Three Essays on Misintermediation

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree
Doctor of Philosophy in the Graduate School of The Ohio State
University

By

Feng Guo, B.S., M.A.
Graduate Program in Department of Economics

The Ohio State University

2012

Dissertation Committee:

J. Huston McCulloch, Advisor
Paul Evans
Pok-sang Lam
Abstract

This dissertation investigates the role of misintermediation in macroeconomic fluctuations. A perennial problem of financial markets is that of maturity mismatching, or misintermediation, a situation in which financial intermediaries fund long-term, illiquid loans with short-term liabilities. McCulloch (1981) concludes that misintermediation can be responsible for business cycles and predicts pro-cyclic behavior of surprises in real interest rates over business cycles. In my dissertation, I study the issue from both theoretical and empirical perspectives. An interpolation method for yield curve fitting is also specified to enable the empirical study of misintermediation.

My dissertation consists of four chapters. In Chapter 1 and 2 I theoretically investigate the mechanism for generating a disequilibrium boom or recession in a finite horizon structural model. In each period, agents decide how much to consume and how much to invest in heterogeneous capital for subsequent periods. If demand and supply happen to coincide in each period, the model will have a unique equilibrium term structure of interest rates. Otherwise, unexpected changes in real interest rates will occur accompanied by the realization of a recession or boom because previous plans cannot be completely corrected as a new period starts. The model is then extended to study the changes of real wage rates and heterogeneous capital prices under either a recession or boom regime. It reveals that misintermediation would not only bring about unexpected
output fluctuations and surprises in interest rates, but also give rise to unanticipated changes in factor prices.

An empirical examination of the misintermediation hypothesis relies on an empirical estimation of the term structure of interest rates. Chapter 3 hence specifies a multiple exponential decay model to fit both U.S. real and nominal term structures following the approach of interpolation. Several estimation methods, including unconstrained/constrained minimization, quadratic programming and iterative least squares, are introduced to estimate the parameters in the objective function according to different curve-fitting purposes. As a comparison, this chapter also proposes a semi-natural cubic spline model to fit the same data set. The results show that the multiple exponential decay model not only gives a parsimonious functional form, but also smoothes through idiosyncratic variations associated with the forward rate curve. In addition, selection of the number of terms/coefficients in an interpolation function governs the overall goodness of fit, which is optimized by three separate statistical tools.

In Chapter 4, an empirical study examines the relationship between unanticipated changes in real interest rates and unexpected fluctuations in real output over time. The former is derived from the U.S. real term structures estimated by the multiple exponential decay interpolation. Specifically, a monthly series of synthetic real consol prices is constructed as a proxy for the price of all future output, changes in which suggest changes in real interest rates over time. To proxy unexpected output fluctuations, I use either a time series of innovations to real GDP or a series of innovations to factor utilization. Statistical results show a negative correlation between the consol price series and the factor-utilization-based series as well as the real GDP-based series. This empirical evidence is consistent with the misintermediation hypothesis.
TO MY BELOVED PARENTS,

… who have been always beside me and backed me up …
Acknowledgments

Five years ago, I was led to the road to be an economist. I am still walking on it right now with this dissertation to be presented. This long journey is wonderful mainly because I have never been alone. I am not sure if there is enough space to acknowledge everyone who has been beside and supported me throughout this journey, but I want to first appreciate my advisor Dr J. Huston McCulloch. This dissertation would not be complete without Hu’s pivotal inputs. He devoted a considerable amount of time and efforts not only to inspire me for research ideas and model construction, but also to give me guidance on a great many details, such as the framework of a paper and professional features of slides. Hu is more than a mentor to me over the course of my naïve pursuit of academic research.

I am also deeply indebted to my dissertation committee members: Dr Paul Evans and Dr Pok-sang Lam. They attended all my presentations, listened to my reports on each research progress, and gave me insightful advices and comments during the last several years. They also invited me to act as an anonymous referee for the Journal of Money, Credit, and Banking, through which I was trained to be a researcher. In addition, many thanks should be spoken out to Dr Stephen McCafferty and Dr Bill Dupor for their attendance to my seminar talks and valuable suggestions. Dr Hajime Miyazaki, the director of the graduate program in economics, petitioned for my tuition coverage in the first year, without which I did not even have a chance to start this journey. I would also
like to appreciate China Scholarship Council (CSC) for their generous financial support from my first day of Ph.D. study.

This dissertation owes a great deal to the intellectual support from my graduate fellow colleagues. Particular thank to Fei Jin for his generous help and inputs, to Xiaoyu Xie for his support in financial market data, to Jaya Dey and Michael Sinkey for sharing ideas and giving comments on my presentations, to Chih-sheng Hsieh, Mithuna Srinivasan, and Saif Mehkari for their helps in my research. Besides, I want to say thanks to my officemates in this last year of graduate study: Zheng Jiang and Huimin Shi. Without their company and encouragement, I can hardly imagine how a year, overwhelmed by teaching, researching, job seeking, and dissertation compiling, would be like.

Last but not the least, I am grateful to my beloved families and friends. My father Xiao Guo and my mother Yingying Zhang not only raised me up, but also backed me up mentally along my journey of academic pursuit. I want to say thanks to my friends Yan Jiang, Shaojuan Liao, Doug & Karen Brown, Feng Zhou, Zhiquan Weng, and the whole economics Ph.D. class. The completion of this dissertation and my doctoral graduation will only be a happy event with their share of my happiness.
Vita

November 10, 1981 ......................... Born - Kunming, China

June 2004 ................................. B.S. Finance, Chongqing University, China
June 2007 ................................. M.A. Economics, Xiamen University, China
August 2008 .............................. M.A. Economics, The Ohio State University, USA
September 2008-present ................. Graduate Teaching Associate, The Ohio State University, USA

Fields of Study

Major Field: Economics
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Heterogeneous Capital and Misintermediation</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Benchmark Heterogeneous Capital Model</td>
<td>10</td>
</tr>
<tr>
<td>2.2.1 Basic Setting</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2 Production Possibility Frontiers</td>
<td>14</td>
</tr>
<tr>
<td>2.2.3 Fisherian Equilibrium and Misintermediation</td>
<td>16</td>
</tr>
<tr>
<td>2.3 An Extension of the Benchmark Model with Labor</td>
<td>27</td>
</tr>
<tr>
<td>2.3.1 Model with Labor and Harvesting Technology</td>
<td>29</td>
</tr>
<tr>
<td>2.3.2 Wage Rate and Capital Price</td>
<td>33</td>
</tr>
<tr>
<td>2.4 Conclusion</td>
<td>42</td>
</tr>
<tr>
<td>3. Estimating the U.S. Term Structure of Interest Rates</td>
<td>44</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>44</td>
</tr>
<tr>
<td>3.2 Identities and Definitions</td>
<td>49</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Descriptive Statistics for Target Time Series</td>
<td>91</td>
</tr>
<tr>
<td>4.2 Linear Regressions for $\Delta p(\Delta t)$ and Three Proxy Series</td>
<td>92</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Capital-specific PPFs for $p_2 = 0$ and $p_2 = p_{\text{max}}$</td>
<td>15</td>
</tr>
<tr>
<td>2.2 Combined PPFs for 4 given $p, p \in [0, p_{\text{max}}]$ and a Complete PPF</td>
<td>17</td>
</tr>
<tr>
<td>2.3 Ex-post PPFs vs Ex-ante PPF</td>
<td>18</td>
</tr>
<tr>
<td>2.4 Term Structure of Real Interest Rates in Period 1</td>
<td>19</td>
</tr>
<tr>
<td>2.5 Optimal Production Plan in Period 1</td>
<td>19</td>
</tr>
<tr>
<td>2.6 Optimal Consumption Plan in Period 1</td>
<td>20</td>
</tr>
<tr>
<td>2.7 Fisherian Term Structure</td>
<td>21</td>
</tr>
<tr>
<td>2.8 Three possible $D_0$ and an equilibrium outcome</td>
<td>23</td>
</tr>
<tr>
<td>2.9 Impending boom and unanticipated rise in real interest rate</td>
<td>24</td>
</tr>
<tr>
<td>2.10 Impending recession and unanticipated fall in real interest rate</td>
<td>25</td>
</tr>
<tr>
<td>2.11 Combined ex-post PPFs for 4 given $p_2 \in (0, p_{\text{max}})$ and an ex-ante PPF</td>
<td>32</td>
</tr>
<tr>
<td>2.12 Ex-post and Ex-ante PPFs for $c_1 = 15$</td>
<td>33</td>
</tr>
<tr>
<td>2.13 Consumption Ratio vs Marginal Rate of Transformation</td>
<td>37</td>
</tr>
<tr>
<td>2.14 Real Wage Rate in Period 2</td>
<td>38</td>
</tr>
<tr>
<td>2.15 Labor Demanded in Harvesting in Period 2</td>
<td>39</td>
</tr>
</tbody>
</table>
2.16 Capital Price of Plows in Period 2 ............................................. 40
2.17 Capital Price of Fields in Period 2 ............................................. 41
2.18 Stock Market Value in Period 2 ................................................. 42
3.1 U.S. Term Structure of Real & Nominal Interest Rates on 07/30/2010 .... 52
3.2 Discount Function v.s. log Discount Function ................................. 54
3.3 E-fold Life ................................................................. 57
3.4 Hump-shapes Generated by Extended Nelson and Siegel .................... 61
3.5 Fitted Negative log Discount Function on 02/26/2010 ...................... 66
3.6 Real U.S. Term Structure on 02/26/2010 by Three Estimation Methods .... 67
3.7 D-W Test for $k$ in Estimating Real U.S. Term Structure on 02/26/2010 . . 69
3.8 Nonparametric Runs Test for $k$ in Estimating Real U.S. Term Structure on 02/26/2010 ............................................... 71
3.9 BIC for $k$ in Estimating Real U.S. Term Structure on 02/26/2010 .......... 72
3.11 Real U.S. Term Structure on 02/26/2010 Estimated by Multiple Exponential Decay Model under Different $k$ ................................. 74
3.12 Unconstrained v.s. Constrained Estimation ................................... 76
4.1 Shift in Real Discount Functions .................................................. 84
4.2 Evolution of TIPS Market from 02/28/1997 to 12/31/2010 ................. 88
4.3 Amortized Loan Prices $P_t^a(m)$ derived by Constant and Time-varying $m$ . . 89
4.4 Monthly $\Delta p(\Delta t)$ over 04/1998–12/2010 ..................................... 90
Chapter 1: Introduction

The recent global financial crisis and economic downturn have raised questions about the stability of financial systems. This is particularly true for the United States where its markets for mortgages, mortgage-backed securities (MBSs), and credit default swaps (CDSs) are widely perceived to be at the root of the crisis (e.g., Gorton (2008, 2009b,a)). Although government sponsored entities (GSEs) and broker-dealer intermediaries (e.g., investment banks) surpassed deposit-taking intermediaries as holders of mortgages in the 1990s (Adrian and Shin 2009), the U.S. commercial banking system still accounted for the bulk of the origination and was left drastically weakened by the crisis. To some extent, the recent turmoil could be thought of as being partially caused by lenders, expecting a continued environment of low interest rates following the implosion of the “tech bubble” in early 2000 and the subsequent collapse of corporate investment, choosing to take on more illiquid financial assets financed with short-term debt.

The situation in which a financial institution holds generally longer maturity (or terms) assets relative to liabilities has been referred to as maturity mismatching. It was a big factor in Savings and Loan (S&Ls) crisis back in 1980s, during which S&Ls were allowed to borrow short-term debts from the public at low passbook interest rates and lend long, making 30 year fixed rate mortgages at a higher rate. A more recent example was Fed’s “Operation Twist”. This program allows Federal Reserve to sell billions of
shorter-term Treasury securities and use the proceeds to buy longer-term Treasury securities. By extending the average maturity of the securities in the Federal Reserve's portfolio, this action may put downward pressure on longer-term interest rates and drive up the short-term interest rates. The consequence is a flatter term structure (than the equilibrium term structure) that Fed manipulates.

Economists have long recognized maturity mismatching as a source of fragility in modern financial systems. [Diamond and Dybvig (1983)] first provided a theoretical explanation in their seminal paper to formalize the possibility that maturity mismatching causes bank runs, and further real economic damages. [Green and Lin (2000)] refined the Diamond-Dybvig model by extending their contractual environment and hence providing policy-makers with a better understanding of financial intermediation. [Young et al. (2010)] in turn give evidence that maturity mismatching became greater starting in the late 1990s.

Maturity mismatching could also serve as a potential interpretation to macro fluctuations. In regards to the recent crisis, economists have widely perceived the role of maturity mismatching in causing the turmoil. [Diamond and Rajan (2009)] note that the “consensus on the proximate causes of the crisis [includes] misallocated resources to real estate, financed through the issuance of exotic new financial instruments [which] were largely financed with short-term debt.” [Farhi and Tirole (2009)] also suggest that the current crisis “is one of wide-scale maturity mismatch.”

However, the Diamond-Dybvig model as well as the subsequent research in the same line are overwhelmingly focused on strategic behavior of different agents and their consequences at micro level, while failing to disentangle the role of maturity mismatching in business fluctuations at macro level. [McCulloch (1981)] filled this gap in one of his
early papers. He attributes the unanticipated changes in interest rate and aggregate excess demand/supply to the break of balanced intermediation, or *misintermediation*. Specifically, a Fisherian model (Fisher [1930]) suggests that an equilibrium term structure of real interest rates is maintained in a world where planned supply of aggregate output would match planned demand of consumption period by period over time. As time moves forward, barring shocks to both production and taste, expectations on future interest rates will be hence perfectly realized, in the sense that forward interest rates will just equal to the future spot interest rates\(^1\). However, since the real world is running based on highly regulated financial institutions and markets, the rise of misintermediation is likely to break down the balance between current plans for future demand and supply. Put differently, even though the present discounted value of expected sum of demand must equal its counterpart in supply, it does not guarantee that planned supply will meet demand period by period throughout into future in a world where government intervention, technological shocks, or “dynamic inconsistency” in tastes exist. If that is the case, then unanticipated falls or rises in real interest rates will be observed along with macroeconomic fluctuations.

Unfortunately, there was neither follow-up research to examine the misintermediation hypothesis empirically\(^2\) nor theoretical endeavor to specify explicit production technologies to investigate behavior of factor prices under misintermediation in a Fisherian model. This becomes a primary motivation for this dissertation research. Such a

\(^1\)In practice, a term premium, either in constant or time-varying style, should be adjusted for forward rates to predict future spot rates.

\(^2\)Until recently, there was no direct data on the real term structure.
research would be an interesting pursuit in that it does not only provide empirical ev-
idence to test a long studied theory, but also give new insights to interpret 1980s’ S&Ls
crisis as well as the recent economic downturn.

In order to fully illustrate misintermediation, I develop a structural model in which
households make plans for intertemporal consumption and production in Chapter 2. The
model assumes a world where preferences are dynamically consistent and without
production uncertainty. The household is able to use heterogeneous intermediate cap-
ital goods to produce output in each period. Different capital goods enable the house-
hold to access different production technologies. Some heterogeneous capital goods
may be produced by the other heterogeneous capital goods as inputs. The household
then faces a tradeoff in allocating capital resources between final output production
or intermediate capital goods production, which in turn brings higher returns on final
outputs.

Chapter 2 then introduces labor and a harvesting-investment decision for house-
holds in an extended model economy. In the benchmark model aforementioned, het-
erogeneous capital is the only class of factor inputs for both capital goods production
(at time \( t - 1 \)) and consumption goods production (at time \( t \)). The introduction of labor,
together with a harvesting technology, then enriches the model by allowing the implica-
tion for the factor prices to be studied. This setup deviates from that of a standard neo-
classical production model, which requires labor only for current consumption goods
production, and has nothing to do with future outputs. It helps explore the implica-
tions of misintermediation for wages rates, capital prices, and of planned employment.
These implications suppose to answer a series of questions. Does a recession lead to
an unexpected rise or fall in the market clearing real wage and/or in market clearing
employment? What happens to the price of capital goods? Do the answers vary with parameter values?

The next important question is whether we really have this pro-cyclic behavior of real interest rate surprises in the U.S. over the last decades. In order to answer this question, I investigate the relationship between unanticipated changes in U.S. real interest rates and excess demand for aggregate output in the following chapters. Since Treasury Inflation-Protected Securities (TIPS) has been issued and traded for more than a decade, this enables us to collect more than 13 years' market data for risk-free real interest rates and to test whether the historical evidence is in accord with the theoretical prediction. This empirical study is a major contribution of this dissertation in that U.S. real interest rates time series data was not available thirty years ago when the misintermediation hypothesis was first put forward by McCulloch [1981].

The first step of the empirical study is to obtain an empirical estimate of U.S. term structure of interest rates. In Chapter 3, I follow the interpolation approach and specify a multiple exponential decay model to fit U.S. term structure of real interest rates. Results show that the multiple exponential decay model not only gives a parsimonious functional form, but also smoothes through idiosyncratic variations associated with the forward rate curve. Several estimation methods are introduced to estimate the parameters in the objective function for different curve-fitting purposes, which broaden the spectrum of the interpolation model’s application.

Given the estimated U.S. term structure of real interest rates, Chapter 4 starts with the construction of a monthly series of synthetic real consol prices, a proxy for the price
of all future output, changes in which over a period of time suggest changes in real interest rates. In a world without distortions, the expectation of that change has to be zero. So any actual changes reflect unanticipated changes in real interest rates.

The other side of the study concerns with the excess supply/demand of aggregate output, imaged by unexpected fluctuations in real output. It is similar to “output gap”, a concept the theory of optimal monetary policy rules tries to measure. In order to obtain this series, a natural idea would find an innovation series to real output by applying auto-regression to real GDP series. However, there are two problems associated with this approach. First, no monthly real GDP series is available in the U.S., which makes the constructed monthly consol price series unable to match up with the output series. Although there are ways to artificially transform monthly series into quarterly or annual series, this manipulation may cause new problems, such as serial correlation stemming from data overlapping. Second, the U.S. real GDP time series appears to be a locally integrated process with up-drift. Since there is no verifiable long-term trend in a unit root process, the innovation to the output gap estimated through GDP time series may not be meaningful. An effective proxy would be Factor Utilization Gap series proposed in a dissertation research by Longbrake (2008), who demonstrates that the new series is free from non-stationary concerns and positively correlated with real GDP fluctuations. In this dissertation, I follow Longbrake (2008) to construct a series of innovation to factor utilization featuring unexpected fluctuations in real output.

