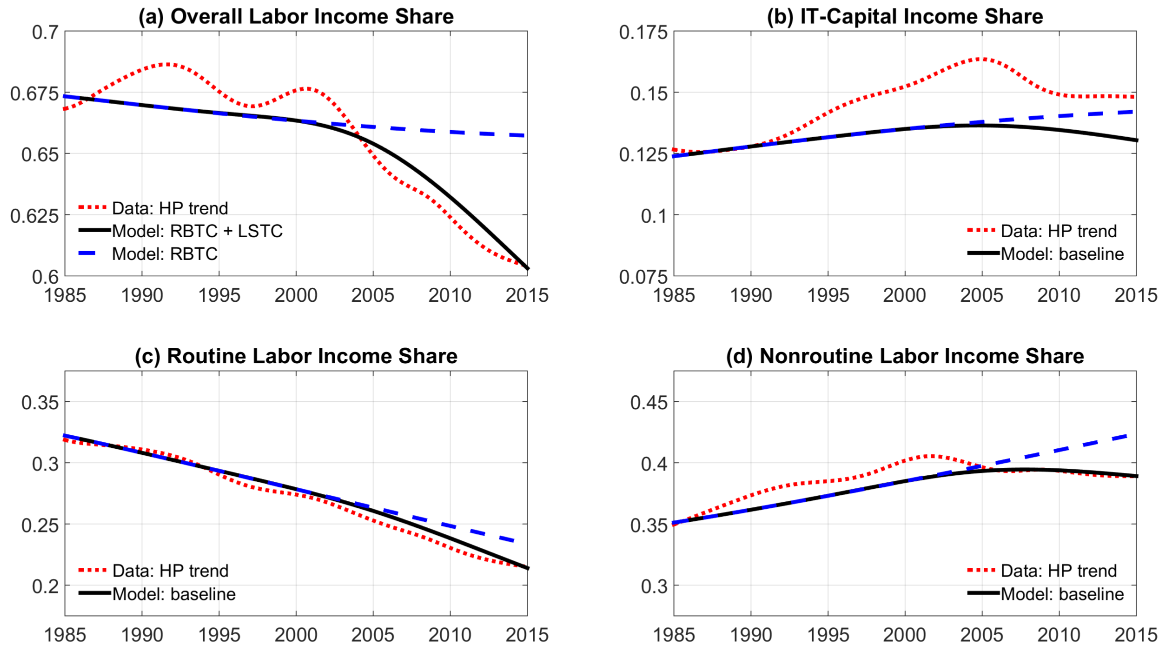


Figure 1.18
Long-Run Factor Income Shares: RBTC + LSTC

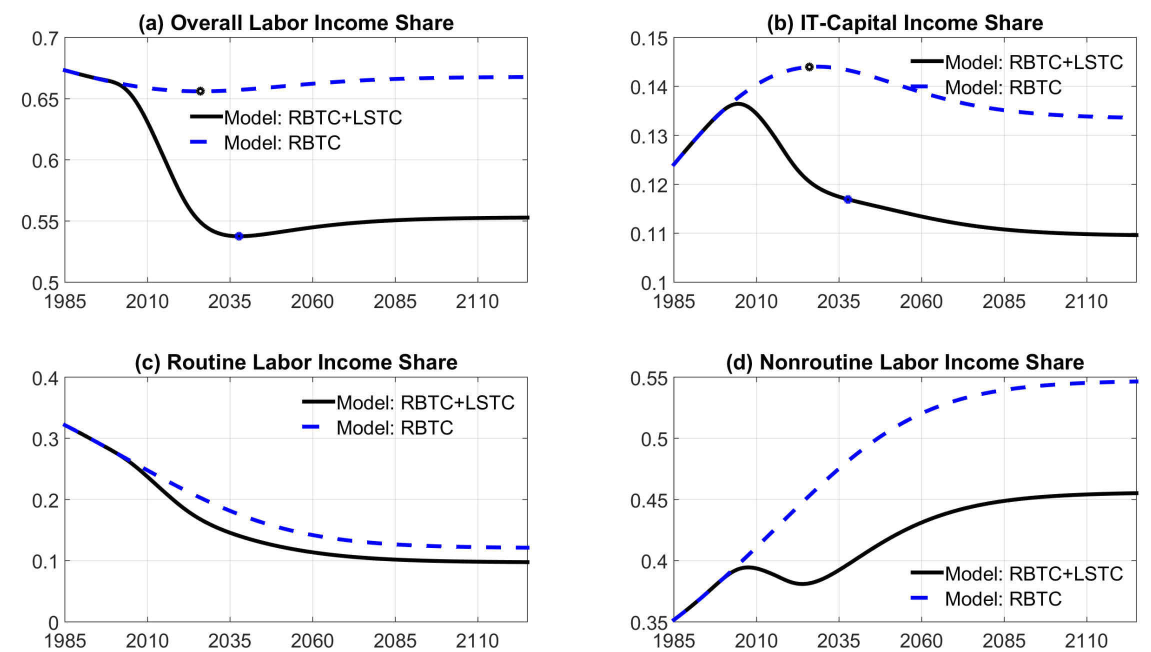
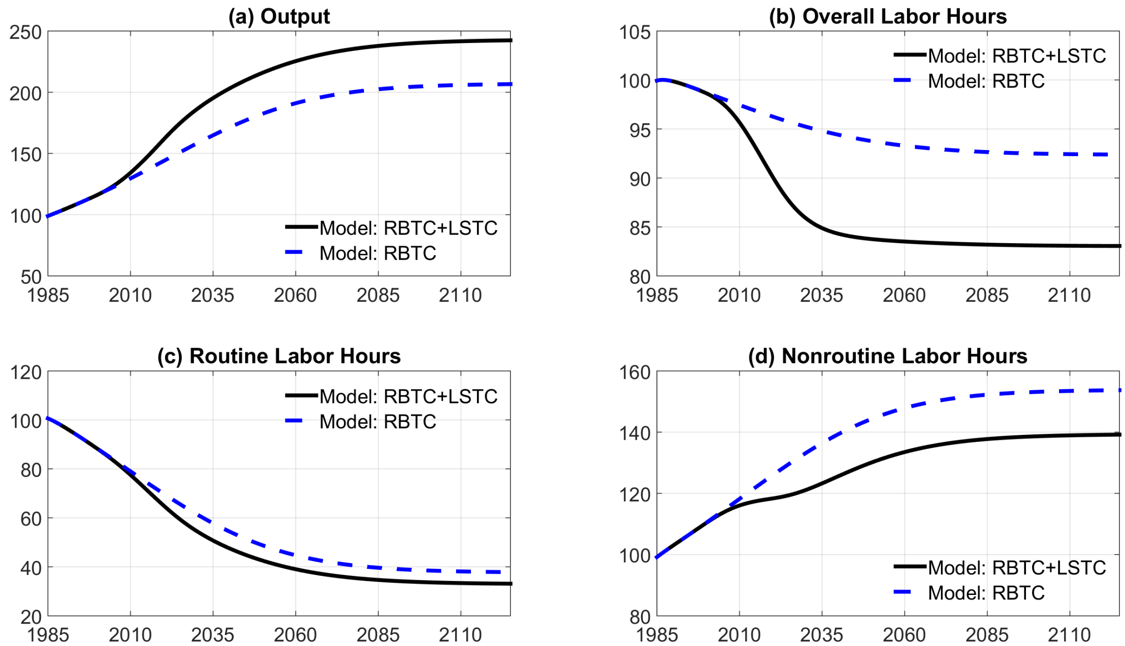


Figure 1.19
Long-Run Output and Labor Hours: RBTC + LSTC



input and the labor income share. Moreover, the model with LSTC generates a decline in the income share of IT capital, which is consistent with the data after 2002.

Figure 1.18 shows that in the long run, the overall labor share will decline much further with prolonged exogenous LSTC than without. Under the specific LSTC process depicted in Figure 1.16, the labor share does turn around before stabilizing at a permanently lower level in the long run. but the timing is delayed. Moreover, the decline of the overall labor share is driven primarily by a decline and then slower growth of the nonroutine labor income share. At the same time, LSTC sharply depresses the income share of IT capital as IM-capital crowds out IT investments. Finally, Figure 1.19 shows that LSTC contributes to output growth while dampening labor inputs on all fronts.

1.7 Conclusion

In this study, I study the impact of IT progress, i.e., Routine-Biased Technological Change (RBTC), on the US labor market. I incorporate unbalanced growth in an otherwise standard neoclassical growth model and study both analytically and quantitatively the role of IT growth/RBTC in driving two well-documented labor market trends: the divergence between routine and non-routine workers in terms of employment and income shares, and the decline of the overall labor income share relative to capital. The first contribution of the paper is to show that RBTC induces a U-shaped growth path for the overall labor share of income, as well as a declining elasticity of substitution between capital and overall labor. On the empirical front, this novel feature of RBTC helps account for the recent decline of labor share in the data, while making an interesting forecast about the future trajectory of the aggregate labor share. On the theoretical front, this finding contributes to the ongoing debate on capital-labor substitutability by providing a dynamic perspective.

Quantitatively, I find that on the aggregate level, RBTC can account for nearly all the divergence between routine and non-routine labor in terms of employment and income shares in the U.S. in the period 1986-2014. RBTC can also account for the mild decline of the overall labor share in the period 1986-2002, but cannot explain the acceleration of the decline after the 2000s. Circumstantial evidence suggests that globalization may have played a much larger role than technology in the 2000s in causing the sharp fall of overall labor share. However, my preliminary analysis within a closed-economy, representative-agent framework is inconclusive on the matter.

Last but not least, when short-run frictions are present and labor adjustments are not instant, there are interesting interactions between secular trends and business cycles. I explore this property of RBTC in the representative-agent model and find that, in the presence of nonconvex routine labor adjustment costs, the model generates a stepwise decline in routine labor hours and large swings in overall labor hours, which is qualitatively consistent with the data. More interestingly, the timing of these inevitable trend adjustments can be significantly affected by aggregate productivity shocks and therefore concentrated in recessions. In future work, I plan to embed RBTC and labor adjustment costs in a heterogeneous firm model to study further the links between the secular trend and business cycles.

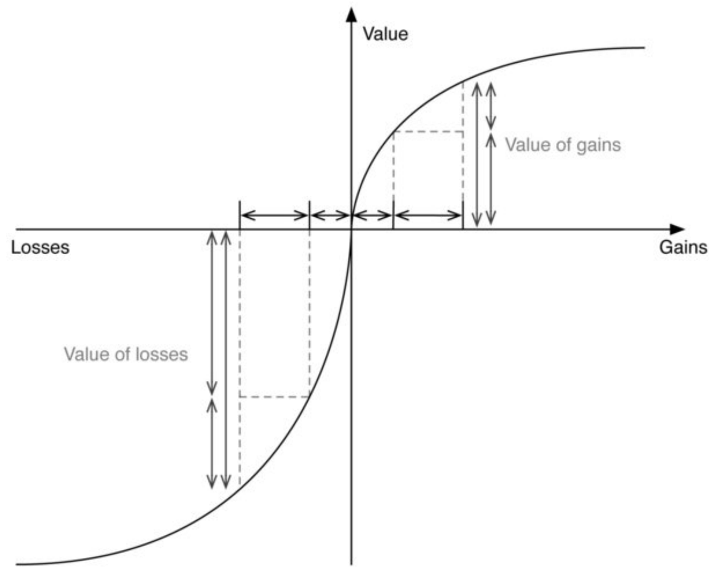
Chapter 2: Loss Aversion and Business Cycle Asymmetries

2.1 Introduction

Economic Research suggests that business cycles are asymmetric (Acemoglu and Scott, 1997; Van Nieuwerburgh and Veldkamp, 2006; among others): contractions are found to be sharper and more violent than expansions (McKay and Reis, 2008); economic activity can stagnate in recessions for a prolonged period of time (Chamley and Gale, 1994); and negative shocks generate more persistent dynamics in aggregate variables than positive shocks (Aruoba, Bocola, and Schorfheide, 2012). However, the standard Dynamic Stochastic General Equilibrium (DSGE) model cannot account for these empirical findings on asymmetries. The goal of this study is to explore the role of loss aversion in generating asymmetric dynamics in consumption and labor hours over business cycles.

