

## ABSTRACT

KUO, CHUN-HUNG. Three Essays on Macroeconometrics. (Under the direction of Atsushi Inoue.)

Chapter 1 discusses the overview of the dissertation and highlights the importance of model misspecification issues in DSGE models. Methodology of approaching these issues is compactly discussed.

Chapter 2 proposes a new algorithm for solving computationally challenging heterogeneous agent (HA) models. The algorithm combines the value function iteration (VFI) method, the endogenous grid points method (EGM), and the explicit aggregation (XPA) method. This new algorithm is robust in the sense that the initial values are irrelevant for finding the optimal policy functions. Moreover, it is also fast because both the root-finding procedure for the household problem and the repeated simulation for the aggregate law of motion are not necessary. The algorithm solves a typical HA model within just a few seconds on an ordinary desktop computer. As a consequence, it could be a proper tool for structural estimation of HA models, which needs to solve the model several thousands times.

Currently, heterogeneous agent (HA) models still rely on calibration for model parameters, which makes the quantitative implications of these models shaky. Since HA models can trace both the dynamics of the cross-sectional consumption (and/or wealth) distributions and those of aggregate variables, a HA model is estimated by utilizing two sources of data: the first one is repeated cross-sectional data from the Consumer Expenditure Surveys (CEX), and the second one is the aggregate data from the National Income and Production Accounts (NIPA) of the United States. In particular, the Laplace-type estimator (LTE) with criterion function from indirect inference is adopted. Therefore, the daunting global optimization task is replaced by the Markov Chain Monte Carlo method. Therefore, Chapter 3 is methodologically oriented, and focuses on proposing a recipe for conducting the structural estimation of HA models, a task that has not yet been done in the literature.

In Chapter 4 we consider a dynamic stochastic general equilibrium (DSGE) model with distortions. Our framework allows us to analyze the quantitative importance of model misspecification and identify the location of model misspecification. To illustrate our framework, we estimate a New Keynesian DSGE model with distortions using US macroeconomic time series. We find that the DSGE model is largely misspecified, especially in the labor market. Specifically, the forecast error variance decomposition (FEVD) exercise reveals that, for all aggregate variables used to estimate the model, the labor market misspecification contributes to more than 40 percent of their total variations.

© Copyright 2012 by Chun-Hung Kuo

All Rights Reserved

Three Essays on Macroeconometrics

by  
Chun-Hung Kuo

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Economics

Raleigh, North Carolina  
2012

APPROVED BY:

---

Douglas K. Pearce

---

Moody T. Chu

---

Nora Traum

---

Atsushi Inoue  
Chair of Advisory Committee

## DEDICATION

To my wife, Hsiao-Fang.

## BIOGRAPHY

Chun-Hung Kuo is a PhD candidate at North Carolina State University. His main research area is applied econometrics and macroeconomics with focus on structural estimation, model misspecification, and related computational methods. He is also interested in wealth and income inequality, and numerical methods for solving dynamic stochastic general equilibrium (DSGE) models.

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my dissertation advisor, Dr. Atsushi Inoue, for his support, patience, and helpful criticism. I would not have been able to complete this dissertation without his step-by-step guidance. I would also like to thank Dr. Moody Chu for sharing his knowledge on numerical analysis. His continuous encouragement enabled me to put myself in a position to finish my dissertation. I would like to express my gratitude to Dr. Pablo Guerron for leading me into the world of dynamic macroeconomics. I am grateful towards Dr. Barbara Rossi for giving me the opportunity to work on a National Science Foundation project as one of my dissertation chapters. I would like to thank Dr. Douglas Pearce and Dr. Nora Traum for their valuable comments that improved my dissertation. At last, I would like to thank my graduate school colleagues, including Steve Tsang, Chien-Yu Huang, Timothy Hamilton, Robert Kane, and Wei Wei, for their friendship and endless support.

## TABLE OF CONTENTS

<b>List of Tables</b> . . . . .	<b>vii</b>
<b>List of Figures</b> . . . . .	<b>viii</b>
<b>Chapter 1 Overview of the Dissertation</b> . . . . .	<b>1</b>
1.1 Motivation . . . . .	1
1.2 The Methodology . . . . .	3
<b>Chapter 2 A New Algorithm for Solving Heterogeneous Agents Models</b> . . . . .	<b>6</b>
2.1 Introduction . . . . .	6
2.2 Literature for Algorithms . . . . .	8
2.3 A Heterogeneous Agents Model . . . . .	10
2.3.1 Model Setting . . . . .	11
2.3.2 Recursive Competitive Equilibrium . . . . .	13
2.4 The Algorithm . . . . .	17
2.4.1 Housekeeping . . . . .	17
2.4.2 Implementation of the Algorithm . . . . .	18
2.5 Numerical Results . . . . .	21
2.5.1 Calibration . . . . .	21
2.5.2 Some Preliminary Results . . . . .	22
2.5.3 Comparison with Existing Algorithms . . . . .	25
2.6 Conclusion . . . . .	28
<b>Chapter 3 Estimation of Heterogeneous Agent Models</b> . . . . .	<b>29</b>
3.1 Introduction . . . . .	29
3.2 Literature Review . . . . .	32
3.2.1 Review of Heterogeneous Agent Models . . . . .	32
3.2.2 Review of Empirical Studies . . . . .	33
3.2.3 Review of Estimation Methods . . . . .	35
3.3 The Model . . . . .	37
3.4 Empirical Methodology . . . . .	41
3.4.1 Some Technical Necessaries . . . . .	41
3.4.2 The Estimation Strategy . . . . .	44
3.5 The Data . . . . .	46
3.5.1 NIPA Data . . . . .	46
3.5.2 CEX Data . . . . .	47
3.6 Empirical Results . . . . .	49
3.7 Conclusion . . . . .	51
<b>Chapter 4 Identifying Sources of Misspecification in DSGE Models</b> . . . . .	<b>52</b>
4.1 Introduction . . . . .	52
4.2 The Model . . . . .	54



4.2.1	The Final Good Firm Problem . . . . .	56
4.2.2	The Intermediate Good Firm Problem . . . . .	57
4.2.3	The Household Problem . . . . .	62
4.2.4	Aggregation Issues . . . . .	68
4.2.5	Government Policy . . . . .	70
4.3	Empirical Methodology . . . . .	70
4.3.1	The Data . . . . .	70
4.3.2	Bayesian Inference . . . . .	71
4.3.3	Prior Settings . . . . .	72
4.4	Empirical Results . . . . .	78
4.4.1	Parameter Estimates . . . . .	78
4.4.2	Impulse Response Functions . . . . .	80
4.4.3	Forecast Error Variance Decompositions (FEVD) . . . . .	84
4.5	Simulation Exercises . . . . .	85
4.5.1	The Model . . . . .	87
4.5.2	Misspecified Models . . . . .	93
4.5.3	The Empirical Strategy and Results . . . . .	94
4.6	Conclusion . . . . .	97
<b>References . . . . .</b>		<b>98</b>
<b>Appendices . . . . .</b>		<b>104</b>
Appendix A Recursive Competitive Equilibrium (RCE) . . . . .		105
Appendix B Equilibrium Conditions . . . . .		107
B.1 Aggregation . . . . .		107
B.2 Steady State . . . . .		110
B.3 Linearized Equilibrium Conditions . . . . .		111
B.4 Misspecified Models for Simulation Exercises . . . . .		116

## LIST OF TABLES

Table 2.1	Parameters Calibration . . . . .	22
Table 2.2	The Euler Equation Residuals . . . . .	25
Table 2.3	Computation Time . . . . .	28
Table 3.1	Meaning of Control Variables . . . . .	49
Table 3.2	Prior Distributions of Structural Parameters . . . . .	50
Table 3.3	Posterior of Structural Parameters . . . . .	51
Table 4.1	Prior and Posterior for the Benchmark Case . . . . .	76
Table 4.2	Priors and Posteriors: Restricted Case . . . . .	77
Table 4.3	Variance Decomposition: Real GDP . . . . .	85
Table 4.4	Variance Decomposition: Real Consumption . . . . .	85
Table 4.5	Variance Decomposition: Real Investment . . . . .	85
Table 4.6	Variance Decomposition: Hours Worked . . . . .	86
Table 4.7	Variance Decomposition: (Quarterly) Inflation Rate . . . . .	86
Table 4.8	Variance Decomposition: Real Wage Rate . . . . .	86
Table 4.9	Variance Decomposition: (Quarterly) Federal Fund Rate . . . . .	86
Table 4.10	DGP and prior distributions . . . . .	95
Table 4.11	Forecast Error Variance Decomposition of Misspecified Model . . . . .	96

## LIST OF FIGURES

Figure 2.1	Individual Policy Functions of the Benchmark Calibration . . . . .	23
Figure 2.2	Simulated Value of Aggregate Capital of Benchmark Calibration . . . . .	24
Figure 4.1	Impulse Response Functions (Monetary Policy Shock) . . . . .	81
Figure 4.2	Impulse Response Functions (Technology Shock) . . . . .	83

# Chapter 1

## Overview of the Dissertation

### 1.1 Motivation

In this chapter, I provide an overview of my dissertation. Broadly speaking, the dissertation is about issues of model misspecification in the context of Dynamics Stochastic General Equilibrium (DSGE) models.

For an empirical economist, the first challenge facing her is usually model specification, before she conducts her econometric analysis. Often she has to answer followings questions: whether some important explanatory variables are excluded, whether the functional form of the regression model is appropriate, whether the assumption about the error term is reasonable, etc. Only when can she properly answer the above questions, her reader start to consider her empirical studies seriously. If she has a dubious model specification but adopts a fancy estimation method, the value of the empirical studies will still be limited. I do not mean that the choice of estimation method is not important; I merely want to to emphasize that the priority of model misspecification is higher than the estimation methods.

Econometricians recognize the consequences of adopting misspecified models, and they have studied the widely, especially in reduced-form regression models. However, studies about model misspecification with respect to DSGE models are still scarce. The scarcity reflects the difficulty

of the issue. Different from traditional reduced-form regression models, the structure of DSGE models is complex. DSGE models might contain dozens of non-linear equations and endogenous variables as well as several exogenous shocks. Their complicated structures make the theoretical econometric investigation difficult, if not impossible.

After three decades' development, economists have invented many DSGE models. One of the main reasons for inventing a DSGE model is to provide an explanation for various economic phenomena. Broadly speaking, when an economist constructs a new model to explain certain aspects of the data, she is providing a new model specification. If the new model can explain that particular aspect of the data, the economist tends to claim that the new model is better than the old one, and suggests, at least implicitly, that the old model should be discarded. This procedure is apt for choosing a better model, if the sole goal of research is providing economic explanation.

However, economists construct their models not just for positive analysis (e.g. what causes the high risk premium). We can also use a DSGE model for normative analysis (e.g. asking what the optimal monetary policy is) as well as for economic forecast. Under these circumstances, the procedure helps less, since we just want to use the model. In contrast, understanding the shortcoming of the model at hand is informative for the user of the model. I do not intend to discredit the above procedure for selecting a better model. I just mean that understanding the limitation of a considered model is informative for some purposes.

As Woodford (2009) argues, both the New classical and the New Keynesian economists now embrace DSGE models, and the DSGE models have become the standard framework for macroeconomic studies. A common feature of existing studies using DSGE models is that economists seek for not only qualitative but also quantitative implications. The requirement for quantitative implications makes parameter estimation inevitable. Along with the rapid development of DSGE models, econometricians have proposed various estimation methods for DSGE models. The existing methods for estimating DSGE models consist of the Generalized Method of Moments (GMM), the Maximum Likelihood Estimation (MLE) method, the Bayesian estimation method, some simulation-based estimation methods, etc.

Now many empirical studies are based the estimated DSGE models such as Christiano et al. (2005), Smets and Wouters (2007), and Justiniano and Primiceri (2008), just name a few. However, to best of my knowledge, when estimating DSGE models, economists have not yet tackled the obvious model misspecification due to the representative agent (RA) assumption. The main justification of the RA assumption is that individuals can fully insure their risks. That is, the markets should be complete. Unfortunately, complete markets are usually rejected in empirical studies. For example, see Cochrane (1991) and Attanasio and Browning (1995). Since RA setting is hard to justify, a serious treatment on model misspecification due to the RA assumption is necessary.

Therefore, in this dissertation I would like to deal with two model misspecification issues in DSGE models. First, I deal with misspecification due to the RA assumption, and estimate the structural parameters of a HA model. Second, I propose a framework for identifying the sources of model misspecification in DSGE models.

## 1.2 The Methodology

To deal with the two misspecification problem raised in the previous section, several econometrics methods are used in the dissertation. I briefly discuss my methodology of using these econometric methods in the dissertation in turn.

When economists estimate RA models, they often rely on likelihood-function-based methods. For example, McGrattan et al. (1997), and Ireland (2001) use MLE, and Smets and Wouters (2007) use the Bayesian estimation method. However, economists still do not know how to express HA models to the state space form, and thus we cannot adopt the likelihood-function-based methods (i.e., MLE and Bayesian method). Moreover, we cannot use GMM to estimate a HA model as well, because the first order conditions of HA models contain both aggregate and individual variables. It is not straightforward to use them together for estimation.

To achieve the goal of estimating a HA model, I combine two estimation methods that are both computational intensive: Indirect Inference and Laplace Type Estimator (LTE). The

main idea of Indirect Inference is to use an auxiliary model to capture the important aspects of observed data. If one can simulate the structural model, then one can also use the simulated data to obtain auxiliary parameter estimates. Indirect Inference chooses structural parameters in such a way that the estimated auxiliary parameters from both observed and simulated data are as close as possible. Indirect Inference is particular useful when the likelihood function of the data is not tractable. It is exactly the situation that we are facing.

The reason of introducing LTE is to overcome the difficulty coming from global optimization problem, if we solely use the Indirect Inference. Chernozhukov and Hong (2003) show that LTE has a computational advantage in dealing with cases where the criterion functions have multiple local modes or are highly non-convex, while the global optimum are well defined. The logic of LTE is to transform a criterion function to a well-defined density function by a Laplace type transformation. The criterion function does not have to be a likelihood function. In my implementation, the needed criterion function stems from indirect inference. Once the deduced density function, or quasi-posterior function, is obtained, we can use the standard Markov Chain Monte Carlo (MCMC) method to conduct random draws with respect to it. Besides, to implement the above hybrid estimation method, I need a algorithm to solve the HA model. I do not use the existing algorithms. Rather, I propose a new algorithm that is fast and robust. Please refer to the next chapter for the detail of the algorithm.

As I explained in previous the section, knowing the limitation of a models is useful for the user of models. However, macroeconomists are often left to wonder if their models are misspecified and, if they are, what parts of models are misspecified, even though model misspecification may be widespread. I propose an empirical framework for this issue. The framework is easy to use, and only the Bayesian estimation method is involved. To illustrate the framework, I consider a New Keynesian DSGE model based on Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007). The model is mildly misspecified in the sense that (i) the cross equation and equilibrium restrictions imposed by the model do not exactly hold in every time period; and (ii) these deviations are zero on average. I assume that the economic agents in the

model (both firms and households) take into account exogenous stochastic processes of deviations when they solve their optimization problems. The variances of deviations are a measure of the degree of misspecification of the New Keynesian DSGE model without a distortion.

After estimating the parameters of the models as well as those of deviation processes, I conduct forecast error variance decompositions (FEVD) to locate possible sources of misspecification. This method is closely related to Chari et al. (2007) who also introduce time-varying “wedges” into a macroeconomic model. There are two main differences. One difference is that they consider a stochastic growth model with wedges (what they call the benchmark prototype economy) while we consider a New Keynesian DSGE model with wedges. Our New Keynesian DSGE model is based on Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007) and incorporates frictions. Thus the distortions reflect model misspecification that cannot be accounted for by the built-in frictions. The other difference is that their analysis is based on calibrating parameter values while mine is based on estimated parameters that take into account possible misspecification.



## Chapter 2

# A New Algorithm for Solving Heterogeneous Agents Models

### 2.1 Introduction

In the literature, many algorithms have already been proposed to solve heterogeneous agent models. Each of them has its own limitations in terms of computation time, numerical instabilities for different parameter combinations, or difficulties in coding. In this chapter, I take these limitations into account and provide a new algorithm that drastically improves the speed and reliability for solving heterogeneous agent models. More importantly, this new algorithm is particularly useful for structural estimation of heterogeneous agent models. The algorithm has three key components: Value Function Iteration (hereafter, VFI), <sup>1</sup> Endogenous Grid points method (hereafter, EGM) of Carroll (2006), and explicit aggregation (hereafter, Xpa) of den Haan and Rendahl (2010).

The basic idea of VFI is follows. One makes an initial guess of the value function of the problem. With the initial guess, one solves the right hand side of the Bellman equation and updates the value function until it converges. The main strength of VFI is that it relies on the

---

<sup>1</sup>Value function iteration a popular algorithm for solving economics model globally, and Aruoba et al. (2006) provide a concise description for it.

contraction mapping theorem and, hence, possesses excellent convergence properties. In contrast, Euler-equation-based methods, which rely on the first-order conditions of the recursive problems to find the solution, cannot assure convergence unless one has good initial conjectures for the solution. Moreover, VFI is suitable to handle the problems with discontinuities or nondifferentiability problems coming from borrowing limits and discrete choices. Euler-equation-based methods usually run into difficulties in these situations. The main drawback of VFI is its speed, and it suffers from a strong case of the curse of dimensionality when the dimensions of state space increase. In the literature, there are many variants of VFI, including discretized VFI, parametrized VFI, or VFI with interpolation. The differences between them are the methods used to keep the information of the value function. Heer and Maussner (2009) give very detail explanations for different versions of value function iteration.

The EGM of Carroll (2006) is an important contribution on solving dynamic economic models.<sup>2</sup> This method changes the time convention of state variables and, by doing so, prevents the intensive computation of the root-finding procedure, which is the most time-consuming part of VFI. Details of the algorithm will be discussed in Section 2.4. By including the idea of EGM, VFI becomes a very attractive algorithm, because one can enjoy the strength of both algorithms, i.e., the convergence property of VFI and the speed improvement of EGM. Barillas and Fernandez-Villaverde (2007) demonstrate how to combine these two algorithms using stochastic neoclassical growth models as examples.

Finally, by including the methodology of Xpa, the new algorithm can obtain the aggregate law of motion of the model by a simple weighted average. Xpa helps one prevent the simulation step of the Krusell-Smith algorithm, which is computationally expensive and prone to numerical instabilities. The combination of VFI, EGM and Xpa makes the new algorithm a reliable and fast tool for solving heterogeneous agent models. In fact, it solves the heterogeneous agent models in just a few seconds. Section 2.2 provides a brief comparison between the new algorithm and existing ones.

---

<sup>2</sup>EGM can be used in many economics models, not just for heterogeneous agent models.

## 2.2 Literature for Algorithms

Krusell and Smith (1998) remains the standard algorithm for solving heterogeneous agent models.<sup>3</sup> The goal of Krusell and Smith (1998) is to find an aggregate law of motion for the aggregate variable that is consistent with the decision on individual capital accumulation. To accomplish this, one needs to first guess an aggregate law of motion in some parametric way. Given this conjecture on the aggregate law of motion, one then solves the individual recursive problem to obtain the policy functions. Next, one uses the deduced policy functions to simulate an economy with a large number of individuals. Not surprisingly, this is done by taking into account both idiosyncratic and aggregate shocks. By simulating the economy this way, the evolution of aggregate capital can be successfully traced out. Using the dynamics of the aggregate capital, researchers can update the aggregate law of motion by simple regression. With the new aggregate law of motion, one can resolve the recursive problem until the implied aggregate law of motion converges.

The structure of the algorithm is easy to understand. However, some of the numerical tools being adopted cause the algorithm to run at a very sluggish pace. This can be explained along these lines. First, they use the primitive value function iteration, which usually performs poorly when the dimension of the state space increases. In their model, the state space is four dimensional. Further, they use cubic splines on the individual capital dimension and polynomial interpolation on the aggregate capital dimension to interpolate the value function off the grid points. This two dimensional interpolation consumes a lot of time. Moreover, as the typical value function iteration, the algorithm has to maximize the right hand side of the Bellman equation for each grid point of the state space. This type of numerical maximization usually takes time and easily breaks down. The more critical problem is that the algorithm resorts to simulation to update the aggregate law of motion. Needless to say, simultaneously simulating the behavior of a large number of individuals takes a long time. Finally, their method needs a good assumption about the initial distribution of the cross-sectional capital distribution.

---

<sup>3</sup>Lucas (2003) says that the algorithm of Krusell and Smith (1998) works like a dream.

Young (2010) provides a slight modification of the Krusell-Smith algorithm. The main ideas are follows. He first implements the VFI given a perception of the aggregate law of motion. This step is exactly the same as that implemented in the Krusell-Smith algorithm. Next, rather than using simulation of a huge amount of individuals, he constructs a histogram to represent the density function of the cross-sectional distribution of individuals' capital holdings. Then, for each grid point on the histogram, he assigns proper probability mass for the histogram of next period through the individual policy functions and the transition matrices of aggregate and individual shocks. This method deals with the density function directly and, hence, does not rely on the law of large numbers. Therefore, this method does not suffer from the sample error of Monte Carlo simulation.<sup>4</sup>

Algan et al. (2008) also provide an algorithm to solve heterogeneous agent models. They use an exponential function to parametrize the cross-sectional distribution of individual capital. With the parametrized distribution functions, they calculate some key moments of the distribution. Furthermore, they provide a parametrized law of motion for those moments, which is in essence similar to the setting of Krusell and Smith (1998). The main innovation of the algorithm is that it does not rely on simulation. All that needs to be done is to provide flexible forms for the cross-sectional distribution and the aggregate law of motion. They then solve the recursive problem many times until the aggregate law of motion converges. Even though the structure of this algorithm is simple, it is very slow in terms of computational speed. This is due to the fact that a precise mapping between key moments and the parameters of the cross-sectional distribution needs to be constructed for it to work. In order to back out this mapping, one needs to solve the given optimization problem. I will show the speed of this algorithm in Section 2.5.

den Haan and Rendahl (2010) make a significant contribution to solving heterogeneous agent models. They find that the aggregate law of motion of the model can be obtained by explicitly aggregating the individual policy functions. Their explicit aggregate law of motion is constructed using the weighted average of individual policy functions. This method does not

---

<sup>4</sup>For the problem of sampling error, readers are referred to Algan et al. (2008).

rely on any simulation, and, thus, is very fast. This algorithm is actually one of building blocks that is incorporated into the algorithm I propose. I will discuss the limitations of their algorithm in section 2.5.

All of the above algorithms are known as global methods. Global methods simply mean that one solves for the policy functions in the whole state space. These methods are suitable for cases in which the policy functions have kinks. If the policy function is smooth (i.e., without kinks), researchers can resort to local methods such as higher order perturbation methods, which are widely used in the representative agent DSGE models. Directly using the perturbation methods in heterogeneous agent models is impossible, since the classical setting of heterogeneous agent models assumes that agents face a potentially binding borrowing constraint that will result in a set of kinked policy functions. In the current literature, some researchers deal with this problem by modifying the traditional heterogeneous agent models and imposing a penalty function on the utility function. The penalty function is constructed such that when individuals decide to accumulate negative capital, their utility level will become very low. By doing so, the individuals will never choose negative individual capital. Examples of using perturbation methods are Preston and Roca (2007) and Kim et al. (2010). These two algorithms are very fast, but they do not really solve the original heterogeneous agent models with borrowing constraints. From the view point of constructing models to mimic reality, the setting of Preston and Roca (2007) and Kim et al. (2010) has nothing wrong. Their models just have different interpretations from the traditional Krusell-Smith model. However, the use of perturbation methods make their algorithms infeasible for the models with discrete choices.

## 2.3 A Heterogeneous Agents Model

Since one of the main purposes of this chapter is to provide a new algorithm for solving incomplete market heterogeneous agents models, I provide below a model to demonstrate the algorithm.

### 2.3.1 Model Setting

The model is a modification of Krusell and Smith (1998) with the following characteristics. First, there are *ex ante* identical individuals of measure one. Second, the asset markets are incomplete, and precisely, there exists only one asset, say capital, to help individuals insure themselves from the underlining uncertainty of the economy. Moreover, individuals are not allowed to hold negative assets. In other words, individuals face an occasionally binding borrowing constraint. Third, the uncertainty of the economy originates from two sources: Aggregate productivity shocks and idiosyncratic labor efficiency shocks. Because of the presence of these two shocks, individual incomes are uncertain, and therefore, agents need to be forward-looking when making consumption and investment decisions.

The differences between Krusell and Smith (1998) and the current model are as follows. Krusell and Smith (1998) allow only two discrete states for both aggregate shocks and idiosyncratic shocks. Precisely, aggregate shocks take two states, either boom or recession; idiosyncratic shocks also take two states, either employed or unemployed. This relatively restrictive setting might be due to computational considerations. On the other hand, I model the exogenous shock process with an AR(1) process, which is more flexible than that of Krusell and Smith (1998). Moreover, I focus on the implementation of a new algorithm, I assume that there is no government sector in the economy, which is also different from the model of den Haan et al. (2010). Extending the current model to incorporate a government sector is easy and straightforward.

I lay out the problems facing individuals and firms and the forcing processes of aggregate and idiosyncratic shocks in turn. For extended details on this type of model, readers should refer to Krusell and Smith (1998), Krusell and Smith (2006) and den Haan et al. (2010).

**Individuals Problem.** Individuals are assumed to live forever. Individuals do not value leisure, and hence, the labor supply is inelastic. Individuals solve a life-time expected utility

maximization problem: <sup>5</sup>

$$\begin{aligned} \max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t), \\ \text{subject to} \\ c_t + k_{t+1} = (1 + r_t - \delta)k_t + w_t e_t, \\ k_{t+1} \geq 0, \\ k_0 \text{ given,} \end{aligned}$$

where  $c_t$  denotes consumption,  $k_t$  stands for individual capital holding,  $\beta$  defines the subjective discount factor, and  $\delta$  refers to the depreciation rate of individual capital. Since individuals cannot hold negative assets, the non-negative constraint is  $k_{t+1} \geq 0$ .  $u(\cdot)$  is the periodic utility function and takes the usual constant relative risk averse (CRRA) form:

$$u(c) = \begin{cases} \frac{c^{1-\gamma}}{1-\gamma} & \text{if } \gamma \neq 1, \\ \ln(c) & \text{otherwise,} \end{cases}$$

where  $\gamma$  is the relative risk aversion coefficient. Interest rate  $r_t$  and wage rate  $w_t$  are determined by the firm's profit maximization.

**Firms Problem.** The firm's production function is the typical Cobb-Douglas one:

$$Y_t = f(z_t, K_t, L_t) = z_t K_t^\alpha L_t^{1-\alpha}$$

where  $z_t$  is the technology level,  $K_t$  designates the aggregate capital input,  $L_t$  is the aggregate labor input, and  $\alpha$  indicates the capital share. The firm maximizes profit in each period and

---

<sup>5</sup>For simplicity of notation, I drop the superscript  $i$ , since the structure of the problem is the same for each individual in the model economy.

state, implying that the following two necessary conditions have to be satisfied:

$$\begin{aligned} r_t &= \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha}, \\ w_t &= (1-\alpha) z_t K_t^\alpha L^{-\alpha}. \end{aligned}$$

**Forcing Process** The aggregate technology shock is represented by an AR(1) process:

$$\ln z_{t+1} = \rho_z \ln z_t + \sigma_z \sqrt{1 - \rho_z^2} \epsilon_{z,t+1}, \quad \epsilon_z \stackrel{iid}{\sim} N(0, 1),$$

where  $\sigma_z$  is the standard deviation of logarithm of technology level and  $\rho_z$  denotes the persistence coefficient of  $\ln z_t$ . Similarly, the idiosyncratic efficient labor shock is parametrized as

$$\ln e_{t+1} = \rho_e \ln e_t + \sigma_e \sqrt{1 - \rho_e^2} \epsilon_{e,t+1}, \quad \epsilon_e \stackrel{iid}{\sim} N(0, 1),$$

where  $\sigma_e$  is the standard deviation of the logarithm of efficient labor, and  $\rho_e$  is the persistence coefficient of  $\ln e_t$ . There are two reasons for this parametrization. First, AR(1) processes can easily be approximated by a first order discrete Markov chain, which is easy to implement (see, Tauchen (1986), and Tauchen and Hussey (1991).) Moreover, one can test the performance of the algorithm by increasing the number of discrete states for both shocks. Second, this parametrization summarizes the forcing process of just two parameters, and therefore, is suitable for our estimation task.

### 2.3.2 Recursive Competitive Equilibrium

As shown in the last subsection, the interest rate and wage rate are functions of aggregate variables such as aggregate capital and aggregate labor, and technology level. Hence, in order to know the prices of the next period, which are necessary for conducting intertemporal decisions, individuals need knowledge about next period aggregate variables. In the representative agent framework, the determination of aggregate capital of the next period is trivial, since ag-



gregate capital is exactly individual capital. That is, the capital accumulation decision of the representative agent affects next period prices directly.