A simple regression analysis demonstrates a significantly negative correlation between the consol price series and factor utilization series. This implies that the series of surprises in real interest rate will be positively correlated with unexpected changes in
real output. The result can be interpreted as empirical evidence consistent with misin-
termediation hypothesis. Nevertheless, as I substitute other real output related series
for the innovation to factor utilization, results become statistically insignificant. The
failure further implies the necessity of a thorough study on other related issues, such as
the adjustment of a term premium in the consol price series.
Chapter 2: Heterogeneous Capital and Misintermediation

2.1 Introduction

This chapter theoretically investigates the mechanism of misintermediation in generating output fluctuations and unexpected factor price changes. Generally, individual agents make production plans to accommodate consumption demand for different points of time in future by following the guide of intertemporal prices, as implied by the term structure of real interest rates. Unfortunately, banks and other financial intermediaries often fund long-term, illiquid loans with short-term liabilities, both of which are unlikely to be perfectly matched up in terms of maturities. If a planned output exceeds the actual demand in future, a recession will set in. When this occurs, real interest rates have to fall to clear the market, and the economy will go through unexpected fluctuations. On the contrary, a boom in future will realize once a planned output fails to meet the actual demand, making real interest rates raise. This also leads to unexpected macro fluctuations.

In a standard neo-classical model, capital is homogeneous over time. This convention does not help generate vanishing intertemporal production possibilities, a problem having long been emphasized by the Austrian School. In this chapter, I introduce

\[\text{Footnote}\]

This could be caused by government subsidies, regulations and other interventions. See McCulloch (1981) for a summary and Gorton (2008) for a thorough analysis on the subprime crisis.
heterogeneous capital to address this problem by proposing a finite horizon structural model with the production technologies based on heterogeneous capital.

For simplicity, I first assume two heterogeneous capital goods and hence two technologies, i.e. a high technology, which gives higher productivity and return, and a low technology, which is the benchmark technology. In order to take advantage of the high technology, however, the household needs to produce the corresponding high technology capital one period ahead by using the low technology capital as an input. The household hence faces a tradeoff in allocating the low technology capital goods between final goods production using low technology capital directly and production of high technology capital goods over time. Its decision is based on the intertemporal prices of output, implied by the term structure of real interest rates, which defines both spot and forward rates. If a prevailing term structure of interest rates happens to clear the excess demand/supply in all periods, the economy would stay at an equilibrium. Otherwise, an unexpected fall or rise in future spot rates has to occur in order to eliminate the excess demand/supply in periods when the ex-ante term structure of real interest rates miscoordinates the planned consumption and production. As a result, the heterogeneous capital model predicts a pro-cyclic behavior of real interest rate surprises. In other words, an unanticipated fall in real interest rates is associated with a recession, while an unanticipated rise in real interest rates and a boom happen together.

The heterogeneous model is then enriched by adding labor and a harvesting technology to investigate the impact of macro fluctuations on factor markets in a mistermediated world. The extension is another contribution of this dissertation in that it makes the model different from that of a standard neo-classical one where labor contributes to current output only. In the extended model, labor in each period is allocated to harvest
(for current output), to plant, and to build heterogeneous intermediate capital goods (the latter two are for future outputs). This setting helps explore the implications of misintermediation for wages rates, capital prices, and of planned employment.

The plan for this chapter is as follows. In the next section, I present the heterogeneous capital model without labor and study the consequences of misintermediation. Section 2.3 extends the benchmark model by introducing labor and a harvesting technology. The focus of this section is in turn on the changes in real wage rates and capital prices according to different phases of an economy. Section 2.4 summarizes and concludes.

### 2.2 Benchmark Heterogeneous Capital Model

#### 2.2.1 Basic Setting

The structural model I develop here extends that of McCulloch (1981). Households make plans for intertemporal consumption and production. There is only one good consumed and produced in each period. Claims on this output at future dates may be discounted and traded for current output. All production takes place within the household, so that this borrowing and lending of the single consumption good is the only trade that takes place. The prices at which households trade claims on future output determine a term structure of real interest rates, which is public knowledge.

To be more specific, I assume that the consumption good is grain. Grain can be consumed directly or used as seeds to plant for harvesting fields in subsequent periods. The planting technologies might differ in that the household may or may not use tools for planting. The household can choose a low planting technology with seeds as the only factor input; or the household can use tools to realize a high planting technology.
in order to obtain a higher productivity, which gives a higher yield of grain outputs in the following periods. However, the tools are not exogenously endowed. The household has to use some grain seeds to produce tools in advance. Therefore, the capital inputs in this model economy are heterogeneous in the sense that there are two kinds of capital i.e. seeds and tools, and two heterogeneous production technologies i.e. low production technology and high production technology. The household thus faces a tradeoff in distributing the harvested grain each period among consumption, seeds for planting, and seeds for tool production\(^4\). In order to keep the model simple enough to illustrate the contraction of production possibility frontiers (PPFs) and its consequent implication for misintermediation, I do not introduce labor in this step, which will be considered in an extended (and hence more realistic) model in the next section.

On a real farm, a household starts with an endowment of grain in Period 1, the very beginning. After reserving some for current consumption, the household can either plant all leftover grain seeds for harvest next period or use a portion to build a plow. Having produced the plow in Period 1, the household would then be able to take advantage of the high planting technology in Period 2 and enjoy a greater harvest return in Period 3. Also in Period 2, this household may choose to give up some harvesting grain to feed oxen so that in Period 3 it will have an even higher productivity yielding higher grain output in Period 4. Having had oxen, the household may think about building a tractor or a robot tractor… This process may proceed to any period in future so that heterogeneous capital goods and thus heterogeneous planting technologies continuously arise on this farm. In contrast to the standard neoclassical model, there is always infinite menu technologies, but it keeps taking longer and longer to put in line. No household

\(^4\)In the simplest 3-period model, tool production does not take place in each period, but only in Period 1.
is able to jump to the most productive technology because the household has to eat in the meanwhile\footnote{Households give up perpetually higher consumption in order to eat simultaneously. So there is always a tradeoff of how much a household consumes right now versus the future.}

In the simplest heterogeneous capital model in this chapter, I assume that the heterogeneous capital goods households can produce/use are plows and grain seeds. In each period with harvesting or endowment grain, households can always have access to a low planting technology, which is just a linear production function with grain seeds $s_t$ as the only capital input:

\[
 f^L(s_t) = \alpha_1 s_t. \tag{2.1}
\]

A household may be also able to access a high planting technology given that it built plows a period ahead. The high planting technology is given by a constant-return-to-scale Cobb-Douglas production function:

\[
 f^H(s_t, p_t) = \alpha_2 s_t^\gamma p_t^{1-\gamma}, \tag{2.2}
\]

in which plows ($p_t$) are built solely by grain seeds. The production technology for $p_t$ is given by:

\[
 p_t = \nu s_{t-1}^\theta \tag{2.3}
\]

where $s_t^\theta$ denotes seeds inputs for plow production.

In this sense, the time $t+1$ harvested crops $h_{t+1}$ is determined by the time $t$ choice of technologies between $f^L$ and $f^H$, and further by the time $t-1$ production of plows $p_t$ as follows:

\[
 h_{t+1} = f^L(s_{t}^L) + f^H(s_{t}^H, p_t). \tag{2.4}
\]

The household in each period divides harvest $h_t$ into consumption $c_t$, seeds for low technology $s_t^L$, seeds for high technology $s_t^H$, and seeds for plow production $s_t^\theta$. The
The resource constraint for the household in time \( t \) is hence given by:

\[
h_t = c_t + s^L_t + s^H_t + s^P_t.
\] (2.5)

In addition, I assume no uncertainty in production in this simplest benchmark model.

I further assume that the households have a 3-period life horizon (Period 1 through 3). A 3-period horizon enables the simplest term structure of interest rates, determining a short-term spot rate, a long-term spot rate, and an implied forward rate. It is the least time horizon in which interest rate surprises can occur. In Period 1, the household enters the economy with an endowment of grain, \( h_1 \), which is distributed according to

\[
h_1 = c_1 + s_1 + s^P_1. \tag{2.5}
\]

The household uses \( s^P_1 \) to build plows, \( p_2 = \nu s^P_1 \theta \), and \( s_1 \) to produce Period 2 harvest, \( h_2 = \alpha_1 s_1 \). This low planting technology is the only planting technology available in Period 1 since the household does not have any endowment plows at beginning. In Period 2, the household receives harvest \( h_2 \), and divides it into consumption \( c_2 \) and planting seeds \( s_2 \). The household does not need to reserve any seeds to produce plows for Period 3 because Period 3 is the last period in which no planting production will take place. Given \( p_2 \) in Period 1, the household in Period 2 has two planting technologies ready to use. The harvest \( h_3 \) is thus taken the following form:

\[
h_3 = \alpha_2 s^H_2 \gamma p_2^{1-\gamma} + \alpha_1 s^L_2
\]

\[
= \alpha_2 s^H_2 \gamma (\nu s^P_1 \theta)^{1-\gamma} + \alpha_1 (s_2 - s^H_2)
\] (2.6)
where $s_2^H$ and $s_2^L$ are portions of $s_2$ reserved for high and low planting technologies respectively. In Period 3, the household consumes the harvested crops ($h_3 = c_3$) and exists the economy.

### 2.2.2 Production Possibility Frontiers

The basic model setting in heterogeneous production technologies gives rise to capital-specific Production Possibility Frontiers (PPFs). In the 3-period heterogeneous capital model, the amount of plows $p_2$ is pre-determined in Period 1. For any given $p_2$ in Period 2, the household then faces a PPF specific to these capital plows. On one extreme, the household may decide to produce zero plows in Period 1 so that it can only access to low planting technology in Period 2. The corresponding PPF, as showed in the left graph of Figure 2.1, implies a pure linear transformation of consumption grain in all three periods, i.e. $c_1$, $c_2$, and $c_3$, due to the linear nature of Eq 2.1. The associated intercepts of the $c_1$, $c_2$, and $c_3$ axes turn out to be $h_1$, $\alpha_1 h_1$, and $\alpha_2^2 h_1$ respectively, which represent the maximum level of consumption the household is able to reach in each period under low planting technology.

However, the household may produce some plows in the expense of a reduction in consumption and planting seeds in Period 1. This decision lets the household enjoy a higher productivity, given by Eq 2.6 and hence a expanded $c_3$ frontier; however, it

---

6The household switches between two technologies according to their respective marginal product of seeds. The actual harvesting function is as follows:

$$h_3 = \begin{cases} \alpha_2 s_2^\gamma \left(\nu s_1^p \theta\right)^{1-\gamma} & s_2 \in [0, s_2^{BE}] \\ \alpha_2 s_2^{BE} \gamma \left(\nu s_1^p \theta\right)^{1-\gamma} + \alpha_1 \left(s_2 - s_2^{BE}\right) & s_2 \in [s_2^{BE}, \alpha_1 h_1] \end{cases}$$

where $s_2^{BE}$ is a break-even value of $s_2$, below which the high planting technology yields a higher marginal product of seeds and the household chooses to use this high technology only.

7The maximum levels of $c_1$, $c_2$ here are globally maximum for any numbers of plows, $p_2$. 

---
results in contracted $c_1$ and $c_2$ opportunities. As $p_2$ increases, the intercept of $c_3$ will monotonically expand to a globally maximum point where the associated number of plows is $p_{max}$. Beyond that point, the $c_3$ intercept of the production frontier (together with $c_1$ and $c_2$ intercepts) starts to contract and eventually falls to 0 as $p_2$ further increases because the household uses up the entire endowment resource $h_1$ to produce plows. For $p_2 = p_{max}$, the capital-specific PPF is illustrated in the right graph of Figure 2.1.

Figure 2.1: Capital-specific PPFs for $p_2 = 0$ and $p_2 = p_{max}$

A set of combined capital-specific PPFs can be depicted in a single graph. Figure 2.2(a) shows such an example of a combination of four capital-specific PPFs in a same graph, each of which associates with a specific $p_2 \in [0, p_{max}]$. The graph demonstrates a contraction in $c_1$ and $c_2$ opportunities as well as an expansion in $c_3$ opportunities as $p$ increases. An envelope of a complete set of capital-specific PPFs, from $p_2 = 0$ to $p_{max}$,
gives a complete PPF as showed in Figure 2.2(b). This complete PPF is the PPF that a household faces in Period 1. In other words, any point on the complete PPF is attainable for the household in Period 1 before a specific number of plows is picked; however, once a decision on $p_2$ is made, the household actually faces a capital-specific PPF, which is a contraction of the initial complete PPF. The 2-dimensional Figure 2.3 illustrates this fact in a more vivid style. The projection of a 3-D complete PPF on a 2-D $(c_2, c_3)$ plane is a vertical "slice" of the 3-D PPFs along the $c_1$ axis. The complete PPF envelops all possible capital-specific PPFs and is tangent to each capital-specific PPF at one point only. Each capital-specific PPF hence turns out to be a contraction of the complete PPF. Since a contracted capital-specific PPF in Figure 2.3 is a PPF a household actually faces in Period 2 (when $p_2$ has already been pinned down), I call this PPF the *ex-post* PPF. The complete PPF is thus the *ex-ante* PPF because this production frontier can no longer be reached in Period 2 except for the tangency point corresponding to the specific choice of $p_2$.

### 2.2.3 Fisherian Equilibrium and Misintermediation

There is a term structure of real interest rates coordinating the supply/demand in the model economy. It implies the intertemporal prices at which consumption goods of different periods trade for each other. I first defines a discount factor $\delta_{t,t+m}$, which gives the present (time $t$) value of $1$ in $m$ periods ahead. The value $(-\log \delta_{t,t+m}/m)$ is thus equal to $r_{t,t+m}$, the real interest rate of a loan which starts at $t$ and matures at $t+m$. More generally, the real interest rate $r_{imm}$ is associated with the discount factor $\delta_{imm}$, at which $c_m$ trades for $c_n$ in period $i$ for $i=1,2,$ and $i \leq m < n$. The subscripts $i,m,$ and $n$ represent a $c_2$-$c_3$ PPF for a particular level of $c_1$. 

---

8It represents a $c_2$-$c_3$ PPF for a particular level of $c_1$. 

16
Figure 2.2: Combined PPFs for 4 given $p, p \in [0, p_{max}]$ and a Complete PPF

represent the time when the interest rate is effective, the time when the loan starts, and the time when the loan matures, respectively. It follows that:

$$\delta_{imn} = \exp(-r_{imn} (t_m - t_n)).$$

(2.7)

In the simplest 3-period economy, the term structure of real interest rates defines 3 real rates, i.e. $r_{112}, r_{113},$ and $r_{123}$. Specifically, $r_{112}$ and $r_{113}$ are the short-term and long-term real interest rate, respectively; $r_{123}$ is an implicit forward rate—the rate of return at which $c_2$ may be rolled over to $c_3$ in Period 1. The forward rate bridges the two explicit spot rates as follows:

$$(t_3 - t_1)r_{113} = (t_2 - t_1)r_{112} + (t_3 - t_2)r_{123}.$$  

(2.8)
Eq 2.8 implies that if a household has $1 worth of Period 1 consumption good $c_1$, then it is equivalent to have $\exp (r_{112} (t_2 - t_1))$ dollars of $c_2$, or $\exp (r_{113} (t_3 - t_1))$ dollars of $c_3$. In the $(c_1, c_2, c_3)$ space, the term structure of real interest rates can be illustrated as the triangular plane $\{r_{112}, r_{113}, r_{123}\}$ as in Figure 2.4. The 3 intercepts are 1, $\exp (r_{112} (t_2 - t_1))$, and $\exp (r_{113} (t_3 - t_1))$ correspondingly. In the graph, the steeper the slope of the triangular plane in absolute value, the higher the corresponding interest rate.

Under the pure expectation hypothesis, the expected forward rates equal to future spot rates, $E_1 r_{223} = r_{123}$. In practice, $E_1 r_{223}$ may be further adjusted by the amount of a term premium.

Given the PPF and the term structure of real interest rates in Period 1, which is public information, competitive profit maximization will lead factor owners to plan to produce the aggregate output supply vector $S_1$, which is the tangency point between the ex-ante
Figure 2.4: Term Structure of Real Interest Rates in Period 1

Figure 2.5: Optimal Production Plan in Period 1
PPF and the triangular plane determined by the given term structure. Figure 2.5 demonstrates the optimal production plan that households make accordingly.

![Figure 2.5: Optimal Production Plan](image)

The same term structure of real interest rates also determines an intertemporal budget constraint for households in Period 1. The households maximize discounted lifetime utility subject to this budget constraint. I assume quasi-concave preferences\(^9\) which are dynamically consistent. In contrast to Diamond-Dybvig model, there is no taste uncertainty. Figure 2.6 shows the optimal consumption plan \(D_1\) that households make accordingly in Period 1. \(D_1\) again is the tangency point between the term structure plane and the indifference surface.

Fisher (1930) argued that there is an equilibrium term structure of real interest rates in such an economy so that the structures of planned consumption and production

---

\(^9\)If the household has diminishing marginal utility, then its indifference curve will be quasi-concave.
would match up over time. If the prevailing term structure of real interest rates in the 3-period economy happens to be in accord with the equilibrium one, then the economy would develop without business fluctuations. As the economy moves forward in time, the planned production will be just equal to the planned consumption. Any miscoordination of the household’s efforts would not arise under the “correct” intertemporal prices.