The notion of loss aversion was first introduced to economics by Kahneman and Tversky (1979) as part of their famous prospect theory. A descriptive model of decision making under uncertainty, prospect theory states that people compare their well-being to a reference point ("reference-dependence") and derive utility (disutility) from the gain (loss). Loss aversion refers to the fact that the disutility from a loss is greater than the utility from a gain of the same amount. In addition, prospect theory asserts that people are risk-loving in the realm of loss and risk-averse in the realm of gain, a phenomenon known as the "reflection effect". Prospect theory can be summarized in an S-shaped, piecewise utility function centered at the reference point with slopes in the local region of loss much steeper than those in the region of gain (Figure 2.1). This utility function is often called "gain-loss utility" in the literature. It contrasts with the traditional consumption utility function, which assigns values to absolute levels of wealth and consumption, does not distinguish between loss and gain, and is everywhere concave.[1]

Figure 2.1
Prospect Theory: Gain-Loss Utility Function



Besides the vast evidence from studies in cognitive and behavioral psychology, the presence of loss aversion in decision making has been confirmed repeatedly in empirical studies in economics. In particular, Camerer (2003) lists ten field phenomena that are inconsistent with Expected Utility (EU) theory and consistent with cumulative prospect theory (i.e. loss aversion and the reflection effect). Shea finds evidence in household (1995a) and aggregate (1995b) data that consumption rises today when expected future income rises but does

[1]The traditional consumption utility can be viewed as taking level zero as a fixed reference point.

not fall today when expected future income falls, suggesting the presence of loss aversion in household and aggregate consumption choices. More recently, Rosenblatt-Wisch (2008) estimates the inter-temporal Euler equation under prospect theory using Generalized Method of Moments (GMM) and finds evidence of loss aversion in aggregate consumption data. Tovar (2009) studies the effects of loss aversion on trade policy, and show that loss aversion can explain why a disproportionate share of trade protection goes to declining industries.

On the theory front, loss aversion has been used to explain various asset pricing anomalies, including the equity premium puzzle and the risk-free rate puzzle. Benartzi and Thaler (1995) build a partial equilibrium model and show that if investors are loss averse and monitor their portfolios frequently (”myopia”), a high equity premium is then required to compensate them for the volatility in stock returns. Barberis, Huang, and Santos (2001) study loss averse investors in a general equilibrium endowment economy. Their model is successful in generating the high mean, excess volatility and predictability of stock returns, as well as the correlation of stock returns with consumption growth observed in the data.[2]

Intuitively speaking, loss aversion has two key implications. The first is that loss averse agents are highly risk averse around the reference point.[3] Under uncertainty, incomplete markets and a frequently updated reference point, a loss averse agent also exhibits strong motives for precautionary saving, as she will optimally choose to build a buffer stock to avoid having consumption fall below the reference level. In other words, loss aversion im-

[2] Barberis, Huang, and Santos (2001) assume time-varying loss aversion. That is, previous performance affects the degree of loss aversion in the current period. This assumption is built on the behavioral evidence of Thaler and Johnson (1990) and is necessary in generating their results.

[3] By assuming time-varying loss aversion, Barberis, Huang, and Santos (2001) extend the prospect theory and have loss averse agents become less risk averse after good investment performance. But the “baseline” loss-aversion still implies monotonically higher risk aversion.

plies a high curvature in the overall utility function. The second key implication of loss aversion is asymmetry: loss averse consumers have asymmetric responses to positive and negative income shocks.[4] Bowman, Minehart, and Rabin (1999) prove these two points in a two-period consumption-saving model. They show that when there is sufficient income uncertainty, a loss-averse person resists lowering consumption in response to bad news about future income, and the resistance is greater than the resistance to increasing consumption in response to good news. These predictions are at odds with the permanent income hypothesis, but are consistent with the empirical evidence on asymmetric consumption behavior found by Shea (1995a,b).

Despite its success in explaining asset pricing anomalies and Bowman, Minehart, and Rabin's early attempt to incorporate it into the consumer choice problem, the application of loss aversion in macroeconomic theory overall is limited. One notable exception is a recent paper by Santoro, Petrella, Pfajfar, and Gaffeo (2014) in which the authors show that loss aversion can be used to explain the asymmetric transmission mechanism of monetary policy over business cycles. Nevertheless, there has not been a systematic treatment of loss aversion in the standard infinite-horizon Real Business Cycle (RBC) framework. In this study, I fill this gap by asking and answering the following questions: Can loss aversion be reconciled with business cycle facts? What are its qualitative and quantitative implications in an otherwise standard RBC framework? Can the fundamental asymmetry built into loss

[4] The other feature of prospect theory, "the reflection effect", implies something distinctively different from loss aversion. In particular, it implies diminishing sensitivity in both gain and loss, and this property mainly affects the course of local (inter-temporal in our context) adjustments. For example, if one uses lagged consumption as the reference point, risk-seeking in the realm of loss means that the agent is better off concentrating a doomed loss in one period, rather than spreading it out over multiple periods; this is in contrast with the region of gain, in which agents are risk averse and prefer to smooth adjustments to take advantage of the high marginal utilities period by period, a property well-understood in the habit formation literature.

aversion help explain (some of) the asymmetries of business cycles found in the data?

To answer these questions, I embed loss aversion into an otherwise standard RBC model. I find that how loss aversion affects business cycle dynamics depends critically on the nature of the reference point. If lagged consumption (i.e. status quo) is used as the reference point, loss aversion dramatically lowers the effective inter-temporal rate of substitution and causes extremely high local risk aversion around the reference point, which is in turn determined in equilibrium. In an otherwise frictionless RBC environment with labor and production, consumers adjust to productivity shocks through the labor-leisure channel and smooth consumption to the extreme. In other words, I find that loss aversion under status-quo-reference (or any reference that is closely correlated with the lagged choice itself) implies excessive consumption smoothing and hence cannot be reconciled with business cycle facts in a simple RBC framework.[5]

In contrast, if the reference point is fixed at a constant level, loss aversion induces a flat region in consumption decision rules and stagnation at the reference level in consumption paths. When the reference point is close to the deterministic steady state consumption, the presence of loss aversion will discourage people from lowering consumption in response to negative shocks, but will not discourage them from raising consumption in response to positive shocks. As a result, consumption will not exhibit excessive smoothness under a fixed reference point. I show that loss aversion in consumption with a fixed reference point improves the simple RBC model in terms of reducing the correlation between output and consumption, and between output and hours. Under this framework, I am able to generate asymmetric impulse responses of consumption to positive and negative shocks.

[5] A similar result holds for habit formation. See Lettau and Uhlig (2000).

I next extend the model to incorporate loss aversion in leisure with a fixed reference point, and investigate the resulting impact on business cycle dynamics. I find that when loss aversion is imposed jointly on the consumption-leisure composite good, it does not affect the intra-temporal margin, and both consumption and labor hours adjust to keep the utility above its reference point. Therefore, loss aversion hardly affects business cycle dynamics and does not induce asymmetric impulse responses of consumption or hours. On the other hand, when loss aversion is imposed on consumption and leisure separately, i.e. double loss aversion, it induces a flat region in the decision rules of both consumption and leisure. With loss aversion in leisure, hours become much less volatile and consumption is slightly more volatile than without loss aversion in leisure. In addition, both consumption and hours exhibit asymmetric impulse responses to positive and negative productivity shocks.

The rest of the chapter is organized as follows. In section 2.2, I formalize the notion of loss aversion and describe the properties of the gain-loss utility. In section 2.3, I study loss aversion in consumption in a simple RBC model and solve the model using two different types of reference points. In section 2.4, I extend the model to incorporate loss aversion in leisure in two ways and repeat the analysis. In section 2.5, I discuss the literature on asymmetric business cycles and how loss aversion can potentially explain business cycle asymmetries. In section 2.6, I offer some concluding remarks.

2.2 Loss Aversion

In this section, I formalize the notion of loss aversion and describe the properties of the gain-loss utility function. For simplicity, I restrict attention to loss aversion in consumption only. It should be obvious that a similar formulation of loss aversion can apply to any choice that affects utility directly, including leisure, which I will discuss later.

To incorporate loss aversion into the traditional consumption-based model, I first assume that the total utility of a loss averse consumer is a weighed sum of two components:

$$U(c,x) = (1-\theta)u(c) + \theta v\left[u(c) - u(x)\right] \tag{2.1}$$

Here $u(.)$ is a standard utility function (i.e. strictly increasing and concave) of consumption c, and $v(.)$ is the gain-loss utility whose properties will be specified below. x is some reference consumption level. Parameter $\theta \in [0,1]$ controls the relative importance of the gain-loss utility, and when $\theta = 0$, we are back to the standard consumption utility case. Following Kőszegi and Rabin (2006)[6], I assume that the gain-loss utility depends on the difference in the relative consumption utility, $u(c) - u(x)$. This assumption provides a normalization to the second term in the utility function to keep it comparable to the first term. An alternative setup is to assume that the gain-loss utility is based directly on the consumption gap: $v(c-x)$. Since preference is ordinal, as long as $v(.)$ and $u(.)$ are strictly increasing the two specifications are equivalent, although the exact specification does affect the interpretation of the parameter θ.

2.2.1 Piecewise-linear gain-loss utility function

Bowman, Minehart, and Rabin (1999) provides a theoretical framework for studying the properties of the gain-loss utility. In particular, they show that a gain-loss utility function $v(.)$ should satisfy the following assumptions:

[6] Kőszegi and Rabin (2006) develop a general class of reference-dependent models which encompasses neoclassical consumption utility with the gain-loss utility of Kahneman and Tversky (1979).

1. *Monotonicity*: $v(z)$ is continuous and strictly increasing for all $z \in \mathcal{R}$, where $v(0) = 0$.

2. *Diminishing sensitivity*: $v(z)$ is twice differentiable for all $z \neq 0$, $v''(z) \leq 0$ for all $z > 0$, and $v''(z) \geq 0$ for all $z < 0$.

3. *Loss aversion*: $v(y)+v(-y) < v(z)+v(-z)$ for all $y > z > 0$, and $\lim_{z \to 0} v'(-z)/v'(z) = \lambda > 1$.

Tversky and Kahneman's cumulative prospect theory (1992) suggests that the gain-loss utility takes the form of a power function:

$$v(z) = \begin{cases} z^{\alpha} & \text{if } z \geq 0 \\ -\lambda(-z)^{\beta} & \text{if } z < 0 \end{cases}$$

where z denotes gain or loss, λ is the loss aversion coefficient, and $\alpha \in (0,1)$ and $\beta \in (0,1)$ control the degree of the reflection effect (i.e. concavity in the region of gain and convexity in the region of loss). Their experimental results suggest that $\alpha = \beta = 0.88$. As theses estimates are not too far away from unity, many applications of prospect theory choose to abstract from curvature to keep the analysis simple (Barberis, Huang, and Santos, 2001). Here I follow suit and make the same assumption. That is, I assume a piecewise-linear function with a kink at the origin:

$$v(z) = \begin{cases} z & \text{if } z \geq 0 \\ -\lambda(-z) & \text{if } z < 0 \end{cases} \tag{2.2}$$

where $\lambda > 1$ is the loss aversion coefficient. In controlled experiments, λ has been estimated to be within the neighborhood of $[2, 2.5]$, and Tversky and Kahneman (1992) suggest a point estimate of $\lambda = 2.25$. Tovar (2009) uses data on trade policy and estimates the loss aversion parameter to be 2, consistent the experimental data.[7]

Under the piecewise linearity assumption, the *marginal* gain-loss utility $v'(.)$ has a very simple binary form:

$$v'(z) = \begin{cases} 1 & \text{if } z > 0 \\ \lambda & \text{if } z < 0 \end{cases} \tag{2.3}$$

which embeds asymmetry. But it is important to note that marginal utility is not defined at $z = 0$, as it jumps from 1 to λ when z becomes negative. This discontinuity is the key to many of the most important implications of loss aversion, as well as being the main source of analytical difficulties involved in solving the model. Nevertheless, assuming piecewise-linearity greatly simplifies the algebra, as this formulation isolates the implications of loss aversion from those of the reflection effect. Moreover, since such a gain-loss utility function is globally concave, it guarantees an interior solution to the optimization problem.

If one further assumes that the reference point, x, is exogenously given, then the marginal utility of consumption implied by the composite utility function in equation 2.1 has a simple representation:

$$\begin{aligned} \frac{\partial U(c,x)}{\partial c} &= [(1-\theta)u'(c) + \theta v'(\Delta u|x)u'(c)] \\ &= \begin{cases} u'(c) & \text{if } c > x \\ [1+\theta(\lambda-1)]\, u'(c) & \text{if } c < x \end{cases} \end{aligned}$$

[7] There is some evidence that the parameter is slightly lower than 2 in aggregate data (Rosenblatt-Wisch, 2008).

where $\Delta u|x = u(c) - u(x)$. Once again, the marginal utility is not defined at $c = x$. This clean representation illustrates the implications of loss aversion: when consumption is well above the reference level, i.e. when one is in the region of gain, the presence of loss aversion does not affect local decisions as the marginal utilities are the same as in the standard case, regardless of the degree of loss aversion; loss aversion only affects marginal utility and decisions when consumption is below the reference level, i.e. when one is in the region of loss. Loss aversion has the strongest influence on local decisions when one is close to the reference point. In particular, at the reference point, the marginal utility jumps from $u'(c)$ to $(1 - \theta + \theta\lambda)u'(c)$, creating potential discontinuities or flat regions in consumption decision rules, as I show below.