However, in the heterogeneous agents framework the determination of prices is more complicated, since the aggregate capital is determined by overall behaviors of heterogeneous agents. In other words, to know current prices, one needs to know the whole current cross-sectional distribution of capital, likewise for the next period prices. Hence, a proper recursive representation of the model should include the whole cross-sectional distribution of individual capital holding into the state space. Krusell and Smith (2006) provide a detail explanation of including the cross-sectional distribution and how the distribution can be modeled by a probability measure. The recursive representation of the model in the previous subsection is:

$$\begin{aligned}
V(k_t, e_t; \Gamma_t, z_t) &= \max_{k_{t+1} \geq 0} \left\{ \frac{c_t^{1-\gamma} - 1}{1-\gamma} + V(k_{t+1}, e_{t+1}; \Gamma_{t+1}, z_{t+1}) \right\}, \\
&\text{subject to} \\
c_t + k_{t+1} &= (1 + r_t - \delta)k_t + w_t e_t, \\
r_t &= \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha}, \\
w_t &= (1 - \alpha) z_t K_t^\alpha L_t^{-\alpha}, \\
\ln z_{t+1} &= \rho_z \ln z_t + \sigma_z \sqrt{1 - \rho_z^2} \epsilon_{z,t+1}, \quad \epsilon_z \stackrel{iid}{\sim} N(0, 1), \\
\ln e_{t+1} &= \rho_e \ln e_t + \sigma_e \sqrt{1 - \rho_e^2} \epsilon_{e,t+1}, \quad \epsilon_e \stackrel{iid}{\sim} N(0, 1).
\end{aligned}$$

With this recursive representation, one can define a recursive competitive equilibrium as follows. A recursive competitive equilibrium is a set of functions, containing the value function, the policy functions and a transition function for the measure of cross-sectional capital holding, such that the following conditions hold:

1. Individuals optimization: Value function  $V(k, e; \Gamma, z)$  solves the Bellman equation and  $g^k(k, e; \Gamma, z)$  is the associated policy function.
2. Firms optimization: Since the firm is representative and faces competitive factor markets,

the pricing functions must satisfy the following marginal productivity conditions

$$\begin{aligned} r_t &= \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha}, \\ w_t &= (1-\alpha) z_t K_t^\alpha L_t^{-\alpha}, \end{aligned}$$

where

$$\begin{aligned} K_t &= \int_{\mathcal{S}} k_t d\Gamma_t, \\ L_t &= \int_{\mathcal{S}} e_t d\Gamma_t. \end{aligned}$$

3. Consistency Condition: the dynamics of the measure  $\Gamma_t$  is governed by

$$\Gamma_{t+1} = H(\Gamma_t, z_t).$$

This transition function is consistent with the individual policy function  $g^k(k_t, e_t; \Gamma_t, z_t)$ .

Unfortunately, this equilibrium definition is not an operatable one, since the measure  $\Gamma$  cannot be implemented in computer because it is an infinite dimensional object. Krusell and Smith (1998) show that even the exact aggregation does not hold in the heterogeneous agent model,<sup>6</sup> but the approximate aggregation does hold. In the heterogeneous agent models (at least in that of Krusell and Smith (1998)), the decision rule is almost linear across individuals' wealth. Hence, holding the mean of individuals' wealth constant, redistribution of capital does not affect the overall capital accumulation. Thus, only the mean of cross-sectional capital distribution affects the mean of next period's mean capital. To understand the next period aggregate capital, which affects next period's prices, current capital is sufficient. Hence, we can just use average capital  $K_t$  to replace the whole measure  $\Gamma_t$ . The respective recursive problem can be rewritten as

---

<sup>6</sup>In economics, aggregation means that the decision rule is linear in the wealth level. Therefore, when all of the individuals have the same decision rule, then the redistribution of wealth level do not affect overall accumulation of wealth.

follows.

$$V(k_t, e_t; K_t, z_t) = \max_{k_{t+1} \geq 0} \left\{ \frac{c_t^{1-\gamma} - 1}{1-\gamma} + V(k_{t+1}, e_{t+1}; K_{t+1}, z_{t+1}) \right\}, \quad (2.1)$$

subject to

$$c_t + k_{t+1} = (1 + r_t - \delta)k_t + w_t e_t \equiv y_t, \quad (2.2)$$

$$r_t = \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha}, \quad (2.3)$$

$$w_t = (1 - \alpha) z_t K_t^\alpha L_t^{-\alpha}, \quad (2.4)$$

$$\ln z_{t+1} = \rho_z \ln z_t + \sigma_z \sqrt{1 - \rho_z^2} \epsilon_{z,t+1}, \quad \epsilon_z \stackrel{iid}{\sim} N(0, 1), \quad (2.5)$$

$$\ln e_{t+1} = \rho_e \ln e_t + \sigma_e \sqrt{1 - \rho_e^2} \epsilon_{e,t+1}, \quad \epsilon_e \stackrel{iid}{\sim} N(0, 1). \quad (2.6)$$

The definition of recursive competitive equilibrium under approximate aggregation is as follows.

1. Individuals optimization: Value function  $V(k, e; K, z)$  solves the Bellman equation and  $g^k(k, e; K, z)$  is the associated policy function.
2. Firms' optimization: The pricing functions satisfy the following marginal productivity conditions

$$r_t = \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha},$$

$$w_t = (1 - \alpha) z_t K_t^\alpha L_t^{-\alpha}.$$

3. Consistency Condition: Aggregate capital is governed by

$$K_{t+1} = H^K(K_t, z_t),$$

which is consistent with the individual policy function  $g^k(k, e; K, z)$ .

Krusell and Smith (2006) provide an extensive explanation and the basis of approximated

aggregation property. Young (2005) examines the robustness of approximated aggregation in detail.

## 2.4 The Algorithm

### 2.4.1 Housekeeping

The model can be summarized by Eq.(2.1), (2.2), (2.3), (2.4), (2.5), and (2.6), where  $y_t$  is the resource on hand in period  $t$ . Before we proceed to the details of the algorithm, there are some points needed to be discussed in advance.

First, I have already imposed the approximated aggregation property on the above problem. That is, I assume that aggregate capital in period  $t+1$ ,  $K_{t+1}$ , can be predicted by the aggregate capital in period  $t$ ,  $K_t$ , and the technology level  $z_t$ . Second, the expectation term of the above Bellman equation can be rewritten as:

$$\tilde{V}(k_{t+1}, e_t; K_t, z_t) = \beta EV(k_{t+1}, e_{t+1}; K_{t+1}, z_{t+1}),$$

since  $e_{t+1}$  is a function of  $e_t$ ,  $z_{t+1}$  is a function of  $z_t$ , and  $K_{t+1}$  is a function of  $K_t$  and  $z_t$ . Thus, the objective function of the above problem can be restated as:

$$V(k_t, e_t; K_t, z_t) = \max_{k_t \geq 0} \left\{ \frac{(y_t - k_{t+1})^{1-\gamma} - 1}{1-\gamma} + \tilde{V}(k_{t+1}, e_t; K_t, z_t) \right\}, \quad (2.7)$$

Third, to solve this problem computationally, one needs to discretized the state space because computers cannot handle the continuous state. As briefly mentioned in Section 2.4, the forcing processes of technology shock  $z_t$  and the idiosyncratic efficient labor shock  $e_t$  can be discretized by the first order Markov chains using the method suggested by Tauchen (1986). Therefore, I denote  $\mathcal{S}_z = \{z^1, z^2, \dots, z^{n_z}\}$  and  $\mathcal{S}_e = \{e^1, e^2, \dots, e^{n_e}\}$  as the discrete state space of  $z_t$  and  $e_t$ , respectively. Moreover, these two discrete Markov chains have transition matrices  $\Pi^z = \{\pi_{i,j}^z\}$  and  $\Pi^e = \{\pi_{i,j}^e\}$ , where  $\pi_{i,j}^z = \text{Prob}(z_{t+1} = z^j | z_t = z^i)$ , and is similar to

$\pi_{i,j}^e$ . By employing these two transition matrices, one can obtain the ergodic distributions of  $z_t$  and  $e_t$ .<sup>7</sup> Next, denote vectors  $\tilde{\pi}^z = [\pi_1^z, \pi_2^z, \dots, \pi_{n_e}^z]'$  and  $\tilde{\pi}^e = [\pi_1^e, \pi_2^e, \dots, \pi_{n_e}^e]'$  as the ergodic distributions of  $z_t$  and  $e_t$ , respectively. Naturally, one also needs to discretize the state spaces of individual capital and aggregate capital. The discrete state space of  $K_t$  is denoted as  $\mathcal{S}_K = \{K^1, K^2, \dots, K^{n_K}\}$ . In order to implement the endogenous grid points method, construct grid points for  $k_{t+1}$  as  $\mathcal{S}_k = \{k^1, k^2, \dots, k^{n_k}\}$ .

## 2.4.2 Implementation of the Algorithm

To solve the above problem, define an operator as follows:

$$V^{(i+1)}(k_t, e_t; K_t, z_t) = \max_{k_{t+1} \geq 0} \left\{ \frac{(y_t - k_{t+1})^{1-\gamma} - 1}{1-\gamma} + \tilde{V}^{(i)}(k_{t+1}, e_t; K_{t+1}, z_t) \right\},$$

subject to Eq. (2.2) to Eq. (2.6). Below is the algorithm procedure.

**Step 0: (Initialization Step)** Construct grid for the state space, i.e., construct  $\mathcal{S}_k, \mathcal{S}_e, \mathcal{S}_z$ , and  $\mathcal{S}_K$ . Denote  $\mathcal{S} = \mathcal{S}_k \times \mathcal{S}_e \times \mathcal{S}_z \times \mathcal{S}_K$ . Set  $i = 0$ . Initialize  $V^{(0)}(k_{t+1}, e_t; z_t, K_t)$  for all points of grid  $\mathcal{S}$ .

**Step 1: (EGM Step)** For each grid point  $k_{t+1} \in \mathcal{S}_k, e_t \in \mathcal{S}_e, z_t \in \mathcal{S}_z$  and  $K_t \in \mathcal{S}_K$ , utilize the first order conditions of the problem

$$(c_t^*)^{-\gamma} = \tilde{V}_{k_{t+1}}^{(i)}(k_{t+1}^*, e_t; K_t, z_t).$$

To compute the derivative, one can use the method proposed by Schumaker (1983).<sup>8</sup> Hence, one can obtain the optimal consumption instantly by

$$c_t^* = \left( \tilde{V}_{k_{t+1}}^{(i)}(k_{t+1}^*, e_t; K_t, z_t) \right)^{-1/\gamma}. \quad (2.8)$$

<sup>7</sup>Use the eigenvector vector to find the ergodic distributions.

<sup>8</sup>Judd's (1988) chapter 6 also explain how to calculate the derivative.

With the optimal consumption,  $c_t^*$ , and the optimal individual capital holding of next period,  $k_{t+1}^*$ , at hand, one can obtain the endogenous resource in period  $t$  from the individual budget constraint:  $y_t^* = c_t^* + k_{t+1}^*$ . Consequently, one can back out the individual capital  $k_t^{endo}$  as follows:

$$k_t^{endo} = \frac{y_t^* - w_t e_t}{(1 + r_t - \delta)}.$$

Note that  $r_t$  and  $w_t$  are functions of  $z_t$  and  $K_t$ , and thus, are known.<sup>9</sup>

Note that the strength of EGM should be obvious here. In this step, for each grid point  $k_{t+1} \in \mathcal{S}_k$ ,  $e_t \in \mathcal{S}_e$ ,  $z_t \in \mathcal{S}_z$  and  $K_t \in \mathcal{S}_K$  one *analytically* solves for  $k_t^{endo}$  through the first order condition (Eq. 2.8) and the budget constraint. Here,  $k_{t+1}$  is already the optimal choice of  $k_t^{endo}$  along with  $e_t$ ,  $z_t$ ,  $K_t$ . In contrast, the traditional methods, which loops over capital of the current period,  $k_t$ , along with the other state variables, require one to utilize certain numerical root-finding procedures to find the optimal capital holding of next period. The EGM simplifies the problem a lot.

Up to now, we have the pair  $\{k_t^{endo}, k_{t+1}^*\}$  for all of the points in  $\mathcal{S}$ . That is, we already obtain the policy function  $k_{t+1}^* = g^k(k_t^{endo}, e_t; K_t, z_t)$ . Note that the first argument of this policy function is not defined in  $\mathcal{S}_k$ . Thus, one needs to do interpolation to adjust the domain of the function. Moreover, for points  $k_t \in \mathcal{S}_k$ , such that  $k_t \leq k_t^{endo}$ , set  $k_{t+1} = k^1$ . That is, the borrowing constraint is binding at those particular points. For simplicity, we call the domain adjusted policy function as  $k_{t+1} = \hat{g}^k(k_t, e_t; K_t, z_t)$ .

---

<sup>9</sup>Note, since the individuals do not value leisure, the aggregate efficient labor supply is always constant and is determined by the ergodic distribution of idiosyncratic labor shocks.

**Step 2: (Xpa Step)** Use the domain adjusted policy function  $\hat{g}^k$  to parametrize the primary auxiliary policy function as follows:

$$k_{t+1} = \begin{cases} \Psi_{0,e^1}(s) + \Psi_{0,e^1}(s)k_t, & \text{if } e = e^1 \\ \Psi_{0,e^2}(s) + \Psi_{0,e^2}(s)k_t, & \text{if } e = e^2 \\ \vdots & \\ \Psi_{0,e^{n_e}}(s) + \Psi_{0,e^{n_e}}(s)k_t, & \text{if } e = e^{n_e} \end{cases},$$

where  $s$  is a vector containing the aggregate state variables, i.e.,  $z_t$  and  $K_t$ . Furthermore, since the proportion of individuals at each idiosyncratic state is given by ergodic distribution  $\tilde{\pi}^e$ , the average policy rule is the weighted average of the primary auxiliary policy function:

$$\begin{aligned} K_{t+1} &= G^K(K_t, z_t) \\ &= \sum_{j=1}^{n_e} \pi_j^e \Psi_{0,e^j}(s) + \sum_{j=1}^{n_e} \pi_j^e \Psi_{1,e^j}(s) K_t. \end{aligned}$$

The logic behind this step is as follows. First, for each idiosyncratic state, we approximate the adjusted policy function  $\hat{g}^k$  by a first order polynomial, which is called the primary auxiliary policy function. Second, we utilize the fact that the proportion of individuals in different idiosyncratic states is exogenously given. Therefore, we can obtain the average policy function through a simple weighted average with respect to those auxiliary policy functions. This average policy function is exactly the aggregate law of motion, since the population size in the model, by definition, is one. Hence, we obtain a aggregate law of motion without relying on the simulation procedure like that adopted in Krusell and Smith (1998). For details of Xpa, refer to den Haan and Rendahl (2010).

**Step 3: (Update Value Function)** For each  $k_t \in \mathcal{S}_k$ ,  $e_t \in \mathcal{S}_e$ ,  $z_t \in \mathcal{S}_z$ , and  $K_t \in \mathcal{S}_K$ , the following equation is utilized to update the value function

$$V^{(i+1)}(k_t, e_t; K_t, z_t) = \left\{ \frac{(y_t - \hat{g}^k(k_t, e_t; z_t, K_t)) - 1}{1 - \gamma} + \tilde{V}^{(i)}(k_{t+1}, e_t; K_t, z_t) \right\}.$$

Here, the maximization operator is discarded since  $\hat{g}^k$  has already been the optimal solution.

**Step 4: (Update Expected Value function)** By using the value function  $V^{(i+1)}(k_t, e_t; z_t, K_t)$  obtained from the last step, one can calculate  $\tilde{V}^{(i+1)}(k_{t+1}, e_t; z_t, K_t)$  by

$$\tilde{V}^{(i+1)}(k_{t+1}, e_t; z_t, K_t) = \beta \sum_{j_e=1}^{n_e} \sum_{j_z=1}^{n_z} \pi_{e_t e^{j_e}} \pi_{z_t z^{j_z}} V^{(i+1)}(\hat{g}^k(k_t), e^{j_e}; z^{j_z}, G^K(K_t, z_t)).$$

Note that a two-dimensional interpolation on  $k_t$  and  $K_t$  directions is needed, since the values of  $V^{(i+1)}(., .; ., .)$  are merely defined on the grid  $\mathcal{S} = \mathcal{S}_k \times \mathcal{S}_e \times \mathcal{S}_z \times \mathcal{S}_K$ , and  $\hat{g}^k(k_t)$  and  $H^K(K_t, z_t)$  are generally not on the grid points.

**Step 5:** If  $\sup_{\mathcal{S}} \left| \tilde{V}^{(i+1)}(k_t, e_t; K_t, z_t) - \tilde{V}^{(i)}(k_t, e_t; K_t, z_t) \right| < \text{predetermined tolerance}$ , the problem is considered to have converged; otherwise,  $i \rightsquigarrow i + 1$  and goes back to step 2.

Note that in the above steps the aggregate law of motion is obtained in each iteration when one updates the value function. Therefore, one just needs to implement the whole value function iteration process once. However, if one adopts the algorithm of Krusell and Smith (1998), one has to conduct the VFI as many times as the numbers of updating the aggregate law of motion.

## 2.5 Numerical Results

### 2.5.1 Calibration

The calibration parameters I choose are as follows. The discount factor  $\beta = 0.9896$  reflects preference in a period of one quarter. The depreciation rate is set to  $\delta = 0.025$ , which is widely



adopted in quarterly data for the U.S. economy.<sup>10</sup> The relative risk aversion coefficient  $\gamma = 1.5$  is a common choice in the literature. The risk aversion coefficient  $\gamma$  signifies people’s reaction towards risk. Similar to most studies in the business cycle literature, capital share  $\alpha$  is set to 0.36. As for the two driving processes, I pick  $\rho_z = 0.5$ ,  $\sigma_z = 0.02$ ,  $\rho_e = 0.5$ ,  $\rho_e = 0.02$ . Moreover, the number of grids for the state variables is set as follows. The number of individual capital is set to  $N_k = 50$ , while the number of aggregate capital is assumed to be  $N_K = 5$ . Both of the driving processes are allowed to take 5 values, i.e.,  $N_z = 5$ ,  $N_e = 5$ . When the difference of successive value functions is less than  $1.0e^{-10}$ , the algorithm stops.

In order to reveal the power of the new algorithm, I also conduct some sensitivity analyses. The sensitivity analyses focus on two parameters: the subjective discount factor  $\beta$  and relative risk aversion coefficient  $\gamma$ . Model 1 changes  $\beta$  to 0.95 from the benchmark value; Model 2 changes  $\gamma$  to 10 from the benchmark value; Model 3 changes both.

Table 2.1: Parameters Calibration

Parameters	$\beta$	$\gamma$	$\delta$	$\alpha$	$\rho_z$	$\sigma_z$	$\rho_e$	$\sigma_e$
Benchmark	0.9896	1.5	0.025	0.36	0.5	0.02	0.5	0.02
Model 1	0.9500	1.5	0.025	0.36	0.5	0.02	0.5	0.02
Model 2	0.9896	10	0.025	0.36	0.5	0.02	0.5	0.02
Model 3	0.9500	10	0.025	0.36	0.5	0.02	0.5	0.02

## 2.5.2 Some Preliminary Results

First of all, I plot some selected policy functions in Figure 2.1. To make the explanation simple, I plot only five policy functions where the aggregate capital is set to  $K = 40.70$ , the aggregate shock is set to  $z = 1$ , and the idiosyncratic shock takes five different values. The result is similar to Krusell and Smith (1998) in that the policy functions are almost linear across the whole domain. Actually, only when the individual capital level is extremely low and the idiosyncratic

<sup>10</sup>See, for example, King et al. (1988) and Plosser (1989).

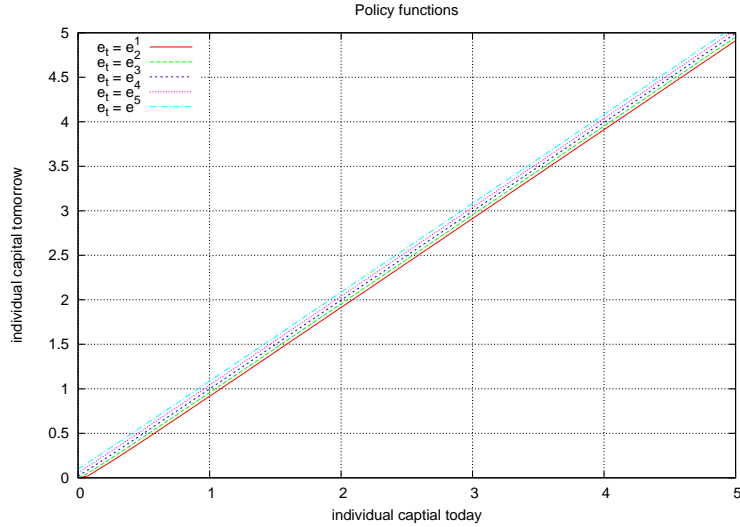


Figure 2.1: Individual Policy Functions of the Benchmark Calibration

efficient labor shock is enormously bad ( $e = e_1$ ), will the policy function look a little nonlinear. Since the policy functions are almost linear, and very few agents are binding, the property of approximated aggregation holds. As explained by Krusell and Smith (2006), since the policy functions are (almost) linear and very few agents are borrowing constrained, the marginal propensity of saving is almost the same across agents. Therefore, the distribution of capital does not matter, and one can simply use the first moment of the capital distribution to predict the aggregate capital of the next period.

As mentioned earlier, the equilibrium of this model has to satisfy the consistency condition. The consistency condition means that the law of motion of aggregate capital implied by the model must mimic the overall outcome of heterogeneous agents. To check this, the procedure is as follows. First, one needs to simulate a huge number of heterogeneous agents who face their own idiosyncratic efficient labor shocks and the same aggregate shocks. Next, calculate the aggregate capital through simulated data. Finally, one needs to compare the simulated capital movement with the implied law of motion from the model.

To do this simulation experiment, the number of heterogeneous agents is set to 10,000. The model is simulated for 2,500 periods, and the first 2,000 periods are discarded to make sure

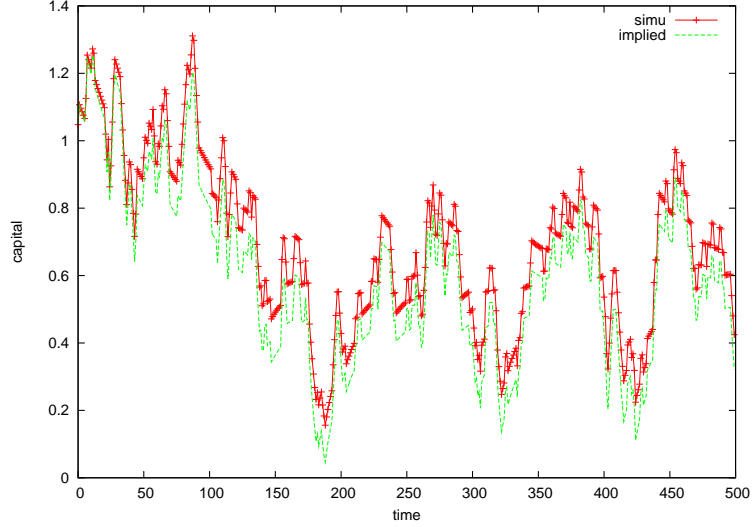


Figure 2.2: Simulated Value of Aggregate Capital of Benchmark Calibration

that the initial state of the distribution does not determine or bias the implied aggregate law of motion. The idiosyncratic shocks and aggregate shocks are simulated from the first order Markov Chains, which reflects the setting of AR(1) processes. The simulation results are shown in Figure 2.2. As can be seen, the implied aggregate law of motion nicely traces the movement of the simulated aggregate law of motion that is solved from the model.

Other than looking at the evolution of aggregate capital, one can also check the so called Euler equation residual to examine whether or not the solution is correct. The Euler equation residual of this model is

$$residual = \frac{c_t - \left\{ \min \left\{ \tilde{V}_k(k_{t+1}, e_t, K_t, z_t), (1 + r_t - \delta)k_t + w_t e_t \right\} \right\}^{-\frac{1}{\gamma}}}{c_t}.$$

One can calculate the residuals on the domain of the state space,  $\mathcal{S}$ . To calculate these residuals, one needs to discretized the state space. For illustration purposes, I do not change the grid numbers of aggregate capital or aggregate shocks; they are still set to 5. However, I increase the number of grids for the individual capital to 1000, which are assumed to be equally-spaced. The results are shown on Table 2.2.

Table 2.2: The Euler Equation Residuals

Norm	Benchmark	Model 1	Model 2	Model 3
Avg $\log_{10}(residual)$	-3.5223	-3.5210	-3.2368	-3.1834
Sup $\log_{10}(residual)$	-0.8299	-0.9538	-0.7781	-0.9098

The interpretation of the Euler equation residuals is as follows. If  $\log_{10}(residual) = -4$ , it simply suggests that the agent will make a mistake of \$1 when he consumes \$10,000. As you can see from Table 2, the Euler equation residuals are pretty small if we examine the residuals from the average norm. However, if we look at the residual from the sup norm, the performance is somewhat unsatisfactory. This is merely a consequence from interpolation. When I calculate the Euler equation residuals, I adopt the linear interpolation method. This interpolation approach is, to some extent, flawed because the policy functions are somewhat nonlinear when individual capital holding is close to zero and idiosyncratic efficient labor shocks are low. This problem can be remedied by using nonlinear interpolation methods such as cubic splines.

### 2.5.3 Comparison with Existing Algorithms

The new algorithm I propose here is in many aspects better than some of the existing algorithms. I briefly discuss the difference between the new algorithm, Krusell-Smith algorithm, and Xpa in this subsection.

**Comparing with Krusell-Smith (1998)** The new algorithm is in essence utilizing the value function iteration, which is identical to Krusell and Smith (1998). However, in their algorithm, Krusell-Smith employs a combination of the Newton method and the bisection method to find the optimum of the right hand side of the Bellman equation. To implement the Newton method, they resort to cubic splines for the dimension of individual capital. The bisection method is used when the individual capital is lower than some level such that they can handle the problem of an occasionally binding borrowing constraint,  $k_t \geq 0$ . The drawback of the cubic spline is that it is a computational intensive method. This becomes a serious problem since for each combination

of  $z_t$ ,  $e_t$ ,  $K_t$  in each value function iteration, one needs to do cubic spline interpolation once, making the computation very sluggish. Moreover, the adoption of bisection to handle a binding constraint also creates problems. The reason is that one usually has no prior knowledge regarding the bracket for the bisection method. This might not be a problem since a proper bracket can be sometimes found by trial-and-error. However, it really becomes a problem if one would like to conduct estimation, since the proper bracket varies with the combination of parameters.

The new algorithm does not need to resort to the Newton method and bisection. This algorithm searches for the optimum of the right hand side of the Bellman equation through the endogenous grid-point method of Carroll (2006). The backward solving structure of EGM makes the search of the optimum of the Bellman equation trivial. Moreover, as explained in step 2 in the last section, the EGM can naturally handle the binding constraint. Therefore, one does not need to resort to the bisection method, and be troubled with the problem of finding the proper bracket.

More importantly, the Krusell-Smith algorithm relies on simulation to update the aggregate law of motion. To come up with the accurate aggregate capital, the number of individuals has to be large; therefore, it usually takes a long time to update the aggregate law of motion. Moreover, to ensure that the simulation is reliable, one needs to provide a good initial distribution for the individual capital holding. If the initial distribution is not accurate enough, one needs to discard a long series of simulation outcomes. This restriction might not be a serious problem if the researchers only need to solve the model once. However, it becomes a serious problem when implementing estimation tasks. In contrast, I follow the spirit of the explicit aggregation (Xpa) method in the new algorithm. Therefore, the aggregate law of motion is obtained instantly through the weighted average of individual policy functions, where the weights comes from the exogenous ergodic distribution of idiosyncratic shocks. Since one can obtain the aggregate law of motion from the policy functions during each iteration of the value function, one actually needs to implement the value function iteration only once. It is a significant improvement compared to the Krusell-Smith algorithm, which needs to undergo value function iteration many times

until the aggregate law of motion converges.

**Comparing with den Haan and Rendahl (2010)** The new algorithm is actually a modification of the explicit aggregation from den Haan and Rendahl (2010). As I explained above, this algorithm adopts explicit aggregation to update the aggregate law of motion. However, the core of solving the models are different. In den Haan and Rendahl (2010), they use the Euler equation based methods and, therefore, need to make assumptions on the shapes of the policy functions in advance. If the original guess for the policy function is not quite good, the algorithm can easily break down. Furthermore, it will become a problem when one needs to implement an estimation task, since one usually does not have any prior knowledge about the policy functions when the underlining parameters change. In contrast, the core of the new algorithm is the value function iteration. The value function iteration method has nice convergence properties, which have been proven mathematically. Theoretically, the value function will always converge since the Bellman equation is a contraction mapping. In practice, all we need to do to implement the new algorithm is to provide a monotonically increasing and concaved value function, which is fairly straightforward. When one does so, the contraction mapping will automatically take care of all other things. This is a significant advantage especially for the estimation task.

As mentioned earlier, having a fast and reliable algorithm is extremely important when one estimates a heterogeneous agent model. I provide a comparison on the speed of the existing algorithms in Table 2.3. There are some points that need to be mentioned regarding this table. First, KS-num is the algorithm of Young (2010), and KS-sim is that of Maliar et al. (2010); they are two different versions of the original Krusell-Smith (1998) algorithm. Param is the algorithm proposed by Algan et al. (2008), and Penal is that of Kim et al. (2010). As you can see, the Krusell-Smith type algorithms are very slow. The Param algorithm takes close to two days to solve the model once. The Penal algorithm is extremely fast, but, as explained above, this method does not solve the traditional heterogeneous agent model with borrowing constraints. Xpa is the explicit aggregation method of den Haan and Rendahl (2010) and its

Table 2.3: Computation Time

Algorithm	Programming Language	Time
KS-num	Fortran	324 minutes
KS-sim	Matlab	310 minutes
Param	Fortran	2739 minutes
Xpa	Matlab	7 minutes
Penal	Matlab	< 1 second!
New Algorithm	Fortran	< 10 second!

