著者 | 松井他 \n学位授与機関 | 東北大学 \n学位授与番号 | 学術院第134号 \n番号 | 20120120001
Tohoku University
Graduate School of Economics and Management

Doctoral Dissertation

Essays on the Empirical DSGE Approach:
Estimation Methodologies and Applications

Tatsuyoshi Matsumae
Assistant Professor
Graduate School of Economics and Management, Tohoku University

August, 2017
## Contents

1 Introduction ................................................. 7
   1.1 Developments and Overviews of Empirical DSGE Approach ................. 7
      1.1.1 Overviews of the Empirical DSGE Approach .......................... 9
      1.1.2 Further Extensions .................................................. 14
      1.1.3 DSGE Models in Policy Institutions .............................. 21
   1.2 Basic Theory and Estimation Method ................................ 22
      1.2.1 Simple DSGE Model .................................................. 23
      1.2.2 Solving the DSGE model .......................................... 32
      1.2.3 Estimating the DSGE model ...................................... 34
   1.3 Organization of the Thesis ....................................... 39

I Estimation Methodologies .................................... 41

2 Role of Measurement Error .................................. 43
   2.1 Introduction ...................................................... 43
      2.1.1 Background ......................................................... 43
      2.1.2 Purposes, Originalities, and Methodologies ..................... 46
      2.1.3 My Contributions ............................................... 47
      2.1.4 Organization of Chapter 2 ..................................... 48
   2.2 Model ............................................................ 48
      2.2.1 Households ......................................................... 48
      2.2.2 Firms ............................................................. 49
      2.2.3 Other Equilibrium Conditions .................................. 50
      2.2.4 AR(1) Shocks and the Measurement Errors ..................... 51
      2.2.5 Structural Form .................................................. 51
   2.3 Estimation Method ............................................... 51
      2.3.1 State Space Representation ..................................... 51
      2.3.2 Bayesian Estimation Method ................................... 53
      2.3.3 Simulation Smoother ............................................. 55
      2.3.4 Measurement Equation Parameter ................................ 57
   2.4 Preliminary Settings and Data ................................ 59
      2.4.1 Observation Variables and Measurement Error .................... 59
      2.4.2 Calibrated Parameters and Prior Settings ...................... 60
      2.4.3 Data .............................................................. 61
   2.5 Results .......................................................... 62
3 Estimation in a Data Rich Environment

3.1 Introduction ......................................................... 75
3.1.1 Background ..................................................... 75
3.1.2 Purposes, Originalities, and Methodologies .................... 83
3.1.3 My Contributions .............................................. 84
3.1.4 Organization of Chapter 3 ..................................... 85
3.2 Data Rich Approach ............................................... 85
3.3 Model .............................................................. 89
3.3.1 Households ...................................................... 90
3.3.2 Firms ............................................................. 91
3.3.3 Miscellaneous Equilibrium Conditions ......................... 92
3.3.4 Persistent Shocks and Forecast Errors ....................... 92
3.3.5 Log-linearized Model ........................................ 92
3.4 Estimation Method ................................................. 93
3.4.1 State Space Model ............................................. 93
3.4.2 Hybrid MCMC Procedure .................................... 93
3.5 Preliminary Settings and Data ................................... 94
3.5.1 Measurement Equation ....................................... 94
3.5.2 Data ........................................................... 96
3.5.3 Calibration and Prior Settings ................................ 97
3.6 Results ............................................................ 97
3.6.1 Model Concepts ................................................. 97
3.6.2 Structural Shocks .............................................. 99
3.6.3 Structural Parameters ....................................... 100
3.6.4 Policy Simulation ............................................ 101
3.6.5 Sources of the Business Cycle ................................ 101
3.6.6 Discussion ..................................................... 104
3.7 Conclusion ......................................................... 105
3.8 Tables and Figures ............................................... 107

II Applications .......................................................... 129

4 Sources of the Great Recession ........................................ 131
4.1 Introduction ......................................................... 131
4.1.1 Background ..................................................... 131
4.1.2 Purposes, Originalities, and Methodologies .................... 134
4.1.3 My Contributions .............................................. 136
4.1.4 Organization of Chapter 4 ..................................... 136
## CONTENTS

4.2 Data Rich Approach with Stochastic Volatility ........................................... 137  
  4.2.1 Introducing Stochastic Volatility with Leverage Effect ........................... 137  
  4.2.2 DSGE Model in a Data Rich Environment .............................................. 138  
  4.2.3 DSGE model with Stochastic Volatility in a Data Rich Environment .......... 140  
4.3 Model ........................................................................................................... 141  
  4.3.1 Financial Friction in Corporate Sector .................................................... 142  
  4.3.2 Financial Friction in Banking Sector ...................................................... 145  
  4.3.3 Incorporating Two Financial Frictions into the DSGE model ................. 147  
4.4 Estimation Method ......................................................................................... 148  
4.5 Preliminary Settings and Data ....................................................................... 150  
  4.5.1 Four Cases ............................................................................................... 150  
  4.5.2 Calibration and Prior Settings .................................................................. 150  
  4.5.3 Data ......................................................................................................... 151  
4.6 Results ........................................................................................................... 152  
  4.6.1 Structural Parameters .............................................................................. 152  
  4.6.2 Structural Shocks and Volatilities ............................................................ 153  
  4.6.3 Historical Decompositions ..................................................................... 155  
  4.6.4 Discussions ............................................................................................... 157  
4.7 Conclusion ..................................................................................................... 157  
4.8 Tables and Figures ......................................................................................... 159  
4.9 Appendix ....................................................................................................... 175  
  4.9.1 Sampling Stochastic Volatility with Leverage Effect ......................... 175  
  4.9.2 Simulation Smoother ................................................................................ 177  
  4.9.3 Sampling Measurement Equation Parameters ...................................... 179  
  4.9.4 Remaining Framework of the DSGE model ........................................... 181  

5 Impacts of Government Spending on Unemployment ........................................ 187  
5.1 Introduction ................................................................................................... 187  
  5.1.1 Background ............................................................................................. 187  
  5.1.2 Purposes, Originalities, and Methodologies ......................................... 189  
  5.1.3 My Contributions .................................................................................... 190  
  5.1.4 Organization of Chapter 5 ...................................................................... 190  
5.2 Model ........................................................................................................... 190  
  5.2.1 Unemployment ......................................................................................... 190  
  5.2.2 Non-wasteful Government Spending ..................................................... 192  
  5.2.3 Effects of Non-wasteful Fiscal Expansions for Unemployment ............ 193  
5.3 Estimation Method ......................................................................................... 195  
  5.3.1 Data and Measurement Equation ............................................................ 195  
  5.3.2 Preliminary Settings ................................................................................ 196  
  5.3.3 Prior Distributions .................................................................................. 197  
5.4 Results ........................................................................................................... 197  
  5.4.1 Estimated Parameters .............................................................................. 197  
  5.4.2 IRFs .......................................................................................................... 198  
  5.4.3 Historical Decomposition ..................................................................... 199  
  5.4.4 Robustness Check ................................................................................... 200  
5.5 Conclusion ..................................................................................................... 201  
5.6 Tables and Figures ......................................................................................... 202
5.7 Appendix .......................................................... 214
5.7.1 Model Description ............................................. 214
5.7.2 Summary of the Model ........................................ 224
5.7.3 Steady States .................................................. 228

6 Conclusion .......................................................... 231
6.1 Chapter 2: Role of Measurement Error ........................................ 231
   6.1.1 Research Questions ........................................... 231
   6.1.2 Methods .................................................... 232
   6.1.3 Results ..................................................... 232
   6.1.4 Conclusions ................................................ 232
   6.1.5 Future Research ............................................. 232
6.2 Chapter 3: Estimation in a Data Rich Environment .............................. 232
   6.2.1 Research Questions ........................................... 232
   6.2.2 Methods .................................................... 233
   6.2.3 Results ..................................................... 233
   6.2.4 Conclusions ................................................ 233
   6.2.5 Future Research ............................................. 234
6.3 Chapter 4: Sources of the Great Recession ...................................... 234
   6.3.1 Research Questions ........................................... 234
   6.3.2 Methods .................................................... 234
   6.3.3 Results ..................................................... 234
   6.3.4 Conclusions ................................................ 235
   6.3.5 Future Research ............................................. 235
6.4 Chapter 5: Impacts of Government Spending on Unemployment .................. 235
   6.4.1 Research Questions ........................................... 235
   6.4.2 Methods .................................................... 236
   6.4.3 Results ..................................................... 236
   6.4.4 Conclusions ................................................ 236
   6.4.5 Future Research ............................................. 236
Chapter 1

Introduction

1.1 Developments and Overviews of Empirical DSGE Approach

This thesis consists of estimation methodologies and applied examples based on the empirical dynamic stochastic general equilibrium (hereinafter, DSGE) approach.

First of all, this introduction briefly outlines the historical background to the empirical DSGE approach. As is well known, the field of the business cycle in macroeconomics has evolved in response to two highly influential criticisms: One is the Lucas critique, and the other is the Sims critique. The former is a criticism on the policy effect evaluation and the latter is a criticism on the measurement method in extracting a certain policy effect, but both criticisms commonly requested to construct a structurally interpretable model, i.e. a model with microeconomic foundations.

Since the two critiques, structural models have been developed with microeconomic foundations to explain the business cycle. The real business cycle (hereinafter, RBC) model is the earliest result: The RBC model explained the business cycle with the unanticipated fluctuations of productivity, but more importantly, explicitly considered the dynamic optimization behaviors of firms and households, and described the business cycle as optimal reactions to unexpected shocks. Speaking of comparison with the conventional models without dynamic optimizations, rather we should emphasize the difference in agents’ responses to “anticipated” shocks. For example, if households anticipate a rise of future productivity, which will cause rises of future real wage and real rental price of capital, because it will raise both future marginal productivities of labor and capital. Then, “current” consumption might go up, if the wealth effect (in anticipation of future income increases, the effect of increasing current consumption to smooth consumption intertemporally) dominates the substitution effect (in anticipation of the future rise in rental price, the effect of accumulating more capital by reducing the current consumption). By contrast, in backward-looking models, the anticipated future shocks do not have any impact on the current behaviors.

The RBC model based on flexible price adjustment, however, cannot reproduce responses of real aggregates such as output and employment against nominal disturbances such as changes in nominal money and nominal interest rate. Then, this model has been extended to a sticky price model as called the new Keynesian model, and it has become a new tool to evaluate the effect of monetary policy. Essentially, the new Keynesian model consists of three fundamental equations: (1) The new IS curve obtained from households’ optimization behaviors mainly explains output fluctuations. (2) The new Keynesian Phillips curve (hereinafter, NKPC) derived from firms’ optimizing pricing behaviors illustrates inflation variations. Both the new IS curve and the NKPC, based on dynamic optimizations, include forward-looking terms, i.e., future expectations influence
CHAPTER 1. INTRODUCTION

the current behaviors. In addition, by introducing Keynesian characteristic, the nominal rigidity, the NKPC (the short-term aggregate supply curve) becomes upward-sloping, thus, nominal money becomes non-neutral. Hence, the new Keynesian model adds another element not included in the RBC model: (3) It is the central bank’s (committed) stabilization policy intended specifically to control private sectors’ future inflation expectations. Usually, a certain monetary policy rule (called the Taylor rule) is assumed to be conducted by the central bank, manipulating nominal interest rate so as to stabilize both output and inflation fluctuations.

With the development of theory, empirical analysis also progressed. First, time series analysis called as the structural vector autoregressive (hereinafter, VAR) model developed where empirical results had been accumulated, supporting the new Keynesian features: The real aggregates respond to nominal disturbances. However, since VAR models are reduced-form models, it is difficult to identify shocks of policy instruments and explain structurally why real variables respond to nominal variables: In other words, the VAR analysis is not a method sufficiently responding to the above two criticisms.

As the next trend, especially since the 2000s, the empirical DSGE approach was established: It is the approach to build a structural model with microeconomic foundations and estimate parameters utilizing not only data information but also information of the structural model itself. The empirical DSGE approach rapidly penetrated not only to the academic circles but also to the practical circles such as central banks and governments of various countries. One of the reasons for this popularization was probably due to the surprising suitability of the DSGE model for data. Thus, the model with the microeconomic foundations not only can explain the sources of the business cycle according to the microeconomic theory, but also can successfully capture actual output and inflation fluctuations. In other words, the empirical DSGE approach has established, since it becomes possible to evaluate or measure a certain policy effect using the estimated DSGE model, which can withstand the two major critiques well.

After this approach has been established, the DSGE model is currently further extended to various dimensions based on the standardized model. Hence, this section will provide a brief survey on the recent directions of the empirical DSGE approach which seems to be important.

Although the recent DSGE models are becoming larger size through numerous extensions, the empirical DSGE approach has the core theory and empirical method that can be called as a common platform. All existing studies of this approach have been developed with the common platform as a starting point. Of course, my research also extends the platform from both the theoretical side and the empirical side. Therefore, this introduction also summarizes the simplest basic theory and empirical method. It would be helpful to show the common platform for explaining the expansion of theoretical and empirical aspects carried out by this thesis.

It should note that, at the beginning of each chapter, I will explain the research outline, background, research purposes, originalities, and methodologies. In particular, the background and purposes that triggered each research will be described in detail to clarify questions on previous studies and summarize the motivation and significance of this research. In addition, there are many collaborative researches in the field of the empirical DSGE approach, and the research of all the chapters in this thesis is also collaborative research. Therefore, my contribution part is briefly summarized.

In sum, first of all, this introduction briefly overviews the historical background and development

1 Other reasons why the DSGE model has become widespread are firstly, good quality textbooks have been prepared such as Woodford (2003), McCandless (2008), Gali (2008) and Walsh (2010), and secondly it has become easily possible to estimate the model with free ware called “Dynare”.
of the empirical DSGE approach. Then, we illustrate the basic theory and estimation method of this approach. Finally, the organization of this thesis is summarized.

1.1.1 Overviews of the Empirical DSGE Approach

Two Critiques and DSGE model

Modern macroeconomics has developed under the following two criticisms on the conventional Keynesian model.

Lucas critique (Lucas 1976): The models without microeconomic foundations cannot identify which change in structural parameter caused the changes in reduced-form parameters: For instance, a change in the policy attitude towards inflation affect almost all changes in reduced-form parameters, but as long as dealing with models without microeconomic foundations, it is not possible to identify whether the policy response to inflation truly change or other structural parameters have simply changed without any change in the monetary policy rule.

Sims critique (Sims 1980): The matrix of the reduced-form equations is so far sparse to identify structural shocks: The traditional macro-models impose incredibly too much restrictions on structural parameters.

In the next subsection, after describing the features of the DSGE model with a simple model, we will consider again the precise meanings of the two critiques above. Importantly, in response to the criticisms above, the DSGE models had been constructed with microeconomic foundations.

Kydland and Prescott (1982) proposed the real business cycle model (RBC model) in which the business cycle is caused by a “real” shock called as the productivity shock, and showed that the actual U.S. business cycle can be well explained by the productivity variations using calibrated parameters (Regarding an example of Japan, see Hayashi and Prescott, 2002). The RBC model methodologies became the turning point of the subsequent business cycle theory: The advantage is, firstly, to describe the business cycle as responses based on dynamic optimization behaviors of the agents in the model (households and firms) against unanticipated structural shocks. Secondly, the model establishes a calibration method that sets parameters so that the second moments of data (variances and covariances) can be reproduced.

The RBC model has been expanded while receiving numerous criticisms. There are two notable criticisms: First, the calibrated RBC model cannot capture actual employment volatile fluctuations against productivity shock (Hansen and Wright, 1992). A rise in productivity raises both wage and real interest rate, which lowers today’s consumption and stimulates today’s labor supply. If today’s wage and real interest rate are higher, then it would be better to work today to saving and enjoy leisure tomorrow. This intertemporal substitution effect of labor supply depends upon the labor disutility parameter. To reproduce the sensitive employment reaction against the productivity shock, the labor disutility parameter must be implausibly small (i.e. the wage elasticity of labor supply must be too high to support from microeconomic evidence).

Another elements incorporated into the RBC model have various dimensions: (1) investment specific technology shock (Greenwood, et al. 1988), (2) government spending shock (Campbell, 1994), (3) multiple sectors (Long and Plosser, 1983) (4) multiple countries (e.g. Backus et al. 1992, Schmitt-Grohe and Uribe, 2003), (5) labor hoarding (Burnside, et al. 1993).
One of the modifications to overcome this criticism is the so-called indivisible labor model (Hansen, 1985). Usually, labor’s response to productivity shock is regarded as a reaction of working hours (intensive margin). Instead, assuming constant working hours, if we regard the labor’s response as a change in the number of employees (extensive margin), we can reproduce actual employment fluctuations.\(^3\)

The second criticism, however, is unavoidable in a sense, and it is difficult to deal with trivial modifications and extensions: A huge amount of literature provides the crucial evidence that real aggregates react to nominal disturbances shown by structural VAR analysis (e.g. Bernanke and Blinder, 1992, Leeper et al. 1996, Sims and Zha, 1998, Christiano et al. 1999). Even if changing monetary policy instrument from M2 to FF rate (Bernanke and Blinder, 1992), or even if examining the estimation accuracy of the impulse responses (Sims and Zha, 1998), or even if relaxing restrictions for monetary policy shock identification (Christiano et al. 1999), all of the literature consistently present a robust result that monetary policy shock affects real aggregates fluctuations. This non-negligible result cannot replicate from the RBC model based on flexible price and wage adjustments.

Long before the advantages and disadvantages of the RBC model have come to light, several literatures have already began attempting to introduce nominal rigidity into dynamic optimization models. Taylor (1979) modeled the nominal rigidity by considering a long-term contract of nominal wages (the so-called “staggered wage” model) where households contract with firms on a nominal wage fixed for two periods and there are two types of households: One could revise the wage contract in the even period and the other in the odd period. Rotemberg (1982) described a nominal price rigidity by specifying a quadratic cost function in firms’ price revisions. Calvo (1983) introduced a nominal price rigidity by giving an exogenous price revision probability. Then, Roberts (1995) showed that the same form of the NKPC can be derived from any setting of the above. Therefore, Yun (1996) rewrote as a calibration-possible discrete-time model based on Calvo (1983), and completed a prototype of the new Keynesian type DSGE model.

**Evidence on the New Keynesian Properties**

With the development of the theory, empirical evidence has also been accumulated and supported the new Keynesian properties such as the short-term upward-sloping aggregate supply curve or nominal rigidities.

**Evidence on the NKPC**: Introducing short-term nominal rigidity will bring the short-term upward-sloping aggregate supply curve (i.e. NKPC). Some empirical studies directly examined

\(^3\)Strictly speaking, this modification implicitly increases in the wage elasticity of labor supply. When labor’s responses are regarded as fluctuations in the number of employees, then unemployment occurs. Suppose that the number of workers employed is \(E_t\), the labor force is \(N_t\), the labor demand is \(L_t\), the constant working hour is \(l_0\), and the utility from leisure is specified by \(\sigma_L\ln(1-l_0)\) (\(\sigma_L\) is a positive parameter). Then, the labor market clearing condition can be represented by \(E_t l_0 = L_t\), and the probability to be hired is \(\frac{L_t l_0}{N_t}\), or equivalently, the probability to be not employed (unemployment rate) is \(1 - \frac{L_t l_0}{N_t}\). Since the working hours is zero in the state of unemployment, the expected utility in household can be written as:

\[
\frac{L_t l_0}{N_t} \cdot \sigma_L \ln(1 - l_0) + \left(1 - \frac{L_t l_0}{N_t}\right) \sigma_L \ln(1)
\]

This is a linear function of \(L_t\): Households become risk neutral against changes in labor supply. In other words, households are willing to change labor supply flexibly against wage changes. That’s why this modification can capture the actual volatile employment fluctuations.
whether the coefficient of output gap in the NKPC is positive or not. The common result of these 
studies is that the NKPC is indeed upward-sloping, but the inflation inertia is also important.\(^4\)

Fuhrer (1997) estimated the hybrid type NKPC (with both forward and backward terms of 
inflation), and confirmed the aggregate supply curve is upward-sloping in the short term. He also 
found, however, the coefficient of the backward term of inflation is relatively large, so he questioned 
price setting behaviors of firms based on dynamic optimization. In response to this result, Gali 
and Gertler (1999) justified the backward term of the NKPC introducing a lagged inflation index 
contract: If firms who cannot revise their prices, then their prices are assumed to slide with previous inflation. Their evidence also showed the importance of the backward term of inflation, by estimating the hybrid-type NKPC via the generalized moments method (hereinafter, GMM).\(^5\)

**Microeconomic evidence on the nominal rigidity:** Microeconomic evidence supporting the 
nominal rigidity also has been provided:

According to the results, the revision frequency is about once every 4.3 months, which is the 
frequency of revision once in 1.5 quarters. Therefore, the probability that the price cannot be 
revised (the so-called Calvo parameter) is about one-third.

Nakamura and Steinsson (2008) also estimated the price revision frequency of CPI in the U.S. 
and reported five fact findings: (i) Price revision frequency is high in the bargain period, (ii) price 
cutting is one-third of price revisions, (iii) Frequency of price hikes is positively correlated with 
inflation, but frequency of cutting is not correlated with inflation, (iv) price revision frequency has 
seasonality, and first quarter has a high frequency, (v) revision probability decreases in a few months 
after price revision.

Those studies commonly show the existence of the short-term nominal price rigidity (see also 

**More Sophisticated Statistical Evidence**

It is impossible to interpret the economic structure (propagation mechanism of monetary policy) 
accurately in empirical studies by VAR model or microeconomic evidence. On the other hand, when 
estimating a single equation for the NKPC or the monetary policy rule by GMM, the endogeneity 
problem remains. Therefore, it had been proceeded to simultaneously estimate all the equations of 
the new Keynesian model.

A likelihood-based estimation method was developed rather than the three-stage least squares 
(3SLS) method adopted for the conventional macro model estimation method. The (log-linearized) 
DSGE model can be expressed in the state space model. Hence, specifying the probability densities 
of structural shocks, we can evaluate the likelihood by the Kalman filter, that is, we can estimate 
the DSGE model via the maximum likelihood (hereinafter, ML) method.\(^6\)

It should be noted that there is one important assumption when estimating the DSGE model 
using the Kalman filter. It is assumed that agents in the model (households, firms, central bank, 
etc.) can exactly observe endogenous variables (output gap, inflation gap, etc.) related to their

\(^4\)Ball (1994) explained the necessity of inflation inertia from another perspective. In the absence of inflation 
persistence (or inflation is jump variable), the boom will occur if the central bank permanently lowers nominal money growth, which is the opposite result of our prediction.

\(^5\)As another method introducing inflation inertia, sticky information models have been proposed. See Mankiw and Reis (2002) and Devereux and Yetman (2003).

\(^6\)The conventional (linearized) Keynesian model can be also represented by the state space model, so we can employ the ML method to estimate the model. See Sargent (1989).
CHAPTER 1. INTRODUCTION

decision making. If this assumption collapses, agents in the model will make their own decisions by predicting endogenous variables, and the likelihood evaluation using the Kalman filter will not be valid.

Subsequently, several empirical studies appeared simultaneously estimate all equations of the DSGE model by the Bayesian technique. If we can evaluate the likelihood, we can also estimate the posterior distribution of parameters by only giving prior information. The advantages of the Bayesian estimation is summarized as follows:

First, estimating simultaneously all of equations of the model could avoid the endogeneity problem without finding out good instrumental variables.

Second, we can simulate the policy effect, considering parameter uncertainties. The Bayesian inference objective is to estimate posterior distributions of structural parameters. If the credible interval of an endogenous variable response against a policy shock includes zero, this policy should be regarded as ineffective for the endogenous variable.

Third, we can also derive historical decompositions of endogenous variables, that is, we can quantitatively evaluate the policy contributions for business cycle.

Finally, the model comparison can be easily carried out by calculating marginal densities, which indicates the fit of the model for data.

**Evidence on the NKPC:** Ireland (2001) is one of the earliest papers estimating the DSGE model by the ML method. He pointed out the importance of the inflation inertia from the fit of the model to data by estimating a new Keynesian model with the Rotemberg type nominal price rigidity. Linde (2005), employing the Monte Carlo simulation, showed that an endogeneity bias is generated if the NKPC alone is estimated by GMM, but disappeared if all of the equations simultaneously estimated by the ML method. Then, he again reported the need for inflation inertia from results by ML method.

**Evidence on the nominal rigidity:** Ireland (2003) confirmed the price nominal rigidity has a crucial role for inflation fluctuations from results using ML method. Rabanal and Rubio-Ramirez (2005) would be one of the earliest studies that applied the Bayesian estimation to the new Keynesian model with Calvo type nominal rigidities. They provided four fact findings: (1) Both price and wage nominal rigidities are important, (2) inflation inertia is also needed, (3) the wage elasticity of labor supply is high, and (4) the Taylor coefficient for inflation is stable.

In this way, the basic theory had been constructed, the method of estimating the model had been established, and supports from various empirical analyzes had been received, which indicated the materials of the empirical DSGE approach seemed to be almost complete. With these simple new Keynesian models, however, it is difficult to capture all the aggregate data fluctuations. First, the simple new Keynesian model ignores the investment, regarded as one of the main factors of the business cycle. Second, the so-called “real” rigidities are not installed. The real rigidities will help to replicate the hump-shaped reactions of consumption and investment to structural shocks, revealed by VAR analysis. Following this, the most important model appeared at the current empirical DSGE approach.

---

To be precise, this paper is not the first attempt to estimate the DSGE model using Bayesian inference. To measure the cost of business cycle, Schorfheide (2000) to the cash-in-advance model and Otrok (2001) to the RBC model had already applied the Bayesian estimation technique.
1.1. DEVELOPMENTS AND OVERVIEWS OF EMPIRICAL DSGE APPROACH

The Standard Empirical DSGE model

It is Christiano et al. (2005, hereinafter CEE) and Smets and Wouters (2003, 2007, hereinafter SW) that are the most influential and remarkable literature in the current empirical DSGE approach.

The CEE model extended the prototype of the simple new Keynesian model mainly in three aspects: First, the model incorporates both price and wage nominal rigidities and lagged inflation indexation contracts (hybrid NKPC in goods market and labor market).\(^8\) Second, the model also introduce “real” rigidities such as habit formation on consumption and adjustment cost of investment. It helps to reproduce the hump-shaped consumption and investment responses against structural shocks.\(^9\) Third, monetary authority controls nominal interest rate according to the extended Taylor rule with nominal interest rate smoothing (i.e. introducing the backward term of the nominal interest rate).

They did not take the likelihood-based estimation strategy. Instead, they first estimated the impulse response functions to the monetary policy shock at the structural VARs which relaxed the identification restrictions proposed by Christiano, et al. (1999). Then, parameters were estimated to match the impulse response functions. The estimation results showed that high nominal rigidities were detected both in price and wage, and the real rigidities such as consumption habit formation and investment adjustment cost were also important. This response matching estimation methodology, however, requires a premise that the estimated impulse responses are sufficiently reliable.

Finally, empirical studies by SW (2003, 2007) triggered that the benchmark of the DSGE model is replaced from the RBC model into the CEE model. Utilizing the Euro area data (SW 2003) and the U.S. data (SW 2007), they estimated the CEE model with Bayesian technique (see also Levin, et al. 2006, Lubik and Schorfheide, 2004, 2007).

In particular, SW (2007) can be said as so ambitious work of having completed almost all of what can be done by the Bayesian inference. Their results showed not only the importance of the nominal rigidity, the plausibility of the NKPC, estimation of monetary policy rule and examination of the policy effect, but also answered from the estimated DSGE model to topics that are still controversial, such as whether the productivity shock boosts or lowers employment, or what the sources caused the Great Moderation.\(^10\)

They first estimated the CEE model as the baseline model by the Bayesian technique. Next, the case without nominal rigidities, the case without inflation inertia, the case without real rigidities, the model of each case was estimated and compared with the baseline model from the fit to the data. As a result, they confirmed that it is important to introduce not only nominal rigidities and inflation inertia but also real rigidities from the view of capturing data.

By showing historical decompositions, they explained the sources of business cycle and inflation fluctuation. According to results, almost all of inflation fluctuations are attributable to the markup variations.

In addition, as applied research, they found that a rise of productivity causes a temporary decline of employment. Moreover, they estimated the model by dividing the observation period and proved

---

\(^8\)Erceg et al. (2000) is one of the first studies to introduce the Calvo type nominal wage rigidity into the DSGE model.

\(^9\)On the structural examination of the consumption habit formation, see also Baukez, et al. (2005).

\(^10\)Against a rise in productivity, some researchers said employment will decline (e.g. Gali 1999), another insisted it will rise (e.g. Christiano, et al. 2004). Regarding the Great Moderation, some researchers stated the monetary tightening policy of the Volcker era was useful (e.g. Clarida, et al. 2000), another regarded it was simply lucky (e.g. Sims and Zha, 2006).
that the Great Moderation since the mid-1980s was not caused by monetary policy but simply due to the declining volatilities of structural shocks.

All the above results are fruitful, but one of the most amazing findings is that both the fit of the CEE model to the data and prediction accuracy of the CEE model were not inferior to the structural VAR models.

The VAR model is an atheoretical model that is not subject to restrictions necessary for shock identification. In other words, the VAR model can be said to be an empirical method to have the data tell the truth as much as possible. On the other hand, the DSGE model is an extremely theory-oriented model. The cross-equation constraints and parameter restrictions are much larger than the VAR model. In other words, estimating the DSGE model is a task of asking how much the story-teller model can reproduce the data. It is natural to think that the DSGE model will lose the VAR model in data fit and data prediction accuracy. Beyond our expectations, however, they showed that the DSGE model also has explanatory and predictive powers.\(^\text{11}\)

If so, since the DSGE model can be structurally interpreted and the realistic plausibility of the model is also guaranteed, it is only necessary to perform policy analysis using this model. As a result, the SW model was established as a standard model of the empirical DSGE approach and spread rapidly and widely.

### 1.1.2 Further Extensions

Even now, the DSGE models show various developments and some studies have revealed defects and limitations of the current standard DSGE model. This subsection summarizes nine dimensions of the main developments: (1) On giving microeconomic foundations to the monetary stabilization policy, (2) tackling the optimal monetary policy with zero lower bound of nominal interest rate, (3) considering the interactions between monetary and fiscal policies, (4) introducing financial friction and (5) search and matching friction in the labor market, (6) extending to an open economy model, (7) endogenizing nominal rigidities, (8) integrating growth model with DSGE model, and (9) examining the effect of anticipated shock.

**Optimal monetary policy:** The standard DSGE model assumes the central bank manipulates the nominal interest rate according to the so-called Taylor rule. A number of authors has also been conducted to give microeconomic foundations to the Taylor rule.

Svensson (1997) would be one of the first studies to explicitly consider the optimal monetary policy so as to maximize welfare (or equivalently, minimizing the welfare loss) under a setting with private agents taking backward looking behaviors (i.e. the conventional back-ward IS curve and Phillips curve). Then, Rotemberg and Woodford (1997) introduces the welfare loss minimizing problem into a simple dynamic optimization model (i.e. the new IS curve and the NKPC) to obtain an optimal monetary policy rule.

The method of considering the optimal monetary policy is as follows (linear-quadratic approach): Given the nominal interest rate, private agents’ optimal reactions to shocks are derived as the new IS curve (households) or the NKPC (firms). So how does the central bank decide the nominal interest rate to raise welfare? First, second-order approximating the utility function around the steady states, the central bank obtains a quadratic welfare loss function. Then, the monetary authority has only to manipulate the nominal interest rate to minimize the welfare loss function subject to

\(^{11}\text{Sensitivity analysis on prior distributions is provided in Del Negro, et al. (2007). See also Onatski and Williams (2010).}\)
private sectors’ reactions such as the new IS curve and the NKPC. As a consequence the Taylor rule is derived as the first order condition of minimizing the quadratic welfare loss function.\footnote{It is still being debated whether the central bank can commit the optimal monetary policy rule. In actual, it is difficult to commit to the optimal monetary policy. Therefore, the optimal “discretion” policy rule has also been proposed that the central bank re-optimizes the nominal interest rate every period. See, e.g. Walsh (2003). Normally, because there is another channel affecting expected inflation, the commitment rule has higher welfare gain than the discretion rule. See also Clarida et al. (2000) which is a canonical paper with easy-to-understand explanations of the optimal monetary policy rule given the quadratic welfare loss function. On the optimal policy in open economy model, see Benigno and Benigno (2003). Fujiwara et al. (2013) considers a simple optimal monetary policy in two country model when the other country is in a liquidity trap.

Recent empirical analysis examines whether there has been a change in the inflation target and evaluates welfare by changing the inflation target based on the quadratic welfare loss function. See, for example, Levine, et al. (2008), Feve, et al. (2010) and Curdia and Finocchiaro (2013).}

Based on this result, the standard DSGE model specifies the behavior of the monetary authority as the Taylor rule and examines the effect of the monetary policy shock.\footnote{Usually, parameters of monetary policy rule are often estimated as if “structural” parameters. But, how much inflation and output variations will reduce welfare depends upon structural parameters, such as CRRA parameter, Calvo parameter, etc. That is, following the linear-quadratic approach faithfully, the Taylor coefficients are also reduced-form parameters represented by highly nonlinear functions of structural parameters. See, for example, Schorfheide (2000) and Lubik and Schorfheide (2004).}

**Nonlinearity: ZLB of nominal interest rate:** One disadvantage of the DSGE model is that the model does not handle nonlinearity well. For instance, if there is irreversibility of investment due to a fixed cost, the optimal timing of investment will be delayed compared with the case without nonlinearity. However, just imposing a nonlinearity of nonnegative investment constraint, it also becomes extremely difficult to solve a model, not only to estimate the model. An example that remarkably expresses this drawback is the zero lower bound (hereinafter, ZLB) constraint of the nominal interest rate.

What makes the problem difficult by adding ZLB constraint on nominal interest rate? If there is a ZLB constraint, the monetary authority must commit to the two timings in advance: When or in what circumstances will the nominal interest rate be dropped to zero, and when or what circumstances will the monetary authority stop the zero interest rate policy. When the nominal interest rate falls to ZLB, the central bank will lose policy tool to manage inflation and aggregate demand on the model.\footnote{Of course, the story will change if we explicitly consider the central bank’s balance sheet (or equivalently, central bank’s budget constraint) and build a model with a channel that quantitative easing leads to a decline in the long-term interest rate. Alternatively, the central bank may raise the national debt outstanding by underwriting government bonds to obtain higher inflation expectation. Later, we will consider the interactions between monetary and fiscal policies.}

If so, it would be best to carry out significant monetary easing and avoid falling into the ZLB before the nominal interest rate is approaching zero. That is, the optimal monetary policy under the ZLB constraint will be represented not by a linear function as the Taylor rule but by a nonlinear function.

Kato and Nishiyama (2005) was just looking at the opposite situation, the so-called “exit strategy”: The economy had already been at zero interest rate, and they considered when the central bank should stop the zero interest rate policy and raise the nominal interest rate. They numerically solve the welfare loss minimization problem under the nonnegative constraint of nominal interest rate. Then, they showed the timing to get out of the zero interest rate should be delayed than the usual Taylor rule, that is, the central bank must be more monetary easier than the Taylor rule.\footnote{Eggertsson and Woodford (2003) would be one of the earliest studies of tackling the optimal monetary policy under the ZLB of nominal interest rate. See also Jung, et al. (2005). Adam and Billi (2006, 2007) considered optimal}
Since the Great Recession, advanced economies including Japan, the U.S. and the Euro economies had simultaneously fallen to zero interest rate, which stimulated studies to derive optimal monetary policy rule under the ZLB constraint. At present, however, there are no established models that can be easily implemented and estimated.

**Interactions between monetary and fiscal policies**: It is important to recognize the link between monetary and fiscal policies through (integrated) government budget constraint. If monetary policy changes inflation, government may change fiscal policy in anticipation of inflationary tax revenue changes. If fiscal policy changes, monetary authority may change the money growth rate so as to cancel the welfare loss from distortionary taxes.

Leeper (1991) tackled formally this topic using a simple monetary model with dynamic optimization. Consider the money-in-utility model where the real money demand is a decreasing function of nominal interest rate. Suppose the central bank makes nominal money supply constant (monetary policy is committed). In addition, suppose a positive government spending shock, which causes a rise of real interest rate. Then, nominal interest rate will also rise and the real money demand will go down. Because nominal money supply is constant, in order to lower the real money supply, price has to jump up. Thus, even though the nominal money supply is constant, inflation may be caused only by fiscal stimulus: In the case of monetary dominance in the Ricardian policy (or the central bank is “active” and government is “passive”),17 when fiscal policy influences real interest rate, the price level can no longer be determined independently of fiscal policy (fiscal theory of price level; FTPL).18 Leeper (1991) showed that if both monetary and fiscal authorities are “active”, then inflation and government debt paths become explosive. If both are “passive”, then price level becomes indeterminate.

Whether the central bank and the government are active or passive is being (although not much yet) examined empirically by the DSGE model in recent years. Davig and Leeper (2011) considered a possibility that active and passive regimes of monetary and fiscal authorities were changing, by using the new Keynesian model. When the fiscal authority is active, then the authority conducts not only an expenditure rule but also a lump-sum tax rule. Their examination strategy has two stages: At the first stage, they estimates monetary and fiscal policy rules considering the regime change (not parameters but rules themselves) and detected the periods during which the regime changed. Their result showed the price level were indeterminate in some periods (both policies were passive). At the second stage, substituting the estimated policy rules into the calibrated new Keynesian model, they investigated the difference of fiscal multipliers among passive or active policy rules. Although there are still rooms for improving estimation method yet, there is no doubt that empirical research considering the interactions of fiscal and monetary policies will continue to develop.

16Of course this topic has been discussed for a long time. See, e.g. “unpleasant monetarist arithmetic” (Sargent and Wallace, 1981).

17This corresponds to the normal setting of the standard DSGE model. The central bank committed to monetary policy rule and the government spending will be covered by lump-sum tax so as to meet the integrated government budget constraint.

18If the (integrated) government budget constraint is regarded as a “equation”, not as an “identity” with respect to price level, the fiscal policy might be an anchor to determine the price level. See also Sims (1994) and Woodford (1995). Uribe (2006) derived the sovereign debt risk endogenously from the integrated government budget constraint. He pointed out the trade-off between the inflation stabilization and the fiscal collapse: The possibility that the government’s real debt outstanding will expand if the central bank stabilizes inflation.
Introducing financial friction: The CEE and SW models do not implement important endogenous shock propagation mechanism due to asset price fluctuations. Normally, collateral constraints are imposed on financing. In this case, the deterioration of the collateral value by a negative asset price shock will raise the borrowing constraints and lower the investment. The standard DSGE models do not introduce this kind of endogenous shock amplification process through asset price fluctuations (financial accelerator mechanism).

Kiyotaki and Moore (1997) first introduced financial friction into the framework of general equilibrium where investor’s borrowing constraints (or collateral constraints) endogenously change according to asset prices. Carlstrom and Fuerst (1997) more explicitly depicted the friction by introducing asymmetric information between lenders and borrowers. In other words, they have provided rigorous microeconomic foundations for what kind of financial transaction environment the collateral constraint occurs.

Then, it is Bernanke, et al. (1999) that integrated these models: Kiyotaki and Moore (1997) considered land as the asset for collateral constraint. But the land volume cannot be adjusted according to the land price. Even if the land price goes down, we cannot lower the land volume, so the collateral value decline will be extremely amplified. In contrast, Carlstrom and Fuerst (1997) introduced financial friction in a setting that allows immediate adjustment of investment (i.e. without adjustment cost). In this case, the amplification mechanism through the asset price becomes extremely small. Bernanke, et al. (1999) modified the Carlstrom and Fuerst model by introducing an investment adjustment cost, which provided a realistic shock amplification mechanism through asset price. Bernanke, et al. (1999) has become one of the benchmarks of the DSGE model incorporating financial friction. Chapter 4 in this theses also employs this model in examining the sources of the Great Recession in the U.S. 19

Introducing search and matching friction: Since the so-called Shimer puzzle (Shimer, 2005), the DSGE models have been developed by expanding search and matching models based on Mortensen and Pissarides (1994) so as to reproduce the high volatilities of unemployment and job vacancies. In the models, the labor market friction that the vacancy does not immediately match with the unemployed is formulated as a matching function, and wage is assumed to be determined endogenously by the Nash bargaining between firms and workers. In recent years, empirical studies have been made to integrate the new Keynesian model with the search and matching model, estimate key parameters (bargaining power of workers and matching function parameters), and try to quantitatively grasp the effect of labor market friction shocks such as bargaining power shock, mismatch shock and so on.

Gertler, et al. (2008) is one of the earliest papers to estimate the model which integrate the SW model with search and matching model. They reformulated the SW model so that the opportunity of wage negotiation is to visit randomly, and the probability that the bargaining opportunity do not arrive corresponds to the Calvo parameter (nominal wage rigidity). In addition, firms are assumed to adjust employment not along the intensive margin (working hour) but along the extensive margin (the number of employee) to replicate high volatilities of unemployment and job vacancies. Their result showed the nominal wage rigidity helped to explain the large volatilities of unemployment: By introducing the wage rigidity, the wage responses to monetary policy shock became moderate while the unemployment responses became larger. They also found the bargaining power shock (a

rise of bargaining power of workers) has no effect for the business cycle (the main source of the business cycle was investment specific technology shock).\textsuperscript{20}

The significant worsening of the unemployment since the Great Recession would have triggered the model refinement of the labor market friction. An adverse financial shock, an unanticipated deterioration of asset price, will raise the firms’ borrowing constraints, which might bring an another channel of reducing employment. Christiano et al. (2011b) estimated a large-scale small open model on Swedish data introducing both financial friction and labor market friction at the same time. The result showed the financial shock (the entrepreneur wealth shock) did not so much affect unemployment variations.\textsuperscript{21}

Open economy model: Extending to an open economy model seems to be one of natural extensions. If we consider interactions of monetary policies among countries (e.g. currency war, policy coordination etc.), we should build a large country model. On the other hand, when home country can be regarded as a price taker in the international financial market, we should build a small country model.

Kollmann (2001) would be one of the first small open economy models, incorporating the Calvo type price and wage nominal rigidities. This calibrated DSGE model can successfully reproduce exchange rate overshooting in response to money supply shocks due to nominal rigidity: This model can quantitatively replicate the volatile variances of the nominal and the real exchange rates as compared with the model without nominal rigidities (corresponding to an open economy RBC model, e.g. Backus, et al. 1992, Schmitt-Grohe and Uribe, 2003).

Adolfson, et al. (2007) estimated, using the Euro area data, a large-scale small open economy model based on the SW model. This paper examined quantitatively the effect of the monetary policy considering a realistic channel through the incomplete exchange rate pass-through: Not only domestic firms, but both domestic importers and exporters are also monopolistic price setters facing the Calvo type nominal rigidities. The price is assumed to be set in the local currency. Under such circumstances, fluctuations in the exchange rate will not be immediately passed on to export goods prices or import goods prices due to nominal rigidities and will not be directly reflected in changes in the trade balance. In other words, even if the monetary easing policy of home country depreciates it’s own currency, the trade balance does not necessarily improve instantly. This large-scale small open economy model has become the prototype of the official DSGE model of Riksbank (Swedish central bank)\textsuperscript{22}

On the large country model, there has already been a canonical paper, Obstfeld and Rogoff (1995) with the new Keynesian characteristics: The dynamic optimization model under monopolistic competition and nominal rigidity.\textsuperscript{23} From the viewpoint of the empirical DSGE model, several papers empirically examines the sources of the business cycle through the channel of the terms-of-

\textsuperscript{20}See also Krause, et al. (2008), Sala, et al. (2008) and Lubik (2009).

\textsuperscript{21}See also Furlanetto and Groshenny (2016). They found, as with Gertler et al. (2008), the mismatch shock had no role for business cycle in normal time, but relatively higher role during the Great Recession.

\textsuperscript{22}See also Adolfson et al. (2011, 2014). Justiniano and Preston (2010) also found the monetary policy of the home country is not so effective for the real exchange rate in Canada. Christiano, et al. (2008) further extended Adolfson, et al. (2007)’s model incorporating financial friction, and examined the effect of the monetary easing policy by the ECB after 2000s. Lubik and Schorfheide (2007) examined a possibility that monetary authorities might respond not only to output gap, inflation gap but also exchange rate in Australia, Canada, New Zealand and the UK.

\textsuperscript{23}There have already been many studies that follow this literature. See, e.g. Betts and Devereux (2000), Corsetti and Pesenti (2001), Benigno and Benigno (2003). Engel (2002) provided surveys of the so-called “new open economy macroeconomics”.

trade and exchange rate.\footnote{See, for example, Kollmann (2013), Punnoose and Peersman (2013).}

**State dependent pricing:** The standard DSGE model usually assumes the Calvo type nominal rigidity where there is a probability that firms cannot revise prices, and this probability is treated as a time-constant structural parameter (Calvo parameter) regardless of the economic situation.

Fernandez-Villaverde and Rubio-Ramirez (2008) estimated the SW model allowing time-varying parameters, and they found (1) the Calvo parameter is not time-constant, and (2) the estimated time-varying Calvo parameter variations are countercyclical with aggregate output fluctuations. The second finding raises a big question in how to formulate the nominal rigidity of the current standard DSGE model.

Previous studies have also attempted to endogenize nominal rigidity, more specifically, to construct a model in which nominal rigidity changes endogenously in response to aggregate output fluctuations (e.g. Rotemberg and Saloner, 1986).

The intuition is straightforward: If the economy is bad, in order to secure profit, firms make high prices keep by collusion. As a result, price adjustment becomes sluggish (nominal rigidity rises in a recession). In contrast, if the economy is good, deviating from the collusion, lowering the price and taking a lot of demand will increase profit. As a consequence, price adjustment becomes flexible (nominal rigidity goes down in a boom). The current standard DSGE model, however, assumes no variations of the nominal rigidity, dealing it with a constant parameter. After all, the nominal rigidity is uncorrelated with aggregate output.\footnote{Unlike Rotemberg and Saloner (1986)’s oligopoly setting, the new Keynesian model assumes monopolistic competition, and there is no interaction with rival firms. Therefore, it is difficult to endogenize nominal rigidity with a tacit collusion. However, in monopolistic competition, there is a merit that it is easy to aggregate. Especially, in the Calvo type nominal rigidity, monopolistic competition makes aggregation very easy.}

Furthermore, following this mechanism, the markup (price over marginal cost) will also rise in a recession and reduce in a boom. Again, the standard DSGE model cannot replicate the countercyclical relationship, since markup variations are also handled exogenously as a shock (the so-called markup shock), resulting in no-correlation with markup and aggregate output. This problem will be considered later in Chapter 2.

In the context of the new Keynesian model, this kind of the model is called as a “state-dependent” pricing model (whereas a model with the time-constant nominal rigidity is called as a “time-dependent” pricing model). The reason why it is troublesome to solve the state-dependent pricing model is that the nominal rigidity changes depending on boom (the low nominal rigidity is profitable) or recession (the high nominal rigidity is profitable), so the firms must also set up the current price by anticipating future aggregate output.

However, this model seems to be important to consider the optimal monetary policy. To reduce the welfare loss due to the nominal rigidity, recession should be strongly avoided with bold monetary easing policy (In contrast, boom should be left because the nominal rigidity becomes low). In other words, an asymmetric optimal monetary policy may be derived depending on economic conditions.

A large number of authors tackle to build a state-dependent pricing model in a general equilibrium framework (See, e.g. Dotsey, et al. 1999, Golosov and Lucas, 2007, Gertler and Leahy, 2008). At the present time, however, we have not yet got an established state-dependent pricing model easily implemented and estimated.

**Integrating growth model with DSGE model:** The new Keynesian model can describe the long-term aggregate supply curve and the short-term aggregate supply curve simultaneously. The
CHAPTER 1. INTRODUCTION

former can be regarded as the steady state of aggregate output (output after price adjustment is over), and the latter corresponds to the output deviations from the steady state (output under price adjustment). Since the long-term output growth after price adjustment can be explained by (new) growth theory, it is a natural attempt to integrate the two models.

Comin and Gertler (2006) is one of the first attempts integrating the growth model with new Keynesian model, by endogenizing the number of variety of intermediate goods. As with Grossman and Helpman (1991)’s settings, the source of growth is to produce ideas of new intermediate goods by the R&D sector. Interestingly, allocating resources from the R&D sector to the production sector during the economic downturn will restore the output gap in the short term but will sacrifice the long-term output growth rate: A trade-off may arise between short-term economic recovery and long-term economic growth decline.

Recently, this model has been drawing attentions as a model explaining the so-called “slow recovery” or “secular stagnation” after the financial crisis in the U.S.\textsuperscript{26} Especially, when the DSGE model with financial friction and the growth model are integrated, the long-term and short-term trade-off will increase. Suppose the R&D sector has to procure external funds for the development activity, but borrowing is constrained by the collateral value due to imperfections in financial markets. Then, a negative shock to the collateral value of the R&D sector increases the borrowing constraints and causes substantial decline of growth. However, shifting workers from the R&D sector to the production sector will help lower the output gap. In the case of the central bank aiming at just stabilizing the output gap, aggregate demand creation through monetary easing might promote the resource allocation to the production sector. Thus, we face a severe trade-off between the short-term stabilization and the long-term growth decline. In other words, long-term nominal money non-neutrality may arise through the channel where monetary policy might induce resource allocation from the R&D sector to the production sector. This channel might be a source of the slow recovery or the secular stagnation in the U.S.

On the other hand, the integration of the two models may also be useful to reproduce countercyclical relationship of output and markup. As the number of intermediate goods firms increases (i.e. the lower market concentration by new entrants), it is difficult to collusion to keep prices high. That is, if the economy is good, the collusion is broken and the markup declines, and if the economy is bad, there is a possibility that the markup will rise as it becomes easy to collusion by exiting incumbents.

In any case, integrating DSGE model with growth model that endogenizes the number of firms (variety of goods) considering the firms’ entrance and exit behaviors is a situation just beginning.\textsuperscript{27}

**Effects of anticipated shock: News shock:** Normally, the DSGE model is estimated as “unexpected” structural shocks lead to the business cycle. However, there are cases where the shock can be expected. For example, news that a firm will constructs new factory or news that a patent has been acquired in new technology will bring about expectations that will increase future productivity. This anticipated shock (called as “news shock”) will influence current consumption and labor decisions through the dynamic optimizations of the agents. That is, the possibility that an expectation-driven business cycle might exist (often referred to as “Pigou cycle” or “animal spirits”).

Fujiwara et al. (2011) would be one of the first papers trying to grasp the news shock quantitatively in the standard DSGE model (CEE model).\textsuperscript{28} They examined the influence of anticipated


\textsuperscript{27}See, for example, Bilbiie et al. (2012) and Bilbiie, et al. (2014)

\textsuperscript{28}See Beaudry and Portier (2004) on the theoretical work of news shock. Beaudry and Portier (2006) found stock price reactions for anticipated TFP shock explained over 50% of business cycle in the U.S. by employing the structural
shocks of productivity on the business cycle (considered up to the fifth periods ahead news shocks from the viewpoint of fit for data). According to the variance decompositions, the impact of the anticipated shock on the business cycle cannot be neglected at all, and especially in inflation variations in the U.S., the contribution of the news on productivity was more than an unexpected productivity shock.\footnote{They also found employment reacts negatively against unexpected productivity shock (in line with Gali 1999) and employment shows positive reaction against expected productivity shocks (consistent with Christiano et al. 2003).}

Empirical studies of the quantitative effects of news shocks would continue to be examined. Especially, the effect of policy news seems to be important. Whether tax increases or changes in monetary policy rule accompanying the replacement of the chairman, there seems to be a high possibility that the market is currently reacting by incorporating future expectations before policies are implemented. Therefore, we should estimate the DSGE models by controlling the influence of the news shocks, then conduct policy simulations.

1.1.3 DSGE Models in Policy Institutions

Since the SW model that can explain the actual data variations in a consistent manner with the microeconomic theory, the DSGE model has been utilized as a useful tool of policy simulations and evaluations by policy institutions, mainly central banks in many countries:

**Bank of Japan**

Three official models have been developed: JEM (Japanese Economy Model) is a large-scale DSGE-VECM (vector error correction model) mixed type model (Fujiwara et al. 2005). Q-JEM (Quarterly JEM) is a large-scale hybrid type model incorporating a pure DSGE model into the core of the VECM model, (Ichigami et al., 2009). M-JEM is a fully pure estimated DSGE model referring to the official model developed by FED (Fueki et al. 2010).

**Federal Reserve Board** (FED)

There are two types official DSGE models: SIGMA (Erceg et al. 2005) is a large-scale calibrated DSGE model, and EDO model (Estimated, Dynamic, Optimization-based model) is an estimated medium-scale DSGE model where the potential output growth rate (technological progress rate) is estimated simultaneously (Edge et al. 2007).

**European Central Bank** (ECB)

Based on a calibrated open economy DSGE model called as NAWM (New Area Wide Model, Christoffel et al. 2007), NAWM Estimated Version is officially published (Christoffel et al. 2008).

**Bank of England** (BOE)

In addition to BEQM (Bank of England Quarterly Model, Harrison et al. 2005) which is the DSGE-VECM mixed model with core/non-core structure, BOE has developed the estimated small open economy DSGE model called as COMPASS (Burgess et al. 2013).

**Bank of Canada** (BOC)

By further developing an calibrated large-scale DSGE model called as ToTEM (Terms-of-Trade Economic Model; Murchison and Rennison 2006), BOC has recently updated it to ToTEM II (Dorich et al. 2013).

**Sveriges Riksbank** (Swedish central bank)

Riksbank has published the official estimated DSGE model called as RAMSES. Adolfson et al. (2007), the prototype of RAMSES, is an extension model of the SW model to the small open

\footnote{VAR model. Schmitt-Grohe and Uribe (2008) illustrated the way of installing the news shock into the DSGE model and examined the effect of news shock on U.S. data by the Bayesian technique.}
Further, Christiano et al. (2011b) has been extended to a large-scaled estimated DSGE model incorporating the financial accelerator mechanism in the financial market and the search and matching friction in the labor market.

**International Monetary Fund (IMF)**

Integrating GEM (Global Economy Model, Bayoumi et al. 2004) with GFM (Global Fiscal Model, Botman et al. 2006), GIMF (Global Integrated Monetary and Fiscal Model) is developed as a large-scale calibrated DSGE model (Kincaid 2008, Kumuhof and Laxton 2007). Also, the estimated DSGE model called as GPM (Global Projection Model, Carabenciov et al. 2008) is constructed in collaboration with CEPR (CEntre Pour la Recherche Economique et ses Applications; Center for economic research and its applications; French institution).

**Organisation for Economic Co-operation and Development (OECD)**

Cacciatorere and Fiori (2010) developed a calibrated DSGE model expanded to a small country open economy model under the incomplete international financial market, and incorporating firms’ entry and exit process and workers’ hired and fired process in a fashion of the Mortensen and Pissarides (1994). The model examines long and short term effects of structural reforms such as relaxation of entrance regulation of firms, employment protection, decline of labor market friction, and reduction of unemployment benefits (Economic Policy Committee, 2011, Working Party No.1 on Macroeconomic and Structural Policy Analysis).

**European Commission**

QUESTIII (Ratto et al. 2008) is a model of DG ECFIN (Directorate General for Economic and Financial Affairs) which can be regarded as the official DSGE model of the government (ECFIN) against the model of the central bank (ECB). Quest III is an extended SW (2003) model by adding liquidity constraint households and detailed fiscal policies, and examines the effect of various fiscal policies by government expenditure, various taxations and income transfers.

**French Ministry for the Economy and Finance (DGTPE, French Directorate General for the Treasury and Economic policy)**

Omega 3 (Carton and Guyon 2007) is a calibrated DSGE model expanded SW (2003) model to a three-country open model and add liquidity constraint households. The two countries are the Euro countries, and the remaining is another currency country. Under the adjustment of the real exchange rate by the incomplete international financial market, Omega 3 simulates the spillover effect to other countries of productivity shock and fiscal expenditure shock due to structural reform of one country in the Euro area.

**Spanish Presidential Economic Bureau (Modelo de Equilibrio Dinamico de la Economia Española)**

MEDEA (Burriel et al. 2010) is an estimated small open economy model with distortionary taxes (capital income tax, labor income tax and consumption tax) which provides simulation results of various fiscal policies.

**Italian Ministry of Economy and Finance (Department of the Treasury)**

A mixed model of the ITEM (Italian Treasury Economic Model) and QUEST III of the European Commission has been officially published (Annicchiarco et al. 2011).

### 1.2 Basic Theory and Estimation Method

This subsection illustrates basic theory and estimation method in the empirical DSGE approach using a simple new Keynesian model.

The basic theory provides three important equations: the new IS curve, the NKPC, and the
monetary policy rule. Essentially, the DSGE model is based on the optimization behavior of each agent in the model (households, firms, central banker, government, etc.). As a natural consequence, each agent will make current decisions taking future expectations into account.

The new IS curve is derived from household optimization behavior. Households smooth their consumption intertemporally by bond trading in respond to real interest rate. The real interest rate at the present period depends upon the inflation expectation at the next period. Thus, households become forward looking: A rise of future inflation anticipations will lower the current real interest rate. A decline in real interest rate will increase in current consumption through the intertemporal substitution effect. It is one of the most important features of the DSGE model that future expectations affect in the current decision makings.

Firms' optimizing pricing behaviors result in a short-term upward sloping aggregate supply curve called as the NKPC. Thus, nominal money is not neutral in the short term: The stabilization policy by the monetary authority will be effective. The NKPC is caused by introducing the nominal rigidity. On the other hand, in the long term when price adjustment is over, a vertical aggregate supply curve is also obtained: Nominal money is neutral in the long term, and the monetary policy is invalid.

The central bank is specified as the Taylor rule that determines nominal interest rate in response to output gap and inflation gap.

Next, a basic method for parameter estimation is summarized. The DSGE model is represented by a linear state space model. Specifying the probability density of the disturbance term called structural shock makes it possible to evaluate the likelihood. Usually, the density is specified as the normal distribution (a linear Gaussian state space model).

In recent research, parameters are estimated based on the Bayesian technique. This subsection also summarizes the likelihood evaluation using the Kalman filter and the method to estimate the posterior distributions of parameters called as the MH algorithm.

1.2.1 Simple DSGE Model

Final goods firms
The final goods firms produce a homogenous good by bundling differentiated intermediate goods \( i \in [0,1] \) and sell the final good to households. Final goods firms maximize profit by controlling purchase amount of intermediate goods \( y_t(i) \), given final goods price \( P_t \), intermediate goods price \( P_t(i) \) and aggregate demand \( y_t \).

\[
\max_{y_t(i)} P_t y_t - \int_0^1 P_t(i) y_t(i) di \\
\text{s.t.} \\
y_t = \left[ \int_0^1 y_t(i) \frac{1 - \lambda}{1 + \lambda} di \right]^{1 + \lambda}, \quad \lambda > 0
\]  

where (1.2) is the Dixit-Stiglitz type production function of the final goods. The profit maximization brings the demand function for an intermediate good \( i \):

\[
y_t(i) = \left( \frac{P_t}{P_t(i)} \right)^{\frac{1 - \lambda}{\lambda}} y_t
\]  

From the result above, intermediate goods firms face downward-sloping demand curve, which implies they have some market power. Also, we can see \( |\partial \ln y_t(i)/\partial \ln P_t(i)| = (1 + \lambda)/\lambda \), thus, \((1 + \lambda)/\lambda\) indicates the price elasticity of intermediate goods demand.
Substituting (1.3) into the production function (1.2), we can obtain:

$$P_t = \left[ \int_0^1 P_t(i)^{-\frac{1}{\lambda}} di \right]^{-\lambda}$$  \hspace{1cm} (1.4)

This is the general price level of this economy. Substituting (1.3) and (1.4) into the objective function (1.1) can confirm the zero profit condition of the final goods firms:

$$\int_0^1 P_t(i)y_t(i) di = P_t y_t$$  \hspace{1cm} (1.5)

**Intermediate goods firms**

The intermediate good firm $i \in [0, 1]$ monopolistically produces a differentiated good $i$ by employing labor and sells it to the final goods firms. The intermediate good firm $i$ faces the following cost minimization problem:

$$\min_{l_t(i)} w_t l_t(i)$$  \hspace{1cm} (1.6)

s.t. $y_t(i) = z_t l_t(i)$  \hspace{1cm} (1.7)

given $w_t$  \hspace{1cm} (1.8)

(1.7) is the production function of intermediate good $i$. $l_t$ is employment, $z_t$ is productivity and $w_t$ is real wage. Let the Lagrange multiplier denoted as $\Psi_t(i)$, then the cost minimization condition yields:

$$\Psi_t(i) = \frac{w_t}{z_t}$$  \hspace{1cm} (1.9)

Thus, we can see $\Psi_t(i)$ corresponds to the real marginal cost of the intermediate good firm $i$.

Next, we consider optimal price settings of intermediate goods firms. Every period, intermediate goods firms shall face a probability, $\xi$, with which they cannot revise their prices (Calvo type nominal rigidity). Also, if prices cannot be optimized, intermediate goods firms have contracted with final goods firms to slide their prices according to the previous inflation (the lagged inflation indexation contract). Then, if the price set at period $t$ is not revised up to $s$ period ahead, we have:

$$P_{t+s}(i) = P_{t+s-1}(i) \pi_{t+s-1} = P_t(i) \prod_{k=1}^{s} \frac{\pi_{t+k-1}}{R_{t+k}} = P_t(i) X_{t,s}$$  \hspace{1cm} (1.10)

where $\pi_t = \frac{P_t}{P_{t-1}}$ is the gross inflation rate. Therefore, the objective function of the intermediate good firm can be represented as follows:

$$\max_{P_t(i)} \mathbb{E}_t \sum_{s=0}^{\infty} \xi^s \left( \prod_{\tau=0}^{s-1} \frac{\pi_{t+\tau+1}}{R_{t+\tau}} \right) \left[ \frac{P_{t+s}(i)}{P_{t+s}} - \Psi_{t+s} \right] y_{t+s}(i)$$  \hspace{1cm} (1.11)

$$\implies \max_{P_t(i)} \mathbb{E}_t \sum_{s=0}^{\infty} \xi^s \left( \prod_{\tau=0}^{s-1} \frac{\pi_{t+\tau+1}}{R_{t+\tau}} \right) \left[ \frac{P_t(i) X_{t,s}}{P_{t+s}} - \Psi_{t+s} \right] \left( \frac{P_{t+s}}{P_t(i) X_{t,s}} \right)^{\frac{1+\lambda}{\lambda}} y_{t+s}$$  \hspace{1cm} (1.12)

where $R_t$ is the gross nominal interest rate, thus $\mathbb{E}_t \frac{R_t}{\pi_{t+1}}$ indicates the (ex-ante) gross real interest rate. The firm discounts the future profit stream with the real interest rate (since the firm can
In case of flexible pricing (ξ), the firm decides today’s price so as to maximize the discounted present value of the future profit stream obtained when the price set today cannot be revised permanently. Notice the term ξ in the objective function: The nominal rigidity affects the firm’s decision making. The firm decides today’s price so as to maximize the discounted present value of the future profit stream.

The optimal condition for price setting of the intermediate good firm can be derived as:

\[ P_t(i)^* = (1 + \lambda) \frac{E_t \sum_{s=0}^{\infty} \xi^s \left( \prod_{t=0}^{s-1} \frac{\pi_{t+s} + 1}{R_{t+s}} \right) \Psi_{t+s} y_{t+s}(i)^*}{E_t \sum_{s=0}^{\infty} \xi^s \left( \prod_{t=0}^{s-1} \frac{\pi_{t+s} + 1}{R_{t+s}} \right) \frac{N_{t+s}}{R_{t+s}} y_{t+s}(i)^*} \]  

(1.13)

In case of flexible pricing (ξ = 0), we have:

\[ \frac{P_t(i)^*}{P_t} = (1 + \lambda) \Psi_t \]  

(1.14)

We can see \( \lambda \) corresponds to (net) markup rate which is inversely proportional to the price elasticity of demand, \( 1 + \frac{1}{\lambda} \). This is the same result as the standard monopoly firm’s pricing.

**Households**

The household \( j \in [0, 1] \) maximizes utility controlling consumption \( c_t(j) \), real money holdings \( m_t(j)(\equiv \frac{M_t}{P_t}) \), real bond holdings \( b_t(j)(\equiv \frac{B_t}{P_t}) \), and labor supply \( l_t(j) \) under the budget constraint and the transversality condition.

\[
\max_{c_t(j), m_t(j), b_t(j), l_t(j)} \left\{ \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s u(c_{t+s}(j), m_{t+s}(j), l_{t+s}(j)) \right\}
\]

s.t. \( c_t(j) + m_t(j) + b_t(j) = w_t l_t(j) + \frac{m_{t-1}(j)}{\pi_t} + \frac{b_{t-1}(j)}{\pi_t} + t_t + \psi_t \)

where \( t_t \) is real lump-sum transfer from government (\( t_t < 0 \) is real lump-sum tax levied by government) and \( \psi_t \) is the real dividend (real profits) from intermediate goods firms. The left-hand side is the expenditure and the right-hand side is the revenue. Solving the budget constraint forward, the transversality condition can be obtained:

\[
\lim_{s \to \infty} \prod_{t=0}^{s-1} \frac{\pi_{t+\tau} + 1}{R_{t+\tau}} \Gamma_{t+s}(j) = 0 \text{ where } \Gamma_{t}(j) \in \{ m_t(j), b_t(j) \}
\]

Thus, the transversality condition imposes the zero discounted present values of real money holdings and real bond holdings at the last period. Here, the utility function is specified as follows:

\[
u(c_t(j), m_t(j), l_t(j)) = \frac{1}{1 - \sigma} c_t(j) - h c_{t-1}(j) + \frac{1}{1 - \sigma_m} m_t(j) - \frac{1}{1 + \sigma L} l_t(j) \]

where \( \sigma(> 0) \) and \( \sigma_m(> 0) \) are CRRA parameters. Since households are risk averse, they dislike fluctuations of consumption and real money holdings, which justifies the stabilization policy by government. The third term represents labor disutility where \( \sigma_L(> 0) \) is the labor disutility parameter. Also, we introduce a real rigidity called as the external habit formation of consumption, which can replicate the hump-shaped consumption response against structural shocks. \( c_t = \int_0^1 c_t(j) \, dj \) depicts the average consumption and the parameter \( h \in [0, 1] \) represents the extent of habit formation. So,
CHAPTER 1. INTRODUCTION

The household is assumed to gain the utility on how much he is consuming from the average consumption level. The first order necessary conditions (FOCs) can be derived as follows:

The Euler equation of consumption:

\[
(c_t(j) - h c_{t-1})^{-\sigma} = \beta E_t \frac{R_t}{\pi_t+1} (c_{t+1}(j) - h c_t)^{-\sigma}
\] (1.15)

The household decides the current consumption so that the current marginal utility of consumption is equalized to the anticipated discounted present value of the marginal utility of consumption at the next period: The intertemporal consumption smoothing condition.

The money demand function:

\[
m_t(j)^{-\sigma_m} = \frac{R_t - 1}{R_t} (c_t(j) - h c_{t-1})^{-\sigma}
\] (1.16)

The real money demand has a common feature with the conventional LM curve: A decreasing function of the nominal interest rate \(R_t\) and an increasing function of an aggregate demand component, \(c_t\).

The labor supply function:

\[
w_t = \frac{l_t(j)^{\sigma_L}}{(c_t(j) - h c_{t-1})^{-\sigma}}
\] (1.17)

The household supplies labor so that the MRS between consumption and labor supply will be equalized to the real wage. We can also see the parameter \(\sigma_L\) indicates the reciprocal of the wage elasticity of labor supply, since \(\frac{\partial \ln l_t(j)}{\ln w_t} = \frac{1}{\sigma_L}\) (Inverse Frisch elasticity).

Government

The policy authority will be assumed to control the nominal interest rate according to output gap and inflation gap (Taylor rule):

\[
\frac{R_t}{R} = (\frac{\pi_t}{\pi})^{\psi_1} (\frac{y_t}{y})^{\psi_2} e_t^R
\] (1.18)

where \(R, y, \pi\) are steady states of gross nominal interest rate, output and inflation, respectively; Output gap and inflation gap are defined as the deviations from the steady states. \(e_t^R (\geq 0)\) is the monetary policy shock which represents an unanticipated deviation from the policy rule. The government budget constraint can be written as follows:

\[
\int_0^1 m_t(j) dj + \int_0^1 b_t(j) dj = \int_0^1 \frac{m_{t-1}(j)}{\pi_t} dj + R_{t-1} \int_0^1 \frac{b_{t-1}(j)}{\pi_t} dj + t_t
\] (1.19)

The left-hand side is the revenue and the right-hand side is the expenditure. The policy authority shall supply money and bonds according to private money and bond demand. In the case of a negative seigniorage or an increase in the outstanding government debt, it shall be covered by a lump-sum tax.
1.2. BASIC THEORY AND ESTIMATION METHOD

Aggregation

Aggregate households’ FOCs can be summarized as follows:

\[(c_t - h c_{t-1})^{-\sigma} = \beta E_t \frac{R_t}{\pi_{t+1}} (c_{t+1} - h c_t)^{-\sigma}\]  \hspace{1cm} (1.20)

\[m_t^{-\sigma_m} = \frac{R_t - 1}{R_t} (c_t - h c_{t-1})^{-\sigma}\]  \hspace{1cm} (1.21)

\[w_t = \frac{l_t^{\sigma_v}}{(c_t - h c_{t-1})^{-\sigma}}\]  \hspace{1cm} (1.22)

Since the intermediate goods firms who can revise prices have common optimal pricing conditions, we can aggregate the optimal price as follows:

\[P_t^* = (1 + \lambda) E_t \sum_{s=0}^{\infty} \xi^s \left( \prod_{\tau=0}^{s-1} \frac{\pi_{t+r}}{\pi_{t+r+1}} \right) \Psi_{t+s} y_{t+s}^*\]  \hspace{1cm} (1.23)

\[y_t^* = \left( \frac{P_t}{P_t^*} \right)^{\frac{1+\lambda}{\lambda}} y_t\]

Since firms that cannot revise prices index their prices associated with the previous inflation, and firms that can revise prices set prices according to the above optimal condition, the general price level (1.4) can be written as:

\[P_t^{-\frac{1}{\lambda}} = \xi (P_{t-1} \pi_{t-1})^{-\frac{1}{\lambda}} + (1 - \xi) P_t^{*-\frac{1}{\lambda}}\]  \hspace{1cm} (1.24)

Next, we consider the aggregate supply. Aggregating the production function (1.7) with respect to \(i\), we have \(\int_0^1 y_t(i) di = z_t \int_0^1 l_t(i) di\). Then, substituting the demand for intermediate goods (1.3) and the labor market clearing condition (\(\int_0^1 l_t(i) di = l_t\)) yields the aggregate supply curve:

\[y_t = \frac{z_t l_t}{v_t}\]  \hspace{1cm} (1.25)

where \(v_t \equiv \int_0^1 \left( \frac{P_t}{P_t^*(i)} \right)^{\frac{1+\lambda}{\lambda}} di\)  \hspace{1cm} (1.26)

where \(v_t\) is the price dispersion that represents the relative price distortion. Since \(v_t > 1\), the result above tells that more labor input is needed to cover the aggregate demand: Due to the Calvo-type nominal rigidity, two types of prices are mixed: Optimized prices and not optimized prices. This relative price distortion also distorts the demand for each intermediate good, and makes resource allocation inefficient.

Next, we consider the aggregate demand. Aggregating profits of intermediate goods firms, we can derive:

\[\int_0^1 P_t(i) y_t(i) di - W_t \int_0^1 l_t(i) di = P_t y_t - W_t l_t \implies \psi_t = y_t - w_t l_t\]  \hspace{1cm} (1.27)

At the first equality we use the zero profit condition of the final goods firms (1.5) and the labor market clearing condition. The total real profit is \(\psi_t = \int_0^1 \psi_t(i) di\). Aggregating households budget
constraints, we can derive:

\begin{align*}
&\int_0^1 c_t(j) dj + \int_0^1 m_t(j) dj + \int_0^1 b_t(j) dj \\
&= w_t \int_0^1 l_t(j) dj + \int_0^1 \frac{m_{t-1}(j)}{\pi_t} dj + R_{t-1} \int_0^1 \frac{b_{t-1}(j)}{\pi_t} dj + t_t \int_0^1 dj + \psi_t \int_0^1 dj \\
\iff& \int_0^1 c_t(j) dj = w_t \int_0^1 l_t(j) dj + \psi_t \\
\iff& c_t = w_t l_t + \psi_t \\
\iff& c_t = y_t
\end{align*}

(1.28) expresses the aggregate demand of the goods market. The first equivalence is derived from substituting the government budget constraint (1.19) into the aggregate household budget constraint. The second equivalence comes from the fact that consumption and labor input do not rely on the subscript \( j \). At the last equivalence, we used the aggregate real profit of the intermediate goods firms (1.27).

Finally, from (1.25) and (1.28), the goods market clearing condition can be obtained as the following equation:

\[ c_t = z_t l_t \]

(1.29)

**Structural shocks**

The DSGE model assumes that exogenous structural shocks cause the business cycle which corresponds to deviations from the steady states of endogenous variables. In this model, we introduce the productivity shock \( z_t \) and the monetary policy shock \( \epsilon^R_t \). These two structural shocks (with mean unity) are assumed to follow the AR(1) processes:

\begin{align*}
\ln z_t &= \rho^Z \ln z_{t-1} + \epsilon^Z_t \\
\ln \epsilon^R_t &= \rho^R \ln \epsilon^R_{t-1} + \epsilon^R_t
\end{align*}

where \( \epsilon^Z_t \) and \( \epsilon^R_t \) are i.i.d. shocks with mean zeros. AR(1) parameters, \( \rho^Z \) and \( \rho^R \), express the inertia of the shocks, which show how long the shocks will prolong.

**Steady states (long-term equilibrium conditions)**

Suppose that there is no technological progress for the sake of simplicity (i.e. \( z_t = z_{t-1} = z \)). Then, the steady states of endogenous variables are calculated as follows:

Using the Euler equation (1.20), the long-run real interest rate is determined by the subjective discounted factor:

\[ \frac{R}{\pi} = \frac{1}{\beta} \]

Letting the net nominal interest rate and net inflation rate denoted as \( R^{net} \) and \( \pi^{net} \), then \( R = 1 + R^{net} \) and \( \pi = 1 + \pi^{net} \). Taking logs for both sides in the equation, \( \ln(1+R^{net}) - \ln(1+\pi^{net}) = -\ln \beta \). Using the first order approximation, we have \( \ln(1+R^{net}) \approx R^{net} \) and \( \ln(1+\pi^{net}) \approx \pi^{net} \). With time-differentiating both sides for the equation \( R^{net} - \pi^{net} = -\ln \beta \), we can derive \( R^{net} = \pi^{net} \). Therefore, we can see the long-term nominal interest rate changes exactly the same as the long-term inflation changes. It should be also noted that the monetary policy rule (1.18) has no role of determining the long-run nominal interest rate.
1.2. BASIC THEORY AND ESTIMATION METHOD

The long-run marginal cost is derived as $\Psi = \frac{1}{1+\lambda}$ from (1.23). Combining the result with (1.9), we can derive the long-run real wage as follows:

$$w = \frac{1}{1+\lambda}z$$  \hspace{1cm} (1.33)

Substituting the result above into the labor supply function (1.22) and using the long-run goods market clearing condition $c = y$ (see (1.28)), we can obtain the long-run aggregate supply curve:

$$y = \left[ \frac{z}{(1+\lambda)(1-h)^{\sigma}} \right]^{\frac{1}{\sigma+\sigma L}}$$  \hspace{1cm} (1.34)

Thus, we can confirm the long-term money neutrality: The long-run aggregate supply curve becomes vertical in $(y, \pi)$ plane, since the steady state of output $y$ does not depend upon the steady state of inflation $\pi$. We can also find the steady state of productivity level $z$ affects the long-run output level. In the case of introducing technological progress, that is, when allowing the growth of $z$, the long-term aggregate supply curve shifts to the right. Substituting the long-run output into the production function (1.7), we can derive the long-run employment level as $l = y/z$.

It should be noted, when introducing technological progress actually, we should specify utility function taking care of the consistency with the balanced growth model. If population growth is not assumed (that is normal assumption in both growth and business cycle models), the utility function should be specified as the log-utility (i.e. $\sigma = 1$): Let the growth rate of technological progress denoted as $\mu_z \equiv \dot{z}$. If we assume $\dot{l} = 0$, then we can see $\dot{y} = \dot{z} = \mu_z$ from the production function (1.7). Substituting this result into the labor supply function (1.22), we can derive $\dot{w} = \sigma \dot{y} = \sigma \mu_z$. Hence, the log-utility ($\sigma = 1$) is needed for the consistency with the balanced growth model.

Also, suppose that both log-utility specifications on consumption and real money holdings (i.e. $\sigma = \sigma_m = 1$, and $h = 0$. Then, from (1.21) and (1.28), the long-term real money demand function can be expressed by:

$$m = \frac{M}{P} = \frac{R}{R-1}y$$

The equation above depicts the long-term aggregate demand curve where government can control the growth rate of nominal money balance defined as $\mu_M = \dot{M}$.

Let us focus on $\dot{\pi}^{net} = 0$ (which implies $\dot{R}^{net} = 0$) and evaluate the long-term inflation $\pi^{net}$. Taking logs for both side of the equation above, time-differentiating, and imposing $\dot{\pi}^{net} = 0$, we can obtain:

$$\pi^{net} = \mu_M - \mu_z$$  \hspace{1cm} (1.35)

The intuition is straightforward: The long-term aggregate supply curve will keep shifting to the right at the growth rate of $\mu_z$, since $y = z/(1+\lambda)$. So, the long-term aggregate demand curve cannot keep constant inflation unless it continues to shift to the right. Hence, if government is to grow the nominal money balance with $\mu_M$, the long-term (net) inflation rate will become the difference between nominal money growth rate and technological progress growth rate. In particular, if government makes the nominal price level keep constant, then the nominal money balance should be supplied at the growth rate equal to the technological progress growth rate, i.e. $\mu_M = \mu_z$.

In sum, if government alter the long-run nominal money balance growth rate, the long-run inflation rate is changed. The change in long-run inflation rate dictates the change in nominal interest rate. But changes in the long-run nominal variables do not influence the long-run real variables.
such as output, consumption, employment, real wage and real interest rate: The long-term money neutrality.

Log-linearized model (short-term equilibrium conditions)

Let us suppose that both the technology progress growth rate and the nominal money growth rate are zero (i.e. \( \mu_s = \mu_M = 0 \)) and make the model log-linearized around steady states. The linear approximation of (1.20) around steady state yields:

\[
\tilde{c}_t = \frac{1}{1 + h} E_t \hat{c}_{t+1} + \frac{h}{1 + h} \hat{c}_{t-1} - \frac{1 - h}{(1 + h)^{\sigma}} \left( \tilde{R}_t - E_t \hat{\pi}_{t+1} \right) \quad (1.36)
\]

Substituting \( \hat{y}_t = \hat{c}_t \) derived from log-linearizing (1.28) into (1.36), we can obtain the so-called new IS equation:

\[
\hat{y}_t = \frac{1}{1 + h} E_t \hat{y}_{t+1} + \frac{h}{1 + h} \hat{y}_{t-1} - \frac{1 - h}{(1 + h)^{\sigma}} \left( \tilde{R}_t - E_t \hat{\pi}_{t+1} \right) \quad (1.37)
\]

The reason the equation above named as the new IS curve comes from the fact that output gap \( \hat{y}_t \) is a decreasing function of the (ex-ante) real interest rate \( (\tilde{R}_t - E_t \hat{\pi}_{t+1}) \). This implies if households anticipate future inflation, then the real interest rate declines, which stimulates the current consumption because of the intertemporal substitution effect. Although the forward looking term, \( E_t \hat{y}_{t+1} \), is a natural consequence of the dynamic optimization, the output gap inertia \( \hat{y}_{t-1} \) also occurs due to the habit formation on consumption. If \( h = 0 \) (no habit), then the backward term of output gap disappears.

Next, we will derive the NKPC. Letting the optimal real price denoted as \( Q_t = \left( \frac{P_t}{n_t} \right) \), we can write (1.24) as the following equation by dividing both sides by \( P_t \):

\[
1 = \xi \left( \frac{\hat{\pi}_{t-1}}{\hat{\pi}_t} \right)^{-1/\lambda} + (1 - \xi)Q_t^{-1/\lambda}
\]

Since \( \pi = 1 \) and \( Q = 1 \) in steady states, log-linearizing the above, we have:

\[
\hat{Q}_t = \frac{\xi}{1 - \xi} (\hat{\pi}_t - \hat{\pi}_{t-1}) \quad (1.38)
\]

The previous inflation \( \hat{\pi}_{t-1} \) remains by the lagged inflation indexation contract.\(^{30}\) Also, dividing both sides of (1.23) by \( P_t \), using \( \frac{P_t X_t}{P_t} = \frac{\pi_t}{\pi_{t+s}} \) and the Euler equation (1.20), we derive:

\[
\prod_{\tau=0}^{s-1} \frac{\hat{\pi}_{t+\tau+1}}{R_{t+\tau}} = \beta^s \frac{\Xi_{t+s}}{\Xi_t} \quad (1.39)
\]

\[
\Xi_t = (c_t - hc_{t-1})^{-\sigma} \quad (1.40)
\]

where \( \beta^s \frac{\Xi_{t+s}}{\Xi_t} \) is the stochastic discount factor (SDF) and \( \Xi_t \) is the marginal utility of consumption. We can write the optimal pricing condition (1.23) as the following expression, using \( Q_t \) and SDF:

\[
Q_t E_t \sum_{s=0}^{\infty} (\xi \beta)^s \frac{\Xi_{t+s}}{\Xi_t} \frac{\pi_t}{\pi_{t+s}} y_{t+s}^* = (1 + \lambda) \left[ E_t \sum_{s=0}^{\infty} (\xi \beta)^s \frac{\Xi_{t+s}}{\Xi_t} \frac{w_{t+s}}{z_{t+s}} y_{t+s}^* \right]
\]

\(^{30}\)Assuming the partial lagged inflation indexation contract, \( P_{t+s} = P_{t+s-1} (X_{t+s})^\chi \) where \( \chi \in (0, 1) \), the linear-approximated optimal real price is \( \hat{Q}_t = \frac{\xi}{1 - \chi \hat{\pi}_{t-1}} \). At this time, if the price does not index to the lagged inflation at all and remains as the same level at the previous period \( (\chi = 0) \), we can see \( \hat{Q}_t = \frac{\xi}{1 - \chi} \hat{\pi}_t \), and the backward term disappears. As a result, it becomes inconsistent with the empirical findings of Furher (1997) and Gali and Gertler (1999) which demonstrated the importance of the backward term of inflation in the NKPC. Here, the perfect lagged inflation indexation is assumed, so \( \chi = 1 \).
Linear-approximating the above, we have:

\[ \hat{Q}_t - \xi \beta E_t \hat{Q}_{t+1} = \xi \beta (E_t \hat{\pi}_{t+1} - \hat{\pi}_t) + (1 - \xi \beta)(\hat{w}_t - \hat{z}_t) \]  

(1.41)

Substituting (1.38) into (1.41) and eliminating \( \hat{Q}_t \) and \( E_t \hat{\pi}_{t+1} \), we can obtain the short-term aggregate supply curve (NKPC):

\[ \hat{\pi}_t = \frac{\beta}{1 + \beta} E_t \hat{\pi}_{t+1} + \frac{1}{1 + \beta} \hat{\pi}_{t-1} + \frac{(1 - \xi)(1 - \xi \beta)}{\xi(1 + \beta)}(\hat{w}_t - \hat{z}_t) \]  

(1.42)

Since the firms decide their price based on the dynamic optimization, the current inflation is affected by the expected future inflation \( E_t \hat{\pi}_{t+1} \). But the NKPC also includes the inflation inertia \( \hat{\pi}_{t-1} \), which is caused by the lagged inflation indexation contract.