2.2.2 Reference points

The nature of the reference point turns out to have a critical effect on how loss aversion affects the dynamics in a general equilibrium model. Unfortunately, the original prospect theory provides little guidance on the choice of reference point, and research on the nature of reference points themselves is limited. Therefore, in this study, I experiment with different specifications of the reference point and explore their distinct implications.

The simplest reference point to use in terms of modeling and computation is a fixed non-zero reference point. Under this specification, the reference point enters the model as a parameter and the value of this parameter matters critically for the direction of asymmetries generated by loss aversion.

The most common assumption on the reference point in existing literature is the status quo. In the standard consumption-saving problem in macroeconomics, the status quo is the previous-period consumption, which is then updated period-by-period. As in the habit formation literature, one can choose to use an internal reference (e.g. lagged own consumption) or an external reference (e.g. lagged aggregate consumption), or a mixture of both. Using lagged consumption in our model has the advantage of being consistent with the habit formation literature, but it adds a state variable and creates some computational issues in our context.[8] I will discuss these issues and solve the model under this type of reference point later in Section 2.3.4.

Before we move on to the next section, note that one can also use lagged *expected* consumption as a reference point (Barberis, Huang, and Santos, 2001; Kőszegi and Rabin, 2006). That is, the current choice poses a gain if it exceeds the previous expectations, and a loss if it falls short. This specification would potentially bring my model closer to the literature on "disappointment aversion" (Gul, 1991), in which the reference point is usually assumed to be the certainty equivalent of future utility. Despite the obvious interest in using this type of reference point, it poses significant computational challenge in a macro DSGE framework. Therefore I leave it for future work.

2.3 Loss Aversion in Consumption

In this section I embed loss aversion in consumption choices into an otherwise standard RBC model and explore its business cycle implications. I use the composite total utility function and piecewise-linear gain-loss utility as discussed in the previous section.

[8]In the finance literature on habit formation, it is commonplace to assume a slow-moving reference point, following Campbell and Cochrane (1999). That is, habits are assumed to be a geometric sum of past consumptions, which introduces persistence in the reference point. However, in the macro-theoretical application of habit formation, in particular, under the "standard" DSGE framework (Christiano, Eichenbaum, and Evans, 2005; Smets and Wouters, 2007), the common assumption on habits is to use a fraction of the lagged consumption, which is updated period-by-period.

2.3.1 The model

Households

A loss-averse representative consumer chooses consumption and labor in each period to maximize her expected lifetime utility, subject to the flow budget constraint:

$$\max_{\{c_t,n_t,k_{t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t \left\{(1-\theta)u(c_t) + \theta v\left[u(c_t) - u(x_t)\right] + \chi h(1-n_t)\right\} \tag{2.4}$$

$$s.t \qquad c_t + k_{t+1} - (1-\delta)k_t = w_t n_t + r_t k_t, \quad \forall t \tag{2.5}$$

Here I assume that the consumer's total utility is comprised of three components: the utility from flow consumption, $u(c_t)$; a gain-loss utility, $v\left[u(c_t) - u(x_t)\right]$, with respect to the difference between the utilities derived from the consumption choice and a reference consumption point, x_t, which will be specified later; and lastly, the utility from leisure, $h(1-n_t)$, which is assumed to be increasing and concave in $(1-n_t)$. The parameter θ governs the weight of the gain-loss utility relative to the pure consumption utility. The rest of the notation is standard: k_t is the capital stock at the beginning of period t, and n_t is the labor hours choice; w_t and r_t denote the marginal returns to labor and capital, respectively, and parameter δ is the capital depreciation rate.

If one assumes that the reference consumption x_t is external, as I do in this study, then the inter-temporal Euler equation and the intra-temporal optimality condition for the consumer's problem are given respectively by:

$$[1-\theta+\theta v'(\Delta u_t|x_t)]\,u'(c_t)=\beta E_t\left\{[1-\theta+\theta v'(\Delta u_{t+1}|x_{t+1})]\,u'(c_{t+1})\,(r_{t+1}+1-\delta)\right\} \tag{2.6}$$

$$w_t=\chi\cdot\frac{h'(1-n_t)}{[1-\theta+\theta v'(\Delta u_t|x_t)]\,u'(c_t)} \tag{2.7}$$

where $\Delta u_t|x_t=u(c_t)-u(x_t)$. Once again, $v'(.)$ is not defined when $u(c_t)=u(x_t)$.

Production

A representative firm operates under Cobb-Douglas production function $f(k_t,n_t;z_t)=z_tk_t^{\alpha}n_t^{1-\alpha}$ and maximizes profits. In equilibrium its optimality conditions pin down prices in factor markets:

$$r_t = z_tf_k(k_t,n_t) \tag{2.8}$$

$$w_t = z_tf_n(k_t,n_t) \tag{2.9}$$

Exogenous TFP shocks

I assume that aggregate TFP follows an $AR(1)$ process:

$$\log(z_t)=\rho_z\log(z_{t-1})+\epsilon_t,\qquad \epsilon_t\sim N(0,\sigma_z^2) \tag{2.10}$$

Finally, the resource constraint of the economy is given by:

$$c_t+k_{t+1}-(1-\delta)k_t=z_tf(k_t,n_t) \tag{2.11}$$

2.3.2 Calibration

Since the main goal of this study is to investigate the properties and implications of loss aversion, I keep the rest of the model entirely standard and use the no-loss-aversion case ($\theta = 0$) as a benchmark when choosing values for the conventional RBC parameters. I then add loss aversion to the model by increasing the value of θ.

Specifically, I first consider the following functional forms for the consumption and leisure utility:

$$u(c_t) = \log(c_t) \tag{2.12}$$

$$h(1-n_t) = \log(1-n_t) \tag{2.13}$$

Since I assume separable utility in consumption and leisure throughout this study, using log utility ensures that the baseline model satisfies the long-run regularity conditions on preferences in the RBC literature (King and Rebelo, 1999).

I assume that the time period in the model is one quarter. There are eight parameters in the model to be calibrated, among which six are conventional preference and technology parameters. These include the discount factor β, capital's share in production α, the capital depreciation rate δ, the weight of the utility from leisure in total utility, χ, and the persistence and volatility levels of the aggregate TFP shock, ρ_z and σ_z. For these parameters I pick values that match my baseline model to the long-run moments in the post-war US data. The specific calibration targets used for each parameter and the parameter values are listed in Table 1. As mentioned before, the practice thus far is entirely standard.

Table 2.1
Calibration of RBC model with Loss Aversion

Parameters	**Values**	**Calibration Targets or Sources**
β: Discount Factor	0.99	Annual real interest rate = 4%
α: Capital Share in Production	0.33	Average labor share of income = 0.67
δ: Capital Depreciation rate	0.025	Annual investment-to-capital ratio = 10%
χ: Relative weight of leisure utility	1.778	Long-run average labor hours $= 0.33$
ρ_z: Persistence of TFP shock	0.974	Empirical estimation
σ_z: Volatility of TFP shock	0.009	Empirical estimation
λ: Degree of loss aversion	2.25	Experimental evidence

Two additional parameters are specific to our model. The first is λ, which governs the additional weight on loss relative to gain within the gain-loss utility. The other parameter θ governs the weight of the gain-loss utility relative to the standard consumption utility. Although these two parameters have distinct economic meanings, their values cannot be individually determined under in my model, where loss aversion affects the marginal utility by a factor of $(1 - \theta + \theta\lambda)$. My calibration strategy here is to use the point estimate of $\lambda = 2.25$ from Tversky and Kahneman (1992) throughout the numerical analysis, and experiment with different values of θ to see how the degree of loss aversion affects the business cycle dynamics.

2.3.3 Fixed Reference point

In this first exercise, I assume the most simplistic form of reference point – a fixed reference point for all consumption choices – and solve the model with value function iteration. In particular, I choose the reference point to be $\bar{x} = 0.7612$, the deterministic steady state level of consumption when there is no loss aversion ($\theta = 0$) and a value that is also equal or close to the mean of the ergodic distribution of consumption in the model economy with or without loss aversion. In other words, in the absence of loss aversion, consumption choices are higher (lower) than this reference level under positive (negative) TFP shocks. I then introduce loss aversion into the model and study how it affects the consumption and hours decisions.

Decision Rules for Consumption and Labor Hours

Figure 2.2
Decision Rules
Loss Aversion in Consumption
($\bar{x} = 0.7612, z_t = 1$)

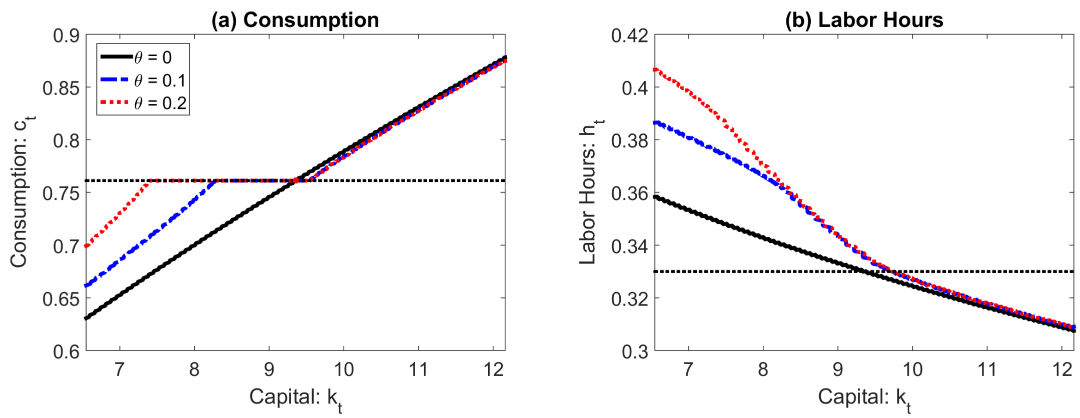
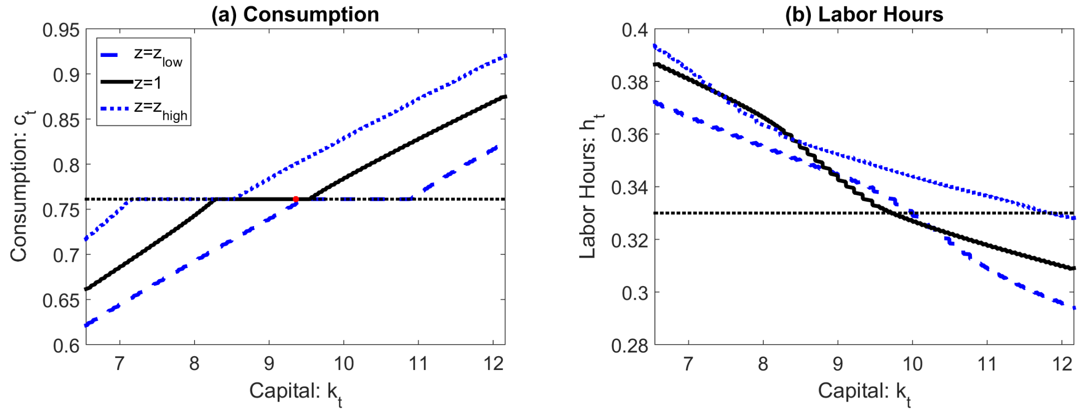


Figure 2.2 shows the consumption and hours decision rules under various degrees of loss aversion and Figure 2.3 shows the decision rules under various TFP shocks. As Figure 2.2 panel (a) shows, loss aversion in consumption induces a flat region at the reference level in the consumption decision rules. This is a striking and intuitive result: loss aversion implies very high local risk aversion around the reference point; given that the reference point is fixed, consumption does not fall below the reference level, i.e. "into the loss region", unless the capital stock is sufficiently low. In other words, for capital not far below

Figure 2.3
Decision Rules
Loss Aversion in Consumption
($\bar{x} = 0.7612, \theta = 0.1$)



the steady state level, the agent keeps consumption at the reference level to avoid incurring a loss in the gain-loss utility. In a model where hours are free to adjust, the agent increases hours (Figure 2.2 panel (b)) as capital falls to maintain the reference consumption level. The marginal benefit of doing so diminishes when capital becomes sufficiently low, and at some point the consumption decision rule resumes its usual shape. The higher the degree of loss aversion, the lower this border-line capital level, and the wider the flat region. In contrast, when capital is high, we see little deviation of the consumption and labor decision rules from the RBC benchmark. This is because loss aversion does not affect inter-temporal choices in the region of gain.

When controlling for the degree of loss aversion (Figure 2.3), consumption is now only weakly increasing in technology levels, due to the presence of the flat regions. Meanwhile, labor hours decision rules are no longer monotonic in technology levels, as hours serve as an adjustment channel for the corresponding consumption choices. As a result, hours appear to be much less pro-cyclical in this model.