\* The model that implements the Penal algorithm is not exactly the same as the others. Moreover, the model used here for comparison is that of den Haan et al. (2010), which is a slightly modification of Krusell and Smith (1998).

speed is acceptable for estimation task. Finally, the algorithm I provide in this chapter is noticeably faster compared to all the existing algorithms.

## 2.6 Conclusion

This chapter proposes a new algorithm for solving computationally challenging heterogeneous agent (HA) models. The algorithm combines the value function iteration (VFI) method, the endogenous grid points method (EGM), and the explicit aggregation (XPA) method. This new algorithm is robust in the sense that the initial values are irrelevant for finding the optimal policy functions. Moreover, it is also fast because both the root-finding procedure for the household problem and the repeated simulation for the aggregate law of motion are not necessary. The algorithm solves a typical HA model within just a few seconds in a ordinary desktop computer. As a consequence, it could be a proper tool for structural estimation of HA models, which needs to solve the model several thousands times.

## Chapter 3

# Estimation of Heterogeneous Agent Models

### 3.1 Introduction

It is well known that macroeconomic models have improved significantly in the modeling strategy and explaining some stylized facts in the past two decades. The development of New Keynesian (NK) models plays an indispensable role. In addition to having a better micro-foundation, NK models incorporate several frictions and thus provide better descriptions of reality than the traditional macroeconomic models. Besides, because of the advance in econometric methods, the implications of NK models can be explored by estimation rather than calibration. In short, estimable NK models make them suitable for answering quantitative questions such as the magnitudes of real effects due to a particular government policy.

The development of macroeconomics conveys two important messages. The first message is about model setting. As explained by Sims (1996), models should be constructed to explain observable real data. NK models incorporate the potential frictions in the real data. Therefore, these models provides more explanatory power for puzzling macro phenomena. For example, Christiano, Eichenbaum, and Evans (2005) provide explanation for the sluggish inflation re-



sponse with respect to monetary shocks. The second message is about the importance of model estimation. Current macroeconomic models are complicated, because they often involve dynamics of many endogenous variables and rational expectations of agents. Thus, the predictions of models are sensitive to the change of parameter values. For example, the duration of real output increases responding to a unexpected technology shock crucially depends on the persistence coefficient of the technology shock. Thus, to obtain a better understanding of the underlying structure of the economy, precise estimation of structural parameters is crucial.

While we have seen huge progress in the field of macroeconomics, up to now, all of the estimated DSGE models resort to the representative agent paradigm. See An and Schorfheide (2007), Smets and Wouters (2007), Christiano et al. (2005) for examples. However, the main justification of a representative agent is that individuals are able to trade their risks in complete markets. Unfortunately, complete markets are usually rejected in empirical studies. In other words, the quantitative implications of these models might be imprecise or biased if the complete markets assumption is far from reality.

To address this shortcoming, economists have already developed several heterogeneous agent (HA) models. See for example, Aiyagari (1994), Huggett (1993), Krusell and Smith (1998), and Quadrini (2000). The existing HA models have incorporated various sources of heterogeneity such as health, age, marriage, unemployment, and preference. These improvements make HA models a strong alternative to the traditional representative agent models. Unfortunately, an important drawback still exists: None of these models is estimated by rigorous econometric methods. In essence, the analysis of HA models still relies on calibration. As noted, calibrated models are not suitable to answer quantitative questions unless accurate parameters are used.

According to An, Chang, and Kim (2009), the representative agent model often fails to represent an equilibrium outcome of a heterogeneous agent economy. The fundamental reason for this is that models are misspecified when one fails to take into account the presence of heterogeneity in the real economy. A direct result of this is that parameter estimations are likely to be misleading. A potential remedy for this particular type of misspecification is to

estimate a model with heterogeneous agents.

Moreover, because of the adoption of the representative agent framework, economists usually use only aggregate data for estimation, such as aggregate output and consumption. This implies that some other useful information is actually left unemployed. For example, the U.S. Bureau of Economic Analysis (BEA) provides a comprehensive household consumption data, known as CEX. This dataset is widely used in analyzing household consumption related issues under partial equilibrium, but rarely used for macroeconomic models under general equilibrium. With HA models, micro-level consumption data can be conveniently utilized in estimation to help us examine and better understand more aspects of the economy.

Thus, the motivation of the chapter is to address the mentioned shortcomings in modern macroeconomic studies. Specifically, we estimate a heterogeneous agent model under general equilibrium by incorporating both aggregate and distributional data. This task is rather ambitious. First of all, HA models are notoriously difficult to solve computationally. Second, unlike the representative agent models, in which the standard procedures for estimating them are fairly established (see Canova (2007) and DeJong and Dave (2007)), the applicable literature does not exist for the estimation of HA models under general equilibrium. Thus, there is a need for us to provide our own procedure for it.

In regards to the first issue, we adopt the algorithm presented in the first chapter, which is not only computational efficient but also robust in the sense that educated guesses for initial values are not necessary. As for the second issue, we combine two simulation-based estimation methods, i.e., Indirect Inference and Laplace-Type estimator, to formulate a hybrid estimation method.

The structure of this chapter are as follows. We review the relevant literature in Section 3.2 and lay out an HA model in Section 3.3. The empirical methodology is discussed in Section 3.4. Section 3.5 contains a short description of how data are being used. Lastly, Section 3.6 covers the empirical results and Section 3.7 concludes the chapter.

## 3.2 Literature Review

### 3.2.1 Review of Heterogeneous Agent Models

Huggett (1993) constructs a heterogeneous agent model in which individual agents face idiosyncratic income shocks. The only asset that individuals use to insure idiosyncratic shocks is the riskless bond.<sup>1</sup> Moreover, he assumes that individuals cannot borrow without limit. That is, every individual faces a potentially binding borrowing constraint. Because of the possibility of low income and the borrowing constraint, individuals exhibit precautionary saving behavior. Under this formulation, Huggett (1993) provides an explanation for the cause of low real interest rates, which has been viewed as a puzzle under the representative agent model.

Aiyagari (1994) considers a production economy with heterogeneous agents, and capital, rather than a riskless bond, is used for insuring idiosyncratic shock. Similar to Huggett (1993), he shows that the interest rate, aggregate capital, and wealth distribution can be coherently determined through the precautionary saving motive, which stems from the borrowing constraint under the heterogeneous agent setting. His paper has now become an indispensable part of modern heterogeneous agent models. However, the prediction about the wealth distribution in the model of Aiyagari (1994) is not satisfactory. The Gini coefficient is too low and the rich accumulate too little wealth, which conflict with what we observe in real data.

Quadri (2000) extends the Aiyagari model to accommodate occupational choice. By including endogenous occupational choice, his model fits surprisingly well with the wealth distribution and socioeconomic mobility of the United States. In his calibrated model, the top 1 percent of rich households holds 24.9 percent of total wealth and the empirical counterpart is 26 percent in the Panel Study of Income Dynamics (PSID) data.

The model of Krusell and Smith (1998) is the first heterogeneous agents model that incorporates aggregate uncertainty. They provide an algorithm based on a property called *approximate aggregation*. Krusell et al. (2009) use a calibrated model to investigate the welfare effect of

---

<sup>1</sup>Hence, it is an incomplete markets model.

removing business cycles, which is a famous issue in economics introduced by Lucas (1987).<sup>2</sup> However, Krusell et al. (2009) find that the welfare gain under their heterogeneous agents model is an order of magnitude larger than that of Lucas. Moreover, their calibrated model shows that there are large differences across groups: Very poor consumers gain a lot, and so do very rich consumers. However, the majority of consumers (middle class) gain little. That is, the improvement of welfare is U-shaped across wealth levels.

The heterogeneous agent models are growing fast now and most of the early literature focused on the heterogeneity of income or wealth. However, heterogeneity of the economy is not limited in this sole dimension. For instance, Palumbo (1999) shows that health status is a key driving process for the saving decision of the elderly, and Cubeddu and Rios-Rull (2003) finds that marital status risk is a larger source of precautionary saving than income risk.

### 3.2.2 Review of Empirical Studies

Gourinchas and Parker (2002) construct a dynamic stochastic model to analyze the life cycle saving behavior of heterogeneous agents. Each agent faces an exogenous, stochastic labor income process. The income process of each agent is a function of agent's characteristics, including education level, occupational type, and age. Moreover, to make the model more flexible, they allow the utility function to be a function of agents' characteristics. That is, the heterogeneity of the economy stems not only from the difference in income level or wealth but also from the diversity in agents' preferences.

They construct average consumption and income profiles across the households of five different educational levels and four occupational groups. The data they use to construct consumption and income profiles are the Consumer Expenditure Survey (CEX), which contains roughly 40,000 households from 1980 to 1993. After constructing the consumption and income profiles, they numerically solve the life-cycle model and generate a theoretical consumption profile. By comparing the theoretical consumption profile to that of the empirical one, they estimate the

---

<sup>2</sup>Lucas uses a representative agent model to quantify the welfare effect of removing the business cycle, and shows that the welfare gain is remarkably small. It only takes one-tenth of a percent of lifetime consumption.

deep parameters of the model through the simulated method of moments (SMM). Besides, Gourinchas and Parker (2002) is methodologically important since it is the first study on structural estimation for models containing a borrowing constraint, which is an indispensable feature of modern incomplete markets, macroeconomic models. Before their paper, economists had no way to explicitly take the precautionary saving motive into account, even when its importance is widely acknowledged. With the improvement in methodology, they provide an explanation for the origin of the hump-shaped consumption pattern over life cycle. Lastly, as explained earlier, the model of Gourinchas and Parker (2002) is a partial equilibrium one. Thus, their framework is only suitable for assessing the consumption functions, but not for macroeconomic issues, such as the welfare consequence of government policies to different groups of people.

Following the work of Gourinchas and Parker (2002), Cagetti (2003) also estimates a life cycle model with heterogeneous agents. The heterogeneity of Cagetti's model comes from two sources: age and education level. The estimation strategy of this paper is similar to that of Gourinchas and Parker (2002); both use the simulated method of moments to estimate the models. However, while the matching targets of Gourinchas and Parker (2002) is the consumption pattern over the life cycle, Cagetti (2003) uses the individual wealth profile as an equivalent counterpart.

The main concern of Cagetti (2003) is to back out the subjective discount factor and relative risk aversion coefficient. His estimation results reveal that the time preference rate decreases with education level. People with higher education are usually more patient than those with lower education. Besides, the estimated risk aversion coefficient is higher than 3 (often higher than 4), which is much higher than the widely accepted level in macroeconomics. In most macroeconomic studies, many researchers adopt log utility, suggesting that the risk aversion coefficient is equal to one.

Unlike the above two papers that adopt the life cycle model setting, Hintermaier and Koeniger (2008) estimate a heterogeneous agent model in which agents live forever. This model is very similar to the that of Aiyagari (1994), but the prices (i.e., interest rate) are exogenously

given and not endogenously determined through the model. Similar to Cagetti (2003), their main goal is to estimate out the subjective discount factor and relative risk aversion coefficient. By taking into account the heterogeneity of race, age, and education level, the model provides a very concise explanation of the US wealth distribution up to the 90th percentile. Their estimation is based on the Survey of Consumer Finances in 1983. Interestingly, even with a different model setting, their estimates for the subjective discount factor and risk aversion coefficients are similar to those of Cagetti (2003);<sup>3</sup> both papers find that people are highly risk averse and impatient. Besides, Hintermaier and Koeniger (2008) also use the simulated method of moments (SMM) to estimate the model. The main difference is that the matching targets are the percentile of wealth distribution rather than the consumption or wealth profile over the life cycle.

### 3.2.3 Review of Estimation Methods

Acknowledging the critical role of precise estimation of structural parameters, economists have responded with different econometric methodologies suitable for backing out these parameters. In general, the methods that are suitable for structural estimation for economic models consist of Generalized Method of Moments (GMM) estimation, Bayesian estimation methods, classical Maximum Likelihood Estimation (MLE), and simulation based estimation methods, such as Simulated Method of Moment (SMM), Indirect Inference, and Simulated Method of Likelihood.

The GMM estimation method was first introduced by Hansen (1982), who generalized the traditional method of moments (MOM) to the over-identifying cases involving serially correlated stochastic processes and established asymptotic properties.<sup>4</sup> As one of the most widely used estimation methods for structural models, GMM has some merits. First, in order to implement GMM estimation, it is necessary to construct moment conditions, which usually can be obtained directly thorough the first order conditions of economic models. Second, GMM estima-

---

<sup>3</sup>Recall, Cagetti (2003) uses the life cycle model but Hintermaier and Koeniger (2008) adopt a dynastic setting.

<sup>4</sup>Here, over-identification means that the number of moment conditions exceeds the number of parameters to be estimated.

tion does not need to specify the distribution of errors behind the models, and hence avoids the model misspecification problem with respect to the error distribution of the model. Two famous examples are Hansen and Singleton (1982) and Christiano and Eichenbaum (1992). While the implementation of GMM is straightforward, it has some limitations. Macroeconomic models usually imply many moment conditions. The determination of which moment conditions are better for estimation is usually a key issue. If one does not choose proper moment conditions, the so called weak identification problem will usually pop up and make the estimation result unreliable or infeasible.<sup>5</sup> Moreover, GMM estimation could have poor small sample properties. That is, if the sample size is not large enough, the estimation results are often biased, see Hansen et al. (1996). Besides, to implement GMM, all the variables in the moment conditions must be observable. If the moment conditions contain some latent variables, GMM estimation will be infeasible. The infeasibility of GMM to handle latent variables limits its application to some macroeconomic models which usually contain some variables that are unobservable in the data but vital for the individual decisions.

Considering the above limitations, the mainstream of structural estimation is now likelihood based methods, such as the maximum likelihood estimation method and Bayesian estimation method. Broadly speaking, one needs to specify the whole data generating process (DGP) along with the distributions of errors under the models to implement likelihood based methods. There are usually two steps to obtain the DGP. Researchers first need to solve the model explicitly in terms of the policy functions. Then, express the model in a state space form through the policy functions. With the state space representation, researchers can write down the likelihood function<sup>6</sup> and, hence, use either maximum likelihood estimation or Bayesian estimation. The main advantage of using the Bayesian estimation method rather than MLE is as follows. By imposing prior distributions of the parameters, the potential weak identification problem of maximum likelihood estimation can be somewhat remedied.<sup>7</sup>

---

<sup>5</sup>Weak identification means that the restrictions imposed to estimate the parameters are not very informative.

<sup>6</sup>DeJong and Dave (2007) give a detail discussion for the procedure.

<sup>7</sup>Of course, the improvement in identification is not costless. The usual responses toward the roles of prior distributions apply here.

These two likelihood based methods are more efficient than GMM, since they use full information of the model. However, they also have some drawbacks. First, the distribution functions of errors in the model have to be fully specified. Moreover, to make the derivation of the likelihood function simple, error terms are usually assumed to be normal; to prevent the so-called stochastic singularity, researchers have to impose some *ad hoc* measurement errors. The introduction of measurement errors might worsen the model mis-specification problem. More importantly, as explained above, the state space representation is needed for writing down the likelihood function. However, sometimes the models are too complicated to easily express the model in the state space form, at least linearly. Heterogeneous agent models are noticeable examples, for example see Gourinchas and Parker (2002).

Likelihood function based methods now are widely adopted in representative agent Dynamic Stochastic General Equilibrium (DSGE) models. For example, McGrattan et al. (1997) and Ireland (2001), Altug (1989) and Ingram et al. (1994) use MLE for many macroeconomic issues. Further, the Bayesian estimation method now has become the standard tool to analyze New Keynesian Models empirically, for instance see Smets and Wouters (2007) and Fernandez-Villaverde and Rubio-Ramirez (2007). Unfortunately, the essence of the heterogeneous agent models makes both GMM estimation and likelihood function based methods infeasible. The first order conditions of heterogeneous agent models contains both the aggregate and individual variables and hence it is not straightforward to use them for estimation. Further, in the heterogeneous agent framework, economists still do not know how to express the model in the state space form. Therefore, researchers usually resort to simulation based estimation methods. For example, Gourinchas and Parker (2002) estimate the consumption function with heterogeneous agents by simulated method of moments.

### 3.3 The Model

As mentioned, macroeconomists have already invented many heterogeneous agent models for different aspects of heterogeneity. Broadly speaking, these models can be classified by two



criteria – the life horizon of agents and whether or not aggregate shocks exist. Based on the first criterion, the life horizon of agents can be either classified as life-cycle based or dynastic based. The life-cycle type models are usually used to deal with issues related to age. For example, Auerbach and Kotlikoff (1987) analyze how a particular policy affects the welfare of different generation, where as Storesletten et al. (2007) study the effects of idiosyncratic shocks and life-cycle aspects on asset prices. On the other hand, some studies use the dynastic setup because researchers do not concern themselves about the issues stemming from age differences.

As for the second criterion - whether or not aggregate shocks exist - the decisive element comes down to whether economists care more about the stationary property or the dynamic aspect of the model. For example, Aiyagari (1994) and Huggett (1993) use heterogeneous agent models to examine the effect of incomplete markets on the saving decisions of individuals within the production and exchange economy. This type of model focuses only on the stationary wealth distribution of the economy, but refrains from the dynamic aspects of the economy. On the other hand, models containing aggregate shocks usually focus on business cycle properties. For example, Chang and Kim (2007) use a heterogeneous agents model with indivisible labor to examine the implication on the labor market due to aggregate shocks, which usually performs poorly in the representative agent model (see Shimer (2009)).

Since the main goal of this chapter is to study the estimation of heterogeneous agent models in a business cycle context without focusing on any age-related issues, we adopt a heterogeneous agent model with dynastic setup and aggregate shocks. For simplicity, we assume that there is no role for government, and hence there exists neither monetary nor fiscal policies. Moreover, we adopt a production economy setting, which is in line with Krusell and Smith (1998).

**Household Problem** There is a continuum of *ex ante* identical forever-living households. Even though being *ex ante* identical, households are actually *ex post* heterogeneous. The reason is that households face idiosyncratic labor efficiency shocks, but there are not complete markets to fully insure these risks. While households cannot fully insure the idiosyncratic risks, they are

able to partially insure them. Specifically, they use the homogeneous physical capital to build a buffer stock of wealth. Therefore, when they encounter a particular bad realization of the efficient labor shock, their consumption level will not fluctuate a lot. Because households face different histories of idiosyncratic shocks and accumulate different levels of capital stocks, they are *ex post* heterogeneous.

Taking the real return of capital ( $r_t$ ) and wage rate ( $w_t$ ) of efficient labor services as given, households solve the following utility maximization problem:

$$\begin{aligned} \max_{\{c_t, a_{t+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t u(c_t) \right\}, \\ \text{s.t. } c_t + a_{t+1} = (1 + r_t - \delta) a_t + w_t e_t, \\ a_{t+1} \geq b, \quad a_0 \text{ given.} \end{aligned}$$

Here,  $c_t$  and  $a_{t+1}$  denote consumption and the desired capital in time  $t$ . The households have two sources of income: net capital income  $r_t a_t$  and efficient labor income  $w_t e_t$ . Moreover, to prevent households from running a Ponzi scheme, it is assumed that households face an exogenously given borrowing limit  $b$ .<sup>8</sup> The period utility function takes the traditional constant relative risk aversion (CRRA) form:<sup>9</sup>

$$u(c) = \begin{cases} \frac{c^{1-\gamma}-1}{1-\gamma} & \text{if } \gamma \neq 1, \\ \ln(c) & \text{otherwise.} \end{cases}$$

Following most business cycle studies, the idiosyncratic efficient labor  $e_t$  is governed by the following first order autoregressive process:

$$\ln e_{t+1} = \rho_e \ln e_t + \varepsilon_{e,t+1}, \quad \varepsilon_{e,t+1} \stackrel{iid}{\sim} N(0, \sigma_e^2).$$

Note that the assumption for efficient labor process is rather simple. In the studies of the consumption function under partial equilibrium, the labor income process is usually more com-

<sup>8</sup>See Aiyagari (1994) for more discussion about the borrowing limit.

<sup>9</sup>This assumption make it become straightforward to use the algorithm we proposed in the previous chapter.

plicated than what we adopt here. Usually, researchers accommodate both permanent and transitory shocks. See Carroll (2007) and Gourinchas and Parker (2002) for a discussion of this issue. Besides, our specification of idiosyncratic shocks is also different from Krusell and Smith (1998), where the idiosyncratic shock takes only two values (0 or 1) for corresponding to employment status.

Two things deserve a discussion here. First, we assume that household's labor supply is inelastic because households do not value leisure.<sup>10</sup> Second, we take a rather simple stand on the sources of heterogeneity. In the model, the source of heterogeneity is solely stemming from the different realizations of efficient labor shocks among households. In other words, we attribute the source of wealth distribution to luck. We recognize that this is a rather restrictive assumption. For example, the origin of the wealth distribution might be based on the attitude of saving or personal preference. See Krusell and Smith (1998) for the effect of preference differences on the wealth distribution. Browning, Hansen, and Heckman (1999) and Heathcote et al. (2009) also provide further discussion on other heterogeneities, such as education and marriage.

**Firm Problem** There is a representative firm operating in perfectly competitive product and factor markets. It produces a homogeneous good by a typical Cobb-Douglas production function

$$y_t = z_t K_t^\alpha L_t^{1-\alpha},$$

where  $K_t$  and  $L_t$  are the aggregate physical capital and efficient labor, respectively.  $z_t$  is the stochastic aggregate productivity shock, and follows an AR(1) process

$$\ln z_{t+1} = \rho_z \ln z_t + \varepsilon_{z,t+1}, \quad \varepsilon_{z,t+1} \stackrel{iid}{\sim} N(0, \sigma_{z,t}^2).$$

---

<sup>10</sup>It is merely a practical assumption. It is possible, however, to generalize the model to allow for elastic labor supply, see Chang and Kim (2007) for instance, with extra computational burden.

Note also that the assumption of continuous support for the aggregate shock is different from Krusell and Smith (1998), but it is in line with Chang and Kim (2007).<sup>11</sup>

The representative firm's profit maximization problem implies the two marginal productivity condition:

$$\begin{aligned}r_t &= (1 - \alpha) z_t K_t^\alpha L_t^{1-\alpha}, \\w_t &= \alpha K_t^\alpha L_t^{1-\alpha},\end{aligned}$$

Since the efficient labor supply of the household is inelastic and governed by a exogenous process, the aggregate labor  $L_t$  is exogenously determined as well. The wage rate and the return of capital are, therefore, mainly determined by the aggregate capital. We will discuss how to calculate the aggregate capital and labor later. Moreover, the equilibrium concept of the model is the Recursive Competitive Equilibrium (RCE). Since the definition of the equilibrium is fairly standard, we leave the discussion of it to Appendix A.

## 3.4 Empirical Methodology

### 3.4.1 Some Technical Necessaries

It is well known that the representative agent DSGE model is not difficult to solve. For example, if first order approximation methods are adopted, the available methods consist of Christiano (2002), Klein (2000), and Sims (2002), just name a few. If researchers are not satisfied with the linearization method, they can even go for some higher order methods, such as Schmitt-Grohe and Uribe (2004). To implement these local approximation methods, researchers usually do not have to worry about any technical details once the steady state of the model is found. Currently, there exist several types of software for this purpose.<sup>12</sup>

---

<sup>11</sup>In the setting of Krusell and Smith (1998), the aggregate shock takes only two values, either boom or recession. As explained before for the case of idiosyncratic shocks, this type of assumption is unrealistic for meaningful business analyses, since one would like to match the model to the real data.

<sup>12</sup>For example, see website of the Dynare project.

However, if the heterogeneous agent framework is adopted, some technical details are necessary to solve the model. Specifically, due to the common assumption of existing borrowing constraints, researchers have to adopt global approximation methods, such as the value function iteration method, the policy function iteration method, or the projection method. To implement these methods, a common prerequisite is to specify the corresponding space for the state variables, such as the upper and lower limits of aggregate and individual capital stock. After the boundaries of state variables are properly determined, one can construct the corresponding grid for the continuous state variables. Besides, since computers can only deal with discrete numbers, exogenous shocks with continuous support also need to be discretized. We discuss these technical details below.

**Grid of the State Variables** Our strategy can be summarize as follows. For any given parameter vector  $\theta$ , we first solve the representative agent version of the model and find the steady state value of capital,<sup>13</sup> denoted as  $k_{ss}(\theta)$ . Then, we construct a equally-spaced grid

$$\mathcal{K} = \{k_1, k_2, \dots, k_{N_k}\},$$

where  $k_1 = (1-\eta)k_{ss}(\theta)$  and  $k_{N_k} = (1+\eta)k_{ss}(\theta)$ , because aggregate capital of the heterogeneous agent economy is not far from that of the representative agent version.<sup>14</sup> In our implementation, we set  $\eta = 0.3$ . As for the boundary of individual capital, we also construct an equally-spaced grid

$$\mathcal{A} = \{a_1, a_2, \dots, a_{N_a}\}$$

where  $a_{N_a} = \nu k_{ss}(\theta)$ . In our implementation, we set  $\nu = 5$  because households in heterogeneous agents economy might accumulate more capital than in the representative agent economy.

---

<sup>13</sup>There is no difference between individual or aggregate capital in the representative agent economy.

<sup>14</sup>See Aiyagari (1994) for further discussion.

**Discretization of the AR(1) process** For any given AR(1) process as follows

$$x_t = \rho x_{t-1} + \epsilon_t, \quad \epsilon_t \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2),$$

we would like discretize this process by a discrete first order Markov Chain. Specifically, one needs to specify a restricted space  $\mathcal{X}_N = \{x_1, x_2, \dots, x_{N_x}\}$  such that the original random variable can only take the values in the element of  $\mathcal{X}_N$ . At the same time, one also needs to construct a transition matrix

$$\Pi^x = \begin{pmatrix} \pi_{11}^x & \cdots & \pi_{1N_x}^x \\ \vdots & \ddots & \vdots \\ \pi_{N_x 1}^x & \cdots & \pi_{N_x N_x}^x \end{pmatrix}.$$

Since this is a transition matrix, the sum of each row should be one

$$\sum_j^{N_x} \pi_{ij}^x = 1, \quad \forall i \in \{1, \dots, N_x\}.$$

In the literature, many methods exist to discretize the AR(1) process, such as Tauchen (1986), Adda and Cooper (2003), and Tauchen and Hussey (1991). In this chapter, we use the method of Rouwenhorst (1995) to discretize the idiosyncratic and aggregate shock processes. Kopecky and Suen (2010) show that this method has several good properties, such as the consistence of the conditional (and unconditional) mean and variance between the discretized and original process.<sup>15</sup> Once we obtain the transition matrix, it is straightforward to calculate the corresponding ergodic distribution  $\tilde{\pi}^x = [\tilde{\pi}_1^x, \dots, \tilde{\pi}_N^x]'$ .<sup>16</sup>

<sup>15</sup>Rouwenhorst's discretization methods are not limited to the univariate AR(1) process. It can also be generalized to the multivariate case. See the paper mentioned in Kopecky and Suen (2010).

<sup>16</sup>The ergodic distribution is the eigenvector of  $(\Pi^x)'$  with unit eigenvalue. In practice, it is convenient to use the iteration method to find ergodic distribution. See Judd (1998) for further detail of finding the ergodic distribution.

Using the above method, it is clear to see that the aggregate efficient labor is simply

$$L_t = \sum_i^{N_e} \exp(\bar{e}_i) \tilde{\pi}_i^e, \quad \forall t,$$

where  $\tilde{\pi}_i^e$  is the  $i^{\text{th}}$  element of the associated ergodic distribution. Thus, aggregate labor is constant over time.

### 3.4.2 The Estimation Strategy

As explained before, we combine two simulation-based estimation methods to estimate the parameters of the heterogeneous agent model. We briefly discuss the idea of these two methods.

**Indirect Inference.** This method is first proposed by Smith (1993), and Gourieroux, Monfort, and Renault (1993) generalized it to allow the presence of exogenous variables. The main idea of Indirect Inference is to use an auxiliary model to capture the important aspects of real data. If one can simulate the structural model, then one can also use the simulated data to obtain auxiliary parameter estimates. Indirect Inference chooses structural parameters in such a way that the estimated auxiliary parameters from both real and simulated data are as close as possible. Indirect Inference is particular useful when the likelihood function of the data is not tractable. This is exactly our case. In a heterogeneous agent model with aggregate shocks, what one obtains is a sequence of distributions rather than variables; thus, it is hard to write down the likelihood function.

The indirect inference estimator is defined by following criterion function

$$L_T(\theta) = \max_{\theta} -\frac{1}{2}T \left[ \hat{\beta}_T - \tilde{\beta}_{ST}(\theta) \right]' \hat{W}_T \left[ \hat{\beta}_T - \tilde{\beta}_{ST}(\theta) \right] \quad (3.1)$$

where  $\hat{\beta}_T$  is the estimates from the auxiliary model, and  $\tilde{\beta}_{ST}$  is the estimates from the simulated data set. The subscripts of  $\hat{\beta}_T$  and  $\tilde{\beta}_{ST}$  indicate the sample size of data used to estimate the auxiliary model.  $\hat{W}_T$  is a positive definite matrix.  $\theta$  is the vector of structural parameters to be

estimated.