However, the market does not necessarily find an equilibrium term structure in a misintermediated world. The Fisherian equilibrium shape of a term structure assumes that each agent in the economy contracts forward all her planned future borrowing
and lending plans. Unfortunately, these Arrow-Debreu contingent claims do not exist in reality. In an actual misintermediation world, financial intermediaries might entice savers into shorter-term deposits while encourage investors into longer-term loans, due to government intervention such as deposit insurance. This would mis-coordinate planned consumption and production and create discrepancy in expectation. The savings and loan crisis of the 1980s and 1990s may serve as an example: investors are building houses to provide housing services 20 years from now, while people don't want housing services right now; people took out loans that they did not think they had to pay back, while banks borrowing them thought they would be repaid. Fed’s “Operation Twist” may serve as another example. The maturity extension program manipulates both the long-term and short-term yields and leads to a flatten yield curve, which may deviate from the Fisherian equilibrium term structure.

Figure 2.7 demonstrates three possible outcomes a manipulated term structure of interest rates may lead to. In the 3-period model, intertemporal budget constraints ensure the sum of the excess demands to be zero, but not that they individually be zero. Since Period 1 excess demand has to be zero\(^{10}\) the \(c_1\) components of the Period 1 supply vector \(S_1\) and demand vector \(D_1\) will coincide, so that \(D_1\) must lie in the plane of points having the same \(c_1\) components as \(S_1\), indicated by the vertical cutting plane in Figure 2.7. It hence follows that the ex-ante PPF in the \((c_2, c_3)\) coordinates is \(P_1P_1\), depicted by the red curve in Figure 2.7. However, misintermediation breaks this balance in the subsequent periods so that the Period 2 and 3 consumption components \((c_2\) and \(c_3)\) in vector \(D_1\) might not necessarily coincide with their counterparts in supply vector \(S_1\), from which three possible locations of \(D_1\) relative to \(S_1\) arise.

\(^{10}\)It will be immediately apparent if demand is a different quantity of \(c_1\) than the quantity supplied during Period 1.
First, the term structure of interest rates in Period 1 might just by accident coincide with the equilibrium term structure, implying a Fisherian equilibrium term structure which lets $D_1 = D^e_1$, the same point as $S_1$ in Figure 2.7. As time moves forward to Period 2, there is no surprise in consumption and production and thus no unanticipated change in real interest rates, i.e. the future spot rate $r_{223}$ will be just equal to the current forward rate $r_{123}$. In Figure 2.8, we notice that the planned consumption and production will be perfectly realized if $D_1 = D^e_1$ even though the ex-post PPF $P_2P_2$ is just a subset of ex-ante PPF $P_1P_1$ (contracted back). This is because the production vector $S_1$ is always the tangency point, common to both ex-ante and ex-post PPFs.

Figure 2.8: Three possible $D_0$ and an equilibrium outcome

Now that we have no priori presumption about the given term structure, it might lead to the second possibility of $D_1$: a point $D^b_1$ to the southeast of $S_1$ as showed in Figure 2.7 and Figure 2.8(a). This possibility is associated with an excess demand, or a
boom, in Period 2. It predicts pro-cyclic behavior in real interest rates according to the proposition as follows.

![Diagram](image)

Figure 2.9: Impending boom and unanticipated rise in real interest rate

**Proposition 2.2.1.** *In Period 1, an impending boom in Period 2 is associated with an unexpected rise in real interest rates.*

**Proof:** In Period 1, a forward rate $r_{123}$ that is higher than its equilibrium value would lead to a demand vector $D_1$ that is to the southeast of the supply vector $S_1$, as the second possibility aforementioned, given the household’s quasi-concave preference and convex ex-ante PPF $P_1P_1$. According to Figure 2.9(a), $D_1$ has a higher $c_2$ component and a lower $c_3$ component than $S_1$. It hence follows that an impending boom in Period 2 (an impending recession in Period 3) will appear in the sense that an excess demand of Period 2 output (an excess supply of Period 3 output) in general becomes apparent in the economy. As time moves forward to Period 2, a value of $r_{223}$, necessarily higher than $r_{123}$,
will be found to digest the excess demand and balance the market again. As showed in Figure (b), Period 2 demand vector $D_2$ and supply vector $S_2$ coincide after an decline in spot rate $r_{223}$ from its previous forward rate $r_{123}$; however, this equilibrium point falls into to a lower indifference curve $U^{b}_2$ and a contracted ex-post PPF $P_2P_2$, which implies an ex-post welfare lost. The rise in $r_{223}$ is a deviation from the expectation hypothesis, so it is unexpected in Period 1. Therefore, an impending boom will be associated with an unanticipated rise in real interest rates. \textit{Q.E.D.}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure.png}
\caption{Impending recession and unanticipated fall in real interest rate}
\end{figure}

Symmetrically, $D_1$ can also be to the northwest of $S_1$, given by $D'_1$, the third possibility of $D_1$ as showed in Figure 2.7 and Figure 2.8-(a). This possibility is associated with an excess supply, or a recession, in Period 2. It also predicts pro-cyclic behavior in real interest rates according to the proposition as follows.

\textsuperscript{11}The interest rate must go up sufficiently high above $r_{123}$ to discourage the household from borrowing.
**Proposition 2.2.2.** In Period 1, an impending recession in Period 2 is associated with an unexpected fall in real interest rates.

**Proof:** In Period 1, a forward rate $r_{123}$ that is lower than its equilibrium value would result in a demand vector $D_1$ that is to the northwest of the supply vector $S_1$, as in the third possibility aforementioned, given the household’s quasi-concave preference and convex ex-ante PPF $P_1P_1$. According to Figure 2.10 (a), $D_1$ has a lower $c_2$ component and a higher $c_3$ component than $S_1$. It hence follows that an impending recession in Period 2 (an impending boom in Period 3) will appear in the sense that an excess supply of Period 2 output (an excess demand of Period 3 output) in general becomes apparent in the economy. As time moves forward to Period 2, a value of $r_{223}$, necessarily lower than $r_{123}$, will be found to eliminate the excess supply and balance the market again. As showed in Figure 2.10 (b), Period 2 demand vector $D_2$ and supply vector $S_2$ coincide after an increase in spot rate $r_{223}$ from its previous forward rate $r_{123}$; however, this equilibrium point falls into to a lower indifference curve $U_r^2$ and a contracted ex-post PPF $P_2P_2$, which implies an ex-post welfare lost. The fall in $r_{223}$ is a deviation from the expectation hypothesis, so it is unexpected in Period 1. Therefore, an impending recession will be associated with an unanticipated fall in real interest rates. *Q.E.D.*

Both the second and third possibilities imply that the heterogeneity of capital in the context of misintermediation bring about technological inefficiency and utility loss. Because the contracted ex-post PPF $P_2P_2$ has a lower transformation elasticity than the ex-ante PPF $P_1P_1$ in a neighborhood of $S_1$, Period 1 production activities appropriate to $S_1$ will be appropriate to that point and to no other point on $P_1P_1$ during Period 1. Once $P_2P_2$ is realized, it is too late to go back and change the production activities conducted.

\[12\] The interest rate must decline sufficiently low below $r_{123}$ to encourage the household to borrow.
during Period 1. An ex-post supply vector $S_2$ will thus be no longer on the frontier of $P_1P_1$ but lie inside this production possibility set in general. Accordingly, the ex-post demand vector $D_2$ will lie on a lower indifference curve, implying an ex-post utility loss. The problem of vanishing intertemporal production possibilities due to the inconvertibility of capital has long been emphasized by the Austrian School. From an Austrian point of view, there were not too many or too few capital goods produced during Period 1, but only the wrong mix. The so-called malinvestment, long emphasized by Austrian economists von Mises and Hayek (1941), leads to business fluctuations (either booms or recessions), which ordinarily entail technological inefficiency and utility loss in a world with financial misintermediation. Put differently, a boom can be just as bad as a recession.

2.3 An Extension of the Benchmark Model with Labor

Even though the above heterogeneous capital model successfully demonstrates the misintermediation hypothesis, the benchmark model is too simple to help investigate the impact of a macro disequilibrium on factor markets. We may want to understand how real wage rates and also prices of heterogeneous capital goods vary with the unexpected changes in real interest rates. This becomes a key motivation to extend the simple benchmark model in order to further study the effects of misintermediation at a macro level.

The benchmark heterogeneous capital model assumes two capital inputs only, grain seeds and plows, the latter of which is produced by the former. Labor does not participate in any production technology. In a more realistic model economy, it makes more
sense to use labor to build plows than seed. Generally, labor has nothing to do with future consumption goods in a neoclassical model, but is only used to produce current consumption goods, which keeps the price of mature fields from being identically 1 in terms of harvested goods. If harvesting is regarded as the production of current consumption goods in a field, while future goods production is planting, then labor should also be used to plant in order to have impact on factor prices. This setting deviates from the neoclassical framework where agents just throw out seeds in the ground—no actual planting activity takes place. On the other hand, a harvesting technology, which is not taken into account in the benchmark model, consists of one of the extra features in the extended model. In sum, labor enters into the extended model via three approaches: 1) labor for harvesting, which keeps the model neoclassical; 2) labor for planting, which generates varying marginal product of factors; and 3) labor for plow production, so as to fit the extended model into the basic setting of the heterogeneous capital framework. A household with 3-period life horizon therefore faces a tradeoff in allocating labor in Period 1: it can either put labor in harvesting for consumption goods in the current period (Period 1), or put labor in planting for consumption in the future period (Period 2), or put labor in building plows to produce consumption in even further future period (Period 3).

The introduction of labor, together with a harvesting technology, has an important implication to factor markets, i.e. real wage rates and capital prices. Since everything is done within households, there is no actual factor market in the hypothetical model economy. Nevertheless, a household has a shadow price for each factor input. For an instance, the real wage rate in Period 1 turns out to be the change in output divided by the change in labor input, which is the marginal product of Period 1 labor. Because the
outputs the household gets are not just $c_1$, but also $c_2$ and $c_3$, the value of this marginal product is the extra $c_1$ the household chooses to produce plus the present value of $c_2$ and $c_3$ it chooses to produce. In this sense, the harvesting-planting decision can affect the marginal product of labor in both activities. We can evaluate a wage rate in either activity. Similarly, the value of a plow in Period 2 is its marginal product in terms of future outputs, which has a present value determined by interest rates. It supposes that plows are going to be worth less if future output is worth less. This section hence focuses on the shadow wage of labor and shadow prices of heterogeneous capitals. I investigate what happens to them when the unexpected change in interest rates affects households’ plans.

2.3.1 Model with Labor and Harvesting Technology

The household is endowed with a constant amount of labor $L$ in each period. It can either use this labor endowment to harvest $(L^h_t)$, to plant fields $(L^f_t)$, or to build plows $(L^p_t)$ so that the labor resource constraint for the household in each period is:

$$L \geq L^h_t + L^f_t + L^p_t.$$  \hfill (2.9)

There are three production technologies associated with the extended model, namely the harvesting technology $f(\cdot)$, the planting technology $g(\cdot)$, and the plow building technology $p(\cdot)$. Comparing to the benchmark model, labor becomes a key ingredient in all three technologies and $f(\cdot)$ is a brand new feature to the model.

At the beginning of each period, the household contributes $L^h_t$ to harvest the mature fields $F_t$ following a constant-return-to-scale Cobb-Douglas production technology:

$$h_t = f(F_t, L^h_t) = \alpha F_t^\eta L^h_t^{1-\eta}.$$  \hfill (2.10)
where $\eta$ is the harvesting labor share. In Period 1, the very beginning, a predetermined endowment fields $F_1 = \bar{F}_1$ is available for harvesting so that the harvesting function at $t = 1$ is simply

$$h_1 = \alpha \bar{F}_1^\eta L_1^{1-\eta}.$$  

The harvesting crops $h_t$ is then divided up into consumption $c_t$ and planting seeds $s_t$ in each period. The resource constraint as follows always holds.

$$h_t = c_t + s_t. \tag{2.11}$$

Seeds $s_t$ are then used to plant (or “produce”) fields $F_{t+1}$, jointly with heterogeneous capital plows $p_t$. I assume that $p_t$, augmenting seeds with a share $\gamma$, enters into another constant-return-to-scale Cobb-Douglas production function as follows. Labor for fields planting $L^f_t$ takes a share $(1 - \theta)$ in that function.

$$F_{t+1} = g(s_t, L^f_t, p_t) = (s_t^\gamma p_t^{1-\gamma})^\theta L_t^f L_t^{1-\theta}. \tag{2.12}$$

In contrast to the benchmark model, I do not distinguish a high planting technology from a low one because it has no significant impact on the implication this extended model supposes to address, but only causes computational complications. The planting technology $g(\cdot)$ in Eq 2.12 is thus uniformly high technology. In Period 1, the very beginning, an amount of predetermined endowment plows $p_1 = \bar{p}_1$ is already available for planting. This gives the planting function at $t = 1$ as

$$F_2 = (s_1^\gamma p_1^{1-\gamma})^\theta L_1^f L_1^{1-\theta}.$$  

Last, the plow production technology is just a one-for-one linear transformation of $L^p_t$ for simplicity:

$$p_t = p(L^p_t) = L^p_t \quad t \geq 2. \tag{2.13}$$

The Period 3 harvest $h_3$ is only for consumption ($h_3 = c_3$) because nothing should be produced at the end.
The remaining parts keep consistent with the benchmark model. The household still has a 3-period life horizon, plows fully depreciate in each period, and the timing is similar to that of the benchmark model. Households' preferences remain quasi-concave in consumption where I leave leisure out for simplicity. Appendix A gives a full description on the timing of the model.

The new setting again gives rise to a series of capital-specific/ex-post PPFs and a complete/ex-ante PPF which envelops all ex-post PPFs over the \((c_1, c_2, c_3)\) space as in Figure 2.11. Again, each ex-post PPF is specific to the amount of plows \((p_2)\) pre-determined in Period 1. The only difference is that in the extended model there is no low planting technology for a household to choose. The household thus has to pick a positive \(p_2\) so as to ensure a positive \(c_3\); otherwise, it has nothing to consume in the last period according to Eq 2.10. A rational household would choose a particular \(p_2 \in (0, p_{\text{max}}]\) and uniquely determine an ex-post PPF. Here \(p_{\text{max}}\) is the value that makes \(c_3\) reach its global maximum value \(c_3^{\text{max}}\). As illustrated in the previous section, any point on the ex-ante PPF (as in Figure 2.11-(b)) is attainable for the household before a particular number of plows is picked; however, once a decision on \(p_2\) is made, the household actually faces an ex-post PPF (as in Figure 2.11-(a)), which is a contraction of the initial ex-ante PPF. Appendix A states the ways to obtain these PPFs.

A similar graph for \((c_2, c_3)\) PPFs can also be obtained at each particular value for \(c_1 \in (0, c_1^{\text{max}})\). Figure 2.12 illustrates 4 individual ex-post PPFs at 4 different \(p\) values and their corresponding ex-ante PPF with a higher transformation elasticity at every point except for the tangency point. Again, the tangency point is the only point on the ex-ante PPF a household can reach in the long-run once \(p\) is pinned down. This tangency

\[14\] However, if \(c_3 = 0, c_1\) and \(c_2\) could reach their global maximum values \(c_1^{\text{max}}\) and \(c_2^{\text{max}}\) respectively as showed in Figure 2.11.
Figure 2.11: Combined ex-post PPFs for 4 given $p_2 \in (0, p_{max})$ and an ex-ante PPF point between the ex-post PPF and its ex-ante PPF is thus the equilibrium point for both short-run and long-run under a Fisherian equilibrium term structure of interest rates. Period 1 term structure of real interest rates \( \{r_{112}, r_{113}, r_{123}\} \) may bring such an equilibrium by accident, or cause an impending boom in Period 2 associated with an unexpected rise in future spot rate \( r_{223} \), or an impending recession associated with an unexpected fall in \( r_{223} \). In another word, Proposition 2.2.1 and 2.2.2 still apply since
the extended model does not change the ‘Austrian’ nature of the heterogeneous capital framework.

![Graph showing the PPFs for different values of p2 and c1 = 15](image)

**Figure 2.12**: Ex-post and Ex-ante PPFs for $c_1 = 15$

### 2.3.2 Wage Rate and Capital Price

The specification of three technologies in Eq 2.11-2.13 enables a scrutiny of the factor market. The implications for changes of real wage rate and asset prices versus changes in real interest rates may help understand what happens in factor market if an excess demand/supply takes place due to misintermediation, and thus gives policy prescription.

In the model, labor is a very flexible factor in production. Not only can a household have labor contribute to current output, but it may also use labor to produce future
output by planting fields and building plows. The plow building-planting-harvesting
decision can thus affect the marginal product of labor (MPL) in all production activities,
which implies the real wage rate. Since efficiency requires that a household allocates
labor to equate its value of marginal product in all the uses, or the pay for all workers
must be the same, one can evaluate the real wage rate in either production activity. Be-
cause a portion of the labor stock is used for harvesting, which directly produces $c_2$, the
real wage can be the MPL in harvesting. Even though the labor would also be allocated
to other production activities, harvesting is still able to show us the wage rate directly
because in that way we do not need to consider future output which needs to be dis-
counted according to the prevailing term structure of interest rates. The real wage rate
I derive here is hence based on a harvesting function Eq 2.10 in Period 2.

$$w_2 = MPL_2^h = \frac{\partial f(F_2, L_2^h)}{\partial L_2^h}$$

$$= \alpha(1 - \eta) \left( \frac{F_2}{L_2^h} \right)^\eta \quad (2.14)$$

According to Eq 2.14, $MPL_2^h$ diminishes as the household increases its Period 2 la-
bor for harvesting. Suppose labor is directly producing current output in the harvest
function, then the marginal product of labor in that activity is the real wage in terms of

15We might be interested in looking at the welfare of labor in terms of all periods of output. In period
2, labor gets a current wage and it is also going to get a wage in period 3 which is discounted by the
interest rate. The present value of wages therefore might go on different directions than the current wage:
even though the current wage goes down, but future wage goes up, workers might be better off if the
present value of wages goes up. Representative agents may be worse off, but labor per se is not necessarily
representative.