Figure 2.4
Simulated Consumption Path
Loss Aversion in Consumption
($\bar{x} = 0.7612$)

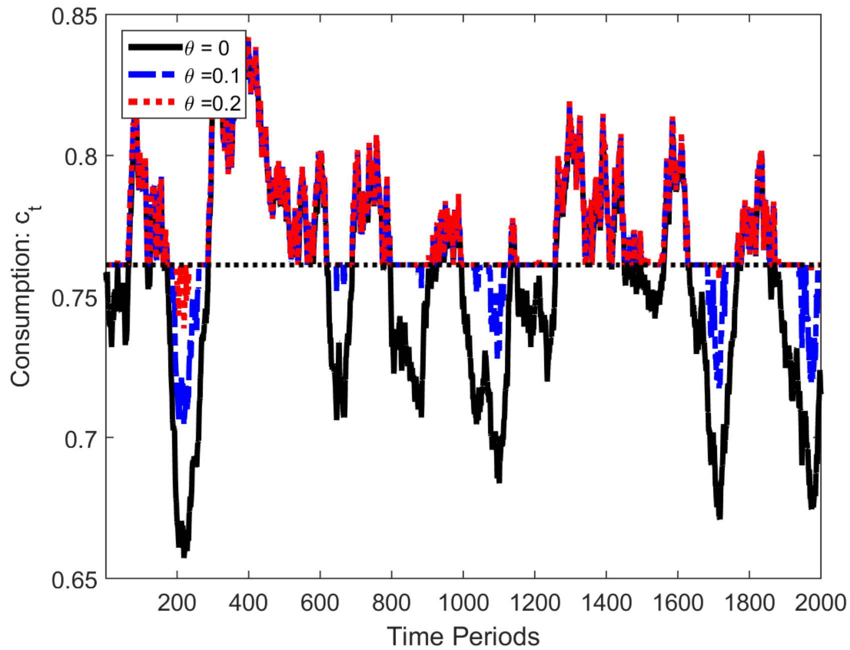


Figure 2.4 depicts a typical consumption path implied by this version of the model. The path features periods of stagnation at the reference level. In addition, consumption tracks the RBC benchmark closely in booms, but rarely falls below the reference level in recessions unless the negative shocks are very large. These features can all be understood from the shape of the decision rules.

Business cycle statistics

Table 2.2 presents business cycle statistics from the U.S. data, the benchmark RBC model ($\theta = 0$), and the model with loss aversion ($\theta > 0$) with a fixed reference point of 0.7612. In calculating the empirical business cycle statistics (column 1), I use quarterly data on U.S. economy in 1948Q1-2013Q1. All data are taken from the Federal Reserve

Economic Data (FRED) website of the St. Louis Federal Reserve. Data are in real per capita terms and detrended using the HP-filter with a smoothing parameter of 1600.

As Columns 3, 4 and 5 in Table 2.2 show, output, consumption, and hours are less volatile under loss aversion, as there is higher effective risk aversion in the economy and consumption is less responsive to negative technology shocks. The relative volatilities of consumption and hours, however, are not much affected when compared to the RBC benchmark. In terms of prices, it is well-known that the simple RBC model is unable to match the moments in data and adding loss aversion does not improve the fit.[9]

In terms of cyclicality, loss aversion lowers the contemporaneous correlation with output of hours significantly and that of consumption and real wages mildly. Note that these changes are in the direction of realism. Lastly, introducing loss aversion in consumption does not affect the persistence of macro variables in the model economy.

Asymmetric impulse responses

Figure 2.5 and Figure 2.6 show the impulse responses of consumption and hours to TFP shocks. In particular, Figure 2.5 plots the impulse response profile of consumption and hours in time to a unit ($\sigma_z = 0.009$) positive TFP shock under various degrees of loss aversion when the reference consumption is fixed at its deterministic steady state level. The figure shows that the greater the degree of loss aversion, the lower the initial and peak responses of consumption and hours.

[9]Loss aversion does not seem to affect the business cycle moments of the real interest rate at all, while both the absolute and relative volatilities of real wages are higher under loss aversion. This can be understood from equation (2.7), $w_t = \chi \cdot \frac{h'(1-n_t)}{[1-\theta+\theta v'(\Delta u_t|x_t)]u'(c_t)}$, the intra-temporal optimality condition. The marginal utility from the gain-loss utility function $v'(.)$ is stepwise around the reference point and hence is volatile over the business cycles, giving rise to volatile wages as well.

Table 2.2
Business Cycle Statistics
Loss Aversion in Consumption with Fixed Reference Point

Variables	US Data	RBC ($\theta = 0$)	Loss Aversion ($\theta = 0.1$)	Loss Aversion ($\theta = 0.2$)	Loss Aversion ($\theta = 0.3$)
Standard Deviation					
Output	1.69	1.69	1.37	1.29	1.28
Consumption	0.90	0.67	0.64	0.59	0.58
Hours	1.90	0.73	0.49	0.56	0.58
Investment	4.70	5.25	4.18	3.97	3.97
Wage	0.90	0.99	1.16	1.21	1.22
Interest Rate	0.40	0.06	0.05	0.05	0.05
TFP	1.20	1.20	1.20	1.20	1.20
Standard Deviation relative to Output					
Output	1.00	1.00	1.00	1.00	1.00
Consumption	0.53	0.40	0.47	0.46	0.45
Hours	1.12	0.43	0.36	0.44	0.45
Investment	2.76	3.12	3.07	3.09	3.11
Wage	0.53	0.59	0.85	0.94	0.96
Interest Rate	0.24	0.04	0.04	0.04	0.04
TFP	0.71	0.71	0.88	0.93	0.94
First-order Auto-Correlation					
Output	0.85	0.72	0.72	0.72	0.72
Consumption	0.79	0.77	0.75	0.75	0.75
Hours	0.90	0.70	0.70	0.71	0.71
Investment	0.87	0.70	0.70	0.70	0.70
Wage	0.73	0.74	0.73	0.73	0.73
Interest Rate	0.42	0.71	0.71	0.71	0.71
TFP	0.75	0.72	0.72	0.72	0.72
Contemporaneous correlation with Output					
Output	1.00	1.00	1.00	1.00	1.00
Consumption	0.76	0.93	0.85	0.84	0.83
Hours	0.88	0.97	0.60	0.35	0.32
Investment	0.79	0.99	0.96	0.96	0.96
Wage	0.10	0.99	0.94	0.90	0.90
Interest Rate	0.00	0.96	0.96	0.96	0.96
TFP	0.76	1.00	0.98	0.96	0.96

Figure 2.5
Impulse Responses to unit Positive shocks
Loss Aversion in Consumption
$(\bar{x} = 0.7612)$

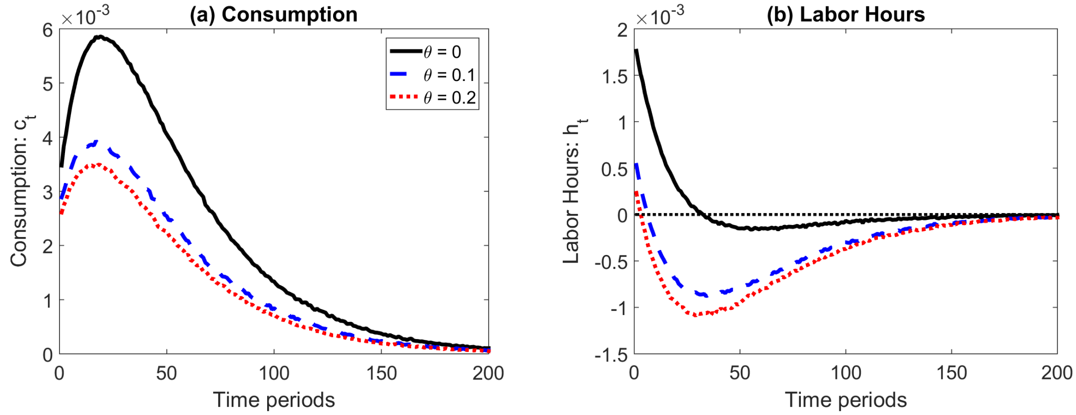


Figure 2.6
Impulse Responses to unit Positive and Negative shocks
Loss Aversion in Consumption
$(\bar{x} = 0.7612)$

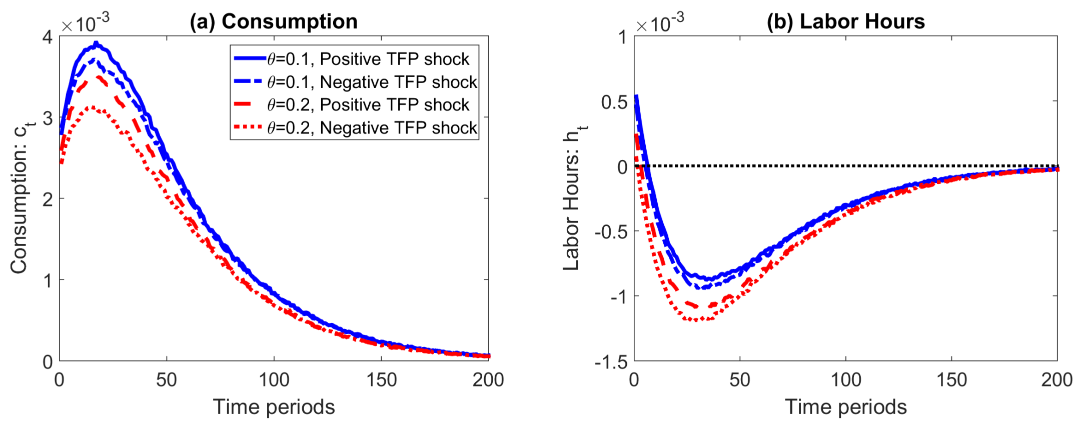
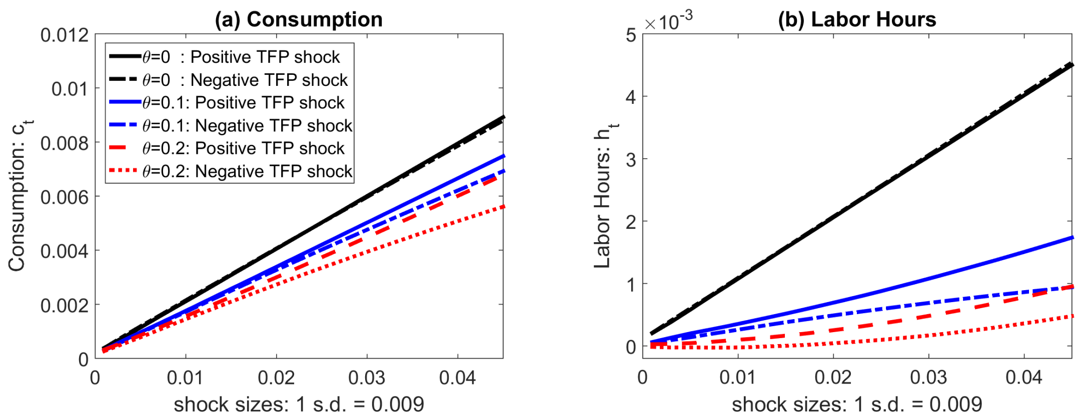


Figure 2.6 provides a clear illustration of asymmetry, the major property that distinguishes loss aversion from habit formation. Recall that a loss averse consumer is less willing to lower her consumption in response to a negative shock than to raise her consumption in response to a positive shock. As a result, the profile of the impulse responses of consumption and hours to a unit negative TFP shock is lower than that to a unit positive

shock, both initially and at peak. [10] In addition, Figure 2.7 clearly shows that the magnitude of asymmetry in the *initial* responses of consumption and labor hours increases with the size of the shock.

Figure 2.7
Initial Responses of Consumption and Labor Hours
Loss Aversion in Consumption
$(\bar{x} = 0.7612)$



Calibration revisited: the Role of the Reference point

The results above are obtained under a fixed reference point of 0.7612, the deterministic steady state of consumption, where positive (negative) shocks generate consumption decisions above (below) the reference point when there is no loss aversion. Under this reference point, loss aversion implies a reluctance to lower consumption in response to negative shocks, and as a result, negative shocks induce lower impulse responses (initial and peak) of consumption and hours than positive shocks. Moreover, the tendency to increase consumption readily but to lower it reluctantly means that the *ex post* ergodic mean of consumption will turn to be higher than the *ex ante* reference point.