To make the economic meanings of the NKPC easier to interpret, we should eliminate the real wage \( \hat{w}_t \) and employment \( \hat{l}_t \) from the equation above, using the following first-order approximated labor supply function (1.43) and production function (1.44):

\[ \hat{w}_t = \sigma L \hat{l}_t + \sigma_1 - h(\hat{y}_t - h \hat{y}_{t-1}) \]  

(1.43)

\[ \hat{y}_t = \hat{z}_t + \hat{l}_t \]  

(1.44)

Then, we can derive another expression of the NKPC:

\[ \hat{\pi}_t = \frac{\beta}{1 + \beta} E_t \hat{\pi}_{t+1} + \frac{1}{1 + \beta} \hat{\pi}_{t-1} + \kappa \left[ \left( \sigma L + \frac{\sigma}{1 - h} \right) \hat{y}_t - \frac{\sigma h}{1 - h} \hat{y}_{t-1} - (1 + \sigma L) \hat{z}_t \right] \]  

(1.45)

where the parameter, \( \kappa \), related to the slope of the NKPC, is written as a function of the Calvo parameter, \( \xi \):

\[ \kappa = \frac{(1 - \xi)(1 - \xi \beta)}{\xi(1 + \beta)} \]  

(1.46)

If \( \xi \in (0, 1) \), thus, the nominal rigidity exists, then \( \kappa > 0 \) and the slope of the NKPC (the coefficient of \( \hat{y}_t \)) has a finite positive value (recall that \( h \in [0, 1) \), \( \sigma > 0 \) and \( \sigma L > 0 \)). Thus, the short-term upward-sloping aggregate supply curve in \( (\hat{y}_t, \hat{\pi}_t) \) plane, which indicates the short-term money non-neutrality due to the nominal rigidity. In other words, the policy authority can increase output with the sacrifice of inflation (the short-term trade-off between output gap and inflation gap). Moreover, the NKPC will become flatter as nominal rigidity increases, since \( \frac{\partial \kappa}{\partial \xi} < 0 \): The non-neutrality of money increases if price becomes more sticky.

Also, we can confirm the NKPC will become vertical without nominal rigidity. If \( \xi \rightarrow 0 \), then \( \kappa \rightarrow \infty \), which leads to the vertical aggregate supply curve: The long-term money neutrality without nominal rigidity. The monetary policy is ineffective if price adjustment is over (in the long-term) or flexible (the RBC model).

The productivity shock, \( \hat{z}_t \), is a major source of inflation fluctuations. Since the coefficient for \( \hat{z}_t \) is negative, when the productivity rises unexpectedly, the aggregate supply curve is shifted down.

Let us move on the dynamics of log-linearized price dispersion. (1.26) can be rewritten as follows, using the price indexed with the lagged inflation and the optimal real price \( Q_t \).

\[ v_t = \int_0^1 \left( \frac{P_{t(i)}}{P_t} \right)^{-\frac{1+\lambda}{\lambda}} di = \xi \left( \frac{\pi_{t-1}}{\pi_t} \right)^{-\frac{1+\lambda}{\lambda}} v_{t-1} + (1 - \xi)Q_t^{-\frac{1+\lambda}{\lambda}} \]  

(1.47)
Log-linearizing the above, we have:
\[
\dot{v}_t = \xi \dot{v}_{t-1} - \xi \frac{1 + \lambda}{\lambda} (\hat{\pi}_{t-1} - \hat{\pi}_t) - (1 - \xi) \frac{1 + \lambda}{\lambda} \dot{Q}_t
\]
Substituting the result of log-linearized optimal real price (1.38) into the above, we can obtain:
\[
\dot{v}_t = \xi \dot{v}_{t-1}
\]  
(1.48)
Suppose the steady state gross inflation is \(\pi = 1\) and the price level at the initial period is on the steady state value, \(P\). Then, \(\dot{v}_t = 0\): The price dispersion does not have no influence on the system of linear difference equations. \(^{31}\)

Finally, log-linearizing the equations (1.18) and (1.21), related to the monetary authority, we can derive:
\[
\hat{R}_t = \psi_1 \hat{\pi}_t + \psi_2 \hat{y}_t + \ln \epsilon_t^R
\]  
(1.49)
\[
-\sigma_m \hat{m}_t = \frac{\beta}{1 - \beta} \hat{R}_t - \frac{\sigma}{1 - h} (\hat{y}_t - h \hat{y}_{t-1})
\]  
(1.50)
(1.49) is the log-linearized monetary policy rule and (1.50) is the log-linearized money demand function. It should be noted how much the nominal money balance by government is to be supplied so as to match with the monetary policy rule is determined from (1.50). This implies the money demand function becomes an auxiliary equation in this system of the simultaneous equations. \(^{32}\)

In summary, we obtain the three fundamental equations in a simple new Keynesian model: The new IS curve, the NKPC and the monetary policy rule. The endogenous variables are output gap \(\hat{y}_t\), inflation gap \(\hat{\pi}_t\) and nominal interest rate \(\hat{R}_t\). And there are two structural shocks, the sources of the business cycle: The productivity shock, \(\hat{z}_t\), and the monetary policy shock, \(\epsilon_t^R\).

\[\dot{y}_t = \frac{1}{1 + h} \hat{y}_{t+1} + \frac{h}{1 + h} \hat{y}_{t-1} - \frac{1 - h}{(1 + h)\sigma} \left(\hat{R}_t - E_t \hat{\pi}_{t+1}\right)\]
(1.51)
\[\hat{\pi}_t = \frac{\beta}{1 + \beta} E_t \hat{\pi}_{t+1} + \frac{1}{1 + \beta} \hat{\pi}_{t-1} + \frac{(1 - \xi)(1 - \xi\beta)}{\xi(1 + \beta)} \left[\sigma_L (\hat{y}_t - \hat{z}_t) + \frac{\sigma}{1 - h} (\hat{y}_t - h \hat{y}_{t-1}) - \hat{z}_t\right]\]
(1.52)
\[\hat{R}_t = \psi_1 \hat{\pi}_t + \psi_2 \hat{y}_t + \epsilon_t^R\]
(1.53)
\[\hat{z}_t = \rho^Z \hat{z}_{t-1} + \epsilon_t^Z\]
(1.54)
\[\epsilon_t^R = \rho^R \epsilon_{t-1}^R + \epsilon_t^R\]
(1.55)

1.2.2 Solving the DSGE model

Up to this point, the FOCs and the resource constraints are log-linearized around steady states and the model is described by a system of linear difference equations. Furthermore, substituting (1.53) into (1.51), we can derive the aggregate demand curve. Define \(k_t \equiv [\hat{Z}_t, \epsilon_t^R]'\), \(x_{t+1} \equiv [E_t \hat{Y}_{t+1}, E_t \hat{\pi}_{t+1}, \hat{Y}_t, \hat{\pi}_t]'\), \(\epsilon_t \equiv [\epsilon_t^Z, \epsilon_t^R]'\), \(\theta = [h, \sigma, \beta, \xi, \sigma_L, \psi_1, \psi_2, \rho^Z, \rho^R]\). Then, we can rewrite the difference equations as a “structural form”:
\[
A(\theta) \begin{bmatrix} k_t \\ x_{t+1} \end{bmatrix} = B(\theta) \begin{bmatrix} k_{t-1} \\ x_t \end{bmatrix} + C(\theta) \epsilon_t
\]  
(1.56)

\(^{31}\)If the steady state of gross inflation \(\pi > 1\) (allowing a nominal money growth) and the partial inflation indexation is assumed, the dynamics of the price dispersion cannot be ignored. See Ascari (2004).

\(^{32}\)It is Romer (2000) that pointed out this fact at the very beginning.
where $k_{t-1}$ is the vector of the predetermined variables, $x_t$ is the vector of the control variables and \( \varepsilon_t \) is the vector of the structural shocks. As an example of the model here, $A(\theta)$ and $B(\theta)$ are $6 \times 6$ matrix, and $C(\theta)$ is $6 \times 2$ matrix.

We solve this simultaneous difference equation, using eigenvalue decomposition, $A^{-1}B = Q^{-1}\Lambda Q$:

\[
\begin{bmatrix}
  k_t \\
x_{t+1}
\end{bmatrix}
= A^{-1}B 
\begin{bmatrix}
  k_{t-1} \\
x_t
\end{bmatrix}
+ A^{-1}C\varepsilon_t
\]

\[= Q \begin{bmatrix}
  k_{t-1} \\
x_t
\end{bmatrix}
+ QA^{-1}C\varepsilon_t
\]

\[= \begin{bmatrix}
  Q_A & Q_B \\
  Q_C & Q_D
\end{bmatrix}
\begin{bmatrix}
  k_{t-1} \\
x_t
\end{bmatrix}
+ \begin{bmatrix}
  \Lambda_s & 0 \\
  0 & \Lambda_x
\end{bmatrix}
\begin{bmatrix}
  Q_A & Q_B \\
  Q_C & Q_D
\end{bmatrix}
\begin{bmatrix}
  k_{t-1} \\
x_t
\end{bmatrix}
+ \begin{bmatrix}
  \Omega_s \\
  \Omega_x
\end{bmatrix}\varepsilon_t
\]

(1.57)

where $QA^{-1}C \equiv [\Omega_s, \Omega_x]'$. $\Lambda$ is the eigenvalue matrix in which $\Lambda_s$ is the matrix with eigenvalues lesser than unity and $\Lambda_x$ with eigenvalues greater than or equal unity. According to the so-called Blanchard-Kahn condition (Blanchard and Kahn, 1980), the characteristics of the solution can be categorized as follows:

- rank($\Lambda_x$) = rank($x$) Suddle stable (the solution is uniquely determined)
- rank($\Lambda_x$) > rank($x$) No solution.
- rank($\Lambda_x$) < rank($x$) Global stable (indeterminancy)

In the following, the solution is assumed to be suddle stable. Eliminating explosive paths with eigenvalues greater than or equal unity from the transversality condition, we can derive the so-called “policy function”:

\[Q_Ck_{t-1} + Q Dx_t = 0\]  
\[\Rightarrow x_t = -Q_D^{-1}Q_Ck_{t-1} = \Phi k_{t-1}\]  

(1.58)  
(1.59)

where $-Q_D^{-1}Q_C \equiv \Phi$. The policy function shows the control variable $x_t$ is optimally determined when the predetermined variable $k_t$ is given. Substituting the policy function into the original system equations, we can obtain the so-called “state transition equation”:

\[Q_Ak_t + Q_Bx_{t+1} = \Lambda_s[Q_Ak_{t-1} + Q_Bx_t] + \Omega_s\varepsilon_t\]
\[\Rightarrow [Q_A + Q_B\Phi]k_t = \Lambda_s[Q_A + Q_B\Phi]k_{t-1} + \Omega_s\varepsilon_t\]
\[\Rightarrow k_t = [Q_A + Q_B\Phi]^{-1}\Lambda_s[Q_A + Q_B\Phi]k_{t-1} + [Q_A + Q_B\Phi]^{-1}\Omega_s\varepsilon_t\]
\[\Rightarrow k_t = Dk_{t-1} + H\varepsilon_t\]  

(1.60)

where $D \equiv [Q_A + Q_B\Phi]^{-1}\Lambda_s[Q_A + Q_B\Phi]$ and $H \equiv [Q_A + Q_B\Phi]^{-1}\Omega_s$. Since components of $\Phi, D, H$ are represented by nonlinear functions of parameters $\theta$, the solution of the DSGE model can be described as follows:

\[k_t = D(\theta)k_{t-1} + H(\theta)\varepsilon_t\]  
\[x_t = \Phi(\theta)k_{t-1}\]  

(1.61)  
(1.62)

Finally, letting a state variable vector defined as $S_t \equiv [k_t', x_t']'$, the solution of the DSGE model can be obtained as the “state equation”:

\[S_t = G(\theta)S_{t-1} + E(\theta)\varepsilon_t\]  

(1.63)
Revisiting Lucas and Sims critiques

By the way, we can see the state equation (1.63) takes the same form as the VAR model. Using the state equation, we can interpret the two influential criticisms introduced at the beginning of this section.

**Lucas critique:** Components of $G(\theta)$ in (1.63), reduced-form parameters, take forms of higher-order and nonlinear functions of structural parameters, $\theta$. Changes in structural parameters such as $\psi_1$ and $\psi_2$ in the monetary policy rule lead to changes in any reduced-form parameters. Similarly, a change in the nominal rigidity parameter $\xi$ also affects any reduced-form parameters. Suppose we observe changes in reduced-form parameters estimated by the unrestricted VAR model. Can we identify the changes come from which changes in structural parameters?

**Sims critique** Components of $E(\theta)$ in (1.63) are similarly higher-order and nonlinear functions of structural parameters, $\theta$. When we estimate the unrestricted VAR model, the estimated residual is derived as $E(\theta)e_t$. Thus, the reduced-form residual must be estimated as a linear combination of structural shocks: In the example here, the residual will be represented by a linear combination of the productivity shock, $\varepsilon_t^p$, and the monetary policy shock, $\varepsilon_t^R$. Essentially, we cannot identify which of the structural shock caused the business cycle only estimating the residuals of the unrestricted VAR model. Thus, we cannot extract only monetary policy shock from the residual. To overcome this shock identification problem, the “structural” VAR model imposes restrictions as much as necessary for identifying structural shocks into $E(\theta)$. Sims criticized that the number of restrictions in the traditional macro-models is too many for identifying structural shocks.

1.2.3 Estimating the DSGE model

The DSGE model can be expressed by the state space model:

$$
S_t = G(\theta)S_{t-1} + E(\theta)e_t
$$

$$
X_t = \Lambda S_t + \epsilon_t
$$

The state equation (1.64) is the solution of the DSGE model that defines the behaviors of agents in the model (firms, households, the policy makers, etc.). $S_t$ is the endogenous variables vector (output gap, inflation gap, nominal interest rate, etc.), $\theta$ is the set of structural parameters (subjective discount factor, nominal rigidity parameters, etc.) and $\epsilon_t$ is the exogenous structural shocks vector (productivity shock, monetary policy shock, etc.). It should be emphasized that, given parameters $\theta$, both $G(\theta)$ and $E(\theta)$ are determined from the model itself. Also, given initial endogenous variables $S_0$ and parameters $\theta$, agents in the model can decide current endogenous variables $S_t$ according to the realized structural shocks $\epsilon_t$.

For example, consider a DSGE model which consists of five endogenous variables: Output gap $\hat{y}_t$, inflation gap $\hat{\pi}_t$, nominal interest rate $\hat{R}_t$, AR(1) productivity shock $\hat{z}_t$, and AR(1) monetary policy shock $\epsilon_t^R$, so the endogenous variables vector is $S_t = [\hat{y}_t, \hat{\pi}_t, \hat{R}_t, \hat{z}_t, \epsilon_t^R]^\prime$. Also, the model has two structural shocks as the sources of the business cycle, i.i.d. monetary policy shock $\varepsilon_t^R$ and i.i.d. productivity shock $\varepsilon_t^p$. In this simple DSGE model, the state equation can be expressed:

$$
\begin{bmatrix}
\hat{y}_t \\
\hat{\pi}_t \\
\hat{R}_t \\
\hat{z}_t \\
\epsilon_t^R
\end{bmatrix}_{5 \times 1}
= 
\begin{bmatrix}
G_{1,1}(\theta) & \cdots & G_{1,5}(\theta) \\
\vdots & \ddots & \vdots \\
G_{5,1}(\theta) & \cdots & G_{5,5}(\theta)
\end{bmatrix}
\begin{bmatrix}
\hat{y}_{t-1} \\
\hat{\pi}_{t-1} \\
\hat{R}_{t-1} \\
\hat{z}_{t-1} \\
\epsilon_{t-1}^R
\end{bmatrix}_{5 \times 1}
+ 
\begin{bmatrix}
E_{1,1}(\theta) & E_{1,2}(\theta) \\
\vdots & \vdots \\
E_{5,1}(\theta) & E_{5,2}(\theta)
\end{bmatrix}
\begin{bmatrix}
\varepsilon_t^R \\
\varepsilon_t^p
\end{bmatrix}_{2 \times 1}
$$

(1.66)
On the other hand, the measurement equation (1.65) is determined by econometricians to match endogenous variables with data. \( X_t \) is the data vector and \( e_t \) is the measurement errors vector. \( \Lambda \) is the so-called selection matrix which represents matching endogenous variables with data. Suppose that an econometrician regards output gap, inflation gap and nominal interest rate as observable, and corresponds these endogenous variables to GDP data \((X^y_t)\), GDP deflater \((X^\pi_t)\) and overnight call rate \((X^R_t)\), respectively. Then, the econometrician assumes the following measurement equation:

\[
\begin{bmatrix}
X^y_t \\
X^\pi_t \\
X^R_t
\end{bmatrix}_{3\times 1} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}_{5\times 5} \begin{bmatrix}
\hat{y}_t \\
\hat{\pi}_t \\
\hat{R}_t \\
\hat{e}_t^y \\
\hat{e}_t^\pi \\
\hat{e}_t^R
\end{bmatrix}_{(3\times 1)} + \begin{bmatrix}
\epsilon_t^y \\
\epsilon_t^\pi \\
\epsilon_t^R
\end{bmatrix}_{(3\times 1)} (1.67)
\]

The state equation determines streams of endogenous variables \( \{S_j\}_{j=1}^t \), given parameters \( \theta \), initial endogenous variables \( S_0 \) and realized structural shocks histories \( \{\varepsilon_j\}_{j=0}^t \).

\[
S_t = G(\theta)^t S_0 + \sum_{j=0}^t G(\theta)^j E(\theta)\varepsilon_j (1.68)
\]

From the measurement equation, we can replicate data generating process, \( \{X_j\}_{j=1}^t \), given \( \{S_j\}_{j=1}^t \) and measurement errors \( \{e_j\}_{j=1}^t \). Using \( \{S_j\}_{j=1}^t \) from the state equation, the data will be generated by the measurement equation:

\[
X_t = \Lambda \left[ G(\theta)^t S_0 + \sum_{j=0}^t G(\theta)^j E(\theta)\varepsilon_j \right] + e_t (1.69)
\]

Therefore, the state space model will be regarded as the data generation process (DGP). In other words, if we have parameters, initial endogenous variables, streams of shocks and measurement errors, we can produce artificial data streams (denoted as \( \{X^a_j\}_{j=1}^t \)).

The measurement error is often ignored. So, the measurement equation is replaced by:

\[
X_t = \Lambda S_t (1.70)
\]

The DGP can be rewritten as follows:

\[
X_t = \Lambda \left[ G(\theta)^t S_0 + \sum_{j=0}^t G(\theta)^j E(\theta)\varepsilon_j \right] (1.71)
\]

First, we specify the probability density of structural shock. In most cases, we assume the normal distribution with mean zero (the linear-Gaussian state space model): Following the example above, \( \varepsilon_t^\Omega \sim N(0,\sigma_\Omega^2)(\Omega \in \{R,Z\}) \). Thus, the econometrician has to estimate shocks volatilities \( \sigma_\Omega^2 \) together with \( \theta \). So, we define the set of parameters to be estimated as \( \Theta \in \{\theta, \sigma_\Omega^2\} \).

Suppose that we give some \( \Theta \). Then, drawing structural shocks from \( N(0,\sigma_\Omega^2) \), we can generate a structural shock sequence \( \{\varepsilon_t\}_{t=1}^T \). Using this sequence, we can generate a stream of endogenous variables \( \{S_j\}_{j=1}^t \) by the state equation. Then, we can generate a stream of artificial data \( \{X^a_j\}_{j=1}^t \).
by the measurement equation. Hence, we should give parameter $\Theta$ so that $X^a_j$ can replicate actual data $X_j$.

The likelihood is the data generating probability, given $\Theta$. In addition, the likelihood can be evaluated by the Kalman filter, since the probability density of structural shock is specified as the normal distribution. In other words, the probability of becoming $X^a_j = X_j$ can be calculated by giving $\Theta$. Therefore, we should estimate parameters to maximize the likelihood. This is an intuition of ML method.\textsuperscript{33}

Likelihood Evaluation using Kalman Filter

Let us evaluate the likelihood $p(X_1, X_2, \ldots, X_T|\Theta) \equiv p(X^T|\Theta)$ using the Kalman filter. The likelihood can be written as follows:

$$ p(X^T|\Theta) = p(X_0|\Theta) \prod_{t=1}^{T} p(X_t|X_{t-1}, \Theta) $$

That is, we need to evaluate $p(X_t|X_{t-1}, \Theta)$, the conditional density of $X_t$ when $X_{t-1}$ and $\Theta$ are given. Using the normal distribution density, the log likelihood can be expressed as follows:

$$ \ln p(X^T|\Theta) = -\frac{Tm}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=0}^{T} \ln |V_{t|t-1}| - \frac{1}{2} \sum_{t=0}^{T} (X_t - X_{t|t-1})' V_{t|t-1} (X_t - X_{t|t-1}) $$

(1.73)

where $m$ is the number of data ($m = \text{rank}(X_t)$), $X_{t|t-1}$ is the predicted value of $X_t$ given $X_{t-1}$ and $\Theta$ (i.e. $E(X_t|X_{t-1}, \Theta)$), and $V_{t|t-1}$ is the variance of the prediction error ($\text{Var}(X_t|X_{t-1}, \Theta) = E(X_t - X_{t|t-1}|\Theta)(X_t - X_{t|t-1}|\Theta)'$).

Suppose that we have initial values and variances of endogenous variables.

$$ E(S_0|X_0, \Theta) \equiv S_{0|0} $$

(1.74)

$$ \text{Var}(S_0|X_0, \Theta) \equiv P_{0|0} $$

(1.75)

Set $t = 1$.

**Step 1. Kalman Prediction**

Given $S_{t-1|t-1}$ and $P_{t-1|t-1}$, we can predict $S_{t|t-1}$ and $P_{t|t-1}$ by the state equation.

$$ S_{t|t-1} = G(\theta)S_{t-1|t-1} $$

(1.76)

$$ P_{t|t-1} = G(\theta)P_{t-1|t-1}G(\theta)' + E(\theta)\Sigma E(\theta)' $$

(1.77)

where $\Sigma \equiv E(\varepsilon_t\varepsilon_t')$

As an example of the simple model above,

$$ \Sigma = \begin{bmatrix} \sigma^2_R & 0 \\ 0 & \sigma^2_\varepsilon \end{bmatrix} $$

(1.78)

Again, note that if $\Theta$ is given, then $\Sigma$, $G(\theta)$ and $E(\theta)$ are determined from the model.

1.2. BASIC THEORY AND ESTIMATION METHOD

Step 2. Evaluation of the conditional likelihood
Using $S_{t|t-1}$ and $P_{t|t-1}$, we can predict data and its variance $X_{t|t-1}$ and $V_{t|t-1}$ by the measurement equation.

$$X_{t|t-1} = \Lambda S_{t|t-1}$$
$$V_{t|t-1} = \Lambda' P_{t|t-1} \Lambda$$

(1.79) (1.80)

From the linear Gaussian assumption, the conditional likelihood can be written as follows:

$$p(X_t|X_{t-1}) = (2\pi)^{-m/2} \left| V_{t|t-1}^{-1} \right|^{1/2} \exp \left[ -\frac{1}{2} (X_t - X_{t|t-1})' V_{t|t-1}^{-1} (X_t - X_{t|t-1}) \right]$$

(1.81)

After observing $X_t$, this conditional likelihood $p(X_t|X_{t-1})$ can be calculated, given $S_{t-1|t-1}$, $P_{t-1|t-1}$, $X_t$ and $\Theta$.

Step 3. Kalman Updating
To reflect the data information at period $t$, endogenous variables and variance are updated from $S_{t-1|t-1}$ and $P_{t-1|t-1}$ to $S_{t|t}$ and $P_{t|t}$.

$$S_{t|t} = S_{t|t-1} + P_{t|t-1} \Lambda V_{t|t-1}^{-1} (X_t - X_{t|t-1})$$
$$P_{t|t} = P_{t|t-1} - P_{t|t-1} \Lambda V_{t|t-1}^{-1} \Lambda' P_{t|t-1}$$

(1.82) (1.83)

where $(X_t - X_{t|t-1})$ is the prediction error. Thus, the coefficient $P_{t|t-1} \Lambda V_{t|t-1}^{-1}$ (called the Kalman gain) indicates how much to correct the prediction error.

Step 4 Set $t = t + 1$ and return to Step 1. Repeat until $t = T$.
Then, we can obtain $\{p(X_t|X_{t-1}, \Theta)\}_{t=1}^T$ and evaluate the log likelihood $\ln p(X^T|\Theta)$ from (1.73).

Estimation via MCMC
If we have parameters $\Theta$, we can evaluate the likelihood, the probability of generating data $X^T$. The situation we placed, of course, is completely opposite: The data $X^T$ is at our hand and we do not know parameters $\Theta$. So, we have to estimate the parameters utilizing the data.

Estimating parameters can be carried out by the ML estimation method just to maximize the log likelihood. But recently, parameters are often estimated via the Bayesian technique, reflecting prior information of parameters. In other words, our estimation target is the posterior distribution of parameters $p(\Theta|X^T)$, taking parameters uncertainties into account.\textsuperscript{34}

After observing the realized data, we have to estimate the posterior distribution of parameters. Using the Bayes’ theorem, the cause (parameter) can be very logically estimated from the result (data): The posterior distribution is determined by both the likelihood $p(X^T|\Theta)$ and the prior distribution $p(\Theta)$:

$$p(\Theta|X^T) \propto p(X^T|\Theta) \cdot p(\Theta)$$

(1.84)

\textsuperscript{34}Another interpretation for parameters to have distributions would be possible. In the Bayesian estimation, a certain “diversity” with regard to agents’ tastes and technologies: For example, consumption habit formation parameter and Calvo type nominal rigidities parameters are estimated as (posterior) distributions. Thus, there assumed to be agents with various tastes and skills, and we regarded that they make decisions according to their respective parameters. When conducting policy simulation using the estimated posterior mean, we are considering the effect on the “average” agent.
The likelihood can be evaluated by the Kalman filter. Hence, by setting the prior distribution, we can calculate the posterior distribution by the Bayes’ theorem. The prior distribution means the parameters information before observing the data. As examples on the prior distributions in the simple DSGE model described above, the Calvo parameter $\xi$ must be in $[0, 1]$, the external habit formation parameter $h$ also must be in $[0, 1]$, and so on.

Usually, the MCMC (Markov Chain Monte Carlo) method is used for estimating the posterior distribution. The MCMC method is, generally speaking, an algorithm for sampling the current parameter $\Theta^j$ under the condition that the previous parameter $\Theta^{j-1}$ was drawn. The MH algorithm (Metropolis-Hastings algorithm) is the method most widely used for estimating the posterior distribution of parameters in the linear Gaussian state space model.

To improve computational efficiency, the estimation is often carried out in the following two stages in practice: In the first stage, we numerically calculate the mode and the Hessian of the log posterior density. In the second stage, we estimate the posterior distribution of parameters with the MH algorithm. Suppose that we have completed the first stage. The second stage is going as follows.

The $j-1$-th sampling result is $\Theta^{j-1}$. When $j=1$ (initial parameter), we use the mode of the log posterior density obtained in the first stage.

**Step 1:** A candidate $\Theta^*$ is drawn from the proposal distribution $N(\Theta^{j-1}, c\Sigma_{\Theta})$.

In other words, the candidate $\Theta^*$ is sampled according to the random walk process, $\Theta^* = \Theta^{j-1} + u^j$ where $u^j \sim N(0, c\Sigma_{\Theta})$: This algorithm is called as the random walk MH algorithm. $\Sigma_{\Theta}$ is the Hessian of the log-posterior density obtained in the first stage. This regulates the step width of the random walk process: Intuitively speaking, imagine a mountain (posterior distribution, targeted here). The Hessian corresponds to the curvature of the mountain. The step width is decreased at the high curvature of the mountain, and increased at the low curvature. In other words, we can efficiently draw the mountain by sampling a large number of parameters around the top of the mountain (high curvature) and sampling a small amount of parameters around the foot of the mountain (low curvature). Also, researchers should adjust the coefficient $c$ so that the acceptance ratio (shown below) will be 20-50%.

**Step 2:** Calculating $\alpha_j$ from (1.85), the $j$-th sampling result $\Theta^j$ is decided by the acceptance-rejection process of (1.86).

\[
\alpha_j = \min \left[ 1, \frac{p(\Theta^*|X^T)}{p(\Theta^{j-1}|X^T)} \right] (1.85)
\]

\[
\Theta^j = \begin{cases} 
\Theta^* & \text{with probability } \alpha_j \\
\Theta^{j-1} & \text{with probability } 1 - \alpha_j 
\end{cases} \quad (1.86)
\]

Return to Step 1 until the sampled parameter distribution converges.

Finally, we have got our estimation target: Discarding sampled parameters before convergence (burn-in), the remaining parameters are estimated posterior distribution of parameters $p(\Theta|X^T)$. We can calculate the required moments of parameters (mean, credible intervals, etc.). Using estimated parameters, we can also estimate (smoothed) structural shocks and (smoothed) endogenous variables from the state space model.

Now, using these estimation results, we can evaluate the effects of monetary and fiscal policies by the impulse response functions, and investigate the sources of the business cycle based on variance and historical decompositions of endogenous variables.
In this thesis, we extend the measurement equation to improve the estimation efficiency, and also expand the state equation (expand the theoretical model) to examine the sources of the business cycle in the U.S. and Japan. Let us turn to the introduction of my research.

1.3 Organization of the Thesis

This thesis can be roughly divided into two categories: The first part is the extension of the estimation methodologies of the empirical DSGE approach (Chapters 2 and 3) and the second part is application examples of the empirical DSGE approach (Chapters 4 and 5).

The first part discusses the extension of the estimation method such as introduction of measurement errors and estimation method making use of a large number of data information.

Chapter 2 considers the role of measurement errors theoretically and empirically in the current empirical DSGE approach. In particular, we investigate how much the SW model depended on ad hoc markup shocks to explain inflation. Then, by introducing the measurement error which is a component which cannot be explained by the model, we examine how the fit of the model to the data improves.

Chapter 3 examines how the data rich estimation method that utilizes a lot of data information improves estimation accuracy of parameters and structural shocks. In addition, we compare the results of historical decompositions of conventional estimation method and data rich estimation method, and consider the difference of both estimation methods with respect to the sources of business cycle and inflation fluctuations.

In the second part, we apply the empirical DSGE approach to the U.S. economy and the Japanese economy.

Chapter 4 expands to the DSGE model which introduced financial friction and time-varying structural shock volatility, and examine the sources of the Great Recession of the U.S. using data rich estimation method. In particular, we consider whether the Great Recession was a bad luck, and the bad luck was caused by deterioration in balance sheet between banks and corporates. Also, we will see if a good policy existed after the financial crisis, and how well the good policy supported the U.S. economy.

Chapter 5 extends to the DSGE model introducing unemployment and non-wasteful government expenditure and quantitatively grasp the effect of fiscal stimulus on unemployment for the Japanese economy. We consider the impact of new channels of government spending such as the Edgeworth complementarities and the productivity effect of public capital on unemployment.

Finally, Chapter 6 summarizes the conclusion and the direction of future research.
Part I

Estimation Methodologies
Chapter 2

Role of Measurement Error

2.1 Introduction

Chapter 2 discusses (i) the problems of the so-called markup shocks in the SW model, considered as the benchmark model of the empirical DSGE approach, and (ii) the role of measurement errors that have not been noticed.


2.1.1 Background

As a major structural shock on the NKPC explaining inflation variations, there is a typical supply shock called as the productivity shock. SW (2003, 2007), in addition to the productivity shock, introduced the so-called markup shock into the log-linearized NKPC, in a one-to-one correspondence with inflation. Considering the simple model in the previous section as an example, it means $\varepsilon^n_m$ called the markup shock is added to the log-linearized NKPC as follows (see (1.52)):

$$\dot{\pi}_t = \frac{1}{1 + \beta} \frac{1}{\hat{\pi}_t + 1 + \kappa \left[ \left( \sigma_L + \frac{\sigma}{1 - h} \right) \hat{y}_t - \frac{\sigma h}{1 - h} \hat{y}_{t-1} - (1 + \sigma_L) \hat{z}_t \right] + \varepsilon^n_m}$$

The problem is the role of markup shock added to the NKPC in this ad hoc manner has played a tremendous role to explain inflation fluctuations. According to variance decompositions of SW (2007) using the U.S. data, most of inflation fluctuations are explained by the price markup shock in the short term and the wage markup shock in the long term (Over 80% of inflation fluctuations are explained by both markup shocks, whether in short or long term). And the contribution of the productivity shock $\hat{z}_t$, another key structural shock on the NKPC, to inflation fluctuations was extremely insignificant.

However, markup shocks on price and wage (from their names) are shocks related to parameters of the price elasticity of intermediate goods demand and the wage elasticity of labor demand, respectively. Normally, the markup rate will increase if the price elasticity of intermediate goods demand (wage elasticity of labor demand) decreases, in other words, if the necessity of intermediate goods (workers) increases and the market power of intermediate goods firms (workers) increases. Therefore, SW (2007) tells that most of inflation fluctuations are caused by markup variations, that is, it is attributed to frequent fluctuations in the market powers of intermediate goods firms and workers (fluctuations in the necessity of intermediate goods and workers).
The reason why the name is the markup shock can be understood from the simple model example. The optimal pricing condition of intermediate goods firms in the case of flexible price adjustment can be written as follows (see (1.14)):

\[
P_t(i)^* = (1 + \lambda)\Psi_t
\]

where \(\lambda\) is the (net) price markup rate, \(\Psi_t\) is the real marginal cost, and the left-hand side is the relative price of the intermediate goods price \(P_t(i)\) against the general price \(P_t\). Since the price adjustment is flexible, the aggregate supply curve becomes vertical. Defining the nominal marginal cost as \(MC_t \equiv P_t\Psi_t\), the inflation after the log-linear approximation around the steady state is expressed as follows:

\[
\hat{\pi}_t = MC_t
\]

In case of flexible price adjustment, if the nominal marginal cost (here, nominal wage) changes, it just means that the inflation will jump to exactly the same change as the nominal marginal cost and the price adjustment will be immediately completed. What SW did is equivalent to adding shock \(\varepsilon^m_t\) to this right-hand side.

\[
\hat{\pi}_t = MC_t + \varepsilon^m_t
\]

Now, the same result can be obtained by adding the following shock to the (gross) markup rate \((1 + \lambda)\) before log-linear approximation:

\[
P_t(i) = (1 + \lambda)e^{\varepsilon^m_t}\Psi_t
\]

From such a point of view, it certainly might be possible to capture it like a shock associated with the gross markup rate. However, when considering another shock called “relative price shock” that distorts the relative price of the intermediate goods price against the general price, the relative price shock could enter into the log-linearized NKPC in the same form as the markup shock (observational equivalent). It is criticized to add a structural interpretation by regarding ad hoc markup shocks as structural shocks (de Walque et al. 2006, Chari et al. 2009). For example, if the shock of \(\varepsilon^R_P\) is added as “relative price shock” as follows, the resulting aggregate supply curve after the log-linear approximation is the same (only the sign of the shock is different):

\[
P_t(i)e^{\varepsilon^R_P} = (1 + \lambda)\Psi_t
\]

Therefore, for the question of what is the source of inflation variations in this model, it can be explained that the change in markup is the source and it can be also explained that fluctuation of a shock distorting the relative price of intermediate goods against general price. In other words, adding shocks easily leads to a problem of lowering the advantage of the empirical DSGE approach that it can provide structural interpretations in line with the theoretical model for output and inflation fluctuations.

Similarly, in the empirical aspect, there is also a contradiction in interpreting that the main source of inflation fluctuations is due to the SW type markup shocks. In the context of the business cycle in macroeconomics, there is a long debate whether output and markup are pro-cyclical or counter-cyclical. According to the current consensus based on empirical analysis, output and markup is counter-cyclical (Bils 1987, Warner and Barsky 1995, Chevalier and Scharfstein 1996).
In this case, it is necessary for the markup to decline (against the nominal marginal cost) during the booming inflation phase.

Consider about the situation where the boom and inflation occur simultaneously due to the negative monetary policy shock ($\varepsilon^R_t < 0$: Monetary easing policy) in a simple model. To meet the empirical consensus that output and markup are counter-cyclical, the markup shock $\varepsilon^m_t$ must respond negative. That is, a positive correlation is required between “structural shocks” (monetary policy shock and markup shock), which have to be independent from each other.

In summary, it is difficult to justify the SW type markup shock as a “structural shock” both theoretically and empirically. So, what kind of shock should we capture inflation fluctuations?

According to data, it is well known that inflation since the 1990s shows a very volatile behavior in both Japan and the U.S. How well can DSGE models or other time series models predict such volatile inflation?

Edge and Gurkaynak (2010) examined predictability for inflation in the U.S. using an official DSGE model of FED (EDO model; Estimated Dynamic Optimization Model). As a result, the prediction accuracy of the DSGE model for recent inflation was extremely low. Unfortunately, however, the low prediction accuracy for inflation is not attributable to the DSGE model’s characteristics that there are a lot of cross-equation constraints and parameter restrictions. According to VAR analysis by Cogley et al. (2010), inflation before the 1990s is smooth fluctuations with inertia, but since the 1990s it shows a noisy and volatile movement and therefore they concluded that it is difficult to predict inflation even in the VAR model. In addition, from the results using dynamic factor model (DFM), which is adopted to measure the diffusion index (DI) in the U.S., inflation after the 1990s is still volatile and it is difficult to predict (Stock and Watson 2007). In the end, even if we employ atheoretical time series models with little equation constraints or parameter restrictions, it is still difficult to capture inflation fluctuations.

Basically, it is the first best to further brush up the structural model expressing the pricing behavior of intermediate goods firms, to derive the aggregate supply curve that output and markup become counter-cyclical and can explain the volatile inflation. However, even reduced-form time series models such as VAR and DFM are difficult to predict inflation. It seems difficult to explain inflation fluctuations by building a structural DSGE model with many equation constraints and parameter restrictions. So, as a second best, we focused on the role of measurement errors.

Introducing measurement errors means to divide the data fluctuations into two components: a component explainable by the model and the other component unexplainable by the model. The state equation expresses any endogenous variable as a linear combination of all structural shocks (and initial values of endogenous variables). For example, fluctuations in the productivity shock of the NKPC systematically spread to various endogenous variables through $E(\theta)$ and $G(\theta)$ (see (1.64)). On the other hand, according to the measurement equation, any measurement error added to a certain data does not affect other data and other endogenous variables. In other words, the measurement error can be regarded as “idiosyncratic component” (or unique factor) that only the data has. Given that current inflation data shows volatile fluctuations and if it is not possible to predict or explain inflation well both in DSGE models and other time series models, we should separate the inflation data fluctuations into two components: The component that can be described by the model and the other component that cannot be explained by the model (measurement error= idiosyncratic component of the data).

There are very few estimated DSGE models with measurement errors: Measurement errors are not introduced in SW (2003, 2007) for the Euro area and the U.S., in Sugo and Ueda (2008) for Japan, or in official estimated DSGE models of various central banks and governments. If
measurement errors are not introduced, it implies to capture inflation data by structural shocks. However, according to SW (2007), the volatile fluctuations of inflation data is unfortunately forced to capture as fluctuations of ad hoc markup shocks which are difficult to interpret structurally.\footnote{Gerron-Quintana (2010) introduced measurement errors into the SW (2007) model and found that estimated parameters are quite sensitive depending upon which data are selected, but they did not consider the problem of the markup shocks. On the other hand, in the VAR model, there is a paper trying to extract policy effects after introducing measurement errors and eliminating factors not explained in the model (Carriero et al. 2015)}

2.1.2 Purposes, Originalities, and Methodologies

When the SW model explains inflation data and wage data, how much it relies upon ad hoc “markup shocks” difficult to add structural interpretations? Is there a possibility that the SW model has not yet been a structural model that can explain inflation data and wage data?

Usually, the Metropolis-Hastings (MH) algorithm has been used as the estimation method for the standard DSGE model. The main target is to estimate structural parameters, with an assumption that observed data and endogenous variables are definitively connected.

However, can endogenous variables (or state variables) explain all fluctuations in data? Even inflation data showing volatile fluctuations that are difficult to predict should we capture all fluctuations with endogenous variables? Should it be assumed that there are unique fluctuations in data that cannot be explained by the model? If so, how do we separate the data specific fluctuations? In other words, how much data variation can be accurately grasped by endogenous variables?

To answer this question, we must consider another method to estimate not only parameters but also endogenous variables with high accuracy. Then, we estimate the SW model by introducing measurement errors which are components unexplainable by the model, and examines how much can it be improved to capture volatile inflation and wage data.

Therefore, this study introduces measurement error and discards the assumption that state variables and data are definitively connected. The measurement error represents the unique component of the data fluctuation. In this case, state variables and measurement errors must be estimated together with the structural parameters.

Boivin and Giannoni (2006) and Kryshko (2010) proposed a hybrid MCMC method for the DSGE model introducing the measurement error with the AR (1) process as a method to improve the estimation accuracy of state variables. The method of Carter and Kohn (1994) is adopted as an estimating method of state variables, but as Chib (2001) pointed out, this method has the problem of interrupting sampling. In response to this problem, this research adopts the simulation smoother proposed by de Jong and Shephard (1995) as an estimating method of state variables and propose a new estimation method of DSGE model with measurement error.

Based on this method, we estimate the SW model with measurement error, and compared the result with the case of the SW model without measurement error.

First of all, as in previous studies, we estimate a model that explains fluctuations in inflation and wage data with price and wage markup shocks using Japanese data. Next, we remove the markup shocks of price and wage, and estimate the model with measurement errors added to inflation and wage data. Then, we compare the fit (the marginal likelihood) for the data of both models and examine which approach can successfully capture the behavior of the data.

Considering the previous empirical results, it is expected that when measurement errors are introduced, most of volatile variations in inflation data would be captured as measurement errors. In addition, it is also expected that our approach with measurement errors would increase in the
marginal likelihood as compared with the conventional approach with markup shocks that affect other endogenous variables.

As a result, there is a negligible difference in structural parameters, structural shocks, and historical decomposition between the two. The result is due to the difference in approaches to capture volatile data movements as structural shocks or as measurement errors. Finally, as a result of model selection, we find that the model introducing the measurement error has a much higher fit to the data than the model without considering the measurement error. In recent years, inflation has shown noisy fluctuation and it becomes difficult to predict. Hence, measurement error should be introduced in the DSGE model, and the hybrid MCMC method with the simulation smoother is useful for the estimation.

There are two originalities on the empirical methodologies: First, we adopt Metropolis within Gibbs when estimating parameters of measurement errors (variances and persistence parameters). Those parameters can be estimated even with the MH algorithms usually used, but adopting this algorithm reduces computational efficiency. Given structural parameters of the model, parameters of measurement errors can be estimated by the (Bayesian) least squares method. Concretely speaking, given structural parameters of the model drawn by the MH algorithm, the conditional densities of measurement errors parameters are well known kernels. So, it is easy to sample those parameters by the Gibbs sampling, which is more computationally efficient than the MH algorithm. Therefore, to improve computational efficiency, the model’s structural parameters are estimated by the MH algorithm and the Gibbs sampler is adopted in estimating parameters of measurement error.

Second, to estimate endogenous variables, we apply the simulation smoother proposed by de Jong and Shephard (1995). The Kalman smoother proposed by Carter and Kohn (1994) is often used for estimating endogenous variables. When the sampling number is large, however, estimates of endogenous variables stacked in the computer become an enormous volume, which increases in the computational burden. Instead, we estimate endogenous variables by the simulation smoother (or disturbance smoother) of de Jong and Shephard (1995). Intuitively, it is only the sampled structural shocks and structural parameters that are stacked on the computer, and it is not necessary to stack sampled endogenous variables, so the computational burden is greatly reduced by this method. We estimate smoothed shocks using sampling results of structural shocks stacked in the computer. Then, using the smoothed shocks and sampled structural parameters, we can calculate the estimates of endogenous variables.

2.1.3 My Contributions

Chapter 2 revised the paper based on Matsumae et al. (2011) which was written with three co-authors.

My main contribution is about the central ideas of this research: The SW model is heavily rely on markup shocks to explain inflation fluctuations. However, since the markup shock is difficult to interpret structurally, a natural question arises the SW model were hardly able to explain anything about inflation variations structurally. Therefore, by introducing measurement errors, unexplainable components from the model, I got an idea to examine how much in actual the SW model were able to explain the inflation variations.

My contribution on the estimation method is two: First, in estimating parameters of measurement errors, I suggested the Gibbs sampling should be utilized from the view of computational efficiency. Second, I found the computational load was very high in estimating endogenous variables by the Kalman smoother. So, I proposed the simulation smoother of de Jong and Shephard (1995) should be introduced.
2.1.4 Organization of Chapter 2

The structure of Chapter 2 is as follows. Section 2 presents the DSGE model to be estimated. Section 3 proposes a new estimation method of the DGSE model with measurement error by the hybrid MCMC method. In particular, a method for smoothing state variables will be described. Section 4 explains the data used and preliminary settings. The results of estimation are shown in Section 5 and the conclusions are summarized in Section 6.

2.2 Model

This study estimates the standard DSGE model proposed by CEE (2005) and SW (2003, 2007). The model installs not only the new Keynesian properties of price and wage nominal rigidities, but also embeds real rigidities such as investment adjustment cost and consumption habit formation. Essentially, equations in the model are based on microeconomic foundations derived from dynamic optimization behaviors of households and firms.

First, we log-linearized the model around steady states to implement the Bayesian technique for estimating structural parameters. Next, we solve the simultaneous difference equation (derive the rational expectation equilibrium, hereinafter REE) using the method of Sims (2002) and combine the solution (the state equation) with the measurement equation to build a linear state space model.

In the following, we will summarize the linearized difference equations shown in SW (2003, 2007) in which the variables with “hat” represent the percent deviations from steady states.

2.2.1 Households

(1) Euler equation on consumption:

\[ \hat{c}_t = \frac{h}{1 + h} \hat{c}_{t-1} + \frac{1}{1 + h} E_t \hat{c}_{t+1} - \frac{1 - h}{(1 + h) \sigma_c} (\hat{R}_t - E_t \hat{\pi}_{t+1}) + \frac{1 - h}{(1 + h) \sigma_c} (u_c^c - E_t u_{t+1}^c) \] (2.1)

where \( \hat{c}_t \) is consumption, \( \hat{R}_t \) is nominal interest rate, \( \hat{\pi}_t \) is inflation, \( u_c^c \) is AR(1) preference shock, and \( E_t \) is expected value operator. \( \sigma_c \) is CRRA parameter (the reciprocal of elasticity of intertemporal substitution on consumption, hereinafter IES), \( h \in [0, 1] \) is habit persistence parameter. If \( h \neq 0 \), current optimal consumption depends not only upon anticipated consumption in the next period but also upon previous consumption from habit formation. The parameter \( h \) determines the weight on the previous and the next period. Furthermore, consumption is a decreasing function of ex-ante real interest rate, but the extent relies not only on IES parameter but also on habit formation parameter. \( u_c^t \) is AR(1) preference shock so that \( u_c^t = \rho^c u_{t-1}^c + \varepsilon_c^t \) where \( \varepsilon_c^t \) is i.i.d preference shock and \( \rho^c \) is persistence parameter.

(2) Euler equation on investment:

\[ \hat{inv}_t = \frac{1}{1 + \beta} \hat{inv}_{t-1} + \frac{\beta}{1 + \beta} E_t \hat{inv}_{t+1} + \frac{\varphi}{1 + \beta} \hat{q}_t + \frac{\beta}{1 + \beta} (E_t u_{t+1}^{inv} - u_t^{inv}) \] (2.2)

where \( \hat{inv}_t \) is investment and \( \hat{q}_t \) is shadow price of capital (Tobin’s marginal q). \( \beta \) is subjective discount factor and \( 1/\varphi \) is investment adjustment cost parameter. Due to adjustment cost of investment, the current optimal investment will also be determined, as with consumption, from both expected investment at the next period and previous investment. \( u_t^{inv} \) is AR(1) investment adjustment cost shock so that \( u_t^{inv} = \rho^{inv} u_{t-1}^{inv} + \varepsilon_t^{inv} \) where \( \varepsilon_t^{inv} \) is i.i.d. investment adjustment cost shock.
2.2. MODEL

and $\rho^{inv}$ is persistence parameter.

(3) Capital price transition equation:

$$q_t = -(\hat{R}_t - E_t \hat{\pi}_{t+1}) + \frac{1 - \tau}{1 - \tau + \overline{r}k} E_t \hat{q}_{t+1} + \frac{\overline{r}k}{1 - \tau + \overline{r}k} E_t \hat{r}_k + \varepsilon^q_t$$ (2.3)

where $\hat{r}_k$ is real return rate of capital, $\varepsilon^q_t$ is equity premium shock, $\tau$ is capital depreciation rate and $\overline{r}k$ is the steady state of real return rate. The shadow price of capital $\hat{q}_t$ becomes a decreasing function of ex-ante real interest rate and an increase function of forward looking terms on real return rate and capital price. Since the return on investment would be changed by equity premium shock, $\varepsilon^q_t$, capital price held by households will also be affected. There is no inertia in equity premium shock, and supposed to be i.i.d. shock.

(4) NKPC (wage)

$$\hat{w}_t = \frac{\beta}{1 + \beta} E_t \hat{w}_{t+1} + \frac{1}{1 + \beta} \hat{w}_{t-1} + \frac{\beta}{1 + \beta} E_t \hat{\pi}_{t+1} - \frac{1}{1 + \beta} \hat{\pi}_t + \frac{\gamma_w}{1 + \beta} \hat{\pi}_{t-1} - \frac{\lambda_w (1 - \beta \xi_w)(1 - \xi_w)}{(1 + \beta)(1 + \lambda_w + \sigma_L)\xi_w} \left[ \hat{w}_t - \sigma_L \hat{l}_t - \frac{\sigma_c}{1 - h} (\hat{c}_t - h \hat{c}_{t-1}) - u^L_t \right] + \varepsilon^w_t$$ (2.4)

where $\hat{w}_t$ is real wage, $\hat{l}_t$ is labor supply, $u^L_t$ is AR(1) labor supply shock and $\varepsilon^w_t$ is wage markup shock. $\xi_w$ is the probability that households cannot revise wage (nominal wage rigidity: the wage Calvo parameter), $\lambda_w$ is the parameter on wage elasticity of labor demand (or wage markup rate parameter). $\sigma_L$ is inverse Frisch elasticity (reciprocal of wage elasticity of labor supply), $\gamma_w$ is the lagged inflation indexation parameter (wage indexation parameter).

The NKPC is the short-term upward-sloping labor supply curve in ($\hat{l}_t$, $\hat{w}_t$) plane due to nominal wage rigidity. The bracket in the last term illustrates a discrepancy of real wage and the MRS between leisure and consumption. The NKPC becomes vertical if $\xi_w \to 0$. In other words, if all households can adjust nominal wages flexibly, labor supply is optimally decided so that the MRS is equalized to real wage. However, since there are households who cannot revise nominal wages, there will be the short-term gap between real wage and the MRS. The current real wage $\hat{w}_t$ also depends on previous real wage. If households cannot set wages optimally, wages not revised are assumed to be linked with previous inflation. $u^L_t$ is AR(1) labor supply shock so that $u^L_t = \rho^L u^L_{t-1} + \varepsilon^L_t$ where $\varepsilon^L_t$ is i.i.d. labor supply shock and $\rho^L$ is persistence parameter.

It should be noted that wage markup (i.i.d.) shock $\varepsilon^w_t$ is introduced in an ad-hoc manner: The shock cannot interpret structurally but will be useful to capture the volatile real wage data.

(5) Capital transition equation:

$$\hat{K}_t = (1 - \tau) \hat{K}_{t-1} + \tau \hat{K}_{inv} \hat{v}_{t-1}$$ (2.5)

where $\hat{K}_t$ is capital. Note that the capital depreciation $\tau$ equals steady state of investment-capital ratio $I/K$.

2.2.2 Firms

(6) Cost minimization condition:

$$\hat{l}_t = -\hat{w}_t + (1 + \psi) \hat{r}_k^t + \hat{K}_{t-1}$$ (2.6)
where $\psi$ is adjustment cost parameter of capital utilization rate. Given capital, labor demand becomes a decrease function of real wage and an increasing function of capital rental rate.

(7) Production function:

$$\hat{y}_t = \phi \hat{a}_t^q + \phi \alpha \hat{K}_{t-1} + \phi \alpha \psi \hat{r}_t^k + \phi (1 - \alpha) \hat{l}_t$$

(2.7)

where $\hat{y}_t$ is output, $\alpha$ is capital income share, $\phi - 1$ is fixed cost share of production and $\hat{a}_t^q$ is the TFP shock.

(8) NKPC (price):

$$\hat{\pi}_t = \beta_1 + \beta_\gamma p E_t \hat{\pi}_{t+1} + (1 - \beta \xi_p)(1 - \xi_p) \left[ \alpha \hat{r}_t^k + (1 - \alpha) \hat{w}_t - \hat{u}_t^g \right] + \varepsilon^p_t$$

(2.8)

where $\hat{\pi}_t$ is inflation, $\varepsilon^p_t$ is price markup shock, $\xi_p$ is the probability that firms cannot revise their prices (nominal price rigidity: the Calvo parameter on price), $\gamma_p$ is lagged inflation indexation parameter (price indexation parameter).

The NKPC is the short-term upward-sloping aggregate supply curve in $(\hat{y}_t, \hat{\pi}_t)$ plane due to nominal price rigidity. The bracket in the last term illustrates a discrepancy of real marginal cost and the marginal productivity. The NKPC becomes vertical if $\xi_p \to 0$. In other words, if all firms can adjust nominal prices flexibly, nominal price is optimally decided so that the marginal cost is equalized to marginal productivity. However, since there are firms who cannot revise nominal prices, there will be the short-term gap between marginal cost and marginal productivity. Similar to real wage, the current inflation is affected by the past inflation. If firms who cannot set prices optimally, prices not revised are assumed to be indexed with previous inflation. If $\gamma_p = 0$, then the NKPC results in a purely forward-looking Phillips curve. In other words, $\gamma_p$ indicates how much the inflation inertia is high.

Again, it should be noted that price markup (i.i.d.) shock $\varepsilon^p_t$ is also introduced in an ad-hoc manner: The shock cannot interpret structurally but will be useful to capture the volatile inflation data.

### 2.2.3 Other Equilibrium Conditions

(9) Goods market clearing condition:

$$\hat{y}_t = (1 - \tau k_y - g_y)\hat{c}_t + \tau k_y \hat{m} \hat{v}_t + \tau k \psi k_y \hat{r}_t^k + g_y \hat{u}_t^g$$

(2.9)

where $k_y$ is steady state of capital-output ratio and $g_y$ is steady state of government expenditure-output ratio. Note that the third term of the right-hand side stands the investment adjustment cost. $u_t^g$ is AR(1) government spending shock so that $u_t^g = \rho^g u_{t-1}^g + \varepsilon_t^g$ where $\varepsilon_t^g$ is i.i.d. government spending shock and $\rho^g$ is persistence parameter.

(10) Monetary policy rule:

$$\hat{R}_t = \rho_m \hat{R}_{t-1} + (1 - \rho_m) \left[ \mu_\pi \hat{\pi}_{t-1} + \mu_y \hat{y}_t \right] + \varepsilon^m_t$$

(2.10)

where $\varepsilon^m_t$ is monetary policy shock, $\mu_\pi$ and $\mu_y$ are Taylor coefficients for inflation gap and output gap, respectively, and $\rho_m$ is interest rate smoothing parameter. The monetary authority is assumed to commit the standard Taylor rule above, manipulating nominal interest rate to stabilize both the inflation gap and output gap. The monetary policy shock $\varepsilon^m_t$ is assumed to be the i.i.d. shock.
2.3. ESTIMATION METHOD

2.2.4 AR(1) Shocks and the Measurement Errors

Five shocks introduced in the simultaneous difference equation are supposed to obey the following AR (1) processes with i.i.d. Gaussian structural shocks.

\[ u_t^c = \rho^c u_{t-1}^c + \varepsilon_t^c, \quad u_t^{inv} = \rho^{inv} u_{t-1}^{inv} + \varepsilon_t^{inv}, \quad u_t^L = \rho^L u_{t-1}^L + \varepsilon_t^L, \]

\[ u_t^a = \rho^a u_{t-1}^a + \varepsilon_t^a, \quad u_t^g = \rho^g u_{t-1}^g + \varepsilon_t^g. \]

where (11) \( \varepsilon_t^c \) is preference shock, (12) \( \varepsilon_t^{inv} \) is investment adjustment cost shock, (13) \( \varepsilon_t^L \) is labor supply shock, (14) \( \varepsilon_t^a \) is TFP shock, and (15) \( \varepsilon_t^g \) is government spending shock. In addition, we introduce six forecasting errors (hereinafter, FE).

\[ \eta_t^\pi = \tilde{\pi}_t - E_{t-1} \tilde{\pi}_t, \quad \eta_t^\nu = \tilde{\nu}_t - E_{t-1} \tilde{\nu}_t, \quad \eta_t^q = \tilde{q}_t - E_{t-1} \tilde{q}_t, \]

\[ \eta_t^{inv} = \tilde{inv}_t - E_{t-1} \tilde{inv}_t, \quad \eta_t^c = \tilde{c}_t - E_{t-1} \tilde{c}_t, \quad \eta_t^r = \tilde{r}_t - E_{t-1} \tilde{r}_t. \]

where (16) \( \eta_t^\pi \) is inflation FE, (17) \( \eta_t^\nu \) is real wage FE, (18) \( \eta_t^q \) is equity premium FE, (19) \( \eta_t^{inv} \) is investment adjustment cost FE, (20) \( \eta_t^c \) is consumption FE and (21) \( \eta_t^r \) is real rental rate FE.

2.2.5 Structural Form

After log-linearizing dynamic optimization conditions and the resource constraints around steady states, we can represent the simultaneous difference equation by the structural form.

\[ AS_t = BS_{t-1} + CE_t + D\eta_t, \quad (2.11) \]

where \( S_t \) is state variables vector: \( S_t = [\hat{y}_t, \hat{\pi}_t, \hat{w}_t, \hat{k}_t, \hat{inv}_t, \hat{q}_t, \hat{r}_t, \hat{L}_t, E_t \hat{\pi}_{t+1}, E_t \hat{w}_{t+1}, E_t \hat{inv}_{t+1}, E_t \hat{c}_{t+1}, E_t \hat{r}_{t+1}, u_t^c, u_t^{inv}, u_t^L, u_t^a, u_t^g] \), \( \varepsilon_t \) is structural shocks vector: \( \varepsilon_t = [\varepsilon_t^c, \varepsilon_t^{inv}, \varepsilon_t^q, \varepsilon_t^L, \varepsilon_t^a, \varepsilon_t^g, \varepsilon_t^{inv} \] \( \eta_t \) is FE vector: \( \eta_t = [\eta_t^\pi, \eta_t^\nu, \eta_t^q, \eta_t^{inv}, \eta_t^c, \eta_t^r] \). Note that \( A, B, C, \) and \( D \) are matrices represented by structural parameters.

2.3 Estimation Method

This section illustrates the estimation method of the state space model incorporating measurement error (hereinafter, ME). Let us denote structural parameters as \( \Theta \), state variables vector as \( S_t \) \( J \times 1 \), structural shocks vector as \( \varepsilon_t \) \( M \times 1 \), and observed variables vector as \( X_t \) \( N \times 1 \). In addition, let us define as \( S^T = \{S_t\}_{t=1}^T \) and \( X^T = \{X_j\}_{j=1}^T \) where \( T \) is the terminal period. We focus on the case where \( J \geq N \).

2.3.1 State Space Representation

We will convert the structural form to the reduced-form by deriving the REE using the method of Sims (2002). The reduced-form of DSGE model with ME is expressed as the state space model consisting of the state equation and the measurement equation. The state variable \( S_t \) is expressed

\(^2\) The case where \( J \leq N \) is called the data rich estimation method which is analyzed by Boivin and Giannoni (2006), Kryshko (2010), Schorfheide et al. (2010), Liboshii et al. (2011) and Nishiyama et al. (2011).

\(^3\) Besides the method of Sims (2002), several methods to solve the model are proposed by the eigenvalue decomposition method by Blanchard and Kahn (1980) and the undetermined coefficients method by Christiano (2002). For details, see Fujiwara and Watanabe (2011).
as an autoregressive model in the state equation, and the state variable explains observation variable \( X_t \) in the measurement equation.

\[
S_t = G(\theta)S_{t-1} + H(\theta)e_t, \quad e_t \sim N(0, Q(\theta)), \quad (2.12)
\]

\[
X_t = \Lambda S_t + e_t, \quad (2.13)
\]

\[
e_t = \Psi e_{t-1} + \nu_t, \quad \nu_t \sim N(0, R), \quad (2.14)
\]

where (2.12) is the reduced-form obtained as the solution of the model, which is called the state equation representing the transition equation of the state variables. (2.13) is called the measurement equation where \( e_t \) is the ME vector \((N \times 1)\) which is assumed to follow the AR(1) process as shown in (2.14). In the principal component analysis (hereinafter, PCA), \( \Lambda S_t \) is called common component, and the ME \( e_t \) is called idiosyncratic component in (2.13). The common component represents variations due to inter-correlation between endogenous variables (that is, the component to be co-moved by structural shocks), and the idiosyncratic component varies independently of other endogenous variables.\(^4\)

We use the terminology “measurement error” for \( e_t \), but in accordance with interpretation of PCA, it means the component that corresponds to “unique” fluctuations of observation variables without depending on structural shocks. Also, not all the state variables are observed, but since only some of them are assumed to be observed, we can formulate the measurement equation as follows:

\[ X_{it} = S_{it} + e_{it}, \quad i = 1, 2, \ldots, N \]

Hence, \( \Lambda \) can be represented by:

\[
\Lambda = \begin{bmatrix} I & 0 \end{bmatrix}
\]

where \( I \) is the identity matrix \((N \times N)\) and \( 0 \) is the zeros matrix \((N \times (J - N))\). The density of the structural shock vector \( \varepsilon_t \) is assumed by \( \varepsilon_t \sim \text{i.i.d. } N(0, Q(\theta)) \), and the density of i.i.d disturbance vector \( \nu_t \) in the ME vector \( e_t \) is \( \nu_t \sim \text{i.i.d. } N(0, R) \). The variance covariance matrices of \( Q \) and \( R \) are positive definite and diagonal. The persistence parameter vector \( \Psi \) in (2.14) is also assumed to be diagonal. Therefore, the MEs are assumed to be uncorrelated in cross section but correlated in time series with AR(1) autocorrelation. Moreover, components of \( G(\theta) \), \( H(\theta) \) and variance covariance matrix of structural shock \( Q(\theta) \) are represented by highly nonlinear functions of structural parameter \( \theta \).

Substituting (2.14) into (2.13), we can write the measurement equation as follows:

\[
(I - \Psi L)X_t = (I - \Psi L)\Lambda(\theta)S_t + \nu_t, \quad \nu_t \sim N(0, R)
\]

where \( L \) is the lag-operator. By defining \( \tilde{X}_t = X_t - \Psi X_{t-1} \) and \( \tilde{S}_t = [S_t' S_{t-1}']' \), we can reformulate the equation above:

\[
\tilde{X}_t = \begin{bmatrix} \Lambda & -\Psi \Lambda \\ \Lambda & \tilde{S}_{t-1} \end{bmatrix} + \nu_t, \quad \nu_t \sim N(0, R), \quad (2.15)
\]

\(^4\)The ME absorbs the influence of model misspecification or the effect of not corresponding state variables to appropriate data. For example, although this study is a closed economy model, data fluctuations due to the exchange rate variations will be regarded as ME fluctuations. As another example, an expectation formation consistent with the model is assumed for predicting model variables at the next period (the rational expectation equilibrium) in DSGE models, but the misspecification of this expectation formation is also absorbed as the ME fluctuations.
2.3. ESTIMATION METHOD

Similarly, we can rewrite the state equation (2.12) as follows:

\[
\begin{bmatrix}
S_t \\
S_{t-1}
\end{bmatrix} =
\begin{bmatrix}
G(\theta) & 0 \\
I & 0
\end{bmatrix}
\begin{bmatrix}
S_{t-1} \\
S_{t-2}
\end{bmatrix} +
\begin{bmatrix}
H(\theta) \\
0
\end{bmatrix}
\begin{bmatrix}
\varepsilon_t \\
\varepsilon_{t-1}
\end{bmatrix},
\varepsilon_t \sim N(0, Q(\theta)),
\] 

(2.16)

where \(I\) is the identity matrix \((J \times J)\). We will explain the estimation method based on the state space model of (2.15) and (2.16), and for the sake of convenience of explaining, we define \(\Gamma\) as \(\Gamma = \{\Psi, R\}\), which represents parameters set of the measurement equation.

2.3.2 Bayesian Estimation Method

We adopt a hybrid MCMC method (also called Metropolis-within-Gibbs) to estimate state space model with ME according to Boivin and Giannoni (2006) and Kryshko (2010). However, it should be noted that applying the simulation smoother in our research is the difference from the previous studies.

One of the advantages in estimating the DSGE model by the hybrid MCMC method is that we can estimate posteriors of not only state variables but also structural shocks. Estimating shocks posteriors could lead to more accurate policy simulations (such as historical decompositions) since we can evaluate credible bands of shocks. In contrast, the MH algorithm usually employed, cannot estimate the shocks posteriors, so we cannot conduct policy simulations with credible intervals, which would be a major disadvantage.

The targets to be estimated in the state space model are structural parameter \(\theta\), measurement equation parameter \(\Gamma\), and state variable \(S^T\). Recall that \(G(\theta), H(\theta), Q(\theta)\) in the state equation are functions of structural parameter \(\theta\). Thus, we have only to estimate \(\theta, \Gamma, S^T\). Bayesian estimation for parameters \(\theta\) and \(\Gamma\) is proceeded as follows:

**Step 1**: Set the joint prior \(p(\theta, \Gamma)\). Note that \(p(\theta, \Gamma) = p(\theta|\Gamma)p(\Gamma)\).

**Step 2**: Using Bayes’ theorem, evaluate the posterior \(p(\theta, \Gamma|X^T)\) from the prior \(p(\theta, \Gamma)\) and the likelihood \(p(X^T|\theta, \Gamma)\).

\[
p(\theta, \Gamma|X^T) = \frac{p(X^T|\theta, \Gamma)p(\theta, \Gamma)}{\int p(X^T|\theta, \Gamma)p(\theta, \Gamma)d\theta d\Gamma}.
\] 

(2.17)

**Step 3**: Calculate the moments of parameters \(\theta\) and \(\Gamma\) from the estimated posterior \(p(\theta, \Gamma|X^T)\).

In the Step 2, however, we cannot draw \(\theta\) and \(\Gamma\) simultaneously from the joint posterior \(p(\theta, \Gamma|X^T)\). But we can draw parameters separately from the following conditional posteriors:

\[p(\theta|\Gamma, X^T),\quad p(\Gamma|\theta, X^T),\]

Hence, by adopting the Gibbs sampler, we will draw structural parameter \(\theta\) and measurement equation parameter \(\Gamma\) from two conditional posteriors above and estimate the joint posterior \(p(\theta, \Gamma|X^T)\).

\(^5\)For the algorithms of the hybrid MCMC method, see chapter 6 of Gamerman and Lopes (2006) and chapter 6 of Nakatsuma (2003).
Furthermore, to evaluate the conditional posterior \( p(\Gamma|\theta, X^T) \), as will be illustrated in the Step 3 below, it must be separated into two conditional posteriors \( p(S^T|\Gamma, \theta, X^T) \) and \( p(\Gamma|S^T, \theta, X^T) \).

For sampling of state variables from the conditional posterior \( p(S^T|\Gamma, \theta, X^T) \), we adopt the simulation smoother. For sampling from the conditional posterior \( p(\Gamma|S^T, \theta, X^T) \), we use the Gibbs sampler. And for sampling from the conditional posterior \( p(\theta|\Gamma, X^T) \), we apply the MH algorithm. That is, the hybrid MCMC method is to mix appropriate algorithms according to parameters.

We can summarize the following five steps in estimating the state space model introducing ME by the hybrid MCMC method:

**Step 1:** Specify initial parameters \( \theta^{(0)} \) and \( \Gamma^{(0)} \), and set the iteration number to \( g = 1 \).

**Step 2:** Given structural parameter \( \theta^{(g-1)} \), we can solve the DSGE model using Sims (2002) method, derive the solution and obtain \( G(\theta^{(g-1)}), H(\theta^{(g-1)}) \) and \( Q(\theta^{(g-1)}) \).

**Step 3:** Draw \( \Gamma^{(g)} \) from \( p(\Gamma|\theta^{(g-1)}, X^T) \).

(3.1): Sampling unknown state variables \( S_t^{(g)} \) from \( p(S^T|\Gamma^{(g-1)}, \theta^{(g-1)}, X^T) \) by using the simulation smoother proposed by de Jong and Shephard (1995) and Durbin and Koopman (2002) (Details will be described in later).

(3.2): Using the sampled state variables \( S^{T(g)} \), drawing the measurement equation parameter \( \Gamma^{(g)} \) from \( p(\Gamma|S^{T(g)}, \theta^{(g-1)}, X^T) \) (Also, details will be explained in later)

**Step 4:** Draw structural parameter \( \theta^{(g)} \) from \( p(\theta|\Gamma^{(g)}, X^T) \) by the MH algorithm.

(4.1): Drawing the candidate \( \theta^{(\text{proposal})} \) from the proposal density \( p(\theta|\theta^{(g-1)}) \), and calculate the acceptance probability as follows.\(^6\)

\[
q = \min \left[ \frac{p(\theta^{(\text{proposal})}|\Gamma^{(g)}, X^T)p(\theta^{(g-1)}|\theta^{(\text{proposal})})}{p(\theta^{(g-1)}|\Gamma^{(g)}, X^T)p(\theta^{(\text{proposal})}|\theta^{(g-1)})}, 1 \right].
\]

(4.2): Accept the candidate \( \theta^{(\text{proposal})} \) with probability \( q \) and reject them with probability \( 1 - q \). If accepted, set \( \theta^{(g)} = \theta^{(\text{proposal})} \), and otherwise, set \( \theta^{(g)} = \theta^{(g-1)} \).

**Step 5:** Set the iteration number \( g = g + 1 \), and return to the Step 2. Repeat until \( g = G \) where \( G \) is the number of sampling.\(^7\)

The detailed calculation method of the Step 3 will be described in the following subsections. Here, we supplement the algorithms in the Steps 1 and 4.

On the initial values settings in the Step 1, as is well known, the MCMC algorithm will converge to the invariant distribution even if we start sampling from any initial values. In case with a large number of parameters, however, we can expect that a considerable number of sampling will be required for convergence. Hence, for the computational efficiency, we set initial values as follows: First, initial values of \( \theta^{(0)} \) can be obtained from the estimated posterior modes of structural parameters by estimating the DSGE model without ME. Given \( \theta^{(0)} \), we estimate state variables

\(^6\)When the solution based on structural parameters sampled from the proposal density becomes indeterminate or no solution, the parameters are not adopted. Sampling from the proposal shall be repeated until the parameters bring the uniqueness of the solution.

\(^7\)The number of sampling \( G \) is determined by referring to convergence diagnostics. We refer the CD (Convergence Diagnostic) statistics proposed by Geweke (1992).
by the simulation smoother. Then, we can derive initial values $S_t^{(0)}$, which are the initial values of the smoothed state variables. Finally, $S_t^{(0)}$ and $X^T$ are used to calculate the initial values of measurement equation parameters $\Gamma^{(0)}$.

Regarding the method of generating parameters from the proposed density in the Step 4, we adopt the random walk MH algorithm according to the previous studies. The candidate $\theta^{(\text{proposal})}$ is generated as follows:
\[
\theta^{(\text{proposal})} = \theta^{(g-1)} + u_t, \quad u_t \sim N(0, c \Sigma),
\]
where $\Sigma$ is variance covariance matrix of the random walk process and $c$ is the adjustment parameter. $\Sigma$ is set to the Hessian $(-l''(\hat{\theta}))$ of the log posterior ($l(\theta) = \ln p(\theta | \Gamma, X^T)$) which has already calculated when obtaining the initial values $\theta^{(0)}$.

Furthermore, when sampling with random walk MH, the detailed balanced condition holds:
\[
p(\theta^{(g-1)} | \theta^{(\text{proposal})}) = p(\theta^{(\text{proposal})} | \theta^{(g-1)})
\]
So, the acceptance probability $q$ can be reduced to the following formula.
\[
q = \min \left[ \frac{f(\theta^{(\text{proposal})})}{f(\theta^{(g-1)}), 1} \right],
\]
The equation above tells the acceptance probability does not depend on the proposal density $p(\theta | \theta^{(g-1)})$. As a result, an advantage of the random walk MH is no need to select a proposal density giving a good approximation of the posterior density. However, the candidate $\theta^{(\text{proposal})}$ should be not so far from the previous sample $\theta^{(g-1)}$, since if the deviation is so large, the acceptance probability $q$ declines, which would reduce the computational efficiency of MCMC. To prevent the decline of $q$, we should adjust the coefficient $c$ to a small value, but if $c$ is too small, we face another problem: The sampling range of $\theta^{(\text{proposal})}$ becomes too narrow. According to Roberts et al. (1997) and Neal and Roberts (2008), the optimal acceptance rate $q$ in random walk MH is about 25%. Therefore, we adjust the coefficient $c$ so that the acceptance probability becomes close to 25%.

### 2.3.3 Simulation Smoother

As mentioned in the Step 3, we adopt the simulation smoother proposed by de Jong and Shephard (1995), in sampling state variables from the conditional posterior $p(S^T | \Gamma^{(g-1)}, \theta, X^T)$.

As mentioned in Section 1, Boivin and Giannoni (2006) and Kryshko (2010) adopted the method of Carter and Koopman (2002) proposed another method of the simulation smoother. The advantage of Durbin and Koopman (2002) is no need to adopt new algorithms. To sample $\nu_t$ and $\varepsilon_t$, however, we have to implement the Kalman filter and the smoother for actual data $X$ and artificial data $X^+$ against the one sampling, thus, a total of two filtering and smoothing processes are required. This study handles a medium-scale DSGE model, so we decided...
and Kohn (1994) in smoothing state variables. However, as pointed out by Chib (2001, p. 3614), Carter and Kohn (1994) has a problem that sampling is interrupted by generating a non-positive definite matrix in drawing the variance covariance matrix of state variables. Hence, the previous studies set an ad hoc variance covariance matrix of state variables in the estimation. On the other hand, since this problem does not occur in the simulation smoother, the estimation method in this study can be said to be a more versatile method in estimating the DSGE model with ME.

To simply express the steps of the algorithm, the state space model of (2.15) and (2.16) is replaced by the following (2.18) and (2.19).

\[
\begin{align*}
\tilde{X}_t &= \tilde{\Lambda} \tilde{S}_t + \nu_t, \quad \nu_t \sim N(0, R), \\
\tilde{S}_t &= \tilde{G} \tilde{S}_{t-1} + \tilde{H} \varepsilon_t, \quad \varepsilon_t \sim N(0, Q(\theta)),
\end{align*}
\]

To conduct the simulation smoother, the Kalman filter is first implemented. The Kalman filter of this state space model consists of the following equations.

\[
\begin{align*}
\eta_t &= \tilde{X}_t - \tilde{\Lambda} \tilde{S}_{t|t}, & F_t &= \tilde{\Lambda} \tilde{P}_{t|t} \tilde{\Lambda}' + R, & K_t &= \tilde{G} \tilde{P}_{t|t} \tilde{\Lambda}' F_t^{-1}, \\
L_t &= \tilde{G} - K_t \tilde{\Lambda}, & \tilde{S}_{t+1|t+1} &= \tilde{G} \tilde{S}_{t|t} + K_t \eta_t, & \tilde{P}_{t+1|t+1} &= \tilde{G} \tilde{P}_{t|t} \tilde{L}_t + \tilde{H} Q(\theta) \tilde{H}',
\end{align*}
\]

where \( \eta_t \) is FE, \( K_t \) is the Kalman gain, \( \tilde{S}_t \) is state variable, \( \tilde{P}_t \) is covariance matrix of state variable. The filtering for \( \tilde{S}_{t|t} \) and \( \tilde{P}_{t|t} \) are proceeded sequentially forward \( (t = 1, 2, \ldots, T) \). Note that initial value \( \tilde{S}_{1|1}, \tilde{P}_{1|1} \) can be derived by solving \( \tilde{X}_1 = \tilde{\Lambda} \tilde{S}_{1|1}, \tilde{P}_{1|1} = \tilde{G} \tilde{P}_{1|1} \tilde{G}' + \tilde{H} Q(\theta) \tilde{H}' \). \(^{10}\)

The subscript \( t|t \) of \( S_{t|t} \) indicates the conditional expected value \( E(S_t | X_1, X_2, \ldots, X_t) \) given \( X_t \) up to the period \( t \).

Next, using the values obtained by the Kalman filter, the simulation smoother calculates \( r_{t-1} \) and \( N_{t-1} \) by proceeding the followings (2.20) and (2.21) to backward \( (t = T, \ldots, 2, 1) \).

\[
\begin{align*}
r_{t-1} &= \tilde{\Lambda}' F_t^{-1} \eta_t - W'_t C_t^{-1} d_t + L_t r_t, \\
N_{t-1} &= \tilde{\Lambda}' F_t^{-1} \tilde{\Lambda} + W'_t C_t^{-1} W_t + L'_t N_t L_t,
\end{align*}
\]

where \( W_t \) and \( C_t \) are derived from the following expressions.

\[
\begin{align*}
W_t &= Q(\theta) \tilde{H}' N_t L_t, \\
C_t &= Q(\theta) - Q(\theta) \tilde{H}' N_t \tilde{H} Q(\theta),
\end{align*}
\]

where \( d_t \) is drawn from \( N(0, C_t) \), and initial values are set to \( r_T = 0 \) and \( N_T = 0 \).

Using these values above, the smoothed structural shock \( \varepsilon_{t|T} \) can be calculated by conducting the following procedure to backward. The subscript \( t|T \) of \( \varepsilon_{t|T} \) represents the conditional expected value \( E(\varepsilon_t | X_1, X_2, \ldots, X_T) \) given \( X_T \) up to the total period \( T \).

\[
\varepsilon_{t|T} = Q(\theta) \tilde{H}' r_t + d_t, \quad d_t \sim N(0, C_t), \quad t = T, \ldots, 2, 1
\]

\(^{10}\)See section 6 of Anderson et al. (1996) and Anderson and Moore (1979) on a numerical solution for finding the initial value \( P_{1|1} \) from \( P_{1|1} = G \tilde{P}_{1|1} G' + H Q(\theta) H' \), a form of discrete Lyapunov equation (or Sylvester equation).
2.3. ESTIMATION METHOD

The generated structural shock $\hat{\varepsilon}_t$ is not only used for estimating state variables, but also for deriving the historical decompositions.\(^\text{11}\)

Finally, the estimated state variable by the simulation smoother can be calculated forwardly as follows.

$$
\hat{S}_{t+1|T} = \hat{G}S_{t|T} + H\hat{\varepsilon}_{t|T}, \quad t = 1, 2, \cdots, T,
$$

(2.22)

where the initial values are given by $\hat{S}_{1|T} = \hat{S}_{1|1} + \hat{P}_{1|1}r_0$. $\hat{S}_{t|T}$ ($t = 1, 2, \cdots, T$) becomes state variables drawn from the conditional posterior $p(S^T | \Gamma g^{-1}, \theta, X^T)$.

2.3.4 Measurement Equation Parameter

Regarding the sampling method of the measurement equation parameter $\Gamma = \{ R, \Psi \}$ from the conditional posterior $p(\Gamma | S^T g, \theta g^{-1}, X^T)$, we adopt the estimation method of the regression model with AR(1) disturbance term proposed by Chib and Greenberg (1994).

Since $R$ and $\Psi$ are assumed to be diagonal, given state variables $S^T$, we can calculate the ME in the $k$-th measurement equation as $e_{k,t} = X_{k,t} - \Lambda_k S_t$. $\Lambda_k$ is a $1 \times J$ vector which corresponds to $k$-th row of $\Lambda$. And the ME follows AR(1) process so that $e_{k,t} = \Psi_{kk}e_{k,t-1} + \nu_{k,t}$ where $\nu_{k,t} \sim$ i.i.d. $N(0, R_{kk})$. $R_{kk}$ and $\Psi_{kk}$ is the $k$-th diagonal component of $R$ and $\Psi$, respectively. Hence, for each measurement equation for $k = 1, \ldots, N$, $\Gamma$ can be independently sampled for each $R_{kk}$ and $\Psi_{kk}$. Proceeding two steps shown below, we can estimate the posteriors of $R_{kk}$ and $\Psi_{kk}$ by the Gibbs sampler, alternately drawing from the conditional posteriors: $p(R_{kk} | \Psi_{kk}, S^T, \theta, X^T)$ and $p(\Psi_{kk}, R_{kk}, S^T, \theta, X^T)$.

In the following, we focus on the $k$-th equation of the measurement equation to simplify the expression.