16Plugging Eq 2.11 and 2.13 in Eq 2.14 gives:

$$w_2 = MPL_2^h = \alpha(1 - \eta) L_2^h \left( \left( \frac{s_1 p_1^{1-\gamma}}{L_1} \right)^\theta L_1^{1-\theta} \right)^\eta$$

$$= \alpha(1 - \eta) L_2^h \left( \left( \frac{\bar{F}_{1}^\eta L_1^{h-\eta}}{p_1^{1-\gamma}} \right)^\theta (L - L_1^p - L_1^h)^{1-\theta} \right)^\eta \quad (2.15)$$
that period’s output only. Therefore, if labor is not used for planting, or there is no planting technology, or there is no future needs for plows, then the real wage rate is constant simply because current labor does not contribute to future outputs. In order to have labor market implication, labor needs to contribute to future output and be used for planting. That is going to vary the amount of labor used for harvesting and change the marginal product of labor.

In addition, there may be an asset/stock market during Period 2, in which households trade shares of plows and fields having crops to be harvested. We can value these assets in terms of consumption goods $c_2$. According to Eq 2.12–2.13, the price/value of plows in Period 2 ($P^p_2$) is the product of real wage rate ($w_2$ or $MPL_h^f$) and the ratio of the marginal product of plows in planting to the marginal product of labor in planting ($MP_p^f/MPL_h^f$):

$$P^p_2 = w_2 \cdot \frac{MP_p^f}{MPL_h^f} = \frac{\partial f(F_2, L_h^f)}{\partial L_h^f} \cdot \frac{\partial g(s_2, L_h^f, p_2)}{\partial p_2} \cdot \frac{\partial g(s_2, L_h^f, p_2)}{\partial L_h^f}$$

$$= \alpha (1 - \eta) \left( \frac{F_2}{L_h^f} \right)^\eta \cdot \theta (1 - \gamma) p_2^{-1} (L - L_h^f)^{1-\theta} \left( (h_2 - c_2) \gamma p_2^{-1-\gamma} \right)$$

$$= \lambda \left( \frac{F_2}{L_h^f} \right)^\eta \cdot \frac{L_h^f}{p_2} (L - L_h^f)^{1-\theta}$$

where the constant term $\lambda = \alpha \theta (1 - \eta)(1 - \gamma)(1 - \theta)^{-1}$. The price/value of fields ($P^f_2$) in the same period is simply the marginal product of fields in harvesting ($MP_f^f$):

$$P^f_2 = MPF_2 = \frac{\partial f(F_2, L_h^f)}{\partial F_2}$$

$$= \alpha \eta \left( \frac{L_h^f}{F_2} \right)^{1-\eta}$$

17One can imagine that there is a farm owning crops and hiring workers to harvest. There is another company who hires workers to build plows and tries to sell them in the next period.
Apparently fields and plows also have diminishing marginal product in their respective capital goods. The value of this hypothetical capital market thus is

\[ P_2 = P_2^p p_2 + P_2^f F_2 \]  

(2.18)

If there are two outputs \( x \) and \( y \), and the price of \( x \) is raised, then the price of \( x \) intensive factor goes up in terms of both outputs. Therefore, suppose \( x \) uses more labor on average and its price goes up, then the price of labor goes up.

The finite elasticity of transformation and Austrian effects allow us to study factor prices changes under inefficient production as well as utility loss. Current factors could produce current or future output, whereas future factors can only produce future output. Therefore, for example, \( c_2 \) must be intensive in current factors. In a recession, current output becomes less valuable in terms of future output. Households ought to foresee a fall in demand for labor. We would then expect a reduction in the real wage. Plows probably fall in value because they contribute to future output only. If the household realizes it planned to plant too much \( c_2 \), then it has to harvest too much in Period 2 and it has to produce less \( c_3 \). Under a recession regime, there is not the expected demand for \( c_2 \). The household may regret the amount of crops it planted.

Does the model implication support this prediction and is it consistent with typical business cycle stories? The relation of factor prices to the intertemporal marginal rate of transformation (MRT) and therefore real interest rates in both the short- and long-run, derived from Eq 2.15–2.18, gives an answer to this question.

I first define MRT between \( c_2 \) and \( c_3 \) to be the slope of a PPF as in Figure 2.12. It is mathematically defined as \(-\frac{\partial c_3}{\partial c_2}\) \( |_{c_1} \). Figure 2.13 shows a negative relationship between the MRT and the production ratio \( c_3/c_2 \): a rise in the MRT leads to a fall in \( c_3/c_2 \) or vice
versa. The slope of each curve in Figure 2.13 turns out to be the elasticity of transformation of the corresponding PPF in Figure 2.12. Mathematically, it is defined as \( \frac{\partial \log(c_3/c_2)}{\partial \log(MRT)} \big|_{c_1} \).

In the short-run, an ex-post PPF has a smaller elasticity of transformation and \( c_3/c_2 \) is more rigid. A short-run curve in Figure 2.13 is hence flatter than the long-run one. The intersection points between a short-run and the long-run curves correspond to each tangency point between the ex-post and the ex-ante PPFs in Figure 2.12. The MRT will also be equal to the price of \( c_3 \) in terms of \( c_2 \), namely how much \( c_3 \) a household can get for one unit of \( c_2 \). In Period 1, it turns out to be the implicit real forward rate \( r_{123} \). At a low value of \( r_{123} \), \( c_3 \) is more expensive: a household has to pay more \( c_2 \) for a unit of \( c_3 \). At a high value of the forward rate, things are just the opposite. Since the whole model is taking interest/forward rates as driving everything, I use MRT to measure the changes in factor prices.
The shadow real wage rate versus MRT is shown in Figure 2.14. The negative relationship between this pair can be interpreted as follows. In order to get more $c_3$, or a high consumption ratio $c_3/c_2$, households have to divert Period 2 labor from harvesting to planting so that the hypothetical labor market has less labor to supply for harvesting ($L_h^2$) in Period 2. It follows that the marginal product of $L_h^2$, or real wage rate, has to increase following a decline in $L_h^2$ according to Eq 2.14. Since a high $c_3/c_2$ is associated with a low MRT, a positive relationship between MRT and $L_h^2$ is present as in Figure 2.15. This in turn leads to a negative relationship between MRT and real wage rate for both the long- and short-runs. The negative relationship further implies that the price of current output is falling relative to future output. Again, since the short-run PPF has lower elasticity of transformation everywhere including the tangency point, the slope of $L_h^2$ v.s. MRT is lower in short-run than in the long-run. The slope of real wage rate v.s. MRT
however turns out to be the opposite: a drop in wage rate is more dramatic in short-run than in long-run as intertemporal prices of current output fall (or MRT raises).

![Figure 2.15: Labor Demanded in Harvesting in Period 2](image)

According to Proposition 2.2.1, an impending boom is associated with a subsequent unexpected rise in MRT. The short-run equilibrium point in a boom deviates from the long-run and falls in the ex-post curve because of the contraction in ex-post PPF. This implies that in Figure 2.14 an equilibrium, the blue point for example, will deviate from its Period 1 plan and fall to the blue short-run real wage curve to the right. A household then ends up with an unexpected fall in real wage rate in a boom in Period 2. On the other hand, the household will go through an unexpected rise in real wage rate if an impending recession dominates. This conclusion contradicts the aforementioned prediction. It is also counterintuitive because we would have expected a rise in the wage rate in a boom, or a fall in a recession.
As for capital (both fields and plows) prices, we can also value them in terms of $c_2$ (and $c_3$). Again, to produce $c_2$ and $c_3$, a household has 4 factors to input: $F_2$, $p_2$, $L_2$, and $L_3$. $L_3$ is down the road, but it is still out there helping explain $c_3$. Therefore, 3 factors are actually taken into account during Period 2: $F_2$, $p_2$, and $L_2$. Since $c_3$ is $L_3$ and $p_2$ intensive, one would expect plows to go up in price ($P^p_2$) as the production ratio $c_3/c_2$ increases. In another word, the price of plows would go down with an increase in MRT as showed in Figure 2.16. On the contrary, $c_2$ is $F_2$ intensive. A rise in $c_2$ outputs, or a rise in MRT, would boost the marginal product of fields because of a higher demand for Period 2 fields to harvest. It hence makes sense to observe an upward sloping curve for fields price $P^F_2$ (v.s MRT) as illustrated in Figure 2.17. In addition, the stock market value in Period 2 as defined by Eq 2.18 turns out be negatively related to MRT according to Figure 2.18. This implies that the value of plows dominates that of fields in the model.
In an impending recession, an unexpected fall in MRT will occur according to Proposition 2.2.2. The equilibrium thus deviates from its Period 1 plan and falls to the ex-post curve because of the contraction in the ex-post PPF. In Figure 2.16, an equilibrium, the blue point for example, will deviate from its Period 1 plan and move up to the blue short-run plow price curve on the left. A household then ends up with an unexpected rise in the plow price in a recession. In Figure 2.17, the blue long-run equilibrium point will jump to the blue short-run field price curve on the left too. A household then realizes an unexpected rise in field prices in Period 2. Households wish they have built more plows than fields back to Period 1 in a recession. Plows hence go up in value relative to fields because households hold too many fields but too few plows. The net effect on the stock market is that its value goes up. In Figure 2.18, the blue long-run equilibrium point will jump to the blue short-run stock market value curve on the left. This results in an unexpected rise in stock market value in a recession. In a boom, everything is just
the opposite. The changes in value of the stock market are also counterintuitive, which contradict our prediction.

![Figure 2.18: Stock Market Value in Period 2](image)

In sum, misintermediation would not only bring about unexpected output fluctuations and surprises in interest rates, but also give rise to unanticipated changes in factor prices. However, the changes in real wage rates and the value of the stock market are inconsistent with the prediction. A more complicated model or different elasticity of substitution may reverse them.

### 2.4 Conclusion

This chapter shows that misintermediation plays a potentially important role in macroeconomic fluctuations. The benchmark heterogeneous capital model demonstrates how
misintermediation subjects an economy to additional, unnecessary interest rate uncertainty, and to inefficiency in the intertemporal production process. The link between unexpected changes in interest rates and output gives a new implication to the recent financial crisis and economic downturn. The introduction of labor in the extended model enables a useful enrichment. It helps explore the implication of misintermediation for factor prices and of planned employment under a recession or boom regime. A harvesting technology that requires some labor at time $t$ in addition to planting activities at $t - 1$ (to obtain fields harvesting at time $t$) further develops the model implication. It keeps the price of maturing fields from being identically 1 in terms of harvested output, and enriches the decisions that need to be made.

Nevertheless, this theoretical analysis has room for further improvement. As for the labor market, we may want to know whether it leads to structural unemployment if real wage rates are sticky. Then, the model implication for changes in real wage rates and the value of the stock market are still inconsistent with our prediction. What is the reason for this inconsistency? Further, some technological uncertainty can be added in the form of shocks to the production functions. What is the equilibrium term premium? What if intermediaries subsidize term structure risk by going for the highest expected return rather than the highest risk-adjusted return? A thorough investigation is quite necessary to answer these questions.
Chapter 3: Estimating the U.S. Term Structure of Interest Rates

3.1 Introduction

The term structure of interest rates is a concept central to both economic theory and financial applications. The interpretation of macroeconomic fluctuation, monetary policy, and pricing and hedging of derivatives should be based on a complete set of interest rates, spanning from the very short term to the longest. Estimating the term structure of interest rates, either in real or nominal terms, is therefore a necessary starting step of any related research or practice. “Term structure” here refers to the discount function, the zero-coupon yield curve, or the forward rate curve since each is a transformation of the others. Basically, they may all be derived from the benchmark yields of riskless Treasury securities at different maturities.

If the Treasury issued bills, notes and bonds with all maturities, then we could simply observe yield curves implied by this full set of the yields and forward rates. In practice, however, the U.S. Treasury has instead issued a limited number of securities with different maturities and coupons. Each of these can be viewed as a basket of zero-coupon securities: one for coupon payment at each semiannual coupon date and one for the principal payment at maturity. In general, we do not have securities at all maturities and hence cannot simply solve for the implied zero-coupon yields. Instead, we must fill
in the “missing” bond yields across the maturity spectrum so as to retrieve a continuous yield curve.

Two lines of research have been trying to recover the “missing” bond yields. One is featured by applying interpolation or curve-fitting methods. The basic idea is to assume the price of any security to be governed by a smoothly fitted discount function, \( \delta(m) \). Fama and Bliss (1987), McCulloch (1975b); McCulloch and Kwon (1993); McCulloch and Kochin (2000), Nelson and Siegel (1987), and Fisher et al. (1995) follow this line, though these studies differ considerably in their respective interpolation models and estimation techniques. The other approach instead focuses on fully specified, dynamic term structure models. In this category, a prevailing approach is to design reduced-form term structure models in which yields are expressed as constant-plus-linear function of some state variable(s); both of the state variables and coefficients follow stochastic processes. This kind of term structure models is known as Affine Term Structure Model (ATSM), traced back to at least Vasicek (1977), Cox and Ross (1985), and further refined by Duffie and Kan (1996), Dai and Singleton (2000), and Piazzesi (2003). A successful element of ATSMs is the cross-asset restrictions imposed on the model to eliminate arbitrage opportunities. Its application also widely extends to other areas, such as swap rates, exchange rates, and associated derivatives.

Nevertheless, a number of desirable properties make the approach of interpolation irreplaceable in studying term structure of interest rates. First of all, interpolation models are independent of any preference forms and economic assumptions. This type of model is pure descriptive in that yield is a function of term to maturity only, without considering a general equilibrium setting or the selection of state variables as it does in
ATSMs. Second, this approach allows simple, parsimonious functional forms to represent a wide range of shapes generally associated with yield curves. Nelson and Siegel (1987) advocate parsimonious yield curve models succeeding in the objective that “the whole term structure of yields can be described more compactly by a few parameters.” Third, the reduced-form function of term to maturity allows easy transformation among the discount function, the zero-coupon yield, the forward rate, and the par bond yield, and thus makes various curves easy to be displayed on a graph for informative purpose. Finally, the interpolation methods do not rely on historical data, but instantaneous market quotes such as The Wall Street Journal has been releasing on each business day. This offers great convenience to practitioners and those researchers who have no access to proprietary databases such as CRSP.

Interpolation typically starts out from specifying a functional form either to approximate a discount function or yield curves, and then estimates the unknown coefficients. An extensive body of works hence has contributed to develop favorable functional forms as well as efficient estimation techniques. Among them, two major approaches exhibit great potential to facilitate high frequency fit with a full spectrum of maturities: one is McCulloch and Kochin (2000); the other is extended Nelson and Siegel (Nelson and Siegel (1987), Svensson (1994)). The former proposes a Quadratic-Natural (QN, henceforth) cubic spline functional form to fit the negative log discount function, whose coefficients are estimated by a well-designed iterative procedure based upon linear least squares. The latter instead fits the forward rate curve via a set of exponential decay functions and nonlinear minimization methods. Gürkaynak et al. (2007) point out that neither of the two approaches wins over the other; the selection of approximating functions

\[18\] Bliss (1996) gives an in-depth survey of this literature.
largely depends on the purpose of fitting a yield curve. They argue that two functional forms differ in the flexibility that each allows the fitted yield curves to have. The cubic spline approach, according to their argument, brings more flexibility to the shape of a yield curve and thus is preferable for financial practitioners who are looking for small pricing anomalies. In contrast, macroeconomists may prefer the more parsimonious exponential decay function because a relatively rigid forward curve that smoothes through idiosyncratic variations helps investigate the fundamental determinants of the term structure of interest rates. Unfortunately, there is no further research trying to integrate the advantages of two approaches and put up with new functional forms. This becomes a major motivation of this chapter.

Specifically, I am answering three questions in constructing a new interpolation model for term structure estimation. First, what target function are we going to select? Second, what kind of interpolation functional form do we specify? Third, how do we estimate the coefficients in the constructed model?

For the target functions, McCulloch (1975b); McCulloch and Kwon (1993) focused on the discount function directly. By modeling a discount function, they derive a linear pricing function for coupon-bearing securities, which makes the estimation largely easier. However, the term structure behavior at long-end maturities will become hard to govern if the discount function is directly targeted. This defect is solved by another approach: instead of targeting on the discount function, the bond pricing function can be specified in terms of the logarithm discount function with well designed end-point conditions. In that way, we can get control over the behavior of yields at long maturities. Nevertheless, modeling a log discount function results in a nonlinear pricing function,

The second dimension of this problem is to specify an interpolation functional form for the target function selected. Again, two approaches are competing in this area. The QN cubic spline approach, introduced by McCulloch and Kochin (2000), features a non-parametric model. It allows substantial flexibility on the shape of yield curves, but we do not need too much most of the time. Nelson and Siegel (1987), Svensson (1994), and Gürkaynak et al. (2007, 2010) in turn specify an exponential decay functional form. In contrast to the cubic spline, the exponential decay form results in a parametric and parsimonious model, which, though depicting a relatively rigid forward rate curve, is able to smooth through idiosyncratic variations generated by individual securities. The two approaches also differ in their long-end behavior. If the log discount function is targeted, the cubic spline functional form needs an extra end-point condition to regulate its long-end behavior; otherwise, the yield curve or the forward rate curve will not necessarily show asymptotical convergence at the long-end. However, no extra condition is required for exponential decay function to control where the curve goes because of its natural property of asymptotic decay.

Estimating the unknown parameters is a relatively less important issue in the literature; however, various approaches do differ substantially in terms of efficiency and effectiveness. In addition, a single estimation method is unable to match all curve-fitting purposes, which call for diversified and specific estimating techniques. Unfortunately, the previous literature fails to address this issue.

In order to answer the three questions, I first propose a multiple exponential decay functional form to fit the negative log discount function. Then, several estimation
methods are introduced to estimate unknown parameters in an objective function according to two distinct curve-fitting purposes. As a comparison, I extend McCulloch cubic spline model to fit the same data set. Since the placement of knots in the previous QN cubic spline approach is arbitrary, I introduce and implement a procedure to pin down the knots based upon the Weighted Least Squares Durbin-Watson (D-W, henceforth) test, the Nonparametric Runs test, and the Bayesian Information Criterion (BIC, henceforth).

Generally, the focus of this chapter is methodological. The remainder of this chapter is organized as follows. Section 3.2 briefly reviews the fundamental definitions of yields and the basic “bond math”. Section 3.3 describes the interpolation model based upon the exponential decay functions I employ to fit yield curves. I also compare to two related model specifications in the same line of research. Section 3.4 discusses the estimation methods for unknown parameters in the interpolation model and shows the results obtained through these methods. Section 3.5 extends the estimation methods in Section 3.4 for financial practitioners who instead concerns portfolio management. Section 3.6 concludes.