---

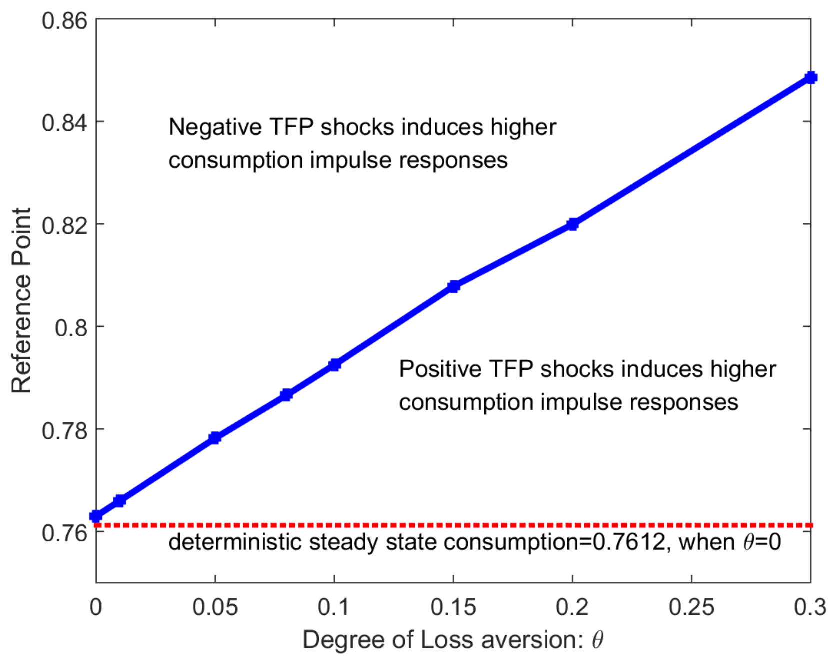
[10] The impulse responses to negative shocks are reversed in sign to help compare the magnitude.

It is worth noting that the direction of the asymmetry can be reversed if the reference point is set higher. In particular, I find that if the reference consumption level is "unrealistically" high such that the model economy only operates in the "loss region", then negative shocks can induce higher impulse responses (initial and peak) than positive shocks. Of course, for a negative unit TFP shock to generate larger impulse responses than a positive unit shock is at odds with the original idea of loss aversion, and thus I do not consider this case in my model. Nonetheless, this feature can be understood from the shape of the decision rules: when the reference point is very high, only under very large positive shocks is the presence of the reference point relevant for local dynamics, and the flat region around the high reference point in the consumption decision rule means that consumption may not rise with large positive shocks but will fall without constraint to negative shocks. As a result, the *ex post* ergodic mean of consumption will be lower than the *ex ante* reference point in this case.

There is a point where the *ex post* ergodic mean of consumption exactly coincides with the *ex ante* reference point and the asymmetry in the consumption and hours impulse responses will be completely muted. I call this critical reference point "the gain-loss-neutral reference point". Figure 2.8 shows the gain-loss-neutral reference points for each level of θ. Unsurprisingly, the greater the degree of loss aversion, the higher the gain-loss-neutral reference point.

To summarize, the direction of the asymmetry implied by loss aversion depends critically on the level of the reference point in this general equilibrium model. In terms of modeling, this is a matter of parametrization. My choice in this study $\bar{x} = 0.7612$, as well as its quantitative implications, is in line with the existing literature on loss aversion

Figure 2.8
"Gain-Loss-Neutral" Reference Point



in consumption-saving models (Bowman, Minehart, and Rabin, 1999) and the empirical evidence (Shea, 1995a,b).

Last but not least, an alternative way to calibrate the model is to introduce an "impatience" condition, i.e. lower β, such that the gain-loss-neutral reference point coincides with the deterministic steady state level of consumption without loss aversion, i.e. 0.7612. In other words, I compare a standard RBC model without loss aversion but with a standard value for discount factor ($\beta = 0.99$) to a model with loss aversion and with a lower discount factor, in which I choose the discount factor so that the *ex post* ergodic mean of consumption (which in the model is also the *ex ante* reference point) is the same as in the standard RBC model. The "impatience" condition is an idea borrowed from the literature on precautionary saving, and the qualitative results under this calibration strategy are similar to those presented before. My results show that the higher the degree of loss aversion, the more impatient the agent needs to be for the gain-loss-neutral reference point to be at 0.7612, which also results in a lower ergodic mean of the aggregate variables in the model economy.

2.3.4 Status quo as Reference point

In the previous section, I showed how loss aversion with a fixed reference point in consumption generates a flat region in the decision rule, and causes asymmetric impulse responses to positive and negative income shocks. The results are intuitive and clear. However, in reality and in the literature, fixed reference points are rare. In this section, I turn to a much more commonly used reference point – the status quo. Specifically, I assume that the loss averse representative consumer uses the one-period-lagged *aggregate* consumption as her reference point in each period: $x_t = C_{t-1}$, following the habit formation literature.

Status quo differs from a fixed reference point in two ways: first, it is a (past) choice itself, and therefore will have dynamic implications in a general equilibrium framework; and second, it is moving period by period. In this section, I first modify the model so that it can be solved using local approximation, and then discuss the results.

Differentiable gain-loss utility function

As mentioned above, in order to apply local approximation around the steady state, the objective function needs to be differentiable at that point. However, the gain-loss utility in the objective function is by definition not differentiable at the reference point, which under the assumption of status-quo reference equals exactly the steady state consumption in equilibrium. To remedy this issue, I use a “smooth-switching” approximation of the gain-loss utility to make it everywhere differentiable, following Rosenblatt-Wisch (2008). I use the following switching function to approximate the loss aversion coefficient:

$$\tilde{\lambda}(z) \approx 1 + \frac{\gamma}{1 + e^{\mu z}} \tag{2.14}$$

where z is the argument inside the gain-loss utility function. Note that the degree of loss aversion $\tilde{\lambda}$ in this formulation is bounded by $[1, \gamma + 1]$. In particular, $\tilde{\lambda} \to 1$ for $z \to \infty$ and $\tilde{\lambda} \to (1 + \gamma)$ for $z \to -\infty$. Therefore, in the approximated gain-loss function, γ is analogous to λ in that it governs the degree of loss aversion. The other parameter μ controls the speed of switching: the higher μ is, the higher the speed of switching. The smooth-switching approximation of the gain-loss utility is therefore given by:

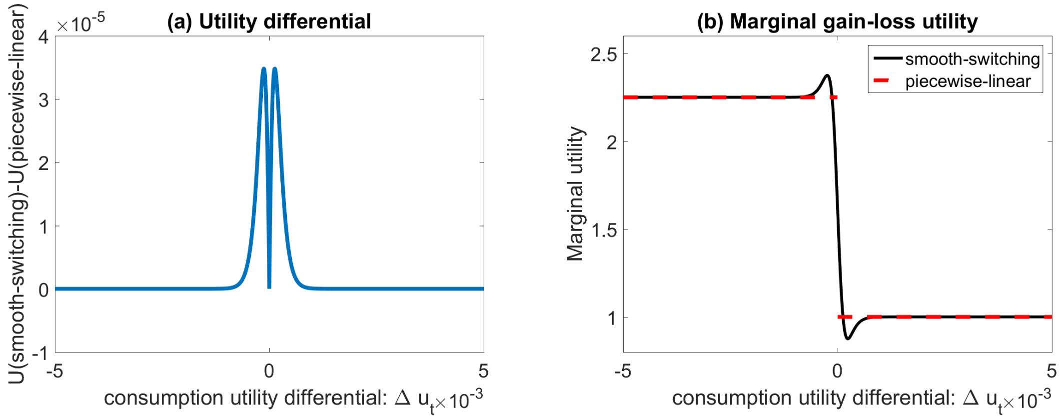
$$\tilde{v}(z) = z \left[1 + \frac{\gamma}{1 + e^{\mu z}} \right] \tag{2.15}$$

Note that $\tilde{v}'(0) = 1 + \frac{\gamma}{2}$, whereas in the exact piece-wise gain-loss utility, $v'(0^+) = 1$ and $v'(0^-) = 1 + \gamma$. However, as we move away from the reference point, a pair of $\gamma = 1.25$ and $\mu = 10000$ gives a very good approximation of the piecewise-linear gain-loss utility with $\lambda = 2.25$: Panel (a) of Figure 2.9 plots the difference between the piece-wise linear gain-loss utility and its smooth-switching approximation. It is clear that the approximation is nearly perfect everywhere except around the reference point, where the difference is still well within 10^{-4}.[11] Panel (b) of Figure 2.9 plots the marginal utility of the piecewise-linear utility function and that of the smooth-switching utility. We can see how the approximation

[11] Under our baseline calibration, the numbers which go inside the gain-loss utility $v(.)$ are roughly of magnitude 10^{-2}. Therefore the smooth-switching function should provide a good approximation when the status quo is used as reference point. This is confirmed by the sensibility of the results.

is successful in capturing the sharp fall of the marginal utility around that point, the most important feature of loss aversion.

Figure 2.9
Piecewise-linear vs. Smooth-switching Utility Functions



Steady state and Business cycle statistics

Now that the gain-loss utility function has been made differentiable everywhere by the smooth-switching approximation, first order conditions pin down a unique steady state for the model economy. Under the benchmark calibration, the higher the degree of loss aversion, the higher the levels of capital, output, and consumption in this approximated steady state. This is a sensible result, as the gain-loss utility increases the overall curvature in the total utility function.

Table 2.3 reports the business cycle statistics of various real aggregate variables under different degrees of loss aversion, with the first column reporting the statistics from the US data. Again, the second column is obtained by solving the benchmark RBC model with local approximation around steady state. The numbers presented there are slightly different from the second column in Table 2.2, which were based on value function iteration. The differences simply reflect the change in the solution method.

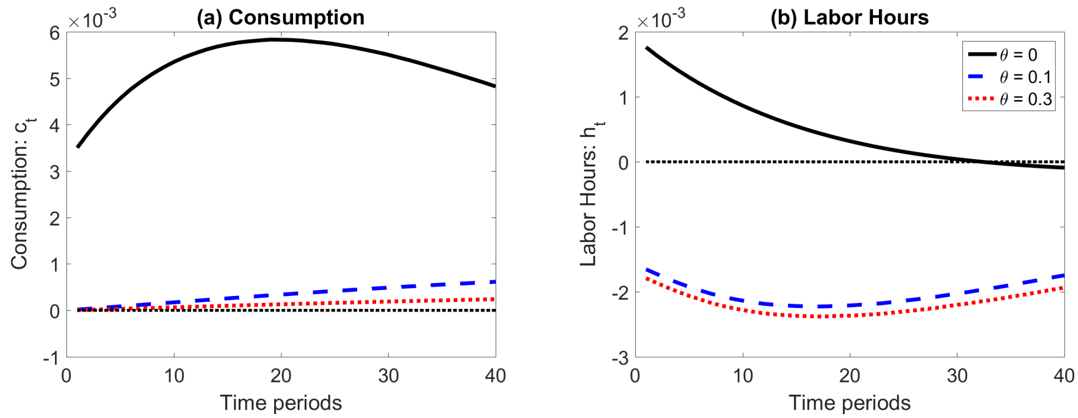
The results suggest that once loss aversion is added to the model, consumption becomes an order of magnitude smoother than without loss aversion. Even with a very low weight such as $\theta = 0.05$, the volatility of output falls sharply, the relative volatility of consumption drops to near zero and consumption becomes acyclical. The intuition behind this strong result is that unlike the case of fixed reference points, the reference points used in this model are determined in equilibrium by inter-temporal choices, in which case loss aversion implies very strong local risk aversion and a very low inter-temporal elasticity of substitution. Not only are loss averse consumers reluctant to lower consumption in response to negative shocks, they are also unwilling to raise consumption in response to positive shocks, because they want to avoid incurring a loss in the next period. In a production economy where the labor-leisure choice is free to adjust, loss averse consumers will optimally adjust to productivity shocks by moving labor input in the opposite direction. As a result, consumption becomes very unresponsive to aggregate shocks in general. Figure 2.10 shows the impulse response functions of consumption and hours with and without loss aversion to a unit positive TFP shock. It is clear that the effect of loss aversion is very strong: the response of consumption drops to near zero on impact and remains essentially flat thereafter. Hours fall sharply in response to higher technology, suggesting very strong income effects. The counter-cyclical labor supply contributes to the low volatility of output, and the relative volatility of hours and wages are both doubled, reflecting the fact that all adjustments to shocks in the economy go through the labor-leisure channel.