First, for each $(R_{kk}, \Psi_{kk})$, we set conjugate priors which leads to Normal-Inverse Gamma:

$$
p(R_{kk}, \Psi_{kk}) = p(R_{kk}) p(\Psi_{kk}) = IG(R_{kk} | s_0, \nu_0) \times N(\Psi_{kk} | \Psi_0, \sigma_{\Psi_0}^2)1_{||\Psi_{kk}|<1}
$$

where $IG$ is the inverse Gamma distribution, $1$ is the indicator function which reacts unity if $\{\}$ is true, and zero, otherwise. Following Kryshko (2010), we set prior distribution parameters to $s_0 = 0.001, \nu_0 = 3, \Psi_0 = 0, \sigma_{\Psi_0}^2 = 1, s_0 = 0.001, \nu_0 = 3, \Psi_0 = 0$ and $\sigma_{\Psi_0}^2 = 1$.

**Step 1:** Sampling $R_{kk}$ from the conditional posterior $p(R_{kk} | \Psi_{kk}, S^T, \theta, X^T)$.

The conditional posterior of $R_{kk}$ is proportional to the inverse Gamma. So we can write the conditional posterior as follows.

$$
p(R_{kk} | \Psi_{kk}, S^T, X^T) \propto p(X^T | S^T, R_{kk}, \Psi_{kk}, \theta)p(R_{kk}). \quad (2.23)
$$

\(^\text{11}\)The historical decomposition is derived from the following way: For the estimated structural shock vector $\hat{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{M1})$, components besides a certain shock $\varepsilon_{st}$, are set to zero, and substituting $(0, 0, \cdots, \varepsilon_{st}, \cdots, 0)$ into $\hat{\varepsilon}_t$ of (2.22), we can examine the influence of the shock for state variables.
The proof is as follows. Since the prior \( p(R_{kk}) \) in the right-hand side of (2.23) is specified by the inverse Gamma,
\[
p(R_{kk}) = \frac{(s_0/2)^{\nu_0/2}}{\Gamma(\nu_0/2)} (R_{kk})^{-\nu_0/2-1} \exp\left(-\frac{s_0}{2R_{kk}}\right)
\times (R_{kk})^{-\nu_0/2-1} \exp\left(-\frac{s_0}{2R_{kk}}\right),
\]
where \( \nu_0 \) and \( s_0 \) are parameters of the prior.

Next, let us derive the conditional likelihood \( p(X^T|S^T, R_{kk}, \Psi_{kk}, \theta) \) in the right-hand side of (2.23). We define:
\[
X_{k,t}^* = X_{k,t} - \Psi_{kk}X_{k,t-1}, \quad S_t^* = S_t - \Psi_{kk}S_{t-1}
\]
Also, we denote \( X_k^* \) and \( S^* \) as \( X_k^* = [X_k^*_1, \ldots, X_k^*_T]' \) (\( T \times 1 \) vector) and \( S^* = [S^*_1, \ldots, S^*_T]' \) (\( T \times J \) matrix). Since \( \Lambda_k \) is known and \( X_k^*, S^* \) are given, i.i.d. shock of measurement equation \( \nu_k \) can be derived as follows.
\[
X_k^* = S^*\Lambda_k' + \nu_k \iff \nu_k = X_k^* - S^*\Lambda_k'.
\]
Hence, the conditional likelihood \( p(X^T|S^T, R_{kk}, \Psi_{kk}, \theta) \) can be expressed by:
\[
p(X^T|S^T, R_{kk}, \Psi_{kk}, \theta) = \left(\frac{1}{\sqrt{2\pi R_{kk}}}\right)^n \exp\left\{-\frac{1}{2R_{kk}} \sum_{t=1}^{T} (X_{k,t}^* - S_t^*\Lambda_k')^2\right\}
\times (R_{kk})^{-T/2} \exp\left(-\frac{\nu_k^T\nu_k}{2R_{kk}}\right)
\]
where \( \nu_k'\nu_k = \sum_{t=1}^{T} (X_{k,t}^* - S_t^*\Lambda_k')^2. \)

Therefore, we can confirm the conditional posterior \( R_{kk} \) is proportional to the inverse Gamma distribution. Thus, we can draw the variance of ME from the conditional posterior of \( R_{kk} \):
\[
R_{kk}|\Psi_{kk}, S^T, \theta, X^T \sim IG(\bar{s}, \bar{\nu}), \quad (2.24)
\]
where \( \bar{s} = s_0 + \nu_k^T\nu_k \), \( \bar{\nu} = \nu_0 + T. \)

**Step 2:** Sampling \( \Psi_{kk} \) from the conditional posterior \( p(\Psi_{kk}|R_{kk}, S^T, \theta, X^T) \)

We can write the conditional posterior \( \Psi_{kk} \) as follows:
\[
p(\Psi_{kk}|R_{kk}, S^T, \theta, X^T) \propto p(X_k^T|S^T, R_{kk}, \Psi_{kk}, \theta)p(\Psi_{kk}),
\]
The ME \( e_{k,t} \) of the likelihood \( p(X_k^T|S^T, R_{kk}, \Psi_{kk}, \theta) \) in the right-hand side of the equation above becomes:
\[
e_{k,t} = X_{k,t} - \Lambda_kS_t,
\]
Let us denote \( e_k = [e_{k,2}, \ldots, e_{k,T}]' \) and \( e_{k,-1} = [e_{k,1}, \ldots, e_{k,T-1}]' \), then the ME can be represented by:
\[
e_k = e_{k,-1}\Psi_{kk} + \nu_k,
\]
2.4 Preliminary Settings and Data

Using $\hat{\Psi}_{kk} = (e'_{k-1}e_{k-1})^{-1}e'_ke_k$, the likelihood can be written by:

$$p(X^T_k | S^T, R_{kk}, \Psi_{kk}, \theta) \propto \exp \left[ -\frac{1}{2R_{kk}} (\Psi_{kk} - \hat{\Psi}_{kk})' e'_{k-1}e_{k-1}(\Psi_{kk} - \hat{\Psi}_{kk}) \right],$$

Since the prior of $\Psi_{kk}$ is the truncated Normal with mean $\bar{\Psi}_0$ and variance $\sigma^2_{\Psi,0}$, we can express:

$$p(\Psi_{kk}) \propto \exp \left[ -\frac{1}{2\sigma^2_{\Psi,0}} (\Psi_{kk} - \bar{\Psi}_0)^2 \right] \times 1_{|\Psi_{kk}| < 1},$$

Thus, the conditional posterior becomes:

$$p(\Psi_{kk} | R_{kk}, S^T, \theta, X^T) \propto \exp \left[ -\frac{1}{2R_{kk}} (\Psi_{kk} - \hat{\Psi}_{kk})' e'_{k-1}e_{k-1}(\Psi_{kk} - \hat{\Psi}_{kk}) \right]$$

$$\times \exp \left[ -\frac{1}{2\sigma^2_{\Psi,0}} (\Psi_{kk} - \bar{\Psi}_0)^2 \right] \times 1_{|\Psi_{kk}| < 1},$$

That is, we can see the conditional posterior density of $\Psi_{kk}$ is proportional to the product of two densities of Normal distribution. Hence, the AR(1) coefficient of ME $\Psi_{kk}$ can be drawn from the truncated Normal:\textsuperscript{12}

$$\Psi_{kk} | R_{kk}, S^T, \theta, X^T \sim N(\bar{\Psi}_{kk}, \bar{\nu}_{kk}) \times 1_{|\Psi_{kk}| < 1}, \quad (2.25)$$

where

$$\bar{\nu}_{kk} = [(R_{kk}(e'_{k-1}e_{k-1})^{-1})^{-1} + (\sigma^2_{\Psi,0})^{-1}]^{-1},$$

$$\bar{\Psi}_{kk} = \bar{\nu}_{kk} [(R_{kk}(e'_{k-1}e_{k-1})^{-1})^{-1}\hat{\Psi}_{kk} + (\sigma^2_{\Psi,0})^{-1}\bar{\Psi}_0].$$

2.4 Preliminary Settings and Data

2.4.1 Observation Variables and Measurement Error

Similar to SW (2003, 2007) and CEE (2005), we choose seven state variables as observation variables $X$: (1) output: $y_t$, (2) consumption: $c_t$, (3) investment: $inv_t$, (4) labor: $l_t$, (5) real wage: $w_t$, (6) inflation: $\pi_t$, and (7) nominal interest rate: $R_t$. These seven observation variables have a one-to-one correspondence with seven state variables. Usually, in estimating the DSGE model without considering ME ($\epsilon_{it}$), observation variable $X_{it}$ and state variable $S_{it}$ were assumed to be identical ($X_{it} = S_{it}$). Instead, we assume the existence of ME in the data, i.e. we set $X_{it} = S_{it} + \epsilon_{it}$. However, the nominal interest rate can be controlled directly by the central bank through the monetary policy. So, ME in the nominal interest rate should be excluded, i.e. $\epsilon_{it}^R = 0$. Hence, MEs are introduced into six observation variables excluding nominal interest rate.

In order to make clear the problem of markup shocks, it may think of an idea to estimate by simultaneously introducing markup shocks and MEs. However, due to the structure of the model, it is difficult to identify those disturbance terms: Recall the NKPC (2.8).

$$\hat{\pi}_t = \frac{1}{1 + \beta\gamma_p} E_t \hat{\pi}_{t+1} + \frac{\gamma_p}{1 + \beta\gamma_p} \hat{\pi}_{t-1} + \frac{(1 - \beta\gamma_p)(1 - \xi_p)}{(1 + \beta\gamma_p)\xi_p} \left[ \alpha \hat{\pi}_t^k + (1 - \alpha)\hat{w}_t - a_t^q \right] + \epsilon_t^p$$

\textsuperscript{12}If $|\Psi_{kk}| \geq 1$ is drawn, then sampling is repeated until the stationary condition $|\Psi|_{kk} < 1$ holds.
where $\varepsilon_p^t$ is the price markup shock. Let $X_{\pi,t}$ and $e_{\pi,t}$ denote inflation data and ME attached to the inflation data, respectively. Then, the measurement equation on inflation can be written as the following expression.

$$X_{\pi,t} = \hat{\pi}_t + e_{\pi,t}$$

We can see that both markup shock $\varepsilon_p^t$ and ME $e_{\pi,t}$ are hanging in the same way for the inflation data $X_{\pi,t}$ (the coefficients of both disturbance terms are unity). The wage markup shock also has the similar problem. Consequently, wage and price markup shocks, $\varepsilon_w^t$ and $\varepsilon_p^t$, are difficult to identify from MEs. So we exclude those two markup shocks when estimating the model with ME. Thus, unlike SW (2003, 2007) and CEE (2005), state variables fluctuations such as wage and inflation are caused by variations in MEs or structural shocks except for markup shocks.\footnote{Note that Justiniano and Primiceri (2008) and Fueki et al. (2010) also replace markup shocks with MEs. However, it differs in that they adopt the MH algorithm but we adopt the hybrid MCMC method with the simulation smoother.}

Our motivation, as emphasized in Section 1, is to consider how much introducing ME affects estimated state variables (especially, wage and inflation) and estimated structural parameters (especially, nominal rigidities). Hence, we will estimate and compare the two cases: “case w/o ME” and “case with ME”. In “case w/o ME”, the DSGE model without ME is estimated which corresponds to normal estimation method as in SW (2003, 2007). On the other hand, in “case with ME”, we estimate the DSGE model introducing the ME. Again, it should be noted that “the case with ME” excludes the two markup shocks and “case w/o ME” includes markup shocks as structural shocks.

### 2.4.2 Calibrated Parameters and Prior Settings

We assume that the seven structural shocks, sources of the business cycle, are independent of each other. Also, the two structural shocks of equity premium shock $\varepsilon_q^t$ and monetary policy shock $\varepsilon_m^t$ are assumed to be i.i.d. shocks. The remaining five structural shocks are assumed to have inertia and follow the AR(1) process: Preference shock $\varepsilon_c^t$, TFP shock $\varepsilon_z^t$, investment adjustment cost shock $\varepsilon_{inv}^t$, labor supply shock $\varepsilon_L^t$, and government expenditure shock $\varepsilon_t^t$. As shown in Table 2.1, the prior for persistence parameter $\rho$ of the AR shock process is specified as the Beta distribution so as to satisfy the stationary condition $\rho \in (0, 1)$, with mean of 0.85, standard error of 0.10, which is a relatively strong prior distribution. Also, since the variance of i.i.d. shock $\varepsilon_t$ is a positive value, the prior is specified as inverse Gamma distribution. The parameters regions in which the solution is not uniquely determined are excluded from the prior distribution.

As with SW (2003, 2007), Levin et al. (2005) and Onatski and Williams (2010), some structural parameters need to be calibrated in advance. We will calibrate parameters according to the previous studies on Japan, the U.S. and Europe.

First, the subjective discount factor $\beta$ is set to 0.99, which means that the steady state of real interest rate is assumed to be 4% at annual rate. The capital depreciation rate $\tau$ is 0.025 per quarter, assuming 10% when converting on an annual basis. From the above setting, the steady state of real rental rate of capital can also be calculated as $\bar{r}_k = \frac{1}{\beta} - (1 - \tau)$. The capital income share $\alpha$ is 0.30, which implies the steady state of labor income share is 70%. The steady state of government expenditure-output ratio $g_y$ is 0.10, and the steady of capital-output ratio $k_y$ is set to 1.50. In addition, because of identification problem, we set the wage markup rate parameter $\lambda_w$ to 0.05 according to Onatski and Williams (2010).

Table 2.1 shows the preliminary settings on the remaining structural parameters. The prior mean is set mainly in accordance with SW (2003), and the standard deviation is set so that the
parameter value covers a reasonable range. For example, the prior means for price and wage Calvo parameters $\xi_p$ and $\xi_w$ roughly follow the estimation result of Gali, et al. (2001) that the average contract period of price and wage is one year so set to 0.75. On the other hand, its standard deviation is set to 0.15 so that the contract period can vary from three quarters to two years. Similarly, the prior mean for the IES ($\sigma_c$) is set to unity. The prior mean of elasticity on capital utilization adjustment cost $\psi$ is set to 0.2, and we set the standard deviation, in which the elasticity can fluctuate up to 0.1, the value reported by King and Rebelo (1999). The prior mean of fixed cost share $\phi$ is 1.45, which is set to the value close to CEE (2005). On the inverse Frisch elasticity ($\sigma_L$), the prior mean is set to 2, but the standard deviation is set up so as to cover a wide range of low values reported in microeconomic evidence up to high values reported in estimation results of DSGE models. Finally, with regard to the prior mean of Taylor coefficients, to guarantee the uniqueness of the solution, parameter $\mu_\pi$ related to inflation is 1.70, and we set 0.8 as the prior mean of the interest rate smoothing parameter $\rho_m$ and 0.125 as the coefficient $\mu_y$ on the output gap.

2.4.3 Data

We use Japanese macroeconomic quarterly data and the estimation period, following Sugo and Ueda (2008), is 1981:Q1 to 1995:Q4 (15 years), excluding the period of the second oil shock and zero interest rate policy. The reason for limiting to this period is based on the fact that the monetary policy rule is linear in the standard DSGE model.

We employ the following data which corresponds to seven observation variables $X_t$: (1) output $y_t$ is real GDP per capita (per unit is one million yen, base year is 1990, seasonally adjusted), (2) consumption $c_t$ is real consumption per capita (unit and others are the same as GDP) calculated as the nominal private final consumption expenditure divided by the GDP deflator and the labor force population, (3) investment $inv_t$ is real investment per capita (unit and others are the same as GDP) calculated as nominal private capital investment divided by GDP deflator and by labor force, (4) labor $l_t$ is a calculated series so that the product of the working hour index and the total employment is divided by the labor force, (5) real wage $w_t$ is a calculated real wage index obtained from dividing nominal wage index by GDP deflator, (6) inflation $\pi_t$ is an annualized growth rate of GDP deflator, and (7) nominal interest rate $R_t$ is annualized uncollateralized call rate. On the other hand, capital stock $K_t$ and capital shadow price $q_t$ are regarded as unobservable state variables as in SW (2003).

We detrend five real series of output $y_t$, consumption $c_t$, investment $inv_t$, labor $l_t$, real wage $w_t$ by taking natural log and removing trend components using Hodrick-Prescott filter. Then, by multiplying the series by 100, we derive the percent deviation from steady states. The two series with percent displayed values of nominal interest rate $R_t$ and inflation $\pi_t$ are detrended by the Hodrick-Prescott filter. It should be noted that there is one drawback by employing the Hodrick-Prescott filter: Detrending the data inconsistent with the balanced growth model. Del Negro et al. (2007) assumed that output, consumption, investment, real wage, capital have a common stochastic trend accompanying technological progress growth rate and simultaneously estimated not only structural parameters but also the technological progress growth rate consistent with the balanced growth theory. When trends are removed by the Hodrick-Prescott filter independently for output, consumption, and investment, trend components are not necessarily common, which could seem not to be desired detrending method since it is inconsistent with the balanced growth theory. However, when estimating the model using Japan’s data, Watanabe and Iiboshi (2007) and Iiboshi (2011) reported that trend changes, that is, structural breaks occurred in the early 1990s. Thus,
we should not apply the method of removing trend components by Del Negro et al. (2007) or SW (2007). Instead, we extract the cycle component of each data by removing the time-varying trend component by the Hodrick-Prescott filter without handling the trend break. Finally, all data was demeaned to make the means zeros. The solid line in Figure 2.1 displays the data $X_t$ used in our estimation.

2.5 Results

This section reports estimation results of the model incorporating ME and compare it with the result of the conventional model without ME. Specifically, we estimate the case w/o ME (using the MH algorithm) and the cases with ME (adopting the hybrid MCMC method), and report estimated parameters, smoothed state variables, historical decompositions and the result of model selection.

In estimating the posterior of parameters, state variables, etc., 300,000 replicates were generated in both the MH algorithm and the hybrid MCMC method. Discarding the first 100,000 replicates as burn-in, required statistics of the posterior are calculated from 200,000 replicates: The posterior mean, the standard deviation (hereinafter, SD), the 90% credible interval (hereinafter, CI), and the standard error (hereinafter, SE) of the posterior mean.

2.5.1 Structural Parameters

First, we report estimated structural parameters in cases with and w/o ME (the posterior mean, SE, SD, and 90% CI). Tables 2.2 (a), (b) show substantial differences in several estimated parameters between the two cases: In particular, the price indexation ($\gamma_p$), the wage indexation ($\gamma_w$) and the Calvo parameter on wage ($\xi_w$). Those posterior means of the case with ME are outside the 90% CI of the case w/o ME. We can see short-term volatile fluctuations on price and wage data, which might cause remarkable differences in estimation results between the case w/o ME where it is explained by markup shocks and the case with ME where it is captured by MEs. We can also find that there are no big differences between two cases in estimated parameters such as $h$, $\sigma_c$, $\sigma_L$, $1/\varphi$, $\phi$ and $\psi$. Those parameters are related to data on output, consumption, investment, and labor. We can confirm those data are relatively stable and persistent fluctuations: Hence, whether the ME affects the estimation result may depend on whether the data fluctuates noisy or not.

Of course, another possibility of causing the difference should be the estimation bias attributable to omitting the ME. In the case where the ME actually exists, if structural parameters are estimated ignoring its existence, the estimates must be biased.

2.5.2 Standard Error, Convergence Diagnostics, Inefficiency Factor

To derive SE of the posterior mean ($SE(\tilde{\theta})$), following Kim, et al. (1998) and Fujiwara and Watanabe (2011), we adopted the calculation method by Parzen window.

$$SE(\tilde{\theta}) = \sqrt{\frac{\gamma(0)}{M}} \left[ 1 + 2 \frac{M}{M - 1} \sum_{i=1}^{BM} K \left( \frac{i}{BM} \right) \frac{\gamma(i)}{\gamma(0)} \right]$$

(2.26)
where $M$ is the number of sampling, $\hat{\gamma}(i)$ is $i$-th order covariance, $K(z)$ is the kernel of Parzen window shown below, where $B_M$ is the band width and set to $B_M = 0.01M$.

\[
\hat{\gamma}(i) = \frac{1}{M} \sum_{k=i+1}^{M} (\theta^{(k)} - \bar{\theta})(\theta^{(k-i)} - \bar{\theta}),
\]

\[
K(z) = \begin{cases} 
1 - 6z^2 + 6z^3, & z \in [0, \frac{1}{2}] \\
2(1-z)^3, & z \in \left[\frac{1}{2}, 1\right]
\end{cases}
\]

Using the SE described above, when judging convergence to invariant distributions in the MCMC method, we refer the following convergence diagnostics (hereinafter, CD) statistics proposed by Geweke (1992).

\[
CD = \frac{\bar{\theta}_A - \bar{\theta}_B}{\sqrt{(SE(\bar{\theta}_A))^2 + (SE(\bar{\theta}_B))^2}}
\]  

(2.27)

where $\bar{\theta}_A$ is the sample mean calculated using the first 20,000 replicates (the first 10% of total samples after discarding), and $\bar{\theta}_B$ is the sample mean derived using the last 100,000 replicates (50% of total samples). $SE(\bar{\theta}_A)$ and $SE(\bar{\theta}_B)$ are calculated by the method of the Parzen window, and the band widths are 200 and 1000, respectively.

The CD statistics becomes asymptotically the standard Normal under the null hypothesis is true that the posterior converges. According to the convergence judgment by the CD statistics, in both cases, the null hypothesis is accepted as 5% significance level with some exceptions such as parameters related to labor supply.

Also, following Fujiwara and Watanabe (2011), to assess how inefficient the sampling method is compared with the random sampling, we calculate inefficiency factor (hereinafter, IF) for each parameter.

\[
IF = 1 + 2 \sum_{i=1}^{\infty} \hat{\rho}(i)
\]  

(2.28)

where $\hat{\rho}(i)$ is $i$-th order autocorrelation of parameters drawn by the MCMC. Let the variance and $i$-th order autocorrelation denote as $\sigma^2$ and $\rho(i)$. Then, the variance of sample mean becomes $\sigma^2 \left(1 + 2 \sum_{i=1}^{\infty} \rho(i) \right) / M$. On the other hand, since there is no autocorrelation in the random sampling, the variance of the sample mean is $\sigma^2 / M$. The IF, the ratio of both, represents the inefficiency of sampling of MCMC against random sampling, and the higher IF indicates the sampling inefficiency by MCMC is high. Since sample autocorrelation up to the infinite order cannot be calculated, we use the calculation method by the Parzen window when calculating IF.

The IF shows high values in both cases, and in particular, parameters related to labor, $\sigma_L$ and $\varepsilon_L$. According to the estimation results of the SW model using the random walk MH algorithm by Chib and Ramamurthy (2010), the sampled parameters of $\sigma_L$ showed high autocorrelation and that IF became more than 2500.

Turning to the nominal rigidities represented by the Calvo parameters $\xi_p$ and $\xi_w$, in the case w/o ME, IFs are 712 and 791 respectively, and in the case with ME, IFs are 579 and 522, respectively. Thus, the case with ME is efficiently sampled.
2.5.3 State Variables

Next, we consider the smoothed state variables in the case with ME. It should be noted that, in the case w/o ME, since observation variables and state variables are definitively connected, state variables cannot be smoothed.

Figure 2.1 shows the smoothed state variables of output, consumption, investment, labor, real wage and inflation.\textsuperscript{14}

As can be seen from Figure 2.1, in output, consumption, investment and labor, there is not a noticeable difference between the observed data and the smoothed state variables, but in real wages and inflation the difference is substantial, which implies that the influence of ME is relatively small in output, consumption, investment and labor, and the influence of ME is relatively large in real wage and inflation.

In the case w/o ME such as SW (2003, 2007) and Sugo and Ueda (2008), the high frequency movements of real wage and inflation data are grasped by price and wage markup shocks, i.e. structural shocks. In the case where ME is taken into account, most of these violent fluctuations are captured by MEs, not by structural shocks, since state variables of wage and inflation are remarkably smoothed. The result is in contrast to the standard DSGE model which has regarded real wage and inflation as a volatile state variables.

Furthermore, it will naturally affect estimated structural parameters related to price and wage whether to regard wage and inflation as persistent state variables or as volatile state variables. As pointed out earlier, it seems to have appeared as a difference between the estimated posteriors between the two cases of price and wage indexations and the Calvo parameter on wage.

In fact, in the case w/o ME, the posterior mean of price indexation is 0.304, while in the case with ME the posterior mean is 0.817. Thus, the inflation was regarded as more persistent state variable in the case with ME. On wage indexation, the posterior mean is 0.395 in the case w/o ME, while 0.678 in the case with ME. Regarding the Calvo parameter on wage, the posterior mean in the case with ME is 0.685, which is higher than 0.355 in the case w/o ME, and the wage is captured as a more sticky state variable in the case with ME. Therefore, smoothed real wage and inflation in the case with ME are consistent with the estimated parameters reported in Table 2.2 (b).

2.5.4 Historical Decompositions

Let us turn to historical decompositions. The historical decomposition in the case w/o ME is an attempt to evaluate the contribution of each structural shock and explain all the historical data movements by structural shocks. Especially, we consider historical decompositions focusing on the inflation data and real wage data where the influence of ME seems to be relatively large.

The upper part (a) in Figure 2.2 shows the historical decomposition of inflation in the case w/o ME, and the lower part (b) corresponds to the case with ME. Here, in historical decomposition in the case w/o ME (the upper part), for convenience, the contributions of markup shocks have been removed. In the cases with ME (the lower part), the contribution of ME is also eliminated. This is to make it easy to compare contributions by seven common structural shocks in both cases.\textsuperscript{15} The dashed line in Figure 2.2 shows the sum of contributions of seven structural shocks.

Focusing on the dashed line in both cases, the basic movement of inflation is roughly the same in the two cases. But in the case w/o ME, even if the influences of markup shocks are removed,

\textsuperscript{14} The nominal interest rate is also an observation variable, but it is not shown in the Figure 2.1, since the data of nominal interest rate is assumed to have no ME.

\textsuperscript{15} The seven common structural shocks are preference shock, productivity shock, investment adjustment cost shock, equity premium shock, labor supply shock, TFP shocks, government expenditure shock, and monetary policy shock.
noisy fluctuations still remain. By contrast, in the case with ME, it is markedly smoothed. Looking at contributions due to each structural shock, we can find big differences among two cases. In the case w/o ME, contributions mainly are brought by preference shock, monetary policy shock, and labor supply shock, whereas in the case with ME, mainly due to investment adjustment cost shock and TFP shock. In sum, both cases are generally in agreement with the underlying movement of inflation, but which structural shock has explanatory power over inflation variations has large differences.

Historical decompositions on real wage are depicted in Figure 2.3. Focusing on the dashed line in both cases, in the case w/o ME, we can see the movement relatively close to the actual wage data, whereas in the case with ME it has been markedly smoothed and almost all the volatile actual wage fluctuations are captured by ME. Looking at contributions due to each structural shock, contributions of preference shock and labor supply shock are main sources in the case w/o ME, and contribution of monetary policy shock and investment adjustment cost shock can be also recognized. In particular, labor supply shock is so volatile, which might make actual wage’s baseline movement volatile in cases w/o ME. By contrast, in case with ME, smoothed wage movements are mainly explained by investment adjustment cost shock and TFP shock. In addition, contributions by labor supply shock are hardly recognized. In summary, the underlying movements of wage are different in both cases and the historical decomposition (especially, contributions of labor supply shock) also changed significantly.

2.5.5 Model Selection

Is the model with ME is better than the model w/o ME? To examine this, we conduct the model selection by the Bayes factor, $B_{10}$, in the following formula.

$$B_{10} = \frac{p(Y | M_1)}{p(Y | M_0)}$$

where $p(Y | M_i)$ is the marginal likelihood of the model $M_i$. On the calculation method of the marginal likelihood using MCMC, we adopt the modified harmonic mean method of Geweke (1999).\footnote{The marginal likelihood $p(Y|M)$ of the model $M$ in the modified harmonic mean method (Geweke,1999) can be calculated by using parameters $(\theta_1, \theta_2, \cdots, \theta_N)$ sampled from the posterior $p(\theta|Y,M)$, according to the following way:

$$p(Y|M) \approx \left[ \frac{1}{N} \sum_{i=1}^{N} \frac{C_g g(\theta_i)}{p(Y|\theta_i,M) \times C_p p(\theta_i|M)} \right]^{-1}$$

where $N$ is the number of sample, $p(Y|\theta,M)$ is the likelihood, $p(\theta|M)$ is the prior, and $g(\theta_i)$ is an any probability density. $C_p$ and $C_g$ are scaling parameters for $p(\theta|M)$ and $g(\theta_i)$. Under any function $g(\theta)$, the modified harmonic mean method by Geweke (1999) employs the following truncated Normal distribution.

$$g(\theta) = \tau^{-1}(2\pi)^{-k/2} |\Sigma|^{-1/2} \exp \left[ -\frac{1}{2} (\theta - \mu)' \Sigma^{-1} (\theta - \mu) \right] \times \mathbf{I} \left[ (\theta - \mu)' \Sigma^{-1} (\theta - \mu) \leq F_{\chi^2_k}^{-1}(\tau) \right]$$

where $k$ is the number of parameters $\theta$, $\mathbf{I}$ is the indicator function that becomes unity if the value in the parentheses is true, and zero, otherwise. $F_{\chi^2_k}^{-1}(\tau)$ is the inverse function of the cumulative $\chi^2$ distribution with degree of freedom $k$. The area of the cumulative density truncated is set to $\tau = 0.95$.

It should be noted, however, in calculating the marginal likelihood of the DSGE model, we must pay attention to the prior $p(\theta_i|M)$ and the scaling parameters $C_p$, $C_g$ of the multivariate Normal $g(\theta)$. In other words, the parameter...}
Table 2.3 shows the log marginal likelihood, its SE and the log Bayes factor \( \log(B_{10}) \). The log Bayes factor for the model with ME against the model w/o ME reveals an extremely large, about 80. Kass and Raftery (1995) suggests if the log Bayes factor exceeds five, the model can be supported very strongly against the other model to be compared from the view of the fit to data.\(^{17}\) Therefore, adding ME to data greatly contributes to increasing the estimation and prediction accuracy of the SW model.

Why the model with ME is superior to the model w/o ME in the marginal likelihood?

The reason is whether to capture the noisy data by structural shocks or by MEs. In the case w/o ME, the data fluctuations are caused by markup shocks, but because markup shocks are structural shocks, shocks variations affect not only wage or inflation, but also other state variables. As markup shocks undertake volatile data movements of real wage and inflation, it will force volatile fluctuations to other state variables as well, resulting in a sacrifice of prediction accuracy as a whole model. On the other hand, in the case with ME, since the volatile data are regarded as MEs, there are no influences for other state variables. For this reason, it is possible to smooth data with high inertia such as output, consumption, investment and labor without forcing volatile movements, resulting in no sacrifice of prediction accuracy as a whole model.

\(^{17}\)In Jeffreys (1961, Appendix B, p. 432), the criteria for model selection by the Bayes factor are shown as follows.
2.6 Conclusion

We propose a generalized estimation method of DSGE model with ME by employing the hybrid MCMC method with the simulation smoother. Boivin and Giannoni (2006) and Kryshko (2010) also proposed the hybrid MCMC method for DSGE model assuming ME with AR (1) process and adopting Carter and Kohn’s (1994) smoothing method. However, there is a problem that sampling is interrupted by generating a non positive definite matrix in drawing variance covariance matrix of state variable according to Carter and Kohn’s (1994) smoothing method. Instead, we can overcome the problem by adopting the simulation smoother proposed by de Jong and Shephard (1995) as a smoothing method of state variables.

Then, we estimate, on Japanese data, the DSGE model with the MEs (using the hybrid MCMC method) and compares it with the result of the DSGE model w/o MEs (using the MH algorithm). From the comparison of the two cases, we find non-negligible differences in estimated structural parameters, structural shocks, and historical decompositions. Furthermore, when we conduct model selection based on the Bayes factor, the model with MEs is strongly supported as compared with the model w/o ME. This indicates that the fit for data and the prediction accuracy by the model with MEs is higher than the model w/o MEs.

As in the conventional DSGE model (w/o MEs), the approach to capture the high frequency movements in inflation or wage by markup shocks implies to push them to structural shocks’ responsibilities, which affects another state variables such as output. On the other hand, the DSGE model incorporating MEs grasps them as ME’s variations without affecting other state variables. This difference may cause the remarkable discrepancy on explanatory power for data between two (marginal likelihoods).

As reported by Stock and Watson (2007), Cogley et al. (2010), Edge and Gurkaynak (2010), inflation shows noisy behavior in recent years and its prediction is becoming difficult. SW model captured the high frequency fluctuations by markup shocks but the shocks cannot be interpreted structurally and the fit of model for data significantly reduces. Given the current situation, the noisy movement of inflation or wage seems to be more realistic by regarding not as structural shocks but as MEs.
### 2.7 Tables and Figures

#### Table 2.1 Prior Distribution

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meanings</th>
<th>Shape</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural Parameter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h$</td>
<td>Habit formation</td>
<td>Beta</td>
<td>0.700</td>
<td>0.100</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>IES</td>
<td>Normal</td>
<td>1.000</td>
<td>0.375</td>
</tr>
<tr>
<td>$\sigma_L$</td>
<td>Inverse Frisch elasticity</td>
<td>Normal</td>
<td>2.000</td>
<td>0.750</td>
</tr>
<tr>
<td>$1/\varphi$</td>
<td>Investment adj. cost</td>
<td>Normal</td>
<td>4.000</td>
<td>1.500</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Fixed cost</td>
<td>Normal</td>
<td>1.450</td>
<td>0.250</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Capital utilization cost</td>
<td>Normal</td>
<td>0.200</td>
<td>0.075</td>
</tr>
<tr>
<td>$\gamma_p$</td>
<td>Price indexation</td>
<td>Beta</td>
<td>0.750</td>
<td>0.150</td>
</tr>
<tr>
<td>$\gamma_w$</td>
<td>Wage indexation</td>
<td>Beta</td>
<td>0.750</td>
<td>0.150</td>
</tr>
<tr>
<td>$\xi_p$</td>
<td>Calvo (price)</td>
<td>Beta</td>
<td>0.750</td>
<td>0.150</td>
</tr>
<tr>
<td>$\xi_w$</td>
<td>Calvo (wage)</td>
<td>Beta</td>
<td>0.750</td>
<td>0.150</td>
</tr>
<tr>
<td><strong>Policy Parameter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_m$</td>
<td>Interest rate smoothing</td>
<td>Beta</td>
<td>0.800</td>
<td>0.100</td>
</tr>
<tr>
<td>$\mu_\pi$</td>
<td>Taylor (inflation)</td>
<td>Normal</td>
<td>1.700</td>
<td>0.100</td>
</tr>
<tr>
<td>$\mu_y$</td>
<td>Taylor (output)</td>
<td>Normal</td>
<td>0.125</td>
<td>0.050</td>
</tr>
<tr>
<td><strong>Shock Persistence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_z$</td>
<td>TFP shock</td>
<td>Beta</td>
<td>0.850</td>
<td>0.100</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>Preference</td>
<td>Beta</td>
<td>0.850</td>
<td>0.100</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>Gov. expenditure shock</td>
<td>Beta</td>
<td>0.850</td>
<td>0.100</td>
</tr>
<tr>
<td>$\rho_L$</td>
<td>Labor supply shock</td>
<td>Beta</td>
<td>0.850</td>
<td>0.100</td>
</tr>
<tr>
<td>$\rho_{inv}$</td>
<td>Inv. adj. cost shock</td>
<td>Beta</td>
<td>0.850</td>
<td>0.100</td>
</tr>
<tr>
<td><strong>Shock SD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_c$</td>
<td>Preference shock</td>
<td>Inverse Gamma</td>
<td>0.200</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_{inv}$</td>
<td>Inv. adj. cost shock</td>
<td>Inverse Gamma</td>
<td>0.100</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_q$</td>
<td>Equity premium shock</td>
<td>Inverse Gamma</td>
<td>0.400</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_z$</td>
<td>TFP shock</td>
<td>Inverse Gamma</td>
<td>0.400</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_p$</td>
<td>Price markup shock</td>
<td>Inverse Gamma</td>
<td>0.150</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_L$</td>
<td>Labor supply shock</td>
<td>Inverse Gamma</td>
<td>1.000</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_w$</td>
<td>Wage markup shock</td>
<td>Inverse Gamma</td>
<td>0.250</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_g$</td>
<td>Gov. expenditure shock</td>
<td>Inverse Gamma</td>
<td>0.300</td>
<td>2.000</td>
</tr>
<tr>
<td>$\varepsilon_m$</td>
<td>Monetary policy shock</td>
<td>Inverse Gamma</td>
<td>0.100</td>
<td>2.000</td>
</tr>
</tbody>
</table>
Table 2.2 (a) W/O MEs: Estimation Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SE</th>
<th>SD</th>
<th>90 % IC</th>
<th>CD</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural Parameter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>0.544</td>
<td>0.003</td>
<td>0.089</td>
<td>[0.400 0.692]</td>
<td>1.489</td>
<td>251</td>
</tr>
<tr>
<td>σ_c</td>
<td>1.498</td>
<td>0.010</td>
<td>0.282</td>
<td>[1.043 1.964]</td>
<td>0.662</td>
<td>255</td>
</tr>
<tr>
<td>σ_L</td>
<td>0.572</td>
<td>0.007</td>
<td>0.080</td>
<td>[0.441 0.698]</td>
<td>6.342</td>
<td>1453</td>
</tr>
<tr>
<td>1/φ</td>
<td>5.326</td>
<td>0.047</td>
<td>1.055</td>
<td>[3.562 7.055]</td>
<td>1.360</td>
<td>396</td>
</tr>
<tr>
<td>ψ</td>
<td>1.951</td>
<td>0.006</td>
<td>0.175</td>
<td>[1.664 2.239]</td>
<td>-0.457</td>
<td>211</td>
</tr>
<tr>
<td>γ_p</td>
<td>0.304</td>
<td>0.005</td>
<td>0.113</td>
<td>[0.123 0.481]</td>
<td>0.594</td>
<td>340</td>
</tr>
<tr>
<td>γ_w</td>
<td>0.395</td>
<td>0.004</td>
<td>0.148</td>
<td>[0.156 0.629]</td>
<td>-0.628</td>
<td>132</td>
</tr>
<tr>
<td>ξ_p</td>
<td>0.768</td>
<td>0.002</td>
<td>0.034</td>
<td>[0.711 0.824]</td>
<td>-0.098</td>
<td>712</td>
</tr>
<tr>
<td>ξ_w</td>
<td>0.355</td>
<td>0.002</td>
<td>0.036</td>
<td>[0.297 0.409]</td>
<td>-5.697</td>
<td>791</td>
</tr>
<tr>
<td><strong>Policy Parameter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ_m</td>
<td>0.870</td>
<td>0.001</td>
<td>0.022</td>
<td>[0.836 0.907]</td>
<td>2.639</td>
<td>315</td>
</tr>
<tr>
<td>µ_c</td>
<td>1.618</td>
<td>0.003</td>
<td>0.103</td>
<td>[1.442 1.779]</td>
<td>1.110</td>
<td>206</td>
</tr>
<tr>
<td>µ_y</td>
<td>0.125</td>
<td>0.002</td>
<td>0.050</td>
<td>[0.041 0.206]</td>
<td>-0.109</td>
<td>381</td>
</tr>
<tr>
<td><strong>Shock Persistence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ_z</td>
<td>0.558</td>
<td>0.003</td>
<td>0.096</td>
<td>[0.400 0.715]</td>
<td>-2.253</td>
<td>219</td>
</tr>
<tr>
<td>ρ_c</td>
<td>0.511</td>
<td>0.004</td>
<td>0.112</td>
<td>[0.328 0.696]</td>
<td>-1.405</td>
<td>265</td>
</tr>
<tr>
<td>ρ_g</td>
<td>0.841</td>
<td>0.002</td>
<td>0.072</td>
<td>[0.730 0.964]</td>
<td>-1.592</td>
<td>229</td>
</tr>
<tr>
<td>ρ_L</td>
<td>0.484</td>
<td>0.011</td>
<td>0.161</td>
<td>[0.230 0.762]</td>
<td>0.806</td>
<td>857</td>
</tr>
<tr>
<td>ρ_inv</td>
<td>0.677</td>
<td>0.009</td>
<td>0.122</td>
<td>[0.476 0.858]</td>
<td>-1.878</td>
<td>968</td>
</tr>
<tr>
<td><strong>Shock SD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ε_c</td>
<td>3.235</td>
<td>0.027</td>
<td>0.610</td>
<td>[2.245 4.192]</td>
<td>2.772</td>
<td>387</td>
</tr>
<tr>
<td>ε_inv</td>
<td>2.465</td>
<td>0.034</td>
<td>0.591</td>
<td>[1.499 3.383]</td>
<td>2.700</td>
<td>664</td>
</tr>
<tr>
<td>ε_q</td>
<td>5.174</td>
<td>0.263</td>
<td>3.477</td>
<td>[0.182 9.249]</td>
<td>-1.866</td>
<td>1156</td>
</tr>
<tr>
<td>ε_z</td>
<td>0.466</td>
<td>0.002</td>
<td>0.046</td>
<td>[0.392 0.538]</td>
<td>-0.212</td>
<td>232</td>
</tr>
<tr>
<td>ε_p</td>
<td>6.683</td>
<td>0.154</td>
<td>2.141</td>
<td>[3.610 9.723]</td>
<td>0.368</td>
<td>1044</td>
</tr>
<tr>
<td>ε_L</td>
<td>9.600</td>
<td>0.025</td>
<td>0.355</td>
<td>[9.120 10.00]</td>
<td>2.826</td>
<td>1024</td>
</tr>
<tr>
<td>ε_w</td>
<td>0.225</td>
<td>0.012</td>
<td>0.168</td>
<td>[0.058 0.421]</td>
<td>1.427</td>
<td>1220</td>
</tr>
<tr>
<td>ε_g</td>
<td>1.948</td>
<td>0.007</td>
<td>0.182</td>
<td>[1.655 2.234]</td>
<td>-0.761</td>
<td>271</td>
</tr>
<tr>
<td>ε_m</td>
<td>0.131</td>
<td>0.001</td>
<td>0.013</td>
<td>[0.110 0.151]</td>
<td>-3.169</td>
<td>261</td>
</tr>
</tbody>
</table>
### Table 2.2 (b) With MEs: Estimation Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SE</th>
<th>SD</th>
<th>90 % CI</th>
<th>CD</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h )</td>
<td>0.524</td>
<td>0.004</td>
<td>0.096</td>
<td>[0.363 0.679]</td>
<td>0.481</td>
<td>335</td>
</tr>
<tr>
<td>( \sigma_c )</td>
<td>1.405</td>
<td>0.013</td>
<td>0.308</td>
<td>[0.920 1.927]</td>
<td>-0.139</td>
<td>346</td>
</tr>
<tr>
<td>( \sigma_L )</td>
<td>0.948</td>
<td>0.005</td>
<td>0.066</td>
<td>[0.848 1.072]</td>
<td>-3.004</td>
<td>1354</td>
</tr>
<tr>
<td>( 1/\varphi )</td>
<td>4.684</td>
<td>0.057</td>
<td>1.170</td>
<td>[2.664 6.483]</td>
<td>1.112</td>
<td>479</td>
</tr>
<tr>
<td>( \phi )</td>
<td>1.814</td>
<td>0.008</td>
<td>0.202</td>
<td>[1.478 2.143]</td>
<td>-0.099</td>
<td>308</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.210</td>
<td>0.003</td>
<td>0.075</td>
<td>[0.084 0.331]</td>
<td>-0.209</td>
<td>249</td>
</tr>
<tr>
<td>( \gamma_p )</td>
<td>0.817</td>
<td>0.006</td>
<td>0.126</td>
<td>[0.638 0.992]</td>
<td>0.438</td>
<td>428</td>
</tr>
<tr>
<td>( \gamma_w )</td>
<td>0.678</td>
<td>0.006</td>
<td>0.169</td>
<td>[0.414 0.950]</td>
<td>2.348</td>
<td>278</td>
</tr>
<tr>
<td>( \xi_p )</td>
<td>0.812</td>
<td>0.002</td>
<td>0.042</td>
<td>[0.748 0.884]</td>
<td>3.044</td>
<td>579</td>
</tr>
<tr>
<td>( \xi_w )</td>
<td>0.685</td>
<td>0.003</td>
<td>0.058</td>
<td>[0.589 0.783]</td>
<td>-0.936</td>
<td>522</td>
</tr>
<tr>
<td><strong>Policy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_m )</td>
<td>0.730</td>
<td>0.004</td>
<td>0.055</td>
<td>[0.639 0.820]</td>
<td>1.858</td>
<td>1025</td>
</tr>
<tr>
<td>( \mu_{\pi} )</td>
<td>1.680</td>
<td>0.003</td>
<td>0.099</td>
<td>[1.516 1.842]</td>
<td>0.216</td>
<td>221</td>
</tr>
<tr>
<td>( \mu_y )</td>
<td>0.062</td>
<td>0.003</td>
<td>0.055</td>
<td>[-0.031 0.147]</td>
<td>3.957</td>
<td>584</td>
</tr>
<tr>
<td><strong>Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_z )</td>
<td>0.756</td>
<td>0.008</td>
<td>0.140</td>
<td>[0.557 0.989]</td>
<td>-3.372</td>
<td>660</td>
</tr>
<tr>
<td>( \rho_c )</td>
<td>0.517</td>
<td>0.005</td>
<td>0.118</td>
<td>[0.323 0.714]</td>
<td>-0.517</td>
<td>310</td>
</tr>
<tr>
<td>( \rho_g )</td>
<td>0.831</td>
<td>0.004</td>
<td>0.095</td>
<td>[0.696 0.980]</td>
<td>-0.807</td>
<td>290</td>
</tr>
<tr>
<td>( \rho_L )</td>
<td>0.854</td>
<td>0.004</td>
<td>0.085</td>
<td>[0.731 0.979]</td>
<td>-1.520</td>
<td>344</td>
</tr>
<tr>
<td>( \rho_{inv} )</td>
<td>0.772</td>
<td>0.005</td>
<td>0.088</td>
<td>[0.645 0.916]</td>
<td>-3.057</td>
<td>535</td>
</tr>
<tr>
<td><strong>Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_c )</td>
<td>2.481</td>
<td>0.042</td>
<td>0.754</td>
<td>[1.274 3.677]</td>
<td>2.176</td>
<td>634</td>
</tr>
<tr>
<td>( \varepsilon_{inv} )</td>
<td>2.490</td>
<td>0.052</td>
<td>0.796</td>
<td>[1.298 3.671]</td>
<td>-1.084</td>
<td>846</td>
</tr>
<tr>
<td>( \varepsilon_q )</td>
<td>3.541</td>
<td>0.205</td>
<td>3.003</td>
<td>[1.970 7.855]</td>
<td>-4.649</td>
<td>939</td>
</tr>
<tr>
<td>( \varepsilon_z )</td>
<td>0.401</td>
<td>0.003</td>
<td>0.057</td>
<td>[0.304 0.487]</td>
<td>-0.082</td>
<td>391</td>
</tr>
<tr>
<td>( \varepsilon_L )</td>
<td>1.033</td>
<td>0.068</td>
<td>0.833</td>
<td>[0.165 2.176]</td>
<td>3.204</td>
<td>1337</td>
</tr>
<tr>
<td>( \varepsilon_g )</td>
<td>0.808</td>
<td>0.032</td>
<td>0.463</td>
<td>[0.147 1.483]</td>
<td>-2.146</td>
<td>953</td>
</tr>
<tr>
<td>( \varepsilon_m )</td>
<td>0.999</td>
<td>0.001</td>
<td>0.012</td>
<td>[0.078 0.118]</td>
<td>-0.673</td>
<td>390</td>
</tr>
</tbody>
</table>

### Table 2.3. Log-Marginal Likelihood and Bayes Factor

<table>
<thead>
<tr>
<th></th>
<th>Modified Harmonic Mean (Geweke, 1999)</th>
<th>SD of Log-Marginal Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>with MEs</td>
<td>-561.206</td>
<td>0.503</td>
</tr>
<tr>
<td>w/o MEs</td>
<td>-643.669</td>
<td>0.176</td>
</tr>
<tr>
<td>log Bayes Factor</td>
<td>82.463</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1 Smoothed State Variables
Figure 2.2 (a) Historical Decomposition: Inflation w/o ME

Figure 2.2 (b) Historical Decomposition: Inflation with ME
2.7. TABLES AND FIGURES

Figure 2.3 (a) Historical Decomposition: Real Wage w/o ME

Figure 2.3 (b) Historical Decomposition: Real Wage with ME
CHAPTER 2. ROLE OF MEASUREMENT ERROR
Chapter 3

Estimation in a Data Rich Environment

3.1 Introduction

Chapter 3 examines how much estimation accuracy of parameters and structural shocks improved by the estimation method utilizing a large number of data.


3.1.1 Background

When estimating a DSGE model, we need to describe the model in the state space model. The state space model consists of state equation and measurement equation.

The agents in the model should observe data necessary for their own decision, know parameters related to their decision making, look at the realized structural shocks at current period and decide their own behaviors (endogenous variables) at this period. In this way, it is well determined how endogenous variables at the previous period transits to current period. Thus, the state equation describes transitions of endogenous variables in this model.\(^1\)

For example, consider an agent in a model called as an intermediate goods firm. He looks at the realized general price, knows the price elasticity of demand for his goods and when next opportunity for price revision will come, and observes realized structural shocks at current period such as demand shocks and supply shocks. Then, he will decide current price of his goods so as to maximize the discounted present value of the future profit streams if the price of this period is left unchanged. By aggregating the results of such pricing behaviors of intermediate goods firms, inflation (endogenous variable) is determined.

\(^1\)Again, it should be noted that the agents in the model are assumed to be able to observe the data precisely. Otherwise, they have to predict the data to make their own decisions, which might cause another interactions between agents and econometricians on predicting data. Without the assumption observing the data precisely by the agents, the likelihood must be evaluated by incorporating agents’ forecasting errors by econometricians, but agents might also take the data predictions by econometricians into account. Then, there might be an incentive to induce econometricians’ data prediction to increase agents’ payoffs and agents may change their own behavior. In this case, the likelihood evaluation by the Kalman filter becomes invalid way.
On the other hand, consider an econometrician who wants to estimate a DSGE model. Of course, he does not know either parameters or realized structural shocks. Moreover, the econometrician does not know exactly which data is being observed when agents in the model decide their actions, although he may find some data likely to correspond. So, for example, the econometrician makes a hypothesis that agents in the model might match GDP data with output gap and match GDP deflator with inflation gap, and so on. That is, the measurement equation shows the correspondence between data and endogenous variables by the econometrician.

Now, the central bank, an agent in the model, observes the output gap and the inflation gap and decides the nominal interest rate, which is an endogenous variable, in accordance with the monetary policy rule. In the case of the above example, the econometrician assumed that the central bank had decided the nominal interest rate by observing the data of GDP and GDP deflator.

Normally, in the measurement equation, one data is associated with one endogenous variable. However, the data observed by agents in the model as endogenous variables such as output gap and inflation gap may not necessarily correspond only to data such as GDP and GDP deflator. For the output gap, they also might refer to data of industrial production index (IIP), in addition to GDP. They might infer the output gap by using the quick estimates (QE) as well as the revised value of GDP. For the inflation gap, they might use CPI data other than GDP deflator. They might also observe core CPI and core core CPI.

In this way, agents in the model may be making their own decisions using various data that seems to have information on output gap and inflation gap. Nonetheless, if econometricians attempt to evaluate policy effects with estimation results based on less data, or if econometricians plan to predict future macroeconomic data, then, there is a risk that erroneous guesses might be brought about.

This remarkable example is the so-called “price puzzle”. Usually, the structural VAR model aims to extract variations depending on monetary or fiscal policy from output data or inflation data. Thus, econometricians want to identify monetary or fiscal policy shock and conduct policy simulations based on the estimated structural VAR model. However, when estimating the structural VAR model using less data such as 6 to 8 (GDP, GDP deflator, FF rate, etc.), econometricians have often observed that price rises against the shock of rise in nominal interest rate, which is the opposite reaction predicted from the theory (price puzzle).

Sims (1992), in response to this puzzle, pointed out that the central bank might predict future inflation from various data and decide the current nominal interest rate taking into account future prospects of inflation as well as the realized current inflation.

Suppose that the central bank implements monetary policy considering future inflation. And consider the situation where the central bank predicted from various data that high inflation would be hit by a certain degree of accuracy in the near future. In this case, the central bank will raise the current policy rate in preparation for future inflation, even if inflation has not been observed at the present period or even if some deflation has been observed at this period. Nevertheless, if the monetary policy rule in the VAR model is estimated without controlling the future projection of inflation by the central bank, the estimated coefficient of inflation at the present period might be close to zero or negative. As a result, econometricians will face the price puzzle that price rises against the monetary tightening policy shock.

Sims (1992) suggested a very simple way to avoid the puzzle: Add data that can control future inflation expectations by the central bank. Specifically, from Sim’s suggestion, the commodity price index should be added to the VAR variable. This data is a representative of leading indicators for price in the U.S. By adding the data, it was shown that the puzzle can be solved in the structural
3.1. INTRODUCTION

VAR model, by controlling the future inflation prediction in the central bank’s policy decision.

The same problem arises even when we want to identify the effect of fiscal policy by the structural VAR model. Suppose, for example, an econometrician wants to measure the effect of tax cuts on output and tax revenue. So, GDP data and tax revenue data are included in the structural VAR model. Of course, what he wants to know is the causality from tax cuts to output. In addition, suppose that the government judged that the current economy is getting worse by looking at the latest GDP statistics and that the current Diet aims to establish a discretionary tax reduction. This situation causes a causality from output to tax cuts. Thus, biased parameters will be estimated due to the endogeneity problem.

Furthermore, the effect of the policy shock is assessed by the response of output and inflation against deviations of the policy instrument from the policy rule.

For example, consider the case where an econometrician wants to see the output response to a tax cut shock deviating from the taxation rule like progressive income tax. In this case, it is necessary to separate the output response into two components: one component endogenously in response to the policy rule (the progressive income tax rule), and the other component in response to shock deviating from the rule. Therefore, in order for the econometrician to know the effect on output against tax cuts, it is necessary to control the endogenous response of output in accordance with the policy rule.

To avoid the identification problem of the fiscal policy shock in the structural VAR model, Blanchard and Perotti (2002) also proposed extremely simple method.

In the example above, consider the situation where the econometrician wants to control the causality from output to tax cuts. If he estimates the VAR model using annual data, the government will cut taxes through the Diet deliberations during the year in response to the current economic downturn. However, if he uses quarterly data, even if the government got information on the current recession, it is unlikely that the tax cut legislation will be passed through the Diet deliberations within three months. Therefore, in order to eliminate the causality from output to tax reduction, Blanchard and Perotti (2002) proposed that the VAR model should be estimated by using high frequency data.

Next, consider the situation where the econometrician wants to control the endogenous reaction of output against the tax rule in tax reduction. This can be addressed by using additional information not adopted as data of the structural VAR model: The econometrician should measure the marginal tax rate according to income class from the institutional information and estimate income elasticity to tax using micro data. Then, it is possible to estimate in advance how much tax cuts will be given to people in each income class and what percentage income of each level responds endogenously to the tax reduction. By eliminating endogenously reacting fluctuations from output data in advance (pretreatment), it is possible to extract the output reaction against the discretionary tax cut shock.

In both approaches, there is a common feature that if we want to identify policy effects, we should give much data that the central bank, government and market participants are using to decide their behaviors.

In the field of time series analysis, when predicting the future data (Stock and Watson 2002a, b), or estimating the monetary policy rule or examining the policy propagation mechanism (Bernanke and Boivin 2003, Bernanke et al. 2005), estimation methods using a large number of data have been proposed, taking account into the situation where agents in the model utilize a lot of data to decide their own actions.\(^2\)

\(^2\)Regarding the developments of this field, Stock and Watson (2010) is conducting a detailed survey. However, Stock
Stock and Watson (2002a, b) proposed the so-called dynamic factor model (DFM) that summarizes information of a large number of data with a small number of factors and estimates transition equations considering the interdependence relationship between factors. Then, they showed that DFM improves the prediction accuracy of data. DFM is also used in estimating the diffusion index (DI), and the estimated DI is referred to when NBER sets the date of turning point of business cycle. DFM is represented as follows:

\[
F_t = G(\theta)F_{t-1} + \varepsilon_t \tag{3.1}
\]

\[
X_t = \Lambda^F F_t + e_t \tag{3.2}
\]

where (3.1) is the state equation of DFM, \(F_t (K \times 1)\) is the factor vector, and \(\varepsilon_t (K \times 1)\) is a disturbance term vector. The coefficient matrix \(G(\theta) (K \times K)\) and the factor \(F_t\) are unknown and to be estimated.

(3.2) is the measurement equation of DFM, \(X_t (N \times 1)\) is the data vector and \(e_t (N \times 1)\) is the measurement error. \(\Lambda^F (N \times K)\) is called as a factor loading matrix, which represents the correspondence between each factor and a large number of data:

\[
\begin{bmatrix}
X_{1,t} \\
X_{2,t} \\
\vdots \\
X_{N,t}
\end{bmatrix}
= \begin{bmatrix}
\lambda_{1,1} & \cdots & \lambda_{1,K} \\
\vdots & \ddots & \vdots \\
\lambda_{N,1} & \cdots & \lambda_{N,K}
\end{bmatrix}
\begin{bmatrix}
F_{1,t} \\
\vdots \\
F_{K,t}
\end{bmatrix}
+ \begin{bmatrix}
e_{1,t} \\
\vdots \\
e_{N,t}
\end{bmatrix}
\tag{3.3}
\]

The parameters to be estimated in DFM are components of \(G(\theta)\), components of \(\Lambda^F\), and the variance covariance matrix of \(\varepsilon_t\) and \(e_t\). DFM has an advantage to reduce large-scale data information with fewer factors, so \(N \gg K\) is assumed.

Stock and Watson (2002a, b) showed that using estimated factors by DFM improves prediction accuracy of data, especially inflation data. In addition, although the factors are usually estimated by the Kalman filter and smoother based on (3.1) and (3.2), they show theoretically the factors can be also estimated approximately by applying the principal component analysis (PCA) for a large number of data.

Recall that, to resolve the price puzzle, we need to incorporate the future inflation expectation by the central bank into the monetary policy rule. Stock and Watson (2002a, b) revealed that using DFM increases prediction accuracy of inflation. Moreover, they also showed the factors can be easily estimated by PCA.

If so, first of all, we should predict future inflation with high precision using factors estimated by PCA. Then, we estimate the monetary policy rule, by adding the predicted value to a regressor. By doing this, we might be able to estimate the monetary policy rule more accurately. Based on this idea, Bernanke and Boivin (2003) estimated the monetary policy rule.

According to the usual monetary policy rule, when the central bank sets the nominal interest rate, it reacts not only to inflation gap but also to output gap. The output gap is defined as what is percent deviation between the actual output (output under the sticky price) and the potential output (output under the flexible price). However, the potential output is a “latent variable” unobservable in the data.

and Watson (2010) focuses on the developments from a statistical point of view, such as the asymptotic characteristics of factors estimated by DFM. Instead, we clarify the problems of the previous studies from the viewpoint of theoretical models such as the identification problem of the monetary policy shock and whether monetary policy rule specification is appropriate or not.
Again, factors are also useful for estimating latent variables such as output gap. Adding estimated output gap by factors should also help to avoid the estimation bias of monetary policy rule. Since output gap corresponds to unemployment rate through the Okun’s law, we should predict unemployment rate data by using factors, then estimate the monetary policy rule using the predicted value of unemployment data as a proxy of output gap.

Utilizing the high prediction accuracy of factors, they estimated the monetary policy rule in the following three stages:

First, factors corresponding to output gap and inflation gap were extracted by PCA using large scale data of Stock and Watson (2002b) such as various industrial production indices and price indices.

Then, they predicted unemployment rate data and CPI data by using estimated factors. Those predicted values are corresponding to the proxy of anticipated future output gap and the expected future inflation gap, which might be used when the central bank conducts the monetary policy.

Finally, they estimated the monetary policy rule by regressing the predicted unemployment rate and the predicted CPI to the nominal interest rate. As a result, the use of factors increased the prediction accuracy of unemployment data and CPI, and when estimating the monetary policy rule using the predicted values, they found the Taylor coefficient for inflation was high during the Chairman Volder.

By further advancing this idea, Bernanke et al. (2005) proposed a FAVAR (Factor-Augmented Vector Autoregressive) model that integrates both DFM and structural VAR. They showed that the FAVAR model is useful not only for estimating the monetary policy rule that incorporate output gap and inflation gap, but also as a way to examine various propagation channels of the monetary policy shock.

The idea of the FAVAR model is quite simple: Use DFM factors for VAR variables. By estimating the state equation considering the interdependence relationship between factors and ordinary VAR variables, the estimation bias of the monetary policy rule could be eliminated.

In the conventional VAR analysis, output gap, inflation gap and nominal interest rate are considered as observable by GDP, GDP deflator and FF rate. Instead, they considered only nominal interest rate as observable, and regarded output gap and inflation gap as “latent variables” (factors estimated from a large number of data). Furthermore, they also consider a factor consisted of another data such as money stock, exchange rate, profit dividend and add the factor to the VAR model:

\[
\begin{bmatrix}
F_t \\
Y_t
\end{bmatrix} = G(\theta) \begin{bmatrix}
F_{t-1} \\
Y_{t-1}
\end{bmatrix} + \varepsilon_t
\]

\[
X_t = \Lambda F_t + \epsilon_t
\]

(3.4) is the state equation in the FAVAR model. \(F_t\) is the factor vector, \(\hat{y}_t\) is output gap, \(\hat{\pi}_t\) is inflation gap, and letting the common factor of data such as money stock, exchange rate, profit dividend is denoted as \(f^t\), then \(F_t = [\hat{y}_t, \hat{\pi}_t, f^t]'\). \(\varepsilon_t\) is the reduced-form shock vector. \(Y_t\) is an observable VAR variable vector, but here, \(Y_t = \hat{R}_t\), because only nominal interest rate \(\hat{R}_t\) can be observed.

Essentially, the reduced-form shock \(\varepsilon_t\), a disturbance term of the state equation, can be represented by a linear combination of structural shocks such as supply shock, demand shock (including fiscal policy shock) and monetary policy shock. Thus, in order to identify the monetary policy shock, we need to extract only the monetary policy shock from the linear combination of structural shocks. However, thanks to the additional assumption that the factor \(F_t\) responds to previous nominal in-
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Interest rate, we can identify the monetary policy shock, since the disturbance term corresponding to the monetary policy rule is only the monetary policy shock.

(3.5) is the measurement equation in the FAVAR model. $X_t$ is the data vector and $e_t$ is the measurement error vector. As with Stock and Watson (2002a, b), $\Lambda^F$ is a factor loading matrix, which is the coefficient matrix of each data when extracting factors.

Then, they estimated the FAVAR model using as many as 120 or more data as $X_t$ and examined impulse responses to a monetary tightening policy shock.

A major advantage of this approach is that it can estimate impulse responses to the monetary policy shock of all the data making up the factor: In other words, by checking the impulse response of each data, it is possible to confirm whether the propagation mechanism of monetary policy shock is consistent with the theory.

There are two methods for estimating the FAVAR model: The first is to extract common factors from a large number of data by PCA, and then, integrate them with an ordinary VAR model. The second is to estimate factors simultaneously by evaluating the likelihood using the Kalman Filter.

In both estimation results, they reported that the price puzzle is eliminated. Moreover, they examined impulse responses of data constituting the factor. Then, they also found money stock will decrease in response to the monetary tightening policy, thus the so-called “liquidity puzzle”, in which money stock is increased against the monetary tightening policy, has been also eliminated. With the rise of nominal interest rate, exchange rate appreciated and profit dividend initially climbed, but later returned to zero. By utilizing such a large amount of data, they confirmed the estimated impulse responses are consistent with the theoretical propagation channels.

That is a rough sketch of development history on the time series analysis approach in estimating policy rules and identifying policy shocks utilizing a large number of data.

These empirical studies seem to be useful for improving the data prediction accuracy, as the purpose of Stock and Watson (2002a, b). If the aim is to increase prediction accuracy, there is no need to ask the mechanism of why this factor is useful for predicting data. However, there are three major problems in estimating policy rules, identifying policy shocks, and conducting policy simulations by statistical approaches using reduced-form models as described the above.

First, whether it is DFM or FAVAR, we cannot understand economic meanings of factors precisely. There is no need to think about the economic meanings of factors if it just aims to improve prediction accuracy. But when considering the propagation mechanisms of how a certain policy spreads through these factors, we have to understand the economic meanings of factors. However, without any structural model information, we cannot investigate the correspondence between factors and model variables. For example, in Bernanke et al. (2005), it is difficult to interpret the economic meaning of $f^1_t$, which is a factor of data such as money stock, exchange rate, and profit dividend.

Next, in the structural VAR model, ad hoc restrictions are required to identify monetary policy shocks. In the state equation, the reduced-form shock is represented by a linear combination of structural shocks through $E(\theta)$ (see (1.64)), but it is necessary to add some constraints to distinguish monetary policy shock from the reduced-form shock. While this is a common problem of the structural VAR model, Bernanke et al. (2005), for example, assumes the IS curve responds to the nominal interest rate at the previous period to identify the monetary policy shock. That is, monetary policy shock is identified by assuming that monetary authority firstly decides nominal interest rate, and then, private agents react to the lagged nominal interest rate. In other words, it is assumed to be zeros for components of $E(\theta)$ other than the monetary policy shock in the equation corresponding to the monetary policy rule. Unless $E(\theta)$ is determined from the model, we need to
impose ad hoc zero restrictions on components of $E(\theta)$ for identifying monetary policy shock.

Finally, there is a discrepancy with DSGE models as to which endogenous variables the central bank responds to. After Sims (1992), the central bank is regarded to conduct monetary policy according to future inflation. That is why previous studies are trying to control inflation expectation by utilizing a lot of data. However, in DSGE models with dynamic optimizations, if the central bank implements his policy according to future inflation expectation, then the economy might become fragile.

In the DSGE model, households (firms) decide on current consumption (current prices) so as to maximize the discounted present value of the future utility streams (future profit streams). As a result, consumption expectation (inflation expectation) at the next period will influence current consumption (inflation) decisions: Consider, for example, the Euler equation (1.36):

$$\hat{c}_t = \frac{1}{1 + h} E_t \hat{c}_{t+1} + \frac{h}{1 + h} \hat{c}_{t-1} - \frac{1 - h}{(1 + h)\sigma} \left( \hat{R}_t - E_t \hat{\pi}_{t+1} \right)$$

The real interest rate is the difference between the current nominal interest rate $\hat{R}_t$ and one step ahead inflation expectation $E_t \hat{\pi}_{t+1}$. According to the Euler equation, if households anticipate high inflation at the next period, which leads to the reduction of current real interest rate, so households will be better off raising current consumption by selling the bonds they hold.

Following Sims (1992), suppose that the central bank also decides current nominal interest rate according to an anticipated inflation at the next period:

$$\hat{R}_t = \phi_\pi E_t \hat{\pi}_{t+1} + \varepsilon_t^R$$

where $\phi_\pi$ is the response coefficient to the expected inflation, and $\varepsilon_t^R$ is the monetary policy shock.

Also, suppose the central bank raised current nominal interest rate in anticipating inflation at the next period. The problem is that the monetary tightening of this period has a role of a signal for households that the central bank is expecting the upcoming inflation. Households also recognize the central bank implements the policy in response to future inflation expectation. Therefore, observing the current tightening policy implies the central bank may have evidence of future inflation, which may cause households’ future inflation anticipations. If the expected future inflation by households is higher than the nominal interest rate raised by the central bank, the real interest rate must be lower than before the monetary tightening: Despite raising nominal interest rate to suppress future inflation, it will trigger future inflation expectation by households, which might reduce current real interest rate. The reduction of real interest rate should lead to an increase in current consumption through the intertemporal substitution effect, which will boost current aggregate demand and will raise inflationary pressures. As a result, inflation will be realized at the next period as expected (self-fulfillingness equilibrium). After all, contrary to the initial prospects of the monetary authority, the economy might become fragile. Using the DSGE terminology, if the private agents are taking forward looking behaviors, and if the monetary authority also adopt the forward looking policy, then there might be a possibility of the indeterminancy problem. \(^3\)

In sum, if the structural model behind the state equation is not explicitly shown, interpreting factors is difficult, the ad hoc constraints necessary for identifying shocks should be imposed, and the specification of the policy rule might not be consistent with the theoretical model.

\(^3\)If private agents do not optimize their behaviors dynamically, it is possible for the central bank to stabilize the economy by conducting forward looking policy. For example, Bernanke et al. (2005) explicitly assumes backward looking IS and Phillips curves.
At the same time, however, the assumption on estimating the DSGE model is also extreme: The endogenous variable corresponds to the data on a one-to-one basis.

In response to this background, Boivin and Giannoni (2006), Schorfheide, et al. (2010), Kryshko (2011) proposed a new estimation method: The data-rich estimation method is to utilize a lot of data and also capture information of the structural model at the same time. The idea is still simple: Integrate DFM with DSGE model.

First, endogenous variables such as output gap and inflation gap of the DSGE model are regarded as latent variables and they are made to correspond to factors of DFM. That is, the state equation of DFM is replaced with the state equation of the DSGE model:

\[
S_t = G(\theta)S_{t-1} + E(\theta)\varepsilon_t \tag{3.6}
\]

\[
X_t = AS_t + \varepsilon_t \tag{3.7}
\]

In this case, since the factor of DFM is associated with the endogenous variable of the DSGE model, it is immediately apparent what the factor indicates. Endogenous variables such as output gap and inflation gap have an obvious model concept and dynamics of endogenous variables are also derived from dynamic optimization of economic agents in the model. By associating factors with endogenous variables, it becomes possible to explicitly grasp the model concept of factors.

In the usual estimation method, endogenous variables and data are made to correspond one to one. As a result, only the number of periods indicated the amount of data information. However, by combining with DFM, it becomes possible to correspond endogenous variables and data one to many. Therefore, by increasing the number of cross section series, even more information can be utilized.

We could expect that an increase in the amount of data information will lead to an improvement in estimation accuracy of parameters and structural shocks: For example, if many price indices are available as data corresponding to inflation gap, it is possible to estimate inflation gap using multiple data information at the same time. By adding cross section series, the estimation efficiency of parameters and shocks volatilities would be improved (the credible interval of the posterior distribution would shrink). Since any endogenous variable is represented by a linear combination of structural shocks, the more accurate estimates of parameters and shocks should also contribute to the more efficient estimates of endogenous variables.

The data rich estimation method has advantages related to the role of the measurement errors in the previous section. For example, we will match the common factor of multiple price indices to a latent variable called as inflation gap. This corresponds to separating fluctuations of multiple price indices into two variations: The first is the variations explained from the model, extracted as the common factors of large data. The common factors are matched to endogenous variables \(S_t\) in the state equation. The other is the variations which are not captured by the model, regarded as the specific variations possessed by each data. The unique factors (or idiosyncratic components) are identified as the measurement errors \(\varepsilon_t\) in the measurement equation. That is, since common factors are extracted using cross-sectional information of a large number of data, as compared with the conventional estimation method of associating endogenous variables with data on a one-to-one basis, the data rich estimation method could be much easier to separate the data into two components: The components can be explained by the model (common factors) and the other components cannot be explained by the model (idiosyncratic factors = measurement errors).

Second, in the DSGE model, how current endogenous variables transit from previous endogenous variables in response to current structural shocks are uniquely determined from the state equation as the solution of the model: For the DSGE model, not only \(G(\theta)\) but also \(E(\theta)\) is uniquely
3.1. INTRODUCTION

derived from the model. In the structural VAR model, ad hoc zero constraints are required for the components of \( E(\theta) \) in identifying the shocks, whereas in the DSGE model, how endogenous variables are represented by structural shocks is determined only by the model information, and no additional assumptions are needed.

Third, the DSGE model can estimate monetary policy rule consistent with private agents behaviors based on dynamic optimizations. Since the uniqueness of the model solution is verified when estimating parameters, the state equation is also uniquely determined. As a result, the stabilization rule by the monetary authority could be specified without conflict with the forward looking behaviors of private sectors, thus it could be possible to eliminate the indeterminancy problem based on the mis-specification of the monetary policy rule.

In addition, as in Bernanke et al. (2005), we can check the impulse response of each data against monetary policy shock by using the estimation result of factor loadings \( \Lambda \) which shows correspondence between factors and data. Therefore, we can also consider various propagation channels of monetary policy by the data rich estimation method.

Based on the above idea, Boivin and Giannoni (2006) estimated the SW model by the data rich estimation method using data sets of at most 90. According to the estimation results, first, as expected, we can confirm improvements of estimation accuracy of structural shocks and endogenous variables, especially estimates of inflation gap became more efficient. Second, the prediction accuracy of one quarter ahead data is improved and the GDP deflator, in particular, becomes more precisely predictable. Third, as a source of inflation variations, the role of the price markup shock has been reduced compared to SW (2007), and the role of the productivity shock increased instead. The third finding may suggest when estimating by utilizing a large number of data information, there might be differences in the determinants of business cycle and inflation fluctuations.

3.1.2 Purposes, Originalities, and Methodologies

This study estimates the SW model by the data rich estimation method using data up to 55 series for the Japanese economy and aims (i) to examine how much estimation accuracy of structural parameters, structural shocks and endogenous variables improves by utilizing a lot of data information and (ii) to consider the sources of output and inflation fluctuations.

There are two originalities: First, this research is the first attempt to apply the data rich estimation method to the standard DSGE model using Japanese data. By utilizing large data information, how much shrinking will be the posterior distribution of parameters? Will the estimation accuracy of inflation gap increase, similarly to Boivin and Giannoni (2006) results in the U.S.? When monetary policy rule is estimated using highly accurate inflation gaps, how will the Taylor coefficients change compared to the results of previous studies? It seems possible to answer those questions by the data rich estimation method.

Also, utilizing both model information and data information will make it easier to distinguish common factors that correspond to endogenous variables with measurement errors that correspond to unique factors of data. Therefore, relating to the discussion in Chapter 2, it becomes possible to ask how much the standard DSGE model can explain the inflation behavior when using a lot of data information.

Second, we examine how much the source of business cycle changes as compared with the conventional estimation method. Boivin and Giannoni (2006) reported shocks that contribute to explain inflation fluctuations have shifted from markup shocks, impossible to interpret structurally, to productivity shock, possible to interpret structurally. This result suggests that the sources of inflation fluctuations may change by using information of a lot of data.
Will the contribution of productivity shock explaining the inflation gap also increase in Japan? If so, do we support the results of Hayashi and Prescott (2002) that concluded Japan’s lost decade since the 1990s is due to a decline in productivity? On the other hand, it has also been qualitatively pointed out the BOJ’s monetary easing policy in the latter half of the 1980s generated asset price bubbles and the monetary tightening policy in the early 1990s aimed at restraining the bubble has become a trigger for Japanese long-lasting recession. How is the monetary policy shock contributing to the explanation of the business cycle around the 1990s?

Given the results of Bernanke and Boivin (2005) that DFM improves the precision of the output gap, the data rich estimation method will also improve the prediction accuracy of output gap. Therefore, it would be worth exploring sources of the business cycle by decomposing highly precise output gap by highly accurate structural shocks.

Compared with the historical decomposition based on the normal DSGE model estimation method, we will examine how much change will occur in the sources of the business cycle by utilizing a lot of data information.

As for the estimation method, unlike Stock and Watson (2006a, b) and Bernanke and Boivin (2005) which extract common factors from a large number of data by PCA in the first stage and then estimate the state equation, similarly to Boivin and Giannoni (2006) and Kryshko (2011), we adopt a method for evaluating the likelihood by the Kalman filter and simultaneously estimating factors.

Also, as in Chapter 2, model parameters are estimated by the MH algorithm, and for the parameters related to the measurement equation Gibbs sampling is used (Metropolis within Gibbs). The difference from Chapter 2 is that the factor loading matrix is also estimated by the Bayesian OLS method.

As a concrete methodology, we estimate the SW model in a data-rich environment using Japanese 55 macroeconomic indicators between 1981:Q1 and 1995:Q4 by the Bayesian technique. We calculate posterior means and credible intervals of (1) endogenous variables, (2) structural shocks and (3) impulse response functions of monetary policy to all model variables, and (4) posterior means of variance decomposition and historical decomposition of data series, as well as the posterior distribution of the structural parameters of the SW model.

We considered the following three cases, in estimating parameters, shocks, and endogenous variables, then comparing historical decompositions of output gap and inflation gap among three cases: Case A: a usual estimation method of matching endogenous variables and data one to one, Case B: the data-rich estimation method of matching endogenous variables and data one to many, and Case C: the data-rich estimation method introducing additional data which do not have corresponding endogenous variables such as exchange rate, money stock, stock price etc., in addition to the data of Case B.

3.1.3 My Contributions

Chapter 3 revised Iiboshi et al. (2015) written with three coauthors.

My main contribution consists of the followings: First, I implemented the codes of the data rich estimation method based on the algorithm by Kryshko (2011).

Second, I interpreted estimation results according to the model characteristics of the SW model. In particular, I considered why the estimated labor supply shock is not so efficient, and why is the increase in the reservation wage as the main source behind the Japan’s lost decade.
3.1.4 Organization of Chapter 3

Chapter 3 is organized as follows. Section 2 addresses the framework and methodology of the data rich approach using the DSGE model. Section 3 describes the SW model. Section 4 explains Bayesian estimation methods in a data-rich approach. Section 5 explains the preliminary settings and data description. Section 6 presents the estimation results. Section 7 concludes.

3.2 Data Rich Approach

DFMs are rapidly developed and applied in many fields of macroeconomics. One of applications is to estimate the DSGE model in a data rich environment. The main idea of data rich estimation is (1) to extract the common factor from a large panel of macroeconomic time series data, then (2) to match the state variable with the extracted common factor. This section illustrates the so-called data rich estimation method proposed by Boivin and Giannoni (2006) and Kryshko (2011), and describes how to utilize a large number of data in estimating DSGE model.

DFM is represented by the state space model consisted of three linear equations: The equation (3.8) is measurement equation, the equation (3.9) is state equation and (3.10) is the transition equations of measurement error. Let $F_t$ denote the $N \times 1$ vector of the unobserved common factor, and $X_t$ denote the $J \times 1$ vector of the large panel of data. Note that the number of data series shall be much larger than the number of state variables in a data rich environment, i.e. $J \gg N$.

$$
\begin{align*}
X_t & = \begin{bmatrix} \Lambda & F_t \end{bmatrix} + e_t, \quad (3.8) \\
F_t & = G(\theta) F_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, Q), \quad (3.9) \\
e_t & = \begin{bmatrix} \Psi_t \end{bmatrix} e_{t-1} + \nu_t, \quad \nu_t \sim \text{i.i.d. } N(0, R), \quad (3.10)
\end{align*}
$$

where $\Lambda$ is the factor loadings matrix ($J \times N$), $e_t$ is the idiosyncratic errors vector which is allowed to be serially correlated as shown in the equation (3.10). $G$ is $N \times N$ matrix, and the common factor $F_t$ follows the AR(1) process (3.9). The matrices $\Psi$, $Q$ and $R$ are assumed to be diagonal in the exact DFM as in Stock and Watson (2005).

DSGE models are also expressed by the state space models as the following three equations:

$$
\begin{align*}
X_t & = \begin{bmatrix} \Lambda & \bar{S}_t \end{bmatrix} + e_t, \quad (3.11) \\
\bar{S}_t & = G(\theta) \bar{S}_{t-1} + H(\theta) \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, Q(\theta)), \quad (3.12) \\
e_t & = \begin{bmatrix} \Psi_t \end{bmatrix} e_{t-1} + \nu_t, \quad \nu_t \sim \text{i.i.d. } N(0, R), \quad (3.13)
\end{align*}
$$

where $X_t$ is the observable variables vector ($J \times 1$), $\bar{S}_t$ is the state variables vector ($N \times 1$), and $e_t$ is the structural shocks vector ($M \times 1$).

DSGE models can be estimated using the Kalman filter, as can DFMs. Thus, we can apply the estimation method for the DFMs to DSGE models. However, the major difference between DFMs and a DSGE models is meanings of parameters. The DSGE model has microeconomic foundations in which agents in the model solve the dynamic optimization problem, given the structural parameters.
The structural parameters can be interpreted as tastes of households or technologies of firms that are not affected by the fiscal and monetary policies. In contrast, DFM’s (reduced-form) parameters are not necessarily interpretable. The solution of the model (the law of motion around steady states) can be expressed by (3.12). We will refer to the elements of $S_t = [x'_t, S'_t]'$ as “model concepts,” e.g. output gap, inflation gap and so on. $z_t$ is a vector of non-predetermined endogenous variables (or jump variables) and $S_t$ is a vector of predetermined endogenous and exogenous variables. $\varepsilon_t$ is a vector of exogenous shocks. $\theta$ is the structural parameters and $G(\theta)$ is a matrix of parameters. Using the DFM framework, the data rich DSGE model can be represented as (3.11), (3.12) and (3.13). In contrast with the regular DSGE model, there is a many-to-many relation between $X_t$ and $\bar{S}_t$, since the matrix $\Lambda$ is $J \times N$ ($J \gg N$) in (3.11), which could grasp the theoretical gap between the data indicator $X_t$ and the model concept $\bar{S}_t$.

In a regular DSGE model, the number of observable variable is lesser than (or equal) that of state variables, i.e. $J \leq N$. In the data rich DSGE model, the number of observable variables is much larger than that of state variables ($J \gg N$) as well as a DFM. From (3.11) and (3.12), the components $\Lambda \bar{S}_t$ will be consistent with microeconomic theory. By contrast, the measurement error $e_t$ indicates a “specific (or unique)” factor of the corresponding data $X_t$. In other words, the measurement error does not follow economic theory nor is it affected by fluctuations of other data and other endogenous variables (model concepts). Because it is allowed to fluctuate freely, the measurement errors might play an important role in removing undesirable relations between the observable variables and the model concept variables caused by model misspecification or the mismatch between model concepts and appropriate observable variables.

In addition, there is also technical requirement to introduce measurement errors. Without measurement errors, the stochastic singularities should be inevitable. Usually, as long as the number of data series does not exceed the number of structural shocks, the DSGE model can be estimated without measurement errors. But if it does not hold, we cannot estimate the model: According to (3.12), state variables (model concepts) will be represented by the linear combination of structural shocks $\varepsilon_t$. Then, the state variable is connected into the data series through (3.11), so data should also be expressed by the linear combination of the structural shocks. Normally, all state variables are not considered observable, thus the number of state variables is larger than the number of data. However, if the number of structural shocks is less than the data series, thus, if the number of structural shocks is less than the number of the state variables, the state variables will be linearly dependent with each other. As a consequence, the variance covariance matrix of state variables becomes singular, so we cannot evaluate the likelihood (stochastic singularity). Especially in a data rich environment, if we do not deal with anything, a large number of data series will have to exceed the number of structural shocks greatly, and we cannot estimate the model due to the problem. Therefore, we need to introduce measurement errors when utilizing a large data set to estimate the DSGE model.