### 3.2 Identities and Definitions

Any pricing problem for fix-income assets starts out from a discount function. Let $\delta(m)$ denote the discount function as a function to maturity $m$. It gives the present value of $1$ to be paid in $m$ years ahead, and it immediately follows that $\delta(0) = 1$, $\delta(m) \geq 0$, and $\delta(\infty) \rightarrow 0$. Although this definition gives the price of a zero-coupon security only, given the discount function, one can price any fix-income asset with coupon payments bearing on it. For example, the price of a hypothetical continuous coupon-bearing bond,
with maturity $m$, coupon rate $c$, and principal payment $1$ at maturity, is calculated as the summation of its individual payments:

$$p = \delta(m) + c \int_0^m \delta(s) \, ds.$$  \hspace{1cm} (3.1)

We can also derive the interest rate of this bond, or the yield, once the discount function is known. Conventionally, three concepts of yield are used to depict the term structure of interest rates. The most fundamental and mathematically simplest concept is the zero-coupon yield. The continuously compounded zero-coupon yield on a zero-coupon bond is defined as:

$$y(m) = -\log(\delta(m)) / m.$$ \hspace{1cm} (3.2)

Except for Treasury bills (T-Bill, henceforth) which are sold at discount with a maturity less than one year, all other Treasury securities, including Treasury notes, Treasury bonds and Treasury Inflation-Protected securities (T-Note, T-Bond, and TIPS respectively, henceforth), are coupon-bearing securities. Hence, another prevailing way to quote coupon-bearing yields is through par bond yields. A par bond yield is the interest rate at which a security with a maturity $m$ would sell just at par (and hence have a coupon-equivalent yield equal to that coupon rate). The continuously compounded par bond yield with maturity $m$ is given by:

$$y^p(m) = \frac{1 - \delta(m)}{\int_0^m \delta(s) \, ds}.$$ \hspace{1cm} (3.3)

A term structure can also be expressed in terms of forward rates. A forward rate is the future yield on a security, at which an investor would agree to make an investment today over a specified period in the future—for $m_2$ years beginning $m_1$ years hence. The continuously compounded zero-coupon return on this future agreement can be expressed
as follows:

\[ f(m_1, m_2) = -\frac{1}{m_2} \log \left( \frac{\delta(m_1 + m_2)}{\delta(m_1)} \right) \]

\[ = -\frac{1}{m_2} [(m_1 + m_2)y(m_1 + m_2) - m_1y(m_1)] \]

Taking the limit as \( m_2 \) goes to zero gives the instantaneous forward rate \( m \) years ahead, which implies the instantaneous return for a future date that an investor would demand today.

\[ f(m) \equiv f(m_1, 0) = y(m) + my'(m) = (-\log \delta(m))' = -\frac{\delta'(m)}{\delta(m)}, \quad (3.4) \]

which implies

\[ y(m) = \frac{1}{m} \int_0^m f(s) \, ds, \quad \text{and} \quad f(m_1, m_2) = \frac{1}{m_2 - m_1} \int_{m_1}^{m_2} f(s) \, ds. \quad (3.5) \]

Eq\[3.1\] through 3.5 show that each of the zero-coupon yield curve, the par yield curve, or the forward rate curve is a transformation of the others. Once the discount function \( \delta(m) \) is specified, we can price a bond and express its continuous term structure of interest rates. In this sense, the estimation of an empirical term structure becomes a problem to specify and estimate a discount function. In another word, any research work toward an estimated term structure of interest rates can be equivalently “translated” into a study of discovering a discount function.

Moreover, the above equations also imply a certain graphical relationship between the zero-coupon yield and the instantaneous forward rate. By way of an analogy, they are similar to average product (AP) and marginal product (MP). If AP curve has a peak, MP curve has to cross through AP via that peak from above, where MP is equal to AP. Similarly, if a zero-coupon yield curve has a peak, an instantaneous forward rate has to cross through that point from above, or vice versa. I depict a representative U.S. term
structure of interest rates in Figure 3.1 to deliver readers a fresh image for curves of the zero-coupon yield, the par bond yield, and the forward rate\textsuperscript{19}

Figure 3.1: U.S. Term Structure of Real & Nominal Interest Rates on 07/30/2010

### 3.3 Model

In this section, I specify several interpolation models for estimating term structure of interest rates. I will discuss the properties and distinctions of each model over the others in the same line of research. The estimation methods and results are in next section.

\textsuperscript{19}I refer to the “instantaneous forward rate” as the “forward rate” henceforth for convenience.
3.3.1 Discount Function v.s. Log Discount Function

According to Section 3.2, the modeling of any yield is based on an explicit discount function. If we, being able to construct a model for the discount function, actually obtain a continuous term structure. However, direct modeling the discount function is not a preferred approach \(^{20}\) because it is hard to regulate the behavior of that discount function, especially at its long-end. Theoretically, a discount function curve should decline monotonically and decay to zero asymptotically since it gives the present value of a security as maturity \(m\) goes up. Yet this is not always the case empirically. An empirical discount function, with coefficients estimated by observed market quotes, is likely to enter the negative region as \(m\) becomes large. In order to avoid this problem, I estimate the (negative) log discount function instead. This alternative approach has a favorable property that it is able to depict an upward-sloping curve beyond the longest maturity on a graph by imposing an appropriate end-point condition so that I can govern its behavior across maturities and make sure what happens out there. For example, McCulloch and Kochin (2000) use a long-end natural condition in a cubic spline model to let the log discount function monotonically increase after the longest maturity outstanding. The difference between a discount function and the corresponding (negative) log discount function is illustrated in Figure 3.2.

Of course, there are advantages to constructing a discount function. Particularly, if the target function is the discount function, the corresponding bond price function becomes linear in unknown parameters, whose estimation is relatively easier. In contrast, if one constructs a log discount function model, the corresponding bond price function is nonlinear, which causes estimating complications. Nevertheless carefully designed

\(^{20}\) Since McCulloch and Kwon (1993), there has not been any research choosing this approach.
estimation methods, such as Iterative Linear Least Square in Section 3.4, will largely reduce the degree of calculation complexity.

3.3.2 Model Specification

Considering the pros and cons of two target functions, this chapter chooses to target on a (negative) log discount function. The general functional form \( \phi(m) \) is given by:

\[
\phi(m_i) \equiv -\log \delta(m_i) = \sum_{j=1}^{k} \beta_j \Psi_j(m_i) \quad i = 1, \ldots, n
\]

s.t. \( \delta(0) = 1 \) or \( \phi(0) = 0, \Psi_j(0) = 0; k \leq n \)  

(3.6)
which is a linear combination of $k$ basis functions $\Psi_j(m)$, with coefficients $\beta$, subject to the condition that the fitted $\phi(m)$ goes through the origin and the total number of its parameters $k$ is no more than $n$, the total number of maturities outstanding.

Two heuristic facts help us specify the form of the general function $\phi(m)$ and basis functions $\Psi_j(m)$. First, the discount function features approximately an exponential decay curve: some shape (monotonic decline, humped, or $S$ shaped) at the short-end and an asymptotical decay at the long-end. The negative log discount function in turn should be (approximately) monotonically increasing. The linear functional form in Eq 3.6 thus keeps in line with this feature. Second, according to the ATSM, the yields $y(m)$ are generated by a differential equation; the forward rate $f(m)$ thus takes an exponential functional form because it turns out to be the solution to $y(m)$ by Eq 3.5. Then, according to Eq 3.4 $f(m)$ is the first derivative of $\phi(m)$, which hence should also take exponential forms. The basis functions $\Psi_j(m)$ therefore would be given by exponential functions.

In addition, there are several properties I expect $\Psi_j(m)$ to have. In order to govern the long-end behavior of these yield curves, the first derivative of one $\Psi_j(m)$ should be one, and all the other $\Psi_j(m)$ die out to zero at infinity; their first derivatives also die out to zero, so that the forward rate curve will have an asymptote and the yield curve have the same asymptote. Given all these features, I specify $\Psi(m)$ as:

$$\Psi_j(m) = \begin{cases} 1 - \exp\left(-\frac{m}{\tau_j}\right) & \text{if } j = 1, \ldots, k - 1 \\ m & \text{if } j = k \end{cases}$$

(3.7)

where parameters $\tau_j$ are constant to maturity $m$ but specific to $\Psi_j(m)$.

The parameters $\tau_j$ in Eq 3.7 associate with a concept called “E-fold Life”. It is defined as the timescale for a quantity to attenuate to $1/e$ of its previous value, just analogue to
“Half Life”. As \( m \) increases from below, \( \exp \left( -\frac{m}{\tau_j} \right) \) will decline to \( 1/e \) as \( m = \tau_j \). It thus follows that different values of \( \tau_j \) correspond different speeds of decay. Small \( \tau_j \) realizes rapid decay in the regressors and therefore will be able to fit curvature at short maturities well while unable to fit excessive curvature over longer maturities. In contrast, a large \( \tau_j \) produces slow decay in the regressors that can only fit curvature over longer maturities. Figure 3.3 illustrates different shapes of \( \Psi(m) \) as \( \tau \) varies. On one extreme, \( \Psi_j(m) \) converges to one at once as \( \tau_j = 0 \). As \( \tau_j \to \infty \), on the other extreme, \( \Psi_j(m) \) takes forever to converge. When \( \tau_j \in [0, \infty) \), the speed of convergence declines as \( \tau_j \) increases, and the function generates distinct curvatures. Since \( k \) determines the number of curvatures the function is able to feature, by properly introducing \( k \) terms of \( \Psi_j(m) \) into \( \phi(m) \), one can fit any number of curvatures for the yield curves at exact maturities needed.

### 3.3.3 Comparison

Basically, most studies of interpolation models for an empirical term structure can be generalized as constructing a similar general form \( \phi(m) \). The difference is the specification of the basis functions \( \Psi(m) \). Here I make a comparison to two related interpolation models: one is \( QN \) cubic spline interpolation; the other is the extended Nelson and Siegel model.

McCulloch and Kochin (2000) fit a \( QN \) cubic spline to a negative log discount function. They select a small number of well-spaced maturities (as knots) from a large pool of T-Notes & Bonds to fit a just-identified term structure. Specifically, \( \nu + 1 \) knots \( \{ \kappa_j \}_{j=0}^\nu \) are selected, where \( \kappa_0 \equiv 0 \) and \( \kappa_\nu \equiv m_n \). In order to let the spline be extrapolated linearly beyond the longest maturity outstanding, they impose a “natural” restriction on
the last knot: $\phi(m)'' = 0$, for $m \in [m_n, \infty)$. At the short-end, they impose a “quadratic” restriction instead, rather than a second natural restriction, to avoid the counterfactual outcomes: $y'(0) = 0$ and $f'(0) = 0$. The quadratic restriction on the first interval is given by: $\phi(m)''' = 0$, for $m \in [0, m_1]$. They thus name their spline model Quadratic-Natural. This cubic spline model, subject to the two restrictions, is specified as:

$$\Psi_j(m) = \theta_j(m) - \frac{\theta''_j(m_n)}{\theta''_{v+1}(m_n)} \theta_{v+1}(m), \quad j = 1, \ldots, \nu$$  \hspace{1cm} (3.8)

where $\theta_1(m) = m$

$\theta_2(m) = m^2$

$\theta_j(m) = \max\{0, (m - \kappa_{j-2})^3\}, j = 3, \ldots, \nu + 1$

There are several disadvantages in this model specification. First, although the spline model is flexible enough to fit various shapes associated with yield curves, the cubic
function tends to be too sensitive to idiosyncratic fluctuation. Idiosyncratic issues arise for reasons, such as liquidity premia, hedging demand, demand for deliverability into future contracts, or repo market specialty. Now that a term structure of interest rates is supposed to exhibit market-wide returns, noise generated by individual securities is definitely undesirable. Second, the two artificially imposed restrictions add extra unwanted features to the model. The original purpose of imposing end-point conditions is to pin down all unknown parameters in an exact-fit term structure model. Unfortunately, the natural restriction results in a linearized yield curve at the long-end, which in turn leads to hypothetical arbitrage opportunities\[21\] The short-end quadratic restriction is actually an expedient in the absence of a second natural restriction. It arbitrarily assumes that the log discount function takes quadratic form within its first interval, which has no supporting evidence.

Admittedly, shortcomings in the QN cubic spline model are inevitable in considering the existence of certain limitations. Back to a decade ago when the method was first proposed, only a handful of distinct maturities in TIPS market were outstanding. This constraint made fitting an over-identified real term structure impossible. McCulloch and Kochin (2000) therefore had to impose a quadratic restriction for an exact fit. In order to ensure the consistency, McCulloch and Kochin (2000) then use the same set of TIPS maturities for the estimate of nominal term structure from over 150 U.S. T-Bonds, T-Notes, and T-Bills outstanding. Later, even though more TIPS securities were issued, the co-authors did not adjust the model so that all estimated data were historically comparable. This gives rise to the problem of biased estimation: with a fixed number of knots, the more securities outstanding, the more biased the estimated term structure

\[21\] Since there is no security with longer maturity beyond the longest one outstanding, the hypothetically existed arbitrage opportunities are not realizable, at least in short-term.
could be. It is because that only those securities selected as knots enter in estimation, while the price information in other securities are simply ignored.

Accordingly, I extend the cubic spline model to partly improve McCulloch and Kochin (2000). First, I select \( k + 1 \) maturities as knots \( \{\kappa_j\}_{j=1}^{k+1} \) by assigning the first knot to be zero, the last equals to the longest outstanding maturity, and the remaining \( k - 1 \) knots evenly distributed in the maturity domain \(^{22}\) Then, I impose a *natural* restriction the same way as in McCulloch and Kochin (2000) but leave the *quadratic* restriction out. Since the *quadratic* restriction at short-end is eliminated, I call this modified model *Semi-natural* cubic spline. The basis function \( \Psi(m) \) is specified as:

\[
\Psi_j(m) = \theta_j(m) - \frac{\theta'_j(m_n)}{\theta''_{k+2}(m_n)} \theta_{k+2}(m), \quad j = 1, \ldots, k + 1
\]  

\[
\text{where} \quad \theta_1(m) = m
\]
\[
\theta_2(m) = m^2
\]
\[
\theta_3(m) = m^3
\]
\[
\theta_j(m) = \max\{0, (m - \kappa_{j-3})^3\}, \quad j = 4, \ldots, k + 2.
\]

Because this spline model is fitted to an over-identified term structure, all outstanding securities are taken into account for estimation; the biased estimation problem thus does not exist in this extended model. However, the hypothetical arbitrage opportunities are still present because we have to impose a long-end *natural* restriction in order not to let a cubic diverge to infinity at large. In a vivid contrast, the multiple exponential decay function does not need an artificial restriction in that an exponential decay is inherent in it, which is one of the advantages associated with this functional form.

\(^{22}\)Mathematically, if we observe \( n \) maturities in the bond market ranked from the shortest, \( \{m_i\}_{i=1}^{n} \), then \( \kappa_1 = 0, \kappa_{k+1} = m_n, \) and \( \kappa_j = m_1 + \frac{(j-1)(m_n - m_1)}{k}, j = 2, \ldots, k \}. \) Except for the two end-point knots, the knots may not be identical to outstanding maturities, but definitely within its domain.
The second closely related model is the extended Nelson and Siegel\textsuperscript{23} This class of models originally targets on the forward rate. G"urkaynak et al. (2007, 2010) summarize the model as:

\[ f(m) = \beta_0 + \beta_1 \exp\left(-\frac{m}{\tau_1}\right) + \beta_2 \left(\frac{m}{\tau_1}\right) \exp\left(-\frac{m}{\tau_1}\right) + \beta_3 \left(\frac{m}{\tau_2}\right) \exp\left(-\frac{m}{\tau_2}\right) \]  \hspace{1cm} (3.10)

In order to be comparable, I integrate \( f(m) \) according to Eq.\textsuperscript{3.4} and obtain another expression in terms of negative log discount function, \( \phi(m) \). The general functional form is identical to Eq.\textsuperscript{3.6} while the specification of basis function \( \Psi(m) \) is given by:

\[
\begin{align*}
\Psi_1(m) &= m \\
\Psi_2(m) &= m \exp\left(-\frac{m}{\tau_1}\right) \\
\Psi_3(m) &= 1 - \exp\left(-\frac{m}{\tau_1}\right) \\
\Psi_4(m) &= m \exp\left(-\frac{m}{\tau_2}\right) + \tau_2 \left(1 - \exp\left(-\frac{m}{\tau_2}\right)\right)
\end{align*}
\]  \hspace{1cm} (3.11)

Several features associated with this model are immediately observed. Firstly, it allows instantaneous forward rates to begin at horizon zero at the level \( \beta_0 + \beta_1 \) and eventually to asymptote to \( \beta_0 \). Second, it is a parsimoniously parametric form, in which a fixed number of \( \beta \) and \( \tau \) are present. In addition, two “humps” are allowed in the forward curve, the locations of which are determined by \( \tau_1 \) and \( \tau_2 \). Figure\textsuperscript{3.4} shows the “hump” shape associated with different values of \( \tau \) in the extended Nelson and Siegel model. They argue that a yield curve often needs two humps, one at short maturities associated with monetary policy expectations and the other at the long-end to capture convexity effects. It is hence apparent that \( \Psi_1(m) \) in Eq\textsuperscript{3.11} determines a long-term horizon level while \( \Psi_2(m) \) to \( \Psi_4(m) \) characterize the short- and medium-term shapes: monotonic, curvatures, or humps.