Our auxiliary model consists of two parts. The first part is a bivariate VAR with lag 2:

$$\begin{bmatrix} \bar{c}_t \\ \bar{y}_t \end{bmatrix} = \begin{bmatrix} \phi_c \\ \phi_y \end{bmatrix} + \begin{bmatrix} \phi_{c,c}^{(1)} & \phi_{c,y}^{(1)} \\ \phi_{y,c}^{(1)} & \phi_{y,y}^{(1)} \end{bmatrix} \begin{bmatrix} \bar{c}_{t-1} \\ \bar{y}_{t-1} \end{bmatrix} + \begin{bmatrix} \phi_{c,c}^{(2)} & \phi_{c,y}^{(2)} \\ \phi_{y,c}^{(2)} & \phi_{y,y}^{(2)} \end{bmatrix} \begin{bmatrix} \bar{c}_{t-2} \\ \bar{y}_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{c,t-2} \\ \epsilon_{y,t-2} \end{bmatrix}, \quad (3.2)$$

where  $\mathbb{Y}_t = [\bar{c}_t, \bar{y}_t]'$ .  $\bar{c}_t$  is the (per capita) aggregate consumption in time  $t$ , and  $\bar{y}_t$  is the (per capita) aggregate personal disposable income at time  $t$ . The second part of our auxiliary model consists of three AR(2) processes on the quantiles of the consumption distribution in time  $t$ .

$$\tilde{c}_{j,t} = \rho_{j,0} + \rho_{j,1}\tilde{c}_{j,t-1} + \rho_{j,2}\tilde{c}_{j,t-2} + \xi_{j,t}. \quad j = 1, 2, 3. \quad (3.3)$$

where  $j = 1, 2, 3$  correspond to the 25<sup>th</sup>, 50<sup>th</sup>, and 75<sup>th</sup> percentile of cross-sectional consumption distribution.

Therefore, we can stack the auxiliary parameters into a column vector<sup>17</sup>

$$\beta = \left[ \phi_{c,c}^{(1)}, \phi_{c,y}^{(1)}, \phi_{c,c}^{(2)}, \phi_{c,y}^{(2)}, \phi_{y,c}^{(1)}, \phi_{y,y}^{(1)}, \phi_{y,c}^{(2)}, \phi_{y,y}^{(2)}, \rho_{25,0}, \rho_{25,1} \right].$$

The bivariate VAR part is very standard.<sup>18</sup> It helps us to trace the joint dynamics of average disposable income and consumption. The AR(2) process is new. Since the heterogeneous agent model can give us the dynamics of the whole consumption distribution, we can use the quantile evolution to trace how the quantiles evolve over time.

Theoretically, this indirect inference estimator allows us to estimate the parameters of the heterogeneous agent model. However, similar to non-linear GMM or MLE, finding the global minimum of the objective function is a difficult task. Thus, we use the Laplace-type estimator (LTE) proposed by Chernozhukov and Hong (2003) to overcome this obstacle.

<sup>17</sup>Don't confuse this auxiliary parameters vector with the subjective discount factor

<sup>18</sup>Smith (1993) uses a bivariate VAR as the auxiliary model.



**Laplace Type Estimator.** LTE is a computationally attractive alternative to the traditional estimators. Chernozhukov and Hong (2003) show that LTE has a computational advantage in dealing with cases where the criterion functions have multiple local modes or are highly non-convex, even though the global optimum are well pronounced.

The logic of LTE is to transform a criterion function to a well-defined density function by a Laplace transformation. The criterion function does not have to be a likelihood function. In fact, the criterion function can be that of GMM, nonlinear IV, empirical likelihood (EL), or classical minimum distance (CMD) methods. Once the deduced density function, or quasi-posterior function, is obtained, researchers can use the standard Markov Chain Monte Carlo (MCMC) method to conduct random draws with respect to it. Thus, LTE is sometimes called the Quasi Bayesian Estimation (QBE) method.

The quasi-posterior function  $p_n(\theta)$  is obtained by the following transformation:

$$p_T(\theta) = \frac{e^{L_T(\theta)}\pi(\theta)}{\int_{\Theta} e^{L_T(\theta)}\pi(\theta)d\theta}.$$

where  $L_n(\theta)$  is the criterion function to be transformed, and  $\pi(\theta)$  is the prior distribution. In our implementation,  $L_T(\theta)$  is the criterion function from Indirect inference, that is, Eq. (3.1).

## 3.5 The Data

As explained in the introduction, this chapter coordinates the aggregate data from NIPA and household level data from Consumption Expenditure Survey (CEX) for our empirical studies. We explain how to reconcile these two data sets in this section.

### 3.5.1 NIPA Data

While NIPA data are widely used in modern business cycle research, the raw NIPA data are not directly applicable for our estimation.<sup>19</sup> First, in the definition of NIPA, GDP data contain both

---

<sup>19</sup>The raw NIPA data can be downloaded from the FRED database of St. Louis Federal Reserve Bank.

net exports and government spending. However, our model contains no foreign or government sector. Second, durable good consumption is classified as part of total consumption in NIPA, but our model does not include durable good consumption. In our model, durable good consumption is better attributed to investment. Therefore, when we use the NIPA data for our estimation, we need to transform it accordingly.

We first define aggregate consumption ( $C_t$ ) as the sum of non-durable good consumption (PCDG) and services (PCESV) of personal consumption expenditure. Investment ( $I_t$ ) is defined as the sum of Gross Private Domestic Investment (GPDI) and Personal Consumption Expenditures on Durable Goods. The aggregate output ( $Y_t$ ) is defined as the sum of aggregate consumption and investment. Aggregate variables are then transformed to real terms by the GDP deflator to obtain real aggregate variables. After that, we divide aggregate variable by non institutional civilian population to get the per capita real aggregate variables (i.e.,  $c_t$ ,  $i_t$ , and  $y_t$ ).

These variables are not stationary, and thus we use the HP filter to remove the trend components of the logarithms of these series. The filtered per capita variables are denoted as  $\bar{c}_t$ ,  $\bar{i}_t$ , and  $\bar{y}_t$ , which are the variables used in (3.2).

### 3.5.2 CEX Data

The Consumption Expenditure Survey (CEX) of the Bureau of Labor Statistics is the most comprehensive consumption data at the household level in the United States. There are two types of survey data in the CEX: Interview Survey and Diary Survey. Diary survey data mainly focus on small, frequently purchased items. On the other hand, the Interview Survey data focuses on the major items of expenses and covers up to 95 percent of the typical household's consumption expenditure. Besides, Interview Survey data is quarterly-based and congruent with most of the business cycle studies. Therefore, we only use it to construct the necessary measures of consumption distribution.

Similar to dealing with NIPA data, we focus on the non-durable goods consumption and

services as our consumption measure. Specifically, our consumption measure contains expenditures on food, alcoholic beverages, tobacco, utilities, personal care, household operations, public transportation, gas and motor oil, apparel, health, education, reading, and miscellaneous expenses.

It is well known that there is a large gap between consumption of CEX and that of NIPA. This is first documented by Slesnick (1992) and also discussed by Heathcote et al. (2010). To simultaneously incorporate CEX and NIPA data, we need a method to fill the gap. Our method can be explained as follows.

Step 1: Deflate the household consumption data by the GDP deflator. Denote the deflated  $i^{th}$  household consumption at time  $t$  as  $C_{it}$ .

Step 2: Divide  $C_{i,t}$  by the number of adults ( $n_{it}$ ) in the household. Denote this per capital consumption as  $c_{i,t}$ .

Step 3: Construct the time  $t$  scale variable by

$$\omega_t = \frac{c_t}{\sum_i n_{i,t} c_{i,t} / \sum_i n_{i,t}},$$

where  $c_t$  is real per capital aggregate consumption in time  $t$ .

Step 4 Obtain the scaled per capita household consumption  $c_{i,t}$  by

$$\hat{c}_{i,t} = \omega_t c_{i,t}$$

The advantage of the above procedure is that the per capita consumption from both the CEX and NIPA will be exactly the same. Our procedure is similar to Laitner and Silverman (2005).

In our heterogeneous agent model, the only heterogeneities are different levels of asset holdings and different idiosyncratic labor productivity shocks. However, there are a lot of demo-

Table 3.1: Meaning of Control Variables

Control	Description
FAMSIZE	family size
ASCOMP <sub>1</sub>	number of males age 16 and over
ASCOMP <sub>2</sub>	number of females age 16 and over
ASCOMP <sub>3</sub>	number of males age 2 to 15
ASCOMP <sub>4</sub>	number of females age 2 to 15
ASCOMP <sub>5</sub>	number of member under 2
PERSLT18	number of children under 18
PERSOT64	number of persons over 64
AGEREF	age of reference person

graphic differences in the data. Hence, we need to control for the differences such that we can map the consumption data into the model. Specifically, we run the following regression

$$\log(\hat{c}_{it}) = \vartheta_0 + \sum_{i=1}^5 \vartheta_i \text{ASCOMP}_i + \vartheta_6 \text{PERSLT18} + \vartheta_7 \text{PERSOT64} + \vartheta_8 \text{AGEREF} + \tilde{c}_{i,t}.$$

The meaning of these control variables are listed in Table 3.1.

Thus, the proper cross-sectional consumption in time  $t$  is represented by the residual consumption ( $\tilde{c}_{i,t}$ ). With this cross-sectional consumption distribution, we can calculate many descriptive statistics such as mean, variance, and of course, percentiles. The percentiles used in Eq. (3.3) are based on this.

### 3.6 Empirical Results

In addition to preventing daunting global minimization problems, the other advantage of LTE is that it can incorporate a prior density function  $\pi(\theta)$  with respect to the vector of structural parameters  $\theta$ . In this chapter, the structural parameters are

$$\theta = [\beta, \gamma, \alpha, \delta, \rho_e, \sigma_e, \rho_z, \sigma_z]'$$

Table 3.2: Prior Distributions of Structural Parameters

Parameter	Description	Prior		
		Type	Para(1)	Para(2)
$\beta$	Subjective discount factor	Beta	0.97	0.01
$\gamma$	Risk aversion coef.	Gamma	1.3	0.5
$\alpha$	Capital share	Beta	0.35	0.05
$\delta$	Depreciation rate	Beta	0.025	0.01
$\rho_e$	Persistence Coef. of $e$ shock	Beta	0.93	0.01
$\sigma_e$	Std. dev. of $e$ innovation	InvGamma	0.2	0.05
$\rho_z$	Persistence Coef. of $z$ shock	Beta	0.78	0.05
$\sigma_z$	Std. dev. of $z$ innovation	InvGamma	0.007	0.002

The prior distributions of structural parameters are listed in Table 3.2.

Since we have the quasi-posterior density function of structural parameters, we use the random walk Metropolis-Hastings (RWM) method to conduct the MCMC draws. This method is widely used in Bayesian estimation for structural estimation. See An and Schorfheide (2007) for details of this method. Our result is based on 50,000 MCMC draws, and the first 30,000 draws are discarded. We discuss the meaning of our estimates below.

First, the posterior median of  $\beta$  is 0.9756. This estimate is lower than that obtained from representative agent models. Since our estimation combines both aggregate and consumption distributional data, this lower estimate must come from the presence of consumption distributional data. Moreover, because this estimate is closer to micro-level studies, it implies that the usual calibration for  $\beta$  in HA model might be set too high. The typical calibration for  $\beta$  is 0.99. See Krusell and Smith (1998) for example. Our empirical result cast doubts on this commonly adopted value.

The posterior median of  $\gamma$  is 1.58. This estimate is located within the commonly accepted range and slightly larger than that of Gourinchas and Parker (2002). It is worth noting that this estimate is much smaller than that of Cagetti (2003), where the risk aversion estimates are usually larger, around 2 to 7. Comparing to Cagetti (2003), we think our estimate realistic because both the aggregate and consumption distributional data are used for estimation.

Table 3.3: Posterior of Structural Parameters

percentile	$\beta$	$\gamma$	$\alpha$	$\delta$	$\rho_e$	$\sigma_e$	$\rho_z$	$\sigma_z$
5th	0.9707	1.2963	0.2700	0.0108	0.8850	0.2792	0.7497	0.0022
50th	0.9756	1.5836	0.4459	0.0180	0.9246	0.4372	0.8233	0.0043
95th	0.9986	1.8608	0.4935	0.0284	0.9517	0.6122	0.9370	0.0104

The estimate of the capital share is 0.4459 which is larger than the widely accepted calibration value, which is around 0.35. One potential explanation for this result is that we assume that capital is the only asset for precautionary saving. However, in the real world, households have many different tools for this purpose, such as bonds, money, and equities.

The median of  $\rho_e$  is 0.9246 which indicate that the idiosyncratic shocks are persistent. Moreover, our estimate for  $\rho_e$  is 0.8233 which is also smaller than the commonly used calibration value, 0.95. This result indicates that when the dynamics of the distribution of consumption are taken into account, the aggregate technology shocks seem to be not very persistent. The implication of this result is that the dynamics of the distribution of consumption tell a different story from that of aggregate data alone.

### 3.7 Conclusion

In this chapter, we estimate a HA model by utilizing two sources of data: the first one is repeated cross-sectional data from the Consumer Expenditure Survey (CEX), and the second one is aggregate data from National Income and Production Accounts (NIPA) of the United States. Our estimation is based on the Laplace-type estimator (LTE) with a criterion function from the indirect inference. This chapter is methodologically oriented, and focuses on proposing a recipe for conducting the structural estimation of HA models, a task that has not yet been done in the literature.

## Chapter 4

# Identifying Sources of Misspecification in DSGE Models

### 4.1 Introduction

In this chapter, we examine the empirical importance of model misspecification in dynamic stochastic general equilibrium (DSGE) models. DSGE models are commonly used in academia and central banks alike and are standard tools for analyzing macroeconomic policies. These macroeconomic models are often highly parameterized to fit better to data, and that makes it more likely that they are misspecified. In a recent paper, Del Negro, Schorfheide, Smets, and Wouters (2007) raise this very issue and show that model misspecification cannot be ignored in policy analyses. While model misspecification may be widespread, macroeconomists are often left to wonder if their models are misspecified and, if they are, what parts of models are misspecified. In this chapter, we propose an empirical framework that is useful for addressing these issues.

To illustrate our framework, we consider a New Keynesian DSGE model based on Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007). Our model is mildly misspecified in the sense that (i) the cross equation and equilibrium restrictions imposed by the model

do not exactly hold in every time period; and (ii) these deviations are zero on average. We assume that the economic agents in the model (both firms and households) take into account exogenous stochastic processes of deviations when they solve their optimization problems. The variances of deviations are a measure of the degree of misspecification of the New Keynesian DSGE model without a distortion. To address the aforementioned issues, we conduct impulse response analyses for different priors on the variances of deviations and forecast error variance decompositions (FEVD) to locate possible sources of misspecification. Our method is closely related to Chari, Kehoe, and McGrattan (2007) who also introduce time-varying “wedges” into a macroeconomic model. There are two main differences. One difference is that they consider a stochastic growth model with wedges (what they call the benchmark prototype economy) while we consider a New Keynesian DSGE model with wedges. Our New Keynesian DSGE model is based on Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007) and incorporates frictions. Thus our distortions reflect model misspecification that cannot be accounted for by the built-in frictions. The other difference is that their analysis is based on calibrating parameter values while ours is based on estimated parameters that take into account possible misspecification.

Another closely related work is Del Negro and Schorfheide (2009). Using the method developed in Del Negro and Schorfheide (2004), they develop a framework for Bayesian estimation of possibly misspecified DSGE models. Specifically, they use DSGE-implied parameters as a prior for vector autoregressive (VAR) models. Their framework allows for model misspecification and produces the posterior distribution of structural parameters as well as the posterior structural impulse responses implied by DSGE-implied priors. Our framework complements theirs in that we can identify where the model is misspecified in our framework.

Our empirical results show that, even though the specification of the New Keynesian DSGE model is much richer than the stochastic growth model considered in Chari, Kehoe and McGrattan (2007), the model is still largely misspecified. Specifically, the forecast error variance decomposition (FEVD) exercise reveals that, for all aggregate variables used to estimate the



model, the labor market misspecification contributes more than 40 percent of their total variations. This suggests that macroeconomists need to model the labor market in DSGE models in a more realistic way. We also find that whether there is a price puzzle or not depends on our prior on the specification of models. When we allow for model misspecification, the price puzzle shows up. However, when we restrict the model to be almost correctly specified, the price puzzle seems to become weaker. That is, disappearance of the price puzzle does not necessarily imply that the model is correctly specified.

The rest of the chapter is organized as follows: Section 4.2 introduces a New Keynesian DSGE model in which wedges are introduced as serially correlated taxes. Section 4.3 describes the solution and estimation methods used as well as the dataset. Section 4.4 shows the empirical results, and Section 4.5 provides 3 simulation exercises to demonstrate our methodology. Section 4.6 concludes the chapter. Appendix B contains some detailed derivations of Section 4.2 and presents the log-linearized versions of the misspecified models on which our simulation exercises are based.

## 4.2 The Model

The model is a micro-founded macroeconometric model which contains a fair amount of deviations from the typical stochastic neoclassical growth model. In line with Christiano et al. (2005), the model contains four real-side frictions: the *internal* habit formation, the investment adjustment cost, the variable capital utilization, and the working capital channel. Altig et al. (2011) and Christiano et al. (2005) discuss thoroughly the justification of these frictions. Following the well-adopted settings of Calvo (1983) and Erceg et al. (2000), the model allows for nominal rigidities in both price and wage. In particular, we presume that there exists a *full* price and wage indexation scheme with respect to the inflation rate when firms and households are not able to set the price and wage at the optimal level.<sup>1</sup>

---

<sup>1</sup>In contrast, Smets and Wouters (2007) and Levin et al. (2006) assume that firms and households adopt partial indexation scheme.

Similar to the setting of Levin et al. (2006), there is no long run growth in the model. Currently, some of the macroeconomic models take the other modeling strategy in that the model economy embeds explicit long-run growth. For example, Smets and Wouters (2007) assume that the economy grows along a deterministic constant growth rate. Altig, Christiano, Eichenbaum, and Linde (2011) assume that both the neutral technology shock and the capital embedded technology shock are random walks with drifts. The advantage of this alternative modeling strategy is that researchers are able to detrend the model and data in exactly the same manner. However, this method is not flawless because it forces us to look at mainly the higher frequency part of the data. The problem stems from the fact that the higher frequency part of the data is usually more noisy. Thus, the estimation results are based on more noisy data, which might make the results less reliable.

Asides from various wedges, there are two structural shocks in the model. First, there is a monetary shock that affects the nominal interest rate. Second, there exists a neutral technology shock in the intermediate firms production function.<sup>2</sup> Moreover, we adopt a *cashless* model setting, which is suggested by Woodford (2003) and popular among many macroeconomists; see Smets and Wouters (2007). In other words, we do not adopt the conventional money-in-utility (MIU) or cash-in-advance (CIA) model settings. This modeling strategy implies that the money market always clears, and the central bank provides sufficient money to fit the potential money demand.

Since our main goal in the chapter is to investigate the sources of model misspecification, which are captured by various wedges, the way we introduce these wedges plays an important role for our study. Generally speaking, various wedges are included into the model in such a way that wedges act like distortionary taxes for firms and households. The motivations of such a setting are two fold. First, we believe these wedges are capable of representing potential information facing firms and households that are not observed by econometricians. This implies that firms and households respond to not only the relative prices of factors and goods observed

---

<sup>2</sup>In the benchmark case of Christiano et al. (2005), only money supply shocks are present.

in the data, but also something else that affects the *effective* relative prices. Second, as explained in Chari et al. (2007), many dynamic macroeconomic models with frictions are equivalent to a *prototype* neoclassical growth model with various wedges. However, in contrast to Chari et al. (2007) where the prototype model is rather simple, the model in our chapter is much more complicated and includes many frictions. Thus, the question that we seek to answer is how severe the model misspecification becomes in a widely adopted DSGE model.

### 4.2.1 The Final Good Firm Problem

At time  $t$ , the representative final good firm produces the homogeneous final good  $Y_t$  by combining a continuum of intermediate goods  $Y_{j,t}$ , using the following technology:

$$Y_t = \left[ \int_0^1 (Y_{j,t})^{\frac{1}{\lambda_f}} dj \right]^{\lambda_f}, \quad (4.1)$$

where  $1 \leq \lambda_f < \infty$  governs the extent of substitution among different intermediate goods.<sup>3</sup> The final good market is perfectly competitive, and the representative firm takes the final good price  $P_t$  and intermediate goods prices  $P_{j,t}$ 's as given. The corresponding profit maximization problem of the final good firm is

$$\max_{\{Y_{j,t}\}_{j \in [0,1]}} P_t Y_t - \int_0^1 \tau_t^y P_{j,t} Y_{j,t} dj,$$

subject to its production function (4.1). In order to measure the extent of model misspecification on intermediate good markets, intermediate goods price wedge,  $\tau_t^y$ , is introduced into the model. The price wedge  $\tau_t^y$  is common across all intermediate goods. Thus, for any particular  $j^{\text{th}}$  intermediate good, the effective price is actually the product of  $\tau_t^y$  and  $P_{j,t}$ . Moreover,  $\tau_{y,t}$  is assumed as known at time  $t$  for the representative final good firm. Besides, it is convenient to

---

<sup>3</sup>This production function is the so-called Stiglitz-Dixit aggregator and widely used in modern New Keynesian models; See Christiano et al. (2005), Schmitt-Grohe and Uribe (2006), and Smets and Wouters (2003). Besides, the production function proposed by Kimball (1995) and later used in Smets and Wouters (2007) is also proper for the model that we consider.

interpret  $(\tau_t^y - 1) P_{j,t} Y_{j,t}$  as a consumption tax levied by the government for the  $j^{\text{th}}$  intermediate input. We assume that the logarithm of  $\tau_t^y$  follows the first order autoregressive process, AR(1):

$$\log \tau_{t+1}^y = (1 - \rho_y) \log \tau^y + \rho_y \log \tau_t^y + \sigma_y \varepsilon_{t+1}^y, \quad \varepsilon_{t+1}^y \stackrel{iid}{\sim} N(0, 1), \quad (4.2)$$

where  $\rho_y$  controls the persistence of the wedge,  $\tau^y$  denotes the steady-state wedge, and  $\sigma_y$  governs the volatility of the wedge. The first order necessary condition with respect to  $Y_{j,t}$  of the problem implies that the demand function for intermediate goods is:

$$Y_{j,t} = \left( \frac{P_t}{\tau_t^y P_{j,t}} \right)^{\theta_f} Y_t, \quad \forall j \in [0, 1] \quad (4.3)$$

where  $\theta_f \equiv \frac{\lambda_f}{\lambda_f - 1}$ . From (4.3), it is clear that the introduction of  $\tau_t^y$  breaks the tight relationship among  $P_t$ ,  $Y_t$ ,  $P_{j,t}$  and  $Y_{j,t}$ . That is, we allow the possibility of the model being misspecified in describing the demand functions of intermediate goods.

#### 4.2.2 The Intermediate Good Firm Problem

In the model, each intermediate good firm is a monopolistic provider of its differentiated good, and able to charge its own price to maximize its monopoly rent. A traditional two-stage procedure utilizing cost minimization and profit maximization is used to reflect the firm's optimal behavior.

**The Cost Minimization Problem.** At time  $t$ , the  $j^{\text{th}}$  intermediate good firm rents capital services  $K_{j,t}$ , and *homogeneous* labor services,  $L_{j,t}$  from perfectly competitive factor markets. Taking the intermediate output level  $Y_{j,t}$ , the nominal capital service rental rate  $R_t^K$ , and the nominal wage rate  $W_t$  as given, it solves the following problem:

$$\min_{L_{j,t}, K_{j,t}} (\tau_t^K R_t^K) K_{j,t} + R_t (\tau_t^L W_t) L_{j,t}$$

subject to the production function

$$Y_{j,t} = \begin{cases} z_t (K_{j,t})^\alpha (L_{j,t})^{1-\alpha} - \Phi, & \text{if } z_t (K_{j,t})^\alpha (L_{j,t})^{1-\alpha} \geq \Phi; \\ 0, & \text{otherwise.} \end{cases}$$

Here,  $\alpha$  is the capital share,  $\Phi$  is the fixed cost of production, and  $z_t$  is the time  $t$  technology shock that is common across intermediate good firms. It is also assumed that there are no exit nor entry decisions for intermediate good firms. The neutral technology shock  $z_t$  follows the AR(1) process

$$\log z_{t+1} = (1 - \rho_z) \log z + \rho_z \log z_t + \sigma_z \varepsilon_{t+1}^z, \quad \varepsilon_{t+1}^z \stackrel{iid}{\sim} N(0, 1). \quad (4.4)$$

It is useful to discuss two issues with this setup. First, two wedges are introduced to handle the model misspecification problem: the first is the capital market wedge  $\tau_t^K$ , and the second is the homogeneous labor market wedge  $\tau_t^L$ . Note that they are assumed to be known for the intermediate good firm at time  $t$ , and thus are contained in the time  $t$  information set of intermediate good firms. As before, these two wedges follow AR(1) processes in logarithms:

$$\log \tau_{t+1}^K = (1 - \rho_K) \log \tau_t^K + \log \tau_t^K + \sigma_K \varepsilon_{t+1}^K, \quad \varepsilon_{t+1}^K \stackrel{iid}{\sim} N(0, 1), \quad (4.5)$$

$$\log \tau_{t+1}^L = (1 - \rho_L) \log \tau_t^L + \log \tau_t^L + \sigma_L \varepsilon_{t+1}^L, \quad \varepsilon_{t+1}^L \stackrel{iid}{\sim} N(0, 1), \quad (4.6)$$

The interpretation of parameters in these two processes are the same as that of the price wedge process in intermediate good markets.

Second, in line with Christiano et al. (2005), there exists a “working capital channel” in the model. That is, at the beginning of each period the intermediate good firms need to borrow from the financial market with nominal interest rate  $R_t$  to pay off the wage bill for hiring labor services. This feature, highlighted in Christiano et al. (1996), Chowdhury et al. (2006) and Ravenna and Walsh (2006), takes into account the possibility that conducting contractionary

monetary policy by raising nominal interest rate might adversely trigger a higher inflation rate because the higher nominal interest rate could raise the marginal costs of production. Moreover, it is helpful in explaining the so-called “price puzzle”. See Christiano et al. (2010) for further discussion. Note that the working capital channel is only imposed on the labor side, not on the capital side.

The Lagrange function of the above problem is

$$\mathcal{L} = (\tau_t^K R_t^K) K_{j,t} + R_t (\tau_t^L W_t) L_{j,t} + \lambda_t \left[ Y_{j,t} - z_t (K_{j,t})^\alpha (L_{j,t})^{1-\alpha} + \Phi \right],$$

and the corresponding first order necessary conditions are

$$\begin{aligned} \tau_t^K R_t^K &= \lambda_t \alpha z_t \left( \frac{K_{j,t}}{L_{j,t}} \right)^{\alpha-1}, \\ R_t \tau_t^L W_t &= \lambda_t (1 - \alpha) z_t \left( \frac{K_{j,t}}{L_{j,t}} \right)^\alpha, \\ Y_{j,t} &= z_t (K_{j,t})^\alpha (L_{j,t})^{1-\alpha} - \Phi. \end{aligned}$$

The first two first order conditions imply that the optimal capital-labor ratio is

$$\frac{K_{j,t}}{L_{j,t}} = \left( \frac{\alpha}{1 - \alpha} \right) \left( \frac{R_t \tau_t^L W_t}{\tau_t^K R_t^K} \right) = \frac{K_t}{L_t}. \quad (4.7)$$

Note that there exists no state variables in this problem, and that each intermediate good firm faces the same cost minimization problem. As a consequence, the ratio is common across firms and equal to the aggregate capital-labor ratio. The nominal marginal cost function is exactly equal to the Lagrange multiplier,

$$\lambda_t = \frac{1}{z_t} \left( \frac{1}{\alpha} \right)^\alpha \left( \frac{1}{1 - \alpha} \right)^{1-\alpha} (R_t \tau_t^L W_t)^{1-\alpha} (\tau_t^K R_t^K)^\alpha,$$

which can be derived by substituting the capital-labor ratio (4.7) into the first order condition

of capital demand (or labor demand). Thus, we have the real marginal cost,

$$s_t \equiv \frac{\lambda_t}{\bar{P}_t} = \frac{1}{z_t} \left( \frac{1}{\alpha} \right)^\alpha \left( \frac{1}{1-\alpha} \right)^{1-\alpha} (R_t \tau_t^L w_t)^{1-\alpha} (\tau_t^K r_t^K)^\alpha, \quad (4.8)$$

where  $w_t \equiv W_t/P_t$  denotes real wage rate, and  $r_t^K \equiv R_t^K/P_t$  is the real capital rental rate.