One of advantages of data rich approach is to identify structural shocks and measurement errors through state space model. It might be hard to identify between them in one-to-one matching of the data and the model concepts, as usual estimation procedure of the DSGE model. However, by connecting the model concept with the common factor extracting from many data, the unexplainable factor by the model (i.e. measurement error) can be easily separated. One of our aims is to verify whether this advantage of data rich estimation methods work successfully.

The structural shocks $\varepsilon_t$ and disturbance terms $\nu_t$ in measurement errors $e_t$ follow normal distributions, $\varepsilon_t \sim \text{i.i.d. } N(0, Q(\theta))$ and $\nu_t \sim \text{i.i.d. } N(0, R)$, respectively. Their variance covariance matrices $Q(\theta)$ and $R$ are positive definite and diagonal matrices. The AR(1) coefficients $\Psi$ in (3.13)
are also assumed to be a diagonal matrix. These assumptions imply that measurement errors $e_t$ are independent of each other in terms of cross section but are dependent on their own lag variables in terms of time series. Finally, it should be noted that, in the DSGE model, the matrices $G(\theta)$, $H(\theta)$ and $Q(\theta)$ are derived as the solution of the model, and become functions of structural parameters $\theta$.

In the DFM framework, state variables $F_t$ are regarded as common factors from the viewpoint of statistics, as shown by (3.8). Hence, there is no need to care about how $F_t$ and $X_t$ are tied, that is, we do not care about the magnitudes and signs of elements of the factor loadings matrix $\Lambda$. In the data rich DSGE framework, data indicators $X_t$ are matched with model concepts $S_t$ in the measurement equation (3.8) and (3.11). In this case, we have to mind (a part of) the specification of the coefficient matrix $\Lambda$ in (3.11). Otherwise, we cannot identify the model concepts $S_{1,t}$.

Let us denote a model concept (for example, output gap) as $\tilde{S}_{1,t}$ and suppose that $\tilde{S}_{1,t}$ is matched with two data indicators $[X_{1,t}^1, X_{2,t}^1]'$ (for instance, GDP data and IIP data). Then, we can express the measurement equation as follows:

$$
\begin{bmatrix}
X_{1,t}^1 \\
X_{2,t}^1
\end{bmatrix} = \begin{bmatrix}
\lambda_1^1 \\
\lambda_2^1
\end{bmatrix} \tilde{S}_{1,t} + \begin{bmatrix}
e_{1,t}^1 \\
e_{2,t}^1
\end{bmatrix}
$$

where $\lambda_1^1$ and $\lambda_2^1$ are factor loadings coefficients, and $e_{1,t}^1$ and $e_{2,t}^1$ are measurement errors. Thus, $X_{1,t}^1$ (GDP data) and $X_{2,t}^1$ (IIP data) have a common factor $\tilde{S}_{1,t}$ (output gap). Now, let us regard the equation above as a regression model in which the two data are regressed by a common factor $\tilde{S}_{1,t}$. In this case, the two data are linearly dependent with each other, so that the multi-collinearity problem arises. As is well known, the problem makes the estimated coefficients, $\lambda_1^1$ and $\lambda_2^1$, to be unstable. Moreover, we need to estimate not only factor loadings parameters but also state variable ($\tilde{S}_{1,t}$) itself. The unstable factor loadings coefficient also causes fragile estimates of state variables.

To avoid the problem above, we impose additional restrictions on factor loadings $\Lambda$. In the above example, we set $\lambda_1^1 = 1$ and $\lambda_2^1$ to be estimated. In other words, output gap ($\tilde{S}_{1,t}$) is fundamentally explained by the GDP data ($X_{1,t}^1$), but if the IIP data ($X_{2,t}^1$) has some additional information on the output gap, the marginal increment can be estimated as $\lambda_{2,1}$.

$$
\begin{bmatrix}
X_{1,t}^1 \\
X_{2,t}^1
\end{bmatrix} = \begin{bmatrix}
1 & 0 \\
\lambda_2^1 & 0
\end{bmatrix} \tilde{S}_{1,t} + \begin{bmatrix}
e_{1,t}^1 \\
e_{2,t}^1
\end{bmatrix}
$$

Next, consider the case where there are two model concepts (for example, output gap and inflation gap). Suppose inflation gap $\tilde{S}_{2,t}$ is fundamentally explainable by GDP deflator $X_{1,t}^2$, but CPI data $X_{2,t}^2$ is also assumed to have some information on inflation gap:

$$
\begin{bmatrix}
X_{1,t}^1 \\
X_{2,t}^1 \\
X_{1,t}^2 \\
X_{2,t}^2
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
\lambda_2^1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & \lambda_2^1 & 0 & 0
\end{bmatrix} \begin{bmatrix}
\tilde{S}_{1,t} \\
\tilde{S}_{2,t}
\end{bmatrix} + \begin{bmatrix}
e_{1,t}^1 \\
e_{2,t}^1 \\
e_{1,t}^2 \\
e_{2,t}^2
\end{bmatrix}
$$

Furthermore, we have an extra data $X_{info}^t$ might affect both model concepts (output gap and inflation gap). For example, stock price data could be regarded as the extra data. Then, the measurement equation can be expressed by:

$$
\begin{bmatrix}
X_{1,t}^1 \\
X_{2,t}^1 \\
X_{1,t}^2 \\
X_{2,t}^2 \\
X_{info}^t
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
\lambda_2^1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & \lambda_2^1 & 0 & 0 & 0 \\
\lambda_{info}^1 & \lambda_{info}^2 & \lambda_{info}^3 & \lambda_{info}^4 & 0
\end{bmatrix} \begin{bmatrix}
\tilde{S}_{1,t} \\
\tilde{S}_{2,t} \\
\tilde{S}_{2,t}
\end{bmatrix} + \begin{bmatrix}
e_{1,t}^1 \\
e_{2,t}^1 \\
e_{1,t}^2 \\
e_{2,t}^2 \\
e_{info}^t
\end{bmatrix}
$$
Again, let us regard the equation above as a regression model. The final equation is \( X_t^{\text{info}} = \lambda_{\text{info}}^1 \bar{S}_{1,t} + \lambda_{\text{info}}^2 \bar{S}_{2,t} + e_t^{\text{info}} \), which implies the coefficient \( \lambda_{\text{info}}^1 \) can be interpreted as the marginal contribution for \( \bar{S}_{1,t} \) (output gap) of explaining \( X_t^{\text{info}} \) (stock price data), by controlling the effect of the \( \bar{S}_{2,t} \) (inflation gap). Conversely, the information on the additional data \( X_t^{\text{info}} \) is reflected to estimate the output gap \( \bar{S}_{1,t} \). In this way, utilizing the extra data could contribute to accurate estimation of model concepts.

Finally, suppose that we have an another model concept \( \bar{S}_{3,t} \) (for example, shadow price of capital) and the information series \( X_t^{\text{info}} \) (stock price) also has some beneficial information for the shadow price. Also, \( X_t^{\text{info}} \) will affect all model concepts (output gap, inflation gap and shadow price of capital). Then, with some rearrangement, we can write the measurement equation as follows:

\[
\begin{bmatrix}
X_1^{\text{sensor}} \\
X_2^{\text{sensor}} \\
X_1^{\text{info}} \\
X_2^{\text{info}} \\
X_3^{\text{info}}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\lambda_2 & 0 & 0 \\
0 & \lambda_2^2 & 0 \\
\lambda_{\text{info}}^1 & \lambda_{\text{info}}^2 & \lambda_{\text{info}}^3
\end{bmatrix}
\begin{bmatrix}
S_{1,t} \\
S_{2,t} \\
S_{3,t}
\end{bmatrix}
+ \begin{bmatrix}
e_1^{\text{t}} \\
e_2^{\text{t}} \\
e_2^{\text{info}}
\end{bmatrix}
\]

where \( \lambda_{\text{info}}^3 \) is the factor loadings coefficient on \( \bar{S}_{3,t} \).

According to this idea, we categorize the data indicator \( X_t \) into two types. The first type is referred to as \textit{sensor series} which is likely to embody a certain model concept and link to just one variable of the concept. In the above example, \( X_1^{\text{sensor}} \) (GDP data), \( X_2^{\text{sensor}} \) (IIP data), \( X_3^{\text{sensor}} \) (GDP deflator) and \( X_4^{\text{sensor}} \) (CPI) correspond to sensor series indicators. The sensor series is the data modeled as to how it relates to the model concepts. The other type is referred to as \textit{information series}. This data is not directly connected to any specific model concept but seems to hold useful information on a number of model concepts. In this example, it is \( X_t^{\text{info}} \) (stock price data). The classification of data indicator depends on how the factor loadings matrix \( \Lambda \) is specified.

Also, unlike the common factor \( F_t \) in DFM, the specification of factor loadings matrix leads to the classification of the model concepts \( S_t \) into two types: One can be considered to directly correspond to data indicators \( (S_t^{\text{link}}) \). In the example above, \( \bar{S}_{1,t} \) (output gap) and \( \bar{S}_{2,t} \) (inflation gap). The other does not directly link to data indicator \( (S_t^{\text{non}}) \). For example, model concepts such as shadow price of physical capital \( (\bar{S}_{3,t}) \) or rental rate of physical capital are not necessarily clear as to which data indicator to match. In sum, to capture types of data indicators and model concepts, we rewrite the measurement equation (3.11) as (3.14):

\[
\begin{bmatrix}
X_t^{\text{sensor}} \\
X_t^{\text{info}}
\end{bmatrix}
= \begin{bmatrix}
\Lambda_1^{\text{sensor}} \\
\Lambda_1^{\text{info}} \\
\Lambda_2^{\text{info}}
\end{bmatrix}
\begin{bmatrix}
S_t^{\text{link}} \\
S_t^{\text{non}}
\end{bmatrix}
+ \begin{bmatrix}
e_t^{\text{t}} \\
e_t^{\text{info}}
\end{bmatrix}
\]

(3.14)

where \( X_t^{\text{sensor}} \) and \( X_t^{\text{info}} \) indicate sensor series and information series, respectively. The matrices \( \Lambda_1^{\text{sensor}}, \Lambda_1^{\text{info}} \) and \( \Lambda_2^{\text{info}} \) are partitioned matrices of the factor loading matrix \( \Lambda \). Let \( S_t^{\text{link}} \) and \( S_t^{\text{non}} \) can be represented by \( N^{\text{link}} \times 1 \) and \( N^{\text{non}} \times 1 \) vectors of model concepts, respectively. And suppose that we will link each model concept in \( S_t^{\text{link}} \) with \( p \) sensor series data, \( X_t^{\text{sensor}} = [X_{1,t}^{\text{sensor}}, X_{2,t}^{\text{sensor}}, \ldots, X_{p,t}^{\text{sensor}}] \). Then, we can express the factor loadings matrix \( \Lambda^{\text{sensor}} \) as follows:
where \( I \) is the identity matrix \((N^{\text{link}} \times N^{\text{link}})\), \( \Lambda_p \) is the diagonal matrix \((N^{\text{link}} \times N^{\text{link}})\), \( p \) is the number of the sensor series matching with each model concept and \( L \) denotes the number of sensor series \( X_t^{\text{sensor}} \), that is, \( L = p \times N^{\text{link}} \). On the other hand, the relationship between the information series \( X_t^{\text{info}} \) and state variables \( \vec{S}_t \) are similar to those in the DFM so that matrix \( \Lambda^{\text{info}} \) has full elements without zeros and no additional restrictions on the sign and magnitudes for \( \Lambda^{\text{info}} \) from economic theory.

The framework can be represented as the element-base as shown below. Basically, the one-to-one matching in the sensor series and the many-to-many matching in the information series. For the first sensor series, \( \lambda \) is set to unity and for the remaining sensor series, to be estimated. For the information series, full elements of the factor loadings matrix are estimated.

\[
\begin{bmatrix}
\text{Output Gap series #1} \\
\text{inflation series #1} \\
\vdots \\
\text{Output Gap series #2} \\
\text{inflation series #2} \\
\vdots \\
\text{Output Gap series #n} \\
\text{inflation series #n} \\
\vdots \\
\text{information series #1} \\
\vdots \\
\text{information series #ni} \\
\end{bmatrix}
\begin{bmatrix}
\text{Output Gap series #1} \\
\text{inflation series #1} \\
\vdots \\
\text{Output Gap series #2} \\
\text{inflation series #2} \\
\vdots \\
\text{Output Gap series #n} \\
\text{inflation series #n} \\
\vdots \\
\text{information series #1} \\
\vdots \\
\text{information series #ni} \\
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
\hat{\gamma}_t \\
\hat{\pi}_t \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\gamma_{n\text{on}} \\
\gamma_{n\text{on}} \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\vec{S}_t \\
\vec{S}_t \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\epsilon_{y1_t} \\
\epsilon_{\pi1_t} \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\epsilon_{y2_t} \\
\epsilon_{\pi2_t} \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\epsilon_{y2_t} \\
\epsilon_{\pi2_t} \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\epsilon_{yn_t} \\
\epsilon_{\pi n_t} \\
\vdots \\
\end{bmatrix} \\
\begin{bmatrix}
\epsilon_{i1_t} \\
\epsilon_{in_t} \\
\vdots \\
\end{bmatrix} \\
\end{bmatrix}
\end{bmatrix}
\]

where \( \hat{\gamma}_t \) and \( \hat{\pi}_t \) are model concepts, output gap and inflation gap, respectively.

### 3.3 Model

This study will estimate the SW model (2003) by employing the data rich estimation method. The model is exactly the same as in Chapter 2. Since we have already illustrated the microeconomic meanings of equations, we will provide additional explanation for the parameter meanings of interest in each equation. Note that the “hat” variable denotes percent derivation from the steady state: i.e. \( \hat{x} = \ln x - \ln x^{ss} \) where \( x^{ss} \) is the steady state.
3.3.1 Households

The Euler equation on consumption:

\[ \hat{c}_t = \frac{h}{1 + h} \hat{c}_{t-1} + \frac{1}{1 + h} E_t \hat{c}_{t+1} - \frac{1 - h}{(1 + h) \sigma_c} (\hat{R}_t - E_t \hat{\pi}_{t+1}) + \frac{1 - h}{(1 + h) \sigma_c} (u^c_t - E_t u^c_{t+1}), \quad (3.1) \]

We have two targets to be estimated in the Euler equation: the elasticity of intertemporal substitution (IES) parameter, \( \sigma_c \) and the (external) habit persistence \( h \).

The IES parameter \( \sigma_c \) is concerned with the elasticity of consumption against (ex-ante) real interest rate. The larger the IES parameter, the higher the degree of risk aversion (the higher the curvature of the utility function), and the households do not like to fluctuate consumption. Consequently, intertemporal substitution effect becomes weakened, which implies that consumption becomes inelastic to real interest rate. This is an important parameter related to the cost of the business cycle.

When \( h = 0 \), the Euler equation will result in a purely forward-looking equation. However, according to structural VAR analysis, consumption shows a hump-shaped reaction to structural shocks. Thus, current consumption should be decided depending on the previous consumption. The habit persistence \( h \) determines how important the backward term is. Note that the interest elasticity of consumption depends not only on the IES (\( \sigma_c \)), but also on the habit persistence.

Furthermore, the preference shock \( (u^c_t) \) is a typical demand shock of the DSGE model. Positive preference shock means that utility from current consumption increases. We will capture the impact of this shock on the business cycle.

The Euler equation on investment:

\[ \hat{\text{inv}}_t = \frac{1}{1 + \beta} \hat{\text{inv}}_{t-1} + \frac{\beta}{1 + \beta} E_t \hat{\text{inv}}_{t+1} + \frac{\varphi}{1 + \beta} \hat{q}_t + \frac{\beta}{1 + \beta} (E_t u^{\text{inv}}_{t+1} - u^{\text{inv}}_t), \quad (3.2) \]

\( \varphi \), represents the elasticity of investment to the shadow price of capital. Since the SW model introduces investment adjustment cost, current investment also depends on the previous investment. Many empirical studies pointed out the investment adjustment cost shock \( u^{\text{inv}}_t \) is one of the main sources in the business cycle.

The transition equation of capital shadow price:

\[ \hat{q}_t = -(\hat{R}_t - E_t \hat{\pi}_{t+1}) + \frac{1 - \tau}{1 - \tau + \tau^k} E_t \hat{q}_{t+1} + \frac{\tau^k}{1 - \tau + \tau^k} E_t \hat{r}^k_{t+1} + \varepsilon^q_t, \quad (3.3) \]

We will calibrate the subjective discount factor \( \beta \) and the capital depreciation rate \( \tau \), which leads to the steady state of capital rental rate, \( \tau^k \) from \( \beta = 1/(1 - \tau - \tau^k) \). Thus, we have no parameters to be estimated in the equation above. However, we have two unobservable model concepts: The capital shadow price \( \hat{q}_t \) and the capital rental rate \( \hat{r}^k_t \). The former directly affects the current investment decision and the latter influences the marginal cost. Needless to say, \( \hat{q}_t \) and \( \hat{r}^k_t \) are estimated to satisfy the above optimal condition (the FOC with respect to capital holdings). \( \varepsilon^q_t \) is the so-called equity premium shock, which is an ad hoc shock as like markup shocks.
3.3. MODEL

NKPC (wage):

\[
\dot{w}_t = \frac{\beta}{1 + \beta} E_t \dot{w}_{t+1} + \frac{1}{1 + \beta} \dot{w}_{t-1} + \frac{\beta}{1 + \beta} E_t \hat{\pi}_{t+1} \frac{1}{1 + \beta} \gamma_w \hat{\pi}_t + \frac{\gamma_w}{1 + \beta} \hat{\pi}_{t-1}
\]

There are three targets to be estimated: Nominal wage rigidity \( \xi_w \) (Calvo parameter on wage), the inertia parameter of real wage \( \gamma_w \) (wage indexation parameter) and the inverse Frisch elasticity \( \sigma_L \).

A high \( \sigma_L \) increases the marginal disutility of labor supply, which induces an increase in household’s reservation wage. Therefore, the labor supply will be inelastic against real wage. This is the key calibration parameter in the RBC model. High wage elasticity of labor supply is needed to reproduce the high volatility of employment. In other words, \( \sigma_L \) needs to be estimated with a low value if the RBC model is valid.

Most empirical studies report high rigidity of nominal wages. Suppose that we estimate \( \xi_w \) as 0.75, which implies the wage revision probability is 25%. Thus, it turns out that wage revisions come on average four quarters. The higher the \( \xi_w \), the flattened labor supply curve (NKPC), fluctuations in labor demand are directly linked to employment fluctuations without wage fluctuations. Also, high \( \gamma_w \) smooths real wage fluctuations. However, it is well known that data fluctuation of real wages is extremely volatile and noisy.

Given higher nominal wage rigidity and wage inertia, the volatile wage data may not be explained by labor demand fluctuations, but shifts in labor supply curve (shift in NKPC) may be a key explanation of real wage fluctuations. It is the labor supply shock, \( u^L_t \), that brings about the shift of NKPC. However, there is a loophole that does not require such a structural interpretation: The wage markup shock \( \varepsilon^w_t \) is added. This shock does not have any direct influence on other endogenous variables, nor does it have structural parameters as a coefficient, so the wage markup shock can capture real wage fluctuation without difficulty.

The capital accumulation equation:

\[
\dot{K}_t = (1 - \tau) \dot{K}_{t-1} + \tau \dot{INV}_{t-1}
\]

The capital, \( \dot{K}_t \), is also unobservable model concept (state variable) in the SW model.

3.3.2 Firms

The cost minimization condition:

\[
\dot{l}_t = -\dot{w}_t + (1 + \psi) \dot{r}_t^k + \dot{K}_{t-1}
\]

where \( \psi = \psi'(1)/\psi''(1) \) is the inverse of elasticity of the capital utilization cost.

Production function:

\[
\dot{y}_t = \phi \dot{u}_t + \phi \alpha \dot{K}_{t-1} + \phi \alpha \psi \dot{r}_t^k + \phi (1 - \alpha) \dot{l}_t
\]

where \( \phi \) represents one plus share of the fixed cost in production. In the standard new Keynesian model, intermediate goods firms are assumed to be monopolistic price setters. Hence, they will enjoy monopoly profit not only in the short term but also in the long term. This will bring new
entrants and complicate the model. CEE (2005) introduces a fixed cost to eliminate the long term excess profit of firms to avoid the entry-exit problem. The productivity shock \( \hat{u}^p_t \) is a representative of supply shock. The shock affects the marginal cost, which directly causes inflation fluctuations.

NKPC (price):

\[
\hat{\pi}_t = \frac{\beta}{1 + \beta\gamma_p}\hat{E}_t\hat{\pi}_{t+1} + \frac{\gamma_p}{1 + \beta\gamma_p}\hat{\pi}_{t-1} + \frac{(1 - \beta\xi_p)(1 - \xi_p)}{(1 + \beta\gamma_p)\xi_p}\left[\alpha\hat{r}^k_t + (1 - \alpha)\hat{w}_t - \hat{u}^a_t\right] + \xi^p_t
\]  

(3.8)

Similar to NKPC on real wage, there are two parameters to be estimated: Nominal price rigidity \( \xi_p \) (Calvo parameter on price) and the inflation inertia parameter \( \gamma_p \). The higher the \( \xi_p \), the flattened the short-term aggregate supply curve (NKPC), fluctuations in aggregate demand are directly linked to output fluctuations without inflation variations. The shifts in NKPC should be essentially caused by the productivity shock \( \hat{u}^p_t \), but again, the (ad hoc) price markup shock might help to grasp the high frequency inflation variations without any structural interpretation.

### 3.3.3 Miscellaneous Equilibrium Conditions

Market clearing condition:

\[
\hat{y}_t = (1 - \tau k_y - g_y)\hat{c}_t + \tau k_y\hat{\pi}_t + \tau k_y \hat{\pi}^k_t + g_y \hat{u}^a_t
\]  

(3.9)

Monetary policy rule:

\[
\hat{R}_t = \rho_m \hat{R}_{t-1} + (1 - \rho_m) [\mu_\pi \hat{\pi}_{t-1} + \mu_y \hat{y}_t] + \xi^m_t
\]  

(3.10)

One of the purposes employing multiple data is to estimate the monetary policy rule precisely. In particular, it is important to examine whether the \( \mu_\pi \) representing the central bank’s attitude towards inflation would change by utilizing a lot of data (the central bank might also use the data when conducting the monetary policy rule).

### 3.3.4 Persistent Shocks and Forecast Errors

Five persistent shocks are characterized by the AR(1) process with i.i.d-normal error terms as in Chapter 2: Preference shock, investment adjustment cost shock, labor supply shock, productivity shock and government spending shock.

### 3.3.5 Log-linearized Model

The log-linearized models are integrated as follows:

\[
\Gamma_0 \hat{S}_t = \Gamma_1 \hat{S}_{t-1} + \Psi \xi_t + \Pi \eta_t,
\]  

(3.22)

where \( \hat{S}_t \) is a vector of endogenous variables: \( \hat{S}_t = [\hat{y}_t, \hat{\pi}_t, \hat{w}_t, \hat{k}_t, \hat{\pi}^k_t, \hat{\pi}_t, \hat{\pi}^r_t, \hat{\pi}^k_t, \hat{\pi}^p_t, \hat{\pi}^g_t, \hat{\pi}^u_t, \hat{\pi}^a_t, \hat{\pi}^m_t] \). \( \xi_t \) is a vector of exogenous shocks: \( \xi_t = [\xi^r_t, \xi^f_t, \xi^p_t, \xi^g_t, \xi^u_t, \xi^m_t] \). \( \eta_t \) is a vector of forecast errors: \( \eta_t = [\eta^r_t, \eta^f_t, \eta^p_t, \eta^g_t, \eta^u_t, \eta^m_t] \). \( \Gamma_0, \Gamma_1, \Psi, \text{ and } \Pi \) are matrices of parameters. Employing Sims’ (2002) method, we can derive the state equation as the solution of the SW model.
3.4 Estimation Method

3.4.1 State Space Model

When we estimate the state space model (3.11), (3.12) and (3.13) by utilizing a large panel data set, the size of the matrix in transition equations (3.12) and (3.13) is equal to the total number of state variables \( \bar{S}_t \) and measurement errors \( e_t \). This data rich estimation framework drastically increases in the size of the matrix as the number of data \( X_t \) increases. To avoid the problem, the transition equations are transformed by eliminating AR(1) process of measurement errors. Substituting (3.14) into (3.11), the measurement equation can be expressed by:

\[
(I - \Psi L)X_t = (I - \Psi L)\Lambda \tilde{S}_t + \nu_t, \quad \nu_t \sim \text{i.i.d. } N(0, R).
\]

where \( L \) is the lag operator. Let us define \( \tilde{X}_t = X_t - \Psi X_{t-1} \) and \( \tilde{S}_t = [\tilde{S}_t', \tilde{S}_{t-1}']' \). Then, the equation above can be rewritten by:

\[
\tilde{X}_t = \left( \begin{array}{c} \Lambda - \Psi \Lambda \\ \bar{A} \end{array} \right) \tilde{S}_t + \nu_t, \quad \nu_t \sim \text{i.i.d. } N(0, R). \tag{3.15}
\]

Thus, data indicator after the transformation can be represented by i.i.d. random disturbances, \( \nu_t \). Similarly, the (state) transition equation (3.9) also can be rewritten by:

\[
\begin{bmatrix} \tilde{S}_t \\ \tilde{S}_{t-1} \\ \bar{S}_t \end{bmatrix} = \begin{bmatrix} G(\theta) & O & \tilde{S}_{t-1} \\ I & O & \tilde{S}_{t-2} \\ \bar{G} & \bar{G} & \bar{S}_{t-1} \end{bmatrix} + \begin{bmatrix} H(\theta) \\ O \\ \bar{H} \end{bmatrix} \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, Q(\theta)). \tag{3.16}
\]

where \( I \) is a \( N \times N \) identity matrix. Using a state space model (3.15) and (3.16), we can apply the data rich estimation method without facing the increase in the size of state variables matrix. For convenience, we define the parameters set on the measurement equation (3.15) as \( \Gamma = \{ \Lambda, \Psi, R \} \). We estimate the parameters \( \Gamma \) as the Bayesian OLS method proposed by Chib and Greenberg (1994). The algorithm the above is basically the same as that in Chapter 2, but the data rich estimation method differs in that the factor loadings coefficients are also estimated by the Bayesian OLS in the Step (3.2).

3.4.2 Hybrid MCMC Procedure

We estimate the data rich DSGE model by employing a hybrid MCMC method (also referred to as the Metropolis within Gibbs algorithm) following Boivin and Giannoni (2006) and Kryshko (2011). However, unlike previous studies, we adopt the simulation smoother to estimate state variables as in Chapter 2. Our estimation procedure consists of the following five steps:

**Step 1**: Specify the initial parameters \( \theta^{(0)} \) and \( \Gamma^{(0)} \). Set the iteration index \( g = 1 \).

\footnote{Schorfheide (2000) is one of the pioneers to adopt MCMC algorithm in estimating DSGE models. In actual, we extend to the hybrid MCMC algorithm based on the computer codes (written by GAUSS) of Schorfheide (2000). The hybrid MCMC algorithm is described in Gamerman and Lopes (2006, see Chapter 6). It should be noted that we also adopt a speed-up algorithm for sampling state variables developed by Jungbacker and Koopman (2008), but we find that the algorithm drops the accuracy of state variables estimates. Hence, we omit the algorithm from our MCMC procedure.}

\footnote{Kim and Nelson (1999) and Bauwns et al. (1999) provide the detail explanation on the Bayesian estimation method of the state space model.}
Step 2: Solve the model numerically at $\theta^{(g-1)}$ based on Sims’ (2002) method and obtain $G(\theta^{(g-1)})$, $H(\theta^{(g-1)})$, and $Q(\theta^{(g-1)})$.

Step 3: Draw $\Gamma^{(g)}$ from $p(\Gamma \mid \theta^{(g-1)}, X^T)$.

3.1: Generate unobserved state variables $S^{(g)}_t$ from $p(S^T \mid \Gamma^{(g)}, \theta^{(g-1)}, X^T)$ using the simulation smoother proposed by de Jong and Shephard (1995).

3.2: Using $S^{T(g)}_t$, generate parameters $\Gamma^{(g)}$ from $p(\Gamma \mid S^{T(g)}, \theta^{(g-1)}, X^T)$ via the Gibbs sampler as with Chib and Greenberg (1994).

Step 4: Draw structural parameters $\theta^{(g)}$ from $p(\theta \mid \Gamma^{(g)}, X^T)$ using MH algorithm:

4.1: Draw the candidate from proposal density $p(\theta | \theta^{(g-1)})$ and calculate the acceptance probability $q$ as in the following substeps.

$$q = \min \left[ \frac{p(\theta^{(\text{proposal})} \mid \Gamma^{(g)}, X^T) p(\theta^{(g-1)} \mid \theta^{(\text{proposal})})}{p(\theta^{(g-1)} \mid \Gamma^{(g)}, X^T) p(\theta^{(\text{proposal})} \mid \theta^{(g-1)})}, 1 \right].$$

4.2: Accept $\theta^{(\text{proposal})}$ with probability $q$ and reject it with probability $1 - q$. Set $\theta^{(g)} = \theta^{(\text{proposal})}$ when accepted, and $\theta^{(g)} = \theta^{(g-1)}$, otherwise.

Step 5: Set the iteration index $g = g + 1$ and return to Step 2 up to $g = G$.

The algorithm is basically the same as that in Section 2, but the data rich estimation method differs in that the factor loadings coefficients $\Lambda$ are also estimated by Bayesian OLS in Step 3.

3.5 Preliminary Settings and Data

3.5.1 Measurement Equation

The observed variables should be regarded as imperfectly measured, i.e., they are contaminated by measurement errors. The measurement errors are unrelated to the model concepts. The usual methods without measurement errors might lead to biased parameters and structural shocks due to the endogeneity problem by omitting the measurement errors. One of the purposes of this research is to remove the measurement errors (unexplainable components by the model) from the data by extracting common factors from a large number of the data and matching the common factors with model concepts (explainable components by the model). That is, the data rich approach would bring unbiased parameters and structural shocks.

To this end, we consider the following four cases where different restrictions are imposed on the measurement equation (3.14), i.e., on the link between the model concepts $S_t$ and the data indicators $X_t$. Table 3.1 summarizes the four cases (Cases SW, A, B, and C).

(1) Case SW (Regular DSGE)

The first case is based on a regular estimation method of the DSGE model with a small data set (just seven indicators) in a fashion of one-to-one matching of data with model concept. This case just replicates the estimation result of SW (2003). The small data consists of real GDP, GDP deflator, nominal interest rate (call rate), investment, consumption, labor and real wage. The seven model concepts $S^{\text{link}}_t$ are assumed to be perfectly observed by data indicators. Hence, there is no measurement error in this model, i.e., $e_t = 0$. The measurement equation (3.14) can be represented as follows:

$$X^{\text{sensor}}_{1,t} = \begin{bmatrix} I & O \end{bmatrix} \begin{bmatrix} S^{\text{link}}_t \\ S^{\text{nom}}_t \end{bmatrix}$$

(3.17)
Thus, the factor loading matrix $\Lambda^{sensor}$ is set to the identity matrix $I$. That is, the usual method can be regarded as an estimation method by constraining $\Lambda^{sensor}$ as $I$. In other words, the data rich method is a generalized estimation method, since it imposes only loose restrictions on $\Lambda^{sensor}$.

Case SW (regular DSGE) assumes that the seven data indicators perfectly can be combined with seven model concepts. Let us denote the seven data indicators as the “primary indicators” $X_{1,t}^{sensor}$.

(2) Case A

Case A uses the same data in the case above, but the seven model concepts are assumed to be “latent” variables due to imperfect observations, i.e., Case A introduces measurement errors. Note that we do not add the measurement error to nominal interest rate, since the central bank can perfectly control the nominal interest rate. This case can be considered as a benchmark case comparing the following two cases: The cases in a data rich environment with measurement errors.

We can express the measurement equation in Case A as follows:

$$X_{1,t}^{sensor} = [I \ O] \begin{bmatrix} \bar{S}_t^{link} \\ \bar{S}_t^{nom} \end{bmatrix} + e_{1,t}. \quad (3.18)$$

where the primary indicators $X_{1,t}^{sensor}$ is the same as those in Case SW and $e_{1,t}$ represents measurement errors.

(3) Case B

Case B corresponds to the data rich estimation method: We add seven new indicators $X_{2,t}^{sensor}$ and each model concept is connected with two data indicators (“sensor” series). In other words, each model concept has two dependent variables if we regard the measurement equation as the regression model. The measurement equation in Case B can be described as:

$$\begin{bmatrix} X_{1,t}^{sensor} \\ X_{2,t}^{sensor} \end{bmatrix} = \begin{bmatrix} I & \Lambda_{21}^{sensor} \\ \Lambda_{21}^{sensor} & O \end{bmatrix} \begin{bmatrix} \bar{S}_t^{link} \\ \bar{S}_t^{nom} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}. \quad (3.19)$$

where the matrix $\Lambda^{sensor}$ is set as $[I \ \Lambda_{21}^{sensor}]'$ and $\Lambda_{21}^{sensor}$ is a diagonal matrix. In addition to the information of $X_{1,t}^{sensor}$, if $X_{2,t}^{sensor}$ has useful information on the model concept, the factor loadings coefficient will be estimated as non-zero. $e_{2,t}$ denotes measurement errors corresponding to the additional indicators $X_{2,t}^{sensor}$.

(4) Case C

Case C estimates the SW model by utilizing the full information of our entire data set. We further add seven indicators $X_{3,t}^{sensor}$ as the sensor series as does in Case B. Moreover, we introduce 34 additional indicators as the “information” series, $X_{t}^{info}$. The relationship between the information series and the seven model concepts is not explicitly considered in the SW model. Thus, we cannot structurally interpret why this data and that model concept were connected. However, the extra information might bring estimation efficiency when extracting the common factors, i.e. estimation efficiency of model concepts (state variables). The measurement equation in Case C can be represented as:

$$\begin{bmatrix} X_{1,t}^{sensor} \\ X_{2,t}^{sensor} \\ X_{3,t}^{sensor} \\ X_{t}^{info} \end{bmatrix} = \begin{bmatrix} I & \Lambda_{21}^{sensor} & \Lambda_{31}^{sensor} & \Lambda_1^{info} \\ \Lambda_{21}^{sensor} & O & \Lambda_{31}^{sensor} & \Lambda_2^{info} \end{bmatrix} \begin{bmatrix} \bar{S}_t^{link} \\ \bar{S}_t^{nom} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \\ e_{4,t} \end{bmatrix}. \quad (3.20)$$
where the matrix $\Lambda^{\text{sensor}}$ is set to $[I, \Lambda^{\text{sensor}}_{21}, \Lambda^{\text{sensor}}_{31}]$, and $\Lambda^{\text{sensor}}_{31}$ is a diagonal matrix. But $\Lambda^{\text{info}}_1$ and $\Lambda^{\text{info}}_2$ are matrices with full elements. $e_{3,t}$ and $e_{4,t}$ are measurement errors corresponding to the indicators $X_{3,t}^{\text{sensor}}$ and $X_{t}^{\text{info}}$, respectively. Following the measurement equation above, the information series does not exclude the possibility of connecting with all model concepts. However, from the burden of the computations, we impose $\Lambda^{\text{info}}_2 = 0$, that is, the information series $X_{t}^{\text{info}}$ is linked only to the observable states (the primary seven model concepts) $\bar{S}_t$. However, it still does not exclude the possibility that the information series can be linked to all seven model concepts.

It should be noted that, in Cases A, B, and C, we replaced the price and wage markup shocks, $\varepsilon^p_t$ and $\varepsilon^u_t$, by the measurement errors, $e^p_t$ and $e^u_t$. The justification of this approach is two-fold:

First, inflation and wage data are known to be quite volatile in movements compared to other macroeconomic series. As pointed out by Stock and Watson (2007) and Edge and Gurkaynak (2010), especially recent inflation cannot be forecastable. Since we have already confirmed the fact in Chapter 2, we decide those variations should be captured not by markup shocks but by measurement errors. Justiniano and Primiceri (2008a) also adopts our strategy.

Second, the key difference between markup shocks and measurement errors is that the former being structural shocks and the latter being non-structural shocks. Suppose that the SW model still has misspecification problem and cannot explain all inflation and real wage variations. Nevertheless, estimating the SW model without measurement errors implies the structural shocks must replicate the volatile inflation and wage data fluctuations. In this case, biased structural shocks will be estimated. According to the state equation, structural shocks affect the model concept, and model concepts are interdependent with each other. Hence, the bias would spread all of the model concepts. In particular, price and wage markup shocks directly undertake noisy data fluctuations. Moreover, when we let the markup shocks undertake data fluctuations, we cannot interpret their economic meanings structurally. Therefore, to avoid the negative spillover effect to the other latent variables (such as output gap), we filter out the volatile movements in the observed data by measurement errors. Replacing markup shocks with the measurement errors will lead to extract the unbiased common factors and focus on structural movements in inflation and wage data.

3.5.2 Data

The estimation period is, as is the same in Chapter 2, from 1981:Q1 to 1995:Q4 following Sugo and Ueda (2008) to exclude the periods during the zero interest rate policy.

We employ at most 55 quarterly Japanese macroeconomic series, consisting of 21 sensor series and 34 information series as summarized in Table 3.2 (a), (b). The seven primary indicators $X_{1,t}^{\text{sensor}}$ are the same as in Chapter 2.

When selecting the additional data, we refer Bernanke and Boivin (2003) and Boivin and Giannoni (2006) that conducted the data rich estimation method in the U.S. We employed IIPs and price indices of various industries for output gap and inflation gap. Especially, since one of the aims is to estimate the monetary policy rule accurately and examine the effects of monetary policy shock, we utilize ten price indices consisting of three sensor series and seven information series. We also adopt the financial data as information series such as stock prices, money stocks and exchange rates. Again, the relationship between the information indicators and the seven model concepts is not certain. However, it is only because we do not make a structural model that takes into account those data, and in actual, the dynamics of the seven model concepts may be affected by those indicators: If the SW model is expanded to an open economy model, the exchange rate has a clear link with real interest rate (model concept) through the interest rate parity condition (through the FOC of foreign bond holdings and domestic bond holdings). If we incorporate the financial friction
3.6 RESULTS

into the model, asset price fluctuations will have an impact on borrowing behavior and will be explicitly related to investment (model concept). If we employ the money-in-utility (MIU) model, the nominal money balance has a direct effect through the substitution effect of real money holdings with consumption (model concept). Utilizing the information indicators, we can extract the useful information on the model concept dynamics.

Finally, we processed all data indicators to be percent deviations from steady states. Unlike Del Negro et al. (2007) or SW (2007), we extract the trend components by Hodrick-Prescott filter, as does with Chapter 2. We recognize the detrending method is inconsistent with the balanced growth theory, but we take the trend-break into account, especially since the beginning of 1990s in Japan.

3.5.3 Calibration and Prior Settings

Seven structural shocks are assumed to be independent of one another. Equity premium shock $\varepsilon^q_t$ and monetary policy shock $\varepsilon^m_t$ are i.i.d. shocks. The remaining five shocks have inertia and follow an AR(1) process. As shown in Table 3.3, the prior settings of AR (1) coefficients $\rho$ are specified by Beta distribution to satisfy the stationarity condition $\rho \in (0,1)$, and their means and standard deviations are set at 0.85 and 0.10, respectively. The volatilities of i.i.d. shocks $\varepsilon_t$ are specified by the inverse Gamma distribution so as to be take positive values.

The parameters calibrations are the same as in Chapter 2: The subjective discount factor is set to 0.99 (4% annual rate). The capital depreciation rate $\tau$ is 0.025 (10% annual rate). Using those calibrated parameters, we can derive the steady state of capital rental rate as $\bar{r}_k = \frac{1}{\beta} - (1 - \tau)$. The capital income share $\alpha$ is set to 0.30. The steady state of government spending against output ratio $g_y$ is 0.1. The capital output ratio $k_y$ is 1.5. In addition, the parameter $\lambda_w$ is set to 0.05, which implies workers will enjoy net markup rate at 5%.

The remaining structural parameters are to be estimated and the prior settings are the same in Chapter 2 shown as in Table 3.3. The prior mean of the Calve parameters of price and real wage, $\xi_p$ and $\xi_w$, are set to 0.75, which implies the average durations of the current pricing are one year. The IES parameter $\sigma_c$ is set equal to one, which corresponds to the log-utility function on consumption. Finally, the prior mean of the Taylor coefficient against inflation $\mu_\pi$ is set to 1.7 to meet the so-called Taylor principle (which ensures the unique solution in a standard settings).

3.6 Results

3.6.1 Model Concepts

We first examine whether utilizing a large number of data improves estimation efficiency of model concepts (common factors), especially inflation gap and real wage. Figures 3.1 (a), (b) and (c) display estimated model concepts and actual observations (primary data indicators $X^\text{sensor}_{1:t}$) in three cases. The solid and dashed lines indicate data and estimates, respectively. The gray shadings are 90% credible intervals. Note that there is no discrepancy between the model concepts and data in Case SW due to no measurement errors.

Five model concepts, except inflation and real wage, are generally consistent with data in Cases A and B. The high frequency variations in real wage and inflation are detected as noise, and their low frequency components are regarded as common factors. It should be emphasized again that the i.i.d. wage markup shock is replaced with the AR (1) measurement error in our data rich approach. Nevertheless, AR(1) measurement error explains the noisy wage variations. Compared to the results of the conventional estimation method where markup shocks extract the high frequency fluctuations,
this replacement brings significant differences in estimating the model concept, in particular, real wage.

Furthermore, Case C shows different pictures that the model concepts of output, consumption, and labor fluctuate considerably compared to data, whereas inflation and wage are much closer to data variations. To clarify the discrepancy across these three cases, Table 3.4 reports the correlations between sensor series and model concepts. In the primary indicators \(X_{1t}^{sensor}\), the correlations of inflation are relatively stable at about 0.38 in all three cases, whereas those of real wage gradually increases from 0.08 (Case A) to 0.38 (Case C). As moving from Case A to C, output (against real GDP data) goes down from 0.98 to 0.65, consumption (against GDP deflator) from 0.99 to 0.33, and labor (labor index data) from 0.92 to 0.76. When comparing the second and third sensor series \(X_{2t}^{sensor}\) and \(X_{3t}^{sensor}\) within Case C, we can see those real variables show stable correlations: Output (0.99), consumption (about 0.50) and labor (about 0.90).

The results simply indicate the choice of indicator matters in extracting common factors: Two types of IIP data were given as additional information for estimating output (estimation of common factor). Then, it reduces the correlation between the estimated common factor and the primary indicator (real GDP). For consumption and labor as well, Cases B and C added similar data (household consumption and extra hours index). As a result, in both model concepts, common factors were extracted from similar data variations added, rather than primary data variations, which results in the reduction of the correlation between the common factor and primary indicator. Although the correlations in investment are stable and high (about 0.90) in all cases, it is just because additional data (Cases B and C) and primary data (Case A) were just like each other. The decrease in the correlations of inflation and wages against the primary indicator also can be demonstrated by the selection of additional data. Simply because we added two resemble data later, the common factor ceased reproducing the original data (the primary indicator).

It is worth noting that reproducing the primary indicators variations by common factors do not mean improvement of estimation accuracy. Otherwise, if the credible interval of the common factor becomes narrowed by additional information, then estimation accuracy of model concepts is improved. So, let us examine whether enriching the information creates narrow credible intervals of estimated model concepts.

Table 3.5 (a) reports the values of model concepts that averaged over the estimation period (the average width of shaded gray in Figure 3.1). Looking at the case with the smallest credible interval, Case B was 4 (output, inflation, consumption and labor), Case A was 2 (real wage and investment), and surprisingly, Case C was nominal interest rate only. However, this does not necessarily mean that the estimation accuracy of the model concept decreases due to additional data information. This is because the estimation accuracy is rising in Case B. The problem turned out to be that the information series added without considering the structural correspondence (i.e. Case C), rather it decreases the estimation accuracy of the model concept. If unnecessarily adding data that are not explicitly considered in the model being estimated, such as stock price, nominal money balance, exchange rate, even if measurement errors are introduced, estimation accuracy of model concepts (endogenous variables) is not improved.

This result may suggest that the SW model still have a problem in introducing the “structural” shock. We excluded ad hoc two markup shocks, and extracted unique factors (idiosyncratic components) of data into measurement errors, but an ad hoc shock of “equity premium shock” still remains in the SW model. This shock is included in the transition equation of the capital shadow price. If the exchange rate, stock price or nominal money balance influences the capital price, and if capital price fluctuations due to its impact are identified as the equity premium shock, even if
measurement errors are introduced, the impact may spread for overall the model concepts.

It also should be noted that our result is contrast with Boivin and Giannoni (2006) in which the estimation accuracy of model concepts (especially, inflation) was improved in Case C. Our estimate result show Case C is the worst for inflation (Case B is the best).

### 3.6.2 Structural Shocks

Structural shocks are sources of estimated common factors (model concepts) variations. That is, if the estimation accuracy of structural shocks is improved by adding data, the estimation accuracy of the model concepts also naturally increases. Hence, let us turn to the estimated structural shocks.

Figures 3.2 (a) and (b) display the estimated structural shocks with credible bands. Here, the benchmark is the result of Case A which makes data and model concept correspond one to one and introduces measurement error. To compare the credible intervals, Case A’s band is displayed in blue and the comparison target case is displayed in red. It should be noted that this is the i.i.d. structural shock, not the AR (1) shock.

As in the model concept, Table 3.5 (b) compares the average values of the credible intervals across cases. Regarding the case where the estimation accuracy of structural shocks is the most improved (the confidence interval most shrinks), Case B is 4 (investment adjustment cost shock, productivity shock, government spending shock and monetary policy shock), Case A is 2 (preference shock, labor supply shock), Case C is only equity premium shock. Given the estimation results of the model concept, this result is not surprising.

Case B improves the estimation accuracy both of supply shock (productivity shock) and demand shock (monetary policy shock) that affect variations such as output, inflation, and labor. That is why Case B brought the most accurate improvement in estimating model concepts of output, inflation and labor. Also, in Figure 3.2 (a), we can see that the red band is obscured by the blue band (with an only exception of labor supply shock).

Case C contributes greatly to improving the estimation accuracy of the equity premium shock, which is clearly shown in Figure 3.2 (b). However, unlike expectations, additional information such as stock price and exchange rate has resulted in suppressing fluctuations of capital price shock itself. The negative shock indicated by Case A and B at 1989 has also disappeared in Case C. Hence, the previous inference is denied from this result: This ad hoc shock does not absorb fluctuations in exchange rates and stock prices, and it does not affect other endogenous variables. That is, deterioration of the estimation accuracy of Case C is not caused by introduction of this structural shock.

Therefore, enriching information does not necessarily reduce the uncertainty of estimated structural shocks. The two reasons why the data rich approach does not contribute to the efficiency of estimation might be: (1) low correlations between data matching with one model concept, (2) low coherence of data fluctuations with replicated fluctuations by the structural model.

Now, let us move on to consider sources of the business cycle. The AR (1) shock also accumulates the inertia of the shock caused by the previous periods, but i.i.d. shock means the shock occurs in each period. Hence, the estimation result of i.i.d. shock is useful for specifying the source that brought the business cycle. We can summarize the observed facts as follows:

First of all, the monetary policy shock turned from positive shock (monetary tightening) in the late 1980s to the early 1990s, and negative shock (monetary easing) since 1991 after the collapse of the asset price bubble. We can see that the Bank of Japan stabilized the economy by monetary tightening that suppresses the bubble and monetary easing due to the collapse of the bubble. On
the other hand, the fiscal authority’s reaction (government spending shock) has not been observed clearly around the 1990s.

Next, we cannot observe an adverse productivity shock at the time of the collapse of the asset price. Our result is not consistent with the result that the decline of productivity has caused the Japan’s lost decade. Rather, in Cases A and B, positive spikes occurred in the second half of the 1980s, which can be said to be a source that brought about a boom in the latter half of the 1980s.

Third, it is the negative investment adjustment cost shock that underpinned the economy in the lost decade. Since the negative shock indicates a decline of the investment adjustment cost, which leads to the efficiency improvement of investment goods production. This negative shock has continued to decline for about two years after the collapse of the asset price bubble, which supported the economy.

Forth, the labor supply shock is very volatile in Cases B and C and the magnitude is also extremely large. Especially in Case C, it shows the deviation from the steady state by about 40% above and below in the 90% credible interval. The labor supply shock affects the marginal disutility of labor supply. A negative labor supply shock signifies a decrease in labor disutility and causes a drop in the reservation wage. Case C detects a large negative spike when generating asset price bubbles. In other words, a decline in reservation wages of workers contributed to the bubble economy.

Finally, the equity premium shock affecting capital goods prices vibrate greatly before and after the bubble periods in Cases A and B, and (surprisingly) both negative and positive spikes are detected immediately after asset price bubble occurrence.

3.6.3 Structural Parameters

Tables 3.6 (a), (b) and (c), summarize the estimated structural parameters.

Let us first look at the estimated nominal price rigidity $\xi_p$. It gradually increases as data information is added from Case A to C, but it is relatively stable at a high value (about 0.8, that is, the price revision duration is five quarters). In other words, whether the data information is added or not, the aggregate supply curve is horizontal. As a result, fluctuations in aggregate demand are passed solely on to output fluctuations without inflation fluctuations. Therefore, inflation as a model concept was extracted as a common factor in the low frequency region. The inflation inertia parameter $\gamma_p$ is unstable: Case SW (0.3), A (0.8), B (0.8) and C (0.5). However, it has become clear at least that inertia is reduced if noise is not removed from high frequency fluctuations of inflation (Case SW). This is inconsistent with previous studies that pointed out the importance of inflation inertia.

Regarding the nominal wage rigidity, as expected, the Calvo parameter on wages $\xi_w$ is also high-level stable (about 0.7). Therefore, the labor supply curve is also horizontal, fluctuations in labor demand are absorbed solely by employment, not reflected in real wage fluctuations. Similarly, the wage persistency parameter is low in CaseSW (0.4) but high in other cases (about 0.7). Since it is consistently high from Case A, this is the result of introducing measurement error extracted noise, rather than reflecting information of large scale data. The real wage fluctuation as a smoothed common factor was reproduced with high inertia parameters. Again, it was confirmed that introducing measurement error had a large influence on estimated parameters.

Let us turn to parameters related to household preferences. The habit formation parameter $h$ is generally stable. (About 0.5) Households will decide the current consumption while thinking the past and the future, each half. This inertia has a role of reproducing the hump-shaped consumption responses to structural shocks.
The IES parameter $\sigma_c$ is also stable (from 1.0 to 1.4). The current consumption will drop by 0.7\% against real interest rate 1\% rise, through the intertemporal substitution effect. Since the value is not so far from unity (implying the log-utility function), the substitution effect should be offset to some extent by the income effect (the effect of raising the current consumption due to the future income increase caused by the real interest rate rise).

On the inverse Frisch elasticity $\sigma_L$, we obtain the results of dividing into two poles. It is less than 0.6 for Cases SW and C, and around 1.0 for Cases A and B. The reciprocal of this parameter corresponds to the elasticity of labor supply to wage. Therefore, for the 1\% real wage increase, the former will increase by 1.6\% of the labor supply increase, the latter will be 1\%. Together with the previous result, the former is more horizontal the labor supply curve.

Finally, let us check the parameters of the monetary policy rule. In any case, there is high interest rate inertia (about 0.8). The Taylor coefficient for inflation is also stable (around 1.6). In response to 1\% inflation, the central bank will respond with an increase in nominal interest rate by 0.3\% (30 basis points). Due to the high interest rate inertia, it means ultimately a 1.6\% rate hike. Today’s $(1 - \rho_m)\mu_\pi$\% hike will be $\rho_m(1 - \rho_m)\mu_\pi$\% next period, $\rho_m^2(1 - \rho_m)\mu_\pi$\% re-coming, and so on. Thus, since the inertia of interest rate will work (and private agents also recognize the inertia), the ultimate effect is $(1 - \rho_m)\mu_\pi + \rho_m(1 - \rho_m)\mu_\pi + \rho_m^2(1 - \rho_m)\mu_\pi + \cdots = \mu_\pi$.

### 3.6.4 Policy Simulation

Figures 3.3 (a), (b) depict the impulse response functions against a nominal interest rate rise of 1\%. What we can see right away is that no price puzzle has occurred in any case: inflation responds negatively to nominal interest rate hikes. It should be noted that the central bank does not decide the current interest rate in anticipating future inflation. Simulation results are based on the monetary policy rule consistent with private agents’ behaviors that predict the future and decide on current behaviors. In addition, as the output also reacts negatively, this is a negative demand shock that shifts the aggregate demand curve to the left.

Besides, consumption and investment show the hump-shaped responses because the SW model installs real rigidities such as consumption habit formation and investment adjustment cost. We can see there is an effect of depressing the output and consumption up to 0.3\% (investment 0.6\%) about half a year after 1\% rate hike (in Cases A and B). As the labor supply curve is horizontal, declines in the aggregate demand are absorbed exclusively by reduction of employment (about 0.5\%), and real wages do not change much.

Now, we move on to verify the estimation accuracy. According to Figure 3.3 (a), it can be seen that the blue band of Case A covers the red band of Case B (except for real wage). Therefore, the estimation accuracy of the policy effect is improved by the additional information of the data. In particular, Case B has high estimation accuracy of model concept such as inflation and output (Figure 3.1 and Table 3.5 (a)), and the credible band of the monetary policy shock also shrank (Figure 3.2 (a) and Table 3.5 (b)). Hence, policy simulation in Case B can be more reliable. However, in Figure 3.3 (b), we see that the Case C’s red band dominates the Case A’s blue band. That is, additional data unrelated to the model concept will degrade the estimation accuracy of the policy effect.

### 3.6.5 Sources of the Business Cycle

Variance decomposition
Let us identify sources of the business cycle. First, we consider the result of variance decomposition that shows which kind of structural shock explains the variance of model concept. Table 3.7 summarizes the long term variance decomposition in three cases. What is common to all cases is that the investment adjustment cost shock plays a very important role in the business cycle. The difference in results in data rich environments is the role of measurement error and the labor supply shock. Especially, Case C is outputting quite peculiar results, so we think about it later.

The output fluctuations depend mostly on the variation of investment adjustment cost shock (46% for Case A and 40% for Case B). This shock is also the biggest fluctuating factor not only for investment but also for labor and wage. In fact, among the seven model concepts, the shock has acquired the first position of the four model concepts variations. A large amount of literature also pointed out the investment adjustment cost shock (or investment specific technology shock) plays a main role in explaining the business cycle. For example, see Sugo and Ueda (2008) and Kaihatsu and Kurozumi (2012b) in Japan, and Fisher (2006) and Justiniano and Primiceri (2008) in the U.S. In particular, this shock is more important than the neutral technology shock (i.e. TFP shock). TFP shock contribution is 18.6% in Case A (No. 2), but Case B is only 4.6% (No. 4). Contribution of demand shock to output still has a large preference shock (15% for Case A and 12% for Case B). The monetary policy shock contributes 1.6% for both Cases A and B. However, it should be noted that this value is a contribution of the shock deviating from the policy rule, without reflecting on the contribution that the central bank endogenously responded to output.

Next, we consider the sources of inflation fluctuations. Similar to the results in Chapter 2, the majority of the inflation fluctuations are explained by noise of measurement error. The inflation variations explained by the measurement error is 70% for Case A and 65% for Case B. So, the model explained it by 30% in Case A and by 35% in Case B. In other words, it was 5% by which utilizing multiple information could further explain the inflation variation. But the contents are quite different. In Case A, 11% is TFP shock. This can be the result supported from this model as well as previous empirical results. Since it is 11% out of the 30%, it is correct by about 33% (as far as the model can explain) that inflation occurred because productivity declined, or deflation occurred because productivity rose. In Case B, the labor supply shock is the main source, accounting for 22%. It is 22% out of the 35%, so it is correct by 60% that inflation occurred because people did not want to work and the reservation wage went up, or deflation occurred because people wanted to work and the reservation wage went down. Also, this argument roughly holds for wage fluctuations in Case B (Instead, Case A explains wage variations mainly by the investment adjustment cost shock).

Finally, let us see the result of Case C. The explanatory power of inflation and wage fluctuations has increased dramatically (over 75%). Indeed, Case C contains wage and inflation fluctuations within the 90% credible interval in Figure 3.1. And the labor supply shock explained inflation fluctuation by 60% and wage fluctuation by 46%. Thus, Case C can explain them nearly 50% by the workers' motivation. Certainly in Figure 3.2 (b), the i.i.d. labor supply shock showed the greatest fluctuation and oscillates 40% above and below the steady state. That is not all. Among the seven model concepts, only the investment caused the labor supply shock to fall to the top position (only 4% difference, the second position barely). In other words, if we seriously follow Case C, all business cycles will depend just upon the motivations of the workers.

\textsuperscript{6}To explain the high-frequency behavior of inflation, it may be important to extend the model such as introducing financial friction and incorporating the influence of volatile fluctuations in asset prices. Then, we should verify how inflation fluctuations can be explained by the financial friction model. However, it should be noted that it is impossible to compare the fit with the data by the marginal likelihood in the SW model and the financial friction model when estimating the financial friction model by newly adding financial data.
3.6. RESULTS

Historical decomposition

Historical decomposition makes it possible to clarify the contribution of shock in more detail. Accumulating the contribution of all the shocks will be a model concept. The discrepancy between the model concept and the data is the measurement error. In other words, it is a task to confirm details of the model concept seen in Figure 3.1 in detail. It should be noted that this contributes not only to the current contribution but also to the inertia of the past shocks. Historical decomposition is also examined centered on Cases A and B.

Figure 3.4 depicts the historical decomposition of output (real GDP data). Sources of the output fluctuations are investment adjustment cost shock and preference shock in both cases. In a data rich environment, labor supply shock is added to them.

Let us pay attention to around the bubble era in the early 1990s. The boom is primarily due to the positive contribution of the preference shock and the investment adjustment cost shock (through the negative shock, see Figure 3.2 (a), (b)). In a data rich environment, the labor supply shock is pulling out the output, so that reservation wage had increased during the bubble period due to positive labor supply shock (increase in labor disutility), which has pushed the economy down (See also historical decomposition of real wage in Figure 3.6 (b)). The collapse of the bubble occurs prior to the fall of the preference shock. After that the negative contribution of the investment adjustment cost shock begins. But we should care about the difference in shock’s inertia (preference shock inertia is about 50%, and investment adjustment cost shock is about 70%).

In the meantime, the monetary authority had made a negative contribution by monetary tightening so as to suppress overheating of the bubble since 1989, and monetary easing since 1991 when the bubble collapsed (see also Figure 3.9 (a) (b)). Hence, the central bank plays a major contribution to stabilization. On the other hand, the fiscal authority is extremely small in contribution to output, regardless of repeated fiscal stimuli after the collapse of the bubble economy (obvious result from the variance decomposition in Table 3.7).

Of course, the results could differ as the models are different. But the result that investment adjustment cost shock is the main source seems quite robust regardless of the model. Sugo and Ueda (2008) also reached the same conclusion, but it is natural because the same model (the SW model) is estimated (though estimation method is different). By the way, what is the investment adjustment cost shock? (Similarly, what is the investment specific technology shock?) We know that it is a shock on the cost function of investment goods production, but what exactly should we imagine? One interpretation is that the shock corresponds to an amplified investment fluctuations due to the agency cost, as pointed out by Justiniano and Primiceri (2008b). Kaihatsu and Kurozumi (2014a,b) incorporated a financial friction into the SW model and examines the impact of the unanticipated firm’s net worth deterioration to output fluctuations. They found it was the main source of the business cycle in the U.S. (Kaihatsu and Kurozumi 2014a) but not in Japan (Kaihatsu and Kurozumi 2014b) and again, they concluded the Japan’s lost decade was caused by an unexpected decline of the investment specific technology.

Inflation is the same fluctuation factor as the output, but in Case A, TFP shock is also a major factor. Instead in Case B, the labor supply shock matters. At the beginning of the bubble era, inflation was caused by positive preference shock and negative investment adjustment cost shock. In

---

7 Kobayashi and Inaba (2006) and Otsu (2011) provided the evidence that wage fluctuations through the labor market friction (called as labor “wedge”) had a key role on output fluctuations by applying the business cycle accounting (BCA) to Japanese data.
Case B, an increase in reservation wage due to a rise in labor disutility will add to the explanatory factor of inflation. On the other hand, in Case A, the deflation pressure can be confirmed by the increase in TFP.

It should be noted that investment adjustment cost shock seems to be a supply shock from the name, but according to this result, this is a demand shock. The decline in the adjustment cost of investment means the improvement of productivity of investment goods. This implies right shift of investment goods supply curve, which lowers investment goods price. This alone is a supply shock. However, the decline of the invest goods price simultaneously triggered demand for investment goods by households. Thus, the right shift of investment goods demand curve also arises at the same time. Looking at the historical decomposition, the adjustment cost shock has made a positive contribution to both output and inflation. In other words, the investment adjustment cost shock is a demand shock.

Finally, monetary tightening after 1989 and monetary easing after 1991 have little effect and are not the main cause of inflation fluctuations.

3.6.6 Discussion

This subsection discusses the usefulness of the data rich estimation method, especially what is happening in Case C. First, we summarize our methodologies and results.

Case B extracted common factors using 21 data series and estimated the model by matching of the common factors with the model concepts. We have seven model concepts and associated one model concept with three data as “sensor” series.

On the other hand, Case C extracted common factors using 55 data series. In the same way as Case B, we matched one model concept with three sensor series (21 series). Regarding the remaining 34 series, we did not have considered explicit link with model concepts, but the data series seem to be related to the concepts. So we employ the 34 series as “information” series and connected each of them with “all” seven model concepts.

In both cases, measurement errors are introduced in all data.

As a result, Case B improved the estimation accuracy of the model concepts and structural shocks. We confirmed that the estimation accuracy increased in four of the seven model concepts and structural shocks.

By contrast, Case C did not improve the estimation accuracy of the seven model concepts and structural shocks. It was also found that the estimation accuracy is lower than that of Case A, which does not reflect additional information on data, rather than Case B.

The difference between both cases is the information series. Information series are data without corresponding model concepts such as stock price, nominal money balance, exchange rate.

But if we introduce a financial friction into the model, the stock price appears on the model as a collateral constraint. Changes in collateral value endogenously through stock price fluctuations will create a channel that affect investment of our model concept. If we change the model to a MIU model, fluctuations in the nominal money balance are directly linked to changes in real money balance. Through the substitution effect of real money balance and consumption, the fluctuation in the nominal money balance influences the consumption of our model concept. If the model is extended to an open economy model, the exchange rate is adjusted according to the interest parity condition, which affects the trade balance, and causes the output fluctuations of our model concept.

In other words, there is a possibility that data variation to be included in the model concept is overlooked by misspecification of the model. Therefore, the basic idea for Case C was to extract information of additional data on model concepts without considering the linkage explicitly.
Why does the estimation accuracy drop in Case C? If we add one data, parameters to be estimated will increase by at least three. The factor loadings coefficient, AR (1) parameter and variance of measurement error. But this is the case for the sensor series (Case B). That is, the correspondence relationship with the model concept is clear in the model.

The information series is associated with all seven model concepts, since the relation to the model concept is not clear. Hence, there are seven factor loading coefficients to be estimated. As a consequence, adding one information series brings nine estimated parameters to be estimated.

Suppose that we added data that is solely irrelevant to the model concepts. In this case, almost all variations in additional data are identified as noise (measurement error) and should not affect common factors. However, the number of parameters to be estimated does not change. In fact, despite not providing any additional information, only the estimated parameters will increase, which leads to deterioration in model concept and structural shock estimation accuracy.  

So how do we choose additional data? Looking at the result of Case B, adding sensor series improves estimation accuracy. Since Case B considers the relationship between the model concept and the data, adding one data brings only three parameters to be estimated.

Therefore, our conclusion is that the data rich estimation method is useful for improving estimation accuracy. However, we should not select data that does not necessarily have a clear relationship with the model concept. Alternatively, when there is no model concept corresponding to additional data, the data rich estimation method should be adopted after expanding the model and explicitly considering the correspondence relationship with the data.

3.7 Conclusion

Is the data rich estimation method with measurement error useful?: Can the method improve estimation accuracy?

Basically, Yes, but there is one condition. The advantage of utilizing a lot of data is prominently shown in Case B. It was revealed that highly accurate endogenous variables (including inflation) and structural shocks can be estimated by matching common factors extracted from a large number of data with model concepts (endogenous variables).

Also, extracting common factors of a large number of data makes it easy to identify unique factors (idiosyncratic components) of data. That is, it is easy to separate measurement errors which cannot be explained by the model in data fluctuations. As a result, it became clear that about 65% of inflation fluctuations are noise.

SW (2007) explained the noise with an ad hoc markup shock. However, if we explain the volatile fluctuations only by the structural shocks without measurement errors, the estimation biases by omitting variables will affect all of the endogenous variables through the state equation. Spreading these estimation biases can be avoided by employing the data rich estimation method introducing measurement errors.

On applied empirical analysis, when the monetary policy rule was estimated with high accuracy, did the Taylor coefficient change? No. It was extremely stable in Japan regardless of estimation methods. The central bank responded with a rate hike of 1.6% against 1% inflation. Was TFP shock the main source of the business cycle in the data rich estimation method? No. The main

---

8 Another possibility is also considered as a reason why the estimation accuracy of Case C has decreased. According to Bai and Ng (2013), the quality of data is crucial for whether or not the estimation accuracy improves in DFM. That is, the estimation accuracy may be sensitive to data selection. Therefore, data such as macroeconomic indicators in China that are important to the Japanese economy are missing, which may have decreased Case C estimation accuracy.
sources are preference shock, investment adjustment cost shock and labor supply shock. It was positive preference shock that towed the boom. Also, negative investment adjustment cost shock was supporting the economy for two years after the collapse of the bubble. Both are demand shocks, not supply shock. Has the monetary authority implemented policies that promote the bubble economy or trigger the collapse? No. It became clear that monetary policy contributed to stabilization by historical decomposition.

As a reservation condition for this method to be useful, additional data should be selected after explicitly considering the model concept and data linkage in the structural model.
### 3.8 Tables and Figures

Table 3.1 Outlines of Based Regular DSGE Model and Three Cases of Data Rich DSGE Models

<table>
<thead>
<tr>
<th></th>
<th>Case SW</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>observable variables; $X_t$</td>
<td>sensor 7</td>
<td>sensor 7</td>
<td>sensor 14</td>
<td>sensor 21</td>
</tr>
<tr>
<td></td>
<td>info 0</td>
<td>info 0</td>
<td>info 0</td>
<td>info 34</td>
</tr>
<tr>
<td></td>
<td>total 7</td>
<td>total 7</td>
<td>total 14</td>
<td>total 55</td>
</tr>
<tr>
<td>measurement error; $e_t$</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>price markup shock; $\varepsilon^p_t$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>wage markup shock; $\varepsilon^w_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Case A, B and C imply three cases of Data-rich DSGE models, in which the measurement errors of observable variables are added in model but the measurement error of interest rate is set to zero. ($e^R_t = 0$).
<table>
<thead>
<tr>
<th>NO</th>
<th>Model Concept</th>
<th>Name of Indicator</th>
<th>Abbreviated Name</th>
<th>Data Proc</th>
<th>FQ</th>
<th>case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Output</td>
<td>National Income (at factor cost) : s.a/GDP deflator</td>
<td>Output (GDP)</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Prices</td>
<td>GDP deflator</td>
<td>Prices</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
</tr>
<tr>
<td>3</td>
<td>Real Wage</td>
<td>nominal wage index: deflated by GDP deflator</td>
<td>Real Wage</td>
<td>HP filter</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Investment</td>
<td>Real Private non-residential investment (Billion yen) : deflated by Chain-type price index (2000=100) : s.a</td>
<td>Investment</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Consumption</td>
<td>Real private demand (Billion yen) : deflated by Chain-type price index (2000=100) : s.a</td>
<td>Consumption</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Interest Rate</td>
<td>Call rate (Uncollaterised Overnight), Average (%)</td>
<td>Interest Rate</td>
<td>HP</td>
<td>APR</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
</tr>
<tr>
<td>7</td>
<td>Labor</td>
<td>work hour index x total employment / labor population</td>
<td>Labor</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Output</td>
<td>Index of Industrial Production -Mining and Manufacturing s.a.</td>
<td>Output 2 (IIP)</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Prices</td>
<td>Consumer Price Index (CPI): General,</td>
<td>Prices 2</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Real Wage</td>
<td>Real Wage Index : Contractual Cash Earnings in all Industries, 30 Employees or more : s.a</td>
<td>Real Wage 2</td>
<td>HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Investment</td>
<td>Business Investment - Total Amount, All industry except finance and insurance industry (million yen)</td>
<td>Investment 2</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Consumption</td>
<td>Consumption Expenditure - General household except Agriculture, Forestry and Fisheries household : yen</td>
<td>Consumption 2</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Interest Rate</td>
<td>Short-term Prime Lending Rate of banks (End of Month) (%)</td>
<td>Interest Rate 2</td>
<td>HP</td>
<td>APR</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Labor</td>
<td>Index of Labor Hours : Extra Working Hours, all industries, 30 Employees or more : s.a.</td>
<td>Labor 2</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Output</td>
<td>Index of Industrial Production -Manufacturing s.a.</td>
<td>Output 3</td>
<td>log + HP</td>
<td>sensor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Prices</td>
<td>Consumer Price Index (CPI): General excluding Fresh Food,</td>
<td>Prices 3</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Real Wage</td>
<td>Real Wage Index : Contractual Cash Earnings in Manufacturing, 30 Employees or more : s.a</td>
<td>Real Wage 3</td>
<td>HP</td>
<td>sensor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Investment</td>
<td>Business Investment - Total Amount, Manufacturing, (million yen)</td>
<td>Investment 3</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Consumption</td>
<td>Consumption Expenditure - Worker's household except Agriculture, Forestry and Fisheries household : yen</td>
<td>Consumption 3</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Interest Rate</td>
<td>Average Contract Interest Rates on Loans and Discounts: Short-Term</td>
<td>Interest Rate 3</td>
<td>HP</td>
<td>APR</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Labor</td>
<td>Index of Labor Hours : Extra Working Hours, Manufacturing, 30 Employees or more : s.a.</td>
<td>Labor 3</td>
<td>log + HP</td>
<td>sensor</td>
<td>sensor</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Data Type: sensor series and information series. Data Proc. is data transformation: log, HP, and Def are logarithm, HP filtered, and differenced, respectively. FQ is data frequency: APR is annual percent rate.
<table>
<thead>
<tr>
<th>NO</th>
<th>Model Concept</th>
<th>Name of Indicator</th>
<th>Abbreviated Name</th>
<th>Data Proc</th>
<th>FQ</th>
<th>case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>Consumption</td>
<td>Real Durable goods consumption (Billion yen): s.a</td>
<td>Consumption 4</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Consumption</td>
<td>Real Semidurable Goods Consumption (Billion yen): s.a</td>
<td>Consumption 5</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Consumption</td>
<td>Real Nondurable Goods Consumption (Billion yen): s.a</td>
<td>Consumption 6</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Consumption</td>
<td>Real Service Consumption (Billion yen): s.a</td>
<td>Consumption 7</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Output</td>
<td>Index of Household Consumption Level - General: s.a</td>
<td>Consumption 8</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Output</td>
<td>Index of Industrial Production -durable consumer goods: s.a</td>
<td>Output 4</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Output</td>
<td>Index of Industrial Production -consumer goods: s.a</td>
<td>Output 5</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Output</td>
<td>Index of Industrial Production -nondurable consumer goods: s.a</td>
<td>Output 6</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Output</td>
<td>Index of Industrial Production -producer goods: s.a</td>
<td>Output 7</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Output</td>
<td>Index of Industrial Production -capital goods: s.a</td>
<td>Output 8</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Labor</td>
<td>Employment Index of Regular Workers (All industries, 30 Employees or more): s.a</td>
<td>Labor 4</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Labor</td>
<td>Unemployment Rate (%): s.a</td>
<td>Labor 5</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Labor</td>
<td>Jobs-to-Applicants Ratio: s.a</td>
<td>Labor 6</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Interest Rate</td>
<td>Average Contract Interest Rates on Loans and Discounts: General</td>
<td>Interest Rate 4</td>
<td>HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Interest Rate</td>
<td>Yields to Subscribers-Industrial Bonds AAA CLASS (12 YEARS) (%)</td>
<td>Interest Rate 5</td>
<td>HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Interest Rate</td>
<td>Yields to Subscribers-Bank Bonds (5 YEARS) (%)</td>
<td>Interest Rate 6</td>
<td>HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Interest Rate</td>
<td>Yields to Subscribers-Bearing Government Bond (10 years) (%)</td>
<td>Interest Rate 7</td>
<td>HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Interest Rate</td>
<td>Long-term Prime Lending Rate (End of Month) (%)</td>
<td>Interest Rate 8</td>
<td>HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Prices</td>
<td>Domestic Wholesale Price Index - All Commodities Average, yen base</td>
<td>Prices 4</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Prices</td>
<td>Domestic Wholesale Price Index - Industrial Products, yen base</td>
<td>Prices 5</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Prices</td>
<td>Domestic Wholesale Price Index - Other Industrial Products, yen base</td>
<td>Prices 6</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Prices</td>
<td>Domestic Wholesale Price Index - Agriculture, forestry and fishery products, yen base</td>
<td>Prices 7</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Prices</td>
<td>Consumer Price Index (CPI): Food,</td>
<td>Prices 8</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Prices</td>
<td>Consumer Price Index (CPI): Housing,</td>
<td>Prices 9</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Prices</td>
<td>Consumer Price Index (CPI): Durable goods,</td>
<td>Prices 10</td>
<td>Def + log + HP</td>
<td>APR</td>
<td>info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Financial Markets</td>
<td>Nikkei Stock Average 225 Selected Stocks (End of Month) (yen)</td>
<td>Stock Price 1</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Financial Markets</td>
<td>Nikkei Stock Average 500 Selected Stocks (End of Month) (yen)</td>
<td>Stock Price 2</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Money Stock</td>
<td>Nikkei Commodity Price Index (17 items) (End of Month)</td>
<td>Commodity Price 1</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Money Stock</td>
<td>Nikkei Commodity Price Index (42 items) (End of Month)</td>
<td>Commodity Price 1</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Money Stock</td>
<td>Money Stock: M2+CD (End of Month): s.a (million yen)</td>
<td>Money Stock (M2+CD)</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>Money Stock</td>
<td>Money Stock: M1 (End of Month): s.a (million yen)</td>
<td>Money Stock (M1)</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>Exchange Rates</td>
<td>Foreign Exchange Rate: Tokyo Interbank, Market Spot rate (17h), end of month (yen/&gt;)</td>
<td>Exchange Rates</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Exchange Rates</td>
<td>Foreign Effective Exchange Rate: Real, BIS method</td>
<td>Exchange Rates</td>
<td>log + HP</td>
<td>info</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Data Type: sensor series and information series. Data Proc. is data transformation: log, HP, and Def are logarithm, HP filtered, and differenced, respectively. FQ is data frequency: APR is annual percent rate.
Table 3.3 Prior Distributions of the Parameters under SW (2003)

<table>
<thead>
<tr>
<th>SW (2003) Prior</th>
<th>parameters</th>
<th>meanings</th>
<th>type</th>
<th>mean</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( h )</td>
<td>habit formation</td>
<td>beta</td>
<td>0.7</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \sigma_c )</td>
<td>IES</td>
<td>normal</td>
<td>1</td>
<td>0.375</td>
</tr>
<tr>
<td></td>
<td>( \sigma_L )</td>
<td>inverse Frisch elasticity</td>
<td>normal</td>
<td>2</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>( 1/\varphi )</td>
<td>inverse adj. cost</td>
<td>normal</td>
<td>4</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>( \phi )</td>
<td>fixed cost share</td>
<td>normal</td>
<td>1.45</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>( \psi )</td>
<td>capital utilization cost</td>
<td>normal</td>
<td>0.2</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>( \gamma_p )</td>
<td>price indexation</td>
<td>beta</td>
<td>0.75</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>( \gamma_w )</td>
<td>wage indexation</td>
<td>beta</td>
<td>0.75</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>( \xi_p )</td>
<td>Calvo price no-revise prob.</td>
<td>beta</td>
<td>0.75</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>( \xi_w )</td>
<td>Calvo wage no-revise prob.</td>
<td>beta</td>
<td>0.75</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>( \rho_m )</td>
<td>policy, lag interest</td>
<td>beta</td>
<td>0.8</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \mu_x )</td>
<td>policy, inflation</td>
<td>normal</td>
<td>1.7</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \mu_y )</td>
<td>policy, output</td>
<td>normal</td>
<td>0.125</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>( \rho_a )</td>
<td>persist, productivity</td>
<td>beta</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \rho_c )</td>
<td>persist, preference</td>
<td>beta</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \rho_g )</td>
<td>persist, gov. expenditure</td>
<td>beta</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \rho_L )</td>
<td>persist, labor supply</td>
<td>beta</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \rho_{inv} )</td>
<td>persist, investment</td>
<td>beta</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_c )</td>
<td>preference shock</td>
<td>inv. gamma</td>
<td>0.2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{inv} )</td>
<td>investment shock</td>
<td>inv. gamma</td>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_q )</td>
<td>equity premium shock</td>
<td>inv. gamma</td>
<td>0.4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_a )</td>
<td>productivity shock</td>
<td>inv. gamma</td>
<td>0.4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_p )</td>
<td>price markup shock</td>
<td>inv. gamma</td>
<td>0.15</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_L )</td>
<td>labor supply shock</td>
<td>inv. gamma</td>
<td>1.0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_w )</td>
<td>wage markup shock</td>
<td>inv. gamma</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_g )</td>
<td>gov. expenditure shock</td>
<td>inv. gamma</td>
<td>0.3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_m )</td>
<td>monetary policy shock</td>
<td>inv. gamma</td>
<td>0.1</td>
<td>2</td>
</tr>
</tbody>
</table>

Notes: This study uses the prior based on SW (2003) except for the S. D. of \( \xi_p \) and \( \xi_w \), which are assumed to be 0.15. The means and variances of Beta distribution: \( X \sim \text{Beta}(\alpha, \beta) \) are derived from the following formula, \( E(X) = \alpha/(\alpha + \beta) \) and \( \text{Var}(X) = \alpha\beta/[(\alpha + \beta)^2(\alpha + \beta + 1)] \). The means and variances of Inverse Gamma distribution: \( X \sim \text{IG}(\alpha, \beta) \) are derived from the following formula, \( E(X) = 1/\beta(\alpha - 1) \) for \( \alpha > 1 \) and \( \text{Var}(X) = 1/\beta^2(\alpha - 1)^2(\alpha - 2) \) for \( \alpha > 2 \).
### Table 3.4 Correlation between Observable Variables and Estimated Model Concepts

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Model Concept</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real GDP</td>
<td>Output</td>
<td>0.975</td>
<td>0.969</td>
<td>0.646</td>
</tr>
<tr>
<td>GDP deflator</td>
<td>Inflation</td>
<td>0.368</td>
<td>0.391</td>
<td>0.378</td>
</tr>
<tr>
<td>$X_{1,t}$ Nominal Wage Index / GDP deflator</td>
<td>Real Wage</td>
<td>0.077</td>
<td>0.297</td>
<td>0.375</td>
</tr>
<tr>
<td>Real Private Non-Residential Investment</td>
<td>Investment</td>
<td>0.998</td>
<td>0.991</td>
<td>0.885</td>
</tr>
<tr>
<td>Real Private Demand</td>
<td>Consumption</td>
<td>0.998</td>
<td>0.926</td>
<td>0.329</td>
</tr>
<tr>
<td>Call Rate</td>
<td>Nominal Rate</td>
<td>0.988</td>
<td>0.989</td>
<td>0.995</td>
</tr>
<tr>
<td>Work Hour Index $\times$ Total Employment</td>
<td>Labor</td>
<td>0.921</td>
<td>0.971</td>
<td>0.762</td>
</tr>
<tr>
<td>IIP - Mining and Manufacturing</td>
<td>Output</td>
<td>0.719</td>
<td>0.999</td>
<td></td>
</tr>
<tr>
<td>CPI: General</td>
<td>Inflation</td>
<td>0.496</td>
<td>0.690</td>
<td></td>
</tr>
<tr>
<td>$X_{2,t}$ Real Wage Index : All Industries</td>
<td>Real Wage</td>
<td>0.222</td>
<td>0.549</td>
<td></td>
</tr>
<tr>
<td>Business Investment - Total Amount</td>
<td>Investment</td>
<td>0.964</td>
<td>0.905</td>
<td></td>
</tr>
<tr>
<td>Consumption Expenditure - General Household</td>
<td>Consumption</td>
<td>0.434</td>
<td>0.550</td>
<td></td>
</tr>
<tr>
<td>Short-term Prime Lending Rate of Banks</td>
<td>Nominal Rate</td>
<td>0.930</td>
<td>0.940</td>
<td></td>
</tr>
<tr>
<td>Extra Working Hours, All Industries</td>
<td>Labor</td>
<td>0.813</td>
<td>0.855</td>
<td></td>
</tr>
<tr>
<td>IIP - Manufacturing</td>
<td>Output</td>
<td>0.998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPI: General excluding Fresh Food</td>
<td>Inflation</td>
<td>0.802</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{3,t}$ Real Wage Index : Manufacturing,</td>
<td>Real Wage</td>
<td>0.711</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business Investment - Manufacturing</td>
<td>Investment</td>
<td>0.961</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption Expenditure - Worker’s household</td>
<td>Consumption</td>
<td>0.508</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Contract Interest Rates Short-Term</td>
<td>Nominal Rate</td>
<td>0.925</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra Working Hours, Manufacturing</td>
<td>Labor</td>
<td>0.945</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.5 Average of 90% Credible Interval of Estimated Model Concepts

a) Estimated Model Concepts

<table>
<thead>
<tr>
<th>Model Concept</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>1.014</td>
<td>1.004</td>
<td>3.706</td>
</tr>
<tr>
<td>Inflation</td>
<td>0.952</td>
<td>0.882</td>
<td>1.941</td>
</tr>
<tr>
<td>Real Wage</td>
<td>0.732</td>
<td>0.796</td>
<td>1.790</td>
</tr>
<tr>
<td>Investment</td>
<td>2.879</td>
<td>3.033</td>
<td>4.844</td>
</tr>
<tr>
<td>Consumption</td>
<td>1.628</td>
<td>1.563</td>
<td>7.427</td>
</tr>
<tr>
<td>Nominal Rate</td>
<td>0.278</td>
<td>0.261</td>
<td>0.217</td>
</tr>
<tr>
<td>Labor</td>
<td>1.375</td>
<td>1.068</td>
<td>3.027</td>
</tr>
</tbody>
</table>

b) Estimated Structural Shocks

<table>
<thead>
<tr>
<th>Structural Shocks</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference Shock</td>
<td>4.327</td>
<td>4.329</td>
<td>7.983</td>
</tr>
<tr>
<td>Investment Shock</td>
<td>6.722</td>
<td>5.878</td>
<td>8.498</td>
</tr>
<tr>
<td>Equity Premium Shock</td>
<td>12.359</td>
<td>7.743</td>
<td>2.625</td>
</tr>
<tr>
<td>Productivity Shock</td>
<td>0.970</td>
<td>0.844</td>
<td>1.584</td>
</tr>
<tr>
<td>Government Spending Shock</td>
<td>2.492</td>
<td>1.906</td>
<td>5.253</td>
</tr>
<tr>
<td>Monetary Policy Shock</td>
<td>0.132</td>
<td>0.119</td>
<td>0.130</td>
</tr>
</tbody>
</table>
### Table 3.6 (a) Estimation Parameters of Data rich DSGE