\textsuperscript{23}Nelson and Siegel (1987), and further developed by Svensson (1994)
The extended Nelson and Siegel and our multiple exponential decay model share some common elements in modeling the $\Psi(m)$; however, they differ considerably in the flexibility each model allows the fitted term structure to have. For the former, the number of the basis functions $k$ is fixed at four, i.e. only four $\Psi_j(m)$ enter the interpolation function with six unknown parameters—four $\beta$’s as well as two $\tau$’s. This results in a rather rigid yield curve for any business day and bond market. As aforementioned, it must assign two humps to an estimated yield curve, which could be redundant for a real term structure and inadequate for a nominal one. The multiple exponential decay function does not have this restriction since it lets any $k$ terms of $\Psi(m)$ enter so as to introduce equal or even richer shapes to the fitted yield curves. Although each $\Psi_j(m)$ in Eq 3.7 is just able to generate a single curvature, by controlling over the parameters $\tau_j$
in a combination of $\Psi_j(m)$ it is able to generate a hump or even more complex shapes. This feature greatly enhances the fitting flexibility of my model. The exact $k$ is determined by the actual market quotes, and thus specific from case to case. In next section, I will discuss how to pin down the parameter $k$.

In sum, the multiple exponential decay function has obvious advantages over other models in the same line of research. Compared to the cubic spline models, it is more parsimonious, more robust to idiosyncratic variation, and free from arbitrage opportunities caused by end-point restrictions. Compared to the extended Nelson and Siegel, my model offers more flexibility to the fitted yield curves. I will further demonstrate these advantages in next section with concrete results.

### 3.4 Estimation and Results

According to the model specification in Eq 3.6 and 3.7, three groups of parameters, $\beta$, $\tau$, and $k$, are unknown. In this section, I introduce the estimation methods for each of the three, and show the estimated term structures. I first estimate $\tau$ and $\beta$ assuming $k$ is given. Then I will discuss the technique for picking an optimal $k$.

#### 3.4.1 Select $\tau_j$

In Gürkaynak et al. (2007, 2010), both $\beta$ and $\tau$ are obtained simultaneously through a (unconstrained) nonlinear minimization. According to their empirical report, computational efficiency matters in nonlinear optimization algorithms. For the multiple exponential decay model, $k$ terms of $\Psi_j(m)$ are followed by $(2k - 1)$ parameters/coefficients
to be pinned down. Direct estimation in a “lump-sum” manner is no longer preferable if
$k$ is large\textsuperscript{24} This fact motivates us to modify their approach.

The idea of my approach comes from the characteristics of the exponential decay function. Section 3.3 concludes that \( \tau_j \) regulates the speed of decay for each exponential function. Different values of \( \tau_j \) thus depict curvatures at different maturities. We know that bond maturities are not uniquely distributed over time horizons from zero to thirty-year, but rather skew to the short maturities: in most cases we will have more short, less medium, and even sparser long maturities outstanding. This arguably follows that there are richer shapes showing up along short- and medium-term. “E-fold Life” then suggests that a larger portion of \( \tau_j \) should be given small values. This logic offers the following way to pin down parameters \( \tau_j \): instead of estimating them directly, I can assign a value for each \( \tau_j \) based on the distribution of maturities outstanding, and then estimate \( \beta_j \) accordingly. Specifically, given \( k \) and a serial of \( n \) maturities, \( m \equiv \{m_i\}_{i=1}^n \), this idea can be realized as follows:

**Step 1**: Let \( \tau_k = \infty \) so that \( \Psi_k(m) = m \) and \( \beta_k \) hence governs the level of \( f(m) \) at infinity \( (m \to \infty) \);

**Step 2**: Given \( 2 \leq k \leq n \), mark the place of each \( k \)-percentile along \( \{m_i\}_{i=1}^n \), \( \{m_l|l = \lfloor nj/k \rfloor, j = 1, \cdots, k-1\} \), where \( \lfloor nj/k \rfloor \) is the round-down integer of \( (nj/k) \);

**Step 3**: Assign \( \{\tau_j = m_l|j = 1, \cdots, k-1, l = \lfloor nj/k \rfloor\} \).

\textsuperscript{24}I follow the estimation method in Gürkaynak et al. (2007) by using optimization toolbox in MATLAB. It takes more than one hour to obtain a locally optimal solution with \( k \geq 5 \). This is definitely not practical if we want to estimate both real and nominal term structures in real time or on each business day during the last decade.
3.4.2 Estimate $\beta_j$

I introduce two techniques to estimate parameter $\beta$ given $k$ and $\tau$ assigned in advance. The first one is Nonlinear Minimization, in which I simply choose $\beta$ to minimize the weighted sum of the squared deviations between the actual prices of outstanding Treasury securities and the estimated ones. The weights are the inverse of the bid-asked price spread of each individual security\textsuperscript{25}. Even though being rather straightforward, this method involves nonlinear optimization algorithms\textsuperscript{26}, which are time-inefficient and sometime can just return locally optimal solutions by, for example, the basic optimization toolbox in MATLAB.

A natural way to bypass this problem is to transform the nonlinear problem to a linear one and use least squares directly. The only difficulty is the lack of coupon prices; however, coupon values can be numerically solved out by iterative procedures. Following McCulloch and Kochin (2000), I introduce the second estimation method—Iterative Linear Least Squares (ILLS, henceforth). Generally speaking, I obtain a first guess of $\beta$ by each security's yield-to-maturity and then estimate $\beta$ iteratively. The coupon value in each iteration is computed by the estimated value $\hat{\beta}$ from the last iteration. Appendix C gives details of the two methods.

Although either of the two methods will return an optimal solution for $\beta$, ILLS has significant advantages over Nonlinear Minimization in terms of computational efficiency. Figure 3.5 shows the estimated real and nominal U.S. (negative) log discount function $\phi(m)$ on February 26, 2010 based on both Nonlinear Minimization and ILLS. On that

\textsuperscript{25}Securities with a large bid-asked price spread are most likely those inactive or off-the-run securities, and hence give less information about the market level of yields, vice versa.

\textsuperscript{26}Although $\beta_j$ is linear in log discount function, the discount function will become exponential in $\beta_j$ according to Eq 3.6 and enter the price function through Eq 3.1. Together with the square, the objective function will be nonlinear and complicated.
business day, 31 TIPS bonds and 198 T-Notes & Bonds were outstanding. According to the data pre-processing procedures in Appendix B, 29 TIPS and 142 T-Notes & Bonds are selected out to estimate $\beta$. I call the selected securities “effective securities” as they are not only taken into account, but actually enter in the term structure estimation. In Figure 3.5, the cross signs give the estimated negative log net price striped of coupon values. The solid curves are fitted negative log discount function curves. Notice that they are two (almost) overlapped curves as results returned by respective estimation method, which suggests the equivalent effectiveness of the two methods. For a same result returned, ILLS however takes far less computational time (shown in the legend areas). I thus recommend ILLS for the term structure estimation and use it to report all empirical results in the remainder of this chapter.

With the fitted log discount functions, I then derive the implied term structure of interest rates. Figure 3.6 depicts the real U.S. term structure on February 26, 2010. The upper left graph (in Figure 3.6) corresponds real log discount function derived by the multiple exponential decay model in Figure 3.5. The other two are given by Semi-natural Cubic Spline and Extended Nelson and Siegel for a comparison.

There are four distinctions among the three fitted term structures. First, the pattern of long horizon decay differs between cubic spline model and models based on exponential functions. As expected, both forward rate and yields remain at a constant after the longest maturity outstanding for spline-fitted term structure, which becomes the source of the hypothetical arbitrage opportunity. In contrast, the multiple exponential decay model and the extended Nelson and Siegel feature an asymptotical decay to infinity for forward rate and another asymptote for yields. Yet yield curves driven by the two exponential-based models do not share an identical asymptote in that the estimate
of $\beta_j$ associated with the common term $\Psi_j(m) = m$ differs between models. The second distinction involves the shape of forward rate curve. Even though all three models generate a hump-shape in the forward rate curve, the curve driven by a cubic spline has a peak at 8 years to maturity while the other two at 12 years to maturity approximately. Moreover, the forward rate curve in the cubic spline model is steeper and has sharper curvatures. Third, the cubic spline does not fit observed data as closely as the other two models. Fourth, even though the multiple exponential decay and the extended Nelson-Siegel 

27 Section 3.3 indicates that this $\beta_j$ actually governs the long-horizon level of convergence.
and Siegel return similar yield curves, the former has only three coefficients in its model specification \((k = 3)\) versus the fixed four in the latter. Also, the multiple exponential decay model has one less term than the cubic spline in this case. This suggests that multiple exponential decay model is able to use the most parsimonious model specification to return equally well-fitted yield curves.

For all the results showed above, I pick \(k = 3\) for the multiple exponential decay model and \(k = 4\) for the cubic spline model. In next subsection, I will explain why I do so and introduce a way to pin down this parameter optimally.
3.4.3 Pinning down $k$

The parameter $k$ indicates the number of coefficients $\beta$ and $\tau$ in either the cubic spline or the multiple exponential decay model. Once $k$ is pinning down, the dimension of a model specification is established. A model with one term only, for instance, fits just a straight line log discount function, whose fitting error with respect to an actual yield curve is likely to exhibit positive serial correlation. On the other hand, too many terms will bring excessive curvatures or humps to the fitted curves, which may in turn cause negative serial correlation to appear its error term. In reality, different days will give different shapes of term structures: some are pretty flat, some are upward-sloping, and still some have more curvature or humps. It thus may hurt to have either redundant or inadequate parameters in an interpolation model. Moreover, we always want a parsimonious model expression to fit the term structure at each point of time. Therefore, we want to find out how many coefficients/terms the target function needs in order to satisfy: (1) the cumulative error will be minimal (overall goodness of fit); (2) error terms should be white noise, without showing any serial correlation. Further, most people who employ this method would not want to conduct a test each time to pin down $k$; rather, they just want to apply a rule that usually works well.

The D-W test is one candidate to test for serial correlation in the errors. Figure 3.7 shows the results of D-W tests for real U.S. term structure on February 26, 2010 derived by the multiple exponential decay model and cubic spline model respectively. I depict a D-W statistic for each $k$ (as showed along the horizontal axis) from $k = 2$ until the D-W statistic exceeds the corresponding upper critical value. Exact 90% confidence interval and mean critical value\textsuperscript{28} are numerically simulated with the given regressor matrix

\textsuperscript{28}Lower, mean and upper critical values are at 5% one tail significance level.
Figure 3.7: D-W Test for $k$ in Estimating Real U.S. Term Structure on 02/26/2010

$\Psi(m)$ and for each value of $k$. They are depicted respectively on the graph. As seen from Figure 3.7, D-W has an upward-sloping trend as $k$ increases. The smallest $k$ whose D-W statistic exceeds its corresponding lower critical value gives the most parsimonious model specification free from serial correlation (at 5% one tale significance level) in error term. This gives an acceptable realization of $k$, or $k_{\text{min}}$. One could also pick the $k$ with D-W statistic closest to its mean critical value. This $k$, or $k_{\text{opt}}$, ensures the lowest probability in terms of serial correlation, but may lead to more regressors and parameters in the model. As seen from the figure, $k = 3$ for the multiple exponential decay model happens to be $k_{\text{min}}$ as well as the $k_{\text{opt}}$ on February 26, 2010. Further, one can pick the third
The value of \( k, k_{\text{max}} \), who has the highest D-W statistic which however is still below its corresponding upper critical value.

However, the D-W test assumes a normal distribution in the error sequence and does not detect higher-order autocorrelation. Since there is no presumption that samples are all drawn from a same normal distribution, the D-W test may not be applicable in each single case. Another candidate for serial correlation test is the Nonparametric Runs Test. It tests the mutual independence of a two-valued sequence, with the null hypothesis that elements in the sequence come in random order against the alternative that they do not. Specifically, a “run” of an error sequence in our problem is a sequence of consecutive elements above or below zero. Figure 3.8 shows the p-values of the Runs test with respect to each \( k \) assigned for both real and nominal U.S. term structures on February 26, 2010. Similar to the D-W test, the smallest \( k \) giving a p-value higher than 5% in Runs test leads to the most parsimoniously acceptable model, while the \( k \) associated with the highest p-value is least likely to have autocorrelation. The former \( k \) for the multiple exponential decay model in Figure 3.8 equals 3, while the latter is 4 or 5. In addition, any \( k \) smaller than 3 or larger than 6 would cause a positively or negatively autocorrelated error term in the multiple exponential decay model.

As for overall goodness of fit, BIC can serve as a standard of measurement. Figure 3.9 shows BIC for each \( k \) in real U.S. term structure fitted by two models respectively. In terms of the multiple exponential decay model, \( k = 3 \) give the lowest BICs and thus the best overall fit. This value of \( k \) coincides with the \( k \) implied by serial correlation test. I hence pin down \( k = 3 \) for the multiple exponential decay model in estimating the real term structure in Section 3.4.2.
I repeat this practice for each month in the recent 13 years and find a stylized fact that $k = \sqrt{n}$ in estimating real U.S. term structures, where $n$ is the number of effective outstanding TIPS securities. The empirical evidence is showed in Figure 3.10. $k_{\text{max}}$, $k_{\text{min}}$, and $k_{\text{opt}}$ in the figure respectively represent the maximal, minimal, and optimal value of $k$ pinned down in estimating real U.S. term structures during 04/1998–12/2010. The dashed line gives the value of $\sqrt{n}$, from which we can see that the line fits closely to $k_{\text{opt}}$ and stays between $k_{\text{max}}$ and $k_{\text{min}}$ most of the time.\(^{29}\) This is especially true for the

\(^{29}\)The rule did not always work. In some months or years, the market was so volatile that $k = \sqrt{n}$ turned out to be a value too small to fit the actual real term structure. The first half of 2008 when the financial turmoil dominated was an example.
multiple exponential decay model in Figure 3.10-(a). With this evidence, we are able to pin down \( k \) directly. The rule itself can also be adjusted seasonally.

What if \( k \) is picked arbitrarily rather than following a rule? In the extended Nelson and Siegel model, \( k \) is fixed at 4 under all circumstances. Although giving a good fit on February 26, 2010 for the real term structure as seen in Figure 3.6, it does not generally guarantee a satisfactory fit in terms of serial correlation for the nominal term structure. As for the multiple exponential decay model or semi-natural cubic spline model, assigning an incorrect value to \( k \), either too small or too large, would also result in fitting failure. Figure 3.11 illustrates fitted Real U.S. term structures on February 26, 2010 with respect to each of successive values of \( k \) from \( k = 2 \). When \( k = 2 \) or 7, the estimated yield
curves either hardly fit the observed data or extrapolate to unrealistic values at the long-end. However, if $k$ is pinned down optimally according to the aforementioned rules, the fitted curves trace each observed data closely from short-term to long and decay asymptotically—with all properties we expect from the model.

### 3.5 Constrained Estimation—An Extension

The estimated term structures offer macroeconomists an overview on the maturity structure of risk-free interest rates. They can use it to interpret monetary policies or macro fluctuations over a period of time. However, financial practitioners may instead need guidance for portfolio selection. Is my method able to meet this demand? The answer is yes if I modify the estimation methods for $\beta$ in Section 3.4.2.
I offer three solutions to help “pick and select” Treasury securities at any point of time. All three are based on the same idea: bounding high/low (outstanding) yields gives high/low (estimated) yield curves so that over-/under-priced securities are screened out from the pool. Agents who want to buy securities may thus keep an eye on a high yield curve to pick those underpriced securities only, and vice versa. Because methods in this part are largely concerned with price constraints, I classify them as Constrained Estimation. The estimation methods in last section, either Nonlinear Minimization or Iterative Linear Least Squares (ILLS), are hence classified as Unconstrained Estimation. Without loss of generality, I illustrate three constrained estimation methods by bounding high yields or underpriced securities which are measured by asked prices.
The first constrained estimation method, Selected Fit, actually follows the methods in the class of the Unconstrained Estimation but only based on a proportion of effective securities. Specifically, I first run an unconstrained estimation using all effective securities to obtain estimated prices $\hat{p}$; then select security $i$ such that $p_i \leq \hat{p}_i$ and run the unconstrained estimation again using these “underpriced” securities only. Although appearing to be easy and straightforward, this estimation ignores a part of market information and thus may suffer from the problem of biased estimation.

The second and third constrained estimation methods are basically modifications to the two methods in the class of the Unconstrained Estimation. The modified Nonlinear Minimization is a constrained minimization that imposes on the original minimization a constraint of nonnegative price deviations. For ILLS we cannot impose a constraint on least squares directly. Instead, I turn to quadratic programming by minimizing the sum of squares of log net prices deviation subject to a nonnegativity constraint. Similar to ILLS, I then iteratively estimate $\beta$ until certain termination condition is met. I call this estimation Iterative Constrained Quadratic Programming (ICQP, henceforth). The details of these constrained estimations can be found in Appendix C.

Figure 3.12 shows three fitted real term structures on the same day, derived from the same interpolation model, but estimated by different methods. It is easy to observe the difference between the Selected Fit and the other two, especially the difference in estimated forward rate curves. Since Constrained Nonlinear Minimization and ICQP avoid the bias problem, I prefer them to the Selected Fit. Further, because ICQP outperforms Constrained Nonlinear Minimization in terms of computational efficiency, I recommend it to financial practitioners in selecting portfolios.
Figure 3.12: Unconstrained v.s. Constrained Estimation

Real U.S. Term Structure on 02/26/2010 Estimated by Multiple Exponential Decay Model under Different Estimation Methods.

3.6 Conclusion

In general, there are three “new” features present in this chapter. First, this chapter proposes a new interpolation model, Multiple Exponential Decay, to fit both the real and nominal term structures of interest rates. Compared to previous interpolation models (either spline-based interpolation or Nelson and Siegel exponential decay models),
this new model is more parsimonious in the model’s specification, more adaptive to a variety of shapes associated with yield curves, and more efficient to smooth through idiosyncratic variations generated by individual securities. Second, this chapter introduces Constrained Estimation, a new estimation method, for portfolio management. Traditionally unconstrained estimation gives a standard term structure useful for understanding the general macroeconomics or business cycles; the constrained estimation however offers an alternative to serve distinct curve-fitting purposes. Third and also the most unique contribution is a new procedure implemented to optimally adjust the dimension of the interpolation functional form. The number of coefficients of my model is not fixed, but adjustable to the actual securities outstanding. In this chapter, I turn to BIC for overall goodness of fit and the D-W test and the Runs test for serial correlation in error terms. These statistical tools enable us to take into account validity, optimality, and parsimoniousness simultaneously in specifying a model. In addition, I extend the McCulloch cubic spline model and compare its fitted yield curves with that of my multiple exponential decay model. I also apply several Unconstrained Estimation methods to estimate the unknown coefficients in the model and conclude that the Iterative Linear Least Squares (ILLS) is the most effective and efficient technique.