Loss aversion does not appear to affect the persistence in any variables other than consumption, nor does it affect cyclical comovements, except for consumption and labor. In

Table 2.3
Business Cycle Statistics
Loss aversion in Consumption with Status quo as Reference point

Variables	US Data	RBC ($\theta = 0$)	Loss Aversion ($\theta = 0.05$)	($\theta = 0.1$)	($\theta = 0.3$)
Calibration: χ		1.78	1.83	1.89	2.11
			Standard Deviation		
Output	1.69	1.70	0.81	0.76	0.73
Consumption	0.90	0.69	0.02	0.01	0.01
Hours	1.90	0.72	0.61	0.69	0.74
Investment	4.70	5.21	3.40	3.19	3.03
Wage	0.90	1.00	1.42	1.44	1.46
Interest Rate	0.40	0.06	0.03	0.03	0.03
TFP	1.20	1.21	1.21	1.21	1.21
			Standard Deviation relative to output		
Output	1	1	1	1	1
Consumption	0.53	0.40	0.03	0.02	0.01
Hours	1.12	0.43	0.76	0.90	1.02
Investment	2.76	3.07	4.19	4.18	4.17
Wage	0.53	0.59	1.75	1.90	2.01
Interest	0.24	0.04	0.04	0.04	0.04
TFP	0.71	0.71	1.49	1.59	1.66
			First-order Auto-correlation		
Output	0.85	0.73	0.72	0.72	0.72
Consumption	0.79	0.79	0.97	0.97	0.97
Hours	0.90	0.72	0.77	0.76	0.76
Investment	0.87	0.79	0.72	0.72	0.72
Wage	0.73	0.75	0.74	0.74	0.74
Interest	0.42	0.72	0.73	0.74	0.74
TFP	0.75	0.73	0.73	0.73	0.73
			Contemporaneous correlation with output		
Output	1	1	1	1	1
Consumption	0.76	0.94	-0.02	-0.04	-0.05
Hours	0.88	0.98	-0.94	-0.95	-0.95
Investment	0.79	0.99	1.00	1.00	1.00
Wage	0.10	0.99	0.99	0.99	0.99
Interest	0.00	0.96	0.93	0.93	0.93
TFP	0.76	1.00	1.00	1.00	1.00

Figure 2.10
Impulse Responses to unit Positive shocks
Loss Aversion in Consumption
Status-quo as Reference point



addition, the negative contemporaneous correlation of hours with output indicates counter-cyclical labor supply in the model. This is not surprising: given that the real wage is pro-cyclical and consumption is completely unresponsive to shocks, the consumer needs to work more hours in recessions in order to maintain the consumption level. In other words, the income effect is so strong under status-quo reference that the labor supply curve bends backwards.

In an experiment not reported here, I use one-period-lagged *own* consumption as the reference point ("internal reference") and repeat the exercise. Note that under this setup, the marginal utility of consumption involves future period terms, as now the consumer takes into account the effect of her current-period choice on the reference point in the future. Unsurprisingly, consumption becomes even smoother than in our baseline model, as the effective inter-temporal elasticity of substitution is even lower.

This version of the model fails to generate quantitatively meaningful asymmetric impulse responses. The reason is obvious: there is excessive smoothness in consumption caused by the agents' reluctance to adjust consumption levels in either direction.

Comparison: Habit formation

When the status quo is used as the reference point, it is instructive to compare loss aversion to habit formation. Two points can be noted. First, when the status quo is used as the reference point, both loss aversion and habit formation lead to excessive consumption smoothing in a frictionless RBC environment: Lettau and Uhlig (2000) embed habit formation into a RBC environment and obtain extremely smooth consumption under a lagged-consumption reference point. The resemblance is not surprising, as both loss aversion and habit formation feature reference-dependence, which under concave preferences generates incentives to smooth not only levels but also changes in levels. Therefore, both loss aversion and habit formation induce extreme local risk aversion around the reference point and a low effective inter-temporal elasticity of substitution.

Second, the most important feature that distinguishes loss aversion from habit formation is the fundamental asymmetry between gains and losses embedded in loss aversion. However, excessive smoothness prevents the display of any meaningful asymmetry under fast-moving reference points. In other words, when the reference point is updated frequently and tracks choices, reference-dependence dominates asymmetry in impulse responses. But when the reference point is fixed at an appropriate level, as shown in the previous section, consumption and hours exhibit quantitatively significant asymmetric impulse responses to shocks. In contrast, habit formation does not imply such an asymmetry.

2.4 Loss Aversion in Consumption and Leisure

2.4.1 Loss aversion in consumption and leisure: Jointly

In the previous sections, I showed how the labor-leisure choice works as an important adjustment channel to allow consumption smoothing under loss aversion in consumption only. In this section, I extend the model by introducing loss aversion to the consumption-leisure composite good and investigate its business cycle implications.

The model

The total utility now takes the following form:

$$U(c, l, x^c, x^l) = (1-\theta)u(c,l) + \theta v[u(c,l) - u(x^c, x^l)] \tag{2.16}$$

where $l = 1 - n$ is leisure, and x^c, x^l are the reference points for consumption and leisure, respectively.

The most significant difference between this version of the model and the previous formulation is that once loss aversion is introduced into consumption and leisure choices "jointly", it no longer affects the contemporaneous rate of substitution between the two, given that the reference consumption and the reference leisure are both external to the agent. In other words, loss aversion no longer enters the intra-temporal optimality condition:

$$\frac{\partial u(c_t, l_t)/\partial l_t}{\partial u(c_t, l_t)/\partial c_t} = w_t \tag{2.17}$$

where w_t is the real wage. Meanwhile, loss aversion continues to affect the inter-temporal choice of the consumption-leisure composite good. The inter-temporal Euler equation is given by:

$$\begin{aligned}[1-\theta+\theta v'(\Delta u_t|x_t)] \quad \cdot \quad & \frac{\partial u(c_t,l_t)}{\partial c_t} = \\ & \beta E_t\left\{[1-\theta+\theta v'(\Delta u_{t+1}|x_{t+1})]\frac{\partial u(c_{t+1},l_{t+1})}{\partial c_{t+1}}(r_{t+1}+1-\delta)\right\}\end{aligned}$$

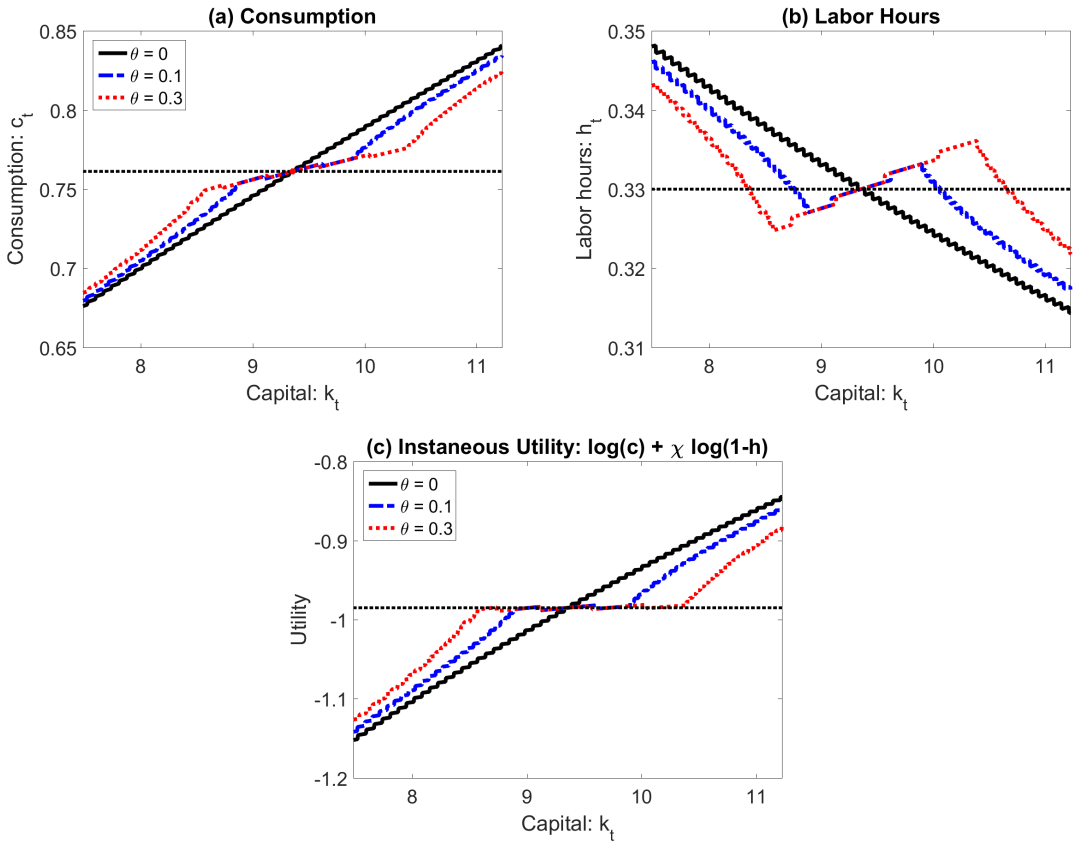
except that now $\Delta u_t|x_t = u(c_t,l_t) - u(x_t^c,x_t^l)$. I assume a fixed reference point for both consumption and leisure so that I can solve the model with value function iteration under exactly the same calibration as before.[12] In particular, I use the steady state levels of consumption and labor hours as the reference point: $\bar{x}^c = 0.7612$ and $1-\bar{x}^l = 0.33$. This implies a reference utility level of $u(\bar{x}^c,\bar{x}^l) = -0.985$.

Results

Panel (a) and (b) of Figure 2.11 show the consumption and labor hours decision rules for various degrees of loss aversion in consumption and leisure jointly. As panel (a) shows, the consumption decision rules become flatter around the reference level, but are not completely flat as in the case of loss aversion in consumption only. This is because under composite loss aversion, the effective reference point is not the reference consumption or the reference leisure, but the reference utility $u(x^c,x^l) = \log(x^c) + \chi log(1-x^l)$. As panel (c) shows, the composite utility $\log(c_t) + \chi\log(1-n_t)$ is indeed completely flat around

[12]Now that loss aversion does not enter the intra-temporal optimality condition, it is easier to evaluate the auxiliary choice matrix of $\{c_t,n_t\}$ given the $\{k_t,z_t\}\times k_{t+1}$ state-choice space, as these values no longer depend on the reference points.

Figure 2.11
Decision Rules
Loss aversion in Consumption and Leisure: Jointly
$(u(\bar{x}^c, \bar{x}^l) = -0.985, z_t = 1)$



this reference point of -0.985.

Since both consumption and labor hours can adjust to keep the composite utility from falling below its reference level and loss aversion does not have much impact on either the consumption or hours decision rules individually, the business cycle dynamics in this model are not much affected by loss aversion, except that the volatility of consumption turns out to be slightly higher and that of output and hours slightly lower than the RBC benchmark (Table 2.4). Consumption and hours are both less pro-cyclical as compared to the RBC benchmark, improving the fit of the model to the data. In addition, we do not see a significant increase in the volatility of wages in this model, as the gain-loss utility does not affect the intra-temporal optimality condition stated in Equation 2.17. Finally, this model fails to generate quantitatively significant asymmetric impulse responses of consumption and hours under the current calibration.

2.4.2 Loss aversion in consumption and leisure: Separately

In the previous section, I introduced loss aversion in the consumption-leisure composite good; in this section, I impose loss aversion on consumption and leisure separately. That is, I assume a reference level of leisure below which the agent feels a sense of loss no matter how high her consumption is.

The model

I assume that the total utility function has the following form:

$$U(c,l,x^c,x^l) = \{(1-\theta_1)u(c) + \theta_1 v\left[u(c) - u(x^c)\right]\} + \chi\left\{(1-\theta_2)h(l) + \theta_2 v\left[h(l) - h(x^l)\right]\right\} \tag{2.18}$$

where $l = 1 - n$ is leisure, and x^c and x^l are reference consumption and reference leisure, respectively. Parameter θ_1 governs the degree of loss aversion in consumption, and θ_2 governs that in leisure. The inter-temporal Euler equation and intra-temporal optimality condition for the consumer's problem are given by:

$$\left[1-\theta_1+\theta_1 v'(\Delta u_t|x_t^c)\right]u'(c_t) = \beta E_t\left\{\left[1-\theta_1+\theta_1 v'(\Delta u_{t+1}|x_{t+1}^c)\right]u'(c_{t+1})\left(r_{t+1}+1-\delta\right)\right\} \tag{2.19}$$

Table 2.4
Business Cycle Statistics
Loss aversion in Consumption and Leisure: Jointly

Variables	US Data	RBC ($\theta = 0$)	Loss Aversion ($\theta = 0.1$)	Loss Aversion ($\theta = 0.2$)	Loss Aversion ($\theta = 0.3$)
			Standard Deviation		
Output	1.69	1.68	1.54	1.49	1.46
Consumption	0.90	0.67	0.84	0.91	0.96
Hours	1.90	0.73	0.57	0.53	0.52
Investment	4.70	5.24	4.25	3.98	3.84
Wage	0.90	0.99	1.05	1.07	1.09
Interest Rate	0.40	0.06	0.06	0.05	0.05
TFP	1.20	1.20	1.20	1.20	1.20
			Standard Deviation relative to output		
Output	1	1	1	1	1
Consumption	0.53	0.40	0.54	0.61	0.65
Hours	1.12	0.43	0.37	0.36	0.36
Investment	2.76	3.11	2.75	2.66	2.63
Wage	0.53	0.59	0.68	0.72	0.75
Interest	0.24	0.04	0.04	0.04	0.04
TFP	0.71	0.71	0.78	0.80	0.82
			First-order Auto-correlation		
Output	0.85	0.72	0.72	0.72	0.72
Consumption	0.79	0.77	0.73	0.72	0.72
Hours	0.90	0.70	0.70	0.68	0.68
Investment	0.87	0.70	0.70	0.70	0.70
Wage	0.73	0.74	0.73	0.73	0.72
Interest	0.42	0.71	0.71	0.71	0.71
TFP	0.75	0.72	0.72	0.72	0.72
			Contemporaneous correlation with output		
Output	1	1	1	1	1
Consumption	0.76	0.93	0.92	0.90	0.88
Hours	0.88	0.97	0.92	0.86	0.80
Investment	0.79	0.99	0.96	0.94	0.92
Wage	0.10	0.99	0.98	0.97	0.96
Interest	0.00	0.96	0.97	0.97	0.97
TFP	0.76	1.00	0.99	0.99	0.98

$$w_t - \chi \cdot \frac{\left[1 - \theta_2 + \theta_2 v'(\Delta h_t | x_t^l)\right] h'(l_t)}{\left[1 - \theta_1 + \theta_1 v'(\Delta u_t | x_t^c)\right] u'(c_t)} \tag{2.20}$$

where $\Delta u_t | x_t^c = u(c_t) - u(x_t^c)$ and $\Delta h_t | x_t^l = h(l_t) - h(x_t^l)$. I continue to use the piecewise-linear gain-loss utility $v(.)$, and solve the model with value function iteration under the same calibration of parameter values. Again, I use the deterministic steady state consumption and labor hours, $\bar{x}^c = 0.7612$ and $1 - \bar{x}^l = 0.33$, as the fixed reference points for consumption and leisure.