#### a) Case A vs. Regular DSGE

<table>
<thead>
<tr>
<th>parameters</th>
<th>Regular DGE model</th>
<th>Data-Rich DGE Model (Case A)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean [ 90% interval ]</td>
<td>mean [ 90% interval ]</td>
</tr>
<tr>
<td><strong>Structural Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h )</td>
<td>0.544 [ 0.400 0.692 ]</td>
<td>0.525 0.004 0.097 [ 0.367 0.687 ]</td>
</tr>
<tr>
<td>( \sigma_c )</td>
<td>1.498 [ 1.043 1.964 ]</td>
<td>1.404 0.012 0.306 [ 0.919 1.922 ]</td>
</tr>
<tr>
<td>( \sigma_L )</td>
<td>0.572 [ 0.441 0.698 ]</td>
<td>1.043 0.009 0.104 [ 0.894 1.219 ]</td>
</tr>
<tr>
<td>( 1/\varphi )</td>
<td>5.326 [ 3.562 7.055 ]</td>
<td>4.689 0.061 1.216 [ 2.661 6.606 ]</td>
</tr>
<tr>
<td>( \phi )</td>
<td>1.951 [ 1.664 2.239 ]</td>
<td>1.811 0.007 0.191 [ 1.499 2.131 ]</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.151 [ 0.033 0.261 ]</td>
<td>0.212 0.003 0.074 [ 0.093 0.335 ]</td>
</tr>
<tr>
<td>( \gamma_p )</td>
<td>0.304 [ 0.123 0.481 ]</td>
<td>0.822 0.006 0.124 [ 0.658 0.995 ]</td>
</tr>
<tr>
<td>( \gamma_w )</td>
<td>0.395 [ 0.297 0.409 ]</td>
<td>0.677 0.005 0.164 [ 0.424 0.949 ]</td>
</tr>
<tr>
<td>( \xi_p )</td>
<td>0.768 [ 0.711 0.824 ]</td>
<td>0.809 0.002 0.038 [ 0.750 0.872 ]</td>
</tr>
<tr>
<td>( \xi_w )</td>
<td>0.355 [ 0.297 0.409 ]</td>
<td>0.693 0.003 0.061 [ 0.592 0.792 ]</td>
</tr>
<tr>
<td><strong>Policy Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_m )</td>
<td>0.870 [ 0.836 0.907 ]</td>
<td>0.722 0.005 0.064 [ 0.623 0.827 ]</td>
</tr>
<tr>
<td>( \mu_\pi )</td>
<td>1.618 [ 1.442 1.779 ]</td>
<td>1.676 0.004 0.103 [ 1.508 1.843 ]</td>
</tr>
<tr>
<td>( \mu_y )</td>
<td>0.125 [ 0.041 0.206 ]</td>
<td>0.063 0.003 0.053 [ -0.033 0.141 ]</td>
</tr>
<tr>
<td><strong>Shock Persistence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_z )</td>
<td>0.558 [ 0.400 0.715 ]</td>
<td>0.753 0.008 0.140 [ 0.556 0.987 ]</td>
</tr>
<tr>
<td>( \rho_c )</td>
<td>0.511 [ 0.328 0.696 ]</td>
<td>0.519 0.005 0.120 [ 0.324 0.717 ]</td>
</tr>
<tr>
<td>( \rho_g )</td>
<td>0.841 [ 0.730 0.964 ]</td>
<td>0.831 0.004 0.097 [ 0.684 0.978 ]</td>
</tr>
<tr>
<td>( \rho_L )</td>
<td>0.484 [ 0.230 0.762 ]</td>
<td>0.864 0.003 0.080 [ 0.752 0.983 ]</td>
</tr>
<tr>
<td>( \rho_{inv} )</td>
<td>0.677 [ 0.476 0.856 ]</td>
<td>0.780 0.005 0.086 [ 0.638 0.916 ]</td>
</tr>
<tr>
<td><strong>S. D. of Shocks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_c )</td>
<td>3.235 [ 2.245 4.192 ]</td>
<td>2.478 0.045 0.778 [ 1.215 3.655 ]</td>
</tr>
<tr>
<td>( \varepsilon_{inv} )</td>
<td>2.465 [ 1.499 3.383 ]</td>
<td>2.591 0.076 1.022 [ 1.275 3.704 ]</td>
</tr>
<tr>
<td>( \varepsilon_q )</td>
<td>5.174 [ 0.182 9.249 ]</td>
<td>3.873 0.209 3.068 [ 0.162 8.112 ]</td>
</tr>
<tr>
<td>( \varepsilon_z )</td>
<td>9.600 [ 9.120 10.000 ]</td>
<td>0.916 0.048 0.602 [ 0.169 1.827 ]</td>
</tr>
<tr>
<td>( \varepsilon_p )</td>
<td>0.225 [ 0.058 0.421 ]</td>
<td>0.000 NA NA NA NA NA</td>
</tr>
<tr>
<td>( \varepsilon_L )</td>
<td>0.466 [ 0.392 0.538 ]</td>
<td>0.396 0.003 0.057 [ 0.303 0.486 ]</td>
</tr>
<tr>
<td>( \varepsilon_w )</td>
<td>6.683 [ 3.610 9.723 ]</td>
<td>0.000 NA NA NA NA NA</td>
</tr>
<tr>
<td>( \varepsilon_g )</td>
<td>1.948 [ 1.655 2.234 ]</td>
<td>0.806 0.034 0.470 [ 0.149 1.528 ]</td>
</tr>
<tr>
<td>( \varepsilon_m )</td>
<td>0.131 [ 0.110 0.151 ]</td>
<td>0.099 0.001 0.012 [ 0.078 0.119 ]</td>
</tr>
</tbody>
</table>

Notes:
(a) The first 100,000 draws of MH algorithm are discarded to guarantee convergence and then the next 200,000 draws are used for calculating the posterior means, the standard errors of the posterior means (S.E.), the standard deviations (S.D.), the 90% intervals and the convergence diagnostic (CD) statistics proposed by Geweke (1992).
(b) The posterior mean is computed by averaging the simulated draws.
(c) S.E. is computed using a Parzen window with a bandwidth of 10,000.
(d) S.D. is computed as the sample standard deviation of the simulated draws.
(e) The 90% intervals refer to 90% posterior probability bands. These bands are calculated using the 5th and 95th percentiles of the simulated draws.
(f) CD is computed using the criterion as follows.

\[
CD = \frac{\hat{\theta}_A - \hat{\theta}_B}{\sqrt{\hat{\sigma}^2_A/n_A + \hat{\sigma}^2_B/n_B}},
\]

where \(\sqrt{\hat{\sigma}^2_A/n_A}\) and \(\sqrt{\hat{\sigma}^2_B/n_B}\) are the standard errors of \(\hat{\theta}_A\) and \(\hat{\theta}_B\), and we set \(n_A = 40,000\) and \(n_B = 100,000\) and compute \(\hat{\sigma}^2_A\) and \(\hat{\sigma}^2_B\) using a Parzen window with bandwidths of 4,000 and 10,000 respectively.
### Table 3.6 (b) Estimation Parameters of Data rich DSGE

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Regular DGE model</th>
<th>Data-Rich DGE Model (Case B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean [90% interval]</td>
<td>mean S.E. S.D. [90% interval]</td>
</tr>
<tr>
<td><strong>Structural Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h$</td>
<td>0.544 [0.400 0.692]</td>
<td>0.528 0.004 0.097 [0.374 0.693]</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>1.498 [1.043 1.964]</td>
<td>1.376 0.012 0.301 [0.879 1.865]</td>
</tr>
<tr>
<td>$\sigma_L$</td>
<td>0.572 [0.441 0.698]</td>
<td>0.955 0.004 0.053 [0.872 1.045]</td>
</tr>
<tr>
<td>$1/\varphi$</td>
<td>5.326 [3.562 7.055]</td>
<td>5.103 0.069 1.242 [3.081 7.149]</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.951 [1.664 2.239]</td>
<td>1.840 0.007 0.191 [1.516 2.138]</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.151 [0.033 0.261]</td>
<td>0.218 0.002 0.071 [0.097 0.333]</td>
</tr>
<tr>
<td>$\gamma_p$</td>
<td>0.304 [0.123 0.481]</td>
<td>0.774 0.007 0.140 [0.516 0.714]</td>
</tr>
<tr>
<td>$\gamma_w$</td>
<td>0.395 [0.297 0.409]</td>
<td>0.705 0.004 0.161 [0.457 0.963]</td>
</tr>
<tr>
<td>$\xi_p$</td>
<td>0.768 [0.711 0.824]</td>
<td>0.838 0.001 0.032 [0.785 0.891]</td>
</tr>
<tr>
<td>$\xi_w$</td>
<td>0.355 [0.297 0.409]</td>
<td>0.698 0.002 0.048 [0.619 0.776]</td>
</tr>
<tr>
<td><strong>Policy Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_m$</td>
<td>0.870 [0.836 0.907]</td>
<td>0.736 0.004 0.053 [0.652 0.822]</td>
</tr>
<tr>
<td>$\mu_\pi$</td>
<td>1.618 [1.442 1.779]</td>
<td>1.648 0.003 0.100 [1.485 1.814]</td>
</tr>
<tr>
<td>$\mu_y$</td>
<td>0.125 [0.041 0.206]</td>
<td>0.074 0.002 0.040 [0.009 0.140]</td>
</tr>
<tr>
<td><strong>Shock Persistence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_z$</td>
<td>0.558 [0.400 0.715]</td>
<td>0.635 0.005 0.117 [0.441 0.830]</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>0.511 [0.328 0.696]</td>
<td>0.538 0.005 0.118 [0.352 0.739]</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>0.841 [0.730 0.964]</td>
<td>0.825 0.005 0.100 [0.677 0.978]</td>
</tr>
<tr>
<td>$\rho_L$</td>
<td>0.484 [0.230 0.762]</td>
<td>0.790 0.003 0.069 [0.678 0.901]</td>
</tr>
<tr>
<td>$\rho_{inv}$</td>
<td>0.677 [0.476 0.856]</td>
<td>0.745 0.003 0.078 [0.629 0.882]</td>
</tr>
<tr>
<td><strong>S.D. of Shocks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_c$</td>
<td>3.235 [2.245 4.192]</td>
<td>2.468 0.034 0.664 [1.405 3.489]</td>
</tr>
<tr>
<td>$\varepsilon_{inv}$</td>
<td>2.465 [1.499 3.383]</td>
<td>2.371 0.039 0.626 [1.370 3.210]</td>
</tr>
<tr>
<td>$\varepsilon_q$</td>
<td>5.174 [0.182 9.249]</td>
<td>1.796 0.149 2.306 [0.142 5.878]</td>
</tr>
<tr>
<td>$\varepsilon_z$</td>
<td>9.600 [9.120 10.000]</td>
<td>5.324 0.130 1.522 [2.958 8.128]</td>
</tr>
<tr>
<td>$\varepsilon_p$</td>
<td>0.225 [0.058 0.421]</td>
<td>0.000 NA NA NA NA</td>
</tr>
<tr>
<td>$\varepsilon_L$</td>
<td>0.466 [0.392 0.538]</td>
<td>0.446 0.003 0.056 [0.355 0.537]</td>
</tr>
<tr>
<td>$\varepsilon_w$</td>
<td>6.683 [3.610 9.723]</td>
<td>0.000 NA NA NA NA</td>
</tr>
<tr>
<td>$\varepsilon_g$</td>
<td>1.948 [1.655 2.234]</td>
<td>0.564 0.028 0.402 [0.130 1.220]</td>
</tr>
<tr>
<td>$\varepsilon_m$</td>
<td>0.131 [0.110 0.151]</td>
<td>0.093 0.000 0.011 [0.075 0.110]</td>
</tr>
</tbody>
</table>

**Notes:**
(a) The first 100,000 draws of MH algorithm are discarded to guarantee convergence and then the next 200,000 draws are used for calculating the posterior means, the standard errors of the posterior means (S.E.), the standard deviations (S.D.), the 90% intervals and the convergence diagnostic (CD) statistics proposed by Geweke (1992).
(b) The posterior mean is computed by averaging the simulated draws.
(c) S.E. is computed using a Parzen window with a bandwidth of 10,000.
(d) S.D. is computed as the sample standard deviation of the simulated draws.
(e) The 90% intervals refer to 90% posterior probability bands. These bands are calculated using the 5th and 95th percentiles of the simulated draws.
(f) CD is computed using the criterion as follows.

\[
CD = \frac{\hat{\theta}_A - \hat{\theta}_B}{\sqrt{\hat{\sigma}^2_A/n_A + \hat{\sigma}^2_B/n_B}},
\]

where $\sqrt{\hat{\sigma}^2_A/n_A}$ and $\sqrt{\hat{\sigma}^2_B/n_B}$ are the standard errors of $\hat{\theta}_A$ and $\hat{\theta}_B$, and we set $n_A = 40,000$ and $n_B = 100,000$ and compute $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_B$ using a Parzen window with bandwidths of 4,000 and 10,000 respectively.
### Table 3.6 (c) Estimation Parameters of Data rich DSGE

**c) Case C vs. Regular DSGE**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Regular DGE model</th>
<th><strong>Data-Rich DGE Model (Case C)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Mean</strong> [90 % Interval]</td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td><strong>Structural Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h )</td>
<td>0.544 [0.400 0.692]</td>
<td>0.693 [0.008 0.112]</td>
</tr>
<tr>
<td>( \sigma_c )</td>
<td>1.498 [1.043 1.964]</td>
<td>0.944 [0.043 0.538]</td>
</tr>
<tr>
<td>( \sigma_L )</td>
<td>0.572 [0.441 0.698]</td>
<td>0.587 [0.004 0.051]</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.151 [0.033 0.261]</td>
<td>0.225 [0.005 0.080]</td>
</tr>
<tr>
<td>( \gamma_p )</td>
<td>0.304 [3.562 7.055]</td>
<td>6.047 [0.145 1.804]</td>
</tr>
<tr>
<td>( \gamma_w )</td>
<td>0.395 [0.297 0.409]</td>
<td>0.645 [0.011 0.180]</td>
</tr>
<tr>
<td>( \xi_p )</td>
<td>0.768 [0.711 0.824]</td>
<td>0.838 [0.005 0.065]</td>
</tr>
<tr>
<td>( \xi_w )</td>
<td>0.355 [0.297 0.409]</td>
<td>0.667 [0.011 0.132]</td>
</tr>
<tr>
<td><strong>Policy Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_m )</td>
<td>0.870 [0.836 0.907]</td>
<td>0.854 [0.005 0.064]</td>
</tr>
<tr>
<td>( \mu_{\pi} )</td>
<td>1.618 [1.442 1.779]</td>
<td>1.623 [0.007 0.109]</td>
</tr>
<tr>
<td>( \mu_y )</td>
<td>0.125 [0.041 0.206]</td>
<td>0.100 [0.004 0.060]</td>
</tr>
<tr>
<td><strong>Shock Persistence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_z )</td>
<td>0.558 [0.400 0.715]</td>
<td>0.767 [0.012 0.152]</td>
</tr>
<tr>
<td>( \rho_c )</td>
<td>0.511 [0.328 0.696]</td>
<td>0.665 [0.009 0.127]</td>
</tr>
<tr>
<td>( \rho_y )</td>
<td>0.841 [0.730 0.964]</td>
<td>0.834 [0.007 0.097]</td>
</tr>
<tr>
<td>( \rho_L )</td>
<td>0.484 [0.230 0.762]</td>
<td>0.744 [0.019 0.228]</td>
</tr>
<tr>
<td>( \rho_{inv} )</td>
<td>0.677 [0.476 0.856]</td>
<td>0.769 [0.010 0.120]</td>
</tr>
<tr>
<td><strong>S. D. of Shocks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_c )</td>
<td>3.235 [2.245 4.192]</td>
<td>3.044 [0.118 1.430]</td>
</tr>
<tr>
<td>( \varepsilon_q )</td>
<td>5.174 [0.182 9.249]</td>
<td>0.797 [0.078 1.104]</td>
</tr>
<tr>
<td>( \varepsilon_p )</td>
<td>9.600 [9.120 10.000]</td>
<td>8.368 [0.061 0.730]</td>
</tr>
<tr>
<td>( \varepsilon_w )</td>
<td>0.948 [1.655 2.324]</td>
<td>2.073 [0.557 0.691]</td>
</tr>
<tr>
<td>( \varepsilon_m )</td>
<td>0.131 [0.110 0.151]</td>
<td>0.108 [0.001 0.015]</td>
</tr>
</tbody>
</table>

**Notes:**

(a) The first 100,000 draws of MH algorithm are discarded to guarantee convergence and then the next 200,000 draws are used for calculating the posterior means, the standard errors of the posterior means (S.E.), the standard deviations (S.D.), the 90% intervals and the convergence diagnostic (CD) statistics proposed by Geweke (1992).

(b) The posterior mean is computed by averaging the simulated draws.

(c) S.E. is computed using a Parzen window with a bandwidth of 10,000.

(d) S.D. is computed as the sample standard deviation of the simulated draws.

(e) The 90% intervals refer to 90% posterior probability bands. These bands are calculated using the 5th and 95th percentiles of the simulated draws.

(f) CD is computed using the criterion as follows.

\[
\text{CD} = \frac{\bar{\theta}_A - \bar{\theta}_B}{\sqrt{\hat{\sigma}^2_A/n_A + \hat{\sigma}^2_B/n_B}},
\]

where \( \sqrt{\hat{\sigma}^2_A/n_A} \) and \( \sqrt{\hat{\sigma}^2_B/n_B} \) are the standard errors of \( \bar{\theta}_A \) and \( \bar{\theta}_B \), and we set \( n_A = 40,000 \) and \( n_B = 100,000 \) and compute \( \hat{\sigma}^2_A \) and \( \hat{\sigma}^2_B \) using a Parzen window with bandwidths of 4,000 and 10,000 respectively.
### Table 3.7 Variance Decomposition of Model Concept for Long Term Shock

<table>
<thead>
<tr>
<th>Shock / Model Concepts</th>
<th>Y</th>
<th>$\pi$</th>
<th>W</th>
<th>Inv</th>
<th>C</th>
<th>R</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference Shock</td>
<td>0.152</td>
<td>0.017</td>
<td>0.013</td>
<td>0.030</td>
<td>0.399</td>
<td>0.101</td>
<td>0.142</td>
</tr>
<tr>
<td>Investment Shock</td>
<td>0.459</td>
<td>0.104</td>
<td>0.304</td>
<td>0.739</td>
<td>0.247</td>
<td>0.303</td>
<td>0.361</td>
</tr>
<tr>
<td>Equity Premium Shock</td>
<td>0.027</td>
<td>0.003</td>
<td>0.009</td>
<td>0.048</td>
<td>0.008</td>
<td>0.010</td>
<td>0.024</td>
</tr>
<tr>
<td>LaborSupply Shock</td>
<td>0.106</td>
<td>0.061</td>
<td>0.033</td>
<td>0.045</td>
<td>0.085</td>
<td>0.126</td>
<td>0.068</td>
</tr>
<tr>
<td>Productivity Shock</td>
<td>0.186</td>
<td>0.111</td>
<td>0.170</td>
<td>0.113</td>
<td>0.154</td>
<td>0.233</td>
<td>0.246</td>
</tr>
<tr>
<td>Government Spending Shock</td>
<td>0.019</td>
<td>0.001</td>
<td>0.003</td>
<td>0.003</td>
<td>0.009</td>
<td>0.007</td>
<td>0.021</td>
</tr>
<tr>
<td>Monetary Policy Shock</td>
<td>0.016</td>
<td>0.003</td>
<td>0.002</td>
<td>0.005</td>
<td>0.014</td>
<td>0.216</td>
<td>0.016</td>
</tr>
<tr>
<td>All Shocks</td>
<td>0.966</td>
<td>0.300</td>
<td>0.534</td>
<td>0.982</td>
<td>0.915</td>
<td>0.996</td>
<td>0.878</td>
</tr>
<tr>
<td>Measurement Error</td>
<td>0.034</td>
<td>0.700</td>
<td>0.466</td>
<td>0.018</td>
<td>0.085</td>
<td>0.004</td>
<td>0.122</td>
</tr>
<tr>
<td><strong>Case B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference Shock</td>
<td>0.123</td>
<td>0.013</td>
<td>0.012</td>
<td>0.025</td>
<td>0.346</td>
<td>0.077</td>
<td>0.114</td>
</tr>
<tr>
<td>Investment Shock</td>
<td>0.399</td>
<td>0.067</td>
<td>0.286</td>
<td>0.685</td>
<td>0.193</td>
<td>0.205</td>
<td>0.311</td>
</tr>
<tr>
<td>Equity Premium Shock</td>
<td>0.011</td>
<td>0.001</td>
<td>0.004</td>
<td>0.019</td>
<td>0.003</td>
<td>0.003</td>
<td>0.009</td>
</tr>
<tr>
<td>LaborSupply Shock</td>
<td>0.350</td>
<td>0.220</td>
<td>0.201</td>
<td>0.198</td>
<td>0.278</td>
<td>0.425</td>
<td>0.264</td>
</tr>
<tr>
<td>Productivity Shock</td>
<td>0.046</td>
<td>0.049</td>
<td>0.046</td>
<td>0.029</td>
<td>0.034</td>
<td>0.090</td>
<td>0.226</td>
</tr>
<tr>
<td>Government Spending Shock</td>
<td>0.008</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
<td>0.003</td>
<td>0.003</td>
<td>0.009</td>
</tr>
<tr>
<td>Monetary Policy Shock</td>
<td>0.016</td>
<td>0.002</td>
<td>0.002</td>
<td>0.005</td>
<td>0.014</td>
<td>0.194</td>
<td>0.014</td>
</tr>
<tr>
<td>All Shocks</td>
<td>0.953</td>
<td>0.352</td>
<td>0.551</td>
<td>0.962</td>
<td>0.967</td>
<td>0.996</td>
<td>0.946</td>
</tr>
<tr>
<td>Measurement Error</td>
<td>0.048</td>
<td>0.648</td>
<td>0.449</td>
<td>0.038</td>
<td>0.129</td>
<td>0.004</td>
<td>0.054</td>
</tr>
<tr>
<td><strong>Case C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference Shock</td>
<td>0.101</td>
<td>0.052</td>
<td>0.079</td>
<td>0.105</td>
<td>0.233</td>
<td>0.109</td>
<td>0.106</td>
</tr>
<tr>
<td>Investment Shock</td>
<td>0.077</td>
<td>0.024</td>
<td>0.102</td>
<td>0.365</td>
<td>0.040</td>
<td>0.026</td>
<td>0.087</td>
</tr>
<tr>
<td>Equity Premium Shock</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.003</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>LaborSupply Shock</td>
<td>0.451</td>
<td>0.596</td>
<td>0.465</td>
<td>0.324</td>
<td>0.469</td>
<td>0.687</td>
<td>0.390</td>
</tr>
<tr>
<td>Productivity Shock</td>
<td>0.015</td>
<td>0.036</td>
<td>0.074</td>
<td>0.019</td>
<td>0.016</td>
<td>0.028</td>
<td>0.161</td>
</tr>
<tr>
<td>Government Spending Shock</td>
<td>0.088</td>
<td>0.037</td>
<td>0.031</td>
<td>0.020</td>
<td>0.056</td>
<td>0.050</td>
<td>0.096</td>
</tr>
<tr>
<td>Monetary Policy Shock</td>
<td>0.039</td>
<td>0.009</td>
<td>0.016</td>
<td>0.028</td>
<td>0.030</td>
<td>0.100</td>
<td>0.040</td>
</tr>
<tr>
<td>All Shocks</td>
<td>0.771</td>
<td>0.753</td>
<td>0.767</td>
<td>0.864</td>
<td>0.844</td>
<td>0.999</td>
<td>0.880</td>
</tr>
<tr>
<td>Measurement Error</td>
<td>0.229</td>
<td>0.247</td>
<td>0.233</td>
<td>0.136</td>
<td>0.156</td>
<td>0.001</td>
<td>0.120</td>
</tr>
</tbody>
</table>
Figure 3.1 (a) Case A: Smoothed States

Figure 3.1 (b) Case B: Smoothed States

Figure 3.1 (c) Case C: Smoothed States
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Figure 3.2 (a) Case A (blue) vs. Case B (red): Structural Shocks

Figure 3.2 (a) Case B (red) vs. Case A (blue): Structural Shocks
Figure 3.2 (b) Case A (blue) vs. Case C (red): Structural Shocks

Figure 3.2 (b) Case C (red) vs. Case A (blue): Structural Shocks
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Figure 3.3 (a) Case A (blue) vs. Case B (red): IRFs to Monetary Policy Shock

Figure 3.3 (a) Case B (red) vs. Case A (blue): IRFs to Monetary Policy Shock
Figure 3.3 (b) Case A (blue) vs. Case C (red): IRFs to Monetary Policy Shock

Figure 3.3 (b) Case C (red) vs. Case A (blue): IRFs to Monetary Policy Shock
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Figure 3.4 (a) Case A: Historical Decomposition: Output

Figure 3.4 (b) Case B: Historical Decomposition: Output

Figure 3.4 (c) Case C: Historical Decomposition: Output
3.8. TABLES AND FIGURES

Figure 3.5 (a) Case A: Historical Decomposition: Inflation

Figure 3.5 (b) Case B: Historical Decomposition: Inflation

Figure 3.5 (c) Case C: Historical Decomposition: Inflation
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Figure 3.6 (a) Case A: Historical Decomposition: Real Wage

Figure 3.6 (b) Case B: Historical Decomposition: Real Wage

Figure 3.6 (c) Case C: Historical Decomposition: Real Wage
3.8. TABLES AND FIGURES

Figure 3.7 (a) Case A: Historical Decomposition: Investment

Figure 3.7 (b) Case B: Historical Decomposition: Investment

Figure 3.7 (c) Case C: Historical Decomposition: Investment
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Figure 3.8 (a) Case A: Historical Decomposition: Consumption

Figure 3.8 (b) Case B: Historical Decomposition: Consumption

Figure 3.8 (c) Case C: Historical Decomposition: Consumption
3.8. TABLES AND FIGURES

Figure 3.9 (a) Case A: Historical Decomposition: Nominal Interest Rate

Figure 3.9 (b) Case B: Historical Decomposition: Nominal Interest Rate

Figure 3.9 (c) Case C: Historical Decomposition: Nominal Interest Rate
CHAPTER 3. ESTIMATION IN A DATA RICH ENVIRONMENT

Figure 3.10 (a) Case A: Historical Decomposition: Labor

Figure 3.10 (b) Case B: Historical Decomposition: Labor

Figure 3.10 (c) Case C: Historical Decomposition: Labor
Part II

Applications
Chapter 4

Sources of the Great Recession

4.1 Introduction

Chapter 4 examines sources of the Great Recession in the U.S. based on an extended DSGE model (i) introducing the balance sheet conditions of banking and corporate sector and (ii) allowing time-varying shocks volatilities.

Note that Chapter 4 revised “Sources of the Great Recession: Bayesian Approach of a Data-Rich DSGE model with Time-Varying Volatility Shocks,” (joint with Iiboshi, H. and Nishiyama, S.-I.), the discussion paper of Economic and Social Research Institute (ESRI Discussion paper series 313, 2015).

4.1.1 Background

The Great Recession began in December 2007 and ended in June 2009, according to the U.S. National Bureau of Economic Research (NBER). The emergence of sub-prime loan losses in 2007 triggered the recession and exposed other risky loans and over-inflated asset prices. With loan losses mounting and the collapse of Lehman Brothers on September 15, 2008, a major panic broke out on the inter-bank loan market. In the recession, the financial crisis played a significant role in the failure of key businesses, declines in consumer wealth estimated in trillions of U.S. dollars, and a downturn in economic activity leading to the 2008-2012 global recession.

The central debate about the origin of the recession has been focused on the respective parts played by the monetary policy and by the practices of private financial institutions. In order to strengthen the financial sector, the Troubled Asset Relief Program (TARP), in which assets and equity are purchased from financial institutions by the U.S. government, was enforced and originally authorized expenditures of $700 billion in October 2008.

Thus, it would be widely recognized and become a qualitative consensus that the solvency and liquidity problems of the financial intermediaries had a key role in the Great Recession. Then, fiscal and monetary authorities responded to this recession by injecting a large amount of public fund into the banking sector to improve the balance sheet. But, how can we quantitatively extract the impact of the bank sector’s balance sheet loss on the recession? Even if we measure the quantitative effect, was the deterioration of the balance sheet of the banking sector the main source of the recession? Why does the negative impact of the banking sector expand so much? How should we quantitatively measure and quantify the policy effect on the recession?

How was the source identified in the past recession that the U.S. experienced and what was the method to measure the impact? Looking back on past economic turning points, the U.S.
economy had experienced the Great Inflation in the mid 1960s to the late 1970s (pre-Volcker era) and the subsequent Great Moderation since the mid 1980s (post-Volcker era). In the context of macroeconomics, there has been a long debate on what was the source of the Great Moderation.

There are two major arguments: One argument insists on “good policy” that the economic stability came as a result of boldly fighting inflation led by the Chairman Volcker at the beginning of the 1980s (Clarida et al. 2000, Bernanke and Boivin 2003), and the other argument claims simply shocks volatilities increased in the Great Inflation (“bad luck”) and decreased in the Great Moderation (“good luck”) (Cogley and Sargent 2005, Sims and Zha 2006, SW 2007, Justiniano and Primiceri 2008).

Clarida et al. (2000) insisted the policy response to inflation was rising during the Volcker era, which contributed to the Great Moderation in the 1980s, estimating monetary policy rule by GMM and examining the change of responses to inflation gap and output gap (changes in the Taylor coefficients). Bernanke and Boivin (2003) also reported that the Taylor coefficient for inflation was high during the Volcker era, estimating monetary policy rule using predicted CPI by DFM’s factor.

On the other hand, Sims and Zha (2006) concluded the main source of the Great Moderation is the decline of shocks volatilities since the 1980s. They estimated a structural VAR model allowing both regime switches of model parameters and shocks volatilities. Thus, they considered both possibilities that model parameters have changed (changes in monetary policy rules: good policy) and shocks volatilities have changed (good luck). As a result, they showed that the latter model was higher fit for the data.

In recent years, the sources of the Great Moderation have been examined using the empirical DSGE model, instead of reduced-form time series models such as DFM or structural VAR model. SW (2007) estimated the DSGE model by dividing the observation period into two terms (Great Inflation: 1966:Q2-1979:Q2, Great Moderation: 1984:Q1-2004:Q4). There are three main findings: First, the shocks volatilities declined in the second period, in particular, reducing volatilities of productivity shock, monetary policy shock and price markup shock. Second, (surprisingly) there was not much changes in both terms regarding steady state value of inflation and parameters of monetary policy rules. Third, the NKPC was flattened out in the second period through the rise in price and wage nominal rigidities (the rise of Calvo parameters).

To further clarify the sources of the Great Moderation, SW (2007) reports three counterfactual simulations: First, calculating GDP and inflation data volatilities in the second period using shocks volatilities estimated in the first period, then extremely volatile data can be reproduced. Second, calculating GDP and inflation data volatilities in the second period using parameters of monetary policy rules estimated in the first period, then the calculated volatilities did not change much. Finally, calculating GDP and inflation data volatilities in the second period using all estimated parameters in the first period, then there also was not change much. From the results of the three experiments, the biggest reason why volatilities in the second period declined was not because the attitude of the monetary authority to inflation and GDP has changed (not a good policy) but because shocks volatilities have declined (but a good luck).

Instead of dividing the observation period, empirical analysis by explicitly incorporating time-varying shocks volatilities (or stochastic volatility; SV) has also been carried out. Thus, this approach is to estimate the SV model using all of the observation data and examine the transition of the estimated time-varying volatilities. Justiniano and Primiceri (2008) argued the decline in the volatility of investment specific technology shock is the main source of the Great Moderation, by introducing SV to the SW model. Liu et al. (2011) reported the fall of the demand shock volatility called financial shock (capital depreciation shock) is an important source of the Great Moderation,
by estimating the SW model that allows regime-switches for shocks volatilities.

However, the SW model has a problem that the friction in the financial market is not explicitly taken into account such as imperfections in capital goods market or bond market (asymmetric information on capital goods or bond transactions), or incompleteness of bond market (there is no market trading for state contingent claims, etc.). Nevertheless, an ad hoc shock, called as the equity premium shock, has been added in the transition equation of capital goods shadow price (Tobin’s q). As with markup shocks, it is impossible to find a structural interpretation of asset price fluctuations in such ad hoc shocks.

The Great Recession is triggered by the financial crisis as symbolized by the collapse of Lehman Brothers in September 2008. In recent years, empirical studies have been reported to extend to the DSGE model explicitly incorporating the financial frictions to examine the source of the Great Recession.

Carlstrom and Fuerst (1997) and Bernanke et al. (1999, BGG) are pioneering theoretical studies that introduced financial friction into the general equilibrium framework. BGG constructed a financial friction model resulting from asymmetric information between banks and firms on the gross return rate of investment projects.

In this model, firms invest in physical capital by receiving loans from banks, and the gross return rate of investment randomly realizes, but the gross return rate is private information for firms: It is observable for firms but unobservable for banks. In this case, firms have an incentive trying to reduce repayment by telling lies that the realized gross return on investment was low. To eliminate this incentive, banks will lend funds at a high rate (spread is added to the risk-free rate) according to firms’ balance sheet. In equilibrium, the firms’ borrowing rate is determined according to the debt ratio (or the leverage ratio) of the firms’ balance sheet and the spread has been shown to be an increasing function of firm’s leverage ratio.

Under such circumstances, a slight decline in firm’s own asset price or a minor damage of capital stock quality will increase the leverage ratio, raises the spread, and the borrowing rate rises, resulting in firms facing high borrowing constraints. As a result, investment will decrease. In addition, the rise in the borrowing rate lowers the net return rate on investment (= gross return rate on investment - borrowing rate). As asset price is the discounted present value of net returns of future investments, declines in net returns will result in a further decline in asset price. A decline in asset price further increases leverage ratio, which will raise the spread and the borrowing rate will rise. As a result, further investment and net return will decline.

In this way, a trivial negative financial shock, such as a decline in asset price will result in a substantial decline in investment by an amplification effect called as a financial accelerator mechanism.

On the other hand, Gertler and Kiyotaki (2010, hereinafter, GK) and Gertler and Karadi (2010) propose a DSGE model introducing a friction into financial transactions between banks and depositors. In this model, depositors are lenders and banks are borrowers, and the financial friction is caused by bank’s moral hazard behavior that banks can convert deposits into their assets. In this case, an incentive constraint should be imposed by depositors so as for bankers to continue the banking business. As a result, depending on the bank’s net worth ratio (= 1 - debt ratio), depositors will discipline banks through the deposit amount. After all, if bank’s net worth ratio declines, spread between deposit rate and risk-free rate will rise, banks will face higher borrowing constraints, and investment will be hampered, similar to BGG, the financial accelerator mechanism will be generated.

Turning to the empirical side, Christensen and Dib (2008) is one of the earliest studies that
estimated the DSGE model with financial friction. They estimated models incorporating BGG type financial friction and without financial friction, using the data during the Great Moderation in the U.S. As a result, they reported the data fits better for models incorporating financial friction.

Kaihatsu and Kurozumi (2014a, b) explored the sources of the Great Recession in the U.S. and the determinants of the lost decade in Japan, similar to Christensen and Dib (2008), by estimating a model with the BGG type financial friction. According to the result, an adverse financial shock (declining asset price) is the main source of the Great Recession in the U.S., but in Japan, as in Hayashi and Prescott (2002), a contraction supply shock (negative investment specific technology shock) was the main factor of the recession.

Recall that BGG is the financial friction model between banks and corporates, and the friction is based on private information of corporate sector, so the incentive constraint is imposed to the corporate sectors’ balance sheet.

Since the recent Great Recession of the U.S. and Japan’s lost decade, however, caused by the collapse of financial intermediaries such as investment banks, commercial banks, securities companies, etc., there is a possibility that the main source was the deterioration of the bank’s balance sheet rather than damaging the corporate balance sheet.

If so, how should we know which deterioration of the balance sheet of the banking sector and the corporate sector was the main factor behind the Great Recession?

To detect the origin of the recession, we need to decompose the economic downturn into two adverse financial shocks. Hirakata et al. (2011) expanded to a model introducing two financial frictions, i.e. one is the friction between banks and corporates and the other is the friction between depositors and banks. The banking sector mediates the supply of funds from depositors to the corporate sector. Thus, bankers are lenders of firms and borrowers from depositors. They introduced agency costs due to asymmetric information, i.e. BGG-type financial friction, in both banking sector financial transactions. Then, they demonstrated the banking sector’s net worth deterioration was a major factor in the substantial decline in investment.

4.1.2 Purposes, Originalities, and Methodologies

The purpose of this research is to investigate the reason for the Great Recession of U.S. as well as Hirakata et al. (2011), by specifying the DSGE model with balance sheet conditions of both banking sector and corporate sector. However, there are four differences from Hirakata et al. (2011):

First, we introduce time-varying shocks volatilities (SV). In the Great Recession, just opposite the discussion of the Great Moderation, it seems to be a natural question whether there is a possibility of “bad luck” which the financial shock volatility has increased. As the Chairman Greenspan described as “once-in-a-century credit tsunami”, we examine the possibility of bad luck that, the financial shock volatility is small in usual time, but during the financial crisis, we unfortunately experienced a negative financial shock with very rare and substantially large volatility.\(^1\)

However, it cannot necessarily be confirmed that the demand side bad luck occurred during the financial crisis (an increase in the negative financial shock volatility), since the main source may be the supply side bad luck (an increase in a negative investment specific technology shock volatility),

\(^1\)Although we will not deal with it, in finding the sources of the Great Recession, the macro-finance approach, in which the magnitudes of the structural shock volatilities affect the model variables after the first-order approximation around the steady states, is also considered to be very useful. In that approach, assuming that the structural shocks follow the Wiener process, if the model is second-order approximated around the steady states, the magnitudes of structural shock volatilities affect the model variables after linear approximation by Ito’s lemma. See, for example, Brunnermeier and Sannikov (2014).
4.1. INTRODUCTION

as Justiniano and Primiceri (2008) demonstrated in the U.S.

Furthermore, a “good policy” might have occurred simultaneously after the financial crisis: The volatility of monetary policy shock might be increasing after the financial crisis since the bold monetary easing policy led by the Chairman Bernanke should be caught as a large deviation from the monetary policy rule (an increase in a negative monetary policy shock volatility).

It should be noted, on the empirical side, since introducing SV makes linear model nonlinear, we estimated SV by employing the particle filter, following Kim et al. (1998).

Second, we examine an asymmetric amplification effect of structural shock. Based on the idea of Liu et al. (2009), we introduced the so-called leverage effect on time-varying structural shocks volatilities. Thus, we consider the correlation between structural shock innovation and volatility innovation:

Suppose that there is a negative correlation between innovations. Then, a negative structural shock will increase in it’s volatility due to the negative correlation. Conversely, a positive structural shock will lower it’s volatility. That is, when a negative leverage effect is detected, a positive shock will reduce it’s shock through declining it’s volatility, but a negative shock becomes more serious shock through expanding it’s volatility. By examining the sign of the correlation, we can verify whether shock amplification process was asymmetric.

Particularly, looking at the data during the financial crisis, we can observe large positive spikes in the spread between corporate borrowing rate and the interbank rate, and also, the spread between deposit rate and risk-free rate. Thus, in addition to capturing the spikes of these spreads with the financial accelerator model, we will also examine whether a serious bad luck existed such that financial shocks volatilities were further amplified through the effective negative leverage effect.

Third, by conducting data rich estimation, we try to identify financial shocks of corporate sector and banking sector. When asking whether the source was balance sheet deterioration of either banking sector or corporate sector, we need to identify financial shocks of both sectors. However, with the usual estimation method, we face a difficulty to identify which financial shock has caused the recession, since we have qualitatively similar responses of endogenous variables to both sectors’ financial shocks.

When we identify a certain shock as supply shock or demand shock, the shock is regarded as supply shock if output and inflation reactions against the shock are opposite (e.g. output reacts positive but inflation reacts negative), as demand shock if it is the same (e.g. output reacts positive, and inflation reacts positive), since supply shock shifts the upward sloping aggregate supply curve and demand shock shifts the downward sloping aggregate demand curve. That is, the qualitative difference in the impulse response is an important information for identifying the shock. However, output and inflation reactions to financial shocks of corporates and banks qualitatively give the same responses. So, it is difficult to identify between the two shocks by the usual estimation method.

To deal with this identification problem, we employ the data rich estimation method proposed in Chapter 3 to identify both sectors’ financial shocks: Estimating the impulse responses with high estimation accuracy narrows the credible intervals of the responses. By the data rich estimation method, we do not miss the slight differences in responses and try to identify two financial shocks. We match common factors of multiple financial data against financial endogenous variables such as corporate borrowing rate and leverage ratio, using financial data such as several Baa corporates’ loan rates and various banks’ leverage ratios. The data rich estimation method will also improve the estimation efficiencies of financial shocks.

Finally, there is a difference in modeling financial frictions. Hirakata et al. (2011) introduced BGG type friction to both financial transactions. On the other hand, we introduce BGG type
friction between banks and corporates, and GK type friction between banks and depositors. In GK type friction, depositors will discipline banks with the deposit amount, observing bank net worth ratio. Following the current financial crisis, BIS is considering a macro-prudential policy called as the countercyclical buffer, loosening banks’ net worth regulation during financial crisis and tightening banks’ net worth regulation in usual time. By introducing GK type friction, we could carry out a counterfactual simulation on macro-prudential policy aimed at stabilizing the financial system. Note that our financial friction model is constructed by Nishiyama et al. (2011).

In summary, chapter 4 constructs an extended DSGE model incorporating balance sheet conditions of bankers and corporates, investigates which sector’s balance sheet deterioration is the source of the recession by employing data rich estimation method, and examines a bad luck that financial shocks volatilities increased in the financial crisis and a good policy that how much monetary easing policy after the crisis supported the U.S. economy by introducing time-varying shocks volatilities.

4.1.3 My Contributions

Chapter 4 revised Iiboshi et al. (2014) written with two coauthors.

My main contribution is to provide an idea to examine a bad luck or a good policy by introducing time-varying shocks volatilities:

According to Nishiyama et al. (2011) prior to this study, negative spikes of net worth shocks were observed both in the corporate sector and the banking sector at the third quarter 2008 when the Lehman Brothers collapsed, demonstrating that the financial crisis in the U.S. was attributable to the balance sheet deteriorations of both sectors. At the same time, however, it was also found that financial shocks volatilities are high even during normal times.

Following this, as well as the discussion of Great Moderation, while the financial shock volatility is low during normal times, there is a possibility that the financial shock volatility expanded during the financial crisis, that is, I got an idea that bad luck should be examined with SV.

In addition, Nishiyama et al (2011) also found that public capital injection into financial institutions called as the Troubled Asset Relief Program (TARP) supported the U.S. economy. TARP was detected as a negative monetary policy shock after the financial crisis, indicating that the monetary easing policy after the financial crisis was effective as a stabilization policy. With the idea to estimate by allowing time-varying volatility for all structural shocks, it would be also possible to verify a good policy through examining whether monetary policy shock volatility expanded after the financial crisis.

4.1.4 Organization of Chapter 4

Chapter 4 is organized as follows: Section 2 provides a framework of the DSGE model including the data rich approach and structural SV shock with leverage effect. Section 3 illustrates an extended DSGE model with bank and corporate sectors balance sheet conditions. Section 4 explains the estimation method. Section 5 describes preliminary settings and data used for the estimation. Section 6 shows the estimation results and examines the sources of the Great Recession. Section 7 concludes.
4.2 Data Rich Approach with Stochastic Volatility

4.2.1 Introducing Stochastic Volatility with Leverage Effect

Was “bad luck” happening at Great Recession? Especially at this time, was there a possibility that the deterioration of the net worth of corporate and banking sectors could become much worse than in normal times? Specifically, did the volatility of capital stock shock increase at this time? To verify this “bad luck”, we need to relax the assumption of structural shock volatility in normal estimation: We need to extend the model shock volatility to a time-varying model.

The solution of the DSGE model can be represented as follows:

\[
S_t = G(\theta) S_{t-1} + E(\theta) \varepsilon_t, \quad (4.1)
\]

where \( S_t \) is a \( N \times 1 \) vector of endogenous variables, \( \varepsilon_t \) is a \( M \times 1 \) vector of exogenous disturbances (structural shocks) and \( \theta \) is structural parameters. Matrices \( G(\theta) \) and \( E(\theta) \) are the function of \( \theta \).

Usually, disturbance terms \( \varepsilon_t \) are assumed to be i.i.d. normal distributions with time-constant variances. Otherwise, this study extends this assumption to the time varying variances as shown below.

\[
\varepsilon_t = \Sigma_t z_t, \quad (4.2)
\]

\[
z_t \sim \text{i.i.d. } N(0, I_M),
\]

\[
\Sigma_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t}, \ldots \sigma_{M,t}),
\]

where \( z_t \) is a \( M \times 1 \) vector with all elements following standard normal distribution. \( I_M \) is a \( M \times M \) identity matrix. \( \Sigma_t \) is standard deviation; or volatility, of disturbance \( \varepsilon_t \), represented by a diagonal matrix with elements \( \sigma_{1,t}, \ldots, \sigma_{M,t} \). The shocks volatilities are assumed to change over time as follows (called “stochastic volatility” model: SV):

\[
\log \sigma_{i,t+1} = \mu_i + \phi_i (\log \sigma_{i,t} - \mu_i) + \eta_{i,t}, \quad i = 1, 2, \ldots, M, \quad (4.3)
\]

\[
\begin{pmatrix} z_{i,t} \\ \eta_{i,t} \end{pmatrix} \sim \text{i.i.d. } N(0, \Omega_i), \quad \Omega_i = \begin{bmatrix} 1 & \rho_i \omega_i \\ \rho_i \omega_i & \omega_i^2 \end{bmatrix} \quad (4.4)
\]

where \( \mu_i \) is the mean of volatility \( \sigma_{i,t} \) of i-th shock. \( \phi_i \) is the persistence parameter of the i-th volatility. This SV model also introduces the so-called “leverage effect” of volatility \( \sigma_{i,t} \) which is the correlation between the sign of disturbance term and the size of volatility. Especially, if the correlation \( \rho \) is negative, then a negative shock \( (z_{i,t} < 0 \text{ thus } \varepsilon_{i,t} < 0) \) causes an increase in the volatility of the shock \( (\eta_{i,t} > 0) \).

Consider a negative shock has occurred such that net worth declines. The “bad luck” is the case where the magnitude of this shock (volatility) is larger at the time of the Great Recession than usual. We will capture it by estimating time varying volatilities \( \sigma_{i,t} \). In addition, we will also verify further misfortunes that increase its volatility with the leverage effect by estimating the correlation coefficient \( \rho \).
4.2.2 DSGE Model in a Data Rich Environment

Estimation in a Data Rich Environment

Similar to Chapter 3, we adopt the data rich approach, integrating the DSGE model with DFM in order to (1) gain the estimation efficiency, especially for structural shocks and volatilities, and (2) identify the two types of financial shocks (corporate and banking sector’s net worth shocks). This approach complements the disadvantages of the DSGE model and DFM with each other.

(1) Disadvantages of DSGE model and Benefits of DFM

In general, data $X_t$ should be regarded as consisting of two components: “comovement” (or systematic) components and “idiosyncratic” components (measurement errors or noise).

\[
data = \text{common (or systematic) component} + \text{idiosyncratic component},
\]

DSGE model describes the systematic components, thus we should remove the noise (measurement error) in estimating the model. But unfortunately, the standard estimation method must face a difficulty to separate the data $X_t$ into the model variable $S_t$ and the measurement error $e_t$, since both two components are unobservable. Accordingly, the standard method assumes there is no idiosyncratic component or measurement errors, i.e., “data = systematic components”. DFM can separate two factors. In this model, comovement component is explained by an economic system dictated by multiple variables with their mutual impacts. The dynamic equation of comovement components might correspond to the VAR model. On the other hand, idiosyncratic components should be expressed as the univariate AR process, since they can be regarded as fluctuating independently without interdependence.

(2) Disadvantages of DFM, and Benefits of DSGE model

DFM focuses on decomposing data into comovement and idiosyncratic errors. However, is it possible to decompose the data only from a statistical method? If we just care about correlation among data, DFM is useful. But if we also think about the causality among data from the economic theory, it is no longer possible only with DFM. Usually, DFM adopts the VAR model describing comovement components. However, it is difficult to interpret structurally VAR coefficients. Instead, DSGE model can express the comovement of multiple variables with causality and theoretical coherence based on microeconomic foundations. That is, we should replace the VAR model with DSGE model in a systematic component in DFM:

\[
\text{comovement (systematic variation)} = \text{genetic correlation},
\]

\[\Rightarrow \text{comovement} = \text{causal association}\]

which resolves drawback of DFM.

(3) Complementarities

Stock and Watson (2002a,b) suggests that DFM can remove efficiently idiosyncratic components by employing a large number of data $X_t$. It also improves accuracy of estimating comovement components $S_t$ and exogenous structural shocks $\varepsilon_t$. If the estimated structural shocks $\varepsilon_t$ with high accuracy can successfully explain the business cycles, which will bring the validity of the DSGE model. In other words, DFM and DSGE are complementary to each other: DSGE needs
to statistical method to separate data into two unobservable components, and DFM needs to the model explaining the economic system. Thus, integrating DFM with DSGE can complement the needs of both. The data rich approach combines the two.\footnote{The data rich framework, same data set is applicable even for DSGE model with different model variables, so that the possibility of model selection among many alternative models emanates. It implies that data rich approach is expected to contribute the evaluation and selection among DSGE models from the point of view of validity of structural shocks and marginal likelihood (or Bayes factor).}

**DFM**

Sargent and Sims (1977) employs DFM in estimating common factors of the business cycle and Stock and Watson (1989) empirically applied DFM in extracting unobserved cycle component as a common factor from many data using the Kalman filter.\footnote{Stock and Watson (2002a,b) developed approximate DFMs using principal component analysis (PCA), extracting several common factors from more than one hundred data and verifying that these factors include useful information on forecasting data. Nowadays, a large number of literature is accumulating on theoretical and empirical studies by DFMs. For example, Boivin and Ng (2005, 2006), Stock and Watson (2002a, b, 2005). The survey of DFMs covering the latest studies is provided by Stock and Watson (2006, 2010). Kose et al. (2003) adopted DFM in extracting common factors of worldwide and regional business cycles.}

DFM is represented by state space models consisted of the following three equations. Let $F_t$ denote the $N \times 1$ vector of unobserved common factor, and $X_t$ denote the $J \times 1$ vector of massive panel of macroeconomic and financial data. Note that $J \gg N$.

\[
\begin{align*}
X_t & = \Lambda F_t + e_t, \quad (4.5) \\
F_t & = G F_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, Q), \quad (4.6) \\
e_t & = \Psi e_{t-1} + \nu_t, \quad \nu_t \sim \text{i.i.d. } N(0, R), \quad (4.7)
\end{align*}
\]

where $\Lambda$ is $J \times N$ matrix of factor loadings, $e_t$ is the idiosyncratic components (or measurement errors) which are allowed to be serially correlated as equation (4.7). $G$ is $N \times N$ matrix, and common factor $F_t$ follows AR process (4.6). Matrices, $\Psi$, $Q$ and $R$ are assumed to be diagonal in the exact DFM as Stock and Watson (2005). The state space model is consisted of two kinds of equations: The measurement equation is (4.5) and the transition equations are (4.6) and (4.7).

DFM decomposes massive panel of macroeconomic and financial data $X_t$ into common components $\Lambda F_t$ and idiosyncratic component $e_t$ in (4.5). However, the common factor $F_t$ cannot be interpreted structurally since the the VAR model is employed as the above equations in (4.6).

**Integrating DFM with DSGE**

The idea of the data rich approach is to extract the common factor $F_t$ from a large number of data $X_t$ and to match the model variable $S_t$ with the common factor $F_t$. DSGE model is expressed by the state space models and estimated using the Kalman filter as well as the DFM. Thus, we can...
apply the framework of the DFM to DSGE model. The data rich DSGE model can be written as:
\[
\begin{align*}
\mathbf{X}_t &= \mathbf{\Lambda} \mathbf{S}_t + \mathbf{e}_t, \quad (4.8) \\
\mathbf{S}_t &= \mathbf{G}(\theta) \mathbf{S}_{t-1} + \mathbf{E}(\theta) \mathbf{e}_t, \quad (4.9) \\
\mathbf{e}_t &= \mathbf{\Psi}_t \mathbf{e}_{t-1} + \mathbf{\nu}_t, \quad (4.10)
\end{align*}
\]

where observable variables \( \mathbf{X}_t \) are a \( J \times 1 \) vector, state variables \( \mathbf{S}_t \) are a \( N \times 1 \) vector, and structural shocks \( \mathbf{e}_t \) are a \( M \times 1 \) vector. In the data rich DSGE model, the number of observable variables is much larger than that of state variables \( (J \gg N) \) as well as DFM. The idiosyncratic components \( \mathbf{e}_t \) \( (J \times 1) \) means measurement errors following AR (1) process in (4.10). (4.8) is measurement equation which splits off components of common factors \( \mathbf{S}_t \) and idiosyncratic components \( \mathbf{e}_t \) from a lot of indicators \( \mathbf{X}_t \). Structural shocks \( \mathbf{e}_t \) and disturbance terms \( \mathbf{\nu}_t \) of measurement errors \( \mathbf{e}_t \) follow normal distributions, i.e., \( \mathbf{e}_t \sim \text{i.i.d. } \mathcal{N}(0, \mathbf{Q}(\theta)) \) and \( \mathbf{\nu}_t \sim \text{i.i.d. } \mathcal{N}(0, \mathbf{R}) \), respectively. And their variance covariance matrix \( \mathbf{Q}(\theta) \) and \( \mathbf{R} \) are positive definite and diagonal matrix. Coefficients \( \mathbf{\Psi} \) of AR(1) process (4.10) is also diagonal matrix. These imply measurement errors \( \mathbf{e}_t \) are independent with each other in terms of cross section but dependent with their lag variables in terms of time series restriction. Matrices \( \mathbf{G}(\theta), \mathbf{E}(\theta) \) and \( \mathbf{Q}(\theta) \) are functions of structural parameters \( \theta \).

### 4.2.3 DSGE model with Stochastic Volatility in a Data Rich Environment

In the previous subsection, structural shocks \( \mathbf{e}_t \) are assumed to follow i.i.d. normal distribution. Otherwise, we introduce the SV model with leverage effects into the data rich DSGE model. Combining the SV model and the data rich DSGE model, our model is represented by:
\[
\begin{align*}
\mathbf{X}_t &= \mathbf{\Lambda} \mathbf{S}_t + \mathbf{e}_t, \quad (4.11) \\
\mathbf{S}_t &= \mathbf{G}(\theta) \mathbf{S}_{t-1} + \mathbf{E}(\theta) \mathbf{e}_t, \quad (4.12) \\
\mathbf{e}_t &= \mathbf{\Psi}_t \mathbf{e}_{t-1} + \mathbf{\nu}_t, \quad (4.13)
\end{align*}
\]

\[
\begin{align*}
\mathbf{e}_t &= \mathbf{\Sigma}_t \mathbf{z}_t, \quad (4.14) \\
\mathbf{z}_t &\sim \text{i.i.d. } \mathcal{N}(0, \mathbf{I}_M), \\
\mathbf{\Sigma}_t &= \text{diag}(\sigma_{1,t}, \sigma_{2,t} \ldots \sigma_{M,t}), \\
\log \sigma_{i,t+1} &= \mu_i + \phi_i(\log \sigma_{i,t} - \mu_i) + \eta_{i,t}, \quad i = 1, 2, \ldots, M, \quad (4.15) \\
\begin{pmatrix} z_{i,t} \\ \eta_{i,t} \end{pmatrix} &\sim \text{i.i.d. } \mathcal{N}(0, \Omega_i), \quad \Omega_i = \begin{bmatrix} 1 & \rho_i \omega_i \\ \rho_i \omega_i & \omega_i^2 \end{bmatrix} \quad (4.16)
\end{align*}
\]

Furthermore, as described in the next subsection, we extend the DSGE model embedding the two financial frictions to examine the sources of the Great Recession: We construct a DSGE model with corporate and banking sectors’ balance sheet conditions due to agency problems. In sum, the
characteristics of our model is (1) introducing the SV model with leverage effect into the DSGE model, (2) utilizing the large panel data to increase in the estimation accuracy of structural shocks and (3) incorporating two financial frictions into the standard DSGE model. Our model can tackle the following questions: Was the “bad luck” (increase in shocks’ volatilities) happening the Great Recession? What is the main source (structural shock) of the Great Recession? Which sector’s balance sheet deterioration (corporate or banking sector) triggered the Great Recession?

State Space Model to be Estimated

We face a difficulty to directly estimate the state space model (4.11), (4.12) and (4.13) utilizing a large panel data, since the size of matrix in transition equations (4.12) and (4.13) is equal to the total number of model variables \( S_t \) and measurement errors \( e_t \). This framework induces a drastically increase of the matrix as the number of data \( X_t \) is increasing. To avoid this situation, we transform to small size for the transition equations by eliminating AR process of measurement errors of (4.10) and expressing from only \( \nu_t \) with i.i.d. process for measurement errors. Substituting (4.13) into (4.11), measurement equation can be transformed to:

\[
(I - \Psi L) X_t = (I - \Psi L) A S_t + \nu_t, \quad \nu_t \sim i.i.d. N(0, R).
\]

where \( L \) is lag operator. By using notations \( \tilde{X}_t = X_t - \Psi X_{t-1} \), and \( \tilde{S}_t = [S_t' S_{t-1}']' \), this equation can be rewritten as:

\[
\tilde{X}_t = \begin{bmatrix}
\Lambda & -\Psi A \\
A & \Lambda
\end{bmatrix}
\begin{bmatrix}
S_t \\
S_{t-1}
\end{bmatrix}
+ \nu_t, \quad \nu_t \sim i.i.d. N(0, R).
\]

In the similar way, transition equation (4.12) is also rewritten as:

\[
\begin{bmatrix}
S_t \\
S_{t-1}
\end{bmatrix}
= \begin{bmatrix}
G(\theta) & 0 \\
I & O
\end{bmatrix}
\begin{bmatrix}
S_{t-1} \\
S_{t-2}
\end{bmatrix}
+ \begin{bmatrix}
E(\theta) & 0 \\
O & E
\end{bmatrix}
\epsilon_t,
\]

where \( I \) is a \( N \times N \) identity matrix. Estimation method of the data rich DSGE model is explained using the state space model (4.17) and (4.18), and we estimate this model using Bayesian Method via MCMC. For convenience, we set parameters of measurement equation (4.17) as \( \Gamma = \{ \Lambda, \Psi, R \} \). And Bayesian estimation of parameters \( \Gamma \) are following Chib and Greenberg (1994) described in Appendix.

4.3 Model

We extend the SW model by embedding the two financial frictions of corporate and banking sectors based on Bernanke et al. (1999, hereinafter, BGG) and Gertler and Kiyotaki (2010, hereinafter, GK). This section is the highlight modeling the frictions in our model. The rest of the model is described in Appendix.
4.3.1 Financial Friction in Corporate Sector

Enterance and Exit of Entrepreneurs

Following BGG (1999), there is a continuum of entrepreneurs indexed by \( j \in [0, 1] \) where each entrepreneur is risk-neutral and has a finite expected horizon.\(^4\) Each entrepreneur faces an exogenous time-varying stochastic survival rate of \( \gamma_{t+1}^E \) from period \( t \) to \( t + 1 \) which is common across all entrepreneurs.\(^5\)

Between period \( t \) and \( t + 1 \), after \( 1 - \gamma_{t+1}^E \) fraction of entrepreneurs have exited from the business, exactly the same amount of new entrepreneurs will enter the business so that the population of entrepreneurs in the economy remains the same (i.e., fraction \( f^E \) of the total members of the household) from period \( t \) to \( t + 1 \). Each entering entrepreneur receives a ‘start-up’ transfer from the household and the total ‘start-up’ transfer from the household will be equal to the constant fraction \( \xi^E \) of aggregate net worth available in the corporate sector, \( n_t^E \), i.e., \( \xi^E n_t^E \). For \( 1 - \gamma_{t+1}^E \) fraction of entrepreneurs who happened to exit the business, they will first sell off the capital they purchased last period and retire all of their debts before maturity. And then, they will transfer their remaining net worth back to the household. The total amount of transfers from exiting entrepreneurs to the household will be \( (1 - \gamma_{t+1}^E) n_t^E \). Accordingly, net transfer, \( \Xi_{t+1}^E \), that the household receives from entrepreneurs at period \( t + 1 \) is \( (1 - \gamma_{t+1}^E - \xi^E) n_t^E \).

Individual Entrepreneur’s Problem

Each entrepreneur produces homogenous intermediate goods, \( y_t(j) \), and they are perfectly competitive when selling their products to retailers. The production function for the intermediate goods is given by

\[
y_t(j) = \omega_t(j) A_t k_t(j)^\alpha l_t(j)^{1-\alpha}, \tag{4.19}
\]

where \( k_t(j) \) is capital inputs and \( l_t(j) \) is labor inputs. The total factor productivity shock (hereafter, TFP shock), \( A_t \), is common across all entrepreneurs. However, following Carlstrom and Fuerst (1997) and BGG (1999), we assume each entrepreneur is subject to an idiosyncratic shock, \( \omega_t(j) \), which is a private information to entrepreneur \( j \) and assumed to be i.i.d. shock with mean equal to one, i.e., \( E[\omega_t(j)] = 1 \).

The balance sheet statement of each entrepreneur at the end of period \( t \) can be expressed as

\[
q_t k_{t+1}(j) = b_t^E(j) + n_t^E(j) \tag{4.20}
\]

where \( q_t \) is the real price of capital, \( k_{t+1}(j) \) is the capital which will be used for production in period \( t + 1 \) but purchased at the end of period \( t \), \( b_t^E(j) \) is the real debt issued at period \( t \) and \( n_t^E(j) \) is the net worth at period \( t \). With the assumption of risk-neutrality and finite planning horizon, net worth itself is never enough in financing the cost of capital purchase and, therefore, each entrepreneur will rely on external financing in equilibrium.

The income statement for entrepreneur \( j \) is specified as follow

\[
n_t^E(j) = p_t^mc(j)y_t(j) - w_t l_t(j) - \frac{R_{t-1}^E(j)}{\pi_t} b_t^E(j) - q_t (1 - \delta) k_t(j) \tag{4.21}
\]
where \( p^{mc}_t(j) \) is the real price of intermediate goods \( j \), \( R^{E-1}_t(j)/\pi_t \) is the real rate of borrowing cost \( (R^{E-1}_t(j) \) is nominal borrowing rate and \( \pi_t \) is inflation rate) and \( \delta \) is capital depreciation rate.

Each entrepreneur entering period \( t \) maximizes her discounted cash flow by choosing capital inputs, labor inputs and debt issuance subject to (4.19), (4.20), and (4.21). The FOCs for each entrepreneur \( j \) are given by

\[
\begin{align*}
E_t \left[ \gamma_{t+1} \frac{R^E_t(j)}{\pi_{t+1}} \right] &= E_t \left[ \gamma_{t+1} \left( \frac{\alpha p^{mc}_{t+1}(j)y_{t+1}(j)}{k_{t+1}(j)} + \frac{(1 - \delta)q_t + 1}{q_t} \right) \right].
\end{align*}
\]

(4.22) equates marginal cost of labor to marginal product of labor and, thus, can be thought of as labor demand function by entrepreneur \( j \). (4.23) equates the expected marginal cost of capital financed by debt to the expected marginal return of capital financed by debt and can be thought of as capital demand function by entrepreneur \( j \). Since stochastic survival rate, \( \gamma_{t+1} \), is uncorrelated to any other shocks in the economy, (4.23) can be further rearranged as

\[
\begin{align*}
E_t \left[ \frac{R^E_t(j)}{\pi_{t+1}} \right] &= E_t \left[ \frac{\alpha p^{mc}_{t+1}(j)y_{t+1}(j)}{k_{t+1}(j)} + \frac{(1 - \delta)q_t + 1}{q_t} \right].
\end{align*}
\]

Under the assumption of risk-neutrality, introduction of stochastic survival rate will not alter the capital demand equation for any entrepreneur \( j \) compared to the case with constant survival rate as in BGG (1999).

### Debt Contract

Each period, entrepreneur \( j \) issues a debt and engages in a debt contract with an arbitrary chosen financial intermediary \( m \) where \( m \) is an indexed number uniformly distributed from 0 to 1. Debt contract is for one period only and if entrepreneur \( j \) needs to issue a debt again next period, another arbitrary financial intermediary \( m' \) will be chosen next period. Following BGG (1999), idiosyncratic TFP shock, \( \omega_t(j) \), is private information of entrepreneur \( j \) that there exists asymmetric information between entrepreneur \( j \) and financial intermediary \( m \). Due to costly state verification, financial intermediary \( m \) cannot observe entrepreneur \( j \)'s output costlessly, but need to incur a monitoring cost to observe it. Entrepreneur \( j \), after observing the project outcome, will decide whether to repay the debt or default at the beginning of period \( t \). If the entrepreneur decides to repay, financial intermediary will receive repayment of \( R^{E-1}_t(j)/\pi_t \) for each unit of credits outstanding, regardless of the realization of idiosyncratic shock. Otherwise, the financial intermediary will pay a monitoring cost to observe \( y_t(j) \) and seize the project outcome from the entrepreneur.

Under the optimal debt contract, BGG (1999) shows that the external finance premium, \( s_t(j) \), to be an increasing function of the leverage ratio. For estimation purpose, we follow Christensen

\footnote{Each entrepreneur is a price-taker in the labor market, financial market, and capital market. At the beginning of period \( t \), each entrepreneur will utilize capital, \( k_t(j) \), and labor input, \( l_t(j) \), to produce the intermediate goods, \( y_t(j) \). Then, they will sell off the intermediate goods to retailers in a perfectly competitive manner and earn the revenue, \( p^{mc}_t(j)y_t(j) \). After earning the revenue, each entrepreneur will pay the labor cost and also repay the debt. Finally, each entrepreneur will sell off a depreciated capital at the capital market. The net income after these activities are captured by \( n^E_t \) and will be a net worth for the entrepreneur \( j \) at the end of period \( t \). Given this net worth, each entrepreneur will plan for the next period and decide how much capital to purchase and how much debt to issue at the end of period \( t \) which appears in the balance sheet equation (4.20).}
and Dib’s (2008) specification of the external finance premium as follow,

\[ s_t(j) = \left( \frac{q_t k_{t+1}(j)}{n_t^E(j)} \right)^\varphi \]  

(4.25)

where parameter \( \varphi > 0 \) can be interpreted as the elasticity of external finance premium with respect to the leverage ratio. In addition, discounting the external finance premium from the borrowing rate \( R_t^E(j) \), the expected risk-adjusted nominal return for financial intermediary \( m \) from the debt contract from period \( t \) to \( t + 1 \) can be expressed as

\[ E_t R_{t+1}^E(m) = \frac{R_t^E(j)}{s_t(j)}. \]  

(4.26)

Aggregation

Since bankruptcy cost is constant-return-to-scale and leverage ratio are equal for all entrepreneur \( j \), the external finance premium is equal across all solvent entrepreneurs in equilibrium, i.e., \( s_t = s_t(j) \) for all \( j \). Since (4.24) holds in aggregate level, the nominal borrowing rates across all solvent entrepreneurs become equal, i.e., \( R_t^E = R_t^E(j) \) for all \( j \). Consequently, because \( R_t^E = R_t^E(j) \) and \( s_t = s_t(j) \) for all \( j \), the expected risk-adjusted nominal return for banker \( m \) becomes equal across all bankers, i.e.,

\[ E_t \left[ R_{t+1}^E(m) \right] = \frac{R_t^E}{s_t} \text{ for all } m. \]  

(4.27)

Next, we derive the law of motion of the aggregate net worth of corporate sector. As for notation, aggregate variable is expressed by suppressing the argument \( j \). Aggregating over income statement (4.21) and taking into account the entrance and exit of entrepreneurs from period \( t \) to \( t + 1 \), we obtain the following aggregate net worth transition equation.

\[ n_{t+1}^E = \gamma_{t+1}^E \left[ \left( r_{t+1}^k - \frac{R_t^E}{\pi_{t+1}} b_t^E \right) + R_t^E n_t^E \right] + \xi^E_{t+1} n_t^E \]  

(4.28)

where \( r_{t+1}^k \) is realized gross return from capital investment at period \( t + 1 \) and is defined as

\[ r_{t+1}^k = \frac{\alpha p_{t+1}^{mc} \bar{y}_{t+1} / k_{t+1} + (1 - \delta) q_{t+1}}{q_t}. \]  

(4.29)

Here, \( \bar{y}_{t+1} \) is the average of project outcomes, \( y_{t+1}(j) \), across all entrepreneurs. Thus, idiosyncratic factor stemming from \( \omega_t(j) \) is averaged away and \( r_{t+1}^k \) only reflects the aggregate factors in the economy. Using entrepreneur’s balance sheet (4.20), the aggregate net worth transition (4.28) can be rearranged as

\[ n_{t+1}^E = \gamma_{t+1}^E \left[ \left( \frac{r_{t+1}^k - \frac{R_t^E}{\pi_{t+1}}}{q_t} \right) q_t k_{t+1} + R_t^E n_t^E \right] + \xi^E_{t+1} n_t^E. \]  

(4.30)

Notice how the realization of \( r_{t+1}^k \) can affect the aggregate net worth next period. Ex-ante, by the rational expectation equilibrium condition (4.24), the expected return from capital investment and borrowing cost are equalized. Ex-post, however, realized return from capital investment can exceed or fall below the borrowing cost depending on the realizations of the aggregate shocks and it affects
the evolution of the aggregate net worth. This is a case where forecast error has an actual effect on the economy. Another factor that affects the evolution of the aggregate net worth is the realization of stochastic survival rate \( \gamma_{t+1}^E \). At the micro-level, \( \gamma_{t+1}^E \) has an interpretation of stochastic survival rate of entrepreneur \( j \) from period \( t \) to \( t + 1 \). At the aggregate level, \( \gamma_{t+1}^E \) is interpreted as an exogenous shock to the aggregate net worth in corporate sector. In our paper, we interpret it as an aggregate corporate net worth shock.

4.3.2 Financial Friction in Banking Sector

Entrance and Exit of Bankers

Following Gertler and Karadi (2011) as well as GK (2010), there is a continuum of bankers indexed by \( m \in [0, 1] \) where each banker is risk-neutral and has a finite horizon. We assume that each banker faces exogenous time-varying stochastic survival rate of \( \gamma_{t+1}^E \) from period \( t \) to \( t + 1 \) which is common to all bankers. By the same token as in corporate sector, the stochastic process of \( \gamma_t^F \) is uncorrelated with any other shocks in the economy and has it mean equal to \( \gamma^F \), i.e., \( E[\gamma_t^F] = \gamma^F \).

After \( 1 - \gamma_{t+1}^E \) fraction of bankers have exit between period \( t \) and \( t + 1 \), exactly the same number of new bankers will enter the banking business from the household. Each banker entering the baking business will receive a 'start-up' transfer from the household, while each banker exiting the business will transfer his net worth back to the household. In aggregate, 'start up' transfer is assumed to be the constant fraction \( \xi^F \) of aggregate net worth available in the banking sector, \( n_t^F \), i.e., \( \xi^F n_t^F \) and the aggregate transfer from the exiting bankers is equal to \( \gamma_{t+1}^E n_t^F \). Thus, net transfer from the banking sector to the household, \( \Xi_t^F \), is equal to \( (1 - \gamma_{t+1}^E - \xi^F) n_t^F \).

Individual Banker’s Problem

We now describe the individual banker’s problem. The treatment here basically follows that of Gertler and Karadi (2011) and perfect inter-bank market version of GK (2010). The balance sheet equation of the individual banker \( m \) is given by

\[
b_t^F(m) = n_t^F(m) + b_t^F(m)
\]

where \( b_t^F(m) \) is the asset of banker \( m \) which is lent out to an arbitrarily chosen entrepreneur \( j \) at period \( t \), \( n_t^F(m) \) is the net worth of banker \( m \), and \( b_t^F(m) \) is the liability of banker \( m \) which is also a deposit made by the household at period \( t \).

By receiving deposits \( b_t^F(m) \) from household at period \( t \), banker \( m \) pledges to pay the deposit rate of \( R_t / \pi_{t+1} \) in real terms next period. As a result of the banking business, the net worth transition for banker \( m \) at period \( t + 1 \) is given by

\[
n_{t+1}^F(m) = r_{t+1}^F(m) b_t^F(m) - r_{t+1} b_{t+1}^F(m)
\]

where \( r_{t+1}^F(m) \equiv R_{t+1}(m) / \pi_{t+1} \) and \( r_{t+1} \equiv R_t / \pi_{t+1} \). Using the balance sheet equation (4.31), the net worth transition equation can be reformulated as follow

\[
n_{t+1}^F(m) = (r_{t+1}^F(m) - r_{t+1}) b_t^F(m) + r_{t+1} n_t^F(m).
\]

As shown by GK (2010), with the agency cost present between banker \( m \) and depositor, the expected spread between \( r_{t+1}^F(m) \) and real deposit rate \( r_{t+1} \) becomes strictly positive, i.e., \( E_t [r_{t+1}^F(m) - r_{t+1}] > 0 \). However, of course, whether the net worth of banker \( m \) increases or decreases next period depends on the realization of \( r_{t+1}^F(m) \).

Given the above net worth transition equation, risk-neutral banker \( m \) will maximize the net worth accumulation by maximizing the following objective function with respect to bank lending,
$b_t^E(m)$,

\[ V_t^F(m) = E_t \sum_{i=0}^{\infty} \beta^i (1 - \gamma_{t+1}^F) \gamma_{t+1,t+1+i}^F \left[ (r_{t+1+i}^F(m) - r_{t+1+i}) b_{t+i}^E(m) + r_{t+1+i} n_{t+i}^F(m) \right] \quad (4.33) \]

where $\gamma_{t+1,t+1+i}^F \equiv \Pi_{j=0}^i \gamma_{t+1+j}^F$. Now, since the expected spread between risk-adjusted bank lending rate and deposit rate is strictly positive, it is in the interest on banker $m$ to lend out infinite amount to an entrepreneur by accepting infinite amount of deposits from the depositor.

In order to avoid the infinite risk-taking by the banker, Gertler and Karadi (2011) and GK (2010) impose a moral hazard/costly enforcement problem between the banker and depositor. Each period, the banker has a technology to divert fraction $\lambda$ of his asset holding to the household and exit from the banking business. However, by doing so, the banker is forced to file bankruptcy and fraction $(1 - \lambda)$ of his asset will be seized by the depositors. Thus, in order for the banker to continue business and depositors to safely deposit their funds to the banker, the following incentive constraint must be met each period,

\[ V_t^F(m) \geq \lambda b_t^E(m). \quad (4.34) \]

In other words, the net present value of the banking business needs to always exceed the reservation value retained by the banker.\(^7\)

Now, assuming that the incentive constraint (4.34) to be binding each period and by maximizing the objective function (4.33) subject to the constraint (4.34), Gertler and Kiyotaki (2010) shows that the value function of the banker can be expressed as follow

\[ V_t^F(m) = \nu_t b_t^E(m) + \eta_t n_t^F(m) \quad (4.35) \]

where

\[ \nu_t \equiv E_t \left[ (1 - \gamma_{t+1}^F) \beta (r_{t+1}^F(m) - r_{t+1}) + \beta \gamma_{t+1}^F \frac{b_{t+1}^E(m)}{b_t^E(m)} \nu_{t+1} \right] \quad (4.36) \]

\[ \eta_t \equiv E_t \left[ (1 - \gamma_{t+1}^F) + \beta \gamma_{t+1}^F \frac{n_{t+1}^F(m)}{n_t^F(m)} \eta_{t+1} \right]. \quad (4.37) \]

Now, from incentive constraint (4.34) and the value function (4.35), it follows that

\[ \frac{b_t^E(m)}{n_t^F(m)} \leq \frac{\eta_t}{\lambda - \nu_t} \equiv \phi_t \quad (4.38) \]

which states that the leverage ratio of banker $m$ cannot exceed the (time-varying) threshold $\phi_t$. By the assumption that incentive constraint to bind every period, in equilibrium, the asset and the net worth by banker $m$ have a following relationship

\[ b_t^E(m) = \phi_t n_t^F(m). \quad (4.39) \]

\(^7\)To see how this constraint binds, consider the case where the banker increases the asset enormously. Then, the reservation value by the banker (right-hand side of inequality (4.34)) will exceed the net present value of the banking business (left-hand side of inequality (4.34)) that the banker will decide to divert the assets to the household. As a stakeholder, the depositors will not allow this reckless behavior by the banker and ask the banker to keep his asset, $b_t^E(m)$, low enough (or, equivalently, by not supplying the deposits beyond the incentive constraint) so that the incentive for the banker to remain in business is met.
4.3. MODEL

Aggregation

Gertler and Karadi (2011) and GK (2010) show that time-varying threshold $\phi_t$ does not depend on banker-specific factors and is common across all bankers. Consequently, from eq. (4.39), aggregate asset and net worth in banking sector can be expressed as

$$b^E_t = \phi_t n^F_t$$  \hspace{1cm} (4.40)

where $b^E_t \equiv \int_0^1 b^E_t(m) dm$ and $n^F_t = \int_0^1 n^F_t(m) dm$. Now, from individual banker’s net worth transition (4.32) and taking into account entrance and exit of bankers, the aggregate net worth transition equation of banking sector is given by

$$n^F_{t+1} = \gamma^F_{t+1} \left[ (\tau^F_{t+1} - r_{t+1}) b^E_t + r_{t+1} n^F_t \right] + \xi^F_t n^F_t$$  \hspace{1cm} (4.41)

where $\tau^F_{t+1}$ stands for the average of realized risk-adjusted returns, $r^F_{t+1}(m)$, across all bankers. From the optimal debt contract specified in (4.26) and using the aggregate condition in (4.27), $\tau^F_{t+1}$ is related to the borrowing rate, external finance premium, and inflation rate as follow

$$\tau^F_{t+1} = \frac{R^E_t}{\pi_{t+1} s_t}.$$  \hspace{1cm} (4.42)

As can be seen from the above equation, idiosyncratic factor pertaining to banker $m$ is averaged away and, thus, realization of risk-adjusted return of banking sector (i.e., $\tau^F_{t+1}$) only depends on aggregate factors in the economy. Now, by using (4.40), the aggregate net worth transition equation becomes

$$n^F_{t+1} = \gamma^F_{t+1} \left[ (\tau^F_{t+1} - r_{t+1}) \phi_t + r_{t+1} \right] n^F_t + \xi^F_t n^F_t$$  \hspace{1cm} (4.43)

4.3.3 Incorporating Two Financial Frictions into the DSGE model

To incorporate the two financial frictions into a stylized DSGE model, we use twelve constraint and FOC equations, which consist of five and seven equations derived in corporate and banking sectors, respectively. The five equations representing the financial friction in the corporate sector are (i) the balance sheet statement of corporate sector (4.20), (ii) the capital demand function (4.24), (iii) the external financial premium (4.25), (iv) the realized gross return from capital investment (4.29), and (v) the aggregate net worth transition equation of corporate sector (4.30). On the other hand, the seven equations expressing the financial friction in the banking sector are (vi) the balance sheet statement of banking sector (4.31), (vii) the dynamics of the weight on the lending volume for the value of the banking business, $\nu_t$, (4.36), (viii) the dynamics of the weight on the bank net worth for the value of the banking business, $\eta_t$, (4.37), (ix) the definition of the threshold, $\phi_t$, (4.38), (x) the banker’s leverage ratio constraint (4.40), (xi) the relationship between the corporate nominal borrowing rate and the risk adjusted nominal lending rate of the banking sector (4.42), and (xii) the aggregate net worth transition equation of the banking sector (4.43).

---

8We have twelve model (or endogenous) variables corresponding to the twelve estimated equations pertain to the financial frictions. These variables are (1) capital, $k_t$, (2) the real price of capital, $q_t$, (3) asset of the corporate sector, $b^E_t$, (4) asset of the banking sector $b^F_t$, (5) the corporate net worth $n^E_t$, (6) the bank net worth $n^F_t$, (7) external financial premium, $s_t$, (8) the gross return from capital investment, $r^E$, (9) time varying weight of lending for the value of banking business, $\nu_t$, (10) time varying weight of bank net worth for the value of banking business, $\eta_t$, (11) the corporate nominal borrowing rate, $R^E_t$, and (12) the risk-adjusted lending rate of banking sector, $R^F_t$. 
To complete our model, we employ the CEE (2005) type medium scale DSGE model described in Appendix with the equations above, as well as structural shocks. We set the following eight structural shocks, each of them having a specific economic interpretation; i.e., (1) TFP shock, (2) preference shock, (3) labor supply shock, (4) investment specific technology shock, (5) government spending shock, (6) monetary policy shock, (7) corporate net worth shock and (8) bank net worth shock. Except monetary policy shock, all of the structural shocks are assumed to follow AR(1) stochastic processes. We devote the following two shocks out of the eight shocks to identifying fundamental factors causing the financial crisis. Corporate net worth shock $\varepsilon_t^E$ is inserted into AR(1) process of the survival rate of the corporate sector $\gamma_t^E$ which is a component of equation (4.30), while bank net worth shock $\varepsilon_t^F$ is done into AR(1) process of the survival rate of the banking sector $\gamma_t^F$ which is that of equation (4.43). The two shocks are given as

Corporate net worth shock : $\gamma_t^E = \rho \gamma_{t-1}^E + \varepsilon_t^E,$

Bank net worth shock : $\gamma_t^F = \rho \gamma_{t-1}^F + \varepsilon_t^F,$

where $\rho$ is for the AR(1) coefficients for respective structural shocks. Both shocks indicating stochastic survival rate for entrepreneurs and bankers at micro-level can be interpreted as net worth shocks for corporate and banking sectors at aggregate level, respectively. Notice that each stochastic disturbance $\varepsilon_t$ is assumed to follow time varying volatility using SV model.