**The Profit Maximization Problem** The intermediate good firms seek to maximize the present value of expected future profit flow. Following Calvo (1983), for each intermediate good firm, the probability of being able to re-optimize its price is  $1 - \xi_p$ , which is independent across firms and common across time. When the intermediate good firm is not able to re-optimize its price, it resets price by the full indexation scheme:  $P_{j,t} = \pi_{t-1} P_{j,t-1}$ . Define  $X_{t,l}$  as

$$X_{t,l} \equiv \begin{cases} \pi_t \times \pi_{t+1} \times \cdots \times \pi_{t+l-1} & \text{for } l \geq 1; \\ 1 & \text{for } l = 0. \end{cases}$$

The time  $t$  profit maximization problem of intermediate goods firms is

$$\begin{aligned} \max_{\bar{P}_t} \mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_p)^l v_{t+l} \left[ \tilde{P}_t X_{t,l} - s_{t+l} P_{t+l} \right] Y_{j,t+l}, \\ \text{s.t. } Y_{j,t+l} = \left( \frac{P_{t+l}}{\tau_{t+l}^y \tilde{P}_t X_{t,l}} \right)^{\theta_f} Y_{t+l}. \end{aligned}$$

$\xi_p^l$  denotes the probability of not being able to set the optimal price at period  $t+l$ , given the price is set in period  $t$ . The constraint of the problem is the intermediate good demand function at time  $t+l$ ,  $Y_{j,t+l}$ . Note that this demand function takes into account the full indexation and the existence of the intermediate good price wedge  $\tau_{t+l}^y$ . The expression within the square brackets is the nominal unit profit. Given that the equities of intermediate good firms are assumed to belong to households,  $v_{t+l}$  is the marginal value of profit for households. Because

$$\frac{\partial Y_{j,t+l}}{\partial \tilde{P}_t} = -\theta_f \frac{Y_{j,t+l}}{\tilde{P}_t},$$

and  $\theta_f/(1-\theta_f) = -\lambda_f$ , the first order condition of the problem can be expressed as

$$\mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_p)^l v_{t+l} Y_{j,t+l} \left( \tilde{P}_t X_{t,l} - \lambda_f P_{t+l} s_{t+l} \right) = 0.$$

Defining the normalized optimal intermediate good price  $\tilde{p}_t \equiv \tilde{P}_t/P_t$  and  $\psi_t \equiv P_t v_t$ , we can rewrite the above expression as

$$\mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_p)^l \psi_{t+l} Y_{j,t+l} \left( \tilde{p}_t \frac{X_{t,l}}{P_{t+l}} - \lambda_f \frac{s_{t+l}}{P_t} \right) = 0. \quad (4.9)$$

By the definitions of  $X_{t,l}$  and gross inflation rate  $\pi_t = P_t/P_{t-1}$ , we have

$$\begin{aligned} \frac{X_{t,l}}{P_{t+l}} &= \frac{\pi_t \times \pi_{t+1} \times \pi_{t+2} \times \cdots \times \pi_{t+l-1}}{P_{t+l}} \\ &= \left( \frac{P_t}{P_{t-1}} \times \frac{P_{t+1}}{P_t} \times \frac{P_{t+2}}{P_{t+1}} \times \cdots \times \frac{P_{t+l-1}}{P_{t+l-2}} \right) / P_{t+l} \\ &= \frac{P_{t+l-1}}{P_{t-1} P_{t+l}} = \frac{1}{P_{t-1} \pi_{t+l}}. \end{aligned}$$

Hence, equation (4.9) becomes

$$\mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_p)^l \psi_{t+l} Y_{j,t+l} \left( \frac{\tilde{p}_t}{\pi_{t+l} P_{t-1}} - \lambda_f \frac{s_{t+l}}{P_t} \right) = 0.$$

Multiplying by  $P_t$  and rearranging terms, we have

$$\tilde{p}_t = \frac{\mathcal{N}_t^p}{\mathcal{D}_t^p}, \quad (4.10)$$

where

$$\begin{aligned} \mathcal{D}_t^p &= \pi_t \mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_p)^l \psi_{t+l} \frac{Y_{j,t+l}}{\pi_{t+l}} = \psi_t Y_{j,t} + \beta \xi_p \mathbb{E}_t \left\{ \frac{\pi_t}{\pi_{t+1}} \mathcal{D}_{t+1}^p \right\}, \\ \mathcal{N}_t^p &= \lambda_f \mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_p)^l \psi_{t+l} Y_{j,t+l} s_{t+l} = \lambda_f \psi_t Y_{j,t} s_t + \beta \xi_p \mathbb{E}_t \left\{ \mathcal{N}_{t+1}^p \right\}. \end{aligned}$$



The second equality is merely a result of the recursive representation and the law of iterated expectations. Imposing the optimal intermediate good demand function at time  $t$ ,  $\mathcal{D}_t^p$  and  $\mathcal{N}_t^p$  become

$$\mathcal{D}_t^p = \psi_t \left( \frac{P_t}{\tau_t^y \tilde{P}_t} \right)^{\theta_f} Y_t + \beta \xi_p \mathbb{E}_t \left\{ \frac{\pi_t}{\pi_{t+1}} \mathcal{D}_{t+1}^p \right\} = \psi_t (\tau_t^y \tilde{p}_t)^{-\theta_f} Y_t + \beta \xi_p \mathbb{E}_t \left\{ \frac{\pi_t}{\pi_{t+1}} \mathcal{D}_{t+1}^p \right\} \quad (4.11)$$

$$\mathcal{N}_t^p = \lambda_f \psi_t \left( \frac{P_t}{\tau_t^y \tilde{P}_t} \right)^{\theta_f} Y_t s_t + \beta \xi_p \mathbb{E}_t \{ \mathcal{N}_{t+1}^p \} = \lambda_f \psi_t (\tau_t^y \tilde{p}_t)^{-\theta_f} Y_t s_t + \beta \xi_p \mathbb{E}_t \{ \mathcal{N}_{t+1}^p \}. \quad (4.12)$$

### 4.2.3 The Household Problem

In the model, there exists a continuum of infinitely living households who are denoted by  $j \in [0, 1]$ . They make a sequence of decisions on consumption, capital accumulation, capital utilization, bond holding and idiosyncratic wage setting. Following Erceg et al. (2000) and Christiano et al. (2005), households are assumed to be monopolistically differentiated labor suppliers, and hence, are able to set their own wage rate. Furthermore, assume that the probability of being able to set the optimal wage rate at each period is  $1 - \xi_w$ , which is independent across households and common over time. On top of that, we assume that the markets are complete because there exists a complete set of Arrow-Debreu type financial assets. Under this assumption, the decisions on consumption, capital accumulation, capital utilization, and bond holding are common across households, even though the optimal wage rates among them are still different. Therefore, we attach household subscript  $j$  only on idiosyncratic variables (i.e.,  $H_{j,t}$ ,  $W_{j,t}$  and  $A_{j,t}$ ) but not on other variables.

The  $j^{\text{th}}$  household maximizes the expected lifetime utility function:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \{ U(C_t - bC_{t-1}) - \Psi(H_{j,t}) \}.$$

$\mathbb{E}_t$  denotes the mathematical expectations operator conditional on information available at time  $t$ , and  $\beta$  represents the subjective discount factor. The period utility function takes the usual

constant relative risk aversion (CRRA) form:

$$\begin{aligned} U(C_t - bC_{t-1}) &= \frac{(C_t - bC_{t-1})^{1-\gamma}}{1-\gamma}, \\ \Psi(H_{j,t}) &= \frac{\psi_h (H_{j,t})^{1+\varphi}}{1+\varphi}. \end{aligned}$$

$\gamma$  is the relative risk aversion coefficient.  $\varphi$  governs the wage elasticity of labor supply.  $\psi_h$  is a weight parameter for the disutility of labor supply.  $b$  is the habit formation parameter.

The period  $t$  budget constraint of the household is

$$B_{t+1} + \tau_t^Y P_t [C_t + I_t + a(u_t)\bar{K}_t] = \tau_t^B R_{t-1} B_t + W_{j,t} H_{j,t} + R_t^K u_t \bar{K}_t + A_{j,t} + D_t + T_t.$$

$C_t$  and  $I_t$  are respectively consumption and investment in period  $t$ .  $B_t$  is the nominal bond with one period maturity.  $R_{t-1}$  is the nominal interest rate which is determined in period  $t-1$ .  $W_{j,t}$  and  $H_{j,t}$  are the idiosyncratic wage rate and the labor supply.  $A_{j,t}$  is the return on holding Arrow-Debreu assets from the previous period. Given that the equities of intermediate good firms are held by households,  $D_t$  is the dividend from intermediate good firms.  $T_t$  denotes the lump sum transfer from the government.  $R_t^K$  is the rental rate of capital service. We assume that households can determine utilization rate  $u_t$  of capital service. In particular, the effective capital services used in the production of intermediate goods is determined by

$$K_t = u_t \bar{K}_t. \tag{4.13}$$

Moreover, different utilization rates implies different costs:  $P_t a(u_t) \bar{K}_t$ . We assume that there exists an investment adjustment cost and that the installed capital evolves by

$$\bar{K}_{t+1} = (1 - \delta) \bar{K}_t + \mathcal{F}(I_t, I_{t-1}), \tag{4.14}$$

with

$$\mathcal{F}(I_t, I_{t-1}) = \left[ 1 - \mathcal{S} \left( \frac{I_t}{I_{t-1}} \right) \right] I_t.$$

The function  $\mathcal{S}(\cdot)$  is restricted to be  $\mathcal{S}(1) = \mathcal{S}'(1) = 0$ . Lastly, two wedges,  $\tau_t^Y$  and  $\tau_t^B$ , are introduced to capture the potential misspecification in the bond and final good markets. They are assumed to follow AR(1) process:

$$\log \tau_{t+1}^Y = (1 - \rho_Y) \log \tau^Y + \rho_Y \log \tau_t^Y + \sigma_Y \varepsilon_{t+1}^Y, \quad \varepsilon_{t+1}^Y \stackrel{iid}{\sim} N(0, 1); \quad (4.15)$$

$$\log \tau_{t+1}^B = (1 - \rho_B) \log \tau^B + \rho_B \log \tau_t^B + \sigma_B \varepsilon_{t+1}^B, \quad \varepsilon_{t+1}^B \stackrel{iid}{\sim} N(0, 1). \quad (4.16)$$

As before, we assume that  $\tau^Y = \tau^B = 1$ . The Lagrange function of the household problem is

$$\begin{aligned} \mathcal{L} = & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \{ U(C_t - bC_{t-1}) - \Psi(H_{j,t}) \\ & + v_t [\tau_t^B R_{t-1} B_t + W_{j,t} H_{j,t} + R_t^K u_t \bar{K}_t + A_{j,t} + D_t + T_t - \tau_t^Y P_t (C_t + I_t + a(u_t) \bar{K}_t) - B_{t+1} \\ & - P_t P_{k',t} (\bar{K}_{t+1} - (1 - \delta) \bar{K}_t - \mathcal{F}(I_t, I_{t-1})) ] \}, \end{aligned}$$

and households make decisions on  $C_t$ ,  $\bar{K}_{t+1}$ ,  $I_t$ ,  $u_t$ ,  $B_{t+1}$ , and  $W_{j,t}$  (and hence,  $H_{j,t}$ ).

**Consumption.** By defining  $\psi_t \equiv P_t v_t$ , the first order condition of consumption is

$$\mathbb{E}_t \{ \mathcal{U}_{c,t} \} = \mathbb{E}_t \{ \tau_t^Y \psi_t \},$$

where

$$\begin{aligned} \mathcal{U}_{c,t} & \equiv \frac{\partial U(C_t - bC_{t-1})}{\partial C_t} + \beta \mathbb{E}_{t+1} \frac{\partial U(C_{t+1} - bC_t)}{\partial C_t} \\ & = (C_t - bC_{t-1})^{-\gamma} - \beta b \mathbb{E}_{t+1} (C_{t+1} - bC_t)^{-\gamma}. \end{aligned}$$

This condition can be further simplified as

$$\mathbb{E}_t \{ (C_t - bC_{t-1})^{-\gamma} - \beta b (C_{t+1} - bC_t)^{-\gamma} \} = \mathbb{E}_t \{ \tau_t^Y \psi_t \}. \quad (4.17)$$

It is clear that the presence of  $\tau_t^Y$  relaxes the tight relationship implied by the traditional consumption Euler equation.

**Capital.** The first order condition of physical capital  $\bar{K}_{t+1}$  is

$$\mathbb{E}_t \{ P_{k',t} \psi_t \} = \beta \mathbb{E}_t \{ \psi_{t+1} [r_{t+1}^K u_{t+1} - \tau_{t+1}^Y a(u_{t+1}) + P_{k',t+1}(1 - \delta)] \}. \quad (4.18)$$

**Investment.** The first order condition of  $I_t$  is

$$\mathbb{E}_t \{ \tau_t^Y \psi_t \} = \mathbb{E}_t \{ P_{k',t} \psi_t \mathcal{F}_{1,t} + \beta P_{k',t+1} \psi_{t+1} \mathcal{F}_{2,t+1} \}, \quad (4.19)$$

where

$$\begin{aligned} \mathcal{F}_{1,t} &\equiv \frac{\partial \mathcal{F}(I_t, I_{t-1})}{\partial I_t} = 1 - \mathcal{S}' \left( \frac{I_t}{I_{t-1}} \right) \frac{I_t}{I_{t-1}} - \mathcal{S} \left( \frac{I_t}{I_{t-1}} \right) \\ \mathcal{F}_{2,t+1} &\equiv \frac{\partial \mathcal{F}(I_{t+1}, I_t)}{\partial I_t} = \mathcal{S} \left( \frac{I_{t+1}}{I_t} \right) \frac{I_{t+1}}{I_t^2} \end{aligned}$$

**Capacity Utilization.** The first order condition of  $u_t$  is

$$\mathbb{E}_t \psi_t \{ r_t^K - \tau_t^Y a'(u_t) \} = 0. \quad (4.20)$$

**Bond Holding.** The first order condition of  $B_{t+1}$  is

$$\mathbb{E}_t \psi_t = \beta \mathbb{E}_t \left\{ \tau_{t+1}^B \frac{\psi_{t+1}}{\pi_{t+1}} R_t \right\}. \quad (4.21)$$

**The Wage Decision.** To reconcile the assumptions that intermediate good firm rent homogeneous labor from a competitive market and that households supply differentiated labor services, following Christiano et al. (2005), we introduce perfectly competitive “labor-packer” firms who hire differentiated household labor services to produce homogeneous labor service by the following technology:

$$L_t = \left[ \int_0^1 (H_{j,t})^{\frac{1}{\lambda_w}} dj \right]^{\lambda_w}. \quad (4.22)$$

Taking the market wage rate  $W_t$  and idiosyncratic wage rates  $W_{j,t}$  of households as given, the labor-packer firm’s profit maximization problem is

$$\max_{\{H_{j,t}\}_{j \in [0,1]}} W_t L_t - \int_0^1 \tau_t^h W_{j,t} H_{j,t} dj.$$

subject to equation (4.22). As usual, a differentiated labor market wedge  $\tau_t^h$  is introduced to capture the potential labor market misspecification, and it is assumed to follow

$$\log \tau_{t+1}^h = (1 - \rho_h) \log \tau_t^h + \rho_h \log \tau_t^h + \sigma_h \varepsilon_{t+1}^h, \quad \varepsilon_{t+1} \stackrel{iid}{\sim} N(0, 1). \quad (4.23)$$

The first order condition of the problem implies that the labor demand facing the household is

$$H_{j,t} = \left( \frac{W_t}{\tau_t^h W_{j,t}} \right)^{\theta_w} L_t, \quad (4.24)$$

where  $\theta_w = \lambda_w / (\lambda_w - 1)$ .

Since the utility function is separable and households cannot re-optimize their wage rate in each period due to the Calvo-pricing setting, households set their optimal idiosyncratic wage  $\tilde{W}_t$  by solving the following problem:

$$\max_{\tilde{W}_t} \mathbb{E}_t \sum_{l=0}^{\infty} (\beta \xi_w)^l (-\Psi(H_{j,t+l}))$$

subject to the labor demand schedule

$$H_{j,t+l} = \left( \frac{W_{t+l}}{\tau_{t+l}^h \tilde{W}_t X_{tl}} \right)^{\theta_w} L_{t+l}$$

and the budget constraint. The first order condition is

$$\mathbb{E}_t \left\{ \sum_{l=0}^{\infty} (\beta \xi_w)^l \left[ \frac{-\partial \Psi}{\partial H_{j,t+l}} \frac{\partial H_{j,t+l}}{\partial \tilde{W}_t} + v_t \left( X_{t,l} H_{j,t+l} + X_{t,l} \tilde{W}_t \frac{\partial H_{j,t+l}}{\partial \tilde{W}_t} \right) \right] \right\} = 0,$$

where

$$\frac{\partial \Psi}{\partial H_{j,t+l}} = \psi_h (H_{j,t+l})^\varphi; \quad \frac{\partial H_{j,t+l}}{\partial \tilde{W}_t} = \frac{-\theta_w H_{j,t+l}}{\tilde{W}_t}.$$

It implies

$$\mathbb{E}_t \left\{ \sum_{l=0}^{\infty} (\beta \xi_w)^l \left[ \frac{\theta_w \psi_h (H_{j,t+l})^{\varphi+1}}{W_t} + v_{t+l} (1 - \theta_w) (X_{t,l} H_{j,t+l}) \right] \right\} = 0.$$

Multiplying  $\tilde{W}_t / (1 - \theta_w)$ , we have

$$\mathbb{E}_t \left\{ \sum_{l=0}^{\infty} (\beta \xi_w)^l \left[ v_{t+l} \tilde{W}_t X_{t,l} H_{j,t+l} - \lambda_w \psi_h (H_{j,t+l})^{\varphi+1} \right] \right\} = 0.$$

Noting that  $\tilde{W}_t = \tilde{w}_t \cdot w_t \cdot P_t$ ,  $\psi_t = P_t v_t$  and  $X_{t,l} / P_{t+l} = (P_{t-1} \pi_{t+l})^{-1}$ , we can rewrite this first order condition as

$$\mathbb{E}_t \left\{ \sum_{l=0}^{\infty} (\beta \xi_w)^l \left[ \psi_{t+l} \tilde{w}_t w_t \pi_t \frac{H_{j,t+l}}{\pi_{t+l}} - \lambda_w \psi_h (H_{j,t+l})^{\varphi+1} \right] \right\} = 0.$$

Therefore, this first order condition can be conveniently expressed as

$$\tilde{w}_t = \frac{\mathcal{N}_t^w}{\mathcal{D}_t^w}, \tag{4.25}$$

where

$$\begin{aligned}\mathcal{D}_t^w &= w_t \pi_t \mathbb{E}_t \left\{ \sum_{l=0}^{\infty} (\beta \xi_w)^l \frac{\psi_{t+l} H_{j,t+l}}{\pi_{t+l}} \right\} = w_t \psi_t H_{j,t} + (\beta \xi_w) \mathbb{E}_t \left\{ \frac{w_t \pi_t}{w_{t+1} \pi_{t+1}} \mathcal{D}_t^w \right\}, \\ \mathcal{N}_t^w &= \lambda_w \psi_h \mathbb{E}_t \left\{ \sum_{l=0}^{\infty} (\beta \xi_w)^l (H_{j,t+l})^{\varphi+1} \right\} = \lambda_w \psi_h (H_{j,t})^{\varphi+1} + (\beta \xi_w) \mathbb{E}_t \{ \mathcal{N}_{t+1}^w \}.\end{aligned}$$

By imposing the labor demand schedules, we get

$$\mathcal{D}_t^w = w_t \psi_t \left( \tau_t^h \tilde{w}_t \right)^{-\theta_w} L_t + \beta \xi_w \mathbb{E}_t \left\{ \frac{w_t \pi_t}{w_{t+1} \pi_{t+1}} \mathcal{D}_t^w \right\}, \quad (4.26)$$

$$\mathcal{N}_t^w = \lambda_w \psi_h \left( \tau_t^h \tilde{w}_t \right)^{-\theta_w(\varphi+1)} L^{\varphi+1} + \beta \xi_w \mathbb{E}_t \{ \mathcal{N}_{t+1}^w \}. \quad (4.27)$$

#### 4.2.4 Aggregation Issues

Strictly speaking, there is no aggregate production function in the economy. We need to associate the aggregate output with the utilization of capital and labor, which by construction are only used by intermediate good firms.

**Aggregation of  $Y_{j,t}$**  Defining unweighted aggregate output  $Y_t^*$  as

$$Y_t^* \equiv \int_0^1 Y_{j,t} dj, \quad (4.28)$$

and imposing intermediate good demand functions (equation 4.3), we have

$$Y_t = \left( \frac{P_t^*}{P_t} \right)^{\theta_f} Y_t^* = (p_t^*)^{\theta_f} Y_t^*, \quad (4.29)$$

where

$$P_t^* \equiv \tau_t^y \left[ \int_0^1 (P_{j,t})^{-\theta_f} dj \right]^{-\frac{1}{\theta_f}}. \quad (4.30)$$

$P_t^*$  is a natural indicator for price dispersion. Moreover, since the capital-labor ratio is common across firms, equation (4.28) implies

$$Y_t^* = z_t K_t^\alpha L_t^{1-\alpha} - \Phi. \quad (4.31)$$

**Aggregation of  $H_{j,t}$ .** Defining

$$H_t \equiv \int_0^1 H_{j,t} dj,$$

and imposing equation (4.24), we have the following relationship

$$L_t = \left(\frac{W_t^*}{\bar{W}_t}\right)^{\theta_w} H_t = (w_t^*)^{\theta_w} H_t, \quad (4.32)$$

where

$$W_t^* \equiv \tau_t^h \left[ \int_0^1 (W_{j,t})^{-\theta_w} dj \right]^{-\frac{1}{\theta_w}}. \quad (4.33)$$

This is an indicator of dispersion among idiosyncratic wage rates.

**Aggregate Output.** Combining equation (4.31), (4.32) and (4.29), we have the following aggregate output,

$$Y_t = (p_t^*)^{\theta_f} (w_t^*)^{(1-\alpha)\theta_w} z_t K_t^\alpha H_t^{1-\alpha} - (p_t^*)^{\theta_f} \Phi. \quad (4.34)$$

This expression shows that in theory the aggregate output of the economy is not only a function of the technology shocks, aggregate capital, aggregate labor, but is also related to the dispersions of intermediate good prices and idiosyncratic wage rates. However, for the linearized version of the model, the dispersions of prices and wages play trivial roles. See Appendix B for further discussion.

**The Aggregate Resource Constraint.** Since this is a closed economy, the resource constraint of the economy is straightforward. Since the total output of the economy can only be distributed to three outlets- consumption, investment, and the real cost of capital utilization-



the resource constraint is

$$C_t + I_t + a(u_t)\bar{K}_t = Y_t \quad (4.35)$$

### 4.2.5 Government Policy

We assume that the government controls nominal interest rates at no resource cost. Specifically, nominal interest rate  $R_{t+1}$  is governed by the following Taylor rule

$$\frac{R_{t+1}}{R} = \left(\frac{R_t}{R}\right)^{\rho_R} \left(\frac{\pi_{t+1}}{\pi}\right)^{\gamma_\pi} \left(\frac{Y_{t+1}}{Y}\right)^{\gamma_Y} \exp(\sigma_R \varepsilon_{t+1}^R), \quad \varepsilon_{t+1}^R \stackrel{iid}{\sim} N(0, 1). \quad (4.36)$$

$\pi$  is the target (gross) inflation rate set by the central bank.  $R$  and  $Y$  are respectively the steady-state nominal interest rate and aggregate output. There are various specifications regarding the Taylor rule, and our specification is in line with that in Del Negro et al. (2007) and Justiniano and Primiceri (2008). We assume that the government always balances its budget by lump sum taxes and subsidies and consumes nothing. Thus, there is no need to explicitly specify the fiscal policy.

## 4.3 Empirical Methodology

### 4.3.1 The Data

We utilize 7 aggregate series of the United States to estimate the model. In particular, they are (per capita) real GDP, (per capita) real consumption, (per capita) real investment, working hours, inflation rate, and Federal funds rate. Since the sample period runs from 1980Q1 to 2004Q4, there are 100 observations in total. Due to the fact that real GDP, real consumption, real investment, and real wage rate are growing over time, we use the Hodrick-Prescott (HP) filter to remove the trend of these series. Moreover, because the inflation rate and Federal fund rate series are stationary, we calculate the corresponding mean values during the sample periods and demean them.

### 4.3.2 Bayesian Inference

Following Del Negro et al. (2007) and Smets and Wouters (2007), we employ the Bayesian method to estimate structural parameters of the model. To implement the Bayesian method, we need to find the likelihood function and impose the prior distributions to obtain the posterior distribution.

We log-linearize the equilibrium conditions around the steady state implied by the model.<sup>4</sup> We obtain a linear system of 20 equations represented by equations (B.7) to (B.26). Correspondingly, there are 20 variables, of which 13 are endogenous, and 7 are exogenous. The endogenous variables are

$$\hat{K}_t, \hat{L}_t, \hat{R}_t, \hat{w}_t, \hat{r}_t^K, \hat{s}_t, \hat{\pi}_t, \hat{u}_t, \hat{I}_t, \hat{C}_t, \hat{P}_{k',t}, \hat{\psi}_t, \hat{Y}_t,$$

and the exogenous variables are

$$\hat{z}_t, \hat{\tau}_t^K, \hat{\tau}_t^L, \hat{\tau}_t^Y, \hat{\tau}_t^B, \hat{\tau}_t^y, \hat{\tau}_t^h.$$

All variables are expressed in log-deviation from the steady-state values, and we use a hat ( $\hat{\cdot}$ ) on the top of variables to denote those changes. Moreover, we adopt the method proposed by Klein (2000) to solve the model. After the linearized policy functions are obtained, we express the system in the state-space form. Thus, we can simply use the Kalman filter to obtain the likelihood function. Following past literature, we assume that the prior distributions are independent. We use the Markov Chain Monte Carlo (MCMC) method to find the posterior distribution. In particular, we use the random-walk Metropolis (RWM) algorithm to obtain the MCMC draws. An and Schorfheide (2007) provide detailed implementation of the algorithm.

We implement two empirical studies to examine the sources and magnitude of model misspecification. The first one is our benchmark case, and the second one is the restricted case. For the benchmark case, we take the stand that the traditional models are indeed misspecified if

---

<sup>4</sup>The procedure to find the steady state are discussed in Appendix B, and the log-linearized equilibrium conditions is listed in Appendix C.

one does not impose the wedges. That is, since the presence of wedges is the defining features of the data, we estimate a model allowing the presence of wedges. In contrast, for the restricted case, we assume that the traditional models are correctly specified. In other words, we impose a strong prior on the processes of wedges to reflect this belief. Thus, the restricted case can also be viewed as a sensitivity test of the benchmark case.

### 4.3.3 Prior Settings

Since the setting of prior distributions plays an important role for conducting the Bayesian estimation method, we discuss how we set these prior distributions. In the benchmark case, we estimate 31 of the 32 structural parameters. Capital depreciation rate  $\delta$  is the only parameter that is not estimated. It is calibrated to  $\delta = 0.025$ , and is consistent with most empirical studies.

#### Utility Related Parameters

The prior distribution of the subjective discount factor  $\beta$  is assumed to have a normal distribution with mean 0.9926 and standard deviation 0.005. Contrary to the existing literature, we estimate it to avoid any systematic bias that arises with assigning the wrong calibrated value.

Since the risk aversion coefficient  $\gamma$  is restricted to be a positive real number, we assume that it follows a Gamma distribution with mean 1 and standard deviation 0.5.

For the same reason, the prior distribution of the risk averse coefficient of labor disutility  $\varphi$  is assumed to have a Gamma distribution. Since the inverse of  $\varphi$  is the Frisch elasticity of labor supply, which is usually estimated to be big in magnitude in most macroeconomic studies, we assume the prior mean of  $\varphi$  as 0.6. The prior standard deviation of  $\varphi$  is set to be 0.3, which reflects that we have little knowledge on the true value of  $\varphi$ . The prior mean and standard deviation of the habit formation parameter  $b$  are respectively set to be 0.75 and 0.15, which is exactly the same as that used in Christiano et al. (2010). This reflects our belief that the degree of habit formation should be high.

## Technology Related Parameters

Since the valid range of capital's share,  $\alpha$ , lies between 0 and 1, its prior distribution is assumed to have a Beta distribution with mean 0.3 and standard deviation 0.05. That is, the prior mean mainly reflects the common estimate for capital's share. This prior distribution is identical to that of Smets and Wouters (2007). It is worth noting that Christiano et al. (2010) do not estimate this parameter and calibrate the capital share at a lower level ( $\alpha = 0.25$ ). The prior distribution of fixed cost  $\Phi$  for producing intermediate goods is assumed to have a Gamma distribution with mean 1.40 and standard deviation 0.5. Once again, this prior distribution rivals that of Smets and Wouters (2007), in which they set a prior mean of 1.25 and a prior standard deviation of 0.15.

In the New Keynesian models, the price markup  $\lambda_f$  and wage markup  $\lambda_w$  play important roles since they control the monopoly powers of intermediate good firms and households. The prior distributions of these two parameters are both assumed to have a Gamma distribution. In particular, the prior mean and standard deviation of  $\lambda_f$  are set to be 1.20 and 0.15. This reflects the fact that the observed price markup is usually not too high. Similarly, the prior distribution of  $\lambda_w$  is set as same as that of  $\lambda_f$ . Note that in the estimation exercises of Christiano et al. (2010) and Smets and Wouters (2007),  $\lambda_w$  is not estimated at all, and is calibrated to be 1.05 and 1.5.

The prior distribution of the intermediate good Calvo-pricing parameter  $\xi_p$  is assumed to have a Beta distribution with mean 0.5 and standard deviation 1.5, which is also identical to that of Christiano et al. (2010). As for  $\xi_w$ , the Calvo-pricing parameter of idiosyncratic wage setting, it is assumed to have the same prior as  $\xi_p$ . Once again,  $\xi_w$  is not estimated in Christiano et al. (2010), and is set to 0.75.

As for the curvature parameter  $\varkappa_s$  of investment adjustment cost, its prior distribution is assumed to have a Gamma distribution with mean 5 and standard deviation 5. We adopt a uninformative prior to reflect the fact that economists have less knowledge about this parame-

ter.<sup>5</sup> Following Christiano et al. (2010), the prior distribution of the curvature parameter  $\varkappa_a$  of the capacity adjustment cost is set to have a Gamma distribution with mean 1 and standard deviation 0.75.