Nonetheless, there is room to further improve this research. My models and estimation methods are able and efficient enough to fit yield curves on any frequency basis given complete information on all outstanding securities in a market. It hence must be of interest to study the time series of cumulative fitting errors generated by a certain combination of model and estimation method. Then I will try to build up a criterion to evaluate each combination and search the best one. Also, I can include more market
factors, such as duration, convexity, and volatility of a security, to design a more practical estimation method for practitioners.
Chapter 4: Testing the Misintermediation Hypothesis

4.1 Introduction

According to the misintermediation hypothesis, the traditional mismatching of asset and liability maturity structures leads to aggregate business fluctuations. The mismatch breaks the link between current plans for future demand and supply, suggesting the possibility that supply of current output will not match planned demand intertemporally. It then follows that an unexpected fall in real interest rates may be observed in the event of a recession, or an unexpected rise in a boom, as predicted by the heterogeneous capital model. The connection between unexpected changes in interest rates and business fluctuations is due to the fact that interest rates play a role as prices governing intertemporal trade for consumption and production. Therefore, in facing an unexpected excess supply/demand, an unanticipated fall/rise in real interest rates has to set in to clear the market.

McCulloch (1981) also predicts this pro-cyclic behavior in real interest rates. Unfortunately, he and other researchers in 1980s were unable to empirically test it simply because a market implying real interest rates was not available until 1997 when the first TIPS bond was issued and traded. We may alternatively turn to nominal interest rates for this study; however, it's hard to tell whether the unexpected changes in nominal
rates come from unanticipated changes in real rates or just pure changes in inflation. Nonetheless, a total of 13 years’ market data for real interest rates has been at hand so far. More importantly, the recent decade witnessed a serious downturn in the U.S. economy. An empirical investigation on the misintermediation theory thus becomes more necessary and interesting.

A naive way to carry out this empirical study would be simply to test the correlation between a series of surprises to a single long-term real interest rate (such as 10- or 30-year yields to maturity on TIPS) and a series of innovations to real GDP. As for the former series, however, I argue that a composite interest rate series holding all information over an entire term structure would be more reliably pro-cyclic than a single (or two) long-term interest rate(s). Given the estimated U.S. term structure of real interest rates, I hence construct a monthly composite series on cumulative excess return on synthetic real amortized loans. As for the series representing excess demand/supply of aggregate real output, I follow the approach in Longbrake (2008) to establish a series of innovations to factor utilization. Specifically, I build up a linear combination of two predicted error decomposition series from AR(3) models, estimated through adaptive least square (ALS), for both employment rate and capacity utilization rate series. As a proxy for unexpected output fluctuation, this series avoids the potential unit root problem that exists in GDP time series and is also positively correlated with innovations to GDP. As a comparison, I also report results obtained by using GDP related innovation

Even if we identify significant changes in nominal interest rates, the changes could stem from the expected inflation rate while real interest rates have no unanticipated change at all. It is even possible that the changes in inflation over-compensate for behavior of real rates so that real rates actually counter-cyclically, rather than pro-cyclically, fluctuate to output, which is opposite to the prediction of the misintermediation hypothesis.
series. Followed by the establishment of these series, I test their correlation respectively and report results later in this section.

4.2 Price of Amortized Loans

It is well known that interest rates are pro-cyclic. In fact, the NBER Business Cycle Dating Committee has previously used various interest rate series to help determine the peaks and troughs of a chronology of the U.S. business cycles. However, a yield curve is not always flat, or without systematic change in shape over business cycles. According to McCulloch (1977), this would cause ambiguous causality in that any single interest rate may either under- or over-expect unanticipated changes in interest rate depending on the actual shape of the yield curve. This fact hence calls for analysis not only on several raw interest rates, but over the entire term structure of interest rates. Therefore, the first step of this empirical study is to obtain a single variable to specify the entire term structure of real interest rates. I use “price of amortized loans” in this chapter.

The spot price for a consol is the present value of a perpetual stream of future output, i.e. coupon payments, as shown in Eq 4.1. It could serve as a proxy for the price of Period 3 consumption \( c_3 \) in terms of Period 2 consumption \( c_2 \) in the 3-period heterogeneous capital model. The real discount factor \( \delta_t(m) \) for maturity \( m \) at time \( t \) is estimated by the

---

31 Primarily long-term Treasury-Bond yields.
32 Kessel (1965) pointed out that short-term rates were relatively high (compared to long-term rate) about cyclical peaks and low about troughs.
33 If the market expects interest rates to rise immediately prior to an expansion, then the yield curve tends to be unusually upward sloping at cyclic troughs. At peaks, however, the yield curve is often found to be humped, followed by market anticipation of a fall in interest rates over the upcoming contraction. This fact, demonstrated by Kessel (1965), implies that market may either under- or over-anticipate cyclical changes in interest rates.
multiple exponential decay interpolation established in Chapter 3.

\[ P_t = \int_0^\infty \delta_t(s) \, ds. \]  
(4.1)

The yield to this perpetuity \( y_t^c \), by \( P_t = \int_0^\infty \exp(-y_t^c s) \, ds \), is just equal to the reciprocal of the consol price and further equivalent to

\[ y_t^c = \frac{1}{\int_0^\infty \delta_t(s) \, ds} = \frac{\int_0^\infty f_t(s) \delta_t(s) \, ds}{\int_0^\infty \delta_t(s) \, ds}, \]  
(4.2)

which can be viewed as a composite real interest rate because it takes into account the whole term structure at time \( t \).

However, we may not want to use this consol price series directly because it may require extrapolation outside observed maturities. If it requires thirty years for a consol to attenuate to one half of its previous value, implied by the estimated \( \delta_t(m) \), then the other half of its value falls beyond thirty years, which is in the extrapolated range. Obviously, putting significant weight on extrapolation causes deterioration in the reliability of the estimated consol price. This concern calls for a more conservative approach. I thus specify another variable, the truncated consol price, or more accurately an amortized loan price, to achieve the same goal.

The amortized loan price, \( P_t^a(m) \), is realized by replacing the integral upper bound \( \infty \) in Eq 4.1 with a long-term maturity \( m \):

\[ P_t^a(m) = \int_0^m \delta_t(s) \, ds, \]  
(4.3)

where \( m \) can be a constant number (e.g. 25 years) or the actual longest maturity at \( t \). Because of this modification, the yields to this amortized loan can no longer be solved

\[ ^{34} \text{The longest maturity we have observed in the historical Treasury market, as showed in Figure 4.2, was typically no longer than 30 years.} \]
by Eq\[4.2\] Instead, since
\[P^a_t(m) = \int_0^m \exp(-y^a_{t,m,s}) \, ds \frac{1}{y^a_{t,m}}[1 - \exp(-y^a_{t,m,m})],\] (4.4)
amortized loan yields \(y^a_{t,m}\) may be solved out numerically once \(P^a_t(m)\) is obtained through Eq\[4.3\] Numerical algorithms and computing programs are based on [Miranda and Fackler (2002)].

At a previous point of time \((t - \Delta t)\) with historical TIPS quotes, I can derive a forward price for an amortized loan with maturity \((m + \Delta t)\) starting from a certain point of time \(t\). As time moves forward to \(t\), I can then derive a spot price for another amortized loan with maturity \(m\) starting immediately at \(t\). The logarithm of the ratio between the spot price and discounted forward price is the excess return of the hypothetical amortized loan market\[35\]
\[\Delta p_t(\Delta t; m) = \log \left( \frac{\int_0^m \delta_t(s) \, ds}{\int_{\Delta t}^{m+\Delta t} \delta_{t-\Delta t}(s) \, ds} \right) \]
\[= \log \left( \int_0^m \delta_t(s) \, ds \right) - \log \left( \int_{\Delta t}^{m+\Delta t} \delta_{t-\Delta t}(s) \, ds \right) + \log (\delta_{t-\Delta t}(\Delta t))\] (4.5)
where \(p(\cdot)\) denotes logarithm of \(P^a(\cdot)\), and \(p^f(\cdot)\) represents a log forward price. Eq\[4.5\] implies the excess return in the log prices over a period of \(\Delta t\) because the last term in the last line of Eq\[4.5\] represents the risk-free interest rate over the same period of time. As a demonstration, Figure\[4.1\] shows the real discount function \(\delta_t(m)\) derived from TIPS market quotes when \(t = 09/30/2010\) as well as another real discount function \(\delta_{t-\Delta t}(m + \Delta t)\) 6 months prior to that date. The shaded area between \(\delta_t(m)\) and \(\delta_{t-\Delta t}(m + \Delta t)/\delta_{t-\Delta t}(\Delta t)\) implies the surprise in prices of the amortized loans I attempted to develop above.

\[35\]A more accurate estimate should adjust for term premia.
4.3 Proxies for Excess Supply/Demand

A large body of research has been working on identifying the deviation of output from its long-term trend, or unexpected output. A simplest way to do so is to estimate AR coefficients to time series of real GDP growth and examine the innovation series, as the approach proposed in [Stock and Watson (2005)]:

$$\Delta y_t = \alpha(L) \Delta y_{t-1} + \varepsilon_t$$  \hspace{1cm} (4.6)

where the regressor is quarterly real GDP growth rate and coefficients are estimated by a AR(4) model\[36\]. A problem associated with this approach, first suggested by [Perron](#).\[37\]

\[36\]The order of autoregression is determined by Akaike info criterion (AIC) and henceforth.
and further mentioned in Stock and Watson (2005), is the time-varying variance of the AR innovation over the past half century. Perron suggests a way to eliminate the heteroskedasticity in output time series by allowing breaks in the series and running regression separately. However, the arbitrary assignment of breaks degrades the general reliability of this method.

Another common approach is the detrending of the GDP time series by running a Hodrick-Prescott (HP) filter. The HP filter decomposes a time series into a growth component and cyclical component. A parameter $\lambda$, controlling over the smoothness of the growth component, is used to penalize variability in this component. However, according to Cogley and Nason (1995), HP filter may generate cycles in random walk series which actually has no cyclic behavior at all. In addition, the use of different arbitrary smoothing parameters generates a large difference in the outcome of the model, which turns out to be an apparent drawback.

A more serious concern in this part is the non-stationary nature of real GDP time series. In a recent dissertation research on output gap, Longbrake (2008) found that U.S. real GDP may follow a unit root process and hence the specification of output gap based on GDP series might be nothing, but a pure “statistical illusion”. Fortunately, he also proposes an alternative approach, Factor Utilization Gap, to substitute for the output gap.

The concept, factor utilization gap, is rooted in growth theory where output is assumed to be a function of labor, capital, and other factors. I assume for this purpose

---

37 Taking a random walk series with up-drift and running a HP filter on it, one may get cycles on the residuals on any maturity by adjusting $\lambda$ even though the original series has no cycle on it.
that the actual production function, taking Cobb-Douglas form, is

\[ Y = AL^\sigma K^{1-\sigma}, \]  

(4.7)

where \( L, K, \) and \( A \) are labor, capital, and TFP respectively; \( \sigma \) is labor share. I further assume a natural production function in which all of the inputs are operating at their natural rates, hence:

\[ \bar{Y} = AL^\sigma K^{1-\sigma}. \]  

(4.8)

After taking logarithms, the difference between Eq 4.7 and 4.8 is just equal to the sum of the labor employment and capital utilization gaps because TFP \( A \) is assumed to be fully utilized and thus crossed out in the log difference. I recognize this log difference as the output gap. Eq 4.9 shows that the decomposed output gap features a linear combination of the employment gap and the capacity utilization gap:

\[ y - \bar{y} = \sigma(l - \bar{l}) + (1 - \sigma)(k - \bar{k}), \]  

(4.9)

where lowercase letters denote logarithms. Longbrake (2008) names this linear combination the Factor Utilization Gap, as a proxy for the output gap derived from GDP series.

Following Longbrake (2008), I specify employment rates \( (n) \) and capacity utilization rates \( (c) \) both as AR(3) process\(^{38}\), which take the form:

\[ n_t = \alpha_{0,t} + \sum_{j=1}^{3} \alpha_{j,t} n_{t-j} + \varepsilon^n_t, \]  

(4.10)

\[ c_t = \beta_{0,t} + \sum_{j=1}^{3} \beta_{j,t} c_{t-j} + \varepsilon^c_t, \]  

(4.11)

\(^{38}\)Longbrake (2008) actually uses AR(4), rather than AR(3). The difference is due to different time span chosen. Following the same rule as in Longbrake (2008) to pin down the optimal lag periods in AR process, I find that the coefficient associated with AR(4) term is not significantly different from zero while AR(3) works better with the new data series I choose for this particular empirical study.
where the employment rate $n$ simply equals to (1 - unemployment rate) and the labor share $\sigma$ is set to 70% as suggested in Longbrake (2008). 

McCulloch (2005) recommends to use ALS in estimating time-vary coefficients in regression models Eq 4.10 and 4.11. ALS is a variant version of Kalman filter, allowing both filter and smoother estimates. I use filter ALS to estimate AR(3) models for both the employment rate and the capacity utilization rate series. A linear combination of two predicted error decomposition series, $\varepsilon_t^n$ and $\varepsilon_t^c$, can be established as:

$$\varepsilon_t = \sigma \varepsilon_t^n + (1 - \sigma) \varepsilon_t^c,$$

(4.12)

I call $\varepsilon_t$ the innovation to factor utilization, a time series proxy for innovation to real GDP growth for the correlation test in the next step. A detailed description on ALS can be found in McCulloch (2005) Appendix I.

### 4.4 Data and Results

Macro NIPA data collects from the Federal Reserve Database (FRED) St. Louis. Because the monthly real GDP data is not available, I use quarterly real GDP (GDPC96, 1947Q1—2010Q4, chained in 2005 dollars) to generate a series of innovations to real GDP growth (Quarter on Quarter). Following Wu (2010), I also use monthly real personal income to approximate monthly real GDP, which is obtained by deflating personal income (PI, 1959M1—2010M12) by the consumption price index (CPIAUCSL, 1959M1—2010M12). As for factor utilization, I compute (1-unemployment rate) (UNRATE, 1948M1—2010M12) for the labor employment rate, and use total index of capacity utilization

---

39Econometrical work in this part is conducted with a GAUSS program downloaded from Professor McCulloch’s website http://www.econ.ohio-state.edu/jhm/jhm.html.
(TCU, 1967M1—2010M12) for the capital employment rate. All time series are seasonally adjusted, and all monthly or quarterly rates are annualized.

Bond market quotes are obtained from Bloomberg system, supplemented by daily bond market data release in The Wall Street Journal online archives. In this research, I exclusively focus on TIPS market quotes, which can be traced back to January 1997; however, I only collect data from April 1998 because the first 30-year long-term TIPS was not issued and traded until then. In each calendar month, I pick the last business day having market transaction as the representative of that month. Therefore, I have a total of 153 observations on TIPS quotes from 1998M4 to 2010M12, during which the number of outstanding securities and the longest maturities are showed in Figure 4.2.

Figure 4.2: Evolution of TIPS Market from 02/28/1997 to 12/31/2010

*40* Thanks to China Everbright Bank Beijing headquarters for offering one month use of their Bloomberg service.
Figure 4.2 indicates that the longest maturities in TIPS have been varying between 22 years and 30 years over the last 13 years. The choice of the integral upper bound $m$ in Eq 4.5 would be subject to pros and cons of the three possible options as follows: (1) constant at $m = 20$, which miss TIPS data with maturities beyond 20 years; (2) constant at $m = 30$, which on the other hand relies partly on extrapolated data at the long-end; (3) a time-varying $m$ equals to the actual longest maturity in each month. Although Option (3) is free from the unfavorable features in (1) and (2), it introduces additional fluctuations in the composite series due to the variation in $m$. In Figure 4.3, both Option (1) and (2) illustrate similar trend in monthly amortized loan prices; the trend given by Option (3), however, fluctuates between that of (1) and (2) as time moves forward. Therefore, I pin down $m$ according to Option (2) ($m = 30$) in this chapter.

Figure 4.3: Amortized Loan Prices $P_t^a(m)$ derived by Constant and Time-varying $m$
According to Eq 4.5, I am hence able to formulate a series of 152 monthly excess returns in amortized loan ($\Delta p(\Delta t)$), in which $\Delta t$ equals to one month. Figure 4.4 illustrates this monthly series. It is expected to be negatively correlated with the proxy series for the aggregate excess supply/demand.

![Figure 4.4: Monthly $\Delta p(\Delta t)$ over 04/1998–12/2010](image)

Then, I formulate three alternative proxy series for the aggregate excess supply/demand: $\varepsilon^{GDP}$ (innovation to real GDP growth) by AR(3), $\varepsilon^{PI}$ (innovation to real personal income) by AR(6), both of which follow Eq 4.6, and $\varepsilon^{FU}$ (innovation to factor utilization) by a linear combination of two error decompositions in AR(3) as in Eq 4.10 and 4.11 and estimated by ALS. Table 4.1 lists descriptive statistics of these data series, including $\Delta p(\Delta t)$. The means of $\Delta p(\Delta t)$, both in the monthly and quarterly series, imply the net term premium associated with the estimated real term structure. The mean is positively

41 I truncate the real GDP series and run a AR regression for $\varepsilon^{GDP}$ using data only since 1986 suggested by Stock and Watson (2005). Therefore, there are only 100 observations in $\varepsilon^{GDP}$.
significant for the monthly series, yet insignificant for the quarterly series. McCulloch (1975a) suggests a constrained estimate of term premium, which supposes to give more precise estimate of the surprise in the amortized loan prices.