4.4 Estimation Method

This study employs a hybrid MCMC (also referred to as Metropolis-within-Gibbs) as our estimation method. Boivin and Giannoni (2006) and Kryshko (2011) estimated the standard DSGE model in a data rich environment by adopting the hybrid MCMC. We further extend the method by (1) incorporating the stochastic volatilities with leverage effect into the shock generating process, (2) adopting the simulation smoother in estimating structural shocks, and (3) introducing the measurement errors with AR(1) process. When detecting the sources of the Great Recession, it becomes important how accurately structural shocks was estimated. It is worthwhile to adopt the hybrid MCMC, especially since the method is possible not only to estimate the posterior of model variables $S_t$ but also to estimate the posterior of structural shocks $\varepsilon_t$, i.e. the method can provide the credible intervals of the structural shocks including corporate and banking sectors’ net worth shocks.

Our estimation targets in the state space model (4.17), and (4.18) are structural parameters $\theta$, parameters in measurement equation $\Gamma (= \{A, \Psi, R\})$, model variables $S^T (= S_1, S_2, \cdots, S_T)$, and stochastic volatilities $H^T (= h_1, h_2, \cdots, h_T)$. For convenience, let log $\sigma_t$ denote as $h_t$, hereinafter. It should be noted that we have only to estimate $\theta, \Gamma, S^T, H^T$ since matrices $G(\theta), E(\theta), Q(\theta)$ is the function of structural parameters $\theta$.

The procedure of implementing the hybrid MCMC to estimate $\theta, \Gamma, H^T$ as the following steps:

**Step I:** Set the priors $\theta, \Gamma, H^T$, i.e. $p(\theta, \Gamma, H^T)$ where $p(\theta, \Gamma, H^T) = p(\theta | \Gamma, H^T) p(\Gamma | H^T) p(H^T)$, since $\theta, \Gamma, H^T$ are assumed to be independent.

**Step II:** Using Bayes’ theorem, evaluate the posterior $p(\theta, \Gamma, H^T | X^T)$ from the prior $p(\theta, \Gamma, H^T)$ and the likelihood $p(X^T | \theta, \Gamma, H^T)$.

$$p(\theta, \Gamma, H^T | X^T) = \frac{p(X^T | \theta, \Gamma, H^T) p(\theta, \Gamma, H^T)}{\int p(X^T | \theta, \Gamma, H^T) p(\theta, \Gamma, H^T) d\theta d\Gamma dH^T}.$$
4.4. ESTIMATION METHOD

Step III: Calculate moments (mean, median, credible band, etc.) from the estimated posterior $\theta, \Gamma, H^T$ from posterior $p(\theta, \Gamma, H^T | X^T)$ using numerical technique.

However, we cannot directly draw parameters from the joint posterior distribution $p(\theta, \Gamma, H^T | X^T)$ of the state space model (4.17), (4.18) in Step II. Thus, using the Gibbs sampler, we obtain the joint posterior $p(\theta, \Gamma, H^T | X^T)$ from the conditional posterior $p(\theta, \Gamma, X^T)$ as below,

$$
p(\theta | \Gamma, H^T, X^T), \quad p(\Gamma | \theta, H^T, X^T), \quad p(H^T | \theta, \Gamma, X^T)
$$

In addition, since parameter $\Gamma$ is dependent on model variables $S_t$, we have to separate two conditional posterior $p(S^T \mid \Gamma, \theta, H^T, X^T)$ and $p(\Gamma \mid S^T, \theta, H^T, X^T)$ from the above conditional posterior $p(\Gamma \mid \theta, H^T, X^T)$ and substitute $S_t$ into the posterior. We also adopt a forward-backward recursion for sampling from $p(S^T \mid \Gamma, \theta, H^T, X^T)$ and $p(H^T \mid \Gamma, \theta, S^T, X^T)$ as a data augmentation method, Gibbs sampling for sampling from $p(\Gamma \mid S^T, \theta, H^T, X^T)$, and MH algorithm for sampling from $p(\theta \mid \Gamma, H^T, X^T)$, respectively. In this way, different algorithms are employed for different parameters in the hybrid MCMC. In sum, we show six steps of hybrid MCMC for estimating the data rich DSGE model as follow.  

Step 1: Specify initial parameters values $\theta^{(0)}$, $\Gamma^{(0)}$, and $H^{T(0)}$. Set iteration index $g = 1$.

Step 2: Solve the DSGE model numerically at $\theta^{(g-1)}$ and obtain matrices $G(\theta^{(g-1)})$, $E(\theta^{(g-1)})$ and $Q(\theta^{(g-1)})$.

Step 3: Draw $\Gamma^{(g)}$ from $p(\Gamma \mid \theta^{(g-1)}, H^{T(g-1)}, X^T)$.

(3.1) Generate model variables $S_t^{(g)}$ and structural shocks $\varepsilon_t^{(g)}$ from $p(S^T, \varepsilon_T \mid \Gamma^{(g-1)}, \theta^{(g-1)}, H^{T(g-1)}, X^T)$ using the simulation smoother by de Jong and Shephard (1995).

(3.2) Generate parameters $\Gamma^{(g)}$ from $p(\Gamma \mid S^T(\cdot), \theta^{(g-1)}, H^{T(g-1)}, X^T)$ based on the sampled draw $S^T(\cdot)$ using the Gibbs sampler by Chib and Greenberg (1994).

Step 4: Draw $H^{T(g-1)}$ from $p(H^T \mid \theta^{(g-1)}, \Gamma^{(g)}, \varepsilon_T^{(g)}, X^T)$.

(4.1) Generate stochastic volatility $H^{T(g)}$ from $p(H^T \mid \Gamma^{(g)}, \theta^{(g-1)}, \varepsilon_T^{(g)}, u^{T(g-1)}, \Phi^{(g-1)}, X^T)$, using a draw of $\varepsilon_T^{(g)}$ at Step 3.1, and the forward-backward recursion by Cater and Kohn (1994).

(4.2) Generate the indicators of the mixture approximation $u^{T(g)}$ using discrete density proposed by Omori et al. (2007).

(4.3) Generate the coefficients $\Phi^{(g)}$ of stochastic volatility process using Metropolis step.

Step 5: Draw parameters $\theta^{(g)}$ from $p(\theta \mid \Gamma^{(g)}, H^{T(g)}, X^T)$ using Metropolis step:

(5.1) Draw the candidate from proposal density $p(\theta | \theta^{(g-1)})$ and, using the sampled draw $\theta^{(\text{proposal})}$, calculate the acceptance probability $q$ as follows.

$$
q = \min \left[ \frac{p(\theta^{(\text{proposal})} \mid \Gamma^{(g)}, H^{T(g)}, X^T)}{p(\theta^{(g-1)} \mid \Gamma^{(g)}, H^{T(g)}, X^T)} \frac{p(\theta^{(g-1)} | \theta^{(\text{proposal})})}{p(\theta^{(\text{proposal})} | \theta^{(g-1)})}, 1 \right].
$$

(5.2) Accept $\theta^{(\text{proposal})}$ with probability $q$ and reject it with probability $1 - q$. Set $\theta^{(g)} = \theta^{(\text{proposal})}$.

---

9Bayesian estimation method using MCMC for the state space model are described in detail in Kim and Nelson (1999) and Bauwens et al. (1999).
when accepted and $\theta^{(g)} = \theta^{(g-1)}$ when rejected.

**Step 6:** Set iteration index $g = g + 1$ and return to Step 2 up to $g = G$.

The detailed algorithm on sampling stochastic volatilities $H_T$ in Step 4, the simulation smoother in Step 3.1, and drawing parameters $\Gamma$ in Step 3.2 are described in Appendix.\(^\text{10}\)

### 4.5 Preliminary Settings and Data

#### 4.5.1 Four Cases

This study considers four alternative cases corresponding to (1) the number of observation variables (11 vs. 40 observable variables) and (2) specification of volatilities (time-constant vs. time-varying volatility) as summarized in Table 4.1. Thus, the four cases consist of (1) utilizing a lot of data or not, and (2) volatilities are time varying or not.

The first case (referred to as Case A) is the standard estimation method with 11 data and i.i.d. normal structural shocks. The second case (Case B) is extended to the data rich approach with 40 data and i.i.d shocks. The third case (Case C) extends to the SV shocks from Case A (11 data with SV shocks). The forth case (Case D) extends to the data rich approach with SV shocks from Case B (40 data with SV shocks).

#### 4.5.2 Calibration and Prior Settings

Some parameters are not identifiable, so calibrated as summarized in Table 4.2. The subjective discount factor $\beta = 0.995$ (annual rate 2%). The profit margin of the retailers is set to be 10% in steady state, which implies the gross markup $\epsilon = 11$. We have no reliable information on the new entry rate of entrepreneurs (i.e., $\xi^E$). So we refer to the calibrated value of GK (2011). The rest of the calibrated values are borrowed from SW (2003), Christensen and Dib (2008), and GK (2011).

Regarding the steady states, most of them are pinned down by equilibrium conditions of the model, but some others need to be calibrated. For the steady state value of external finance premium, we follow the calibrated value of Christensen and Dib (2008). For the steady state corporate borrowing rate, we employ sample mean of Moody’s Baa-rated corporate bonds yields. Similarly, we use sample mean on the steady state of corporate leverage ratio.

---

\(^{10}\)Here, we explain Steps 1 and 5. On setting of initial values $\theta^{(0)}$ and $\Gamma^{(0)}$ in Step 1, we first estimate the posterior mode of structural parameters $\theta$ in a regular DSGE model without measurement errors. Then, we set as initial value $\theta^{(0)}$. Second, implementing the simulation smoother of state variables $S_t$ using $\theta^{(0)}$, we obtain the initial value $S_t^{(0)}$. Finally, initial values $\Gamma^{(0)}$ of measurement equations are obtained by OLS using $S_t^{(0)}$ and $X_T^T$. Next, on estimating structural parameters $\theta$ from proposal density in Step 5.1, we adopt the random walk MH algorithm. Proposal density $\theta^{\text{proposal}}$ is generated as

$$\theta^{\text{proposal}} = \theta^{(g-1)} + u_t, \quad u_t \sim \mathcal{N}(0, c\Sigma),$$

where $\Sigma$ is variance covariance matrix of the random walk process, and $c$ is the adjustment coefficient. The matrix $\Sigma$ is the Hessian $-\nabla^2 l(\hat{\theta})$ of log posterior distribution $l(\theta) = \ln p(\theta|\Gamma, X_T^T)$ when obtaining initial value $\theta^{(0)}$. The acceptance rate $q$ can be calculated as follows.

$$q = \min \left[ \frac{f(\theta^{\text{proposal}})}{f(\theta^{(g-1)})}, 1 \right],$$

We adjust the coefficient $c$ so that the acceptance rate is close to around 25%.
The prior settings are reported in Table 4.3. Following the calibration by BGG, we set \( \varphi = 0.05 \) which governs the external finance premium on the corporate sector. AR(1) shock persistence parameters are set to 0.5 for all and standard error to 1%, except for monetary policy shock (since a change of policy rate for more than 25 basis point is rare). By the same token, the prior mean is set to 1% for most of the measurement errors, except for the data related to interest rates.

4.5.3 Data

The estimation period is from 1985:Q2 to 2012:Q2, since we exclude the periods when the monetary policy regime might change (especially around the end of the 1970’s and early 1980’s; i.e., pre and post regimes by Volcker and Greenspan, see Clarida et al. 2000, Lubik and Schorfheide 2004, and Boivin 2005) and structural changes may occur (at the Great Moderation which began in mid-1980’s, see Bernanke 2004, Stock and Watson 2002, Kim and Nelson 1999, and McConnell and Perez-Quiros 2000). Another reason is data availability: The charge-off rates for banks are available only from 1985Q1.

Cases A and C have 11 series: 
1. output, \( y_t \),
2. consumption, \( c_t \),
3. investment, \( i_k t \),
4. inflation, \( \pi_t \),
5. real wage, \( w_t \),
6. labor input, \( l_t \),
7. nominal interest rate, \( R_t \),
8. nominal corporate borrowing rate, \( R^E_t \),
9. external finance premium, \( s_t \),
10. corporate leverage ratio, \( q_t k_t / n^E_t \), and
11. bank leverage ratio, \( b^E_t / n^F_t \).

The first seven series are used in a large literature (see, e.g., SW 2003, 2007). The remaining four data are financial data: 
8. Entrepreneur’s nominal borrowing rate, \( R^E_t \), is the yield on Moody’s Baa-rated corporate bonds, detrended via Hodrick-Prescott filter.
9. the external financial premium, \( s_t \), we employ the charge-off rates for all banks credit and issuer loans, measured as an annualized percentage of uncollectible loans. The charge-off rate is demeaned to be consistent with our model variable. 
10 and 11. The two leverage ratios, \( q_t k_t / n^E_t \) and \( b^E_t / n^F_t \) are calculated as their total asset divided by their net worth, respectively. Taking natural log for both leverage ratios and then we demean for entrepreneur’s leverage ratio and detrend banking sector leverage ratio by Hodrick-Prescott filter, considering Basel Capital Accord Revision.

Cases B and D correspond to the data rich approach where we use 40 data in all. The additional 29 series consists of 18 macroeconomic data and 11 financial data. Following Boivin and Giannoni (2006), we select additional 18 macroeconomic data. The additional financial data is focused on bank sectors’ data (bank leverage) to identify the bank net worth shock. Three additional bank sectors’ data: (i) the core capital leverage ratio, (ii) the domestically chartered commercial banks’ leverage ratio and (iii) the leverage ratio of brokers and dealers. It should be noted that as the leverage ratio we use corresponds to the reciprocal of the commonly-used ratio, i.e., bank asset over bank equity. On the external financial premium, we collect charge-off rates on loans data

\[ \text{11(1) Output is real GDP less net export. (2) Consumption and (3) investment are normalized respectively to personal consumption expenditures and fixed private domestic investment. Following Altig et al. (2003), SW (2003), and Boivin and Giannoni (2006), the nominal series for consumption and investment are deflated with the GDP deflator. (6) The labor input corresponds to hours worked per person. Average hours of nonfarm business sector are multiplied with civilian employment to represent the limited coverage of the nonfarm business sector, compared to GDP, as in SW (2003), and Boivin and Giannoni (2006). (5) The real wage is normalized with the hourly compensation for the nonfarm business sector, divided by the GDP deflator. We express these six series as percent deviations from steady states consistently with model concepts, taking the natural logarithm, extracting the linear trend by an OLS regression, and multiplying the resulting detrended series by 100. (4) Inflation measures are obtained by taking the first difference of the natural logarithm of the GDP deflator, and multiplied by 400 for expressing the annualized percentages. (7) The nominal interest rate is the effective Federal funds rate. Both inflation and the interest rate are detrended via Hodrick-Prescott filter (penalty parameter is 1600), indicating time-varying targeting inflation rate.\]

\[ \text{12The core capital leverage ratio represents tier 1 (core) capital as a percent of average total assets. Tier 1 capital consists largely of equity. We use the reciprocal of the core capital leverage ratio. Taking natural log, we detrended} \]
from three different institutions, which are detrended in the same way as the above. The detail is described in Data Appendix.

4.6 Results

This section reports our estimation results and especially focuses on key structural parameters, SV shocks, and historical decompositions of four principal variables: (1) output, (2) investment, (3) bank leverage and (4) corporate borrowing rate. In particular, bank leverage and corporate borrowing rate play significant roles in the Great Recession. Then, we discuss and remark on the sources of the Great Recession. Our results are based on estimated posterior distributions from 300,000 draws using the hybrid MCMC algorithm.\textsuperscript{13}

4.6.1 Structural Parameters

The estimated parameters of Cases A and B are summarized in Table 4.9, and those of Cases C and D are in Table 4.10. The estimated parameters on the SV models are in Table 4.11. We focus on interpreting seven key structural parameters, i.e., parameters related to the financial friction, nominal rigidities and monetary policy rule. Table 4.4 collects the key parameters in four cases to compare with one another. The parenthesis in the table indicates the 90\% credible interval of the posterior distribution.

First of all, we consider two estimated parameters involved in the financial friction of the corporate sector; $\kappa$ and $\varphi$. $\kappa$ indicates the coefficient of the quadratic adjustment cost of investment (on the investment Euler equation, see Appendix). $\varphi$ is the elasticity of the external financial premium.

According to Table 4.4, the posterior mean of $\kappa$ in Case B (the data rich approach with time-constant volatility shocks) is around 0.88, whereas those in the rest cases are between 0.56 and 0.63. The high $\kappa$ in Case B means adjusting investment becomes more costful. Suppose that an adverse corporate net worth shock occurs. Then, capital price goes down. Even if corporate sector observes low capital price, he cannot immediately reduce investment due to high real rigidity (adjustment cost). This excess supply of investment will further lower capital price. Further decline of capital price raises corporate leverage ratio strongly, so corporate sector faces more severe borrowing constraint. As a result, investment has to be significantly lowered. Thus, higher $\kappa$ plays a role of promoting the shock amplification mechanism through changes in capital price.

Looking at the posterior mean of $\varphi$, a rise of 1\% increase in corporate leverage ratio raises the spread $s_t$ (external financial premium) by about 0.03\% in Cases A and B (time-constant volatility), or by about 0.04\% in Cases C and D (with SV shocks). Notice that the posterior mean, 0.04 in the cases with SV shocks (Cases C and D) exceed the upper bound of the 90\% credible interval of Cases A and B. If the corporate leverage increases, in the case of SV, a further rise in spread equivalent to an interest rate hike of 1 basis point will put pressure on the corporate sector. Of course, higher $\varphi$ more directly helps shock amplification mechanism than higher $\kappa$.

There are no parameters exactly corresponding to the agency cost of banking sector, so let us consider the influence of bank net worth from the estimation result of structural shock in the next

\textsuperscript{13}300,000 iterations are implemented using the MH within Gibbs. One sample is drawn out of every 10 replicates to reduce the impact of autocorrelations. The posterior distributions are store up total 30,000 samples. Then, we discard first 10,000 samples, and the remaining 20,000 samples are used for calculating moments of posterior distributions.
Next, we check price and wage nominal rigidities (Calvo parameters on price $\theta_P$ and wage $\theta_W$). The nominal price rigidity is about 0.8 (except Case B) indicating price revision duration is about five quarters. The nominal wage rigidity is roughly 0.5 which implies wage revision duration is a half year. In times of severe depression with financial frictions, rather high nominal rigidity may rescue the economy somewhat, since monetary easing under constant price has the effect of reducing the real interest rate significantly. This will have the effect of lowering the borrowing constraints of banking and corporate sectors.

Finally, the Taylor coefficients in the monetary policy rule are stable in all cases (on the monetary policy rule, see Appendix). The interest rate smoothing parameter $\rho^R$ is between 0.61 and 0.67, Taylor coefficient for inflation gap $\mu^\pi$ is around 2.8 to 3.0 and for output gap $\mu^Y$ is tiny such as 0.006 through 0.010. Thus, FED is very conservative as far as the average attitude of the estimation period is concerned: Aggressively reacts for inflation gap, while not so for output gap. However, the volatilities of monetary policy shock are largely different among the four cases. We will see that time-varying volatilities of monetary policy shock rapidly increase in the period of the Great Recession.

4.6.2 Structural Shocks and Volatilities

Figures 4.1 (a), (b) show the posterior mean and 90% credible interval of the eight structural shocks in Cases A and B (with time-constant volatility shocks), whereas Figures 4.2 (a), (b) are those in Cases C and D (with time-varying volatility shocks). The panels (a) are with 11 data in which the deep blue solid lines for posterior means and light blue shades for the 90% intervals and the panels (b) with 40 data (data rich estimation) in which the deep red solid lines for posterior means and light red shades for the 90% intervals. Figure 4.3 depicts posterior means and 90% intervals of time varying volatilities in Cases C and D.

With the impression that we looked at estimated shocks (Figures 4.1 and 4.2), we have at least four findings:

First, in all cases during the financial crisis, as expected, negative spikes are observed in both banks and corporate net worth shocks. Similarly, positive labor supply shocks can be also observed at the same time (except Case A).

Second, in the cases with SV (Cases C and D), the fluctuations of the posterior mean of shocks are small in peacetime, but large spikes are observed only during the financial crisis. On the other hand, in the cases without SV (Cases A and B), fluctuations in posterior mean of shocks are also observed during normal times. This tendency is particularly noticeable in the net worth shocks. The result could be understood that the restriction of time-constant volatility makes the large shock at the financial crisis smoothed over all the estimation period.

Third, the credible intervals in Cases C and D get higher shrinkage than those in Cases A and B, i.e. shocks estimated accurately in cases with SV. In particular, the red shades in Case D covers almost all of area of blue shade in Case C. This result may imply that the assumption imposed on the shock (time-constant volatility) was more restrictive than the constraint of the data information.\textsuperscript{14}

Finally, related to the second finding, structural shocks estimated with a large number of data (red shades) seem to fluctuate with bigger swing than those (blue shade) of the standard approach.

\textsuperscript{14}In the DFM on which the data rich estimation method relies, there is a characteristic that factors are very smoothed by a large number of data information. In the financial crisis where large structural shock spikes are observed, it seems a natural result that the volatile data fluctuations are better explained by introducing time-varying shock volatilities than by matching the smoothed factors with high frequency data.
CHAPTER 4. SOURCES OF THE GREAT RECESSION

Next, let us focus on the two net worth shocks pertaining to the financial frictions in banking and corporate sectors. Table 4.5 provides the timings of the peaks of the two shocks during the financial crisis. At first, the banking net worth shocks have the exactly same peak at 2008:Q3 for all cases. In this period, i.e., September and October 2008, several major financial institutions were either failed, acquired under duress, or subject to government takeover. These financial institutions include Lehman Brothers, Merrill Lynch, Fannie Mae, Freddie Mac, Washington Mutual, Wachovia, Citi group, and AIG. On the other hand, the timings of the peak of corporate net worth shock are not consistent and divided into two periods, i.e., 2009:Q1 in Cases A and B, and 2009Q2 in Cases C and D. We can remark corporate net worth shocks have peak after banking sector shocks hit peak, whatever the case.

We also examine the estimation accuracy of the eight shocks using average range of 90% credible interval over the all sample period. Table 4.6 shows the result. If we observe the 90% interval ranges are smaller, then we can regard the shocks are estimated more precisely. Compared among the four cases, five average intervals of shocks out of eights are smaller in Cases C and D than in Cases A and B. These five shocks are (1) preference, (2) banking net worth, (3) labor supply, (4) government spending and (5) monetary policy shocks. The intervals in the former three shocks are around half in the two cases with time-varying volatility shocks against the other two cases with time-constant ones. The credible interval of government spending shock averagely shrink to about one eighth by adopting SV shocks. The results suggest that time-constant shocks volatilities might be misspecified and shocks follow time-varying volatilities.

Figure 4.3 draws estimated time-varying shocks volatilities in Cases C and D. Surprisingly, the seven shocks volatilities are very similar in both cases (one exception is government spending shock). The six shocks except preference and labor supply shocks are very stable and level off between 1990Q1 and 2007Q3. The preference and labor supply shocks might play an important role of the boom around 2003 to 2005 (we can confirm the negative labor supply shocks, especially in Cases B and D around 2003 to 2005 in Figure 4.1).

What we should pay attention to here is volatilities of corporate and banking sectors’ net worth shocks. Was there a “bad luck” in the sense that the volatilities of the two financial shocks expanded? During the financial crisis of 2007 to 2009, the volatilities of both banking and corporate net worth, investment, TFP, and monetary policy shocks rapidly increased. The estimates show that the magnitudes of volatilities in this period seem extraordinary. So, regarding the previous question, “Yes”. We clearly confirm the negative spikes in both i.i.d net worth shocks (Figure 4.2) and large expansions in both shocks volatilities (Figure 4.3) during the Great Recession.

We should also examine the estimation accuracy of stochastic volatilities. Table 4.7 reports average 90 % intervals of SVs over the entire sample period in the two cases. As seen from the table as well as Figure 4.3, there are no differences of means of interval ranges between Cases C (with 11 data) and D (with 40 data). Thus, additional information does not necessarily improve the estimation accuracy. Together with the results in Table 4.6, it may imply that the constraint on the shock process is more important than the data constraints.

Let us turn to discuss the leverage effects of SV shocks, that is, we examine a possibility of further “bad luck’’. Table 4.11 summarizes the results of the parameters in the SV model. The leverage effect is determined by the sign of the correlation coefficient $\rho_{\sigma}$. If $\rho_{\sigma}$ is negative, the shock has leverage effect which induces the negative shock at the present period amplifies its volatility at the next period, and vice versa. Table 4.8 sums up the sign of the correlation coefficient $\rho_{\sigma}$ in terms of 90% credible interval. The mark “-” indicates negative $\rho_{\sigma}$ (leverage effect) at 90% credible degree
of posterior probability, while the mark “+” does positive $\rho_\sigma$ (opposite leverage effect). The mark “0” means we judge no leverage effect since zero is within 90% interval. According to empirical financial studies, the leverage effects are often observed, e.g. in stock price data. Our question is whether bank and corporate net worth shocks have the leverage effect. In other words, did we face a further “bad luck” that an adverse net worth shock leads to expand its volatility during the Great Recession? According to our result, “No”. The leverage effect cannot be detected in both net worth shocks.

Finally, we consider the monetary policy in the period of the Great Recession. It should be noted that we estimate the linear Taylor rule for the sample period including QE1 (round 1 of quantitative easing by FED, between 2008Q4 and 2010Q2) and QE2 (2010Q4 to 2011Q2). Monetary policy shocks in Figures 4.1 and 4.2 seem to have two big negative spikes after 2007. The first negative spike is observed at 2007Q4 when BNP Paribas announcement impacts on global financial market. And the second one is observed at 2008Q3 immediately before an unconventional monetary policy (QE1) was conducted by the FED. In particular, the magnitudes of these two negative shocks are distinguished in the cases with time-varying volatility as Figure 4.2. Figure 4.3 also captures rapidly appreciation of these volatilities of policy shocks in the period between 2007Q4 and 2008Q3. Table 4.8 shows monetary policy has “opposite” leverage effect over the entire sample periods, that is, tightening policy is likely to be conducted more boldly without hesitation, while easing policy might be done more carefully, according to the results with 90 % credible degree of posterior probability. Nevertheless, the conservative FED took tremendous monetary easing policies in the Great Recession. So, did “good policy” exist? “Yes”. We can confirm the strongly negative monetary policy shock with extremely high volatility from immediately after the financial crisis.

### 4.6.3 Historical Decompositions

Let us investigate the sources of the Great Recession by historical decompositions. We focus on the decompositions on four observable variables; (1) real GDP as output gap, (2) gross private domestic investment as investment, (3) Moody’s bond index (corporate Baa) as corporate borrowing rate, (4) commercial banks leverage ratio as bank leverage ratio (described in detail in Data Appendix). Figures 4.4 to 4.7 draw four decompositions from 2000Q1 to 2012Q2, and light blue shades denote the period of Great Recession (2007Q3 to 2009Q2). To facilitate visualization and focus on contributions of two financial frictions, technology and monetary policy shocks for the recession, we collect the remaining four miscellaneous shocks as one bundle in these figures.

The recession stories of our financial friction model are as follows: bank and corporate balance sheets get worse by adverse net worth shocks. Through the amplification effect of capital price decline, the leverage of the bank rises and the corporate borrowing rate increases. Both increase bank and corporate borrowing constraints. As a result, investment declines sharply and production is getting cold.

At first, we consider real activities (output and investment). Figures 4.4 and 4.5 show historical decompositions of real GDP and gross private domestic investment, respectively. The results of real activities are qualitatively similar, that is, the signs of shocks’ contributions are the same in all cases. But, the magnitudes of the shocks’ contributions are different among cases. Case A (standard estimation method) tells the TFP shock is the source of the U.S. “business cycle”(output fluctuation). Almost all of boom, recession and slow recovery are explained by TFP shock. This is partly correct. In all cases, it has already been detected that TFP shock was contributing negatively to real activities since around 2005, despite being the shock with high inertia (in all cases, over 0.95, Tables 4.9 and 4.10). Although it is not a big negative spike, we can confirm negative shocks
continued from early 2005 in the i.i.d. TFP shock (Figures 4.1 and 4.2). The second position is the corporate net worth shock. Especially before the recession period, big positive contributions of this shock was driving the economy. So, the Great recession is partly due to the lack of the large driving force.

On the other hand, the remaining three cases (Cases B, C and D with extended estimation method) provides different stories. The corporate net worth shock is no longer the role of the economic towing, but turned into a role that pulls the leg of the economic recovery greatly after the recession. Instead, bank net worth shock accounts for relatively bigger place of downturn of investment and output in the period. It should be noted that corporate net worth shock has high inertia (about 0.9, except Case C, Tables 4.9 and 4.10), but the peak time of the i.i.d. shock was still after the bank’s net worth deterioration (Table 4.5). More interestingly, the bank’s net worth shock has contributed greatly to the economic recovery right after the recession period (in Cases B and D). What happened to the balance sheet of the banking sector? Troubled Asset Relief Program (TARP), in which U.S. government purchased assets and equity of banking sector up to $700 billion in October 2008. TARP works and prominently improves bank’s balance sheet. As a result, the bank’s positive net worth shock became powerfully one of the main sources of driving the economy after the Great Recession.

Recall that the model is the same although we change the estimation method or shock generation process. It is worth noting that the results of considering the main factors of the business cycle are so different. In all cases, we estimate the same model that introduced financial friction. Nevertheless, Case A detects different source (TFP shock) as the main source of the recession (although it is consistent with previous studies). The policy response of monetary authority is different depending on demand shock or supply shock on the source of the recession. In the case of a supply shock like TFP shock, the economic recovery will be accompanied by a sacrifice of inflation and it will be a cautious response. On the other hand, in the case of a demand shock like net worth shock, it is not necessary to sacrifice inflation, so it is possible to take drastic policy response.

Figure 4.6 decomposes corporate borrowing rate (corporate Baa of Moody’s bond index). According to figures, a sharp rise of the rate might be derived from mainly negative bank net worth shock as well as a fall of TFP shock, whereas positive corporate net worth shock contributed to the fall of own borrowing rate in the recession. But after the recession, corporate net worth shock turns to be remarkably negative, which seriously deteriorates its balance sheet and accounts for large portion of rise of the rate after the recession. On the other hand, TARP work well and make bank net worth shock turns to positive, and this makes remarkable contributions to the relaxation of corporate borrowing constraints after 2010:Q1. In particular, we can see these findings in Cases A, B, and D.

Finally, we examine the historical decompositions on bank leverage ratio in Figure 4.7 (commercial banks leverage ratio. Again, it should be noted that we defined the leverage ratio as the reciprocal of the commonly-used ratio, i.e., bank asset over bank net worth). The tremendous positive spike in 2008:Q3 was caused by the damage of the net worth of bank and corporate sectors in any case. Since the bank’s net worth shock is extremely low inertia (from 0.02 to 0.20, Table 4.9), the positive spike of leverage matches the negative peak time of i.i.d. shock (Table 4.5). Soon after the recession, an improve of bank balance sheet by TARP rapidly lowers the bank leverage (Cases A, B and D). However, negative corporate net worth shock makes corporate balance sheet much worse, leading to raise bank leverage about a year.
4.6.4 Discussions

Overall, we can make three important observations based on our empirical results.

First, as for the timing of the financial shocks during the period of Great Recession shown in Figures 4.1 and 4.2, we observed that the bank net worth shock occurred earlier than the corporate net worth shock. Putting it differently, two financial shocks did not occur concurrently, but the corporate net worth shock occurred just shortly after the bank net worth shock. This timing pattern (not concurrent, but proximate timing) may points to the possibility of endogenous relationship between the balance sheet conditions of the banking sector and the corporate sector. For instance, in reality, it is possible for the corporate sector to hold financial sector’s equity as an asset and the devaluation of the financial sector’s asset may affect the balance sheet condition of the corporate sector. However, our model does not allow the corporate sector to hold the asset fully in the form of physical capital and further assumes the two financial shocks to be independent with each other. Thus, it is inappropriate to interpret the endogenous relationship between two financial shocks in the context of the model assumed in our study. Yet, the timing of the two financial shocks during the Great Recession is worth noting.

Second, through the historical decomposition results shown in Figure 4.4 to Figure 4.7, we observed that the corporate net worth shock during the Great Recession to be relatively weak in Case A, compared to those in Case B, C, and D. This results may turn out the possibility of underestimation of the importance of corporate net worth shock when the model is estimated by a plain-vanilla Bayesian estimation method, i.e., without data rich estimation or stochastic volatility. Moreover, an accurate estimation of corporate net worth shock during the Great Recession is crucially important in accounting for the recovery of the U.S. economy in recent years. For instance, in Case A, a slow recovery of output is mainly accounted by negative productivity shock, while in Case B, C, and D, it is mainly accounted by a prolonged negative corporate net worth shock. A slow recovery of the U.S. economy after the Great Recession remain as an important puzzle and persuasive explanation of this puzzle calls for an accurate estimation of the structural shocks. For accurate estimation of the structural shocks (especially for corporate net worth shock), data rich estimation with stochastic volatility may be more reliable than a plain-vanilla Bayesian estimation method.

Third, another important observation from the historical decomposition results is the behavior of bank net worth shock. Bank net worth shock declines sharply during the Great Recession and is the main source of the sharp decline in output and investment as shown in Figures 4.4 and 4.5. But then, right after the Great Recession period, bank net worth shock quickly reverses its direction and contributes positively to output and investment. Considering the timing of this reversal, it is quite possible that the TARP is behind this reversal. In other words, TARP may have successfully countered the negative bank net worth shock. Interpreting further, considering the positive contribution of bank net worth shock to output and investment right after the Great Recession period, TARP may be one of the major reasons in stopping the spell of the Great Recession and contributing to the recovery (albeit weak) of the U.S. in recent years.

4.7 Conclusion

What shocks triggered in the Great Recession? How deteriorated bank and corporate balance sheets affected the recession? What kind of channels did the damage to the balance sheets have caused negative effects on real activities? Did we face with ”bad luck” that was hit by big negative shocks
not occurring in peacetime? Did there be any further “bad luck” that adverse shocks would expand their volatilities? Was ”good policy” implemented in the recession?

To clarify the channel affecting the business cycle by balance sheet loss, we first extended the standard DSGE model by embedding financial frictions in both banking and corporate sectors. The model installed two types of agency problems between borrowers and lenders: One is the agency cost between corporate sector (borrower) and bank sector (lender) due to asymmetric information. The other is the agency cost between banks (borrower) and depositors (lender) caused by bank’s moral hazard/costly verification problem. To verify the possibility of big shocks that would not occur under normal times, we introduced time-varying volatilities into structural shocks. We also considered the effect that the bad shock itself expands the magnitude of its own shock. Then, to improve the estimation accuracy of the shocks causing the business cycle, we adopted an estimation method that makes use of a large number of data information. According to our results, the replies to the questions are as follows:

What shocks triggered in the Great Recession? The trigger was deteriorating net worth of banking sector (2008:Q3). The worsening of the net worth of the corporate sector followed it. A shock in which worker’s reservation wage rose also occurred at the same time. However, the problem of productivity decline has already occurred before that.

How deteriorated bank and corporate balance sheets affected the recession? The rise in banking leverage ratio and the rise in corporate borrowing rate increased the borrowing constraints of both sectors and reduced investment and output.

What kind of channels did the damage to the balance sheets have caused negative effects on real activities? The channels are shock amplification effects due to agency costs by asymmetric information and moral hazard. The spread equivalent to four basis points was added to the corporate borrowing rate against 1% increase in corporate leverage.

Did we face with ”bad luck” that was hit by big negative shocks not occurring in peacetime? Yes. The volatilities of the negative net worth stock shocks of banks and corporate sectors were greatly expanding at that time.

Did there be any further “bad luck” that adverse shocks would expand their volatility? No. The leverage effects were not detected.

Was ”good policy” implemented in the recession? Yes. Since a large negative spike in monetary policy shock and an expansion of its volatility were observed, usually conservative central bank implemented bold monetary easing during the recession. Also, if the improvement in the balance sheet of the banking sector after the recession was due to TARP, fiscal and monetary authorities played a major role in the economic recovery.

Finally, the introduction of time-varying volatility and the estimation method that utilizes a lot of data are extremely important. Because even if the model does not change, the sources of the business cycle changes accordingly. If we misjudge the sources, there might be the possibility of bringing wrong policy responses.
4.8 Tables and Figures

Table 4.1: Specifications of Four Alternative Cases

<table>
<thead>
<tr>
<th></th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observation</td>
<td>11</td>
<td>40</td>
<td>11</td>
<td>40</td>
</tr>
<tr>
<td>Model Variable to Obs.</td>
<td>1 to 1</td>
<td>1 to 4</td>
<td>1 to 1</td>
<td>1 to 4</td>
</tr>
<tr>
<td>Structural Shock</td>
<td>i.i.d. Normal</td>
<td>i.i.d. Normal</td>
<td>SV with Leverage</td>
<td>SV with Leverage</td>
</tr>
</tbody>
</table>

Note: Item “Number of Observation” in the first column denotes the number of data indicators used for estimating the model of each case. Item “Model Variable to Obs” denote the ratio what number of observations per one model variable are adopted. In the case of a standard DSGE model, we adopt one to one matching between model variables and observations. In data rich approach, one to many matching are adopted between model variables and observations. Item “Structural Shock” denotes specification of stochastic process of shocks. SV is abbreviation of stochastic volatility.
### Table 4.2: Calibrated Parameters and Key Steady States

<table>
<thead>
<tr>
<th>Calibrated Param.</th>
<th>Description</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Discount factor</td>
<td>0.995</td>
<td>Our setting</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Depreciation rate</td>
<td>0.025</td>
<td>Christensen and Dib (2008)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Capital share</td>
<td>0.33</td>
<td>Gertler and Kiyotaki (2010)</td>
</tr>
<tr>
<td>$\gamma^E_{ss}$</td>
<td>Survival rate of entrepreneur in steady state</td>
<td>0.972</td>
<td>Christensen and Dib (2008)</td>
</tr>
<tr>
<td>$\gamma^F_{ss}$</td>
<td>Survival rate of banker in steady state</td>
<td>0.972</td>
<td>Gertler and Kiyotaki (2010)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Bank’s participation constraint parameter</td>
<td>0.383</td>
<td>Gertler and Kiyotaki (2010)</td>
</tr>
<tr>
<td>$\psi^w$</td>
<td>Wage markup</td>
<td>0.05</td>
<td>Smets and Wouters (2003)</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Elasticity Substitution of intermediate goods</td>
<td>11</td>
<td>Our setting</td>
</tr>
<tr>
<td>$\xi^E$</td>
<td>New entrepreneur entry rate</td>
<td>0.003</td>
<td>Our setting</td>
</tr>
<tr>
<td>$\xi^F$</td>
<td>New banker entry rate</td>
<td>0.003</td>
<td>Gertler and Kiyotaki (2010)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key Steady State</th>
<th>Description</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$mc_{ss}$</td>
<td>Steady state marginal cost</td>
<td>$\frac{\epsilon}{1-\epsilon}$</td>
<td>-</td>
</tr>
<tr>
<td>$S_{ss}$</td>
<td>Steady state external financial premium</td>
<td>1.0075</td>
<td>Christensen and Dib (2008)</td>
</tr>
<tr>
<td>$rr^E_{ss}$</td>
<td>Steady state corp. borrowing rate (real, QPR)</td>
<td>1.0152</td>
<td>From data (1980Q1-2010Q2)</td>
</tr>
<tr>
<td>$rr^F_{ss}$</td>
<td>Steady state bank lending rate (real, QPR, ex-premium)</td>
<td>$rr^E_{ss}/S_{ss}$</td>
<td>-</td>
</tr>
<tr>
<td>$rr_{ss}$</td>
<td>Steady state real interest</td>
<td>$\frac{1}{\beta}$</td>
<td>-</td>
</tr>
<tr>
<td>$\nu_{ss}$</td>
<td>Steady state Nu</td>
<td>$\frac{(1-\gamma^F_{ss})\beta(rr^E_{ss}-rr_{ss})}{(1-\gamma^F_{ss})}$</td>
<td>-</td>
</tr>
<tr>
<td>$\eta_{ss}$</td>
<td>Steady state Eta</td>
<td>$\frac{1-\gamma^F_{ss}}{1-\beta\gamma^F_{ss}}$</td>
<td>-</td>
</tr>
<tr>
<td>Lev$^F_{ss}$</td>
<td>Steady state leverage ratio of banker</td>
<td>$\frac{\xi^F_{ss}}{\lambda^F_{ss}}$</td>
<td>-</td>
</tr>
<tr>
<td>$K_{ss}/N^E_{ss}$</td>
<td>Steady state leverage ratio of entrepreneur</td>
<td>1.919</td>
<td>From data (1980Q1-2010Q2)</td>
</tr>
<tr>
<td>$K_{ss}/Y_{ss}$</td>
<td>Steady state capital/output ratio</td>
<td>$\frac{\alpha \cdot mc_{ss}}{rr^E_{ss}-[1-\delta]}$</td>
<td>-</td>
</tr>
<tr>
<td>$I_{ss}/Y_{ss}$</td>
<td>Steady state investment/output ratio</td>
<td>$\frac{\delta K_{ss}/Y_{ss}}{Y_{ss}}$</td>
<td>-</td>
</tr>
<tr>
<td>$G_{ss}/Y_{ss}$</td>
<td>Steady state government expenditure/output ratio</td>
<td>0.2</td>
<td>Gertler and Kiyotaki (2010)</td>
</tr>
<tr>
<td>$C_{ss}/Y_{ss}$</td>
<td>Steady state consumption/output ratio</td>
<td>$1 - I_{ss}/Y_{ss} - G_{ss}/Y_{ss}$</td>
<td>-</td>
</tr>
</tbody>
</table>
### Table 4.3: Prior Settings of Structural Parameters

<table>
<thead>
<tr>
<th>Structural Parameters</th>
<th>Description</th>
<th>Density</th>
<th>Prior Mean</th>
<th>Prior SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>Investment adjustment cost</td>
<td>Gamma</td>
<td>1.000</td>
<td>0.500</td>
</tr>
<tr>
<td>$h$</td>
<td>Habit formation</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\sigma^C$</td>
<td>IES of consumption</td>
<td>Gamma</td>
<td>1.500</td>
<td>0.500</td>
</tr>
<tr>
<td>$\sigma^L$</td>
<td>Inverse Frisch elasticity of labor supply</td>
<td>Gamma</td>
<td>1.500</td>
<td>0.500</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>Elasticity of premium to leverage ratio</td>
<td>Inv. Gamma</td>
<td>0.050</td>
<td>4.000</td>
</tr>
<tr>
<td>$\iota_P$</td>
<td>Price indexation</td>
<td>Beta</td>
<td>0.500</td>
<td>0.100</td>
</tr>
<tr>
<td>$\iota_W$</td>
<td>Wage indexation</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\theta_P$</td>
<td>Calvo parameter for goods pricing</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\theta_W$</td>
<td>Calvo parameter for wage setting</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_R$</td>
<td>Monetary policy persist. param.</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\mu_Y$</td>
<td>Taylor coefficient for inflation</td>
<td>Gamma</td>
<td>1.500</td>
<td>0.500</td>
</tr>
<tr>
<td>$\mu_Y$</td>
<td>Taylor coefficient for output gap</td>
<td>Gamma</td>
<td>0.500</td>
<td>0.250</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Persistence Parameters for Structural Shocks</th>
<th>Description</th>
<th>Density</th>
<th>Prior Mean</th>
<th>Prior SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_A$</td>
<td>Persistent parameter for TFP shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_C$</td>
<td>Persistent parameter for preference shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_K$</td>
<td>Persistent parameter for investment tech. shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_E$</td>
<td>Persistent parameter for entrepreneur net worth shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_F$</td>
<td>Persistent parameter for banking sector net worth shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_G$</td>
<td>Persistent parameter for government expenditure shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
<tr>
<td>$\rho_L$</td>
<td>Persistent parameter for labor supply shock</td>
<td>Beta</td>
<td>0.500</td>
<td>0.250</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard Errors for Structural Shocks</th>
<th>Description</th>
<th>Density</th>
<th>Prior Mean</th>
<th>Prior SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_A$</td>
<td>SE of TFP shock</td>
<td>Inv. Gamma</td>
<td>0.707</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_C$</td>
<td>SE of preference shock</td>
<td>Inv. Gamma</td>
<td>0.707</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_E$</td>
<td>SE of entrepreneur net worth shock</td>
<td>Inv. Gamma</td>
<td>0.707</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_F$</td>
<td>SE of banking sector net worth shock</td>
<td>Inv. Gamma</td>
<td>0.707</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_G$</td>
<td>SE of government expenditure shock</td>
<td>Inv. Gamma</td>
<td>0.707</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_K$</td>
<td>SE of investment specific technology shock</td>
<td>Inv. Gamma</td>
<td>1.000</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_L$</td>
<td>SE of labor supply shock</td>
<td>Inv. Gamma</td>
<td>0.707</td>
<td>4.000</td>
</tr>
<tr>
<td>$\epsilon_R$</td>
<td>SE or monetary policy shock</td>
<td>Inv. Gamma</td>
<td>0.224</td>
<td>4.000</td>
</tr>
</tbody>
</table>
Table 4.4: Posterior Estimates of Key Structural Parameters

<table>
<thead>
<tr>
<th>Parameters for Financial Friction in Corporate Section</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>0.614</td>
<td>0.877</td>
<td>0.564</td>
<td>0.562</td>
</tr>
<tr>
<td></td>
<td>[0.547, 0.689]</td>
<td>[0.818, 0.938]</td>
<td>[0.498, 0.632]</td>
<td>[0.470, 0.661]</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.027</td>
<td>0.025</td>
<td>0.039</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>[0.024, 0.030]</td>
<td>[0.023, 0.026]</td>
<td>[0.032, 0.045]</td>
<td>[0.036, 0.046]</td>
</tr>
</tbody>
</table>

Parameters for Nominal Rigidities

<table>
<thead>
<tr>
<th>Parameters for Nominal Rigidities</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_P$</td>
<td>0.854</td>
<td>0.374</td>
<td>0.804</td>
<td>0.760</td>
</tr>
<tr>
<td></td>
<td>[0.811, 0.895]</td>
<td>[0.305, 0.440]</td>
<td>[0.763, 0.846]</td>
<td>[0.697, 0.822]</td>
</tr>
<tr>
<td>$\theta_W$</td>
<td>0.589</td>
<td>0.428</td>
<td>0.623</td>
<td>0.516</td>
</tr>
<tr>
<td></td>
<td>[0.531, 0.649]</td>
<td>[0.351, 0.500]</td>
<td>[0.544, 0.703]</td>
<td>[0.452, 0.580]</td>
</tr>
</tbody>
</table>

Parameters for Monetary Policy Rule

<table>
<thead>
<tr>
<th>Parameters for Monetary Policy Rule</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_R$</td>
<td>0.670</td>
<td>0.643</td>
<td>0.653</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>[0.581, 0.758]</td>
<td>[0.582, 0.707]</td>
<td>[0.605, 0.698]</td>
<td>[0.590, 0.675]</td>
</tr>
<tr>
<td>$\mu_\pi$</td>
<td>2.805</td>
<td>2.820</td>
<td>2.989</td>
<td>2.986</td>
</tr>
<tr>
<td></td>
<td>[2.767, 2.842]</td>
<td>[2.790, 2.848]</td>
<td>[2.979, 2.998]</td>
<td>[2.977, 2.995]</td>
</tr>
<tr>
<td>$\mu_Y$</td>
<td>0.006</td>
<td>0.010</td>
<td>0.006</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>[0.000, 0.014]</td>
<td>[0.000, 0.020]</td>
<td>[0.000, 0.013]</td>
<td>[0.000, 0.018]</td>
</tr>
</tbody>
</table>

Note: The parenthesis in the table indicates 90% credible interval of structural parameters. 300,000 iterations are implemented using algorithm of MH within Gibbs described in Section 4.4. We sample one draw out of every 10 replicates and discard first 10,000 samples. The remaining 20,000 samples are used for calculating moments of the posterior distributions.

Table 4.5: Timings of Peaks of the Financial Shocks

<table>
<thead>
<tr>
<th>Structural Shock</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corp. Net Worth</td>
<td>2009Q1</td>
<td>2009Q1</td>
<td>2009Q2</td>
<td>2009Q2</td>
</tr>
<tr>
<td>Bank Net Worth</td>
<td>2008Q3</td>
<td>2008Q3</td>
<td>2008Q3</td>
<td>2008Q3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stochastic Volatilities</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corp. Net Worth</td>
<td>-</td>
<td>-</td>
<td>2009Q2</td>
<td>2009Q2</td>
</tr>
<tr>
<td>Bank Net Worth</td>
<td>-</td>
<td>-</td>
<td>2009Q3</td>
<td>2009Q3</td>
</tr>
</tbody>
</table>
4.8. TABLES AND FIGURES

Table 4.6: Average Ranges of 90% Credible Interval of Structural Shocks over the entire sample periods

<table>
<thead>
<tr>
<th>Structural Shocks</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFP</td>
<td>0.635</td>
<td>0.353</td>
<td>0.528</td>
<td>0.539</td>
</tr>
<tr>
<td>Preference</td>
<td>1.593</td>
<td>1.633</td>
<td>1.058</td>
<td>0.824</td>
</tr>
<tr>
<td>Corp. Net Worth</td>
<td>0.141</td>
<td>0.148</td>
<td>0.246</td>
<td>0.216</td>
</tr>
<tr>
<td>Bank Net Worth</td>
<td>1.902</td>
<td>1.433</td>
<td>0.886</td>
<td>0.907</td>
</tr>
<tr>
<td>Government Expenditure</td>
<td>2.207</td>
<td>2.018</td>
<td>0.417</td>
<td>0.322</td>
</tr>
<tr>
<td>Investment</td>
<td>0.983</td>
<td>0.236</td>
<td>0.575</td>
<td>1.107</td>
</tr>
<tr>
<td>Labor Supply</td>
<td>2.516</td>
<td>3.133</td>
<td>1.447</td>
<td>1.430</td>
</tr>
<tr>
<td>Monetary Policy</td>
<td>0.121</td>
<td>0.178</td>
<td>0.127</td>
<td>0.126</td>
</tr>
</tbody>
</table>

Note: This table reports the average of the difference between the upper and the lower bounds of 90% credible interval of the structural shock over the entire sample periods (1985Q2-2012Q2), depicted in Figures 4.1 and 4.2.

Table 4.7: Average Ranges of 90% Credible Interval of Stochastic Volatilities in the entire sample periods

<table>
<thead>
<tr>
<th>Structural Shocks</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFP</td>
<td>0.498</td>
<td>0.384</td>
</tr>
<tr>
<td>Preference</td>
<td>0.906</td>
<td>0.857</td>
</tr>
<tr>
<td>Corp. Net Worth</td>
<td>0.243</td>
<td>0.219</td>
</tr>
<tr>
<td>Bank Net Worth</td>
<td>1.043</td>
<td>0.908</td>
</tr>
<tr>
<td>Government Expenditure</td>
<td>0.709</td>
<td>0.769</td>
</tr>
<tr>
<td>Investment</td>
<td>0.604</td>
<td>0.592</td>
</tr>
<tr>
<td>Labor Supply</td>
<td>1.743</td>
<td>1.378</td>
</tr>
<tr>
<td>Monetary Policy</td>
<td>0.095</td>
<td>0.095</td>
</tr>
</tbody>
</table>

Note: This table reports the average value in the entire sample periods (1985Q2-2012Q2) of the difference between the upper bound and the lower bound of 90% credible interval on the stochastic volatiliy for the structural shock depicted in Figure 4.3.

Table 4.8: Leverage Effect of Structural Shocks: Correlation between the Sign of Shock and its Volatility

<table>
<thead>
<tr>
<th>Structural Shocks</th>
<th>Case C</th>
<th>Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFP</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Preference</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Corp. Net Worth</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Bank Net Worth</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Government Expenditure</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Investment</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Labor Supply</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Monetary Policy</td>
<td>0</td>
<td>+</td>
</tr>
</tbody>
</table>

Note: The mark “-” indicates negative of $\rho_{\sigma}$ (leverage effect) at 90% credible degree of posterior probability, while the mark “+” does positive of $\rho_{\sigma}$ (opposite leverage effect) in similar way. The mark “0” implies that we do not judge the sign of $\rho_{\sigma}$ and leverage effect of each shock because zero is within 90% credible interval of $\rho_{\sigma}$. 
Table 4.9: Posterior Estimates: Case A and Case B

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case A</th>
<th></th>
<th>Case B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>κ</td>
<td>0.614 0.043 [0.547 0.689]</td>
<td></td>
<td>0.877 0.038 [0.818 0.938]</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>0.464 0.045 [0.396 0.537]</td>
<td></td>
<td>0.597 0.040 [0.535 0.661]</td>
<td></td>
</tr>
<tr>
<td>σC</td>
<td>1.628 0.036 [1.578 1.688]</td>
<td></td>
<td>1.404 0.032 [1.356 1.451]</td>
<td></td>
</tr>
<tr>
<td>σL</td>
<td>0.939 0.071 [0.819 1.052]</td>
<td></td>
<td>0.417 0.072 [0.323 0.524]</td>
<td></td>
</tr>
<tr>
<td>ϕ</td>
<td>0.027 0.002 [0.024 0.030]</td>
<td></td>
<td>0.025 0.001 [0.023 0.026]</td>
<td></td>
</tr>
<tr>
<td>τp</td>
<td>0.521 0.027 [0.478 0.566]</td>
<td></td>
<td>0.358 0.017 [0.330 0.386]</td>
<td></td>
</tr>
<tr>
<td>τW</td>
<td>0.422 0.009 [0.408 0.437]</td>
<td></td>
<td>0.450 0.007 [0.440 0.459]</td>
<td></td>
</tr>
<tr>
<td>θP</td>
<td>0.854 0.026 [0.811 0.895]</td>
<td></td>
<td>0.374 0.041 [0.305 0.440]</td>
<td></td>
</tr>
<tr>
<td>θW</td>
<td>0.589 0.037 [0.531 0.649]</td>
<td></td>
<td>0.428 0.048 [0.351 0.500]</td>
<td></td>
</tr>
<tr>
<td>ρR</td>
<td>0.670 0.055 [0.581 0.758]</td>
<td></td>
<td>0.643 0.038 [0.582 0.707]</td>
<td></td>
</tr>
<tr>
<td>μπ</td>
<td>2.805 0.025 [2.767 2.842]</td>
<td></td>
<td>2.820 0.018 [2.790 2.848]</td>
<td></td>
</tr>
<tr>
<td>μY</td>
<td>0.006 0.005 [0.000 0.014]</td>
<td></td>
<td>0.010 0.007 [0.000 0.020]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case A</th>
<th></th>
<th>Case B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ρA</td>
<td>0.975 0.007 [0.964 0.986]</td>
<td></td>
<td>0.975 0.005 [0.966 0.983]</td>
<td></td>
</tr>
<tr>
<td>ρC</td>
<td>0.636 0.093 [0.504 0.788]</td>
<td></td>
<td>0.088 0.053 [0.004 0.166]</td>
<td></td>
</tr>
<tr>
<td>ρK</td>
<td>0.391 0.044 [0.320 0.462]</td>
<td></td>
<td>0.998 0.001 [0.996 0.999]</td>
<td></td>
</tr>
<tr>
<td>ρE</td>
<td>0.907 0.022 [0.873 0.944]</td>
<td></td>
<td>0.976 0.012 [0.959 0.996]</td>
<td></td>
</tr>
<tr>
<td>ρF</td>
<td>0.031 0.024 [0.000 0.064]</td>
<td></td>
<td>0.016 0.011 [0.000 0.031]</td>
<td></td>
</tr>
<tr>
<td>ρG</td>
<td>0.798 0.047 [0.733 0.864]</td>
<td></td>
<td>0.671 0.012 [0.652 0.686]</td>
<td></td>
</tr>
<tr>
<td>ρL</td>
<td>0.933 0.041 [0.876 0.995]</td>
<td></td>
<td>0.967 0.009 [0.953 0.982]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case A</th>
<th></th>
<th>Case B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>eA</td>
<td>0.564 0.043 [0.492 0.629]</td>
<td></td>
<td>0.398 0.030 [0.347 0.447]</td>
<td></td>
</tr>
<tr>
<td>eC</td>
<td>1.475 0.161 [1.242 1.716]</td>
<td></td>
<td>1.729 0.189 [1.441 1.986]</td>
<td></td>
</tr>
<tr>
<td>eK</td>
<td>0.238 0.016 [0.212 0.265]</td>
<td></td>
<td>0.286 0.020 [0.254 0.318]</td>
<td></td>
</tr>
<tr>
<td>eE</td>
<td>0.787 0.072 [0.689 0.918]</td>
<td></td>
<td>1.423 0.042 [1.358 1.491]</td>
<td></td>
</tr>
<tr>
<td>eF</td>
<td>0.757 0.057 [0.690 0.843]</td>
<td></td>
<td>0.890 0.058 [0.811 0.979]</td>
<td></td>
</tr>
<tr>
<td>eG</td>
<td>0.520 0.050 [0.439 0.603]</td>
<td></td>
<td>0.895 0.119 [0.751 1.102]</td>
<td></td>
</tr>
<tr>
<td>eL</td>
<td>0.881 0.110 [0.722 1.060]</td>
<td></td>
<td>1.383 0.040 [1.325 1.448]</td>
<td></td>
</tr>
<tr>
<td>eR</td>
<td>0.228 0.016 [0.201 0.255]</td>
<td></td>
<td>0.245 0.019 [0.215 0.274]</td>
<td></td>
</tr>
</tbody>
</table>

Note: 300,000 iterations are implemented using MH within Gibbs described in Section 4.4. We sample one draw out of every 10 replicates and discard first 10,000 samples. The remaining 20,000 samples are used for calculating moments of the posterior distributions. Items SD and 90% CI denote the standard deviations and 90% credible intervals of the posterior distributions of the structural parameters, respectively.
### Table 4.10: Posterior Estimates: Case C and Case D

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case C</th>
<th></th>
<th>Case D</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>90% CI</td>
<td>Mean</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.564</td>
<td>0.041</td>
<td>[0.498, 0.632]</td>
<td>0.562</td>
</tr>
<tr>
<td>$h$</td>
<td>0.334</td>
<td>0.038</td>
<td>[0.271, 0.396]</td>
<td>0.221</td>
</tr>
<tr>
<td>$\sigma_C$</td>
<td>1.630</td>
<td>0.012</td>
<td>[1.613, 1.649]</td>
<td>1.605</td>
</tr>
<tr>
<td>$\sigma_L$</td>
<td>0.819</td>
<td>0.021</td>
<td>[0.786, 0.855]</td>
<td>0.597</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.039</td>
<td>0.004</td>
<td>[0.032, 0.045]</td>
<td>0.041</td>
</tr>
<tr>
<td>$\iota_P$</td>
<td>0.397</td>
<td>0.014</td>
<td>[0.376, 0.422]</td>
<td>0.503</td>
</tr>
<tr>
<td>$\iota_W$</td>
<td>0.475</td>
<td>0.002</td>
<td>[0.472, 0.479]</td>
<td>0.489</td>
</tr>
<tr>
<td>$\theta_P$</td>
<td>0.804</td>
<td>0.025</td>
<td>[0.763, 0.846]</td>
<td>0.760</td>
</tr>
<tr>
<td>$\theta_W$</td>
<td>0.623</td>
<td>0.049</td>
<td>[0.544, 0.703]</td>
<td>0.516</td>
</tr>
<tr>
<td>$\rho_R$</td>
<td>0.653</td>
<td>0.029</td>
<td>[0.605, 0.698]</td>
<td>0.632</td>
</tr>
<tr>
<td>$\mu_\pi$</td>
<td>2.989</td>
<td>0.006</td>
<td>[2.979, 2.998]</td>
<td>2.986</td>
</tr>
<tr>
<td>$\mu_\nu$</td>
<td>0.006</td>
<td>0.006</td>
<td>[0.000, 0.013]</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Case C</td>
<td></td>
<td>Case D</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>----------------</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>90% CI</td>
<td>Mean</td>
</tr>
<tr>
<td>$\rho_A$</td>
<td>0.989</td>
<td>0.006</td>
<td>[0.981, 0.999]</td>
<td>0.956</td>
</tr>
<tr>
<td>$\rho_C$</td>
<td>0.819</td>
<td>0.037</td>
<td>[0.757, 0.877]</td>
<td>0.909</td>
</tr>
<tr>
<td>$\rho_K$</td>
<td>0.127</td>
<td>0.050</td>
<td>[0.038, 0.202]</td>
<td>0.776</td>
</tr>
<tr>
<td>$\rho_E$</td>
<td>0.333</td>
<td>0.131</td>
<td>[0.107, 0.518]</td>
<td>0.918</td>
</tr>
<tr>
<td>$\rho_F$</td>
<td>0.192</td>
<td>0.011</td>
<td>[0.174, 0.209]</td>
<td>0.167</td>
</tr>
<tr>
<td>$\rho_G$</td>
<td>0.655</td>
<td>0.006</td>
<td>[0.646, 0.664]</td>
<td>0.619</td>
</tr>
<tr>
<td>$\rho_L$</td>
<td>0.924</td>
<td>0.053</td>
<td>[0.844, 0.991]</td>
<td>0.982</td>
</tr>
</tbody>
</table>

Note: 300,000 iterations are implemented using MH within Gibbs described in Section 4.4. We sample one draw out of every 10 replicates and discard first 10,000 samples. The remaining 20,000 samples are used for calculating moments of the posterior distributions. Items SD and 90% CI denote the standard deviations and 90% credible intervals of the posterior distributions of the structural parameters, respectively.
Table 4.11: Posterior Estimates of Parameters of SVs: Case C and Case D

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>90% CI</th>
<th>Mean</th>
<th>SD</th>
<th>90% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameters of Stochastic Volatilities for TFP Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_A$</td>
<td>0.373</td>
<td>0.099</td>
<td>[0.215, 0.500]</td>
<td>0.338</td>
<td>0.120</td>
<td>[0.158, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_A}$</td>
<td>0.059</td>
<td>0.307</td>
<td>[-0.490, 0.507]</td>
<td>0.347</td>
<td>0.390</td>
<td>[-0.186, 0.989]</td>
</tr>
<tr>
<td>$\phi_A$</td>
<td>0.737</td>
<td>0.168</td>
<td>[0.530, 0.986]</td>
<td>0.509</td>
<td>0.184</td>
<td>[0.213, 0.810]</td>
</tr>
<tr>
<td>$\mu_A$</td>
<td>0.442</td>
<td>0.035</td>
<td>[0.378, 0.491]</td>
<td>0.429</td>
<td>0.049</td>
<td>[0.349, 0.501]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Preference Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_C$</td>
<td>0.479</td>
<td>0.027</td>
<td>[0.456, 0.500]</td>
<td>0.476</td>
<td>0.023</td>
<td>[0.447, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_C}$</td>
<td>0.357</td>
<td>0.190</td>
<td>[0.043, 0.658]</td>
<td>0.481</td>
<td>0.141</td>
<td>[0.226, 0.696]</td>
</tr>
<tr>
<td>$\phi_C$</td>
<td>0.934</td>
<td>0.059</td>
<td>[0.854, 0.990]</td>
<td>0.958</td>
<td>0.037</td>
<td>[0.919, 0.990]</td>
</tr>
<tr>
<td>$\mu_C$</td>
<td>0.656</td>
<td>0.075</td>
<td>[0.544, 0.764]</td>
<td>0.933</td>
<td>0.055</td>
<td>[0.844, 1.026]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Corporate Net Worth Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_E$</td>
<td>0.434</td>
<td>0.053</td>
<td>[0.357, 0.500]</td>
<td>0.412</td>
<td>0.073</td>
<td>[0.303, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_E}$</td>
<td>0.418</td>
<td>0.221</td>
<td>[0.066, 0.785]</td>
<td>0.280</td>
<td>0.329</td>
<td>[-0.217, 0.869]</td>
</tr>
<tr>
<td>$\phi_E$</td>
<td>0.803</td>
<td>0.124</td>
<td>[0.627, 0.990]</td>
<td>0.758</td>
<td>0.186</td>
<td>[0.493, 0.990]</td>
</tr>
<tr>
<td>$\mu_E$</td>
<td>0.149</td>
<td>0.007</td>
<td>[0.139, 0.162]</td>
<td>0.194</td>
<td>0.013</td>
<td>[0.173, 0.212]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Bank Net Worth Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_F$</td>
<td>0.450</td>
<td>0.034</td>
<td>[0.404, 0.500]</td>
<td>0.445</td>
<td>0.041</td>
<td>[0.395, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_F}$</td>
<td>0.161</td>
<td>0.231</td>
<td>[-0.216, 0.543]</td>
<td>0.218</td>
<td>0.199</td>
<td>[-0.132, 0.498]</td>
</tr>
<tr>
<td>$\phi_F$</td>
<td>0.854</td>
<td>0.121</td>
<td>[0.685, 0.990]</td>
<td>0.894</td>
<td>0.066</td>
<td>[0.804, 0.990]</td>
</tr>
<tr>
<td>$\mu_F$</td>
<td>0.665</td>
<td>0.051</td>
<td>[0.598, 0.769]</td>
<td>0.893</td>
<td>0.050</td>
<td>[0.783, 0.959]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Government Expenditure Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_G$</td>
<td>0.266</td>
<td>0.050</td>
<td>[0.182, 0.327]</td>
<td>0.440</td>
<td>0.048</td>
<td>[0.373, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_G}$</td>
<td>0.384</td>
<td>0.350</td>
<td>[-0.152, 0.990]</td>
<td>0.044</td>
<td>0.367</td>
<td>[-0.536, 0.670]</td>
</tr>
<tr>
<td>$\phi_G$</td>
<td>0.603</td>
<td>0.286</td>
<td>[0.178, 0.990]</td>
<td>0.517</td>
<td>0.246</td>
<td>[0.071, 0.891]</td>
</tr>
<tr>
<td>$\mu_G$</td>
<td>0.505</td>
<td>0.028</td>
<td>[0.461, 0.548]</td>
<td>0.570</td>
<td>0.031</td>
<td>[0.519, 0.627]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Investment Specific Technology Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_K$</td>
<td>0.449</td>
<td>0.038</td>
<td>[0.399, 0.500]</td>
<td>0.452</td>
<td>0.063</td>
<td>[0.335, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_K}$</td>
<td>0.304</td>
<td>0.318</td>
<td>[-0.228, 0.841]</td>
<td>0.128</td>
<td>0.246</td>
<td>[-0.215, 0.540]</td>
</tr>
<tr>
<td>$\phi_K$</td>
<td>0.450</td>
<td>0.257</td>
<td>[0.001, 0.838]</td>
<td>0.219</td>
<td>0.214</td>
<td>[0.000, 0.548]</td>
</tr>
<tr>
<td>$\mu_K$</td>
<td>0.496</td>
<td>0.023</td>
<td>[0.457, 0.528]</td>
<td>0.406</td>
<td>0.049</td>
<td>[0.324, 0.476]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Labor Supply Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_L$</td>
<td>0.446</td>
<td>0.043</td>
<td>[0.386, 0.500]</td>
<td>0.482</td>
<td>0.016</td>
<td>[0.458, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_L}$</td>
<td>-0.163</td>
<td>0.308</td>
<td>[-0.781, 0.254]</td>
<td>0.232</td>
<td>0.178</td>
<td>[-0.071, 0.517]</td>
</tr>
<tr>
<td>$\phi_L$</td>
<td>0.779</td>
<td>0.263</td>
<td>[0.291, 0.990]</td>
<td>0.903</td>
<td>0.084</td>
<td>[0.813, 0.990]</td>
</tr>
<tr>
<td>$\mu_L$</td>
<td>1.010</td>
<td>0.116</td>
<td>[0.870, 1.207]</td>
<td>1.461</td>
<td>0.078</td>
<td>[1.351, 1.580]</td>
</tr>
<tr>
<td><strong>Parameters of Stochastic Volatilities for Monetary Policy Shock</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_R$</td>
<td>0.422</td>
<td>0.045</td>
<td>[0.355, 0.493]</td>
<td>0.464</td>
<td>0.034</td>
<td>[0.407, 0.500]</td>
</tr>
<tr>
<td>$\rho_{\sigma_R}$</td>
<td>0.156</td>
<td>0.238</td>
<td>[-0.268, 0.520]</td>
<td>0.479</td>
<td>0.211</td>
<td>[0.122, 0.797]</td>
</tr>
<tr>
<td>$\phi_R$</td>
<td>0.763</td>
<td>0.109</td>
<td>[0.589, 0.948]</td>
<td>0.727</td>
<td>0.122</td>
<td>[0.540, 0.941]</td>
</tr>
<tr>
<td>$\mu_R$</td>
<td>0.099</td>
<td>0.006</td>
<td>[0.089, 0.106]</td>
<td>0.112</td>
<td>0.013</td>
<td>[0.092, 0.131]</td>
</tr>
</tbody>
</table>

Note: 300,000 iterations are implemented using MH within Gibbs described in Section 4.4. We sample one draw out of every 10 replicates and discard first 10,000 samples. The remaining 20,000 samples are used for calculating moments of the posterior distributions. Items SD and 90% CI denote the standard deviations and 90% credible intervals of the posterior distributions of the structural parameters, respectively.
## Data Appendix

### Case A and Case D: The standard one-to-one matching estimation method

<table>
<thead>
<tr>
<th>No.</th>
<th>Variables</th>
<th>Code</th>
<th>Series description</th>
<th>Unit of data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
<td>6</td>
<td>Interest rate: Federal Funds Effective Rate</td>
<td>% per annum</td>
<td>FRB</td>
</tr>
<tr>
<td>2</td>
<td>Y_t</td>
<td>5</td>
<td>Real gross domestic product (excluding net export)</td>
<td>Billion of chained 2000</td>
<td>BEA</td>
</tr>
<tr>
<td>3</td>
<td>C_t</td>
<td>5*</td>
<td>Gross personal consumption expenditures</td>
<td>Billion dollars</td>
<td>BEA</td>
</tr>
<tr>
<td>4</td>
<td>I_t</td>
<td>5*</td>
<td>Gross private domestic investment - Fixed investment</td>
<td>Billion dollars</td>
<td>BEA</td>
</tr>
<tr>
<td>5</td>
<td>\pi_t</td>
<td>8</td>
<td>Price deflator: Gross domestic product</td>
<td>2005Q1 = 100</td>
<td>BEA</td>
</tr>
<tr>
<td>6</td>
<td>w_1</td>
<td>2</td>
<td>Real Wage (Smets and Wouters)</td>
<td>1992Q3 = 0</td>
<td>SW (2007)</td>
</tr>
<tr>
<td>7</td>
<td>L_t</td>
<td>1</td>
<td>Hours Worked (Smets and Wouters)</td>
<td>1992Q3 = 0</td>
<td>SW (2007)</td>
</tr>
<tr>
<td>8</td>
<td>RE_b</td>
<td>6</td>
<td>Moody’s bond indices - corporate Baa</td>
<td>% per annum</td>
<td>Bloomberg</td>
</tr>
<tr>
<td>9</td>
<td>Lev_F</td>
<td>7</td>
<td>Commercial banks leverage ratio</td>
<td>Total asset/net worth ratio</td>
<td>FRB</td>
</tr>
<tr>
<td>10</td>
<td>Lev_E</td>
<td>3</td>
<td>Nonfarm nonfin corp business leverage ratio</td>
<td>Total asset/net worth ratio</td>
<td>FRB</td>
</tr>
</tbody>
</table>

### Case B and Case D: The data-rich estimation method

<table>
<thead>
<tr>
<th>No.</th>
<th>Variables</th>
<th>Code</th>
<th>Series description</th>
<th>Unit of data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Y_2</td>
<td>4</td>
<td>Industrial production index: final products</td>
<td>Index 2007 = 100</td>
<td>FRB</td>
</tr>
<tr>
<td>13</td>
<td>Y_3</td>
<td>4</td>
<td>Industrial production index: total index</td>
<td>Index 2007 = 100</td>
<td>FRB</td>
</tr>
<tr>
<td>14</td>
<td>Y_t</td>
<td>4</td>
<td>Industrial production index: products</td>
<td>Index 2007 = 100</td>
<td>FRB</td>
</tr>
<tr>
<td>15</td>
<td>C_2</td>
<td>5*</td>
<td>PCE excluding food and energy</td>
<td>Billions of dollars</td>
<td>BEA</td>
</tr>
<tr>
<td>16</td>
<td>C_5</td>
<td>5</td>
<td>Real PCE, quality indexes; nondurable goods</td>
<td>Index 2005 = 100</td>
<td>BEA</td>
</tr>
<tr>
<td>17</td>
<td>C_4</td>
<td>5</td>
<td>Real PCE, quality indexes; services</td>
<td>Index 2005 = 100</td>
<td>BEA</td>
</tr>
<tr>
<td>18</td>
<td>I_2</td>
<td>5</td>
<td>Real gross private domestic investment</td>
<td>Billions of Chained 2005</td>
<td>BEA</td>
</tr>
<tr>
<td>19</td>
<td>I_5</td>
<td>5*</td>
<td>Gross private domestic investment: fixed nonresidential</td>
<td>Billions of dollars</td>
<td>BEA</td>
</tr>
<tr>
<td>20</td>
<td>I_k</td>
<td>5</td>
<td>Manufacturers’ new orders: nondefense capital goods</td>
<td>Millions of dollars</td>
<td>DOC</td>
</tr>
<tr>
<td>21</td>
<td>p_2</td>
<td>8</td>
<td>Core CPI excluding food and energy</td>
<td>Index 2005 = 100</td>
<td>BEA</td>
</tr>
<tr>
<td>22</td>
<td>p_3</td>
<td>8</td>
<td>Price index - PCE excluding food and energy</td>
<td>Index 2005 = 100</td>
<td>BEA</td>
</tr>
<tr>
<td>23</td>
<td>p_4</td>
<td>8</td>
<td>Price index - PCE - Service</td>
<td>Index 2005 = 100</td>
<td>BEA</td>
</tr>
<tr>
<td>24</td>
<td>w_2</td>
<td>4*</td>
<td>Average hourly earnings: manufacturing</td>
<td>Dollars</td>
<td>BLS</td>
</tr>
<tr>
<td>25</td>
<td>w_3</td>
<td>4*</td>
<td>Average hourly earnings: construction</td>
<td>Dollars</td>
<td>BLS</td>
</tr>
<tr>
<td>26</td>
<td>w_4</td>
<td>4*</td>
<td>Average hourly earnings: service</td>
<td>Dollars</td>
<td>BLS</td>
</tr>
<tr>
<td>27</td>
<td>L_2</td>
<td>4</td>
<td>Civillian Labor Force: Employed Total</td>
<td>Thous.</td>
<td>BLS</td>
</tr>
<tr>
<td>28</td>
<td>L_1</td>
<td>4</td>
<td>Employees, nonfarm: total private</td>
<td>Thous.</td>
<td>BLS</td>
</tr>
<tr>
<td>29</td>
<td>L_4</td>
<td>4</td>
<td>Employees, nonfarm: goods-producing</td>
<td>Thous.</td>
<td>BLS</td>
</tr>
<tr>
<td>30</td>
<td>RE_2</td>
<td>6</td>
<td>Bond yield: Moody’s Baa industrial</td>
<td>% per annum</td>
<td>Bloomberg</td>
</tr>
<tr>
<td>31</td>
<td>RE_3</td>
<td>6</td>
<td>Bond yield: Moody’s A corporate</td>
<td>% per annum</td>
<td>Bloomberg</td>
</tr>
<tr>
<td>32</td>
<td>RE_4</td>
<td>6</td>
<td>Bond yield: Moody’s A industrial</td>
<td>% per annum</td>
<td>Bloomberg</td>
</tr>
<tr>
<td>33</td>
<td>Lev_F</td>
<td>9</td>
<td>Core capital leverage ratio PCA all insured institutions</td>
<td>Core capital/total asset</td>
<td>FDIC</td>
</tr>
<tr>
<td>34</td>
<td>Lev_E</td>
<td>7</td>
<td>Domestically chartered commercial banks leverage ratio</td>
<td>Total asset/net worth</td>
<td>FRB</td>
</tr>
<tr>
<td>35</td>
<td>Lev_3</td>
<td>7</td>
<td>Brokers and dealers leverage ratio</td>
<td>Total asset/net worth</td>
<td>FOP</td>
</tr>
<tr>
<td>36</td>
<td>Lev_F</td>
<td>3</td>
<td>Nonfarm nonfinancial non-corporate leverage ratio</td>
<td>Total asset/net worth</td>
<td>FOF</td>
</tr>
<tr>
<td>37</td>
<td>Lev_3</td>
<td>3</td>
<td>Nonfarm corporate leverage ratio</td>
<td>Total asset/net worth</td>
<td>FRB</td>
</tr>
<tr>
<td>38</td>
<td>s_2</td>
<td>1</td>
<td>Charge-off rate on all loans and leases all commercial banks</td>
<td>% per annum</td>
<td>FRB</td>
</tr>
<tr>
<td>39</td>
<td>s_3</td>
<td>1</td>
<td>Charge-off rate on all loans all commercial banks</td>
<td>% per annum</td>
<td>FRB</td>
</tr>
<tr>
<td>40</td>
<td>s_4</td>
<td>1</td>
<td>Charge-off rate on all loans banks 1st to 100th largest by assets</td>
<td>% per annum</td>
<td>FRB</td>
</tr>
</tbody>
</table>

Note: The format is: series number; transformation code; series description; unit of data and data source. The transformation codes are: 1 - demeaned; 2 - linear detrended; 3 - logarithm and demeaned; 4 - logarithm, linear detrend, and multiplied by 100; 5 - log per capita, linear detrended and multiplied by 100; 6 - detrended via HP filter; 7 - logarithm, detrended via HP filter, and multiplied by 100; 8 - first difference logarithm, detrended via HP filter, and multiplied by 400; 9- the reciprocal number, logarithm, detrended via HP filter, and multiplied 100. A * indicate a series that is deflated with the GDP deflator. “PCE” and “SW (2007)” in this table denote personal consumption expenditure and Smets and Wouters (2007), respectively.
Figure 4.1 Structural Shocks in Cases A and B
4.8. TABLES AND FIGURES

Figure 4.2 Structural Shocks in Cases C and D
Figure 4.3 Stochastic Volatilities in Cases C and D
Figure 4.4 Historical Decomposition: Output
Figure 4.5 Historical Decomposition: Investment
Figure 4.6 Historical Decomposition: Corporate Borrowing Rate
CHAPTER 4. SOURCES OF THE GREAT RECESSION

Figure 4.7 Historical Decomposition: Bank Leverage Ratio
4.9 Appendix

4.9.1 Sampling Stochastic Volatility with Leverage Effect

Step 4 in Section 4.4 employs the algorithm of Omori et al. (2007) which is the extension of Kim et al. (1998) toward a SV model with leverage effect. This subsection is based on Justiniano and Primiceri (2008) who employed Kim et al. (1998) for drawing the stochastic volatilities.

According to Omori et al. (2007), the key idea of MCMC algorithm of a SV model with leverage effect is to obtain a draw from an approximate linear and Gaussian state space form such as

\[
\begin{pmatrix}
\sigma^*_t \\
h_{t+1}
\end{pmatrix} = 
\begin{pmatrix}
\mu + \phi(h_t - \mu) \\
\mu
\end{pmatrix} + 
\begin{pmatrix}
z^*_t \\
v_t
\end{pmatrix},
\]

(4.44)

\[
\left\{ \begin{pmatrix}
z^*_t \\
v_t
\end{pmatrix} \right| d_{i,t}, u_{it} = k, \rho_i, \omega_i \right. = 
\begin{pmatrix}
m_k + v_k \zeta_t \\
d_t \rho \omega (a_k + b_kv_k \zeta_t) \exp(m_k/2) + \omega \sqrt{1 - \rho^2} \zeta^*_t
\end{pmatrix},
\]

(4.45)

\[
\begin{pmatrix}
\zeta_t \\
\zeta^*_t
\end{pmatrix} \sim i.i.d. N(0, I_2),
\]

where \(\sigma^*_i,t = \log \sigma_{i,t} = h_{i,t} + z^*_i,t, h_{it} = \log \sigma_{i,t},\) and \(z^*_i,t = \log (z_{i,t}^2)\). And \(d_{i,t}\), and \(\eta_{i,t}\) are denoted as

\[d_{i,t} = I(z_{i,t} \geq 0) - I(z_{i,t} < 0),\]
\[\eta_{i,t} = (h_{i,t} - \mu) - \phi(h_{i,t-1} - \mu),\]

where, \(I(\cdot)\) is an indicator function which indicates \(d_{i,t} = 1\) when \(z_{i,t} > 0\), or otherwise: \(d_{i,t} = -1\).

Suppose that the MCMC algorithm has implemented iteration \(g\), generating samples \(\Phi^{(g)}_i (= (\phi_i, \rho_i, \omega_i))\) and \(H^{T,(g)}\). In iteration \(g+1\), the following four steps are used to a set of new draws.

**Step 1: Draw the structural shocks \(\varepsilon^{(g+1)}_t\).**

In order to generate a new sample of stochastic volatilities, we need to obtain a new sample of structural shocks. This can be done using simulation smoother developed by de Jong and Shephard (1995).