### Parameters of the Taylor Rule

To reflect the fact that the Federal Funds rate is persistent, the smooth parameter  $\rho_R$  in the Taylor rule is set to follow a Beta distribution with mean 0.8 and standard deviation 0.15. As for the response parameter  $\gamma_\pi$  with respect to inflation in the Taylor rule, a normal distribution with mean 1.5 and standard deviation 0.3 is adopted as its prior distribution. This prior is less informative compared to that of Christiano et al. (2010). As for the response parameter  $\gamma_Y$  with respect to the output gap, it is also assumed to have a normal distribution with mean 0.5 and standard deviation 0.3. Finally, the prior distribution of the target (quarterly) gross inflation rate  $\pi$  is set to be normally distributed with mean 1.006 and standard deviation 0.01.<sup>6</sup>

### Other Parameters

Since most studies find the technology shock to be persistent, the prior distribution of the persistent parameter  $\rho_z$  of the technology shock is assumed to have a normal distribution with mean 0.8 and standard deviation 0.15. The prior distribution of the standard deviation parameter,  $\sigma_z$ , in the technology process is set to be an inverse Gamma random variable with mean 0.1 and standard deviation 2.0. This prior distribution is the same as that used in Smets and Wouters (2007).

The parameters for the wedge processes play a prominent role for examining the model misspecification issue, for which we have little knowledge. Therefore, we assign relatively loose priors on them. Specifically, for the persistence parameters in these wedge processes, the prior means are all set to be 0.5 and standard deviations are set to be 0.2. Moreover, they are assumed

---

<sup>5</sup>This prior distribution is slightly different from that in Christiano et al. (2010), where the parameters are set to 12 and 8, respectively.

<sup>6</sup>It implies that the annualized net inflation rate is 2.4 percent.

to behave as Beta random variables. As for the standard deviation parameters, such as  $\sigma_K$ ,  $\sigma_L$ , and etc. , they are assumed to be inverse Gamma distributed with mean 0.1 and standard deviation 2.0.

Similar to Smets and Wouters (2007), our wedge processes act like reduced form shocks. The priors are set in the same way to reflect the lack of knowledge on them. Table 4.1 summarizes the above discussion.

### **Priors for the Restricted Case**

As explained before, we would like to impose a strong belief that the traditional models are essentially correctly specified. Therefore, we impose all of the persistence coefficients of the wedges to be 0. More importantly, since we impose stronger priors such that all prior means of the wedge standard deviations (e.g.,  $\sigma_K$ ,  $\sigma_L$ ,...) are small. In particular, we set prior means of  $\sigma_K$ ,  $\sigma_L$ ,  $\sigma_Y$ ,  $\sigma_B$ ,  $\sigma_y$ , and  $\sigma_h$  to be 0.01, which are one-tenth those in the benchmark case. Table 4.2 summarizes the prior distributions of the parameters in the restricted case.

Table 4.1: Prior and Posterior for the Benchmark Case

Parameter	Source	Prior			Posterior	
		Type	Mean	Std Dev	Mean	Std. Dev
$\beta$	CTW	Normal	0.9926	0.0050	0.9926	0.0000
$\gamma$		Gamma	1.0000	0.5000	4.5763	0.7794
$\varphi$		Gamma	0.6000	0.3000	1.2153	0.0718
$\alpha$	SW	Normal	0.3000	0.0500	0.2994	0.0026
$\Phi$	SW	Gamma	1.4000	0.5000	0.4653	0.1074
$\lambda_f$	CTW	Gamma	1.2000	0.1500	1.5843	0.1397
$\lambda_w$		Gamma	1.2000	0.1500	1.0076	0.0035
$b$	CTW	Beta	0.7500	0.1500	0.1916	0.0405
$\xi_p$	CTW	Beta	0.5000	0.1500	0.6179	0.0565
$\xi_w$		Beta	0.5000	0.1500	0.4088	0.0755
$\varkappa_s$		Gamma	5.0000	5.0000	5.5180	1.0052
$\varkappa_a$	CTW	Gamma	1.0000	0.7500	2.2974	0.2565
$\gamma_\pi$	CTW	Normal	1.5000	0.3000	1.4035	0.1006
$\gamma_Y$	CTW	Normal	0.5000	0.3000	0.0612	0.0302
$\pi$		Normal	1.0060	0.0100	1.0060	0.0001
$\rho_R$	CTW	Beta	0.8000	0.1500	0.6955	0.0510
$\rho_z$		Beta	0.8000	0.1500	0.9308	0.0156
$\rho_K$	SW	Beta	0.5000	0.2000	0.8957	0.0599
$\rho_L$	SW	Beta	0.5000	0.2000	0.8852	0.0368
$\rho_Y$	SW	Beta	0.5000	0.2000	0.0963	0.0751
$\rho_B$	SW	Beta	0.5000	0.2000	0.6945	0.0650
$\rho_y$	SW	Beta	0.5000	0.2000	0.5036	0.0642
$\rho_h$	SW	Beta	0.5000	0.2000	0.7051	0.0615
$\sigma_R$	SW	InvGamma	0.1000	2.0000	0.0040	0.0003
$\sigma_z$	SW	InvGamma	0.1000	2.0000	0.0047	0.0004
$\sigma_K$	SW	InvGamma	0.1000	2.0000	0.0416	0.0063
$\sigma_L$	SW	InvGamma	0.1000	2.0000	0.0096	0.0018
$\sigma_Y$	SW	InvGamma	0.1000	2.0000	0.0294	0.0036
$\sigma_B$	SW	InvGamma	0.1000	2.0000	0.0232	0.0085
$\sigma_y$	SW	InvGamma	0.1000	2.0000	0.0041	0.0005
$\sigma_h$	SW	InvGamma	0.1000	2.0000	0.0069	0.0005

Table 4.2: Priors and Posteriors: Restricted Case

Parameter	Source	Prior			Posterior	
		Type	Mean	Std Dev	Mean	Std Dev
$\beta$	CTW	Normal	0.9926	0.0050	0.9926	0.0000
$\gamma$		Gamma	1.0000	0.5000	1.8847	0.1789
$\varphi$		Gamma	0.6000	0.3000	0.0920	0.0481
$\alpha$	SW	Normal	0.3000	0.0500	0.2995	0.0026
$\Phi$	SW	Gamma	1.4000	0.5000	0.2272	0.0660
$\lambda_f$	CTW	Gamma	1.2000	0.1500	1.6163	0.1009
$\lambda_w$		Gamma	1.2000	0.1500	1.0145	0.0044
$b$	CTW	Beta	0.7500	0.1500	0.0428	0.0173
$\xi_p$	CTW	Beta	0.5000	0.1500	0.7006	0.0204
$\xi_w$		Beta	0.5000	0.1500	0.5276	0.0395
$\varkappa_s$		Gamma	5.0000	5.0000	1.5445	0.1703
$\varkappa_a$	CTW	Gamma	1.0000	0.7500	0.7791	0.1327
$\gamma_\pi$	CTW	Normal	1.5000	0.3000	1.8515	0.0761
$\gamma_Y$	CTW	Normal	0.5000	0.3000	0.0917	0.0340
$\pi$		Normal	1.0060	0.0100	1.0060	0.0001
$\rho_R$	CTW	Beta	0.8000	0.1500	0.8319	0.0526
$\rho_z$		Beta	0.8000	0.1500	0.9695	0.0063
$\sigma_R$		InvGamma	0.1000	2.0000	0.0064	0.0006
$\sigma_z$		InvGamma	0.1000	2.0000	0.0055	0.0004
$\sigma_K$		InvGamma	0.0100	0.0010	0.0346	0.0028
$\sigma_L$		InvGamma	0.0100	0.0010	0.0415	0.0025
$\sigma_Y$		InvGamma	0.0100	0.0010	0.0228	0.0017
$\sigma_B$		InvGamma	0.0100	0.0010	0.0101	0.0011
$\sigma_y$		InvGamma	0.0100	0.0010	0.0083	0.0006
$\sigma_h$		InvGamma	0.0100	0.0010	0.0080	0.0007



## 4.4 Empirical Results

In this section, we discuss our empirical results. We focus only on some relevant parameters and mainly on the benchmark case. For comparison purposes, we digress to the restricted case whenever necessary. Our empirical result is based on 32,000 MCMC draws.

### 4.4.1 Parameter Estimates

First, the point estimate of  $\gamma$  (risk aversion coefficient of consumption) in terms of the posterior mean is around 4.58, and the corresponding standard deviation of posterior density is 0.78. The result of posterior mean is significantly different from most calibrated settings, which is usually assumed to be 1.

However, if we look at the restricted case, the point estimate of  $\gamma$  reduces to around 1.88, which is somewhat closer to the estimate from Levin et al. (2006).<sup>7</sup> The significant difference between the benchmark and restricted cases indicates that model misspecification tremendously affects the point estimates of the risk aversion coefficient in the consumption utility. This is important because different values for the risk aversion coefficient imply different consumption patterns over time. In either benchmark or restricted case, our estimates are different from that of Smets and Wouters (2007), where their point estimate is around 1.39.

As for the risk aversion coefficient ( $\varphi$ ) of labor disutility, we obtain sharply different point estimates for the benchmark and the restricted cases. In the benchmark case, the point estimate of  $\varphi$  is 1.215, but in the restricted case, it is merely 0.092. Since the difference between the two cases is induced only by the prior settings for various wedges, it once again demonstrates the importance of model misspecification on estimation results. Since the inverse of  $\varphi$  is the Frisch elasticity of labor supply, which is important for understanding labor market phenomena, further investigation on this parameter is necessary and warranted. We believe that our benchmark estimate on this parameter is reliable because it is more or less in line with the estimates from the past literature.

---

<sup>7</sup>Levin et al. (2006) estimate  $\gamma = 2.19$ .

The point estimates of  $\lambda_f$  are respectively 1.58 and 1.61 in the benchmark and restricted case, which are larger than that of Christiano et al. (2010). Nonetheless, it is important to note that Christiano et al. (2010) use the Bayesian version of matching impulse response functions estimation method, which is essentially a limited information estimation method. On the other hand, our results are based on the pure Bayesian method, which is a full information method.

Our estimation results for  $\lambda_w$  between benchmark and restricted cases are similar to each other: they are 1.007 and 1.0145. This indicates that the monopoly power of households in determining the wage rate is small.

The presence of habit formation in consumption is the hallmark of modern New Keynesian models. In our benchmark case, the habit formation parameter ( $b$ ) is merely 0.19, which is much smaller than that in Christiano et al. (2005), where the point estimate of  $b$  is 0.65 in their benchmark case. However, our estimate is close to that of Levin et al. (2006). Note that the point estimate of  $b$  in the restricted case is extremely small (0.042), indicating that habit formation hardly exists in consumption. We think this particular result is dubious since almost all empirical studies in the field find evidence of habit formation, even though the level of habit formation varies.<sup>8</sup>

As for the Calvo-pricing parameter ( $\xi_p$ ) of intermediate goods, its point estimate is 0.70, which is broadly in line with other studies. For instance, Guerron-Quintana's (2011) result is 0.68. Our point estimate implies that firms adjust their prices, on average, every 3.1 quarters. In addition, the point estimate of  $\xi_w$  (the Calvo-pricing parameter with respect to the households wage decision) is 0.56, which is also within the range of existing studies. In the restricted case, the magnitudes of these two parameters are similar to the benchmark case.

As for the persistence parameter  $\rho_z$ , our estimation results convey an interesting story, which is broadly consistent with the finding of Del Negro and Schorfheide (2009). In their misspecified model, the estimate of  $\rho_z$  is very closed to 1, but in their DSGE-VAR case, where the model misspecification is controlled, the  $\rho_z$  becomes much smaller. We have a similar finding:

---

<sup>8</sup>For example, in Guerron-Quintana (2011), the point estimate of  $b$  is as high as 0.93.

the point estimate is 0.93 when model misspecification is controlled for and is 0.9695 when model misspecification is not accounted for. The minor difference in the level of  $\rho_z$  has major implications on the convergence rate of the economy.

As for the parameters of the Taylor rule, we find that the point estimate of  $\gamma_\pi$ , 1.403, is much bigger than  $\gamma_Y$ , 0.061, indicating that the central bank only moderately responds to the output gap but actively responds to inflation. This finding is consistent with existing studies, e.g., Del Negro and Schorfheide (2009), Guerron-Quintana (2011), and Smets and Wouters (2007).

The meanings of various wedge processes will become clear when we discuss the forecast error variance decomposition.

#### 4.4.2 Impulse Response Functions

In this section, we briefly discuss the impulse response functions. Note that we focus only on two structural shocks because we do not impose structural interpretations on the wedges.

Fig. 4.1 illustrates the impulse response functions when the economy faces a contractionary monetary policy shock. At face value, these impulse response functions are consistent with the general intuition. In particular, after a contractionary monetary shock, real GDP and consumption initially decrease and then gradually return to their steady state levels. More importantly, we find strong evidence of a *price puzzle*. That is, different from the common intuition, the inflation rate increases after a contractionary monetary shock. In the empirical study of Christiano et al. (2010), they also find that the inflation rate goes up after a contractionary monetary shock, but only temporarily. In contrast to their study, our empirical results indicate that the adverse increase of inflation is persistent. In fact, the inflation rate takes more than two years to return to its steady state value.

Our results suggest that there is no sluggish response in real GDP, which is different from Christiano et al. (2005). We find that real GDP drops down instantly when the federal funds rate goes up. This result is surprising since our model is very similar to Christiano et al. (2005). There are two possible explanations for this finding. First, the difference might stem from the

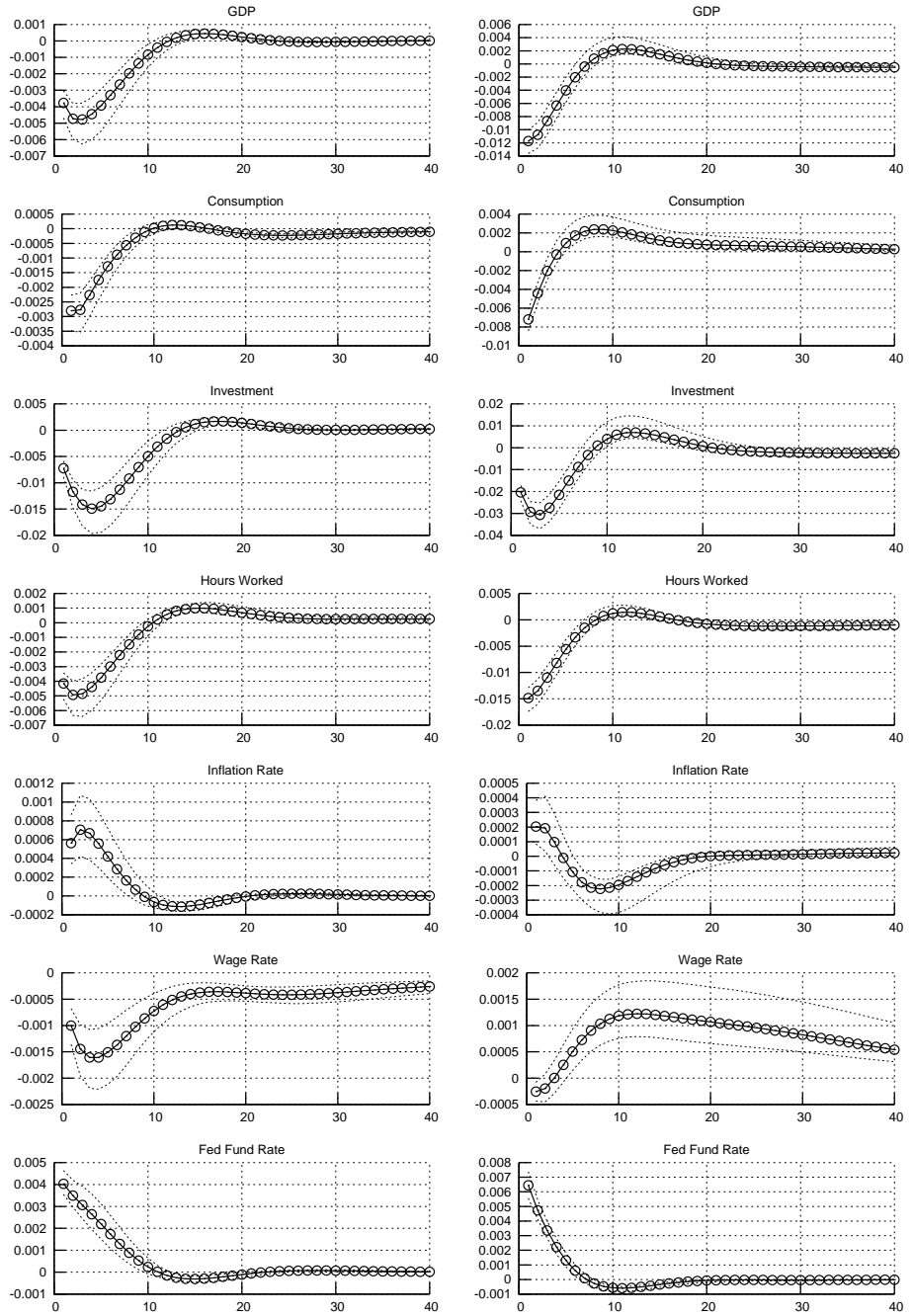


Figure 4.1: Impulse Response Functions (Monetary Policy Shock)

fact that we employ a different estimation method. Christiano et al. (2005) use the matching impulse response functions method while we adopt the Bayesian method. The second possible explanation for the difference is that our benchmark case effectively takes into account the model misspecification. Since our point estimate of habit formation parameter  $b$  is low, we do not observe the hump-shaped impulse response function for consumption. This finding is different from the results of Christiano et al. (2005) and Smets and Wouters (2007). Given that their results are based on the assumption of a correctly specified model, our results call for further investigation of the habit formation parameter within the context that the model might be potentially misspecified.

Fig. 4.2 displays the impulse response functions when the economy faces a technology shock. The left panels are for the benchmark case and the right panels are for the restricted case. In general, they have similar patterns and the impulse response functions are consistent with the general intuition. Facing a positive technology shock, real GDP, consumption, investment, hours worked, and the wage rate first increase and then gradually fall back to their steady-state levels. Furthermore, the inflation rate and Federal funds rate initially decrease and then gradually return to their steady-state levels. The main difference between the benchmark case and the restricted case is on how long it would take the economy to return to the pre-shocked levels.

With respect to real GDP, it takes about 5 years to return to the steady-state level in the benchmark case, but almost 10 years in the restricted case. We find similar durations for consumption and investment.

The most interesting part of these impulse response functions rests on the dynamic path of the wage rate. In the restricted case, the wage rate takes an extremely long time to return to its steady-state level. Of course, such a long duration can be explained by the fact that the technology shock is highly persistent.

One of main purposes of building monetary DSGE models is to predict the outcome of the economy facing different shocks, both qualitatively and quantitatively. Our results highlight the

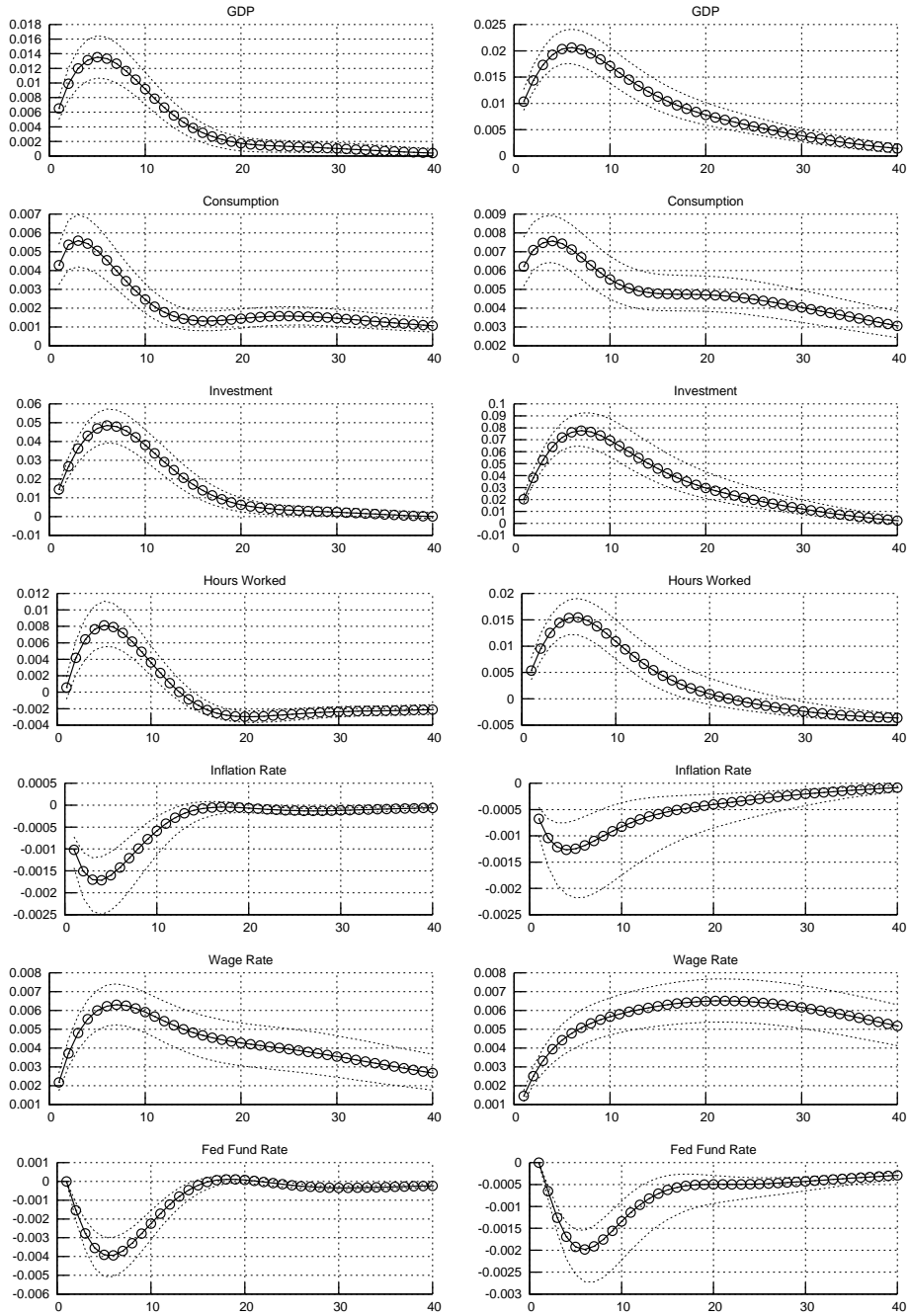


Figure 4.2: Impulse Response Functions (Technology Shock)

need for more exploration on model misspecification problems.

### 4.4.3 Forecast Error Variance Decompositions (FEVD)

In this section, we discuss the forecast error variances decomposition. The discussion focuses on the 100 period ahead FEVD. This is because the relative effects of shocks or wedges become stable after such a long period.

Table 4.3 shows the forecast variance decomposition with respect to real GDP. First, monetary and technology shocks only account for less than 30 percent of the total variation. In other words, the sum of wedges contributes to about 70 percent of the variation. This result clearly indicates that the model misspecification problem is severe. This is rather surprising given that we are using a state-of-the-art DSGE model. It also shows that the cross-equation restrictions implied in the traditional model are obviously too restrictive. Moreover, the homogeneous labor wedge (the  $L$  wedge) plays the most important role in explaining the total variation. At the same time, the idiosyncratic labor wedge (the  $h$  wedge) also plays a non-trivial role. In sum, these two labor market wedges account for about 47.48 percent of the total variation. Therefore, if one would like to use the DSGE model to predict future output, it will be improper to ignore the potential misspecification of labor markets.

The FEVD of consumption is reported in Table 4.4. First, the two structural shocks account for less than 25 percent of the variation, and the effects of monetary policy shock are trivial. Second, although the two labor wedges (i.e.,  $L$  wedge and  $h$  wedge ) are still important, their contribution to consumption variation is slightly smaller, when compared to the case of real GDP. At the same time, the  $K$  and  $Y$  wedges account for 17.87 and 19.43 percent of the variation. It is also clear to see that the importance of the  $L$  wedge decreases as the forecast periods increase.

Table 4.5 reports the FEVD with respect to investment. Again, the two structural shocks account for about 30 percent of the variation. Different from the FEVD of consumption, the  $K$  wedge and the  $Y$  wedge account for much less variation. Instead, the two labor market wedges

Table 4.3: Variance Decomposition: Real GDP

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	5.442	29.586	1.190	47.987	7.257	2.123	5.745	0.666
h=10	2.933	30.174	1.479	44.427	8.662	4.961	3.070	4.290
h=40	2.343	28.393	5.985	38.278	7.816	4.619	2.609	9.954
h=100	2.298	27.995	7.426	37.725	7.689	4.520	2.581	9.763

Table 4.4: Variance Decomposition: Real Consumption

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	5.603	23.790	12.055	36.616	13.298	4.743	3.505	0.387
h=10	3.299	22.610	14.335	31.945	17.179	5.970	2.121	2.538
h=40	2.517	22.009	16.755	26.853	19.516	5.050	1.783	5.514
h=100	2.364	22.077	17.865	25.984	19.432	4.829	1.684	5.761

(*L* wedge and *h* wedge) dominate again. Table 4.6 shows the FEVD with respect to hours worked. The two structural shocks account for less than 20 percent of total variation. In all, our results show that the labor market is severely misspecified.

## 4.5 Simulation Exercises

This section provides simulation evidence on how the proposed methodology helps to detect the the source of the model misspecification. Specifically, a variant of the New Keynesian model, which is akin to Clarida et al. (1999) and Woodford (2003), is exploited as the data generating process. Since the structure of the model is typical, we only briefly address the key features of the model. Step-by-step derivations for this type of model can be found in Galí (2008) and

Table 4.5: Variance Decomposition: Real Investment

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	4.846	30.234	0.785	51.474	5.495	1.077	5.307	0.779
h=10	2.592	30.693	1.357	46.668	6.975	4.461	2.822	4.427
h=40	1.934	28.409	6.540	39.257	6.139	4.296	2.225	11.196
h=100	1.896	27.976	7.921	38.686	6.125	4.205	2.199	10.989



Table 4.6: Variance Decomposition: Hours Worked

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	7.527	10.166	2.327	63.752	2.520	2.934	9.683	1.088
h=10	4.271	13.117	2.874	58.649	2.526	5.923	5.953	6.684
h=40	3.223	13.520	8.553	46.326	7.637	5.435	4.367	10.936
h=100	3.015	14.423	9.373	44.255	8.753	5.211	4.083	10.882

Table 4.7: Variance Decomposition: (Quarterly) Inflation Rate

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	2.607	14.624	0.237	29.506	4.935	7.385	35.542	5.160
h=10	1.996	17.649	0.636	32.625	6.915	11.218	23.956	5.001
h=40	1.961	17.362	1.360	31.405	6.863	10.963	23.002	7.081
h=100	1.946	17.300	1.862	31.237	6.836	10.876	22.894	7.047

Table 4.8: Variance Decomposition: Real Wage Rate

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	2.479	21.259	3.245	39.840	11.703	4.366	9.339	7.766
h=10	1.354	24.097	5.866	39.250	15.695	6.132	2.894	4.708
h=40	0.736	26.607	10.688	31.558	20.218	4.534	1.322	4.335
h=100	0.651	25.984	14.674	28.998	19.612	4.052	1.162	4.863

Table 4.9: Variance Decomposition: (Quarterly) Federal Fund Rate

	<i>R</i> shock	<i>z</i> shock	<i>K</i> wedge	<i>L</i> wedge	<i>Y</i> wedge	<i>B</i> wedge	<i>y</i> wedge	<i>h</i> wedge
h=4	23.565	11.774	0.565	25.913	4.900	11.331	15.082	6.866
h=10	10.137	16.339	1.481	33.096	7.950	18.751	8.081	4.163
h=40	9.297	16.004	2.474	31.632	8.264	18.814	7.648	5.863
h=100	9.244	16.051	2.636	31.529	8.331	18.709	7.607	5.891

Walsh (2003) among others.

### 4.5.1 The Model

Following the convention of New Keynesian DSGE models, it is assumed that a representative final good firm transforms a continuum of intermediate goods  $Y_{j,t}$  into a homogeneous final good  $Y_t$  by the following technology:

$$Y_t = \left( \int_0^1 Y_{j,t}^{\frac{1}{\lambda_f}} dj \right)^{\lambda_f}. \quad (4.37)$$

Taking as given the prices of intermediate goods ( $P_{j,t}$ ) and the final good ( $P_t$ ) as given, the final good firm solves the profit maximization problem:

$$\max_{Y_{j,t}} P_t Y_t - \int_0^1 P_{j,t} Y_{j,t} dj,$$

subject to Eq. (4.37). The first order condition of the above problem implies that the demand functions facing intermediate good firms are

$$Y_{jt} = \left( \frac{P_{j,t}}{P_t} \right)^{-\frac{\lambda_f}{\lambda_f-1}} Y_t.$$

The intermediate goods firms are monopolistic providers of differentiated goods, and hence, are allowed to set their own product prices. However, since they are not allowed to do so in every period, the intermediate goods firms also take into account the probability of not being able to re-optimize their prices. Specifically, the model follows Calvo (1983) in assuming that the probability of resetting the the optimal prices is  $1 - \xi$  and is independent across time and firms. As a consequence, when an intermediate goods firm is able to re-optimize its price, the corresponding pricing problem can be expressed as follows:

$$\max_{\tilde{P}_t} E_t \sum_{l=0}^{\infty} (\beta\xi)^l v_{t+l} \left\{ \tilde{P}_t \tilde{Y}_{t+l} - P_{t+l} s_{t+l} \right\} \quad (4.38)$$

subject to

$$\tilde{Y}_{t+l} = \left( \frac{\tilde{P}_t}{P_t} \right)^{-\frac{\lambda_f}{\lambda_f-1}} Y_{t+l}.$$

The constraint represents intermediate good demand in period  $t+l$  with the price of that good set in period  $t$ .  $s_{t+l}$  indicates the real marginal cost in period  $t+l$ , which is common across firms and independent of the level of production.<sup>9</sup> Besides, because households hold equities of intermediate good firms,  $v_{t+l}$  is the marginal value of per unit of money for households at time  $t+l$ . Lastly, the presence of  $\xi^l$  in the objective function indicates the probability of not being able to re-optimize the price after  $l$  periods. If the  $j^{\text{th}}$  firm cannot reset its price in period  $t$ , the effective price for its product is the price of the previous period, i.e.,  $P_{j,t-1}$ . The first order condition for this profit maximization is

$$E_t \sum_{l=0}^{\infty} (\beta\xi)^l v_{t+l} \tilde{Y}_{t+l} \left\{ \tilde{P}_t - \lambda_f P_{t+l} s_{t+l} \right\} = 0.$$

It is straightforward to see that when firms are absent from the price rigidity,  $\xi = 0$ , the optimal price in each period is a markup of nominal marginal cost. In the current model, the markup is simply  $\lambda_f$ .