Results

Figure 2.12
Decision Rules
Loss aversion in Consumption and Leisure: Separately
$(\bar{x}^c = 0.7612, 1 - \bar{x}^l = 0.33, z_t = 1)$

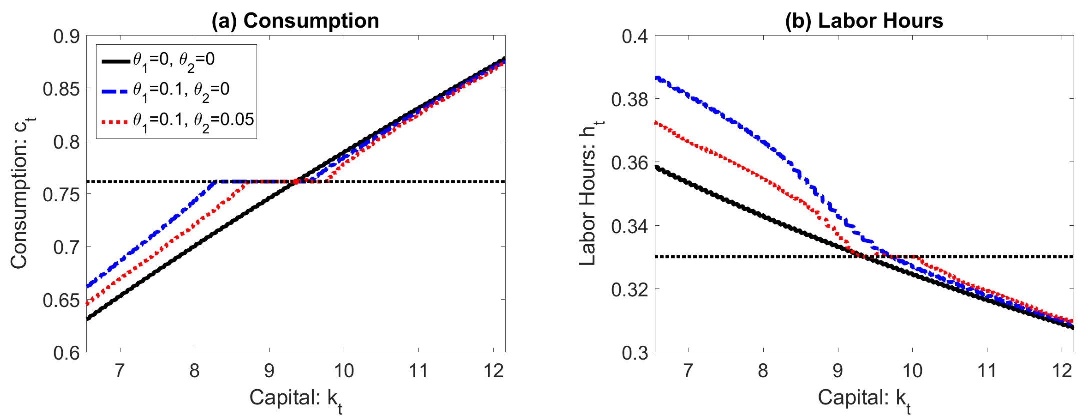
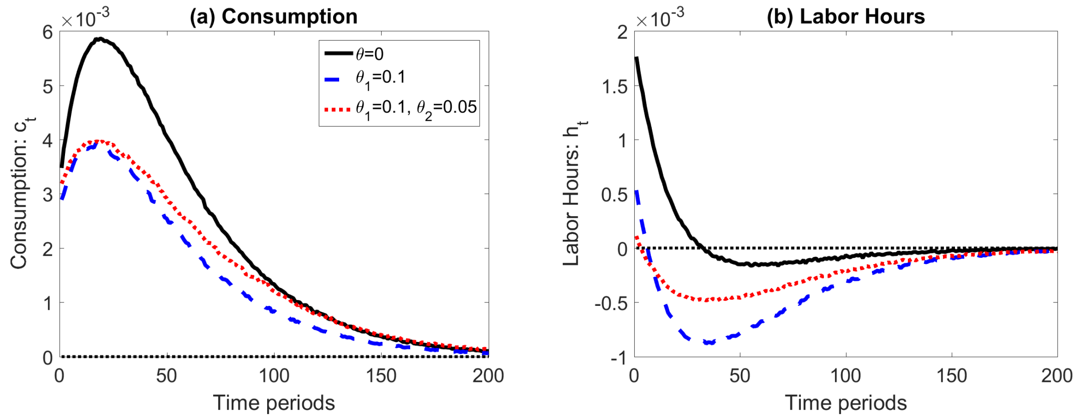


Figure 2.12 shows the consumption and labor hours decision rules and Figure 2.13 shows the impulse responses. Not surprisingly, introducing loss aversion in leisure separately induces a flat region in the hours decision rules around its reference level for exactly the same reason as before. As a result, hours are less responsive to shocks in this version of the model than without loss aversion in leisure (e.g. $\theta_1 = 0.1, \theta_2 = 0$). Consumption

Figure 2.13
Impulse Responses to unit Positive shocks
Loss aversion in Consumption and Leisure: Separately
$(\bar{x}^c = 0.7612, 1 - \bar{x}^l = 0.33)$



features a narrower flat region around the reference level and is slightly more responsive to shocks than when $\theta_2 = 0$. The reason is that hours no longer works as an adjustment channel under loss aversion in leisure.

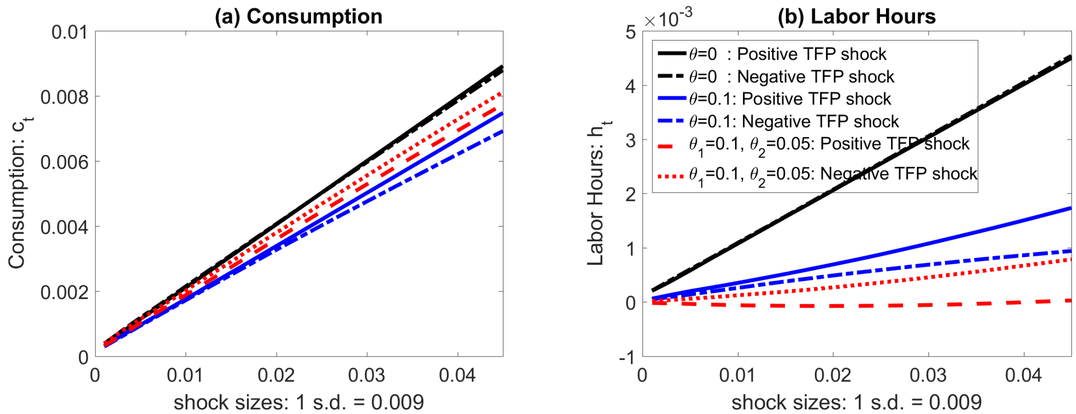
In terms of business cycle statistics (Table 2.5), hours become significantly less volatile and less pro-cyclical as compared to the case of loss aversion in consumption only. Consumption becomes slightly more volatile because hours no longer adjust freely. Real wages are more volatile under loss aversion as the marginal utilities of the gain-loss-utilities of both consumption and leisure affect the intra-temporal optimality condition (22), which determines the real wage in the equilibrium.

Figure 2.14 shows the initial impulse responses of consumption and hours to shocks of different sizes. As it shows, negative shocks induce higher responses of both consumption and hours than positive shocks in this version of the model under the current calibration. To understand this, suppose first that a negative shock hits and consumption and hours both fall: consumption falls into its "loss" region and leisure rises into its "gain" region

Table 2.5
Business Cycle Statistics
Loss aversion in Consumption and Leisure: Separately

Variables	US Data	RBC $\theta = 0$	Loss Aversion $\theta_1 = 0.1$, $\theta_2 = 0$	$\theta_1 = 0.1$, $\theta_2 = 0.05$	$\theta_1 = 0.1$, $\theta_2 = 0.1$	$\theta_1 = 0.1$, $\theta_2 = 0.2$
			Standard Deviation			
Output	1.69	1.69	1.29	1.25	1.23	1.21
Consumption	0.90	0.67	0.63	0.70	0.74	0.74
Hours	1.90	0.73	0.44	0.32	0.13	0.09
Investment	4.70	5.25	4.02	3.72	3.42	3.37
Wage	0.90	0.99	1.19	1.19	1.19	1.19
Interest Rate	0.40	0.06	0.05	0.05	0.04	0.04
TFP	1.20	1.20	1.20	1.20	1.20	1.20
			Standard Deviation relative to Output			
Output	1.00	1.00	1.00	1.00	1.00	1.00
Consumption	0.53	0.40	0.49	0.55	0.61	0.61
Hours	1.12	0.43	0.34	0.26	0.10	0.07
Investment	2.76	3.12	3.12	2.97	2.79	2.78
Wage	0.53	0.59	0.93	0.95	0.97	0.98
Interest Rate	0.24	0.04	0.04	0.04	0.04	0.04
TFP	0.71	0.71	0.93	0.96	0.98	0.99
			First-order Auto-Correlation			
Output	0.85	0.72	0.72	0.72	0.72	0.72
Consumption	0.79	0.77	0.74	0.73	0.72	0.71
Hours	0.90	0.70	0.70	0.69	0.61	0.59
Investment	0.87	0.70	0.70	0.71	0.71	0.71
Wage	0.73	0.74	0.73	0.73	0.72	0.72
Interest Rate	0.42	0.71	0.71	0.71	0.72	0.72
TFP	0.75	0.72	0.72	0.72	0.72	0.72
			Contemporaneous correlation with Output			
Output	1.00	1.00	1.00	1.00	1.00	1.00
Consumption	0.76	0.93	0.82	0.84	0.85	0.85
Hours	0.88	0.97	0.37	0.31	0.37	0.27
Investment	0.79	0.99	0.95	0.94	0.93	0.92
Wage	0.10	0.99	0.94	0.97	1.00	1.00
Interest Rate	0.00	0.96	0.96	0.97	0.97	0.97
TFP	0.76	1.00	0.98	0.98	1.00	1.00

Figure 2.14
Initial Responses of Consumption and Labor Hours
Loss Aversion in Consumption and Leisure: Separately
$(\bar{x}^c = 0.7612, 1 - \bar{x}^l = 0.33)$



as falling hours imply rising leisure. When a positive shock hits, hours rise and leisure falls, but falls less than it rises under negative shocks, due to loss aversion. As hours rises less, consumption rise less under positive shocks than it falls under negative shocks. In other words, under the current calibration, the loss aversion in leisure dominates that in consumption, resulting in the pattern of asymmetry shown in Figure 2.14.

2.5 Discussion

Having discussed the dynamic implications of loss aversion under various specifications in a simple RBC framework, I now discuss the literature on asymmetric business cycles and how loss aversion may help explain asymmetry.

Following the discussion in Van Nieuwerburgh and Veldkamp (2006), asymmetries in business cycles can be categorized into three types. The first type is "level asymmetry" or deepness asymmetry, which refers to asymmetry in the unconditional distribution of de-trended aggregate variables. In terms of modeling, level asymmetry usually results from some type of external constraint. For example, Hansen and Prescott (2005) show that booms are not as large a deviation from trend as recessions due to firm capacity constraints. Kocherlakota (2000) assumes that credit constraints also bind during booms but not recessions. As a result, large negative shocks can lead to large cuts in production since agents cannot borrow, while positive shocks are attenuated using savings.

The second type is "delay asymmetry", which refers to asymmetry in the timing and duration of booms and recessions, in the sense that some aggregate variables, such as output, may stagnate at the trough of an otherwise symmetric cycle. Chamley and Gale (1994) generate delay asymmetry using irreversible investment in a game-theoretic environment. In their model, at low levels of production, firms wait to produce until they have learned other firms' investment, which leads to an overall delay in economic recovery. In the labor search-and-matching literature, Mortensen and Pissarides (1994) features different rates of job creation and job destruction. In their model, job destruction occurs immediately once the value to the firm and the worker of being matched is negative, but job creation takes place only with some probability. Thus employment can fall quickly and violently, but it must expand slowly.

As this study shows, loss aversion provides a promising alternative mechanism to generate both level asymmetry and delay asymmetry in business cycles. Specifically, this study shows that with a fixed reference point, loss aversion in consumption alone, or in consumption and leisure separately, can generate asymmetric impulse responses (initial and peak) of consumption and hours to income shocks, i.e. "level asymmetry". Also, loss aversion causes stagnation of consumption at the reference level, which could be regarded as "delay" in recessions. Of course, the asymmetries I generate in this study are preliminary in the sense that they have shown clear patterns but do not yet fit the empirical data. However, the potential of loss aversion to generate meaningful asymmetric dynamics similar to those found in the data should be evident.

Lastly, the third type of asymmetry documented in the literature is “growth rate asymmetry”, which refers to the unconditional distribution of changes in aggregate variables. Growth rate asymmetry means that increases and decreases in aggregate variables have different distributions. Aruoba, Bocola, and Schorfheide (2012) used a structural quadratic auto-regression and found that negative shocks generate more persistent dynamics in aggregate variables than positive shocks in the years after 1984. In terms of modeling, growth rate asymmetry is usually generated by assuming some type of asymmetric learning process. Based on the findings of this study, loss aversion does not seem to generate significant asymmetry in the convergence rate of the consumption impulse responses (i.e. positive and negative impulse responses take about the same time to return to the steady state.) However, it would be interesting to explore the possible implications of loss aversion for growth asymmetry in a more complex model framework.