**Step 2: Draw the stochastic volatilities \(H^{T,(g+1)}\) with leverage effect**

With a draw of shocks in hand, nonlinear measurement equations (4.2) in Section 4.2, which is represented as eq.(4.46) for each structural shock, can be easily converted in linear one such as (4.47) by squaring and taking logarithms of every elements. This induces the following approximating state space model (4.47) and (4.48).

\[
\varepsilon_{i,t} = \sigma_{i,t} z_{i,t}, \ i=1,2,\cdots, M,
\]

(4.46)
\[ \tilde{\varepsilon}_{i,t} = 2h_{i,t} + z_{i,t}^*, \quad (4.47) \]

\[ h_{i,t} = \mu + \phi(h_{i,t-1} - \mu) + \nu_{i,t}, \quad \nu_{i,t} \sim \text{i.i.d. } N(0, \omega_i^2) \quad (4.48) \]

where \( \tilde{\varepsilon}_{i,t} = \log((\varepsilon_{i,t}^2 + \bar{c})^2) \); \( \bar{c} \) is the offset constant (set to 0.001); \( h_{i,t} = \log(\sigma_{i,t}) \) and \( z_{i,t}^* = \log(z_{i,t}^2) \). M is the number of structural shocks. Since the squared shocks \( \varepsilon_{i,t}^2 \) is very small, an offset constant is used to make the estimation procedure more robust. Eqs. (4.47) and (4.48) are linear, but non-Gaussian state space form, because \( z_{i,t}^* \) are distributed as a log \( \chi^2(1) \). In order to transform the system in a Gaussian state space form, a mixture of normal approximation of the log \( \chi^2(1) \) distribution is used, as described in Kim et al. (1998) and Omori et al. (2007). A draw of \( z_{i,t}^* \) is implemented from the mixture normal distribution given as

\[ f(z_{i,t}^*) = \sum_{k=1}^{K} q_k f_N(z_{i,t}^* | u_{i,t} = k), \quad i = 1, \cdots, M, \quad (4.49) \]

where \( u_{i,t} \) is the indicator variable selecting which member of the mixture of normals has to be used at period \( t \) for shock \( i \). And \( q_k \) is the probability of \( u_{i,k} = k \); \( q_k = \Pr(u_{i,t} = k) \), and \( f_N(\cdot) \) denotes the probability density function of normal distribution. Omori et al (2007) select a mixture of ten normal densities \( (K = 10) \) with component probabilities \( q_k \), means \( m_k \), and variances \( v_k^2 \), for \( k = 1, 2, \cdots, 10 \), chosen to match a number of moment of the log \( \chi^2(1) \) distribution. The constant \( \{q_k, m_k, v_k^2\} \) are reported as Table blow.

<table>
<thead>
<tr>
<th>( K )</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k )</td>
<td>( q_k )</td>
</tr>
<tr>
<td>1</td>
<td>0.00609</td>
</tr>
<tr>
<td>2</td>
<td>0.04775</td>
</tr>
<tr>
<td>3</td>
<td>0.13057</td>
</tr>
<tr>
<td>4</td>
<td>0.20674</td>
</tr>
<tr>
<td>5</td>
<td>0.22715</td>
</tr>
<tr>
<td>6</td>
<td>0.18842</td>
</tr>
<tr>
<td>7</td>
<td>0.12047</td>
</tr>
<tr>
<td>8</td>
<td>0.05591</td>
</tr>
<tr>
<td>9</td>
<td>0.01575</td>
</tr>
<tr>
<td>10</td>
<td>0.00115</td>
</tr>
</tbody>
</table>

Using generator of the mixture normal distribution above, the system has an approximate linear and Gaussian state space form. Therefore, a new draw of the stochastic volatilities \( H_T^{(g+1)} \) can be obtained recursively with standard Gibbs sampler for state space form using the algorithm of Carter and Kohn (1994).

**Step 3:** Draw the indicators of the mixture approximation \( u_T^{(g+1)} \).
In the case of SV with leverage effect, we need to modify the indicator $u_{i,t}$ for the mixture normal described in Step 2, compared with Justiniano and Primiceri (2008). We follow the algorithm proposed by Omori et al. (2007), and obtain a new draw of indicators $u_{i,t}$ which is generated conditional on $e_{i,t}^{(g+1)}, H^{T,(g+1)}$ by independently sampling each from the discrete density defined by

$$
\pi(u_{i,t} = k \mid e_{i,t}, h_{it}, \Phi) \propto \pi(u_{i,t} = k \mid \sigma_{i,t}^*, d_{it}, h_{it}, \Phi) \propto \pi(u_{i,t} = k \mid z_{i,t}^*, \eta_{i,t}, d_{it}, \Phi)
$$

$\propto q_k v_k^{-1} \exp \left\{ -\frac{(z_{i,t}^* - m_k)^2}{2v_k^2} - \frac{[\eta_{i,t} - d_{it} \rho_i \omega_i \exp (m_k/2)\{a_k + b_k (z_{i,t}^* - m_k)\}]^2}{\omega_i^2 (1 - \rho_i^2)} \right\}$ (4.50)

**Step 4: Draw the coefficients $\Phi_i^{(g+1)} (= (\phi_i, \rho_i, \omega_i))$ of stochastic volatility processes.**

Having generated a sample $H^{T,(g+1)}$, we sample the elements of vector $\Phi_i^{(g+1)}$ from the density

$$p(\Phi_i \mid \sigma_{i,t}^*, d_{it}, u_{i,t} \Phi_i) \propto p(\sigma_{i,t}^* \mid d_{it}, u_{i,t} \Phi_i)p(\Phi_i).$$

The density $p(\sigma_{i,t}^* \mid d_{it}, u_{i,t} \Phi_i)$ is found from the output of Kalman filter recursion applied to the state space model (4.44) and (4.45). For the sampling we rely on the Metropolis-Hasting algorithm with a proposal density based on random walk such as

$$\theta^{(\text{proposal})} = \theta^{(g-1)} + u_t, \quad u_t \sim N(0, c \Sigma),$$

where $c$ is an adjustment constant.

### 4.9.2 Simulation Smoother

Step 3.1 of algorithm of data rich DSGE described in Section 4.4 employs simulation smoother (de Jong and Shephard, 1995) which generate sampling of model variables $S_t$ from conditional posterior distribution, $p(S_t \mid \Gamma^{(g-1)}, \theta, X_t).$ On the other hand, Boivin and Giannoni (2006), and Kryshko (2011) employ smoothing method proposed by Carter and Kohn (1994). But their method does not apply only to sample positive definite matrix as variance covariance matrix of state variables so that their method discontinue on the way of sampling in MCMC pointed out by Chib (2001, p.3614). As a result, Kryshko (2011) transforms to ad hoc variance covariance matrix of state variables. To avoid this problem, our algorithm employs simulation smoother instead of Carter and Kohn’s (1994) algorithm. Accordingly, our algorithm accomplishes generalization of estimating data rich DSGE model.

To simplify representation of algorithm of simulation smoother, we rewrite state space model of (4.17) and (4.18) described in Section 4.2 into (4.51), and (4.52) as below.

---

15Another simulation smoother has been invented by Durbin and Koopman (2002). The advantage of their method is to make code easily because of using existing Kalman smoother and not coding new algorithm, while simulation smoother of Carter and Kohn (1994) and de Jong and Shephard (1995) need to made new code of their algorithm. However, since our model is medium-size DSGE model and it requests long computing time for MCMC processing, we adopt more speeding algorithm of de Jong and Shephard (1995), instead of Durbin and Koopman (2002).
\( \tilde{X}_t = \tilde{\Lambda}\tilde{S}_t + \nu_t, \quad \nu_t \sim N(0, R), \) \hfill (4.51)

\( \tilde{S}_t = \tilde{G}\tilde{S}_{t-1} + \tilde{E}\varepsilon_t, \quad \varepsilon_t \sim N(0, Q(\theta)), \) \hfill (4.52)

The following four steps are conducted to generate a new draw of model variables.

**Step 1: Kalman filter for state space model is implemented.**

Kalman filter is represented as

\[
\eta_t = \tilde{X}_t - \tilde{\Lambda}\tilde{S}_{t|t}, \quad F_t = \tilde{\Lambda}\tilde{P}_{t|t}\tilde{\Lambda}' + R, \quad K_t = \tilde{G}\tilde{P}_{t|t}\tilde{\Lambda}'F_t^{-1},
\]

\[
L_t = \tilde{G} - K_t\tilde{\Lambda}, \quad \tilde{S}_{t+1|t+1} = \tilde{G}\tilde{S}_{t|t} + K_t\eta_t, \quad \tilde{P}_{t+1|t+1} = \tilde{G}\tilde{P}_{t|t}L_t + \tilde{E}Q(\theta)\tilde{E}',
\]

where \( \eta_t \) is forecasting errors, \( K_t \) is Kalman gain, \( \tilde{P}_t \) is variance covariance matrix of state variables \( \tilde{S}_t \). Filtering of \( \tilde{S}_{t|t}, \tilde{P}_{t|t} \) iterates forward for period \( t = 1, 2, \cdots, T \). And for initial value \( \tilde{S}_{1|1}, \tilde{P}_{1|1} \), we set \( \tilde{X}_1 = \tilde{\Lambda}\tilde{S}_1 \), and \( \tilde{P}_{1|1} = \tilde{G}\tilde{P}_{1|1}\tilde{G}' + \tilde{E}Q(\theta)\tilde{E}' \), where subscript \( t|t \) of \( \tilde{S}_t \) denotes conditional expected value of \( \tilde{S}_t \) up to information on \( X_1, \cdots, X_t \), thus, \( E(\tilde{S}_t|X_1, X_2, \cdots, X_t) \).

**Step 2: Generate values of \( r_{t-1}, N_{t-1} \) by implementing simulation smoother.**

This algorithm is iterated backward from period: \( t = T, \cdots, 2, 1 \) using values obtained from Kalman filter, as following equations (4.53), (4.54).

\[
r_{t-1} = \tilde{\Lambda}'F_t^{-1}\eta_t - W_tC_t^{-1}d_t + L_tr_t, \quad \hfill (4.53)
\]

\[
N_{t-1} = \tilde{\Lambda}'F_t^{-1}\tilde{\Lambda} + W_tC_t^{-1}W_t + L_t'N_tL_t, \quad \hfill (4.54)
\]

where \( W_t \) and \( C_t \) are obtained from the equations such as

\[
W_t = Q(\theta)\tilde{E}'N_tL_t,
\]

\[
C_t = Q(\theta) - Q(\theta)\tilde{E}'N_t\tilde{E}Q(\theta),
\]

and random variable \( d_t \) is generated from \( N(0, C_t) \). Initial value \( r_T \) and \( N_T \) are set at \( r_T = 0 \), and \( N_T = 0 \).
Step 3: Smoothing of structural shocks $\hat{\varepsilon}_{t|T}$, are implemented backward iteration using the equation (4.55).

Subscript $t|T$ of $\hat{\varepsilon}_{t|T}$ denotes expected value conditional on total sample period such as $E(\varepsilon_t|X_1, X_2, \ldots, X_T)$.

$$\hat{\varepsilon}_{t|T} = Q(\theta)\hat{E}' r_t + d_t \quad d_t \sim N(0, C_t), \quad t = T, \ldots, 2, 1$$  
(4.55)

Step 4: Generate model variables $\hat{S}_t$ by forward iteration of the equation (4.56).

$$\hat{S}_{t+1|T} = \hat{G}\hat{S}_t + \hat{E}\hat{\varepsilon}_{t|T}, \quad t = 1, 2, \ldots, T,$$

where initial value $\hat{S}_{1|T}$ is obtained from $\hat{S}_{1|T} = \hat{S}_{1|1} + \hat{P}_{1|1} r_0$.

The algorithm described above is procedure generating model variables $\hat{S}_t (t = 1, 2, \ldots, T)$ from conditional posterior distribution $p(S^T|\Gamma^{(g-1)}, \theta, X^T)$ which is implemented in Step 3.1 of Section 4.4.

### 4.9.3 Sampling Measurement Equation Parameters

In Step 3.2 of MCMC algorithm in Section 4.4, we sample parameters $\Gamma = \{\Lambda, R, \Psi\}$ of measurement equation obtained from (4.11) and (4.13). To do so, (4.11) is transformed by substituting (4.13) into it as

$$(I - \Psi L) X_t = (I - \Psi L) \Lambda S_t + \nu_t, \quad \nu_t \sim i.i.d. N(0, R),$$

where $I$ denotes identity matrix. The sampling of parameters $\Gamma = \{\Lambda, R, \Psi\}$ from conditional posterior distribution $p(\Gamma|S^{T(g)}, \theta^{(g-1)}, X^T)$ given the unobserved model variables $S^T$ and deep parameters $\theta$, is conducted following the approach by Chib and Greenberg (1994) who proposed Bayesian estimation method of linear regression model with AR (1) errors such like (4.11) and (4.13).

For estimating above model, Chib and Greenberg (1994) divided it into two linear regression models. First, by using notations, $X^*_t = X_{k,t} - \Psi_{kk}X_{k,t-1}$, and $S^*_t = S_{k,t} - \Psi_{kk}S_{k,t-1}$ where subscript $k$ is $k$-th indicator of data set $X_t$, above equation is represented as

$$X^*_t = \Lambda S^*_t + \nu_t, \quad \nu_t \sim i.i.d. N(0, R),$$

Second, by using notation $e_{k,t} = X_{k,t} - \Lambda_k S_t$ which means measurement errors, the equation is also rewritten as

$$e_k = \Psi_{kk} e_{k,-1} + \nu_k,$$

where $e_k = [e_{k,2}, \ldots, e_{k,T}]'$, $e_{k,-1} = [e_{k,1}, \ldots, e_{k,T-1}]'$. We sample parameter $(\Lambda, R)$ given parameter $\Psi$ from the first equation, and parameter $\Psi$ given $(\Lambda, R)$ from the second equation sequentially based on the following two-step algorithm.

**Step 1.** Sampling $(\Lambda_k, R_{kk})$ from conditional posterior distribution $p(\Lambda_{kk}, R_{kk} | \Psi_{kk}, S^T, \theta, X^T)$ for estimating equation

$$X^*_t = \Lambda S^*_t + \nu_t, \quad \nu_t \sim i.i.d. N(0, R).$$
The posterior density of \((\Lambda, R_{kk})\) given the unobserved state variables \(S^T\) and deep parameters \(\theta\) is represented as

\[
p(\Lambda_k, R_{kk} | \Psi_{kk}, S^T, X^T) \propto p(X^T | S^T, \Lambda_k, R_{kk}, \Psi_{kk}, \theta) p(\Lambda_k, R_{kk}),
\]

where \(p(X^T | S^T, \Lambda_k, R_{kk}, \Psi_{kk}, \theta)\) is likelihood function and \(p(\Lambda_k, R_{kk})\) is prior density.

As shown by Chib and Greenberg (1994), the above likelihood function is proportional to a Normal-Inverse-Gamma density as

\[
p(X^T | S^T, \Lambda_k, R_{kk}, \Psi_{kk}, \theta) \propto p_{\text{NIG}}(\Lambda_k, R_{kk} | \hat{\Lambda}_k, (S^{*'}S^{*})^{-1} S^{*'}X^*_k, s, T - N - 2)
\]

Since above prior \(p(\Lambda_k, R_{kk})\) is assumed to be Normal-Inverse-Gamma \(p_{\text{NIG}}(\Lambda_k, R_{kk} | \Lambda_{k,0}, M_{k,0}, s_0, \nu_0)\), the resulting conditional posterior density is also Normal-Inverse-Gamma as following.

\[
p(\Lambda_k, R_{kk} | \Psi_{kk}, S^T, X^T) \propto p_{\text{NIG}}(\Lambda_k, R_{kk} | \hat{\Lambda}_k, (S^{*'}S^{*})^{-1} S^{*'}X^*_k, s, T - N - 2)
\]

\[
\times p_{\text{NIG}}(\Lambda_k, R_{kk} | \Lambda_{k,0}, M_{k,0}, s_0, \nu_0)
\]

\[
\times p_{\text{NIG}}(\Lambda_k, R_{kk} | \bar{\Lambda}_k, \bar{M}_k, \bar{s}, \bar{\nu})
\]

where

\[
\bar{M}_k = M_{k,0} + (S^{*'}S^{*})^{-1} S^{*'}X^*_k
\]

\[
\bar{\Lambda}_k = \bar{M}_k^{-1} (M_{k,0} \Lambda_{k,0} + (S^{*'}S^{*}) \hat{\Lambda}_k)
\]

\[
\bar{s} = s_0 + s + (\lambda_{k,0} - \hat{\lambda}_k)' [M_{k,0}^{-1} + (S^{*'}S^{*})^{-1}]^{-1} (\lambda_{k,0} - \hat{\lambda}_k)
\]

\[
\bar{\nu} = \nu_0 + T
\]

and \(\Lambda_{k,0}, M_{k,0}, s_0, \) and \(\nu_0\) are parameters of the prior density.

We sample factor loading \(\Lambda_k\) and the variance of measurement error \(R_{kk}\) sequentially from

\[
R_{kk} | \Psi_{kk}, S^T, \theta, X^T \sim IG(\bar{s}, \bar{\nu})
\]

\[
\Lambda_k | R_{kk}, \Psi_{kk}, S^T, \theta, X^T \sim N(\bar{\Lambda}_k, R_{kk}, \bar{M}_k^{-1})
\]
Step 2. Sampling $\Psi_{kk}$ from conditional posterior distribution $p(\Psi_{kk}|\Lambda, R_{kk}, S^T, \theta, X^T)$ for estimating equation of measurement errors

$$e_k = \Psi_{kk} e_{k,-1} + \nu_k.$$ 

The conditional posterior density $p(\Psi_{kk}|\Lambda, R_{kk}, S^T, \theta, X^T)$ is given as

$$p(\Psi_{kk}|\Lambda, R_{kk}, S^T, \theta, X^T) \propto p(X^T_k|S^T, \Lambda, R_{kk}, \Psi_{kk}, \theta) p(\Psi_{kk}),$$

where $p(X^T_k|S^T, \Lambda, R_{kk}, \Psi_{kk}, \theta)$ is likelihood function and $p(\Psi_{kk})$ is prior density. Then, above likelihood function is proportional to the normal density such as

$$p(X^T_k|S^T, R_{kk}, \Psi_{kk}, \theta) \propto \exp \left[ -\frac{1}{2R_{kk}} (\Psi_{kk} - \hat{\Psi}_{kk})' e_{k,-1} e_{k,-1}' (\Psi_{kk} - \hat{\Psi}_{kk}) \right].$$

And above prior density of coefficient of AR (1) errors $\Psi_{kk}$ is also normal density but truncated at less than unity because dynamic of errors keep to be stationary. So, prior densty is assumed to be such as

$$p(\Psi_{kk}) \propto \exp \left[ -\frac{1}{2\sigma_{\Psi,0}^2} (\Psi_{kk} - \Psi_0)^2 \right] \times 1_{\{|\Psi_{kk}|<1\}},$$

where $1_{\{|\Psi_{kk}|<1\}}$ denotes indicator function which is unity if $|\Psi_{kk}| < 1$, otherwise zero.

The conditional posterior density is proportional to a product of above two normal densities, and represented as

$$p(\Psi_{kk}|R_{kk}, S^T, \theta, X^T) \propto \exp \left[ -\frac{1}{2R_{kk}} (\Psi_{kk} - \hat{\Psi}_{kk})' e_{k,-1} e_{k,-1}' (\Psi_{kk} - \hat{\Psi}_{kk}) \right] \times \exp \left[ -\frac{1}{2\sigma_{\Psi,0}^2} (\Psi_{kk} - \Psi_0)^2 \right] \times 1_{\{|\Psi_{kk}|<1\}}.$$ 

Hence, we sample coefficient of AR (1) errors $\Psi_{kk}$ from truncated normal such as

$$\Psi_{kk}|R_{kk}, S^T, \theta, X^T \sim N(\hat{\Psi}_{kk}, \bar{V}_{\Psi_{kk}}) \times 1_{\{|\Psi_{kk}|<1\}},$$

where

$$\bar{V}_{\Psi_{kk}} = \left[ (R_{kk}(e_{k,-1}' e_{k,-1})^{-1})^{-1} + (\sigma_{\Psi,0}^2)^{-1} \right]^{-1},$$

$$\hat{\Psi}_{kk} = \bar{V}_{\Psi_{kk}} \left[ (R_{kk}(e_{k,-1}' e_{k,-1})^{-1})^{-1} \hat{\Psi}_{kk} + (\sigma_{\Psi,0}^2)^{-1} \Psi_0 \right].$$

4.9.4 Remaining Framework of the DSGE model

This subsection describes the remaining structure of our DSGE model in Section 4.3.
Household Sector

There is a continuum of members in the household where the total population measures to one. Within the household, there are fractions of $f^E$ entrepreneurs, $f^F$ financial intermediaries (or “bankers”), and $1 - f^E - f^F$ workers. Entrepreneurs engage in a business where they produce intermediate goods and transfer the net worth back to the household when they exit from the business. Now, each financial intermediary manages a bank where it accepts the deposits from the household sector and lends to entrepreneurs. When financial intermediaries exit from their business, they also transfer their net worth back to the household sector. Finally, remaining fraction of the members of the household become workers. Workers supply labor input to earn wage and they transfer their wage earnings to the household each period. Within the household, each member shares the risk perfectly.

The representative household maximizes her expected discounted sum of utility over time and their objective function is specified as follow;

$$E_t \sum_{i=0}^{\infty} \beta^i \lambda_{t+i} \left[ \frac{(c_{t+i} - hC_{t+i-1})^{1-\sigma^c}}{1-\sigma^c} - \lambda_{t+i}^{L} \frac{(l_{t+i})^{1+\sigma^L}}{1+\sigma^L} \right]$$  \hspace{1cm} (4.57)

where $\beta$ is the discount rate, $h$ is the habit persistence, $\sigma^c$ is the inverse of intertemporal elasticity of substitution, $c_t$ is final goods consumption, $C_{t-1}$ represents the external habit formation, $\sigma^L$ is the inverse of Frisch labor supply elasticity and $l_t$ is the supply of aggregate labor by workers. Now, there are two structural shocks embedded in the function. $\lambda_t^L$ represents an intertemporal preference shock, while $\lambda_t^H$ represents labor disutility shock relative to consumption.

Next, turning to the budget constraint of the representative household, they make a deposit, $b_t$, at period $t$ and earn real interest rate, $R_t/\pi_{t+1}$, next period where $R_t$ is risk-free gross nominal interest rate at period $t$ and $\pi_{t+1}$ is gross inflation rate at period $t+1$. In addition, the household pays lump sum tax of $\tau_t$ to the government. Now, they receive a lump-sum transfer of wage incomes from workers which is expressed as $\int_0^1 w_t(x)l_t(x)dx$, where $w_t(x)$ and $l_t(x)$ are real wage and labor supply by individual worker $x$, respectively. Finally, the household earns the combined dividend of $\Xi_t^{div}$ from retailers, earns the net transfer of $\Xi_t^E$ from entrepreneurs, and the net transfer of $\Xi_t^F$ from bankers each period. Thus, the representative household’s budget constraint at period $t$ can be expressed as, in real terms, as follow,

$$c_t + b_t = \frac{R_{t-1}}{\pi_t} b_{t-1} - \tau_t + \Xi_t^{div} + \Xi_t^E + \Xi_t^F. \hspace{1cm} (4.58)$$

Consumption and Deposit Decision The first-order conditions (hereafter, FOCs) of the household with respect to $c_t$ and $b_t$ as follows;

$$\zeta_t^H = \chi_t^c (c_t - hC_{t-1})^{-\sigma^c} \hspace{1cm} (4.59)$$

$$\zeta_t^H = \beta E_t \zeta_{t+1}^H \frac{R_t}{\pi_{t+1}}. \hspace{1cm} (4.60)$$

where $\zeta_t^H$ is Lagrangian multiplier associated with the budget constraint. (4.59) is the FOC of consumption which equates the marginal utility of consumption to the shadow price of the final goods. (4.60) is the FOC of deposit decision.

---

16Here, the real wage set by worker $x$ is defined as $w_t(x) \equiv W_t(x)/P_t$, where $W_t(x)$ stands for the nominal wage set by worker $x$ and $P_t$ stands for the price index of final goods. The formulation of $W_t(x)$ and $P_t$ will be described later in this section.
4.9. APPENDIX

Wage Setting Decision by Workers Following Erceg, et al. (2000), each worker indexed by \( x \in [0,1] \) supplies differentiated labor input, \( l_t(x) \), monopsonistically and sells this service to the labor union who is perfectly competitive.\(^{17}\) Each worker sets his nominal wage according to Calvo style sticky price setting where fraction \( \theta^w \) of the entire workers cannot freely adjust the wages at their discretion. For fraction \( \theta^w \) of workers, the partial indexation of the nominal wage is assumed.\(^{18}\) Due to the perfect risk-sharing assumed in the model, each worker maximizes the objective function \( (4.57) \) by choosing the amount of individual labor supply, \( l_t(x) \), while taking the amount of consumption, \( c_t \), as given. Under this setting, \( (1 - \theta^w) \) fraction of workers maximize their objective function by setting the nominal wage, \( \tilde{W}_t \), such that

\[
E_t \sum_{i=0}^{\infty} \beta^i (\theta^w)^i \left[ \frac{\tilde{W}_t}{P_{t+i}} \left( \frac{P_{t+1+i}}{P_{t+i}} \right)^{\theta^w} \chi_{t+i}^c \left( c_{t+i} - h c_{t+i-1} \right)^{1-\sigma^c} - (1 + \psi^w) \chi_{t+i}^{L} \left( l_{t+i}(x) \right)^{\sigma^L} \right] l_{t+i}(x) = 0. \tag{4.61}
\]

The law of motion of the aggregate wage index can be shown to be as follow,

\[
W_t^{-1/\psi^w} = \theta^w \left[ W_{t-1} \left( \frac{W_{t-1}}{W_{t-2}} \right)^{\psi^w} \right]^{-1/\psi^w} + (1 - \theta^w) \tilde{W}_{t-1}^{-1/\psi^w}. \tag{4.62}
\]

Finally, the real wage index in the economy is defined as \( w_t \equiv W_t/P_t \).

Capital Production Sector

Capital producers are identical, perfectly competitive, and risk neutral. They purchase \( l_t^k \) units of final goods from the retailer, convert them to \( i_t^k \) units of capital goods, and combine them with existing capital stock, \( (1 - \delta) k_t \), to produce new capital stock, \( k_{t+1} \). Capital producers will, then, sell off new capital stock to entrepreneurs in a perfectly competitive manner. Capital producers have linear production technology in converting final goods to capital goods. In addition, they will incur quadratic investment adjustment cost when they change the production capacity of capital goods from previous period. Each capital producer maximizes the expected discounted cash flow with respect to \( i_t^k \).\(^{19}\) The FOC is given by

\[
q_t = \frac{1}{A_t^k} \left[ 1 + \kappa \left( \frac{i_t^k}{i_{t-1}^k} - 1 \right) + \frac{\kappa}{2} \left( \frac{i_t^k}{i_{t-1}^k} - 1 \right)^2 \right] - \beta \kappa \left( \frac{i_{t+1}^k}{i_t^k} - 1 \right)^2. \tag{4.63}
\]

\(^{17}\)The labor union transforms labor services to an aggregate labor input, \( l_t \), using the Dixit and Stiglitz type aggregator function. The factor demand function for \( l_t(x) \) is given by \( l_t(x) = (W_t(x)/W_t)^{-1 + \psi^w} l_t \), where \( \psi^w \) is the wage markup, \( W_t(x) \) is the nominal wage set by worker \( x \) and \( W_t \) is the aggregate nominal wage index which is given as \( W_t = \left[ \int_0^1 W_t(x)^{-1/\psi^w} dx \right]^{-\psi^w} \).

\(^{18}\)The lagged inflation indexation is specified as \( W_t(x) = (P_{t-1}/P_{t-2})^{\psi^w} W_{t-1}(x) \) where \( \psi^w \) controls the degree of nominal wage indexation to past inflation rate.

\(^{19}\)The profit function for each capital producer at period \( t \) can be expressed as follows,

\[
E_t \sum_{i=0}^{\infty} \beta^i \left\{ q_{t+i} i_{t+i}^k - \frac{1}{A_{t+i}^k} \left[ i_{t+i}^k + \frac{\kappa}{2} \left( \frac{i_{t+i+1}^k}{i_{t+i}^k} - 1 \right)^2 \right] \right\}
\]
where $A_t^k$ is the investment-specific technology shock common across all capital producers and $\kappa$ is the investment adjustment cost parameter. Finally, aggregate capital accumulation equation is given by

$$k_{t+1} = i_t^k + (1 - \delta) k_t. \quad (4.64)$$

**Retailing Sector**

Retailers $z \in [0, 1]$ purchase intermediate goods from the entrepreneur at perfectly competitive price and resell them monopolistically in the retail market.\(^{20}\) We assume Calvo type sticky price setting for the retailer where, for any given period $t$, fraction $\theta^p$ of the entire retailers cannot freely revise their prices. Further, $\theta^p$ fraction of the retailers who did not receive a ‘signal of price change’ will partially index their nominal prices to lagged inflation rate of price index.\(^{21}\) Under this setting, for $(1 - \theta^p)$ fraction of the retailers who received a ‘price changing signal’ at period $t$, they maximize their expected discounted sum of profits by setting the nominal price, $\tilde{p}_t$, such that

$$E_t \sum_{i=0}^{\infty} \beta^i (\theta^p)^i \left[ \frac{\hat{p}_t}{P_{t+i}} \left( \frac{P_{t-1+i}}{P_{t-1}} \right)^{\theta^p} - \left( \frac{\epsilon}{\epsilon - 1} \right) \tilde{p}_{t+i}^{mc} \right] y_{t+i}(z) = 0. \quad (4.65)$$

From the definition of aggregate price index, the law of motion of $P_t$ can be shown to be as follow,

$$(P_t)^{1-\epsilon} = \theta^p \left[ \frac{P_{t-1}}{P_{t-2}} \right]^{\theta^p} + (1 - \theta^p) \tilde{p}_t^{1-\epsilon}. \quad (4.66)$$

**The Rest of the Economy**

In closing the model, we describe the rest of the model structure here. The central bank is assumed to follow a standard Taylor-type monetary policy rule,

$$\hat{R}_t = \rho^R \hat{R}_{t-1} + (1 - \rho^R) \left[ \mu^\pi \hat{\pi}_t + \mu^\gamma \hat{Y}_t \right] + \epsilon^R_t \quad (4.67)$$

where $\rho^R$ controls the magnitude of interest smoothing, $\mu^\pi$ is Taylor coefficient in response to inflation gap, $\mu^\gamma$ is Taylor coefficient in response to output gap, and $\epsilon^R_t$ is i.i.d. monetary policy shock.

The government budget constraint is simply specified as

$$g_t = \tau_t. \quad (4.68)$$

The government expenditure, $g_t$, is financed solely by lump-sum tax, $\tau_t$. In our model, we simply assume that the government expenditure to follow stochastic AR(1) process.

Finally, the market clearing condition for final goods is given as follow,

$$Y_t = c_t + i_t^k + g_t. \quad (4.69)$$

---

\(^{20}\)The demand function for retail goods sold by retailer $z$ is given by $y_t(z) = (P_t(z)/P_t)^{-\epsilon} Y_t$, where $Y_t$ is aggregated final goods, $p_t(z)$ is nominal price of retail goods $y_t(z)$, $P_t$ is aggregate price index of final goods, and $\epsilon$ is the price elasticity of retail goods. Specifically, aggregated final goods, $Y_t$, and the aggregate price index, $P_t$, are given as follows: $Y_t \equiv \left[ \int_0^1 y_t(z)^{\epsilon/(\epsilon-1)} dz \right]^{1/\epsilon}$ and $P_t \equiv \left[ \int_0^1 P_t(z)^{\epsilon/(\epsilon-1)} dz \right]^{1/\epsilon}$.

\(^{21}\)The lagged inflation indexation is specified as $p_t(z) = (P_{t-1}/P_{t-2})^{\theta^p} p_{t-1}(z)$ where $\theta^p$ controls for the magnitude of price indexation to past inflation rate.
Structural Shocks in the Model

There are eight structural shocks in the model, each of them having a specific economic interpretation as below. Except for monetary policy shock, all of the structural shocks are assumed to follow AR(1) stochastic processes where $\rho$ is for the AR(1) coefficients for respective structural shocks.

- **TFP shock**: $\hat{A}_t = \rho^A \hat{A}_{t-1} + \varepsilon^A_t$
- **Preference shock**: $\hat{\chi}_t = \rho^c \hat{\chi}_{t-1} + \varepsilon^c_t$
- **Labor supply shock**: $\hat{\chi}_t^L = \rho^L \hat{\chi}_{t-1}^L + \varepsilon^L_t$
- **Investment specific technology shock**: $\hat{A}_t^K = \rho^K \hat{A}_{t-1}^K + \varepsilon^K_t$
- **Government spending shock**: $\hat{g}_t = \rho^G \hat{g}_{t-1} + \varepsilon^G_t$
- **Monetary policy shock**: $\varepsilon^R_t$
- **Corporate net worth shock**: $\hat{\gamma}_t^E = \rho^E \hat{\gamma}_{t-1}^E + \varepsilon^E_t$
- **Bank net worth shock**: $\hat{\gamma}_t^F = \rho^F \hat{\gamma}_{t-1}^F + \varepsilon^F_t$

Notice that each stochastic disturbance $\varepsilon_t$ including monetary policy shock is assumed to follow time varying volatility using SV model as mentioned in Section 4.2.
Chapter 5

Impacts of Government Spending on Unemployment

5.1 Introduction

Chapter 5 examines the quantitative effect of government spending on unemployment in Japan. The question is quite simple: Does government spending improve unemployment, if so, how big is it?

Note that Chapter 5 revised the paper based on “Impacts of Government Spending on Unemployment: Evidence from a Medium-scale DSGE Model,” (joint with Hasumi, R.), the discussion paper of Economic and Social Research Institute (ESRI Discussion paper series 329, 2016).

5.1.1 Background

This study investigates qualitative and quantitative effects of government spending for unemployment, based on an estimated medium-scale new Keynesian model.

Figure 5.1 shows unemployment rate from 1980Q2 to 2012Q4 in Japan. From the 1980s to the beginning of the 1990s, the unemployment rate stabilized at a low level (the lowest is 2.1%). After the burst of the so-called bubble economy, however, the unemployment rate increased to more than 5% (5.4% at peak), even though Japan was following expansionary fiscal policy as well as monetary easing policy.

After the so-called financial crisis of 2007-2008, the increase of unemployment rate in the U.S. induces a number of studies to (i) investigate whether an increase of government spending improves unemployment or not, and (ii) examine how large fiscal stimuli influences unemployment, extending a standard DSGE model (e.g. SW, 2007).

So, does government spending improve unemployment? Surprisingly, contradictory results are reported against simple questions about whether government spending improves unemployment and if so how much it improves.

Monacelli et al. (2010) confirmed fiscal stimulus improves unemployment by estimating a structural VAR model, and reproduced the result by constructing a search type DSGE model with friction in the labor market (see also Mayer et al. 2010, Campolmi et al. 2011). Conversely, Bruckner and Pappa (2012) showed the opposite result that government spending worsens unemployment by estimating a structural VAR model (see also Yuan and Li, 2000).

A standard story in considering the effect of government spending on unemployment is as follows
(Linnemann and Schabert, 2003): Government spending creates aggregate demand, but households refrain from consuming because they anticipate future tax increases (negative wealth effect). Suppose a situation where real wage adjustment is sluggish. Government expenditure raises labor demand thanks to aggregate demand creation (which is partly offset by the negative wealth effect), while increasing labor supply due to consumption and labor substitution. The effect of the expansionary fiscal policy on unemployment depends on the difference between the increase in labor demand and the increase in labor supply. Government expenditure improves unemployment if the increase in labor demand dominates the increase in labor supply, and deteriorates if dominated.

There are three points to think about the effect of government expenditure on unemployment. The first is how much government spending creates aggregate demand, that is, the size of the fiscal multiplier. Secondly, what is the channel for which government spending creates aggregate demand? Lastly, what kind of unemployment mechanism should be introduced into the general equilibrium framework.

First, the empirical results of Japan and the U.S. show that the fiscal multiplier has declined in recent years. Bilbiie et al. (2008) demonstrated that the effects of fiscal policy are declining in the U.S. since the 1980s, by estimating the DSGE model that introduced the Rule-of-Thumb household (hereinafter referred to as RoT household). For the reason, they pointed out that the decline in RoT households, monetary policy becoming more aggressive (Greenspan era), and the increase in government debt outstanding.

These three factors that bring down the fiscal multiplier are closely related to one another: Government procures funds by the direct underwriting of government bonds by the central bank (aggressive monetary easing), government expends while accumulating debt outstanding (increase in debt outstanding by the non-Ricardian policy), households recognize they will be drastically taxed in the future (decline in the RoT households), and the negative wealth effect is generated so as to offset the demand creation by the expansionary fiscal policy. That’s why the fiscal multiplier declined.

Note that Bilbiie et al. (2008) does not include the U.S. zero interest rate policy era in the estimation period. Christiano et al. (2011a) pointed out, in the zero interest rate policy era without crowding out effect, the fiscal multiplier will increase in the U.S.

Meanwhile, among the three factors that lead to a decline in the fiscal multiplier, the increase in government debt outstanding is a serious issue in Japan. According to the VAR analysis, fiscal policy after the 1980s was the non-Ricardian policy (Miyazaki, 2010, Ito et al. 2011, Ko and Morita, 2011). Furthermore, with the huge amount of government debt outstanding, the fiscal multiplier has also declined since the middle of the 1990s including the zero interest rate policy era (Ko and Morita, 2015).

Regarding channels where government spending creates aggregate demand, recently non-wasteful government expenditure channels are paying attention.

The demand creation channel that government expenditure induces consumption of RoT households (Coenen and Straub 2005, Gali et al. 2007) is a wasteful expenditure in which government consumption does not directly affect welfare. As mentioned above, the effect of demand creation through this channel has been declining in recent years due to the decline in the proportion of RoT households.

On the other hand, empirical results have been reported in which private demand is stimulated by a positive government consumption shock (Blanchard and Perotti 2002). As a mechanism behind this, a new channel has been proposed that government consumption directly increases the marginal utility of household consumption, and households have an incentive to smooth their consumption
5.1. INTRODUCTION

intertemporally, thus increasing their own consumption (the Edgeworth complementarities). The DSGE model has been constructed that takes into account the new channel, and confirmed the Edgeworth complementarities (Bouakez and Rebei 2007, Iwata 2013, Feve et al. 2013).

In addition, another channel has been proposed in which production is directly increased by public capital accumulation through government investment (Linnemann and Schabert 2006). This is a contribution to the supply side rather than demand creation by government expenditure. However, there are channels that indirectly create aggregate demand: Public capital accumulation increases productivity, deflationary pressures arise, monetary easing is triggered, which is the path that creates aggregate demand.

Finally, as a method to introduce unemployment in the DSGE framework, two major approaches are currently proposed. One approach is to explicitly introduce friction in the labor market and estimate the search type DSGE model (Gertler et al. 2008, Christiano et al. 2011b). The other approach is to introduce unemployment by high real wages based on workers’ market power (Gali 2011, Gali et al. 2012, hereinafter GSW, Casares et al. 2014).

The search type DSGE model can consider an additional channel that government expenditure increases aggregate demand, which lowers search costs, increases labor participation rates and worsen the unemployment rate. However, modeling of the richer labor market dramatically increases the number of structural equations and parameters, so there are difficulties that non-identifiable parameters increase and we have no choice but to rely on calibration.

On the other hand, in the case of introducing unemployment by GSW, there is no channel of endogenous increase in labor participation rate due to government expenditure, but there is a channel of endogenous rise of labor supply through the substitution between consumption and labor. Also, since modeling of labor market is simple, parameter identification problem does not arise in estimation, so it is possible to quantitatively evaluate the effect of government spending on unemployment.

5.1.2 Purposes, Originalities, and Methodologies

This study quantitatively verifies whether Japanese fiscal stimulus improves unemployment and how much improvement is made, by the empirical DSGE approach.

First, as a channel for fiscal stimulus to create demand, we focus on non-wasteful government spending that government expenditure directly affects household utility and firm’s productivity. Specifically, we assume two channels: Complementarities between government consumption and private consumption, and productivity effect of public capital accumulation. By introducing the non-wasteful government expenditure channel that directly affects utility and production, it is also possible to welfare comparisons among various government expenditure rules.

Second, following Corsetti et al. (2012), we consider the so-called spending reversal rule which cuts government expenditure with the increase in debt outstanding. By estimating the elasticity of spending cuts on the increase in debt outstanding, we will examine the non-Ricardian policy in the general equilibrium framework and consider the impact on the fiscal multiplier.

Third, as a way to introduce unemployment into the general equilibrium framework, we adopt GSW type unemployment based on workers’ market power. This research focuses on quantitatively grasping the effect of fiscal stimulus on unemployment rather than clarifying the mechanism of unemployment. Introducing unemployment with the search DSGE model is difficult to estimate due to parameter identification problem since labor market is modeled too rich. With GSW type unemployment, we can use labor market modeling in the standard DSGE model, so we can easily estimate key parameters and evaluate the quantitative effect of government expenditure on unemployment.
From the above, Chapter 5 aims to demonstrate the impact of government spending on unemployment to the Japanese economy, by introducing unemployment to the general equilibrium framework in a simple way of GSW, considering the influence of the increase in debt outstanding on government expenditure, incorporating channels such as the Edgeworth complementarities and the productivity effect of public capital.

5.1.3 My Contributions

Chapter 5 revised Matsumae and Hasumi (2016).

My main contributions are as follows: I expanded the SW model by introducing GSW type unemployment and non-wasteful government expenditure. In addition, I interpreted structurally about the estimation result of the Edgeworth complementarities, the productivity effect of public capital, and the contribution of government expenditure shock to unemployment fluctuations.

5.1.4 Organization of Chapter 5

Chapter 5 is organized as follows. Section 2 presents our model. Section 3 explains our estimation method and Section 4 reports estimation results. Section 5 concludes this paper.

5.2 Model

We incorporate unemployment and non-wasteful government spending into the standard medium-scale DSGE model (e.g. SW 2007). This section focuses on explaining how to introduce unemployment and non-wasteful government spending and on illustrating how non-wasteful fiscal expansions may affect unemployment. The entire model is described in Appendix.

5.2.1 Unemployment

Following the GSW framework, we consider a large household with a continuum of members represented by the unit square and indexed by a pair \((j, h) \in [0, 1] \times [0, 1]\). The first dimension indexed by \(j \in [0, 1]\) represents a differentiated skill in which a given household member is specialized. The second dimension indexed by \(h \in [0, 1]\) indicates member’s labor disutility. We can think intuitively of the first dimension as labor unions and the second dimension as members within each union.

Unions have market power due to their differentiated skills indexed by \(j \in [0, 1]\), but they are assumed to face nominal wage rigidities a la Calvo in line with Erceg et al. (2000). Therefore, unions set their nominal wages, taking the nominal stickiness into consideration. It should be noted that setting wages simultaneously determines employment from labor demand for union \(j\).

Members within each union have different labor disutilities with uniformly distributed as \(h \in [0, 1]\). Given the nominal wage determined by each union, members decide to work or not taking their labor disutilities into consideration. In addition, we assume the full risk sharing of consumption across members: Members can enjoy consuming with the same level.

Then, the preference of a member \(h\) (who has a disutility \(h\)) in any union \(j\) at period \(t\) can be written by

\[
\zeta_t \ln \left( \tilde{C}_{j,t} - \theta \tilde{C}_{t-1} \right) - 1_t(j, h) c_t^h x_t^h A_H h^{\sigma_h} \tag{5.1}
\]

\(\tilde{C}_{j,t}\) stands for consumption of member \(h\) in union \(j\) and \(\tilde{C}_t \equiv \int_0^1 \tilde{C}_{j,t} dj\) stands for aggregate consumption. The term \(\theta \tilde{C}_{t-1}\) indicates (external) habits on consumption and the parameter \(\theta \in\)
5.2. MODEL

(0, 1) depicts the importance of the habit formation. \(1_t(j, h)\) is the indicator function, which takes a value equal to one if the member \(h\) is employed at period \(t\), and zero otherwise. It is worth noting that the indicator function means members decide to work with fixed hours (normalized as unity) or not. \(\chi^h_t\) stands the endogenous preference shifter defined as the following equation:

\[
\chi^h_t \equiv \frac{Z_{\chi,t}}{C_t - \theta C_{t-1}},
\]

(5.2)

\[
Z_{\chi,t} = Z_{\chi,t}^{1-v} \left(C_t - \theta C_{t-1}\right)^v,
\]

(5.3)

This preference specification leads marginal labor disutility decreases during (aggregate) consumption booms. Two structural shocks are embedded: \(\zeta^c_t\) is the preference shock and \(\zeta^h_t\) is the labor supply shock. \(A_H\) is the scale parameter and \(\sigma_h\) is the inverse Frisch elasticity.

Let \(H_{j,t}\) be defined as employment of union \(j\). Then, aggregating the member’s utility regarding \(h\), we derive utility of union \(j\) at period \(t\) as follows:

\[
\zeta^c_t \ln \left(\hat{C}_{j,t} - \theta \hat{C}_{t-1}\right) - \zeta^h_t \chi^h_t A_H \int_0^{H_{j,t}} h^{1+\sigma_h} dh
\]

\[
= \zeta^c_t \ln \left(\hat{C}_{j,t} - \theta \hat{C}_{t-1}\right) - \zeta^h_t \chi^h_t A_H \frac{H_{j,t}^{1+\sigma_h}}{1 + \sigma_h}
\]

(5.4)

Thus, the preference of union \(j\) falls into the standard functional form.

Now, we explain how to introduce unemployment into the medium-scale DSGE model. Because of full risk sharing of consumption, the marginal utility of consumption becomes common across members. Let the marginal utility of consumption denoted by \(\varphi^c_t\). Given the (real) wage \(w_{j,t}\), the member \(h\) is willing to work as long as the real wage is greater than the marginal rate of substitution (MRS) between labor supply and consumption:

\[
(1 - \tau^h_t) w_{j,t} \geq \frac{\zeta^h_t \chi^h_t A_H h^{\sigma_H}}{\varphi^c_t}
\]

(5.5)

\(\tau^h_t\) stands for labor income tax rate. The left-hand side (LHS) indicates the marginal benefit of labor supply after tax adjustment. The right-hand side (RHS) is MRS which corresponds to reservation wage of member \(h\). Letting the marginal supplier of union \(j\)’s member be denoted by \(L_{j,t}\), we have:

\[
(1 - \tau^h_t) w_{j,t} = \frac{\zeta^h_t \chi^h_t A_H L_{j,t}^{\sigma_H}}{\varphi^c_t}
\]

(5.6)

As mentioned before, union \(j\) decides nominal wage \(W_{j,t}\) which simultaneously determines employment \(H_{j,t}\) from the labor demand for union \(j\). Let aggregate employment denoted by \(H_t\) determined by unions and let aggregate labor supply denoted by \(L_t\) determined by members. Then, unemployment rate \(U_t\) is defined as the following equation:

\[
U_t \equiv \frac{L_t - H_t}{L_t}
\]

(5.7)

Figure 5.2 (a) illustrates the occurrence of unemployment in our model where three curves are depicted: Labor demand curve, marginal revenue (MR) curve and MRS curve (labor supply curve). Consider three members \(\{a, b, c\}\) with different reservation wages (equivalently, different labor disutilities). Member \(a\) has the lowest reservation wage (the lowest labor disutility), member \(b\) has the medium one and member \(c\) has the highest one.
In a steady state, unions should set wage at $A$ so as to match MR curve with MRS curve and aggregate employment $H$ is simultaneously determined. Given the wage at $A$, members $a$ and $b$ are willing to work since the wage determined by the union is higher than their reservation wages. Meanwhile, member $c$ enjoys leisure since the wage determined by the union is lower than her reservation wage. In other words, member $c$ is *voluntarily* unemployed. Thus, given the wage at $A$, aggregate labor supply is determined at $L$.

The difference between $L$ (aggregate labor supply) and $H$ (aggregate employment) corresponds to *involuntary* unemployment. In Figure 5.2 (a), for instance, the reservation wage of member $b$ is lower than the wage at $A$ but the member $b$ is not employed.

### 5.2.2 Non-wasteful Government Spending

**Edgeworth complementarities**

Government consumption is assumed to directly affect household’s utility as the following way:

$$\tilde{C}_{j,t} = C_{j,t} + \nu_g G^c_t$$  \hspace{1cm} (5.8)

$\tilde{C}_{j,t}$ consists of private consumption $C_{j,t}$ and government consumption $G^c_t$. The parameter $\nu_g$ governs qualitative and quantitative influence of government consumption for private consumption. Equation (5.8) indicates household gains utility not from private consumption $C_{j,t}$ but from the above composite consumption $\tilde{C}_{j,t}$. Thus, household wants to smooth intertemporally the composite consumption from the Euler equation.

Suppose that government increases consumption. If the parameter $\nu_g$ is negative, an increase in government consumption $G^c_t$ leads to a decrease in the composite consumption $\tilde{C}_{j,t}$. Then, households will increase private consumption $C_{j,t}$ to keep the composite consumption constant along with the intertemporal consumption smoothing condition (Strictly speaking, an increase in government consumption raises the marginal utility of private consumption at the present period). Thus, a negative $\nu_g$ causes a cyclical comovement between private consumption and government consumption, which is the so-called *Edgeworth complementarities* (hereafter, EC).\(^1\) If $\nu_g$ is positive, a counter-cyclical comovement is shown, which implies government consumption is substitutes to private consumption. If $\nu_g$ is zero, there is no comovement, thus government consumption is independent of private consumption.

Examples on the EC are government spending to Medicare and education service. Fiorito and Kollintzas (2004) empirically investigate the complementarities between government consumption and private consumption in the Euro area, and they find that the government spending to the merit goods such as Medicare and education service becomes a complement to private consumption.\(^2\)

\(^1\)On the functional specification of the composite consumption, $\tilde{C}_t$, we can consider more general functional form: $\tilde{C}_t = \left[\phi C^c_t + (1 - \phi) (G^c_t)^{\theta c} \right]^{1/\theta_c}$. If $\theta_c \to 1$, then $\tilde{C}_t \to \phi C^c_t + (1 - \phi) G^c_t$ (linear function). On the other hand, if $\theta_c \to 0$, then $\tilde{C}_t \to C^c_t (G^c_t)^{1-\phi}$ (Cobb-Douglas type function). In specifying $\tilde{C}_t$ as the CES aggregator, however, we face a difficulty to identify two structural parameters, i.e. $\theta_c$ and $\phi$. In fact, Coenen et al. (2013) specifies $\tilde{C}_t$ as the CES aggregator, but they calibrate the private consumption share, $\phi_c$, in CES aggregator due to the difficulty to identify it. Following Ni (1995), Iwata (2013) and Feve et al. (2013), we specify the bundled consumption, $\tilde{C}_t$, as the linear function since we can easily recognize whether the government consumption is complements or substitutes to the private consumption from the sign of the parameter, $\nu_g$.

\(^2\)They also find government spending for general public goods (i.e. national defense, public security service, etc.) is not a compliment to private consumption. See also Sakai et al. (2015).
Productive public capital

Public capital accumulated by government investment is assumed to improve the productivity of private firms. An intermediate good firm \( j \) produces a differentiated good \( Y_{j,t} \) (\( j \in [0,1] \)), using capital \( \tilde{K}_{j,t} \) and labor input \( H_{j,t} \):

\[
Y_{j,t} = \epsilon_t \tilde{K}_{j,t}^\alpha (z_t H_{j,t})^{1-\alpha} - z_t^+ \Theta
\]  

(5.9)

\( \epsilon_t \) stands for neutral technology shock, \( z_t \) stands for labor-augmented technology, \( z_t^+ \) stands for a scaling variable, \( \alpha \) is capital income share and \( \Theta \) is fixed cost.

There are two types of capital in this economy. One is the effective private capital \( u_{j,t} K_{j,t} - 1 \) where \( u_{j,t} \) is the capital utilization rate. The other is the public capital \( K^g_t \) accumulated by the government.

\[
\tilde{K}_{j,t-1} = (K_{t-1}^g)^{\alpha_g} (u_t \tilde{K}_{j,t-1})^{1-\alpha_g}
\]  

(5.10)

\( \alpha_g \) is the marginal productivity of public capital for the bundled capital, \( \tilde{K}_{j,t-1} \). It should be noted that the productivity of the public capital for output can be expressed as \( \alpha \times \alpha_g \) from (5.9) and (5.10).

If \( \alpha_g \) is positive, then public capital accumulation by the government operates positive externalities, which takes the form of an exogenous increase in the productivity of private firms. Improvement of productivities via government investment causes a reduction of marginal costs of private firms. Therefore, if \( \alpha_g > 0 \), we call \( K^g_t \) productive public capital (hereafter, PPC).

Finally, public capital is accumulated by government investment as follows:

\[
K_t^g = (1 - \delta_g)K_{t-1}^g + \zeta_{t,i} G_{t}^i.
\]  

(5.11)

\( \delta_g \) is the depreciation rate of public capital and \( \zeta_{t,i} \) is government investment specific technology shock.

The remaining parts of the model are in line with the standard medium-scale DSGE model (e.g. SW 2007), embedding nominal price and wage rigidities, investment adjustment cost, monetary policy rule, and so on. Our model consists of 49 equations and 15 structural shocks. The entire model is described in Appendix. Tables 5.1 and 5.2 report endogenous variables and structural shocks.

Now, we turn to the illustration on effects of non-wasteful fiscal expansions on unemployment.

5.2.3 Effects of Non-wasteful Fiscal Expansions for Unemployment

How do non-wasteful fiscal expansions affect unemployment? Here, we intuitively explain mechanisms that non-wasteful fiscal expansions may bring additional channels for improvements of unemployment.

Suppose that the economy is in a steady state at the initial period. Real wage, aggregate employment, and aggregate labor supply are determined at \( A, H \) and \( L \), respectively. Unemployment

---

4Several ways of introducing productive public capital are suggested by previous studies. Coenen et al. (2013) specifies the capital production function as a CES aggregator. Iwata (2013) specifies an increasing return to scale production function of output such that \( y_t = \epsilon_t (u_t k_{t-1})^\alpha H_{t-1}^{1-\alpha} (k^g_t)^{\alpha_g} \). The specification in this paper is the constant returns to scale production function (5.9) and the Cobb-Douglas type capital production function (5.10) because of the difficulty in identifying the parameter \( \alpha_g \) in the estimation.
$U$ is depicted as the difference between $L$ and $H$. In addition, for the sake of simplicity of illustration, real wage is assumed to be a constant at $A$ at least in the short-term due to nominal wage and price rigidities.

Figure 5.2 (b) shows effects of fiscal expansions without EC and PPC (“wasteful” government spending) on unemployment. The standard story goes as follows: Fiscal expansions create aggregate demand, which induces an increase in labor demand. This effect is depicted by the shift of the labor demand curve to the right, which is shown as (i) in Figure 5.2 (b). But forward-looking households will decrease consumption because of anticipation of future tax increases (negative wealth effect). The decrease in private consumption lowers labor demand, which is depicted by the shift of labor demand curve to the left (shown as (ii) in Figure 5.2(b)). Thus, the effect of increase in aggregate demand will be partly cancelled out by the negative wealth effect. In addition, the decrease in private consumption also causes an incentive to work more (an increase in labor supply), which is shown by the shift of labor supply curve to the right (shown as (iii) in Figure 5.2 (b)). Meanwhile, real wage adjustment is sluggish due to nominal rigidities. Here, real wage is assumed to remain at $A$. As a result, aggregate employment is determined at $H'$, aggregate labor supply is determined at $L'$ and unemployment is determined by the difference $L'$ and $H'$. If the increase in labor demand dominates the increase in labor supply, then fiscal stimuli decrease unemployment. Otherwise, fiscal stimuli increase unemployment.

Now, we consider effects of non-wasteful fiscal expansions on unemployment in Figure 5.2 (c). Under “non-wasteful” government spending, several channels are added to the previous story: Fiscal expansions create aggregate demand. Forward-looking households will decrease private consumption from the negative wealth effect. The decrease of private consumption causes an increase in labor supply. Up to this point, effects of fiscal stimuli to unemployment are the same as the previous story.

Suppose that the parameter $\nu_g$ in equation (5.8) is negative, which corresponds to the case with EC. Then, an increase in government consumption stimulates private consumption because of EC. The additional channel of EC to private consumption is depicted by the shift of labor demand to the right, which brings an improvement in unemployment. This channel is shown as (iv) in Figure 5.2(c). Furthermore, from equation (5.6), the increase in private consumption leads to a decrease of labor supply under nominal rigidities: The increase of private consumption decreases marginal utility of consumption $\varphi_c^t$. This raises the RHS in (5.6), that is, MRS between labor supply and consumption. On the other hand, real wage, the LHS in (5.6), is fixed due to nominal rigidities. Thus, to recover the equality of (5.6), members must work less (a decrease of labor disutility). This channel is shown as (v) in Figure 5.2 (c).

Suppose that the parameter $\alpha_g$ in equation (5.10) is positive, which corresponds to the case with PPC. Then, an increase in government investment improves the productivities of private firms, which implies a decrease in marginal costs of private firms. Forward-looking private firms set their prices by taking future marginal costs into account under nominal stickiness. Thus, accumulation of PPC delivers a decrease in inflation, which triggers monetary easing policy. Therefore, the “crowd-out” effect will be weakened, which stimulates both private consumption and private investment. These effects are realized by the shift of labor demand to the right (shown as (iv) in Figure 5.2 (c)), which improves unemployment. It is worth noting that the channel through PPC has relatively longer effects than the channel through EC: Due to price adjustment sluggishness, the effect through the decrease in future inflation will be delayed but long-lasting.

Through the channels of non-wasteful fiscal expansions, aggregate labor supply and aggregate employment are determined at $L''$ and $H''$, respectively. As a result, non-wasteful government
spending may help unemployment to be reduced via fiscal stimuli.

Of course, the above illustration relies heavily on an extreme assumption that real wage stays at the same level. A rise of real wage leads to an increase in unemployment (and vice versa). The sluggishness of real wage adjustment depends on the strengths of price and nominal wage rigidities. It is also important how the monetary authority reacts to variations of inflation and output: An increase in private consumption through the channel of EC will cause inflation. If the central banker is a strong fighter of inflation, the increase in private consumption might be dominated by a crowding out effect from a monetary tightening policy.

Therefore, we need to estimate structural parameters to examine qualitatively and quantitatively the effects of non-wasteful fiscal expansions on unemployment in the Japanese economy. The next section describes our estimation methodologies.

5.3 Estimation Method

We estimate parameters using Japanese macroeconomic quarterly data from 1980Q2 to 2012Q4. It should be noted that the estimation period includes zero-interest-rate policy periods. We argue about this point in section 5.4. This section provides a data description and illustrates how to implement the Bayesian estimation method.

5.3.1 Data and Measurement Equation

We regard the following 14 variables as observable: GDP, consumption, private investment, government consumption, government investment, nominal bond holdings, unemployment, nominal interest rate, wage inflation, inflation, investment goods inflation, consumption tax rate, corporate income tax rate, and labor income tax rate. Let data denoted as \( \Omega_{\text{data}}^t \), \( \Omega \in \{Y, C, I, G^c, G^i, B, U, R, W, P, P^i, \tau^c, \tau^k, \tau^h, N\} \) where \( N_{\text{data}}^t \) indicates the labor force data.

Real variables are expressed as per capita variables in our model. Since the model is a closed economy, we use GDP data excluding net export.

Regarding the construction of tax rate data, we follow Mendoza et al. (1994). Figure 5.3 shows three distortionary tax rates constructed in this paper. Consumption tax rate data is higher than normal due to the inclusion of specific higher taxed goods such as tabacco, alcohol, etc. Corporate income tax rate data also becomes higher than actuality since we use the operating surplus as the denominator in the calculation formula.

Since there are 15 structural shocks for 14 observable variables, we do not face the stochastic singularity problem in evaluating the likelihood. It is pointed out, however, that DSGE models have only low prediction power for price and wage inflation due to high volatilities and difficulties capturing those volatile variations of price and wage inflation by additional “structural” shocks. Therefore, we add measurement errors to nominal wage inflation, price inflation, and investment goods price inflation, denoted as \( \varepsilon_{\omega}^t, \varepsilon_{\Pi}^t, \) and \( \varepsilon_{\Pi^i}^t \), respectively.

Finally, the measurement equation is defined as the following, in which the “hat” indicates the percent deviation from the steady state:

\[
\hat{\omega}_t = \omega_t + \varepsilon_{\omega}^t,
\]
has 3.53% from the sample mean. Since we can derive endogenously the steady state of employment, we can calibrate several parameters, following CTW and GSW. Table 5.3.2 Preliminary Settings

Prior to the parameter estimation, we calibrate several parameters, following CTW and GSW. Table 5.4 summarizes calibrated parameters in this paper.

The subjective discount factor, $\beta$, is set to 0.995. Since $\bar{R}/\bar{\Pi} = \mu_{z+}$ in steady state, this calibration implies that the real interest rate in a steady state is set to 4% annual rate under zero trend inflation and zero technological progress, i.e. $\bar{\Pi} = \mu_{z+} = 1$. It should be noted that we estimate the trend inflation, $\bar{\Pi}$, and the growth rate of technological progress, $\mu_{z+}$, to be consistent with structural and measurement equations. Both $\lambda$ and $\lambda_w$ are set to 1.20, which implies steady states of mark-up rates for price and wage settings are 20% and 25.6%, respectively. On the private and public capital (quarterly) depreciation rates, $\delta$ and $\delta_g$, we employ sample means and set them to 2.31% and 0.81%, respectively.

In addition, we calibrate two parameters due to the difficulties of identifying the estimate: the persistency parameter of the endogenous preference shifter, $v$, and the parameter on the investment adjustment cost, $\tau''$. On the former parameter, $v$, we employ the estimation result of GSW and set it to 0.02. The latter, $\tau''$, is calibrated as 2.58, which comes from the posterior mean of CTW.

Additionally, we calibrate several ratios, tax rates, and unemployment in steady states so as to match the sample means. Steady states for government consumption to GDP ratio, $\frac{g_c}{y}$, government investment to GDP ratio, $\frac{i_c}{y}$, and debt to GDP ratio, $\frac{b}{y}$, are set to 14.3%, 5.12% and 193.4%, respectively. We also employ sample means to calibrate three distortionary tax rates in a steady state, where consumption tax rate, corporate income tax rate and labor income tax rate are set to 6.95%, 48.7% and 25.6%, respectively. Similarly, the steady state of unemployment, $U$, is calibrated as 3.53% from the sample mean. Since we can derive endogenously the steady state of employment, $H$, from structural equations, the steady state of the desirable labor supply, $L$, is derived by $L = \frac{H}{1-U}$. The details of evaluating steady states for endogenous variables are described in Appendix.

5These calibrations also imply that price elasticity for intermediate goods demand, $|d\ln Y_{j,t}/d\ln P_{j,t}|$ $(\forall j \in (0,1))$, and wage elasticity for skilled labor demand, $|d\ln L_{j,t}/d\ln W_{j,t}|$ $(\forall j \in (0,1))$ are set to $\lambda \frac{\lambda}{\lambda + \frac{\lambda}{\lambda}} = \frac{\lambda}{\lambda + \frac{\lambda}{\lambda}} = 6.$
5.4. RESULTS

5.3.3 Prior Distributions

Table 5.5 reports the prior distributions employed in this paper. For the choice of the prior distribution, we mostly refer to the estimation results of CTW, GSW, and Iwata (2013).

Prior mean of the parameter of EC, \( \nu_g \), is set to zero. Prior means of \( \alpha_g \) and \( \alpha \) (capital income share) are set to 0.20 and 0.40, respectively, which implies the prior mean for the marginal productivity of public capital \((\alpha \times \alpha_g)\) is set to 0.08. Regarding parameters of fiscal policy rules, prior means on spending reversal rules, \( \phi_{g,c} \) and \( \phi_{g,i} \), are set to 0.50, and persistency parameters, \( \rho_{g,c} \) and \( \rho_{g,i} \), are set to 0.95. On parameters of monetary policy rules, prior means of Taylor coefficients on inflation and output are set to 1.5 and 0.125, respectively. It should be noted that the Taylor coefficient on inflation, \( \phi_\Pi \), is transformed into \( \phi_\Pi,0 > 0 \) which satisfies \( \phi_\Pi = 1 + \phi_\Pi,0 \) to ensure the Taylor principle (equivalently, to avoid the indeterminacy problem).

Posterior distributions of structural parameters are estimated via Markov chain Monte Carlo (MCMC) method. We sample four separate chains for 52,000 replicates each, discarding the first 2,000 replicates. Thus, posterior means of parameters are calculated by 200,000 replicates.

5.4 Results

This section presents estimation results. First of all, we report posterior means of structural parameters. Then, we show impulse response functions (hereafter, IRFs), focusing on four structural shocks. Finally, we provide historical decompositions on output, inflation and unemployment.

5.4.1 Estimated Parameters

Table 5.5 reports posterior distributions of structural parameters. We can confirm convergence on all of the estimated parameters from the CI (convergence inference) proposed by Gelman and Rubin (1992). The average acceptance rate across four chains is about 39%.

First of all, we take notice of estimated parameters on EC and PPC The posterior mean of \( \nu_g \) is estimated to a negative value \((-0.023)\), which implies that government consumption is a compliment to private consumption. The magnitude, however, is so small and EC is still ambiguous since the 90% credible interval includes zero.\(^7\) The posterior means of \( \alpha_g \) and \( \alpha \) are 0.155 and 0.386, respectively. This result implies the marginal productivity of public capital \((\alpha \times \alpha_g)\) is calculated as about 0.06.\(^8\)

We now turn to the estimation results of parameters on fiscal policy rules. On parameters of the spending reversal rules, \( \phi_{g,c} \) and \( \phi_{g,i} \) are estimated as 0.289 and 0.496, respectively. On persistency parameters, \( \rho_{g,c} \) and \( \rho_{g,i} \) are 0.970 and 0.958, respectively. It should be pointed out that one government spending is partly canceled out by the reduction of the other spending due to debt accumulation. However, the canceling out effect is negligibly small, since estimation results imply that 1% increase of deficit reduces only by 0.0087% on government consumption and by 0.0197% on government investment.

Regarding monetary policy rule, the Taylor coefficient on inflation, \( \phi_\Pi \), is 1.532 (from \( \phi_\Pi,0 = 0.532 \)) and the Taylor coefficient on output, \( \phi_y \), is 0.035. The interest rate smoothing parameter, \( \rho_R \), is estimated by 0.507.

\(^6\)If CI is below 1.20, then the corresponding parameter can be regarded as to be converged.

\(^7\)The magnitude is really smaller than that of Iwata (2013) where \( \nu_g \) is estimated to -0.416.

\(^8\)This productivity is slightly higher than that of Iwata (2013) where that is estimated to 0.046.
Finally, parameters on nominal rigidities, $\xi_p$ and $\xi_w$, are estimated as 0.560 and 0.489, respectively. These results indicate average durations to remain the same price and wage are about 2.27 quarters on price and about 1.95 quarters on wage.

5.4.2 IRFs

Figures 5.4 (a)-(d) depict impulse response functions (hereinafter, IRFs) for government investment shock, government consumption shock, monetary policy shock, and neutral technology shock. All of IRFs are responses against a 1% shock and calculated by using posterior means of parameters.

**IRFs for government investment shock**

First of all, we take notice of the IRFs for a 1% increase of government investment depicted in Figure 5.4 (a). The blue line corresponds to the case with PPC, i.e. when $\alpha_g = 0.155$, and the red line corresponds to the case without PPC, i.e. when $\alpha_g = 0$. A positive government investment shock with PPC leads to a relatively lower inflation rate since public capital accumulation gradually reduces real marginal cost. According to our result, the inflation rate on the blue line falls below the red line after eight quarters. Relatively lower inflation loosens monetary tightening policy (after 12 quarters), which delivers relatively higher consumption and investment (after 15 quarters). After all, a positive government investment shock persistently enhances output in the case with PPC.

However, there are no differences between two cases on effects of government investment for unemployment: In both cases, 1% stimulus arising from government investment brings an improvement of unemployment by 0.06%. Relatively lower inflation through the increase of productivity causes real wage to be relatively higher in the case with PPC than in the case without PPC. The higher real wage, however, partly hurts employment: The increase in labor demand from higher output is cancelled out by the decrease in employment from higher real wage.

**IRFs for government consumption shock**

Next, we consider the effect of government consumption. Figure 5.4 (b) reports IRFs for a 1% increase in government consumption. The blue line corresponds to the case with EC, i.e. $\nu_g = -0.023$, and the red line corresponds to the case without EC, i.e. $\nu_g = 0$. In the case with EC, an increase of government consumption directly boosts private consumption (the crowd-in effect), whereas the crowd-out effect also appears in both cases since the fiscal stimulus causes inflation, which leads to an increase in real interest rate through the monetary tightening policy. Our results indicate the latter effect completely dominates the former one. Since the magnitude of $\nu_g$ is so small, almost all of the same responses are shown among both cases. As a result, we can see a 1% increase of government consumption reduces unemployment by 0.18% in both cases.

After all, we can see that fiscal stimuli improve unemployment, but almost all of contributions are explained by the creation of aggregate demand. The channels of non-wasteful government spending play little role in improvement of unemployment.

**IRFs for monetary policy shock**

Figures 5.4 (a) and (b) indicate that both fiscal stimuli can be regarded as demand shocks since these shocks generate a cyclical comovement of output with inflation. The monetary authority is assumed to raise inevitably nominal interest rate against inflation caused by fiscal stimuli, which worsens unemployment through the crowd-out effect. So we turn to the effect of monetary policy. Figure 5.4
5.4. RESULTS

(c) shows IRFs for 1% monetary easing policy: IRFs are responses against unanticipated 1% reduction in short-term nominal interest rate. Unemployment is improved by 0.15% by the reduction of 100 basis points of nominal interest rate. This result might suggest that policy coordination matters in evaluating the effects of fiscal stimuli on unemployment. Again, it should be noted that zero interest rate periods are included during our estimation periods. Since the crowding out effect will become smaller under policy coordination between fiscal expansions and zero-interest-rate policy, we will discuss the robustness of our estimation results in the next subsection.

IRFs for neutral technology shock

Finally, we examine the effect of productivity improvement. Figure 5.4 (d) reports IRFs for a 1% positive neutral technology shock. The response of unemployment for a productivity improvement is still ambiguous since totally opposite evidence is shown by the previous empirical studies: Gali (1999) and GSW show a positive response, whereas Christiano et al. (2004) report a negative one. The results of this paper indicate a positive response of unemployment against an increase of productivity (about 1.1% increase at peak) and unemployment returns to the steady state after five quarters, which is in line with the result of GSW. Real rigidities are key factors, as are nominal rigidities: An increase in productivity (a decrease of marginal cost) delivers a deflation and triggers a reduction of real interest rate through monetary easing policy, which stimulates aggregate demand from the intertemporal substitution effect. However, consumption and investment cannot respond immediately due to habit persistence and adjustment cost. These sluggish responses of aggregate demand require real wage adjustment to clear the labor market, but the real wage also cannot be adjusted quickly because of nominal rigidities. After all, the temporary excess supply caused by a positive productivity shock is resolved by the reduction in employment.

5.4.3 Historical Decomposition

Output

Figure 5.5 (a) shows historical decomposition on output. The variations of output are mostly explained by four structural shocks: neutral technology shock, investment specific technology (hereafter, IST) shock, labor supply shock and preference shock.

It should be noted that there are three findings on the determinants of the variations on output: First, neutral technology shock is not a key factor explaining the recession during the so-called lost decade. Second, preference shock switches from positive to negative contributions after the bursting of the bubble economy. Finally, IST shock is one of the key factors of depression after the 2000s including periods of financial crisis of 2007-2008.

The first finding is in contrast to Hayashi and Prescott (2002), who conclude that a decline in productivity plays an important role in the long-lasting recession. On the other hand, the above two findings are in line with Kawamoto (2005), in that there is a possibility that productivity continued growing even during the so-called lost decade and that the decline of labor utilization rate (a proxy of a demand component) is a key determinant of the recession. Regarding the third finding, the negative contribution of IST shock might be related to negative financial shocks.9

---

9Justiniano and Primiceri (2008) find that an increase of the volatility of IST shock is a main source of the recession during the financial crisis in the U.S. economy. They also point out there is a possibility that IST shock includes the financial shock. See also Kaihatsu and Kurozumi (2014).
Inflation

Figure 5.5 (b) depicts historical decomposition on inflation. We can recognize that serious deflation after the late 1990s is mainly attributed to negative contributions of IST shock, preference shock, and labor supply shock. After the beginning of the burst of the bubble economy, IST shock contributes negatively to inflation. Then, preference shock and labor supply shock follow the negative contributions of IST shock after 2000. It can be pointed out that the negative preference shock is a key factor on deflation during the financial crisis of 2008-2009.

Unemployment

Figure 5.5 (c) shows the historical decomposition on unemployment. There are two main driving forces of unemployment variations: Neutral technology shock and labor supply shock. As mentioned above, a positive neutral technology shock makes employment decrease due to real and nominal rigidities. In addition, a negative labor supply shock indicates a decrease in the reservation wage, which causes an increase in aggregate labor supply. The combination of positive neutral technology shock with negative labor supply shock leads to serious unemployment, especially after 2000.

Figure 5.5 (d) displays government contributions for unemployment. We can see a negative comovement of government consumption shock with unemployment after the bursting of the bubble economy. Since the late 1990s and especially after the financial crisis of 2007-2008, the positive government consumption shock plays a significant role in the reduction of unemployment rate by more than 1.5%.

5.4.4 Robustness Check

The periods of zero-interest-rate policy are included during our estimation period. The policy will suppress the crowding out effect, which will enhance the effects of expansionary fiscal policies (see e.g. Christiano et al. 2011). Thus, the inclusion of periods of zero-interest-rate policy may support our main findings, i.e. fiscal stimuli improve unemployment. In this subsection, we estimate the model in the periods when the nominal interest rate does not bind zero and check the robustness of our results.

Figure 5.6 (a) shows the nominal interest rate from 1980Q2 to 2012Q4. While the Bank of Japan (BoJ) adopted the zero-interest-rate policy (with quantity easing policy) after March 2001, the BoJ has already derived a short-term nominal interest rate (uncollateral overnight call rate) to 0.15% at February 1999. Therefore, we estimate the model in the following sub-sample periods: 1980Q2-2001Q1 and 1980Q2-1998Q4.

Table 5.6 shows posterior means and 90% credible intervals in full sample and two sub-sample periods. We can see almost all parameters are surprisingly robust even if we include or exclude zero interest rate periods. Posterior means of EC and PPC, $\nu_g$ and $\alpha_g$, without including zero interest rate periods are also in 90% credible intervals of full sample. Thus, IRFs for fiscal stimuli are almost all the same as those of full sample and our results do not necessarily rely on the reduction of crowding out effects in zero interest rate periods.

This robustness of estimated parameters under including zero interest rate periods is consistent with Hirose and Inoue (2016), who investigate estimation bias when zero interest rate periods are included via the following Monte Carlo experiments: Given true distributions of parameters, data including zero interest rate are artificially generated by a new Keynesian model. Then, they estimate parameters without corresponding to zero interest rate periods and check the estimation bias. As a result, they find the robustness of estimated parameters, which might support our results.
5.5. CONCLUSION

However, Table 5.6 also shows contradictionary results with Hirose and Inoue (forthcoming). We find reductions of almost all estimated shock volatilities when zero interest rate periods are included. In contrast, Hirose and Inoue (2016) report an upward bias on the estimated volatility of the monetary policy.

Binding at zero lower bound of nominal interest rate indicates deviations from the monetary policy rule, and deviations are regarded as monetary policy shocks. Thus, including zero interest rate periods will brings upward estimation bias for the volatility of monetary policy shock. This interpretation, however, is appropriate if the central bank keeps the monetary policy rule strictly in positive interest rate periods, which is a crucial assumption of Hirose and Inoue (2016) in generating artificial data.

Figure 5.6 (b) shows a scatter plot between inflation (annual rate, horizontal axis) and nominal interest rate (annual rate, vertical axis). We can make two observations: First, deviations from the monetary policy rule seem to be high even in positive interest rate periods. Second, the variance of inflation seems to become smaller after 2001Q2. That is, there might be a possibility that the central bank does not strictly keep the Taylor rule in positive interest rate periods. The volatile variations of inflation in positive interest rate periods are captured as relatively higher volatility of the monetary policy shock. Therefore, estimating parameters including data after 2001Q2 when variations of inflation become relatively moderate might reduce volatility of the monetary policy shock even if the monetary authority keeps nominal interest rate at zero. The reason: our results show the decline of variance of monetary policy shock when zero interest rate periods are included.

5.5 Conclusion

Introducing unemployment and non-wasteful government spending into a medium-scale DSGE model, this study examines the effect of government spending on unemployment in the Japanese economy. There are four findings: (i) 1% increase of government consumption reduces unemployment by 0.18%. (ii) 1% stimulus arising from government investment brings an improvement of unemployment by 0.06%. (iii) Since the late 1990s and especially after the financial crisis of 2007-2008, the positive government consumption shock plays a significant role in the reduction of unemployment rate by more than 1.5%. (iv) We also find the “crowd-in” channel where fiscal stimuli induce private consumption and investment to increase through the EC and PPC does not have much influence on unemployment variations.

There are remaining issues in this study: First, as mentioned before, several papers suggest different methods to introduce non-wasteful government spending under alternative specifications of utility and production functions. Thus, we should check the robustness of our result to ensure that the channels of non-wasteful fiscal expansions do not play important roles for variations of unemployment. Second, we introduce not only non-wasteful government spending but also tax rules. If we attempt to simulate the effect of a tax increase, we should estimate parameters, taking news shocks into account. Third, we should reconsider the validity of modeling unemployment based on the market power of workers. While our model predicts higher market power of workers will raise the unemployment rate, Christiano (2011) points out that the hypothesis is rejected in Japan. Thus, it should be necessary to model the labor market more richly, for instance, by embedding search-matching friction where firms have some bargaining power.