We assume that the intermediate good firms use labor as the sole input of production and that the production function of intermediate goods is linear with a stochastic technology shock  $z_t$ .<sup>10</sup>

$$Y_{j,t} = z_t L_{j,t}. \tag{4.39}$$

---

<sup>9</sup>The reason for common real marginal cost will be clear when we discuss the cost minimization problem of intermediate good firms.

<sup>10</sup>Notice that while the assumption of only one input for production is unrealistic, it is sufficient for demonstrating our main points. Actually, this assumption is widely used in the literature of New Keynesian models. See Clarida et al. (1999) and An and Schorfheide (2007), among others.

In terms of the deviation form, the technology shock  $z_t$  follows an AR(1) process

$$\hat{z}_{t+1} = \rho_z \hat{z}_t + \sigma_z \varepsilon_{t+1}^z, \quad \varepsilon_{t+1}^z \sim N(0, 1). \quad (4.40)$$

Taking the nominal interest  $R_t$  and nominal wage rate  $W_t$  as given, the  $j^{\text{th}}$  intermediate good firm solves the following cost minimization problem:

$$\max_{L_{j,t}} R_t W_t L_{j,t}$$

subject to its production function (4.39). The first order necessary condition implies that the real marginal cost function of the firm is

$$s_t = \frac{R_t W_t}{z_t},$$

which is a function independent of the level of production,  $Y_{j,t}$ . This independence property stems from the assumption of a linear production function. In other words, the production function is constant returns to scale. If the production function is rather decreasing return to scale, the real marginal cost will depend on the level of  $Y_{j,t}$  in general, see Galí and Gertler (1999) for instance.

Notice that there exists a working capital channel in the model. This channel reflects the fact that many firms need to borrow money in advance to pay the variable input costs for production. Following Christiano et al. (2005), we assume that the price of borrowing for a wage bill is the nominal interest rate, and thus,  $R_t$  enters the objective function of the firm's cost minimization problem. While this feature is usually absent from the simple New Keynesian models, some studies have shown that it improves the quantitative prediction of New Keynesian models. The reason for introducing the working capital channel as a feature of the data generating process is to examine whether or not model misspecification can be detected if a model without working capital channel is adopted.

In this model, a representative household receives utility from consuming final goods and disutility from providing its labor services to intermediate goods firms. Following most real business cycle studies, the household utility function is assumed to be both time separable and separable between consumption and labor. Thus, the representative household maximizes the expected discounted lifetime utility function:

$$\max_{\{C_t, L_t\}_{t=0}^{\infty}} E_t \sum_{t=0}^{\infty} \beta^t \left\{ \frac{C_t^{1-\gamma}}{1-\gamma} - \frac{L_t^{1+\varphi}}{1+\varphi} \right\},$$

subject to the household budget constraint:

$$P_t C_t + B_{t+1} = R_{t-1} B_t + W_t L_t + D_t + T_t.$$

The household has several sources of income in time  $t$ . First, the household collects return from holding government bonds with gross nominal interest rate  $R_t$ . Second, by providing labor services the household obtains labor income at the overall nominal wage rate  $W_t$ . Third, the household receives dividends from holding equities of intermediate goods firms. Finally, there exists a lump-sum transfer (or taxation) from the government. The household uses its income to buy  $C_t$  units of the final good under the competitive price  $P_t$  and to hold government bond  $B_{t+1}$ , which facilitates consumption smoothing across time. Notice that the capital accumulation decision is excluded from the current model. The accompanying first order conditions of the household problem are

$$\begin{aligned} v_t P_t &= C_t^{-\gamma}, \\ v_t &= \beta E_t \{v_{t+1} R_t\}, \\ L_t^\varphi &= v_t W_t, \end{aligned}$$

where  $v_t$  is the Lagrange multiplier associated with the household budget constraint at time  $t$

and represents the marginal value of (money) income.<sup>11</sup>

The government plays two roles in this simple model. First, it acts as the sole provider of bond assets. Second, it consumes certain amounts of final goods for non-productive purposes. The government budget constraint can be expressed as

$$G_t + R_{t-1}B_t = B_{t+1} + T_t,$$

where  $T_t$  is the lump sum taxation (subsidies) to the household in time  $t$ . We assume that the taxation policy is passive. That is, when the total expense of government is higher than its total income, the government levies a lump sum tax; otherwise, it provides a lump sum subsidy.

In terms of log deviation from steady state values, the government determines the nominal interest rate by the Taylor rule:

$$\hat{R}_{t+1} = \rho_1 \hat{R}_t + \rho_2 \hat{R}_{t-1} + (1 - \rho_1 - \rho_2) \left\{ \gamma_\pi \hat{\pi}_{t+1} + \gamma_Y \hat{Y}_{t+1} \right\} + \sigma_R \varepsilon_{t+1}^R, \quad \varepsilon_{t+1}^R \sim N(0, 1). \quad (4.41)$$

The two lag terms of the nominal interest rate in this government policy rule capture the fact that the short term nominal rate is quiet persistent.<sup>12</sup> The terms in the curly brackets capture the systematic part of the interest rate policy with respect to output and the inflation rate. Moreover, the random variable  $\varepsilon_t^R$  is designed to capture the non-systematic part of the nominal rate, and one can treat the nominal interest rate as a potential shock facing agents as well.

We assume that the government spending is stochastic, and governed by

$$G_t = \zeta_t Y_t,$$

---

<sup>11</sup>Recall that we have incorporated  $v_t$ 's into the profit maximization problem.

<sup>12</sup>Most authors working on NK models usually assume only one lagged term in the interest rate policy rule. See Smets and Wouters (2007) for example. However, some studies include more lags to describe the path of the nominal rate. For example, Clarida et al. (2000) incorporate more than one lag to make the nominal rate dynamics flexible.

We re-parametrize  $\zeta_t$  as

$$g_t \equiv \frac{1}{1 - \zeta_t},$$

and assume that in terms of log deviation,  $g_t$  follows an AR(1) process

$$\hat{g}_{t+1} = \rho_g g_t + \sigma_g \varepsilon_{t+1}^g, \quad \varepsilon_{t+1}^g \sim N(0, 1). \quad (4.42)$$

Since there is no capital accumulation in this model and final output is only consumed by households and the government, the resource constraint of the economy can conveniently be expressed as

$$\hat{Y}_t = \hat{C}_t + \hat{g}_t.$$

After combining the first order conditions of households' and firms' problems, one can simplify the the system to the following two equations

$$\hat{Y}_t = E_t \left\{ \hat{Y}_{t+1} \right\} - \frac{1}{\gamma} \left( \hat{R}_t - E_t \hat{\pi}_{t+1} \right) + (1 - \rho_g) \hat{g}_t \quad (4.43)$$

$$\hat{\pi}_t = \kappa \left( \hat{R}_t + (\gamma + \varphi) \hat{Y}_t - \gamma \hat{g}_t - (\varphi + 1) \hat{z}_t \right) + \beta E_t \hat{\pi}_{t+1}, \quad (4.44)$$

where  $\kappa = \frac{(1-\xi)(1-\beta\xi)}{\xi}$ . Eq. (4.43) is the dynamic IS (DIS) curve, and Eq. (4.44) is the New Keynesian Phillips Curve (NKPC).

The DIS curve links the current output level to the expected future output and real interest rate. The presence of expected future output comes from the consumption smoothing nature of households. The effect of the real interest rate ( $\hat{R}_t - E_t \hat{\pi}_{t+1}$ ) on current output depends crucially on the elasticity of intertemporal substitution (EIS, and it is  $\frac{1}{\gamma}$ ). Therefore, in order to precisely quantify the extent of monetary policy in affecting the real economy, a precise estimation of  $\gamma$  is critical.

As for the NKPC, the coefficient  $\kappa$  indicates the extent of inflation rate responding to the change of the real marginal cost. When the price is completely flexible (i.e the Calvo pricing coefficient approaches zero),  $\kappa$  approaches one, and the inflation rate will fully reflect the change

of real marginal cost. Besides, the real marginal cost also depends on the output level, the government spending shock, and the technology shock because they affect the real wage rate of the economy under general equilibrium. The presence of expected future inflation rate comes from the forward-looking consideration of firms due to Calvo pricing.

In sum, the model proposed in this subsection consists of 5 linearized equations, containing Eq. (4.40), (4.41), (4.42), (4.43) and (4.44), and there are 5 variables (i.e.,  $\hat{Y}_t$ ,  $\hat{\pi}_t$ ,  $\hat{R}_t$ ,  $\hat{g}_t$ ,  $\hat{z}_t$ ) in the system, correspondingly.

### 4.5.2 Misspecified Models

The logic behind our simulation exercise is simple. One can imagine that an econometrician would like to estimate the deep parameters of the model in the previous subsection. However, the econometrician does not perfectly understand the true model, and hence, the models for estimation are somehow misspecified. Specifically, three misspecified models are considered, and let us call them  $\mathcal{M}(1)$ ,  $\mathcal{M}(2)$ , and  $\mathcal{M}(3)$ .<sup>13</sup>

$\mathcal{M}(1)$       The working capital channel is missing. That is, the nominal interest rate is assumed to not affect the real marginal cost of an intermediate good firm.

$\mathcal{M}(2)$       The monetary policy rule is misspecified. There is only one lag included to characterize the smooth behavior of the nominal interest rate. In other words,  $\rho_2$  is absent from the estimation.

$\mathcal{M}(3)$       The government spending shock is excluded.

While the econometrician makes mistakes in specifying the correct models for estimation, he recognizes this problem and tries to incorporate some time-varying wedges into the model. Specifically, he introduces 3 wedges into the model to detect the potential sources of model misspecification. First, he introduces a labor demand wedge  $\tau_t^L$  into the cost minimization

---

<sup>13</sup>The linearized equilibrium conditions are list in Appendix B.



problem. Thus, the cost minimization problem of intermediate good firms becomes

$$\min_{L_{j,t}} \tau_t^L R_t W_t L_{j,t},$$

subject to (4.39). Essentially, this wedge is introduced to capture the potential misspecification with respect to the labor demand of intermediate good firms. Second, a final good wedge  $\tau_t^Y$  and a bond demand wedge  $\tau_t^B$  are incorporated into the household budget constraint, and thus, the constraint is modified to

$$\tau_t^Y P_t C_t + B_{t+1} = \tau_t^B R_{t-1} B_t + W_t L_t + D_t + T_t.$$

We assume that the time-varying wedges follow the AR(1) process. That is, for  $x \in \{L, Y, B\}$

$$\hat{\tau}_{t+1}^x = \rho_x \hat{\tau}_t^x + \sigma_x \varepsilon_{t+1}^x, \quad \varepsilon_{t+1}^x \sim N(0, 1). \quad (4.45)$$

### 4.5.3 The Empirical Strategy and Results

The empirical strategy for our simulation exercise are as follows. First, we use the true model to generate a sequence of observables, consisting of output, inflation, and the nominal interest rate. The sample size is 100, reflecting the limited sample size in most business cycle studies. The generated sample is kept as the true data for the three misspecified models. The parameter values for the data generating process are summarized in Table 4.10.

We use the Bayesian estimation method to estimate these three misspecified models. To do so, we need to specify the priors of the parameters, including those of the wedge processes. The prior distributions of parameters are summarized in Table 4.10. We assume that the persistence coefficients of the wedge processes (i.e.,  $\rho_L$ ,  $\rho_Y$  and  $\rho_B$ ) follow the Gamma distribution with mean value 0.5 and standard deviation 0.2. These wide prior distributions prevent us from obtaining unrealistic estimates driven by the restricted prior distributions. Under the same

Table 4.10: DGP and prior distributions

Parameter	DGP	Prior Distribution <sup>a</sup>		
		Type <sup>b</sup>	Para(1)	Para(2)
$\beta$	0.99	N	0.992	0.005
$\gamma$	1.50	G	1.300	0.300
$\varphi$	1.00	G	0.800	0.300
$\xi$	0.67	B	0.700	0.050
$\gamma_\pi - 1$	0.80	G	1.000	0.050
$\gamma_Y$	0.50	G	0.400	0.050
$\rho_1$	0.40	B	0.450	0.100
$\rho_2$	0.30	B	0.400	0.100
$\rho_z$	0.70	B	0.750	0.100
$\rho_g$	0.70	B	0.750	0.100
$\sigma_R$	0.10	IG	0.300	2.000
$\sigma_z$	0.10	IG	0.300	2.000
$\sigma_g$	0.10	IG	0.300	2.000
$\rho_L$	0.00	B	0.500	0.200
$\rho_Y$	0.00	B	0.500	0.200
$\rho_B$	0.00	B	0.500	0.200
$\sigma_L$	0.00	IG	0.100	2.000
$\sigma_Y$	0.00	IG	0.100	2.000
$\sigma_B$	0.00	IG	0.100	2.000

<sup>a</sup> Para(1) and Para(2) indicate the mean and standard deviation of the prior distribution, respectively.

<sup>b</sup> In the third column, 'N', 'G', 'B', 'IG' are abbreviation of Normal, Gamma, Beta, and Inverse Gamma distributions, respectively.

consideration, the prior distributions of  $\sigma_L$ ,  $\sigma_Y$  and  $\sigma_B$  are all set to be quite disperse. <sup>14</sup>

Finally, the estimation procedure is the same as that in Section 4.3.

The forecast error variance decomposition of the misspecified models are listed in Table 4.11. For the  $\mathcal{M}(1)$  case, none of three wedges can detect the model misspecification due to the missing working capital channel. This is because the FEVD indicates that the contribution of the three wedges to variations of observables are essentially ignorable in the long run. This result implies that if the econometrician uses the misspecified model to estimate the parameters, he might conclude that the model he used is very close to the true model. This is not good news. However, it reveals how hard it is to detect the model misspecification in DSGE models.

<sup>14</sup>Note also that prior distributions of structural parameters are close to the true values; and thus, we do not have to worry about the consequences of improper priors.

Table 4.11: Forecast Error Variance Decomposition of Misspecified Model

Observable	$\mathcal{M}(1)$					
	$R$	$z$	$g$	$\tau_L$	$\tau_Y$	$\tau_B$
Nominal Rate	93.885	5.072	0.024	0.085	0.002	0.931
Output	76.448	5.819	10.131	0.092	1.392	6.117
inflation rate	63.811	30.405	1.505	1.324	0.046	2.910

Observable	$\mathcal{M}(2)$					
	$R$	$z$	$g$	$\tau_L$	$\tau_Y$	$\tau_B$
Nominal Rate	87.743	8.020	0.932	0.174	0.161	2.970
Output	65.389	8.524	10.288	0.163	3.087	12.549
Inflation rate	44.249	37.438	0.373	1.884	0.501	15.556

Observable	$\mathcal{M}(3)$					
	$R$	$z$	$g$	$\tau_L$	$\tau_Y$	$\tau_B$
Nominal Rate	93.926	5.539	0	0.102	0.264	0.170
Output	78.979	6.765	0	0.115	11.374	2.767
Inflation rate	63.836	32.869	0	1.657	0.215	1.423

As for the  $\mathcal{M}(2)$  case, it is clear to see that the bond wedge ( $\tau^B$ ) explain significant portions of the variation in output and the inflation rate. This result gives the econometrician very useful information that there is something wrong, because wedges are economically meaningless. If the econometrician looks at his model carefully, he will find that the bond wedge affects the price of bonds. Therefore, our methodology helps the econometrician think whether or not the model provide a good description of the relative price of bonds, which is the nominal interest rate.

As for the  $\mathcal{M}(3)$  case, the final good wedge ( $\tau^Y$ ) contributes more than 10 percent of the total variation of output. Again, this result should not exist if the model is correctly specified. This FEVD table gives the econometrician clear information that the specification for consumption or output is problematic.

To sum up, our methodology can detect two of three model misspecifications. It is worth noticing that our simulation exercises are tougher than implementing this method to real data. The reason is because in our DGP there are only three structural shocks. However, when one

implements this method to the real economy, in which various shocks exist, and our methodology will be very informative as shown in Section 4.3.

## 4.6 Conclusion

In this chapter, we provide a framework to examine the magnitude and possible sources for model misspecification. Our framework treats model misspecification as distortionary taxes. Our empirical results show that the state-of-the-art DSGE models are still highly misspecified, especially in the labor market. We also find that the severity of the price puzzle and parameter estimates depend heavily on whether or not we allow for model misspecification. This suggests that incorrectly believing the model is correctly specified may lead to misleading results.

In addition, the labor market misspecification contributes to more than 40 percent of the total variation in almost all of the economic aggregates. Thus, our empirical results call for further investigation on the structure of labor markets.

## REFERENCES

- J. Adda and R. W. Cooper. *Dynamic Economics: Quantitative Methods and Applications*, volume 1 of *MIT Press Books*. The MIT Press, 2003.
- S. R. Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–84, August 1994.
- Y. Algan, O. Allais, and W. J. Den Haan. Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control*, 32(3): 875–908, March 2008.
- D. Altig, L. Christiano, M. Eichenbaum, and J. Linde. Firm-specific capital, nominal rigidities and the business cycle. *Review of Economic Dynamics*, 14(2):225–247, April 2011.
- S. Altug. Time-to-build and aggregate fluctuations: Some new evidence. *International Economic Review*, 30(4):889–920, November 1989.
- S. An and F. Schorfheide. Bayesian analysis of dsge models. *Econometric Reviews*, 26(2-4): 113–172, 2007.
- S. An, Y. Chang, and S.-B. Kim. Can a representative-agent model represent a heterogeneous-agent economy. *American Economic Journal: Macroeconomics*, 1(2):29–54, July 2009.
- S. B. Aruoba, J. Fernandez-Villaverde, and J. F. Rubio-Ramirez. Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control*, 30(12): 2477–2508, December 2006.
- O. P. Attanasio and M. Browning. Consumption over the life cycle and over the business cycle. *American Economic Review*, 85(5):1118–37, December 1995.
- F. Barillas and J. Fernandez-Villaverde. A generalization of the endogenous grid method. *Journal of Economic Dynamics and Control*, 31(8):2698–2712, August 2007.
- M. Browning, L. P. Hansen, and J. J. Heckman. Micro data and general equilibrium models. In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 8, pages 543–633. Elsevier, October 1999.
- M. Cagetti. Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics*, 21(3):339–53, July 2003.
- G. A. Calvo. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398, September 1983.
- F. Canova. *Methods for Applied Macroeconomic Research*. Princeton University Press, 2007.
- C. D. Carroll. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320, June 2006.

- Y. Chang and S.-B. Kim. Heterogeneity and aggregation: Implications for labor-market fluctuations. *American Economic Review*, 97(5):1939–1956, December 2007.
- V. V. Chari, P. J. Kehoe, and E. R. McGrattan. Business cycle accounting. *Econometrica*, 75(3):781–836, 05 2007.
- V. Chernozhukov and H. Hong. An mcmc approach to classical estimation. *Journal of Econometrics*, 115(2):293–346, August 2003.
- I. Chowdhury, M. Hoffmann, and A. Schabert. Inflation dynamics and the cost channel of monetary transmission. *European Economic Review*, 50(4):995–1016, May 2006.
- L. J. Christiano. Solving dynamic equilibrium models by a method of undetermined coefficients. *Computational Economics*, 20(1-2):21–55, October 2002.
- L. J. Christiano and M. Eichenbaum. Current real-business-cycle theories and aggregate labor-market fluctuations. *American Economic Review*, 82(3):430–50, June 1992.
- L. J. Christiano, M. Eichenbaum, and C. Evans. The effects of monetary policy shocks: Evidence from the flow of funds. *The Review of Economics and Statistics*, 78(1):16–34, February 1996.
- L. J. Christiano, M. Eichenbaum, and C. L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, February 2005.
- L. J. Christiano, M. Trabandt, and K. Walentin. Dsge models for monetary policy analysis. In B. M. Friedman and M. Woodford, editors, *Handbook of Monetary Economics*, volume 3 of *Handbook of Monetary Economics*, chapter 7, pages 285–367. Elsevier, October 2010.
- R. Clarida, J. Gali, and M. Gertler. The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37(4):1661–1707, December 1999.
- R. Clarida, J. Galí, and M. Gertler. Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics*, 115(1):147–180, February 2000.
- J. H. Cochrane. A simple test of consumption insurance. *Journal of Political Economy*, 99(5):957–76, October 1991.
- L. Cubeddu and J.-V. Rios-Rull. Families as shocks. *Journal of the European Economic Association*, 1(2-3):671–682, 04/05 2003.
- D. N. DeJong and C. Dave. *Structural Macroeconometrics*. Princeton University Press, 2007.
- M. Del Negro and F. Schorfheide. Priors from general equilibrium models for vars. *International Economic Review*, 45(2):643–673, 05 2004.
- M. Del Negro and F. Schorfheide. Monetary policy analysis with potentially misspecified models. *American Economic Review*, 99(4):1415–50, September 2009.

- M. Del Negro, F. Schorfheide, F. Smets, and R. Wouters. On the fit of new keynesian models. *Journal of Business & Economic Statistics*, 25:123–143, April 2007.
- W. J. den Haan and P. Rendahl. Solving the incomplete markets model with aggregate uncertainty using explicit aggregation. *Journal of Economic Dynamics and Control*, 34(1):69–78, 2010. ISSN 0165-1889. Computational Suite of Models with Heterogeneous Agents: Incomplete Markets and Aggregate Uncertainty.
- W. J. den Haan, K. L. Judd, and M. Juillard. Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty. *Journal of Economic Dynamics and Control*, 34(1):1–3, January 2010.
- C. J. Erceg, D. W. Henderson, and A. T. Levin. Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics*, 46(2):281–313, October 2000.
- J. Fernandez-Villaverde and J. F. Rubio-Ramirez. Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies*, 74(4):1059–1087, October 2007.
- J. Galí. *Monetary Policy and the Open Economy*. Introductory Chapters. Princeton University Press, 2008.
- J. Gali and M. Gertler. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222, October 1999.
- C. Gourieroux, A. Monfort, and E. Renault. Indirect inference. *Journal of Applied Econometrics*, 8(S):S85–118, Suppl. De 1993.
- P.-O. Gourinchas and J. A. Parker. Consumption over the life cycle. *Econometrica*, 70(1):47–89, January 2002.
- P. A. Guerron-Quintana. The implications of inflation in an estimated new keynesian model. *Journal of Economic Dynamics and Control*, 35(6):947–962, June 2011.
- L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–54, July 1982.
- L. P. Hansen and K. J. Singleton. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, 50(5):1269–86, September 1982.
- L. P. Hansen, J. Heaton, and A. Yaron. Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics*, 14(3):262–80, July 1996.
- J. Heathcote, K. Storesletten, and G. L. Violante. Quantitative macroeconomics with heterogeneous households. *Annual Review of Economics*, 1(1):319–354, 05 2009.
- J. Heathcote, F. Perri, and G. L. Violante. Unequal we stand: An empirical analysis of economic inequality in the united states: 1967-2006. *Review of Economic Dynamics*, 13(1):15–51, January 2010.

- B. Heer and A. Maussner. *Dynamic General Equilibrium Modeling: Computational Methods and Applications*. Springer, 2 edition, 2009.
- T. Hintermaier and W. Koeniger. Incomplete markets and the evolution of the us consumer wealth distribution. Technical report, EUROPEAN UNIVERSITY INSTITUTE, 2008.
- M. Huggett. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969, 1993.
- B. F. Ingram, N. R. Kocherlakota, and N. E. Savin. Explaining business cycles: A multiple-shock approach. *Journal of Monetary Economics*, 34(3):415–428, December 1994.
- P. N. Ireland. Sticky-price models of the business cycle: Specification and stability. *Journal of Monetary Economics*, 47(1):3–18, February 2001.
- A. Justiniano and G. E. Primiceri. The time-varying volatility of macroeconomic fluctuations. *American Economic Review*, 98(3):604–41, June 2008.
- S. H. Kim, R. Kollmann, and J. Kim. Solving the incomplete market model with aggregate uncertainty using a perturbation method. *Journal of Economic Dynamics and Control*, 34(1):50–58, January 2010.
- M. S. Kimball. The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking*, 27(4):1241–77, November 1995.
- R. G. King, C. I. Plosser, and S. T. Rebelo. Production, growth and business cycles : I. the basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195–232, 1988.
- P. Klein. Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control*, 24(10):1405–1423, September 2000.
- K. Kopecky and R. Suen. Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics*, 13(3):701–714, July 2010.
- P. Krusell and A. Smith. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, 1998.
- P. Krusell and A. Smith. *Quantitative macroeconomic models with heterogeneous agents*. Ninth World Congress, Cambridge University Press, New York, NY, 2006.
- P. Krusell, T. Mukoyama, A. Sahin, and J. Anthony A. Smith. Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 12(3):393–402, July 2009.
- J. Laitner and D. Silverman. Estimating life-cycle parameters from consumption behavior at retirement. NBER Working Papers 11163, National Bureau of Economic Research, Inc, Mar. 2005.



- A. T. Levin, A. Onatski, J. Williams, and N. M. Williams. Monetary policy under uncertainty in micro-founded macroeconomic models. In *NBER Macroeconomics Annual 2005, Volume 20*, NBER Chapters, pages 229–312. National Bureau of Economic Research, Inc, December 2006.
- R. E. Lucas. *Models of Business Cycles*. Basil Blackwell, New York, 1987.
- R. E. Lucas. Macroeconomic priorities. *American Economic Review*, 93(1):1–14, March 2003.
- L. Maliar, S. Maliar, and F. Valli. Solving the incomplete markets model with aggregate uncertainty using the krusell-smith algorithm. *Journal of Economic Dynamics and Control*, 34(1):42–49, January 2010.
- E. R. McGrattan, R. Rogerson, and R. Wright. An equilibrium model of the business cycle with household production and fiscal policy. *International Economic Review*, 38(2):267–90, May 1997.
- M. G. Palumbo. Uncertain medical expenses and precautionary saving near the end of the life cycle. *Review of Economic Studies*, 66(2):395–421, April 1999.
- C. I. Plosser. Understanding real business cycles. *Journal of Economic Perspectives*, 3(3):51–77, Summer 1989.
- B. Preston and M. Roca. Incomplete markets, heterogeneity and macroeconomic dynamics. (13260), July 2007.
- V. Quadrini. Entrepreneurship, saving and social mobility. *Review of Economic Dynamics*, 3(1):1–40, January 2000.
- F. Ravenna and C. E. Walsh. Optimal monetary policy with the cost channel. *Journal of Monetary Economics*, 53(2):199–216, March 2006.
- Rouwenhorst. Asset returns and business cycles. In T. Cooley, editor, *Frontiers of Business Cycle Research*, pages 294–330. 1995.
- S. Schmitt-Grohe and M. Uribe. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control*, 28(4):755–775, January 2004.
- S. Schmitt-Grohe and M. Uribe. Optimal fiscal and monetary policy in a medium-scale macroeconomic model. In *NBER Macroeconomics Annual 2005, Volume 20*, NBER Chapters, pages 383–462. National Bureau of Economic Research, Inc, 2006.
- L. L. Schumaker. On shpa-preserving quadratic spline interpolation. *SIAM Journal of Numerical Analysis*, 20:854–864, 1983.
- R. Shimer. Convergence in macroeconomics: The labor wedge. *American Economic Journal: Macroeconomics*, 1(1):280–97, January 2009.

- C. A. Sims. Macroeconomics and methodology. *Journal of Economic Perspectives*, 10(1):105–20, Winter 1996.
- C. A. Sims. Solving linear rational expectations models. *Computational Economics*, 20(1-2): 1–20, October 2002.
- D. T. Slesnick. Aggregate consumption and saving in the postwar united states. *The Review of Economics and Statistics*, 74(4):585–97, November 1992.
- F. Smets and R. Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123–1175, 09 2003.
- F. Smets and R. Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606, June 2007.
- A. A. Smith. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8(S):S63–84, Suppl. De 1993.
- K. Storesletten, C. Telmer, and A. Yaron. Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548, October 2007.
- G. Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181, 1986.
- G. Tauchen and R. Hussey. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, 59(2):371–96, March 1991.
- C. E. Walsh. *Monetary Theory and Policy, 2nd Edition*, volume 1 of *MIT Press Books*. The MIT Press, 2003.
- M. Woodford. *Interests and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, 2003.
- M. Woodford. Convergence in macroeconomics: Elements of the new synthesis. *American Economic Journal: Macroeconomics*, 1(1):267–79, January 2009.
- E. R. Young. Approximate aggregation. *Computing in Economics and Finance 2005* 141, Society for Computational Economics, Nov. 2005.
- E. R. Young. Solving the incomplete markets model with aggregate uncertainty using the krusell-smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, 34(1):36 – 41, 2010. ISSN 0165-1889. Computational Suite of Models with Heterogeneous Agents: Incomplete Markets and Aggregate Uncertainty.