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon^{GDP}$</th>
<th>$\varepsilon^{PI}$</th>
<th>$\varepsilon^{FU}$</th>
<th>$\Delta p(\Delta t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>monthly</td>
<td>quarterly</td>
<td>monthly</td>
<td>quarterly</td>
</tr>
<tr>
<td>N</td>
<td>100</td>
<td>610</td>
<td>521</td>
<td>152</td>
</tr>
<tr>
<td>Mean</td>
<td>0.00</td>
<td>-0.00</td>
<td>-0.05</td>
<td>3.50</td>
</tr>
<tr>
<td>(s.e.)</td>
<td>(0.205)</td>
<td>(0.285)</td>
<td>(0.053)</td>
<td>(1.565)</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.05</td>
<td>7.03</td>
<td>1.21</td>
<td>19.23</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.33</td>
<td>-0.77</td>
<td>-0.59</td>
<td>-0.75</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.58</td>
<td>13.00</td>
<td>6.47</td>
<td>5.98</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>3.22</td>
<td>2598.95</td>
<td>292.40</td>
<td>69.88</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.20)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

Table 4.1: Descriptive Statistics for Target Time Series

In order to test the relationship between $\Delta p(\Delta t)$ and each of the three proxy series for the excess supply/demand, I run linear regressions for the monthly $\Delta p(\Delta t)$ with both $\varepsilon^{PI}$ and $\varepsilon^{FU}$, and for the quarterly $\Delta p(\Delta t)$ with $\varepsilon^{GDP}$ as follows

$$\Delta p(\Delta t) = \beta_0 + \beta_1 x + \epsilon$$  \hspace{1cm} (4.13)$$

where regressor $x$ is $\varepsilon^{GDP}$, $\varepsilon^{PI}$, or $\varepsilon^{FU}$. I report the results in Table 4.2. The significantly negative $\beta_1$ coefficient between each of $\varepsilon^{GDP}$ and $\varepsilon^{FU}$ and the dependent variable $\Delta p(\Delta t)$ suggests a negative correlation in these two pairs. This is in conformity with the misintermediation hypothesis, even though the Durbin-Watson test associated with $\varepsilon^{FU}$
Table 4.2: Linear Regressions for $\Delta p(\Delta t)$ and Three Proxy Series

<table>
<thead>
<tr>
<th></th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$R_t^2$</th>
<th>D.W. stat</th>
<th>J-B on $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon^{GDP}$</td>
<td>3.02</td>
<td>-1.55*</td>
<td>0.0624</td>
<td>1.790</td>
<td>1.594</td>
</tr>
<tr>
<td></td>
<td>(1.796)</td>
<td>(-2.080)</td>
<td>(0.356)</td>
<td>(0.45)</td>
<td></td>
</tr>
<tr>
<td>$\epsilon^{PI}$</td>
<td>3.95</td>
<td>0.29</td>
<td>0.0002</td>
<td>2.086</td>
<td>50.810**</td>
</tr>
<tr>
<td></td>
<td>(1.742)</td>
<td>(1.017)</td>
<td>(0.668)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>$\epsilon^{FU}$</td>
<td>3.01</td>
<td>-2.86*</td>
<td>0.0235</td>
<td>1.172**</td>
<td>42.868**</td>
</tr>
<tr>
<td></td>
<td>(1.926)</td>
<td>(-2.146)</td>
<td>(0.000)</td>
<td>(0.00)</td>
<td></td>
</tr>
</tbody>
</table>

Parentheses under $\beta$ give $t$-stats, while parentheses under D.W. and Jarque-Bera give respective $p$-value. *=5%. **=1%.

implies a serial correlation in the error term and the Jarque-Bera test rejects the iid normality assumption in the same error term. Although the positive $\beta_1$ in the regression between $\epsilon^{PI}$ and $\Delta p(\Delta t)$ is opposite to our expectation, it is not significant and thus fails to account for any meaningful relationship between the aggregate excess demand/supply and the real interest rate.

In short, results in Table 4.2 give empirical evidence that supports the theory I develop in Section 2.2.

4.5 Conclusion

The empirical evidence presented in this chapter supports the existence of pro-cyclic behavior in U.S. real interest rate surprises. The empirical study however has room for further improvement. Several issues in the statistical tests requires further investigation.

First, the assumption of iid normality in several innovation series is rejected according to Table 4.1 Except for $\epsilon^{GDP}$ and quarterly $\Delta p(\Delta t)$, all the other series reject the null
hypothesis in the Jarque-Bera test that the data are from a normal distribution since cor-
responding $p$-values are lower than a 5% significant level. This fact implies that most of
innovation or surprise series disobey the Gaussian White Noise assumption. It may im-
ply that the AR models are mis-specified, or I need to further look into the ARCH or
GARCH structure of the error term in AR models for a more correctly specified innova-
tion series. Further, a more frustrating fact is that the Gaussian normality assumption
is also rejected in the error term of the linear regression between $\Delta p(\Delta t)$ and $\varepsilon^F_U$ (Ta-
ble 4.2), which weakens the conclusion of the existence of a negative correlation in this
pair of time series.

Second, we observe a positive mean in monthly $\Delta p(\Delta t)$ in Table 4.1 which implies
an approximately 3.5% annual average excess return in amortized loan, a similar level
as 10-year TIPS yield. In other words, the expected return on zero-maturity is expected
to be zero. This observation reminds us to adjust a positive term premium in the real
term structure. It then further calls for a more detailed study on the shape of the term
premium associated with the real term structure.

Last but not the least, it would be interesting to investigate the lead-lag relationships
between the output innovation series and the excess return series. In short, a thorough
investigation on the empirical relationship between real output and real interest rate is
needed.
Bibliography


Appendix A: Model Timing and Production Optimization

The Production Possibility Frontiers (PPF) in Section 2.3 are pinned down according to the specification of production technologies in Eq 2.9–2.13. In the 3-period model, the timing proceeds as follows:

**Period 1** \((t = 1)\): A household enters the economy with endowments fields \(\bar{F}_1\) and plows \(\bar{p}_1\). It devotes labor \(L^h_1\) to harvest \(\bar{F}_1\) by

\[
h_1 = f(F_1, L^h_1) = \alpha F^n_1 L^{1-n}_1. \tag{A.1}
\]

The harvested crops \(h_1\) are allocated to consumption \(c_1\) and seed \(s_1\). The household then uses \(s_1\), the endowment of plows \(\bar{p}_1\), and labor \(L^f_1\) to produce fields \(F_2\) for harvesting in \(t = 2\):

\[
F_2 = g(s_1, L^f_1, \bar{p}_1) = (s_1^{\gamma_1} \bar{p}_1^{1-\gamma_1})^{\theta} L^{1-\theta}_1. \tag{A.2}
\]

The household also devotes labor \(L^p_1\) to produce plows for \(F_3\) production in Period 2:

\[
p_2 = L^p_1. \tag{A.3}
\]

In order to reach the production frontier, the household has to employ all the factors, including labor so that

\[
L = L^h_1 + L^f_1 + L^p_1. \tag{A.4}
\]
Period 2 ($t = 2$): The household first devotes labor $L^h_2$ to harvest $F_2$ according to

$$h_2 = f(F_2, L^h_2) = \alpha F^n_2 L^{h_1 - \eta}_2.$$  \hspace{1cm} (A.5)

$h_2$ is allocated to consumption $c_2$ and seed $s_2$. The household then uses $s_2$, plows $p_2$, and labor $L^f_2$ to produce fields $F_3$ for harvesting in $t = 3$:

$$F_3 = g(s_2, L^f_2, p_2) = (s_2^\gamma p_2^{1-\gamma})^\theta L^{f_1 - \theta}_2.$$ \hspace{1cm} (A.6)

The household does not produce any plows in this period because no planting production will take place in the next period ($t = 3$) as the household will then exit the economy. Therefore,

$$L = L^h_2 + L^f_2.$$ \hspace{1cm} (A.7)

Period 3 ($t = 3$): The household uses up all of its labor endowment ($L = L^h_3$) to harvest $F_3$:

$$h_3 = f(F_3, L^h_3) = \alpha F^n_3 L^{h_1 - \eta}_3.$$\hspace{1cm} (A.8)

It consumes $h_3$ ($h_3 = c_3$) and exits the economy.

A PPF over the $(c_1, c_2, c_3)$ space is a combination of optimal solutions $\{c_1, c_2, c_3\}$ provided that the households exploit all the factors and endowments. In order to pin down the ex-ante PPF, I first solve out $p_{max}$ and $c_{max}$ by

$$\max_{p_2} c_3 = \alpha F^n_3 L^{h_1 - \eta}_3$$

s.t. $F_3 = (s_2^\gamma p_2^{1-\gamma})^\theta L^{f_1 - \theta}_2$

$$L = L^h_3; \quad s_2 = h_2; \quad s_1 = h_1.$$ \hspace{1cm} (A.9)
Since $c_1^{max} = h_1 = f(\bar{F}_1, L_1^h = L)$ according to Eq\[A.1\] the $c_2$ ex-ante production frontier in Figure 2.11-(b) for each level of $c_1 \in [0, c_1^{max}]$ and $L_1^p = 0$ is determined by solving

$$\max_{L_1^h} c_2 = \alpha F_2^\eta L_2^{1-\eta}$$

$$\text{s.t. } F_2 = (s_2^\gamma p_1^{1-\gamma})^\theta L_1^{f1-\theta}$$

$$L = L_2^h; \quad L = L_1^h + L_1^f; \quad s_1 = h_1 - c_1$$

(A.10)

Then, the $c_3$ ex-ante production frontier in Figure 2.11-(b) is the optimal $c_3$ for each attainable $\{c_1, c_2\}$ combination:

$$\max_{\{L_1^h, L_1^f, L_2^f\}} c_3 = \alpha F_3^\eta L_3^{h1-\eta}$$

$$\text{s.t. } F_3 = (s_2^\gamma p_2^{1-\gamma})^\theta L_2^{f1-\theta}$$

$$L = L_2^h; \quad L = L_2^h + L_2^f; \quad L = L_1^h + L_1^f + L_1^p; \quad s_2 = h_2 - c_2; \quad s_1 = h_1 - c_1$$

$$L_1^p \in [0, p_{max}]; \quad c_1 \in [0, \alpha \bar{F}_1^\eta L_1^{1-\eta}]; \quad c_2 \in [0, \max\{\alpha F_2^\eta L_2^{h1-\eta}\}|c_1]$$

(A.11)

The parameters are pre-assigned as $\alpha = 5$, $\eta = .3$, $\gamma = .3$, $\theta = 4$, $\bar{F}_1 = 10$, $\bar{p}_1 = 1$, and $L = 10$. Jointly solving A.9–A.11 gives the ex-ante PPF over the $(c_1, c_2, c_3)$ space as showed in Figure 2.11-(b).

As for the ex-post PPFs, the intercept of $c_1$ is pinned down according to $c_1 = f(\bar{F}_1, L_1^h = L - L_1^p)$, where $L_1^p = p_2 \in (0, p_{max}]$. Solving A.10–A.11 again, with one more constraint $L_1^p = p_2$, returns the ex-post PPFs at different levels of $p_2 \in (0, p_{max}]$ as illustrated in Figure 2.11-(a). The optimal solutions for $L_1^h$, $L_2^h$, and $L_1^p$ in the ex-ante PPF problem help generate the real wage rate (Eq 2.14) and capital prices (Eq 2.16 and 2.17).
Appendix B: Data Preprocessing for Yield Curve Fitting

The real U.S. term structure is derived from TIPS market quotes, while nominal U.S. term structure is driven by T-Bills, T-Notes, and T-Bonds. Bond market quotes are collected from Bloomberg service and supplemented by The Wall Street Journal archive on its website.

Prior to fitting the term structure, the original data needs to be preprocessed. First, I convert all quoted fractional prices to decimal ones. As convention, Treasury security market in the U.S. quotes “clean prices”, which are not the actual prices in transactions. I hence compute “dirty prices” by adding in accrued interest. For the initial guess in the Iterative Linear Least Squares (ILLS), all quoted semiannually compounded yield-to-maturities \( B \) are then converted to continuously compounded yields \( R \) by:

\[
R = 2 \log \left( 1 + \frac{B}{2} \right)
\]  

(B.1)

The remaining work is to select effective securities for estimation. For TIPS, I exclude TIPS bonds with less than one year to maturity from the data set. This is because TIPS are indexed by inflation with a 2.5-month indexation lag. This makes their quoted yields behave erratically as they mature. In addition, there is strong seasonal fluctuation in the CPI, and hence a significant seasonal in short-term real rates, which I have not take into account. Once going beyond one year, we will not see the seasonal as much, and the
2.5-month lag is not as significant. For the conventional nominal Treasury securities, I basically employ all quoted T-Bills and T-Bonds, and most T-Notes with maturities larger than one year. However, if there are more than one security with an identical maturity, I will select the one(s) having the more (most) recent issue date. That one is usually more active and with smaller bid-asked price spread.
Appendix C: Estimation methods

In this appendix, I give details of the estimation methods in Section 3.4.2 and Section 3.5.

Nonlinear Minimization returns the estimated coefficient vector \( \hat{\beta} \) by minimizing the weighted sum of the squared deviations between the actual prices of Treasury securities outstanding and the estimated ones as follows:

\[
\min_{\{\beta_j\}_{j=1}^k} \sum_{i=1}^n \left( \frac{p_i - \hat{p}_i}{d_i} \right)^2
\]  

where 
\[
d_i = |p_{i,a} - p_{i,b}|
\]
\[
p_i = \frac{p_{i,a} + p_{i,b}}{2}
\]
\[
\hat{p}_i = \sum_{\ell=0}^{h_i} \frac{c_i}{2} \delta \left( \frac{\ell}{2} + m_i - \frac{h_i}{2} \right) + 100\delta(m_i)
\]  

\[
\delta(\cdot) = \exp \left( - \frac{1}{k} \sum_{j=1}^k \beta_j \Psi_j(\cdot) \right)
\]
\[
h_i = \lfloor 2m_i \rfloor
\]

where \( d_i \) is the bid-asked price spread\(^{42} \) \( m_i \) is the maturity for the security \( i \), \( c_i/2 \) is the security-specific semiannual coupon rate, and \( p_{i,b}, p_{i,a} \) and \( p_i \) are the observed bid, asked and mean price respectively. The unknown coefficients \( \beta_j \) enter the objective function

\(^{42}\)The market quotes do not report any bid-asked price spread smaller than 1/32. For those securities with equal bid and asked prices, I suppose the price spread is 1/64 instead.
through the discount function $\delta(\cdot)$ firstly, and then through predicted prices $\hat{p}_i$. The summation in $\hat{p}_i$ gives the present value of total coupon payments of a security from the transaction day to maturity. Since coupon is paid in every six months, each coupon payment actually has a maturity $\left(\frac{\ell}{2} + m_i - \frac{|2m_i|}{2}\right)$ for $\ell = 0, \ldots, |2m_i|$ and hence corresponds a new $\delta(\cdot)$. This set-up adds complexity to computation in that $\beta$ enters all $\delta(\cdot)$s.

An alternative approach is Iterative Linear Least Squares (ILLS), which greatly enhances the computational efficiency. The basic idea of this method goes as follows. If the coupon rates were all zero or the value of the coupons were all known, the estimation problem would be simply an exact fit or linear regression to log discount function. However, except for T-Bills, all Treasury securities are coupon-bearing securities. It hence is easier to solve the problem iteratively by using the last iteration’s discount function to evaluate the coupons before estimating $\beta$ of this iteration to price net of coupons in log. This can be initialized by using each bond’s yield-to-maturity, as given in Eq B.1, as a first guess to evaluate the value of the coupons. I give a brief description of this method as follows. For further details, please refer to McCulloch and Kochin (2000).

**Step 1:** Estimate $\beta^0$ by OLS where $y_i$ is the observed yield-to-maturity of the $i^{th}$ security;

$$m_i y_i = \sum_{j=1}^{k} \beta_j^0 \Psi_j (m_i) + \epsilon_i^0 \quad i = 1, \ldots, n \quad (C.3)$$

**Step 2:** Given $\beta^q$, evaluate the coupons and corresponding net prices by $\delta^q(\cdot)$;

$$\hat{p}_i^{\text{net}(q)} = p_i - \sum_{\ell} c_{i\ell} \exp \left( -\Psi \left( \frac{\ell}{2} + m - \frac{|2m|}{2} \right) \beta^q \right) = \delta^q (m_i) \quad (C.4)$$
Step 3: Estimate $\beta^{q+1}$ by OLS through the following regression:

$$- \log \left( \hat{p}^{\text{net}(q)}_i \right) = \sum_{j=1}^{k} \beta^{q+1}_j \Psi_j(m_i) + \varepsilon^{q+1}_i \quad i = 1, \ldots, n$$  \hspace{1cm} (C.5)

Step 4: Repeat Step 2 & 3 and compute the associated zero-coupon yield $\hat{y}_t^q$ in each iteration until $\max \left\{ |\hat{y}_t^{q+1} - \hat{y}_t^q| \right\} \leq \epsilon$, where $\epsilon = .001$ in this dissertation and $t = 1, \ldots, 481$ represents each month from 0 out to 40 years.\footnote{Uniformly extrapolate to 40 years.}

Constrained Nonlinear Estimation is similar to the (unconstrained) Nonlinear Estimation except for an extra price constraint. The problem is given as:

$$\min_{\{\beta_j\}_{j=1}^k} \sum_{i=1}^{n} \left( \frac{p_i - \hat{p}_i}{d_i} \right)^2$$  \hspace{1cm} \text{s.t.} \quad p_i - \hat{p}_i \geq 0 \quad \text{for} \quad i = 1, \ldots, n$$  \hspace{1cm} (C.6)

where all notations keep in line with the unconstrained Nonlinear Minimization.

The constrained version of ILLS is Iterative Constrained Quadratic Programming (ICQP). It minimizes the square of log net price deviation subject to the constraint that the deviation is nonnegative:

$$\min_{\beta} \frac{\varepsilon' \varepsilon}{2}$$  \hspace{1cm} \text{s.t.} \quad \varepsilon \geq 0$$  \hspace{1cm} (C.7)

where

$$\varepsilon_i = \log p^{\text{net}}_i - \log \hat{p}^{\text{net}}_i$$

$$= \log \left( p_i - \sum_{\ell} \frac{c_i}{2} \delta \left( \frac{\ell}{2} + m_i - \left\lfloor \frac{2m_i}{2} \right\rfloor \right) \right) - \log \left( 1 + \frac{c_i}{2} \right) + \Psi(m_i) \beta$$

The remaining steps are identical to ILLS.