2.6 Conclusion

This chapter studies the business cycle implications of loss aversion in an otherwise standard RBC framework. My results indicate that the nature of the reference point plays an important role in determining how loss aversion affects business cycles. In the baseline model of loss aversion in consumption only, loss aversion induces excessive consumption smoothing under status-quo reference, which mutes the fundamental asymmetry implied by loss aversion. By contrast, under a fixed reference point, loss aversion induces a flat region in the consumption decision rules and stagnation in consumption paths around the reference level, which lead to both moderate consumption smoothing and asymmetric impulse responses to productivity shocks.

In terms of business cycle statistics, introducing loss aversion into a choice variable (e.g. consumption and/or hours) lowers the volatility of that choice variable in business cycles. In my modified RBC model, loss aversion improves the fit of the model to the data by lowering the contemporaneous correlation of consumption and hours with output. Loss aversion does not affect the persistence of aggregate variables.

Finally, one can ask whether it is appropriate to apply a behavioral element, like loss aversion, to a representative agent model directly, as I do in this study. In some way, this comes down to a question of aggregation, and to answer that question, one needs to build a heterogeneous-agent model and study the aggregation dynamics. However, the computation involved in solving such a model is likely to be difficult.

If we leave the issue of aggregation aside for the moment, and allow ourselves to view loss aversion as no more than a modeling tool that economists can utilize to study the aggregate economy, then the question becomes simple: can loss aversion offer a working mechanism in macro models to generate the asymmetries we observe in business cycles? The answer to this question is yes. As discussed above, loss aversion has the potential of generating both level asymmetry and delay asymmetry in DSGE models, which currently is largely missing in the existing literature. Although it remains unclear whether loss aversion will generate realistic predictions for asymmetric business cycles after incorporating it into a more complex model framework and fitting the model to the data, this study offers a reason to be hopeful.

Bibliography

AARONSON, S., T. CAJNER, B. FALLICK, F. GALBIS-REIG, C. SMITH, AND W. WASCHER (2014): "Labor Force Participation: Recent Developments and Future Prospects," *Brookings Papers on Economic Activity*, 2014(2), 197–275.

ACEMOGLU, D. (2002): "Directed Technical Change," *The Review of Economic Studies*, 69(4), 781–809.

——— (2010): "When Does Labor Scarcity Encourage Innovation?," *Journal of political economy*, 118(6), 1037–1078.

ACEMOGLU, D., AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 4, 1043–1171.

ACEMOGLU, D., G. GANCIA, AND F. ZILIBOTTI (2015): "Offshoring and Directed Technical Change," *American Economic Journal: Macroeconomics*, 7(3), 84–122.

ACEMOGLU, D., AND V. GUERRIERI (2008): "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 116(3), 467–498.

ACEMOGLU, D., AND P. RESTREPO (2015): "The Race between Man and Machine: Implications of Technology Growth, Factor Shares and Employment," Discussion paper, Massachusetts Institute of Technology.

ACEMOGLU, D., AND A. SCOTT (1997): "Asymmetric Business Cycles: Theory and Time-Series Evidence," *Journal of Monetary Economics*, 40(3), 501–533.

ARELLANO, C., L. MALIAR, S. MALIAR, AND V. TSYRENNIKOV (2013): "Envelope Condition Method Versus Endogenous Grid Method for Solving Dynamic Programming Problems," *Economics Letters*, 120(2), 262–266.

ARUOBA, S. B., L. BOCOLA, AND F. SCHORFHEIDE (2012): "A New Class of Nonlinear Time Series Models for the Evaluation of DSGE Models," *University of Maryland, manuscript*.

AUTOR, D., AND D. DORN (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market," *American Economic Review*, 103(5), 1553–1597.

AUTOR, D. H., D. DORN, AND G. H. HANSON (2015): "Untangling Trade and Technology: Evidence from Local Labour Markets," *The Economic Journal*, 125(584), 621–646.

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2006): "The Polarization of the U.S. Labor Market," *American Economic Review*, 96(2), 189–194.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, pp. 1279–1333.

BARBERIS, N., M. HUANG, AND T. SANTOS (2001): "Prospect Theory and Asset Prices," *The Quarterly Journal of Economics*, 116(1), 1–53.

BEAUDRY, P., D. A. GREEN, AND B. SAND (2013): "The Great Reversal in the Demand for Skill and Cognitive Tasks," *NBER Working Paper*, (w18901).

BENARTZI, S., AND R. H. THALER (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *The Quarterly Journal of Economics*, 110(1), 73–92.

BLINDER, A. S., AND A. B. KRUEGER (2013): "Alternative Measures of Offshorability: A Survey Approach," *Journal of Labor Economics*, 31(S1), S97 – S128.

BOWMAN, D., D. MINEHART, AND M. RABIN (1999): "Loss Aversion in A Consumption-Savings Model," *Journal of Economic Behavior & Organization*, 38(2), 155–178.

BRIDGMAN, B. (2014): "Is Labors Loss Capitals Gain? Gross versus Net Labor Shares," *Bureau of Economic Analysis (June 2014)*.

BRYNJOLFSSON, E., AND A. MCAFEE (2014): *The Second Machine Age: Work, Progress, and Prosperity in A Time of Brilliant Technologies*. WW Norton & Company.

CABALLERO, R. J., AND E. M. ENGEL (1993): "Microeconomic Adjustment Hazards and Aggregate Dynamics," *The Quarterly Journal of Economics*, 108(2), 359–383.

CABALLERO, R. J., E. M. ENGEL, AND J. HALTIWANGER (1997): "Aggregate Employment Dynamics: Building from Microeconomic Evidence," *American Economic Review*, pp. 115–137.

CAMERER, C. F. (2003): “Prospect Theory in the Wild: Evidence from the Field,” *Advances in Behavioral Economics*, pp. 148–161.

CAMPBELL, J. Y., AND J. H. COCHRANE (1999): “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *The Journal of Political Economy*, 107(2), 205–251.

CHAMLEY, C., AND D. GALE (1994): “Information Revelation and Strategic Delay in a Model of Investment,” *Econometrica: Journal of the Econometric Society*, pp. 1065–1085.

CHIRINKO, R. S. (2008): “σ: The Long and Short of It,” *Journal of Macroeconomics*, 30(2), 671–686.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): “Nominal Rigidities and the Dynamic Effects of A Shock to Monetary Policy,” *Journal of Political Economy*, 113(1), 1–45.

COOPER, R., J. HALTIWANGER, AND J. L. WILLIS (2015): “Dynamics of Labor Demand: Evidence from Plant-Level Observations and Aggregate Implications,” *Research in Economics*, 69(1), 37–50.

COOPER, R., AND J. L. WILLIS (2009): “The Cost of Labor Adjustment: Inferences from the Gap,” *Review of Economic dynamics*, 12(4), 632–647.

CORTES, G. M., N. JAIMOVICH, C. J. NEKARDA, AND H. E. SIU (2014): “The Micro and Macro of Disappearing Routine Jobs: A Flows Approach,” Discussion paper.

DICECIO, R. (2009): “Sticky Wages and Sectoral Labor Comovement,” *Journal of Economic Dynamics and Control*, 33(3), 538–553.

DORN, D. (2009): “Essays on Inequality, Spatial Interaction, and the Demand for Skills,” Ph.D. thesis, University of St. Gallen.

EDEN, M., AND P. GAGGL (2014): “The Substitution of ICT Capital for Routine Labor: Transitional Dynamics and Long-Run Implications,” *SSRN 2432313*.

ELSBY, M. W., B. HOBIJN, AND A. ŞAHIN (2013): “The Decline of the US Labor Share,” *Brookings Papers on Economic Activity*, 2013(2), 1–63.

FERNALD, J. G. (2014): “Productivity and Potential Output Before, During, and After the Great Recession,” *NBER Working Paper*, (w20248).

FOOTE, C. L., AND R. W. RYAN (2015): "Labor-Market Polarization over the Business Cycle," *NBER Macroeconomics Annual*, 29(1), 371 – 413.

GAGGL, P., AND S. KAUFMANN (2014): "The Cyclical Component of Labor Market Polarization and Jobless Recoveries in the US," Working Papers 14.03, Swiss National Bank, Study Center Gerzensee.

GOLLIN, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110(2), 458–474.

GOOS, M., A. MANNING, AND A. SALOMONS (2014): "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 104(8), 2509–2526.

GORDON, R. J. (2010): "Revisiting U.S. Productivity Growth Over the Past Century with A View of the Future," *NBER Working Paper*.

GUL, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, pp. 667–686.

HANSEN, G. D., AND E. C. PRESCOTT (2005): "Capacity Constraints, Asymmetries, and the Business Cycle," *Review of Economic Dynamics*, 8(4), 850–865.

HAWKINS, W., R. MICHAELS, AND J. OH (2013): "The Joint Dynamics of Capital and Employment at the Plant Level," 2013 Meeting Papers 1189, Society for Economic Dynamics.

JAIMOVICH, N., AND H. SIU (2012): "The Trend is the Cycle: Job Polarization and Jobless Recoveries," *NBER Working Paper*, (w18334).

JUDD, K. L., L. MALIAR, S. MALIAR, AND R. VALERO (2014): "Smolyak Method for Solving Dynamic Economic Models: Lagrange Interpolation, Anisotropic Grid and Adaptive Domain," *Journal of Economic Dynamics and Control*, 44, 92–123.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, pp. 263–291.

KARABARBOUNIS, L., AND B. NEIMAN (2014): "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 129(1), 61–103.

KING, R. G., AND S. T. REBELO (1999): "Resuscitating Real Business Cycles," *Handbook of Macroeconomics*, 1, 927–1007.

KOCHERLAKOTA, N. (2000): "Creating Business Cycles Through Credit Constraints," *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(3), 2–10.

KŐSZEGI, B., AND M. RABIN (2006): "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, pp. 1133–1165.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68(5), 1029–1053.

LEHN, C. V. (2015): "Labor Market Polarization, the Decline of Routine Work, and Technological Change: A Quantitative Evaluation," Discussion paper, Society for Economic Dynamics.

LETTAU, M., AND H. UHLIG (2000): "Can Habit Formation Be Reconciled with Business Cycle Facts?," *Review of Economic Dynamics*, 3(1), 79–99.

MALIAR, L., S. MALIAR, J. TAYLOR, AND I. TSENER (2015): "A Tractable Framework for Analyzing a Class of Nonstationary Markov Models," Discussion paper.

MCKAY, A., AND R. REIS (2008): "The Brevity and Violence of Contractions and Expansions," *Journal of Monetary Economics*, 55(4), 738–751.

MICHAELS, G., A. NATRAJ, AND J. VAN REENEN (2014): "Has ICT Polarized Skill Demand? Evidence from Eleven Countries Over Twenty-Five Years," *Review of Economics and Statistics*, 96(1), 60–77.

MORIN, M. (2014): "Technology Adoption and the Evolution of the Labor Market," Discussion paper.

MORTENSEN, D. T., AND C. A. PISSARIDES (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *The review of economic studies*, 61(3), 397–415.

ROGNLIE, M. (2015): "Deciphering the Fall and Rise in the Net Capital Share," in *Brookings Papers on Economic Activity, Conference Draft, March*, pp. 19–20.

ROSENBLATT-WISCH, R. (2008): "Loss Aversion in Aggregate Macroeconomic Time Series," *European Economic Review*, 52(7), 1140–1159.

SANTORO, E., I. PETRELLA, D. PFAJFAR, AND E. GAFFEO (2014): "Loss Aversion and the Asymmetric Transmission of Monetary Policy," *Journal of Monetary Economics*, 68, 19–36.

SHEA, J. (1995a): "Myopia, Liquidity Constraints, and Aggregate Consumption: A Simple Test," *Journal of Money, Credit and Banking*, pp. 798–805.

——— (1995b): "Union Contracts and the Life-Cycle/Permanent-Income Hypothesis," *American Economic Review*, pp. 186–200.

SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

TEVLIN, S., AND K. WHELAN (2003): "Explaining the Investment Boom of the 1990s," *Journal of Money, Credit and Banking*, 35(1), 1–22.

THALER, R. H., AND E. J. JOHNSON (1990): "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice," *Management Science*, 36(6), 643–660.

TOVAR, P. (2009): "The Effects of Loss Aversion on Trade Policy: Theory and Evidence," *Journal of International Economics*, 78(1), 154 – 167.

TVERSKY, A., AND D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5(4), 297–323.

VAN NIEUWERBURGH, S., AND L. VELDKAMP (2006): "Learning Asymmetries in Real Business Cycles," *Journal of Monetary Economics*, 53(4), 753–772.