## APPENDICES

## Appendix A

# Recursive Competitive Equilibrium (RCE)

A proper recursive representation of the model should include the whole cross-sectional distribution of individual capital holding into the state space. The original individuals problem can be rewritten as the following recursive form:

$$\begin{aligned} V(k, e; \Gamma, z) &= \max_{c, k' \geq 0} \left\{ \frac{c^{1-\gamma} - 1}{1-\gamma} + V(k', e'; \Gamma', z') \right\} \\ \text{st} \quad c + k' &= (1 + r - \delta)k + we \\ \Gamma' &= \mathbf{F}(\Gamma, z) \end{aligned}$$

where  $V(k, e; \Gamma, z)$  is the value function and  $\mathbf{F}$  is a operator which governs the transition of probability measure  $\Gamma(k, e)$ .

The recursive competitive equilibrium consists of a value function,  $V(k, e; \Gamma, z)$ , a consumption policy function  $g^c(k, e; \Gamma, z)$ , a capital policy function,  $g^{k'}(k, e; \Gamma, z)$ , aggregate capital  $K(\Gamma, z)$ , aggregate labor  $L(\Gamma, z)$ , interest rate  $r(\Gamma, z)$ , wage rate  $w(\Gamma, z)$  and the law of motion  $\mathbf{F}$  for the probability measure, such that the following conditions holds:

1. Individuals optimize: Given  $r(\Gamma, z)$  and  $w(\Gamma, z)$ , policy functions,  $g^{k'}(k, e; \Gamma, z)$  and  $g^c(k, e; \Gamma, z)$ ,

solve the value function  $V(k, e; \Gamma, z)$  in the above Bellman equation for all  $(\Gamma, z)$ .

2. The representative firm optimizes profits:

$$\begin{aligned} r(\Gamma, z) &= z [K(\Gamma, z)]^{\alpha-1} [L(\Gamma, z)]^{1-\alpha}, \\ w(\Gamma, z) &= z [K(\Gamma, z)]^{\alpha} [L(\Gamma, z)]^{-\alpha}, \end{aligned}$$

for all  $(\Gamma, z)$ .

3. The good market clears:

$$\int \left\{ g^c(k, e; \Gamma, z) + g^{k'}(k, e; \Gamma, z) \right\} d\Gamma = F(z, K(\Gamma, z), L(\Gamma, z)) + (1 - \delta)K(\Gamma, z)$$

for all  $(\Gamma, z)$ .

4. Factor Markets clear:

$$K(\Gamma, z) = \int k d\Gamma, \text{ and } L(\Gamma, z) = \int e d\Gamma,$$

for all  $(\Gamma, z)$ .

5. Consistency Condition: the dynamics of the measure  $\Gamma_t$  is governed by

$$\Gamma = H(\Gamma, z) \dots$$

This transition function is consistent with the individual policy function  $g^k(k_t, e_t; \Gamma_t, z_t)$ .

# Appendix B

## Equilibrium Conditions

### B.1 Aggregation

Here, we demonstrate how to link the prices of intermediate goods with that of final good, and hence economy wide inflation. At the same time, we show the relation between the wage rate and inflation rate. We also define two dispersion measures for the intermediate good prices and idiosyncratic labor wages.

#### Price level

Substituting equation (4.3) into (4.1), we have the the relationship between the final good price and intermediate good prices:

$$P_t = \tau_t^y \left[ \int_0^1 (P_{j,t})^{\frac{1}{1-\lambda_f}} dj \right]^{1-\lambda_f}.$$

Hence, we can rewrite the ab

$$\begin{aligned} \frac{P_t}{\tau_t^y} &= \left[ \int_0^1 (P_{j,t})^{\frac{1}{1-\lambda_f}} dj \right]^{1-\lambda_f} \\ &= \left[ (1 - \xi_p) \left( \tilde{P}_t \right)^{\frac{1}{1-\lambda_f}} + \int_{S^c} (P_{j,t})^{\frac{1}{1-\lambda_f}} dj \right]^{1-\lambda_f} \end{aligned}$$

where  $S^c$  denotes a index set for firms who are not able to re-optimize their price. Note also that

$$\int_{S^c} (P_{j,t})^{\frac{1}{1-\lambda_f}} dj = \int f_{t-1,t}(\omega) [P(\omega)\pi_{t-1}]^{\frac{1}{1-\lambda_f}} d\omega,$$

where  $f_{t-1,t}(\omega)$  denotes the ‘number’ (density) of firms that had price  $P(\omega)$  in time  $t-1$  and were not able to re-optimize in time  $t$ . The presence of  $\pi_{t-1}$  is because of full indexation. Calvo-pricing implies

$$f_{t-1,t}(\omega) = \xi_p f_{t-1}(\omega),$$

where  $f_{t-1}(\omega)$  is the total *number* of firms with price  $P(\omega)$  in time  $t-1$ . Hence,

$$\begin{aligned} \int_{S^c} (P_{j,t})^{\frac{1}{1-\lambda_f}} dj &= \int f_{t-1,t}(\omega) [P(\omega)\pi_{t-1}]^{\frac{1}{1-\lambda_f}} d\omega \\ &= \xi_p (\pi_{t-1})^{\frac{1}{1-\lambda_f}} \int f_{t-1}(\omega) [P(\omega)]^{\frac{1}{1-\lambda_f}} d\omega \\ &= \xi_p (\pi_{t-1})^{\frac{1}{1-\lambda_f}} \int (P_{j,t-1})^{\frac{1}{1-\lambda_f}} dj \\ &= \xi_p (\pi_{t-1})^{\frac{1}{1-\lambda_f}} \left( \frac{P_{t-1}}{\tau_{t-1}^y} \right)^{\frac{1}{1-\lambda_f}}. \end{aligned}$$

Hence, we have the following expression

$$P_t = \tau_t^y \left[ (1 - \xi_p) (\tilde{P}_t)^{\frac{1}{1-\lambda_f}} + \xi_p \left( \pi_{t-1} \frac{P_{t-1}}{\tau_{t-1}^y} \right)^{\frac{1}{1-\lambda_f}} \right]^{1-\lambda_f}.$$

Dividing both sides by  $P_t$  and recalling  $\tilde{p}_t = \tilde{P}_t/P_t$ , we have

$$1 = \tau_t^y \left[ (1 - \xi_p) (\tilde{p}_t)^{\frac{1}{1-\lambda_f}} + \xi_p \left( \frac{\pi_{t-1}}{\pi_t} \frac{1}{\tau_{t-1}^y} \right)^{\frac{1}{1-\lambda_f}} \right]^{1-\lambda_f}. \quad (\text{B.1})$$

## Wage Rate

Substituting equation (4.24) into equation (4.22), we have the relationship

$$W_t = \tau_t^h \left[ \int_0^1 (W_{j,t})^{\frac{1}{1-\lambda_w}} dj \right]^{1-\lambda_w}.$$

Under similar derivation as the case of  $P_t$ , the above equation implies

$$W_t = \tau_t^h \left[ (1 - \xi_w) (\tilde{W}_t)^{\frac{1}{1-\lambda_w}} + \xi_w \left( \pi_{t-1} \frac{W_{t-1}}{\tau_{t-1}^h} \right)^{\frac{1}{1-\lambda_w}} \right]^{1-\lambda_w}.$$

Dividing both side by  $W_t$  and recalling  $\tilde{w}_t = \tilde{W}_t/W_t$  and  $w_t = W_t/P_t$ , it further becomes

$$1 = \tau_t^h \left[ (1 - \xi_w) (\tilde{w}_t)^{\frac{1}{1-\lambda_w}} + \xi_w \left( \frac{\pi_{t-1} w_{t-1}}{\pi_t} \frac{1}{w_t} \frac{1}{\tau_{t-1}^h} \right)^{\frac{1}{1-\lambda_w}} \right]^{1-\lambda_w}. \quad (\text{B.2})$$

## Prices Dispersion

Equation (4.30) implies

$$P_t^* = \tau_t^y \left[ (1 - \xi_p) (\tilde{P}_t)^{-\theta_f} + \xi_p \left( \pi_{t-1} \frac{P_{t-1}^*}{\tau_{t-1}^y} \right)^{-\theta_f} \right]^{-\frac{1}{\theta_f}}.$$

Since  $p_t^* = P_t^*/P_t$ , it further becomes

$$p_t^* = \tau_t^y \left[ (1 - \xi_p) (\tilde{p}_t)^{-\theta_f} + \xi_p \left( \frac{\pi_{t-1} p_{t-1}^*}{\pi_t} \frac{1}{\tau_{t-1}^y} \right)^{-\theta_f} \right]^{-\frac{1}{\theta_f}}. \quad (\text{B.3})$$

## Wage Dispersion

Equation (4.33) implies

$$W_t^* = \tau_t^h \left[ (1 - \xi_w) (\tilde{W}_t)^{-\theta_w} + \xi_w \left( \pi_{t-1} \frac{W_{t-1}^*}{\tau_{t-1}^h} \right)^{-\theta_w} \right]^{-\frac{1}{\theta_w}}$$



It further becomes

$$w_t^* = \tau_t^h \left[ (1 - \xi_w) (\tilde{w}_t)^{-\theta_w} + \xi_w \left( \frac{\pi_{t-1}}{\pi_t} \frac{w_{t-1}}{w_t} \frac{w_{t-1}^*}{\tau_{t-1}^h} \right)^{-\theta_w} \right]^{-\frac{1}{\theta_w}} \quad (\text{B.4})$$

## B.2 Steady State

In this appendix, we explain how to find the steady state of the model. Since we presume that the model is on average correctly specified, hence in steady state we have  $\tau^i = 1$ ,  $\forall i \in \{L, K, Y, y, h, B\}$ . Our model has no long run growth, and thus, we can simply normalized the steady-state technology level to  $z = 1$ . Note that equation (B.1), (B.2), (B.3) and (B.4) imply  $\tilde{p} = \tilde{w} = p^* = w^* = 1$ . We assume the target (gross) inflation rate  $\pi$  is exogenously given. We also assume that at steady state the capital utilization  $u = 1$ . We normalized the steady steady labor (ie,  $L=1$ ).

Since the steady state inflation rate is  $\pi$ , equilibrium condition for bond demand ( equation 4.21) implies the steady-state nominal interest rate  $R = \frac{\pi}{\beta}$ . Equilibrium condition for investment (equation 4.19) implies  $P_{k'} = 1$ , since  $\mathcal{F}_{1,t}(I, I) = 1$  and  $\mathcal{F}_{2,t+1}(I, I) = 0$ . Moreover, the equilibrium condition of capital accumulation (equation 4.18) implies that the steady-state real capital rental rate is  $r^K = \frac{1}{\beta} - (1 - \delta)$ .

Since  $\tilde{p} = 1$ , equation (4.10), (4.11) and(4.12) implies steady state real marginal cost is  $s = \frac{1}{\lambda_f}$ . Since we have already steady-state values of  $s$ ,  $R$  and  $r^K$ , the real marginal cost function (equation 4.8) implies that the wage rate of homogeneous labor services is  $w = (s/A)^{\frac{1}{1-\alpha}}/R$ , where  $A = (1/\alpha)^\alpha (1/1-\alpha)^{1-\alpha} (r^K)^\alpha$ .

By equation (4.7), define  $\tilde{\phi} \equiv \left( \frac{\alpha}{1-\alpha} \right) \left( \frac{Rw}{r^K} \right)$ . Hence, we have  $K = \tilde{\phi}L$ . Besides, since we impose full capital utilization in the steady sate (ie,  $u = 1$ ), equation (4.13) implies  $K = \bar{K}$ . Since  $\mathcal{S}(1) = \mathcal{S}'(1) = 0$ , capital accumulation equation (equation 4.14) implies the steady state investment is  $I = \delta K$

Since  $w^* = 1$ , equation (4.32) implies  $L = H$ . Since  $p^* = w^* = z = 1$ , equation (4.34) implies

that the steady-state aggregate output is  $Y = K^\alpha - \Phi$ . Because, we assume  $a(u) = a(1) = 0$ , resource constraint of the economy (equation 4.35) implies  $C = Y - I$ . With the value of steady-state consumption, the consumption Euler equation (4.17) implies the marginal utility of consumption is  $\psi = \frac{(1-\beta b)}{(1-b)^\gamma} C^{-\gamma}$ . Moreover, since  $\tilde{w} = 1$ , equation (4.25), (4.26) and (4.27) imply  $\psi_h = \frac{w\psi}{\lambda_w}$ . Now, we have all steady-state values.

### B.3 Linearized Equilibrium Conditions

Here, we demonstrate the log-linearized equilibrium condition of the model. Basically, we adopt the following notation convention. For any variable  $x_t$ , we denote  $\hat{x}_t$  as the log-deviation from its steady state  $x$ :

$$\hat{x}_t \equiv \log \frac{x_t}{x}.$$

All of steady-state values have been discuss in the previous section.

Equation (B.1) can be linearized as

$$(1 - \xi_p) \hat{p}_t = \xi_p (\hat{\pi}_t - \hat{\pi}_{t-1}) - \hat{\tau}_t^y + \xi_p \hat{\tau}_{t-1}^y. \quad (\text{B.5})$$

Equation (B.2) can be linearized as

$$(1 - \xi_w) \hat{w}_t = \xi_w (\hat{\pi}_t - \hat{\pi}_{t-1}) + \xi_w (\hat{w}_t - \hat{w}_{t-1}) - \hat{\tau}_t^h + \xi_w \hat{\tau}_{t-1}^h. \quad (\text{B.6})$$

Equation (B.3) implies

$$\hat{p}_t^* = (1 - \xi_p) \hat{p}_t + [\hat{\tau}_t^y - \xi_p (\hat{\pi}_t - \hat{\pi}_{t-1}) - \xi_p \hat{\tau}_{t-1}^y] + \xi_p \hat{p}_{t-1}^*.$$

Imposing equation (B.5), above equation becomes

$$\hat{p}_t^* = \xi_p \hat{p}_{t-1}^*.$$

Hence, if we assume  $\hat{p}_0^* = 0$ , then  $\hat{p}_t^* = 0$  for all  $t$ . Even  $\hat{p}_0^* \neq 0$ ,  $\hat{p}_t = 0$  for large  $t$ , since  $\xi_p \in [0, 1)$ .

Equation (B.4) can be linearized as

$$\hat{w}_t^* = (1 - \xi_w) \hat{w}_t + \left[ \hat{\tau}_t^h - \xi_w (\hat{\pi}_t - \hat{\pi}_{t-1}) - \xi_w (\hat{w}_t - \hat{w}_{t-1}) - \xi_w \hat{\tau}_{t-1}^h \right] + \xi_w \hat{w}_{t-1}^*.$$

Imposing (B.6), we have

$$\hat{w}_t^* = \xi_w \hat{w}_{t-1}^*.$$

Similar argument implies  $\hat{w}_t^* = 0, \forall t$ .

**Capital Labor Ratio.** Linearization of equation (4.7) is

$$\hat{K}_t = \hat{L}_t + \hat{R}_t + \hat{w}_t - \hat{r}_t^K + \hat{\tau}_t^L - \hat{\tau}_t^K. \quad (\text{B.7})$$

**Real Marginal Cost function.** Linearization of equation (4.8) gives us

$$\hat{s}_t + \hat{z}_t = (1 - \alpha) \left( \hat{R}_t + \hat{w}_t \right) + \alpha \hat{r}_t^K + (1 - \alpha) \hat{\tau}_t^L + \alpha \hat{\tau}_t^K. \quad (\text{B.8})$$

**(Price) Phillip Curve.** Note that equation (4.11) can be linearized as

$$\hat{\mathcal{D}}_t^p = (1 - \beta \xi_p) \left[ \hat{\psi}_t - \theta_f \left( \hat{\tau}_t^y + \hat{p}_t \right) + \hat{Y}_t \right] + (\beta \xi_p) \mathbb{E}_t \left\{ \hat{\pi}_t - \hat{\pi}_{t+1} + \hat{\mathcal{D}}_{t+1}^p \right\},$$

since in the steady state  $\psi Y = (1 - \beta \xi_p) D$ . Similarly, equation (4.12) can be linearized as

$$\hat{\mathcal{N}}_t = (1 - \beta \xi_p) \left[ \hat{\psi}_t - \theta_f \left( \hat{\tau}_t^y + \hat{p}_t \right) + \hat{Y}_t + \hat{s}_t \right] + (\beta \xi_p) \mathbb{E}_t \left\{ \hat{\mathcal{N}}_{t+1}^p \right\},$$

since in the steady state  $\psi Y = (1 - \beta \xi_p) D$  and  $s = 1/\lambda_f$ . Combining above two equations with  $\hat{p}_t = \hat{\mathcal{N}}_t - \hat{\mathcal{D}}_t$ , which is the linearization of equation (4.10), we have

$$\hat{p}_t = (1 - \beta \xi_p) s_t + (\beta \xi_p) \mathbb{E}_t \left\{ (\hat{\pi}_{t+1} - \hat{\pi}_t) + \hat{p}_{t+1} \right\}.$$

Furthermore, substituting equation (B.5) into above equation and rearranging terms, we have the following Phillips curve

$$\begin{aligned}\hat{\pi}_t &= \frac{1}{1+\beta}\hat{\pi}_{t-1} + \frac{\beta}{1+\beta}\mathbb{E}_t\{\hat{\pi}_{t+1}\} + \left(\frac{1-\beta\xi_p}{1+\beta}\right)\left(\frac{1-\xi_p}{\xi_p}\right)\hat{s}_t \\ &\quad - \left(\frac{1}{1+\beta}\right)\hat{\tau}_{t-1}^y + \left(\frac{\beta\xi_p^2+1}{(1+\beta)\xi_p}\right)\hat{\tau}_t^y - \frac{\beta}{1+\beta}\mathbb{E}_t\{\hat{\tau}_{t+1}^y\}.\end{aligned}\quad (\text{B.9})$$

**Capital Evolution Equation.** Equation (4.13) implies  $\hat{K}_t = \hat{u}_t + \hat{K}_t$ . Moreover, equation (4.14) implies  $\hat{K}_{t+1} = (1-\delta)\hat{K}_t + \delta\hat{I}_t$ . Therefore, we have

$$\left(\hat{K}_{t+1} - \hat{u}_{t+1}\right) = (1-\delta)\left(\hat{K}_t - \hat{u}_t\right) + \delta\hat{I}_t, \quad (\text{B.10})$$

since we assume  $\mathcal{S}(1) = \mathcal{S}'(1) = 0$ .

**Consumption Euler Equation.** The equation (4.17) can be linearize

$$\hat{C}_t = \frac{b}{1+\beta b^2}\hat{C}_{t-1} + \frac{\beta b}{1+\beta b^2}\mathbb{E}_t\hat{C}_{t+1} - \left(\frac{(1-\beta b)(1-b)}{\gamma(1+\beta b^2)}\right)\left(\hat{\psi}_t + \hat{\tau}_t^Y\right). \quad (\text{B.11})$$

**Capital Supply Equation.** Equation (4.18) implies

$$\hat{P}_{k',t} + \hat{\psi}_t = \beta\mathbb{E}_t\left\{r^K\left(\hat{\psi}_{t+1} + \hat{r}_{t+1}^K\right) + (1-\delta)\hat{P}_{k',t+1}\right\}, \quad (\text{B.12})$$

since we assume  $a(u) = a(1) = 0$ .

**Investment Equation.** Equation (4.19) implies

$$\hat{P}_{k',t} = \varkappa_s\mathbb{E}_t\left\{\left(\hat{I}_t - \hat{I}_{t-1}\right) - \beta\left(\hat{I}_{t+1} - \hat{I}_t\right)\right\} + \hat{\tau}_t^Y \quad (\text{B.13})$$

where  $\varkappa_s = \mathcal{S}''(1)$ .

**Capital Utilization Equation.** Equation (4.20) implies

$$\mathbb{E}_t \{ \hat{r}_t^K - \hat{r}_t^Y - \varkappa_a \hat{u}_t \} = 0 \quad (\text{B.14})$$

where  $\varkappa_a = a''(1)/a'(1)$ , a parameter to be estimated.

**Bond Demand Equation.** Equation (4.21) can be linearized as

$$\hat{\psi}_t = \beta \mathbb{E}_t \left\{ \hat{\psi}_{t+1} + \hat{R}_t - \hat{\pi}_{t+1} + \hat{\tau}_{t+1}^B \right\}. \quad (\text{B.15})$$

**(Wage) Phillips Curve.** Equation (4.27) can be linearized as

$$\hat{\mathcal{N}}_t^w = (1 - \beta \xi_w) \left[ -\theta_w (\varphi + 1) \left( \hat{w}_t + \hat{\tau}_t^h \right) + (\varphi + 1) \hat{L}_t \right] + (\beta \xi_w) \mathbb{E}_t \left\{ \hat{\mathcal{N}}_t^w \right\}.$$

Equation (4.26) can be linearized as

$$\hat{\mathcal{D}}_t^w = (1 - \beta \xi_w) \left[ \hat{w}_t + \psi_t - \theta_w \left( \hat{w}_t + \hat{\tau}_t^h \right) + \hat{L}_t \right] + (\beta \xi_w) \mathbb{E}_t \left\{ (\hat{w}_t - \hat{w}_{t+1}) + (\hat{\pi}_t - \hat{\pi}_{t+1}) + \hat{\mathcal{D}}_t^w \right\}.$$

Therefore, we have

$$\hat{\hat{w}}_t = (1 - \beta \xi_w) \left[ -\hat{w}_t - \psi_t - \theta_w \varphi \left( \hat{w}_t + \hat{\tau}_t^h \right) + \varphi \hat{L}_t \right] + (\beta \xi_w) \mathbb{E}_t \left\{ (\hat{w}_{t+1} - \hat{w}_t) + (\hat{\pi}_{t+1} - \hat{\pi}_t) + \hat{\hat{w}}_{t+1} \right\},$$

since  $\hat{\hat{w}}_t = \hat{\mathcal{N}}_t^w - \hat{\mathcal{D}}_t^w$ , which is the linearization of equation(4.25). Imposing equation (B.6), we have

$$\begin{aligned} (\theta_a + \theta_b + \beta) \hat{w}_t &= \theta_a \hat{w}_{t-1} + \beta \mathbb{E}_t \hat{w}_{t+1} + \theta_a \hat{\pi}_{t-1} - (\theta_a + \beta) \hat{\pi}_t + \beta \mathbb{E}_t \{ \hat{\pi}_{t+1} \} \\ &\quad + \theta_b \varphi \hat{L}_t - \theta_b \hat{\psi}_t - \theta_a \hat{\tau}_{t-1}^h + \left( \frac{\theta_a}{\xi_w} - \theta_b \theta_w \varphi + \beta \xi_w \right) \hat{\tau}_t^h - \beta \mathbb{E}_t \hat{\tau}_{t+1}^h, \end{aligned} \quad (\text{B.16})$$

where  $\theta_a = 1 + (1 - \beta \xi_w) \theta_w \varphi$  and  $\theta_b = (1 - \beta \xi_w) \left( \frac{1 - \xi_w}{\xi_w} \right)$ .

**The Aggregate Output.** Equation (4.32) implies  $\hat{H}_t = \hat{L}_t$  since  $\hat{w}_t^* = 0, \forall t$ . Therefore, from now on, all  $\hat{H}_t$  will be replaced by  $\hat{L}_t$  without mentioning again. Equation (4.34) implies

$$\hat{Y}_t = (1 + \phi) \left[ \hat{z}_t + \alpha \hat{K}_t + (1 - \alpha) \hat{L}_t \right], \quad (\text{B.17})$$

where  $\phi = \Phi/Y$ .

**The Resource Constraint.** Equation (4.35) implies

$$s_c \hat{C}_t + \delta s_k \hat{I}_t + r^K s_k \hat{u}_t = \hat{Y}_t, \quad (\text{B.18})$$

where  $s_c = C/Y, s_K = K/Y$ .<sup>1</sup>

**Taylor Rule** Taking logarithm on the both side of equation (4.36), we obtain the linearized Taylor rule

$$\hat{R}_{t+1} = \rho_R \hat{R}_t + \left[ \gamma_\pi \mathbb{E}_t \hat{\pi}_{t+1} + \gamma_Y \mathbb{E}_t \hat{Y}_{t+1} \right] + \sigma_R \varepsilon_{t+1}^R. \quad (\text{B.19})$$

**Technology shocks and Wedges.** Equations (4.4), (4.5), (4.6), (4.15), (4.16), (4.2), and (4.23) can all be expressed as the log deviation form. For the sake of completeness of equilibrium

---

<sup>1</sup>Recall that  $a'(1) = r^k$ .

conditions, we show them below

$$\hat{z}_{t+1} = \rho_z \hat{z}_t + \sigma_z \varepsilon_{t+1}^z, \quad (\text{B.20})$$

$$\hat{\tau}_{t+1}^K = \rho_K \hat{\tau}_t^K + \sigma_K \varepsilon_{t+1}^K, \quad (\text{B.21})$$

$$\hat{\tau}_{t+1}^L = \rho_L \hat{\tau}_t^L + \sigma_L \varepsilon_{t+1}^L, \quad (\text{B.22})$$

$$\hat{\tau}_{t+1}^Y = \rho_Y \hat{\tau}_t^Y + \sigma_Y \varepsilon_{t+1}^Y, \quad (\text{B.23})$$

$$\hat{\tau}_{t+1}^B = \rho_B \hat{\tau}_t^B + \sigma_B \varepsilon_{t+1}^B, \quad (\text{B.24})$$

$$\hat{\tau}_{t+1}^y = \rho_y \hat{\tau}_t^y + \sigma_y \varepsilon_{t+1}^y, \quad (\text{B.25})$$

$$\hat{\tau}_{t+1}^h = \rho_h \hat{\tau}_t^h + \sigma_h \varepsilon_{t+1}^h. \quad (\text{B.26})$$

## B.4 Misspecified Models for Simulation Exercises

In this appendix, we list the equilibrium conditions with respect to each misspecified models.

### $\mathcal{M}(1)$ Model:

In this misspecified model, the working capital channel is absent. The corresponding linearized equilibrium conditions are

$$\begin{aligned} \hat{Y}_t &= E_t \left\{ \hat{Y}_{t+1} \right\} - \frac{1}{\gamma} \left( \hat{R}_t - E_t \hat{\pi}_{t+1} \right) + (1 - \rho_g) \hat{g}_t - \frac{1}{\gamma} \left\{ (1 - \rho_Y) \hat{\tau}_t^Y + \rho_B \hat{\tau}_t^B \right\}, \\ \hat{\pi}_t &= \kappa \left( (\gamma + \varphi) \hat{Y}_t - \gamma \hat{g}_t - (\varphi + 1) \hat{z} \right) + \beta E_t \hat{\pi}_{t+1} + \kappa \left( \hat{\tau}_t^Y + \hat{\tau}_t^L \right), \\ R_t &= \rho_1 \hat{R}_t + \rho_2 \hat{R}_{t-1} + (1 - \rho_1 - \rho_2) \left\{ \gamma_\pi \hat{\pi}_{t+1} + \gamma_Y \hat{Y}_{t+1} \right\} + \sigma_R \varepsilon_{t+1}^R, \end{aligned}$$

accompanying with the wedge processes, Eq. (4.45). Notice that in this case the nominal interest  $R_t$  does not enter the NKPC since that working capital channel is ignored.

### $\mathcal{M}(2)$ Model:

In this case, the monetary policy rule is misspecified in the sense that there is only one lag of nominal interest rate is incorporated rather than two lags. The corresponding linearized conditions are

$$\begin{aligned}\hat{Y}_t &= E_t \left\{ \hat{Y}_{t+1} \right\} - \frac{1}{\gamma} \left( \hat{R}_t - E_t \hat{\pi}_{t+1} \right) + (1 - \rho_g) \hat{g}_t - \frac{1}{\gamma} \left\{ (1 - \rho_Y) \hat{\tau}_t^Y + \rho_B \hat{\tau}_t^B \right\}, \\ \hat{\pi}_t &= \kappa \left( \hat{R}_t + (\gamma + \varphi) \hat{Y}_t - \gamma \hat{g}_t - (\varphi + 1) \hat{z} \right) + \beta E_t \hat{\pi}_{t+1} + \kappa \left( \hat{\tau}_t^Y + \hat{\tau}_t^L \right), \\ \hat{R}_t &= \rho_1 \hat{R}_t + (1 - \rho_1 - \rho_2) \left\{ \gamma_\pi \hat{\pi}_{t+1} + \gamma_Y \hat{Y}_{t+1} \right\} + \sigma_R \varepsilon_{t+1}^R,\end{aligned}$$

accompanying with the wedge processes, Eq. (4.45).

### $\mathcal{M}(3)$ Model:

In this case, the econometrician do not include the government spending shock, which is one of the fundamental shock of the true model. The linearized equilibrium conditions are

$$\begin{aligned}\hat{Y}_t &= E_t \left\{ \hat{Y}_{t+1} \right\} - \frac{1}{\gamma} \left( \hat{R}_t - E_t \hat{\pi}_{t+1} \right) - \frac{1}{\gamma} \left\{ (1 - \rho_Y) \hat{\tau}_t^Y + \rho_B \hat{\tau}_t^B \right\}, \\ \hat{\pi}_t &= \kappa \left( \hat{R}_t + (\gamma + \varphi) \hat{Y}_t - (\varphi + 1) \hat{z} \right) + \beta E_t \hat{\pi}_{t+1} + \kappa \left( \hat{\tau}_t^Y + \hat{\tau}_t^L \right), \\ \hat{R}_{t+1} &= \rho_1 \hat{R}_t + \rho_1 \hat{R}_{t-1} + (1 - \rho_1 - \rho_2) \left\{ \gamma_\pi \hat{\pi}_{t+1} + \gamma_Y \hat{Y}_{t+1} \right\} + \sigma_R \varepsilon_{t+1}^R,\end{aligned}$$

accompanying with the wedge processes, Eq. (4.45).