10On the importance of anticipated shocks for aggregate fluctuations, see Fujiwara et al. (2011) and Hirose and Kurozumi (2012). See also Benjamin et al. (2013), which examines the effect of fiscal news for business cycle fluctuations in the U.S. economy.
## 5.6 Tables and Figures

<table>
<thead>
<tr>
<th>Endogenous variable</th>
<th>Name</th>
<th>Observability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ( \Pi_t )</td>
<td>Inflation</td>
<td>*</td>
</tr>
<tr>
<td>2 ( \Pi^w_t )</td>
<td>Wage inflation</td>
<td>*</td>
</tr>
<tr>
<td>3 ( \Pi^{\text{opt}}_t )</td>
<td>Optimal inflation</td>
<td></td>
</tr>
<tr>
<td>4 ( w^{\text{opt}}_t )</td>
<td>Optimal wage inflation</td>
<td></td>
</tr>
<tr>
<td>5 ( \Pi_{\kappa,t} )</td>
<td>Inflation indexation</td>
<td></td>
</tr>
<tr>
<td>6 ( \Pi_{\kappa,w} )</td>
<td>Wage inflation indexation</td>
<td></td>
</tr>
<tr>
<td>7 ( K^p_t )</td>
<td>Numerator of NKPC (price)</td>
<td></td>
</tr>
<tr>
<td>8 ( F^p_t )</td>
<td>Denominator of NKPC (price)</td>
<td></td>
</tr>
<tr>
<td>9 ( K^w_t )</td>
<td>Numerator of NKPC (wage)</td>
<td></td>
</tr>
<tr>
<td>10 ( F^w_t )</td>
<td>Denominator of NKPC (wage)</td>
<td></td>
</tr>
<tr>
<td>11 ( w_t )</td>
<td>Real wage</td>
<td></td>
</tr>
<tr>
<td>12 ( q_t )</td>
<td>Tobin’s Q (shadow price of capital)</td>
<td></td>
</tr>
<tr>
<td>13 ( R_t )</td>
<td>Nominal interest rate</td>
<td></td>
</tr>
<tr>
<td>14 ( r^p_t )</td>
<td>Real rental price of capital (intermediate goods firms)</td>
<td></td>
</tr>
<tr>
<td>15 ( p^k_t )</td>
<td>Real rental price of capital (capital producing firms)</td>
<td></td>
</tr>
<tr>
<td>16 ( mc_t )</td>
<td>Real marginal cost</td>
<td></td>
</tr>
<tr>
<td>17 ( y_t )</td>
<td>Output</td>
<td></td>
</tr>
<tr>
<td>18 ( c_t )</td>
<td>Consumption</td>
<td></td>
</tr>
<tr>
<td>19 ( \tilde{c}_t )</td>
<td>Bundled consumption</td>
<td></td>
</tr>
<tr>
<td>20 ( i_t )</td>
<td>Private investment</td>
<td></td>
</tr>
<tr>
<td>21 ( g^c_t )</td>
<td>Government consumption</td>
<td></td>
</tr>
<tr>
<td>22 ( g^i_t )</td>
<td>Government investment</td>
<td></td>
</tr>
<tr>
<td>23 ( k_t )</td>
<td>Private capital</td>
<td></td>
</tr>
<tr>
<td>24 ( k^p_t )</td>
<td>Public capital</td>
<td></td>
</tr>
<tr>
<td>25 ( \tilde{k}_t )</td>
<td>Capital</td>
<td></td>
</tr>
<tr>
<td>26 ( u_t )</td>
<td>Capital utilization rate</td>
<td></td>
</tr>
<tr>
<td>27 ( H_t )</td>
<td>Employment</td>
<td></td>
</tr>
<tr>
<td>28 ( \varphi^c_t )</td>
<td>Marginal utility on consumption</td>
<td></td>
</tr>
<tr>
<td>29 ( \chi^h_t )</td>
<td>Endogenous preference shifter 1</td>
<td></td>
</tr>
<tr>
<td>30 ( \tilde{\chi}_{t} )</td>
<td>Endogenous preference shifter 2</td>
<td></td>
</tr>
<tr>
<td>31 ( L_t )</td>
<td>Desirable labor supply</td>
<td></td>
</tr>
<tr>
<td>32 ( U_t )</td>
<td>Unemployment rate</td>
<td></td>
</tr>
<tr>
<td>33 ( \Lambda_{t,t+1} )</td>
<td>Stochastic discount factor</td>
<td></td>
</tr>
<tr>
<td>34 ( b_t )</td>
<td>Real bond holdings</td>
<td></td>
</tr>
<tr>
<td>35 ( v_t )</td>
<td>Real profit (intermediate goods firms)</td>
<td></td>
</tr>
<tr>
<td>36 ( v^k_t )</td>
<td>Real profit (capital producing firms)</td>
<td></td>
</tr>
<tr>
<td>37 ( \Pi_t )</td>
<td>Trend inflation</td>
<td></td>
</tr>
<tr>
<td>38 ( \mu_{z,t} )</td>
<td>Labor augmented technological progress</td>
<td></td>
</tr>
<tr>
<td>39 ( \mu_{\psi,t} )</td>
<td>Investment specific technological progress</td>
<td></td>
</tr>
<tr>
<td>40 ( \mu_{z^+,t} )</td>
<td>Neutral technological progress</td>
<td></td>
</tr>
<tr>
<td>41 ( \epsilon_t )</td>
<td>AR(1) neutral technology shock</td>
<td></td>
</tr>
<tr>
<td>42 ( \zeta^c_t )</td>
<td>AR(1) preference shock</td>
<td></td>
</tr>
<tr>
<td>43 ( \zeta^h_t )</td>
<td>AR(1) labor supply shock</td>
<td></td>
</tr>
<tr>
<td>44 ( \zeta^i_t )</td>
<td>AR(1) investment specific technology shock</td>
<td></td>
</tr>
<tr>
<td>45 ( \zeta^g,i_t )</td>
<td>AR(1) government investment specific technology shock</td>
<td></td>
</tr>
<tr>
<td>46 ( \tau^c_t )</td>
<td>AR(1) consumption tax shock</td>
<td></td>
</tr>
<tr>
<td>47 ( \tau^k_t )</td>
<td>AR(1) corporate income tax shock</td>
<td></td>
</tr>
<tr>
<td>48 ( \tau^h_t )</td>
<td>AR(1) labor income tax shock</td>
<td></td>
</tr>
<tr>
<td>49 ( \tau_t )</td>
<td>AR(1) lump-sum tax shock</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Endogenous Variables
Table 5.2: Structural Shocks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Unit</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_{t, data}$</td>
<td>Real GDP (excluding net export)</td>
<td>a billion yen</td>
<td>SNA</td>
</tr>
<tr>
<td>$C_{t, data}$</td>
<td>Real private consumption</td>
<td>a billion yen</td>
<td>SNA</td>
</tr>
<tr>
<td>$I_{t, data}$</td>
<td>Real private investment</td>
<td>a billion yen</td>
<td>SNA</td>
</tr>
<tr>
<td>$G_{t, data}$</td>
<td>Real government consumption</td>
<td>a billion yen</td>
<td>SNA</td>
</tr>
<tr>
<td>$G_{t, data}$</td>
<td>Real government investment</td>
<td>a billion yen</td>
<td>SNA</td>
</tr>
<tr>
<td>$B_{t, data}$</td>
<td>Real government bond</td>
<td>a billion yen</td>
<td>Bank of Japan</td>
</tr>
<tr>
<td>$W_{t, data}$</td>
<td>Compensation of Employees</td>
<td>a billion yen/a thousand</td>
<td>SNA and Statistic Bureau, MIC</td>
</tr>
<tr>
<td>$U_{t, data}$</td>
<td>Unemployment</td>
<td>%</td>
<td>Statistic Bureau, MIC</td>
</tr>
<tr>
<td>$R_{t, data}$</td>
<td>Overnight call rate</td>
<td>%</td>
<td>Bank of Japan</td>
</tr>
<tr>
<td>$P_{t, data}$</td>
<td>GDP deflator</td>
<td>2005 year=100</td>
<td>SNA</td>
</tr>
<tr>
<td>$P_{t, data}$</td>
<td>Investment deflator</td>
<td>2005 year=100</td>
<td>SNA</td>
</tr>
<tr>
<td>$\tau_{t, data}$</td>
<td>Consumption tax rate</td>
<td>%</td>
<td>-</td>
</tr>
<tr>
<td>$\tau_{t, data}$</td>
<td>Corporate income tax rate</td>
<td>%</td>
<td>-</td>
</tr>
<tr>
<td>$\tau_{t, data}$</td>
<td>Labor income tax rate</td>
<td>%</td>
<td>-</td>
</tr>
<tr>
<td>$N_{t, data}$</td>
<td>Labor force</td>
<td>a thousand</td>
<td>Statistic Bureau, MIC</td>
</tr>
</tbody>
</table>

Table 5.3: Data
### Chapter 5. Impacts of Government Spending on Unemployment

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.995</td>
<td>Subjective discount factor</td>
<td>CTW (2013)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.20</td>
<td>Gross price mark-up rate</td>
<td>CTW (2013)</td>
</tr>
<tr>
<td>$\lambda_w$</td>
<td>1.20</td>
<td>Gross wage mark-up rate</td>
<td>CTW (2013)</td>
</tr>
<tr>
<td>$v$</td>
<td>0.020</td>
<td>Endogenous preference shifter parameter</td>
<td>GSW (2011)</td>
</tr>
<tr>
<td>$S''$</td>
<td>2.580</td>
<td>Private investment adjustment cost parameter</td>
<td>CTW (2013)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.023</td>
<td>Private capital depreciation rate</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$\delta_g$</td>
<td>0.008</td>
<td>Public capital depreciation rate</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$p^w$</td>
<td>0.143</td>
<td>Government consumption to GDP ratio</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$p^g$</td>
<td>0.051</td>
<td>Government investment to GDP ratio</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$y^g$</td>
<td>1.934</td>
<td>Debt to GDP ratio</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$\tau_c$</td>
<td>0.070</td>
<td>Consumption tax rate</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$\tau_k$</td>
<td>0.487</td>
<td>Corporate income tax rate</td>
<td>Sample mean</td>
</tr>
<tr>
<td>$\tau_h$</td>
<td>0.256</td>
<td>Labor income tax rate</td>
<td>Sample mean</td>
</tr>
</tbody>
</table>

Table 5.4: Calibrated Parameters

![Figure 5.1 Unemployment Rate in Japan](image.png)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior</th>
<th>Posterior</th>
<th>90% credible interval</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi_p$</td>
<td>0.500</td>
<td>0.025</td>
<td>0.560 0.023 [0.522, 0.598]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\xi_w$</td>
<td>0.500</td>
<td>0.025</td>
<td>0.489 0.025 [0.449, 0.531]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.600</td>
<td>0.100</td>
<td>0.820 0.041 [0.748, 0.882]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\kappa_w$</td>
<td>0.600</td>
<td>0.100</td>
<td>0.958 0.012 [0.938, 0.975]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\alpha_y$</td>
<td>0.200</td>
<td>0.100</td>
<td>0.155 0.060 [0.061, 0.259]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.400</td>
<td>0.100</td>
<td>0.386 0.025 [0.345, 0.428]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\nu_y$</td>
<td>0.000</td>
<td>0.100</td>
<td>-0.023 0.100 [-0.185, 0.143]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.200</td>
<td>0.100</td>
<td>0.732 0.028 [0.685, 0.776]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\sigma_H$</td>
<td>3.000</td>
<td>0.100</td>
<td>3.151 0.097 [2.991, 3.310]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_a$</td>
<td>0.200</td>
<td>0.075</td>
<td>0.789 0.141 [0.572, 1.038]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{g,c}$</td>
<td>0.950</td>
<td>0.010</td>
<td>0.970 0.007 [0.958, 0.980]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{g,i}$</td>
<td>0.950</td>
<td>0.010</td>
<td>0.958 0.008 [0.944, 0.970]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_{g,c}$</td>
<td>0.500</td>
<td>0.100</td>
<td>0.289 0.063 [0.195, 0.400]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\phi_{g,i}$</td>
<td>0.500</td>
<td>0.100</td>
<td>0.496 0.080 [0.371, 0.632]</td>
<td>1.02</td>
</tr>
<tr>
<td>$\rho_R$</td>
<td>0.500</td>
<td>0.050</td>
<td>0.507 0.046 [0.432, 0.582]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_{1,0}$</td>
<td>0.500</td>
<td>0.050</td>
<td>0.532 0.049 [0.456, 0.614]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_y$</td>
<td>0.125</td>
<td>0.050</td>
<td>0.035 0.012 [0.018, 0.057]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\Pi$</td>
<td>1.002</td>
<td>0.005</td>
<td>0.998 0.004 [0.992, 1.005]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\mu_{z,+}$</td>
<td>1.002</td>
<td>0.001</td>
<td>1.002 0.001 [1.000, 1.003]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\mu_{z}$</td>
<td>1.000</td>
<td>0.001</td>
<td>0.999 0.001 [0.997, 1.000]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{u_z}$</td>
<td>0.500</td>
<td>0.050</td>
<td>0.390 0.041 [0.323, 0.459]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{u_{g,\psi}}$</td>
<td>0.500</td>
<td>0.050</td>
<td>0.496 0.049 [0.415, 0.578]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_u$</td>
<td>0.500</td>
<td>0.050</td>
<td>0.427 0.041 [0.359, 0.496]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>0.950</td>
<td>0.005</td>
<td>0.958 0.004 [0.950, 0.965]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{c,c}$</td>
<td>0.950</td>
<td>0.005</td>
<td>0.955 0.005 [0.947, 0.963]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\rho_{c,H}$</td>
<td>0.950</td>
<td>0.010</td>
<td>0.944 0.009 [0.928, 0.958]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{c,i}$</td>
<td>0.800</td>
<td>0.010</td>
<td>0.804 0.010 [0.787, 0.819]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{c,i}$</td>
<td>0.800</td>
<td>0.075</td>
<td>0.801 0.075 [0.670, 0.913]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{c,c}$</td>
<td>0.500</td>
<td>0.075</td>
<td>0.517 0.074 [0.396, 0.637]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\rho_{r_z}$</td>
<td>0.500</td>
<td>0.075</td>
<td>0.807 0.032 [0.753, 0.857]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_{r_h}$</td>
<td>0.500</td>
<td>0.075</td>
<td>0.615 0.067 [0.501, 0.722]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>0.500</td>
<td>0.075</td>
<td>0.824 0.028 [0.777, 0.868]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_{\nu_z}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0036 0.0005 [0.0028, 0.00438]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\sigma_{\nu_{g,\psi}}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0029 0.0004 [0.0024, 0.00357]</td>
<td>1.04</td>
</tr>
<tr>
<td>$\sigma_{\rho}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0084 0.0017 [0.0060, 0.01146]</td>
<td>1.02</td>
</tr>
<tr>
<td>$\sigma_{\nu_z}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0030 0.0004 [0.0024, 0.00368]</td>
<td>1.03</td>
</tr>
<tr>
<td>$\sigma_{\nu_{g,\psi}}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0083 0.0015 [0.0062, 0.01105]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_{\rho}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0127 0.0023 [0.0093, 0.01691]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_{\nu_z}$</td>
<td>0.150</td>
<td>0.100</td>
<td>0.0162 0.0023 [0.0127, 0.02030]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_{\nu_{g,\psi}}$</td>
<td>0.150</td>
<td>0.100</td>
<td>0.3285 0.2953 [0.0926, 0.85982]</td>
<td>1.08</td>
</tr>
<tr>
<td>$\sigma_{\rho}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0026 0.0003 [0.0021, 0.00315]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_{\nu_z}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0032 0.0004 [0.0025, 0.00391]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\sigma_{\nu_{g,\psi}}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0027 0.0003 [0.0022, 0.00325]</td>
<td>1.02</td>
</tr>
<tr>
<td>$\sigma_{\rho}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0109 0.0023 [0.0077, 0.01523]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\sigma_{\nu_z}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0036 0.0005 [0.0029, 0.00448]</td>
<td>1.01</td>
</tr>
<tr>
<td>$\sigma_{\nu_{g,\psi}}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0049 0.0006 [0.0040, 0.00601]</td>
<td>1.00</td>
</tr>
<tr>
<td>$\sigma_{\rho}$</td>
<td>0.150</td>
<td>0.500</td>
<td>0.0030 0.0004 [0.0024, 0.00370]</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Table 5.5: Prior and Posterior distributions
### Table 5.6: Robustness Check

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>90% credible interval</td>
<td>Mean</td>
<td>90% credible interval</td>
</tr>
<tr>
<td>$\xi_1$</td>
<td>$[0.522, 0.598]$</td>
<td>$0.528$</td>
<td>$[0.488, 0.567]$</td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>$[0.449, 0.531]$</td>
<td>$0.489$</td>
<td>$[0.449, 0.530]$</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>$[0.748, 0.882]$</td>
<td>$0.790$</td>
<td>$[0.699, 0.868]$</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>$[0.938, 0.975]$</td>
<td>$0.946$</td>
<td>$[0.919, 0.969]$</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>$[0.061, 0.259]$</td>
<td>$0.146$</td>
<td>$[0.050, 0.262]$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$[0.345, 0.428]$</td>
<td>$0.361$</td>
<td>$[0.313, 0.412]$</td>
</tr>
<tr>
<td>$\nu_1$</td>
<td>$[-0.185, 0.143]$</td>
<td>$-0.021$</td>
<td>$[-0.183, 0.139]$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>$[0.685, 0.776]$</td>
<td>$0.753$</td>
<td>$[0.698, 0.802]$</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>$3.151$</td>
<td>$2.991, 3.310$</td>
<td>$3.113$</td>
</tr>
<tr>
<td>$\rho_{g,i}$</td>
<td>$0.789$</td>
<td>$0.572, 1.038$</td>
<td>$0.621$</td>
</tr>
<tr>
<td>$\rho_{g,i}$</td>
<td>$0.970$</td>
<td>$0.958, 0.980$</td>
<td>$0.961$</td>
</tr>
<tr>
<td>$\rho_{g,i}$</td>
<td>$0.958$</td>
<td>$0.944, 0.970$</td>
<td>$0.956$</td>
</tr>
<tr>
<td>$\phi_{g,i}$</td>
<td>$0.289$</td>
<td>$0.195, 0.400$</td>
<td>$0.385$</td>
</tr>
<tr>
<td>$\phi_{g,i}$</td>
<td>$0.496$</td>
<td>$0.371, 0.632$</td>
<td>$0.465$</td>
</tr>
<tr>
<td>$\phi_{g,i}$</td>
<td>$0.507$</td>
<td>$0.432, 0.582$</td>
<td>$0.495$</td>
</tr>
<tr>
<td>$\phi_{g,i}$</td>
<td>$0.532$</td>
<td>$0.456, 0.614$</td>
<td>$0.527$</td>
</tr>
<tr>
<td>$\phi_{g,i}$</td>
<td>$0.035$</td>
<td>$0.018, 0.057$</td>
<td>$0.042$</td>
</tr>
<tr>
<td>$\mathbb{H}$</td>
<td>$1.002$</td>
<td>$1.000, 1.003$</td>
<td>$1.002$</td>
</tr>
<tr>
<td>$\mu_1^*$</td>
<td>$0.999$</td>
<td>$0.997, 1.000$</td>
<td>$0.999$</td>
</tr>
<tr>
<td>$\mu_1^*$</td>
<td>$0.390$</td>
<td>$0.323, 0.459$</td>
<td>$0.414$</td>
</tr>
<tr>
<td>$\mu_1^*$</td>
<td>$0.496$</td>
<td>$0.415, 0.578$</td>
<td>$0.495$</td>
</tr>
<tr>
<td>$\rho_3$</td>
<td>$0.427$</td>
<td>$0.359, 0.496$</td>
<td>$0.457$</td>
</tr>
<tr>
<td>$\nu_1$</td>
<td>$0.585$</td>
<td>$0.950, 0.965$</td>
<td>$0.955$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.944$</td>
<td>$0.928, 0.958$</td>
<td>$0.948$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.804$</td>
<td>$0.787, 0.819$</td>
<td>$0.803$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.801$</td>
<td>$0.670, 0.913$</td>
<td>$0.799$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.517$</td>
<td>$0.396, 0.637$</td>
<td>$0.503$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.807$</td>
<td>$0.753, 0.857$</td>
<td>$0.719$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.615$</td>
<td>$0.501, 0.722$</td>
<td>$0.535$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.824$</td>
<td>$0.777, 0.868$</td>
<td>$0.770$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0036$</td>
<td>$0.0028, 0.00438$</td>
<td>$0.0054$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0029$</td>
<td>$0.0024, 0.00357$</td>
<td>$0.0046$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0084$</td>
<td>$0.0060, 0.01146$</td>
<td>$0.0119$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0030$</td>
<td>$0.0024, 0.00368$</td>
<td>$0.0046$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0083$</td>
<td>$0.0062, 0.01105$</td>
<td>$0.0120$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0127$</td>
<td>$0.0093, 0.01601$</td>
<td>$0.0165$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0162$</td>
<td>$0.0127, 0.02030$</td>
<td>$0.0219$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.3285$</td>
<td>$0.0926, 0.85982$</td>
<td>$0.563$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0026$</td>
<td>$0.0021, 0.00315$</td>
<td>$0.0041$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0032$</td>
<td>$0.0025, 0.00391$</td>
<td>$0.0050$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0027$</td>
<td>$0.0022, 0.00325$</td>
<td>$0.0041$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0109$</td>
<td>$0.0077, 0.01523$</td>
<td>$0.0152$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0036$</td>
<td>$0.0029, 0.00448$</td>
<td>$0.0054$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0049$</td>
<td>$0.0040, 0.00601$</td>
<td>$0.0071$</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>$0.0030$</td>
<td>$0.0024, 0.00370$</td>
<td>$0.0046$</td>
</tr>
</tbody>
</table>
5.6. TABLES AND FIGURES

Figure 5.2 (a) Unemployment

Figure 5.2 (b) Effects of Wasteful Fiscal Expansions
CHAPTER 5. IMPACTS OF GOVERNMENT SPENDING ON UNEMPLOYMENT

Figure 5.2 (c) Effects of Non-Wasteful Fiscal Expansions

Figure 5.3 Distortionary Tax Rates
5.6. TABLES AND FIGURES

Figure 5.4 (a) IRFs to Government Investment Shock

Figure 5.4 (b) IRFs to Government Consumption Shock
CHAPTER 5. IMPACTS OF GOVERNMENT SPENDING ON UNEMPLOYMENT

Figure 5.4 (c) IRFs to Monetary Policy Shock

Figure 5.4 (d) IRFs to Neutral Technology Shock
Figure 5.5 (a) Historical Decomposition: Output

Figure 5.5 (b) Historical Decomposition: Inflation
CHAPTER 5. IMPACTS OF GOVERNMENT SPENDING ON UNEMPLOYMENT

Figure 5.5 (c) Historical Decomposition: Unemployment

Figure 5.5 (d) Historical Decomposition (Government Spending Shock): Unemployment
Figure 5.6 (a) Inflation and Nominal Interest Rate

Figure 5.6 (b) Inflation vs. Nominal Interest Rate
5.7 Appendix

5.7.1 Model Description

This subsection describes the entire model: The dynamic optimization conditions on firms and households, fiscal and monetary policy rules and market clearing condition.

Firms

We can classify firms into three types according to type of goods: Final goods firms, intermediate goods firms, and capital-producing firms. The final goods firms purchase intermediate goods, produce homogenous final goods by bundling differentiated intermediate goods, and sell the final goods to households and government. Intermediate goods firms purchase labor from households and capital from the capital producing firms, produce the differentiated intermediate goods using labor and capital as factors of production, and sell the intermediate goods to the final goods firms. The capital-producing firms rent private capital from households, produce homogenous capital goods by combining private capital with public capital, and sell the homogenous capital goods to the intermediate goods firms.

Final Goods Firms

Final goods firms bundle the differentiated intermediate goods $Y_{j,t}$, $(j \in [0, 1])$, and produce homogenous final goods. Given the final good price, $P_t$, and the intermediate good price, $P_{j,t}$, the final good firm solves the following profit maximization problem:

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

s.t. $Y_t = \left[ \int_0^1 Y_{j,t}^{\frac{1}{\lambda}} dj \right]^{\lambda}$, where $\lambda \geq 1$.

From the first order condition (hereafter FOC), the demand function for an intermediate-good produced by the firm $j$ is obtained as:

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

Substituting the demand function above into the production function, the price of final goods can be expressed as:

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

Intermediate Goods Firms

An intermediate good firm $j$ produces a differentiated good $Y_{j,t}$, $(j \in [0, 1])$, using capital $K_{j,t}$ and labor input $H_{j,t}$:

$$Y_{j,t} = \epsilon_t K_{j,t-1}^{\alpha} (z_t H_{j,t})^{1-\alpha} - z_{t}^{\alpha} \Theta$$

$\epsilon_t$ stands the neutral technology shock. The scaling variable, $z_{t}^{\alpha}$, is defined as

$$z_{t}^{\alpha} = \Psi_t^{\frac{1-\alpha}{1-\alpha}} z_t.$$
Then, the (gross) balanced growth rate can be expressed as $\mu_{x,t} \equiv \frac{z^t_{t+1}}{z^t_{t-1}}$. The growth rates of output, consumption and real wage become $\mu_{x,t}$. Meanwhile, previous papers point out that investment specific technological (IST) progress has an important role for variations of output (e.g. Greenwood et al. 2000). This implies private capital and investment have a different growth rate from the above balanced growth rate. The IST progress is captured by a scaling variable, $\Psi_t$, and the scaling variable for private capital and investment can be expressed as $z^t_i \Psi_t$.

We can remove trend components using these scaling variables. The detrended production function of the intermediate good can be expressed as follows:

$$y_{j,t} = \epsilon_t \left( \frac{1}{\mu_{x,t}} \right)^{\alpha} \left( \frac{\tilde{k}_{j,t-1}}{z^t_{t+1}} \right)^{\alpha} H_{j,t}^{1-\alpha} - \Theta,$$

where $y_{j,t} \equiv Y_{j,t} - z^t_{x,t}$, $\tilde{k}_{j,t} \equiv \frac{\tilde{K}_{j,t}}{z^t_{t+1} \Psi_t}$.

Hereafter, model variables in scaled form are denoted by small letters.

Given the capital rental price, $R_k^t$, and nominal wage, $W_t$, intermediate good firms solve the cost minimization problem. Then, we obtain the following real marginal cost, $mc_t$ in scaled form:

$$mc_t = \frac{1}{\epsilon_t} \left( \frac{r_k^t}{w_t} \right)^{\alpha} \left( \frac{\psi_t R_k^t}{P_t} \right)^{1-\alpha},$$

where $w_t \equiv \frac{W_t}{z^t_{x,t}} P_t$, $r_k^t \equiv \frac{\Psi_t R_k^t}{P_t}$.

The FOC to the cost minimization problem gives:

$$\frac{\tilde{k}_{j,t-1}}{H_{j,t}} = \mu_{x,t} \mu_{x,t} \frac{1}{1-\alpha} \frac{w_t}{r_k^t}.$$

From this result, the capital-labor ratio does not depend on the subscript $j$. Therefore, real marginal cost and optimal capital-labor ratio are the same across intermediate goods firms.

Since intermediate goods firms have some market power in the intermediate goods market, they set prices to maximize their profits. We introduce Calvo-type nominal rigidities for price settings of intermediate goods firms. Intermediate goods firms can optimize their prices with a probability $1 - \xi_p$, and the firms that cannot revise their price index those prices partially with an inflation rate at the previous period, $\Pi_{t-1}$, and with the trend inflation rate, $\bar{\Pi}_t$, controlled by the central bank. Therefore, given $P_t$ and $Y_t$, the intermediate good firm $j$ sets price so as to maximize the following objective function:

$$\max_{P_{j,t}} E_t \sum_{i=0}^{\infty} (\beta \xi_p)^i \Lambda_{t+i} \left[ (P_{j,t+i} - P_{t+i} mc_{t+i}) Y_{j,t+i} - z^t_{t+i} P_{t+i} \Theta \right],$$

where $P_{j,t} = \Pi_{\kappa,t} P_{j,t-1}$, $\Pi_{\kappa,t} = (\Pi_{t-1})^\kappa (\bar{\Pi}_t)^{1-\kappa}$, $\Lambda_{t+i} = \left( \frac{P_{j,t}}{P_t} \right)^{\frac{1}{1-\kappa}} Y_t$.

$\beta \in (0,1)$ stands for the subjective discount factor, and $\Lambda_{t+i}$ stands for the stochastic discount factor. The fixed cost, $\Theta$, is introduced to ensure that profits of the intermediate good firms are zero in a steady state. $\kappa \in [0,1]$ represents the inflation indexation parameter.
Since the firms that can revise their prices set the same price among them, FOCs in scaled forms are derived as follows:

\[
\Pi_t^{opt} = \lambda \left( \frac{K^p_t}{F^p_t} \right) = \frac{P_t^{opt}}{P_t}
\]  

\[
K^p_t = mc_t y_t + \beta \xi_p E_t \Lambda_{t,t+1} \mu_{z^+,t+1} (\Pi_{t+1})^{1 - \frac{1}{1-\alpha}} (\Pi_{\kappa,t+1})^{\frac{1}{1-\alpha}} K^p_{t+1}
\]

\[
P^p_t = y_t + \beta \xi_p E_t \Lambda_{t,t+1} \mu_{z^+,t+1} (\Pi_{t+1})^{\frac{1}{1-\alpha}} (\Pi_{\kappa,t+1})^{1 + \frac{1}{1-\alpha}} F^p_{t+1}
\]

where \( \Pi_t^{opt} = \frac{P_t^{opt}}{P_{t-1}}, \Pi_t = \frac{P_t}{P_{t-1}}, y_t = \frac{Y_t}{z^+_t}, \mu_{z^+,t} = \frac{z^+_t}{z^+_{t-1}}. \)

In a flexible price economy, i.e., \( \xi_p = 0 \), we can see the optimal flexible price equals to mark-up \( \lambda \) times real marginal cost. The price level of final goods and inflation rate can be derived as follows:

\[
P_t = \left[ \int_0^1 (P_{j,t})^{\frac{1}{1-\alpha}} dj \right]^{1-\lambda} = \left[ (1 - \xi_p) (P_{t}^{opt})^{\frac{1}{1-\alpha}} + \xi_p [(\Pi_{\kappa,t}) P_{t-1}]^{\frac{1}{1-\alpha}} \right]^{1-\lambda}.
\]

\[\Leftrightarrow \Pi_t = \left[ (1 - \xi_p) (\Pi_t^{opt})^{\frac{1}{1-\alpha}} + \xi_p (\Pi_{\kappa,t})^{\frac{1}{1-\alpha}} \right]^{1-\lambda}.
\]  

(5.20)

The new Keynesian Phillips curve consists of equations (5.17)-(5.20).

**Capital-Producing Firms**

There are two types of capital in this economy. One is the effective private capital \( u_t K_{t-1} \) where \( u_t \) is capital utilization rate. The other is the public capital \( K^g_{t-1} \) accumulated by government. Capital-producing firms produce a homogenous capital good \( K_{t-1} \), bundling the effective private capital with public capital. Given the rental price of effective private capital, \( P^k_t \), the profit maximization problem of capital producing firm can be expressed as:

\[
\max_{u_t, \kappa_{t-1}} R^k_t K_{t-1} - P^k_t u_t K_{t-1} - z^+_t \Psi_t \Theta^k
\]

s.t. \( \kappa_{t-1} = (K^g_{t-1})^{\alpha_g} (u_t K_{t-1})^{1 - \alpha_g} \)  

(5.22)

\( \Theta^k \) stands for the fixed cost and (5.22) is the production function of capital where \( \alpha_g \) is the marginal productivity of public capital for the bundled capital, \( K_{t-1} \). It should be noted that the productivity of the public capital for output can be expressed as \( \alpha \times \alpha_g \) from (5.13) and (5.22). If \( \alpha_g \) is positive, then the public capital accumulation by government operates positive externalities to capital-producing firms in the form of increasing their productivities exogenously. From the production function of output (5.13), the public capital accumulation indirectly increases the productivity for the intermediate goods firms. Therefore, if \( \alpha_g > 0 \), we call \( k^g_t \) productive public capital.

The FOC in scaled form is derived as:

\[
p^k_t = (1 - \alpha_g) \frac{\kappa_{t-1}}{u_t K_{t-1}} r^k_t
\]

where \( p^k_t = \frac{\Psi_t P^k_t}{P_t} \).

Notice that \( p^k_t = r^k_t \) if \( \alpha_g = 0 \). When public capital is wasteful, i.e. \( \alpha_g = 0 \), there are no entrants to capital-producing business since their profits become negative due to the fixed cost \( \Theta^k \).
Households

Following the GSW framework, we consider a large household with a continuum of members represented by the unit square and indexed by a pair \((j, h) \in [0, 1] \times [0, 1]\). The first dimension indexed by \(j \in [0, 1]\) represents a differentiated skill in which a given household member is specialized. The second dimension indexed by \(h \in [0, 1]\) indicates member’s labor disutility.

Households have market power due to a differentiated skills indexed by \(j \in [0, 1]\), but they are assumed to face nominal wage rigidities a la Calvo, in line with Erceg et al. (2000). Members within each household have different labor disutilities with uniformly distributed as \(h \in [0, 1]\). Given the nominal wage determined by each household, members decide to work or not taking their labor disutilities into consideration. In addition, we assume full risk sharing of consumption across members: Thanks to sharing labor income equally, members can enjoy consuming with the same level.

Then, the preference of a member \(h\) (who has a disutility \(h\)) in any household \(j\) at period \(t\) can be written by

\[
\zeta^c_t \ln \left( \tilde{C}_{j,t} - \theta \tilde{C}_{t-1} \right) - 1_t(j, h) \zeta^h_t \chi^h_t A_H h^{\sigma_h} \tag{5.24}
\]

\(\tilde{C}_{j,t}\) stands for consumption of member \(h\) in union \(j\) and \(\tilde{C}_t \equiv \int_0^1 \tilde{C}_{j,t} dj\) stands for aggregate consumption. The term \(\theta \tilde{C}_{t-1}\) indicates (external) habits on consumption and the parameter \(\theta \in (0, 1)\) depicts the importance of habit formation. \(1_t(j, h)\) is the indicator function which takes a value equal to one if the member \(h\) is employed at period \(t\), and zero otherwise. It is worth noting that the indicator function means members decide to work with fixed hours (normalized as unity) or not. \(\chi^h_t\) stands the endogenous preference shifter defined as the following equation:

\[
\chi^h_t \equiv \frac{Z_{\chi,t}}{\tilde{C}_t - \theta \tilde{C}_{t-1}}, \tag{5.25}
\]

\[
Z_{\chi,t} = Z_{\chi,t-1}^{1-v} \left( \tilde{C}_t - \theta \tilde{C}_{t-1} \right)^v, \tag{5.26}
\]

This preference specification leads marginal labor disutility decreases during (aggregate) consumption booms. Two structural shocks are embedded: \(\zeta^c_t\) is the preference shock and \(\zeta^h_t\) is the labor supply shock. \(A_H\) is the scale parameter and \(\sigma_h\) is the inverse Frisch elasticity.

Let \(H_{j,t}\) be defined as employment of household \(j\). Then, aggregating the member’s utility regarding \(h\), we derive utility of union \(j\) at period \(t\) as follows:

\[
\zeta^c_t \ln \left( \tilde{C}_{j,t} - \theta \tilde{C}_{t-1} \right) - \zeta^h_t \chi^h_t A_H \int_0^{H_{j,t}} h^{\sigma_h} dh
= \zeta^c_t \ln \left( \tilde{C}_{j,t} - \theta \tilde{C}_{t-1} \right) - \zeta^h_t \chi^h_t A_H \frac{H_{j,t}^{1+\sigma_h}}{1 + \sigma_h} \tag{5.27}
\]

Thus, the preference of household \(j\) falls into the standard functional form.

**Objective Function of Household**

At the beginning at period \(t\), household \(j\) owns two types of assets: bond \(B_{j,t-1}\) and capital stock \(K_{j,t-1}\). Household \(j\) maximizes the following discounted present value of her utility, controlling consumption \(C_{j,t}\), bond holdings \(B_{j,t}\), investment \(I_{j,t}\), capital stock \(K_{j,t}\) and capital utilization rate
nominal wage and $R$ is substitutes for private consumption. If $\nu$ and government consumption, which is the so-called $u$ from intermediate good firms and capital-producing firms, and $\nu$ of households is positive when $\nu$ is negative, an increase in government consumption raises marginal utility of private consumption at the present period. As a result, households respond with an increase in private consumption for an increase in government consumption: A negative $\nu$ causes a cyclical comovement between private consumption and government consumption, which is the so-called Edgeworth complementarities (hereafter, EC). If $\nu$ is positive, a counter-cyclical comovement is shown, which implies government consumption is substitutes for private consumption. If $\nu$ is zero, government consumption is independent of private consumption.

In addition, $V(G_t^c)$ is assumed to be satisfied by $\partial V/\partial G_t^c > 0$ to ensure that the marginal utility of households is positive when $\nu$ is negative, following Karras (1994), Ganelli and Tervala (2009) and Iwata (2013).

Equation (5.29) is household $j$’s budget constraint at period $t$. The household gains labor income, dividends $V_{j,t}^c$ and $V_{j,t}^k$ from intermediate good firms and capital-producing firms, and interest payments from bonds and capital holdings. On the other hand, the household expends its income as consumption, investment and bond holdings. $P_t^c$ is investment goods price, $W_{j,t}$ is nominal wage and $R_t$ is nominal interest rate. $\tau_t^c$, $\tau_t^h$, $\tau_t^k$, $T_t$ are consumption tax rate, labor income tax rate, corporate income tax rate and (nominal) lump-sum transfer, respectively. (5.30) is the capital accumulation equation. $\delta$ is depreciation rate, $S$ is investment adjustment cost where $S(\mu, \mu') = S'(\mu, \mu') = 0$. $\zeta_t$ stands for the investment specific technology shock. The rental price of effective private capital is $P_t^i a(u(j,t))$. $a(u(j,t))$ stands for capital utilization cost where $a(\cdot) = 0$ in a steady state and $a'(\cdot) > 0$.

**FOCs of Households**

Let Lagrange multipliers for the budget constraint be denoted by $\Lambda_{j,t}$ and for the capital accumulation equation denoted by $Q_{j,t}$. FOCs with respect to $C_{j,t}$, $B_{j,t}$, $I_{j,t}$, $K_{j,t}$ and $u_{j,t}$ are derived as follows:

$$\Lambda_t P_t (1 + \tau_t^c) = \frac{\zeta_t^c}{C_t - \theta C_{t-1}},$$

$$\Lambda_t = \beta E_t R_t \Lambda_{t+1},$$

$$\Lambda_t P_t^i = \zeta_t^i Q_t \left[ 1 - S \left( \frac{I_t}{I_{t-1}} \right) - S' \left( \frac{I_t}{I_{t-1}} \right) \frac{I_t}{I_{t-1}} \right] + \beta E_t Q_{t+1} \zeta_t^i S' \left( \frac{I_{t+1}}{I_t} \right) \left( \frac{I_{t+1}}{I_t} \right)^2,$$
\[ Q_t = \beta E_t \left( (1 - \delta)Q_{t+1} + \Lambda_t (1 - \tau^k) \left( P^k_{t+1} u_{t+1} - P^i_{t+1} a(u_{t+1}) \right) \right) \]  
(5.36)

\[ P^k_t = P^i_t a'(u_t) \]  
(5.37)

From equation (5.33), the Lagrange multiplier is marginal utility of consumption. Using equations (5.33) and (5.34), we can derive a consumption Euler equation. Let the stochastic discount factor be defined as the following equation:

\[ E_t \Lambda_{t,t+i} \equiv E_t \frac{\Lambda_{t+i}}{\Lambda_t}. \]

\( \Lambda_{t,t+i} \) represents the marginal rate of intertemporal substitution between consumption at period \( t \) and period \( t + i \).

FOCs in scaled forms of households are expressed as follows:

\[ \varphi^c(1 + \tau^c) = \frac{\zeta^c}{\bar{c}_t - \theta \bar{c}_{t-1} \mu_{z+t, t}}, \]  
(5.38)

\[ \beta E_t \Lambda_{t,t+1} = \frac{1}{R_t} \]  
(5.39)

\[ 1 = \zeta^i \left[ 1 - S \left( \mu_{z,t} \bar{\mu}_{z,t} \frac{i_t}{i_{t-1}} \right) - S' \left( \mu_{z,t} \bar{\mu}_{z,t} \frac{i_t}{i_{t-1}} \right) \mu_{z+1,t} \bar{\mu}_{z+1,t} \frac{i_{t+1}}{i_t} \right] q_t \]

\[ + \beta E_t \Lambda_{t,t+1} \frac{\Pi_{t+1}}{\mu_{z,t+1}} \zeta^i \left[ \frac{\Pi_{t+1}}{\mu_{z,t+1}} S' \left( \mu_{z+1,t} \bar{\mu}_{z+1,t} \frac{i_{t+1}}{i_t} \right) \right] \left( \mu_{z+1,t} \bar{\mu}_{z+1,t} \frac{i_{t+1}}{i_t} \right)^2 q_{t+1} \]

\[ q_t = \beta E_t \Lambda_{t,t+1} \frac{\Pi_{t+1}}{\mu_{z,t+1}} \left[ (1 - \delta) q_{t+1} + (1 - \tau^k_{t+1}) \left( P^k_{t+1} u_{t+1} - P^i_{t+1} a(u_{t+1}) \right) \right] \]  
(5.40)

\[ \bar{p}^k_t = a'(u_t), \]  
(5.41)

where \( \varphi^c \equiv z^+_{t} P_t \Lambda_t, \bar{c}_t \equiv \frac{\bar{c}_t}{z^+_{t}}, q_t \equiv \frac{\Psi_t Q_t}{P_t \Lambda_t}. \)

The stochastic discount factor in scaled form can be derived as follows:

\[ E_t \Lambda_{t,t+1} = E_t \frac{\varphi^c_{t+1}}{\Pi_{t+1} \mu_{z+t, t+1} \varphi^c_t}. \]  
(5.43)

The budget constraint in scaled form can be obtained as the following equation:

\[ (1 + \tau^c) c_t + i_t + b_t + \tau_t = (1 - \tau^k) w_t H_t + (1 - \tau^k) \left[ \left( P^k_t u_t - a(u_t) \right) \frac{k_{t-1}}{\mu_{z+t, t} \bar{\mu}_{z+t, t}} + v_t + v^k_t \right] + \frac{1}{\mu_{z+t, t}} \frac{R_{t-1}}{R_t} b_{t-1} \]  
(5.44)

where \( b_t \equiv \frac{B_t}{z^+_{t} P_t}, w_t \equiv \frac{W_t}{z^+_{t} P_t}, v_t \equiv \frac{V_t}{z^+_{t} P_t}, \tau_t \equiv \frac{T_t}{z^+_{t} P_t}. \)

\( b_t \) stands for bond holdings, \( w_t \) stands for real wage, \( v_t \) stands for real dividend and \( \tau_t \) stands for real lump-sum transfer from government. Finally, households satisfy the following transversality conditions:

\[ \lim_{i \to \infty} \beta \Lambda_{t,t+i} \mu_{z+t, t+i} \Pi_{t+i} (b_{t+i} + k_{t+i}) = 0. \]  
(5.45)
CHAPTER 5. IMPACTS OF GOVERNMENT SPENDING ON UNEMPLOYMENT

Wage Setting

Following Erceg et al. (2000), employment agencies bundle household’s differentiated skill $H_{j,t}$ into a homogenous aggregate labor $H_t$, and sell the bundled labor to intermediate goods firms. Given the nominal wage $W_t$ and the differentiated skill $j$’s nominal wage $W_{j,t}$, the employment agency maximizes the following profit:

$$\max_{H_{j,t}} \ W_t H_t - \int_0^1 W_{j,t} H_{j,t} dj, \ \text{s.t.} \ H_t = \left[ \int_0^1 H_{j,t} \frac{1}{\lambda_w} dj \right]^{\lambda_w}. $$

FOC leads to the demand function for differentiated labor $j$:

$$H_{j,t} = \left( \frac{W_{j,t}}{W_t} \right)^{\frac{\lambda_w}{1-\lambda_w}} H_t.$$ 

Substituting this into bundling technology, the nominal wage level is derived by

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

Household $j$ can revise it’s nominal wage with a probability $1 - \xi_w$, otherwise it indexes it’s wage with past wage inflation rate and trend inflation rate. Household $j$’s objective function is as follows:

$$\max_{W_{j,t}} \ E_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \left[ -\xi_{t+i}^h \chi_{t+i}^h A_H \frac{H_{j,t+i}^{1+\sigma_H}}{1+\sigma_H} + \Lambda_{j,t+i}(1-\tau_{t+i}) W_{j,t+i} H_{j,t+i} \right] $$

where

$$W_{j,t} = \Pi_{\kappa,t}^w W_{j,t-1}, \quad \Pi_{\kappa,t}^w \equiv (\Pi_{t-1}^w)^{\kappa_w} (\Pi_t)^{1-\kappa_w}. $$

$$H_{j,t} = \left( \frac{W_{j,t}}{W_t} \right)^{\frac{\lambda_w}{1-\lambda_w}} H_t. $$

where $\Pi_t$ is trend inflation and $\kappa_w \in (0, 1)$ is wage inflation persistency. Letting $w_{t,\text{opt}}$ denote optimal wage and $\Pi_{t,\text{opt}}^w$ denote optimal inflation rate, we can derive the following FOCs in scaled forms:

$$w_{t,\text{opt}} = \left\{ A_H \lambda_w \frac{K_t^w F_t^w}{F_t^w} \right\}^{\frac{1-\lambda_w}{1-\lambda_w - \lambda_w \sigma_H}} $$

$$K_t^w = \zeta_t^h \chi_t^h \pi_t^h \frac{\lambda_w (1+\sigma_H)}{1-\lambda_w} H_t^{1+\sigma_H} + \beta \xi_w (\Pi_{\kappa,t}^w) \frac{\lambda_w (1+\sigma_H)}{1-\lambda_w} (\mu_{t+1}^z \Pi_{t+1}^{1-\kappa_w})^{-\frac{1}{1-\lambda_w}} K_{t+1}^w $$

$$F_t^w = \varphi_t^c (1-\tau_t^h) w_t^{\frac{\lambda_w}{1-\lambda_w}} H_t + \beta \xi_w (\Pi_{\kappa,t+1}^w) \frac{1}{1-\lambda_w} (\mu_{t+1}^z \Pi_{t+1})^{\frac{1}{1-\lambda_w}} F_{t+1}^w $$

where $w_{t,\text{opt}} \equiv \frac{W_{t,\text{opt}}}{z_t^P P_t}$, $w_t \equiv \frac{W_t}{z_t^P P_t}$.
The endogenous preference shifters in scaled forms can be written by

$$
\chi^h_t = \frac{z_{\chi,t}}{\tilde{c}_t - \theta \tilde{c}_{t-1} \frac{1}{\mu_{z^+,t}}}
$$

(5.50)

$$
z_{\chi,t} = \left( \frac{z_{\chi,t-1}}{\mu_{z^+,t}} \right)^{1-v} \left( \frac{\tilde{c}_t - \theta \tilde{c}_{t-1} \frac{1}{\mu_{z^+,t}}}{v} \right)
$$

(5.51)

where $z_{\chi,t} \equiv \frac{Z_{\chi,t}}{z^+_t}$.

Nominal wage level and wage inflation rate can be derived as follows:

$$
W_t = \left[ (1 - \xi_w) \left( W_t^{opt} \right)^{1-\lambda_w} + \xi_w \left( \Pi_t^{w, opt} W_{t-1} \right)^{1-\lambda_w} \right]^{\frac{1}{1-\lambda_w}}.
$$

$$
\iff \quad \Pi_t^w = \left[ (1 - \xi_w) \left( \Pi_t^{w, opt} \right)^{1-\lambda_w} + \xi_w \left( \Pi_t^{w, opt} \right)^{1-\lambda_w} \right]^{\frac{1}{1-\lambda_w}}.
$$

(5.52)

The wage-type new Keynesian Phillips curve consists of equations (5.46), (5.47), (5.48), (5.49) and (5.52).

**Unemployment**

Following GSW, a member $h$ in household $j$ has a differentiated skill $j$ and the labor disutility $\zeta^h_t \chi^h_t A_H h^{\sigma_H}$. Given the real wage $w_{j,t}$, we consider the desirable labor supply $L_{j,t}$ of the member $h$ within the household $j$. As long as the real wage is greater than the MRS between labor and consumption, the member is willing to work:

$$
(1 - \tau^h_t) w_t \geq \zeta^h_t \chi^h_t A_H h^{\sigma_H} \varphi_{j,t}
$$

$\tau^h_t$ stands for labor income tax rate. The left-hand side indicates the marginal benefit of labor supply after tax adjustment. The right-hand side is MRS, which corresponds to reservation wage of member $h$.

Thanks to the consumption sharing assumption, the marginal utility of consumption becomes common across members represented as $\varphi_t^c$. Letting the marginal supplier of union $j$’s member be denoted by $L_{j,t}$, we have:

$$
(1 - \tau^h_t) w_{j,t} = \frac{\zeta^h_t \chi^h_t A_H L_{j,t}^{\sigma_H}}{\varphi_{j,t}^c}
$$

It is worth noting that the above labor supply condition is explicitly derived as the FOC wrt. $H_{j,t}$ given nominal wage, since household members are assumed to be price takers.

Then, using an endogenous preference shifter (5.32), household $j$’s labor supply $L_{j,t}$ is derived as follows:

$$
(1 - \tau^h_t) w_{j,t} = (1 + \tau^h_t) \frac{\zeta^h_t}{\varphi_t^c} z_{\chi,t} A_H L_{j,t}^{\sigma_H}
$$

Let aggregate employment be denoted by $H_t$ and let aggregate labor supply be denoted by $L_t$. Then, unemployment rate $U_t$ is defined as the following equation.

$$
U_t \equiv \frac{L_t - H_t}{L_t}
$$

(5.53)
Fiscal and Monetary Authorities

Fiscal Authority

Government purchases final goods at price $P_t$. $G_t^e$ and $G_t^i$ are government consumption and investment, respectively. Let $B_t$ denote the government nominal bond issued at the beginning of the period $t$. Then, government budget constraint can be written by

$$P_t G_t^e + P_t^i G_t^i + R_{t-1} B_{t-1} + T_t = \tau_t^c P_t C_t + \tau_t^h W_t H_t + \tau_t^k \left[ \left( P_t^k u_t - P_t^i a(u_t) \right) K_{t-1} + V_t + V_t^k \right] + B_t.$$  \hspace{1cm} (5.54)

where $g_t^c = \frac{G_t^c}{z_t^c}$, $g_t^i = \frac{G_t^i}{z_t^i \Psi_t}$.

Public capital is accumulated by government investment as follows:

$$K_t^g = (1 - \delta_g) K_{t-1}^g + \zeta_t^g \cdot G_t^i.$$  \hspace{1cm} (5.55)

where $k_t^g = \frac{K_t^g}{z_t^g \Psi_t}$.

$\delta_g$ is depreciation rate of public capital and $\zeta_t^g$ is government investment specific technology shock. It should be noted that government investment and public capital have the same growth rate as private investment and private capital, i.e. the growth rate is $z_t^c \Psi_t$. Following Corsetti et al. (2012) and Iwata (2013), government spending rules are given by

$$\ln \frac{g_t^c}{g^c_t} = \rho_{g^c} \ln \frac{g_{t-1}^c}{g_t^c} - \left( 1 - \rho_{g^c} \right) \phi_{g^c} \ln \frac{b_t}{b} + \varepsilon_{t}^{g^c}$$  \hspace{1cm} (5.56)

In case of $\phi_{g^c} > 0$, government consumption (government investment) is determined so as to smooth deficits (the so-called “spending reversals” rule). $\varepsilon_{t}^{g,i}$ are government consumption and investment shocks, respectively. Furthermore, tax rules are assumed to be represented by the following AR(1) processes:

$$\tau_t^c = \rho_{\tau_c} (\tau_{t-1}^c - \tau^c) + \epsilon_t^{\tau_c}$$  \hspace{1cm} (5.58)

$$\tau_t^k = \rho_{\tau_k} (\tau_{t-1}^k - \tau^k) + \epsilon_t^{\tau_k}$$  \hspace{1cm} (5.59)

$$\tau_t^h = \rho_{\tau_h} (\tau_{t-1}^h - \tau^h) + \epsilon_t^{\tau_h}$$  \hspace{1cm} (5.60)

$$\tau_t = \rho_{\tau} (\tau_{t-1} - \tau) + \epsilon_t^{\tau}$$  \hspace{1cm} (5.61)

Monetary Authority
Following CTW, we assume that monetary authority controls nominal interest rate according to the following rule:

\[
\ln \frac{R_t}{R} = \rho_R \ln \frac{R_{t-1}}{R} + (1 - \rho_R) \left[ \ln \frac{\Pi_t}{\Pi} + \phi_{\Pi} \ln \frac{\Pi_t}{\Pi} + \phi_y \ln \frac{y_t}{y} \right] + \epsilon^R_t \tag{5.62}
\]

\( \phi_{\Pi} \) and \( \phi_y \) are Taylor parameters and \( \rho_R \in (0, 1) \) stands nominal interest rate smoothing parameter. \( \epsilon^R_t \) is monetary policy shock and assumed as iid shock with zero mean.

**Aggregation and Resource Constraint**

Aggregating the production function of the differentiated intermediate good, we derive aggregate supply of homogenous good as the following equation:

\[
y_t = \frac{1}{\tilde{p}_t} \left[ \epsilon_t \left( \frac{1}{\mu_{z,t} \psi_t} \right)^\alpha \left( \tilde{k}_{t-1} \right)^\alpha \left( \frac{H_t}{\tilde{w}_t} \right)^{1-\alpha} - \Theta \right] \tag{5.63}
\]

where \( \tilde{p}_t \equiv \int_0^1 \left( \frac{P_i, t}{P_t} \right)^{\frac{\lambda}{1-\lambda}} d j \) and \( \tilde{w}_t \equiv \int_0^1 \left( \frac{W_i, t}{W_t} \right)^{\frac{\lambda_w}{1-\lambda_w}} d j \) (5.64)

\( \tilde{p}_t \) and \( \tilde{w}_t \) are price and wage dispersions, respectively. Price and wage dispersions satisfy the following recursive equation:

\[
\tilde{p}_t = (1 - \xi_p) \left( \frac{1 - \xi_p \left( \frac{\Pi_{\psi,t}}{\Pi_t} \right)^{\frac{1}{1-\lambda}}} {1 - \xi_p} \right)^\lambda \tilde{p}_{t-1} + \xi_p \left( \frac{\Pi_{\psi,t}}{\Pi_t} \right)^{\frac{1}{1-\lambda}} \tilde{p}_{t-1}.
\] (5.65)

\[
\tilde{w}_t = (1 - \xi_w) \left( \frac{1 - \xi_w \left( \frac{\Pi_{w,t}}{\Pi_t} \right)^{\frac{1}{1-\lambda_w}}} {1 - \xi_w} \right)^{\lambda_w} \tilde{w}_{t-1} + \xi_w \left( \frac{\Pi_{w,t}}{\Pi_t} \right)^{\frac{1}{1-\lambda_w}} \tilde{w}_{t-1}.
\] (5.66)

Using household and government budget constraints ((5.44) and (5.54)), we have following the resource constraint equation:

\[
c_t + i_t + g_i^c + g_i^i = w_t H_t + (p_{k,t}^k u_t - a(u_t)) \frac{\tilde{k}_{t-1}}{\mu_{z, t} \mu_{\psi, t}} + v_t + v_t^k
\]

The real profit of intermediate goods firms can be derived as:

\[
v_t = y_t - \left( w_t H_t + r_t^k u_t r_t^k - \frac{\tilde{k}_{t-1}}{\mu_{z, t} \mu_{\psi, t}} \right)
\] (5.67)

The real profit of capital producing firm can be expressed as:

\[
v_t^k = \alpha_g \frac{\tilde{k}_{t-1}}{\mu_{z, t} \mu_{\psi, t}}
\] (5.68)

Substituting those profits into the resource constraint equation, and using the FOC of capital-producing firm, \( p_{k,t}^k u_t k_{t-1} = (1 - \alpha_g) r_t^k \tilde{k}_{t-1} \), we can obtain the following resource constraint of goods market:

\[
y_t = c_t + i_t + g_i^c + g_i^i + a(u_t) \frac{\tilde{k}_{t-1}}{\mu_{z, t} \mu_{\psi, t}}.
\] (5.69)
5.7.2 Summary of the Model

This subsection summarizes specifications of functional forms on adjustment costs, non-linear simultaneous difference equations and structural shocks.

Specifications of Functional forms

Capital utilization adjustment cost:

\[
a(u_t) = \frac{1}{2} \sigma_b \sigma_a u_t^2 + \sigma_b (1 - \sigma_a) u_t + \sigma_b \left( \frac{\sigma_a}{2} - 1 \right) \\
a'(u_t) = \sigma_b \sigma_a u_t + \sigma_b (1 - \sigma_a)
\]

Private investment adjustment cost:

\[
S(x_t) = \frac{1}{2} \left[ \exp \left( \sqrt{S''(x_t - \mu_z + \mu \psi, t)} \right) + \exp \left( -\sqrt{S''(x_t - \mu_z + \mu \psi, t)} \right) - 2 \right] \\
S'(x_t) = \frac{1}{2} \sqrt{S''} \left[ \exp \left( \sqrt{S''(x_t - \mu_z + \mu \psi, t)} \right) - \exp \left( -\sqrt{S''(x_t - \mu_z + \mu \psi, t)} \right) \right]
\]

where \(x_t = \mu_z + \mu \psi, t \frac{i_t}{t-1}\)

Firms

Real marginal cost:

\[
mc_t = \frac{(r_k^t)^\alpha w_t^{1-\alpha}}{\epsilon_t \alpha^\alpha (1 - \alpha)^{1-\alpha}} 
\]

Cost minimization condition:

\[
\frac{\bar{k}_{t-1}}{\bar{H}_t} = \frac{\mu_{z+} \mu \psi, t}{1 - \alpha} \frac{w_t}{r_k^t} 
\]

New Keynesian Phillips Curve (price):

\[
\Pi_{opt}^t = \lambda K_t \\
\Pi_t = \lambda K_t + (1 - \xi_p) (\Pi_{opt}^t)^{1-\lambda} + \xi_p \Pi_{\kappa, t+1}^{1-\lambda} \\
\Pi_{\kappa, t} = \Pi_{t-1}^{1-k} \\
\Pi_{t} = \left[ (1 - \xi_p) (\Pi_{opt}^t)^{1-\lambda} + \xi_p \Pi_{\kappa, t}^{1-\lambda} \right]^{1-\lambda}
\]

Production function for capital goods:

\[
\bar{k}_{t-1} = (\bar{k}_{t-1})^\alpha g (u_t \bar{k}_{t-1})^{1-\alpha g}
\]

Rental price of private capital goods:

\[
p_t^k = (1 - \alpha_g) \frac{\bar{k}_{t-1}^{1-\alpha g}}{u_t \bar{k}_{t-1}}
\]
5.7. APPENDIX

Household

Edgeworth complementarities between private and government consumption:
\[ \bar{c}_t = c_t + \nu_g g_t^c \] (5.79)

Marginal utility of consumption:
\[ \varphi_t^c(1 + \tau_t^c) = \frac{\zeta_t^c}{\bar{c}_t - \theta \bar{c}_{t-1} \mu_{z,+,t}} \] (5.80)

Euler equation on consumption:
\[ \beta \mathbf{E}_t \Lambda_{t,t+1} = \frac{1}{R_t} \] (5.81)

Stochastic discount factor:
\[ E_t \Lambda_{t,t+1} = E_t \frac{\varphi_{t+1}^c}{\Pi_{t+1} \mu_{z,+,t+1} \varphi_t^c} \] (5.82)

Capital accumulation equation:
\[ k_t = \frac{1}{\mu_{z,+,t} \mu_{\psi,t}} (1 - \delta) k_{t-1} + \zeta_t^c \left[ 1 - S \left( \mu_{z,+,t} \mu_{\psi,t} \frac{i_t}{i_{t-1}} \right) \right] i_t \] (5.83)

FOC wrt. private investment:
\[ 1 = \zeta_t^c \left[ 1 - S \left( \mu_{z,+,t} \mu_{\psi,t} \frac{i_t}{i_{t-1}} \right) - S' \left( \mu_{z,+,t} \mu_{\psi,t} \frac{i_t}{i_{t-1}} \right) \left( \mu_{z,+,t} \mu_{\psi,t} \frac{i_t}{i_{t-1}} \right) \right] q_t \]
\[ + \beta E_t \Lambda_{t,t+1} \Pi_{t+1}^{i_c} \left( \mu_{z,+,t+1} \mu_{\psi,t+1} \frac{i_{t+1}}{i_t} \right)^2 q_{t+1} \] (5.84)

FOC wrt. capital:
\[ q_t = \beta E_t \Lambda_{t,t+1} \frac{\Pi_{t+1}}{\mu_{\psi,t+1}} \left[ (1 - \delta) q_{t+1} + (1 - \tau_{t+1}^k) \left( \rho_{t+1}^k u_{t+1} - a(u_{t+1}) \right) \right] \] (5.85)

FOC wrt. capital utilization rate:
\[ \rho_{t}^k = a'(u_t) \] (5.86)

New Keynesian Phillips curve (wage):
\[ w_t^{opt} = \left\{ A_H \lambda w \frac{K_t^w}{K_t^w} \right\}^{1-\lambda_w} w_t^{\alpha_w} \] (5.87)
\[ K_t^w = \zeta_t^h \lambda w w_t \frac{\lambda_w(1+\sigma_h)}{1-\lambda_w} H_t^{1+\sigma_h} + \beta \xi_w (\Pi_{k,t+1}^w)^{\lambda_w(1+\sigma_h)} (\mu_{z,+,t+1} \Pi_{t+1})^{-\lambda_w} K_{t+1}^w \] (5.88)
\[ F_t^w = \varphi_t^c(1 - \tau_t^h) w_t^{1-\lambda_w} H_t + \beta \xi_w (\Pi_{k,t+1}^w)^{1-\lambda_w} (\mu_{z,+,t+1} \Pi_{t+1})^{-1} F_{t+1}^w \] (5.89)
\[ w_t = \left\{ (1 - \xi_w)(w_t^{opt})^{1-\lambda_w} + \xi_w \left( \Pi_{k,t+1}^w \Pi_t \right)^{1-\lambda_w} \right\}^{1-\lambda_w} \] (5.90)
\[ \Pi_{k,t}^w = (\Pi_{t-1}^w)^{\kappa_w} (\mu_{z,+,t} \Pi_t)^{1-\kappa_w} \] (5.91)
Endogenous preference shifter:
\[
\chi_{t}^{h} = \frac{z_{\chi,t}}{\bar{c}_{t} - \theta \bar{c}_{t-1} \frac{1}{\mu_{z+,t}}}
\]
(5.92)
\[
z_{\chi,t} = \left( \frac{z_{\chi,t-1}}{\mu_{z+,t}} \right)^{1-v} \left( \bar{c}_{t} - \theta \bar{c}_{t-1} \frac{1}{\mu_{z+,t}} \right)^{v}
\]
(5.93)
Desirable labor supply:
\[
(1 - \tau_{t})w_{t} = (1 + \tau_{t}^{c}) \left( \frac{z_{h,t}}{z_{t}} \right) z_{\chi,t} L_{t}^{\sigma_{h}}
\]
(5.94)
Unemployment rate:
\[
U_{t} = \frac{L_{t} - H_{t}}{L_{t}}
\]
(5.95)

**Fiscal and Monetary Authorities**

Public capital accumulation equation:
\[
k_{t}^{g} = \frac{1}{\mu_{z+,t} \mu_{\Psi,t}}(1 - \delta_{g})k_{t-1}^{g} + \epsilon_{t}^{g_{i}} g_{t}^{i}
\]
(5.96)
Fiscal policy rule:
\[
\ln \frac{g_{t}^{c}}{g^{c}} = \rho_{g} \ln \frac{g_{t-1}^{c}}{g^{c}} - (1 - \rho_{g}) \phi_{g} \ln \frac{b_{t}}{b} + \epsilon_{t}^{g_{c}}
\]
(5.97)
\[
\ln \frac{g_{t}^{i}}{g^{i}} = \rho_{g} \ln \frac{g_{t-1}^{i}}{g^{i}} - (1 - \rho_{g}) \phi_{g} \ln \frac{b_{t}}{b} + \epsilon_{t}^{g_{i}}
\]
(5.98)
Monetary policy rule:
\[
\ln \frac{R_{t}}{R} = \rho_{R} \ln \frac{R_{t-1}}{R} + (1 - \rho_{R}) \left[ \ln \frac{\bar{\Pi}_{t}}{\Pi_{t}} + \phi_{\bar{\Pi}} \ln \frac{\bar{\Pi}_{t}}{\Pi_{t}} + \phi_{y} \ln \frac{y_{t}}{y} \right] + \epsilon_{t}^{R}
\]
(5.99)

**Aggregation and Resource Constraint**

Aggregate supply:
\[
y_{t} = \epsilon_{t} \left( \frac{1}{\mu_{z+,t} \mu_{\Psi,t}} \right)^{\frac{\alpha}{\alpha}} k_{t-1}^{g} H_{t}^{1 - \alpha} - \Theta
\]
(5.100)
Resource constraint:
\[
y_{t} = c_{t} + i_{t} + g_{t}^{c} + g_{t}^{i} + \frac{1}{\mu_{z+,t} \mu_{\Psi,t}} a(u_{t}) k_{t-1}
\]
(5.101)
Budget constraint of government:
\[
g_{t}^{c} + g_{t}^{i} + \frac{1}{\mu_{z+,t} \mu_{\Psi,t}} R_{t-1} b_{t-1} = \tau_{t}^{c} c_{t} + \tau_{t}^{h} w_{t} H_{t} + \tau_{t}^{k} p_{t}^{k} u_{t} - a(u_{t}) \frac{k_{t-1}}{\mu_{z+,t} \mu_{\Psi,t}} + \tau_{t}^{k} (v_{t} + v_{t}^{k}) + b_{t} + \tau_{t}
\]
(5.102)
5.7. APPENDIX

Aggregated profit (intermediate goods firms):
\[ v_t = (1 - mc_t)y_t - mc_t\Theta \quad (5.103) \]

Aggregated profit (capital producing firms):
\[ v^k_t = \alpha_g \frac{\bar{k}_{t-1}}{u_tk_{t-1}}r^k_t - \Theta^k \quad (5.104) \]

Auxiliary Variables

Wage inflation:
\[ \frac{w_t}{w_{t-1}} = \frac{\Pi^{w}_t}{\mu^{z+}_{t}\bar{\Pi}_t} \quad (5.105) \]

Growth rate of technological progress:
\[ \mu^{z+}_t = \frac{\alpha}{\mu^{\psi}_t} \mu^{z}_t \quad (5.106) \]

Structural Shocks

Neutral technological progress shock:
\[ \ln \mu^{z+}_t - \ln \mu^{z}_t = \rho^{\mu^{z}_t} (\ln \mu^{z+}_{t-1} - \ln \mu^{z}_t) + \epsilon^{\mu^{z}_t}_t \quad (5.107) \]

Investment specific technological progress shock:
\[ \ln \mu^{\psi}_t - \ln \mu^{\psi}_t = \rho^{\mu^{\psi}_t} (\ln \mu^{\psi}_{t-1} - \ln \mu^{\psi}_t) + \epsilon^{\mu^{\psi}_t}_t \quad (5.108) \]

Trend inflation shock:
\[ \ln \bar{\Pi}_t - \ln \bar{\Pi} = \rho^{\bar{\Pi}_t} (\ln \bar{\Pi}_{t-1} - \ln \bar{\Pi}) + \epsilon^{\bar{\Pi}_t}_t \quad (5.109) \]

Neutral technology shock:
\[ \ln \epsilon^{\epsilon}_t = \rho^{\epsilon^{\epsilon}_t} \ln \epsilon^{\epsilon}_{t-1} + \epsilon^{\epsilon}_t \quad (5.110) \]

Preference shock:
\[ \ln \zeta^{c}_t = \rho^{\zeta^{c}_t} \ln \zeta^{c}_{t-1} + \epsilon^{\zeta^{c}_t}_t \quad (5.111) \]

Labor supply shock:
\[ \ln \zeta^{h}_t = \rho^{\zeta^{h}_t} \ln \zeta^{h}_{t-1} + \epsilon^{\zeta^{h}_t}_t \quad (5.112) \]

Private investment specific technology shock:
\[ \ln \zeta^{i}_t = \rho^{\zeta^{i}_t} \ln \zeta^{i}_{t-1} + \epsilon^{\zeta^{i}_t}_t \quad (5.113) \]

Government investment specific technology shock:
\[ \ln \zeta^{g,i}_t = \rho^{\zeta^{g,i}_t} \ln \zeta^{g,i}_{t-1} + \epsilon^{\zeta^{g,i}_t}_t \quad (5.114) \]
Consumption tax shock:
\[ \tau^c_t - \tau^c = \rho_{\tau^c} (\tau^c_{t-1} - \tau^c) + \epsilon^c_t \]  
(5.115)

Corporate income tax shock:
\[ \tau^k_t - \tau^k = \rho_{\tau^k} (\tau^k_{t-1} - \tau^k) + \epsilon^k_t \]  
(5.116)

Labor income tax shock:
\[ \tau^h_t - \tau^h = \rho_{\tau^h} (\tau^h_{t-1} - \tau^h) + \epsilon^h_t \]  
(5.117)

Lump-sum transfer shock:
\[ \tau_t - \tau = \rho_{\tau} (\tau_{t-1} - \tau) + \epsilon^T_t \]  
(5.118)

5.7.3 Steady States

This subsection explains how we compute steady states of endogenous variables. Following CTW, we impose the following assumptions in the derivation of steady states: First, we assume \( u = 1 \), thus the capital utilization rate is 100% in steady state. Second, we assume that excess profits of firms are zero in steady state, i.e. \( v = v^k = 0 \).

**Endogenous Variables that Have Explicit Solutions**

From equations (5.72)-(5.76), we can derive the following equations:
\[ \Pi = \Pi^{opt} = \Pi_k = \Pi \]  
(5.119)
\[ mc = \frac{1}{\lambda} \]  
(5.120)

Using (5.81) and (5.82), we have the steady state value of the real interest rate:
\[ \frac{R}{\Pi} = \frac{\mu_\pi^+}{\beta} \]  
(5.121)

Since the shadow price of capital, \( q \), becomes 1 in steady state from (5.84), we can derive the real rental price of capital in steady state from (5.85).
\[ p^k = \frac{1}{1 - \tau^k} \left[ \frac{\mu_\pi^+ \mu_\psi}{\beta} - (1 - \delta) \right] \]  
(5.122)

**Numerical Calculation**

Since we cannot derive explicit solutions on the rest of the endogenous variables, we employ the following numerical calculation method.

Let \( X^{(j)} \) denote the value of \( X \) at the \( j \)-th iteration. Given \((r^k)^{(j)}, (g^c)^{(j)}, (g^i)^{(j)}\) and \(H^{(j)}\), we can derive the following solutions: From (5.70) and (5.120),
\[ w^{(j)} \leftarrow \left[ \frac{\alpha (1 - \alpha)^{1 - \alpha}}{\lambda (r^k)^{(j)}} \right]^{\frac{1}{1-\alpha}} \]  
(5.123)
From (5.71),
\[
\left( \frac{\tilde{k}}{H} \right)^{(j)} \leftarrow \mu_{z+} \mu_{\psi} \frac{\alpha}{1 - \alpha} \frac{w^{(j)}}{(r^k)^{(j)}}
\]  
(5.124)

From (5.96),
\[
(k^g)^{(j)} \leftarrow \frac{(g^i)^{(j)}}{1 - \frac{\delta}{\mu_{z+} + \mu_{\psi}}}
\]  
(5.125)

From (5.77) and (5.83),
\[
i^{(j)} \leftarrow \left( 1 - \frac{1 - \delta}{\mu_{z+} + \mu_{\psi}} \right) \left[ \left( \frac{\tilde{k}}{H} \right)^{(j)} H^{(j)} \right]^{\frac{1}{\alpha g}}
\]  
(5.126)

From (5.103) and (5.120), the zero excess profit assumption requires \( \Theta = (\lambda - 1)y \). Substituting this into (5.100), we have
\[
y^{(j)} \leftarrow \frac{1}{\lambda} \left[ \left( \frac{\tilde{k}}{H} \right)^{(j)} \right]^{\alpha} H^{(j)}.
\]  
(5.127)

Then, from (5.101),
\[
c^{(j)} \leftarrow y^{(j)} - \left[ i^{(j)} + (g^c)^{(j)} + (g^i)^{(j)} \right]
\]  
(5.128)

From (5.79),
\[
c^{(j)} \leftarrow c^{(j)} + \nu_g (g^c)^{(j)}
\]  
(5.129)

From (5.93)
\[
z^{(j)} \leftarrow \left( \frac{1}{\mu_{z+}} \right)^{\frac{1 - \theta}{\alpha}} \left( 1 - \frac{\theta}{\mu_{z+}} \right) \tilde{c}^{(j)}
\]  
(5.130)

From (5.94),
\[
L^{(j)} \leftarrow \left[ \frac{1 - \tau_h w^{(j)}}{1 + \tau c \tilde{z}^{(j)}} \right]^{1/\sigma_h}
\]  
(5.131)

Using (5.123)-(5.131), we can derive \((r^k)^{(j+1)}\), \((g^c)^{(j+1)}\), \((g^i)^{(j+1)}\) and \(H^{(j+1)}\) according to the following equations:
\[
(r^k)^{(j+1)} \leftarrow \frac{1}{1 - \alpha_g} \left[ \left( \frac{\tilde{k}}{H} \right)^{(j)} \frac{H^{(j)}}{(k^g)^{(j)}} \right]^{\frac{\alpha_g}{1 - \alpha_g}} p^k
\]  
(5.132)

\[
(g^c)^{(j+1)} \leftarrow \left( \frac{g^c}{y} \right) y^{(j)}
\]  
(5.133)

\[
(g^i)^{(j+1)} \leftarrow \left( \frac{g^i}{y} \right) y^{(j)}
\]  
(5.134)

\[
H^{(j+1)} \leftarrow L^{(j)}(1 - U)
\]  
(5.135)
The first equation is derived from (5.71), (5.78) and (5.77). The second and third equations are derived by definitions. Note that \((g^c/y)\) and \((g^i/y)\) are given as the sample means. The fourth equation comes from (5.95). Note that \(U\) is given as the sample mean.

Given \((r^k)\) (\(j\)), \((g^c)\) (\(j\)), \((g^i)\) (\(j\)) and \(H\) (\(j\)), we have explicit solutions through (5.123)-(5.131). However, we cannot derive explicit solutions for \(r^k\), \(g^c\), \(g^i\) and \(H\), since we face highly non-linear simultaneous equations described as (5.132)-(5.135). Therefore, we have to rely on the following numerical calculation method:

1. Guess the initial values of \((r^k)\) (\(1\)), \((g^c)\) (\(1\)), \((g^i)\) (\(1\)) and \(H\) (\(1\)) and set \(j = 0\).
2. Calculate (5.123)-(5.131).
3. Derive \((r^k)\) (\(j+1\)), \((g^c)\) (\(j+1\)), \((g^i)\) (\(j+1\)) and \(H\) (\(j+1\)) from (5.132)-(5.135).
4. Calculate \(|\Phi^{(j+1)} - \Phi^{(j)}| < \varepsilon\) where \(\Phi \in \{r^k, g^c, g^i, H\}\) and \(\varepsilon > 0\) indicates a convergence criterion.
5. If the criteria are not satisfied, then set \(j = j + 1\) and go back to the second step. Otherwise, we have numerical solutions on steady states of \(\Phi\) and endogenous variables from (5.123) to (5.131).

**Auxiliary Variables and Endogenously Determined Parameters**

Steady states on the rest of endogenous variables can be derived as follows. It should note that several parameters are determined endogenously.

From (5.80), \(\varphi^c = \left[(1 + \tau^c) \left(1 - \frac{\theta}{\mu_z} \right) \tilde{c}\right]^{-1}\). From (5.92) and (5.93), \(\chi^h = \mu_z^{-(1-v)/v}\). By definition, we have the steady state value of bond from \((b/y) \times y\) where debt to GDP ratio, \((b/y)\), is given as the sample mean. Then, the lump-sum transfer is determined from government budget constraint:

\[
\tau = g^c + g^i + \left(\frac{R}{\mu_z + \Pi} - 1\right) b - \left(\tau^c c + \tau^h w H + \tau^k p^k \frac{k}{\mu_z + \mu_w}\right)
\]

From (5.87)-(5.91), \(\Pi^w = \Pi^w_\kappa = \mu_z + \Pi\) and \(A_H = [w \varphi^c(1 - \tau^h)/[\lambda_w \chi^h \Pi^w_\kappa]]\). Thus, the scaling parameter, \(A_H\), is endogenously determined. In our specification on the utilization cost function, \(\alpha'(1) = \sigma_b\). Then, (5.86) requires \(\sigma_b = p^k\) to be satisfied. Since \(p^k\) is given by (5.122), the private capital adjustment cost parameter, \(\sigma_b\), is also endogenously determined. Finally, fixed costs on goods and capital production are also determined endogenously as follows: \(\Theta = (\lambda - 1)y\) and \(\Theta^k = \alpha_g \frac{\rho^k}{\mu_z + \mu_w k}\) from (5.103) and (5.104).
Chapter 6

Conclusion

This thesis extended the estimation methods (chapters 2 and 3) and analysed the business cycle in the U.S. and Japan (chapters 4 and 5) based on the empirical DSGE approach. The first part of this thesis extended estimation methods (1) to introduce measurement errors and (2) to make use of a large number of data information. In the second part, we apply the empirical DSGE approach (1) to examine the sources of the Great Recession in the U.S. and (2) to measure the impact of the government spending on unemployment in Japan. This chapter outlines the research questions, methods, results, conclusions and summarizes the future directions of the research.

6.1 Chapter 2: Role of Measurement Error

6.1.1 Research Questions

The SW model, current standard DSGE model, explained inflation and wage data, heavily relying upon ad hoc “markup shocks” difficult to add structural interpretations. The research questions in Chapter 2 are as follows: When components unexplained by the model (measurement errors) are separated from the data, how much the SW model can explain inflation and wage data? Rather than forcing the model to explain all the data, may introducing measurement errors increase the explanatory power of model for data?

The standard estimation method assumes observed data and endogenous variables are definitively connected. However, the data should be regarded as consisting of two components: One is a systematic component and the other is an idiosyncratic component. The former is a component determined by the interdependence relationship between macroeconomic variables. We agree to describe this component by a structurally interpretable DSGE model, and we call the component as “endogenous variables (or model variables)” such as output gap, inflation gap, etc. But the latter is simply noise. The model should not take care of the noise fluctuations (measurement error). So, we need to decompose data into endogenous variable and measurement error.

When decomposing data into the two components, the problem is that both endogenous variable and measurement error are unobservable. Thus, we propose a method to estimate endogenous variables with high accuracy. Then, we estimate the SW model by introducing measurement errors, and examines how much can it be improved to capture volatile inflation and wage data.
6.1.2 Methods

We propose a generalized estimation method of DSGE model with measurement errors by employing the hybrid MCMC method with the simulation smoother. The method proposed by previous studies has a problem that sampling is interrupted by generating a non positive definite matrix in drawing variance covariance matrix of endogenous variables according to standard smoothing method. Instead, we can overcome the problem by adopting the simulation smoother as a smoothing method of endogenous variables. Then, we estimate, on Japanese data, the SW model with the measurement errors and compares it with the result of the SW model without measurement errors.

6.1.3 Results

It turns out that the estimation results of structural parameters, structural shocks, and historical decompositions to be considerably different between the model with measurement errors and without measurement errors. Also, we compare two models using Bayes factor and find the model with measurement errors to be favored very strongly which implies that the model with measurement errors to have significantly higher forecasting power over the model without measurement errors.

6.1.4 Conclusions

Inflation and wage data show noisy behaviors in recent years and predictions are becoming difficult. The SW model captured the high frequency fluctuations by markup shocks but the shocks cannot be interpreted structurally and the fit of model for data significantly reduces. Given our results, after removing measurement errors, inflation and wage data should be explained by structural shocks. In other words, noise should be explained by measurement errors.

6.1.5 Future Research

Our results showed that the standard DSGE model still cannot explain inflation and wage data structurally. According to microeconomic evidence, there is a consensus that markup and output are countercyclical. One of the important tasks is to construct a structural model that simultaneously satisfies the countercyclical relationship while explaining the volatile movements of inflation and wage data.

6.2 Chapter 3: Estimation in a Data Rich Environment

6.2.1 Research Questions

Boivin and Giannoni (2006) integrated DSGE model with DFM as a data rich approach and they show the estimation accuracy of endogenous variables, especially inflation, is improved by using a lot of data information.

The research questions in Chapter 3 are as follows: First, when applying the data rich approach to Japan, how much shrinking will be the posterior distributions of structural shocks? Will the estimation accuracy of inflation gap increase, similar to Boivin and Giannoni (2006) in the U.S.?

Second, as applied empirical research, when monetary policy rule is estimated using highly accurate inflation gap, how will the Taylor coefficients change compared to the results of previous studies? Even if we add data, will productivity shock still explain well the Japans lost decade?
How is the monetary policy shock contributing to the explanation of the business cycle around the 1990s?

6.2.2 Methods

The standard approach matches one endogenous variable with one endogenous variable with many data. We estimate the SW model by the data rich estimation method using up to 55 Japanese data and aims (i) to examine how much estimation accuracy of structural shocks and endogenous variables improves by utilizing a lot of data and (ii) to consider the sources of output and inflation fluctuations. Then, we compared the results of the data rich approach and the standard estimation approach.

We adopt the hybrid MCMC as with Boivin and Gianonni (2006), but our method has a difference: We estimate the posterior of not only endogenous variables but also structural shocks by employing the simulation smoother.

6.2.3 Results

By utilizing multiple data, five endogenous variables out of sevens has been accurately estimated, i.e., the credible intervals become more shrink than the standard estimation approach. Inflation gap has been estimated with high accuracy, and output gap also does. Variations in endogenous variables are brought about by structural shocks. As expected, estimation accuracy of five structural shocks out of sevens has been also improved. Especially, the shocks with improved estimation accuracy included both of supply shock (TFP shock) and demand shock (monetary policy shock) that affect variations in output gap and inflation gap.

On the applied side, we find the estimated Taylor coefficients are stable regardless of estimation methods. The effect of monetary tightening policy is depressing the output and consumption up to 0.3% (investment 0.6%) about half a year after 1% rate hike. Variance and historical decompositions show investment adjustment cost shock plays a significant role in explaining the business cycle in Japan, not TFP shock. Most of variations in inflation were regarded as measurement error and the remaining fluctuations were mainly due to labor supply shock.

6.2.4 Conclusions

Is the data rich estimation method with measurement errors useful? Can the method improve estimation accuracy? Basically, “Yes”, but there is one condition. It was revealed that highly accurate endogenous variables (including inflation) and structural shocks can be estimated by matching common factors extracted from a large number of data with model concepts (endogenous variables).

On applied empirical analysis, when the monetary policy rule was estimated with high accuracy, did the Taylor coefficient change? No. It was extremely stable in Japan regardless of estimation methods. The central bank responded with a rate hike of 1.6% against 1% inflation. Was TFP shock the main source of the business cycle in the data rich estimation method? No. The main sources are preference shock, investment adjustment cost shock and labor supply shock. It was positive preference shock that towed the boom. Also, negative investment adjustment cost shock was supporting the economy for two years after the collapse of the bubble. Both are demand shocks, not supply shock. Has the monetary authority implemented policies that promote the bubble economy or trigger the collapse? No. The monetary policy contributed to stabilization according to the historical decomposition.
As a reservation condition for this method to be useful, additional data should be selected after explicitly considering the endogenous variable and data linkage in the structural model.

6.2.5 Future Research

The data rich estimation method has the result that the estimation efficiency decreases when we add the data not necessarily corresponding to endogenous variables. One of the tasks is to verify what the estimation accuracy was lowered and what kind of data should be selected not by actual data but by Monte Carlo experiment.

6.3 Chapter 4: Sources of the Great Recession

6.3.1 Research Questions

Chapter 4 examined the sources of the Great Recession in the U.S. Especially, we addressed the following questions: What shocks triggered in the Great Recession? How deteriorated bank and corporate balance sheets affected the recession? What kind of channels did the damage to the balance sheets have caused negative effects on real activities? Did we face with "bad luck" that was hit by big negative shocks not occurring in peacetime? Did there be any further "bad luck" that adverse shocks would expand their own volatilities? Was "good policy" implemented in the recession?

6.3.2 Methods

To make clear the channel affecting the business cycle by balance sheet loss, we first extended the standard DSGE model by embedding financial frictions in both banking and corporate sectors. The model installed two types of agency problems between borrowers and lenders: One is the agency cost between corporate sector (borrower) and bank sector (lender) due to asymmetric information. The other is the agency cost between banks (borrower) and depositors (lender) caused by bank’s moral hazard/costly verification problem.

To examine the possibility of big shocks that would not occur under normal times, we introduced time-varying volatilities into structural shocks (stochastic volatility model: SV). We also considered the effect that the bad shock itself expands the magnitude of its own shock (leverage effect). In addition, to improve the estimation accuracy of the shocks, we adopted the data rich estimation method utilizing 40 data including multiple financial data to identify net worth shocks.

6.3.3 Results

There are six main findings: (1) Both bank and corporate net worth shocks show negative spikes during the financial crisis. The peak time of the negative bank net worth shock was at 2008:Q3. Then, corporate net worth shock had followed it. (2) Volatilities of both bank and corporate net worth shocks and monetary policy shock rapidly increased during the recession. (3) The leverage effect cannot be detected in both net worth shocks. (4) Monetary policy shock had two big negative spikes: 2007:Q4 and 2008:Q3. (5) Bank net worth shock had contributed greatly to the economic recovery right after the recession period. (6) The main sources of the business cycle are different according to estimation methods. Especially, in the standard estimation method, a decline of TFP shock was the source of the recession, but in other cases with SV or the data rich estimation method, an adverse bank net worth shock triggered the recession.
6.3.4 Conclusions

According to our results, the replies to the research questions are as follows: What shocks triggered in the Great Recession? The trigger was deteriorating net worth of banking sector (2008:Q3). The worsening of the net worth of the corporate sector followed it. How deteriorated bank and corporate balance sheets affected the recession? The rise in banking leverage ratio and the rise in corporate borrowing rate increased the borrowing constraints of both sectors and reduced investment and output. What kind of channels did the damage to the balance sheets have caused negative effects on real activities? The channels are shock amplification effects due to agency costs by asymmetric information and moral hazard. The spread equivalent to four basis points was added to the corporate borrowing rate against 1% increase in corporate leverage.

Did we face with "bad luck" that was hit by big negative shocks not occurring in peacetime? Yes. The volatilities of the negative net worth shocks of bank and corporate sectors were greatly expanding at that time. Did there be any further “bad luck” that adverse shocks would expand their volatilities? No. The leverage effect was not detected. Was "good policy" implemented in the recession? Yes. Since two large negative spike in monetary policy shock and expansions of its volatility were observed, the central bank implemented bold monetary easing during the recession. Also, if the improvement in the balance sheet of the bank sector after the recession was due to TARP (Troubled Asset Relief Program), fiscal and monetary authorities played a major role in the economic recovery.

Finally, introducing time-varying volatility and estimating with a lot of data are important. Because even if the model is the same, the sources of the business cycle changes accordingly. If we misjudge the sources, there might be the possibility of bringing wrong policy responses.

6.3.5 Future Research

We observed that the corporate net worth shock occurred just shortly after the bank net worth shock. This timing pattern may suggest the possibility of endogenous relationship between the balance sheet conditions of the banking and the corporate sectors. For instance, it is possible for the corporate sector to hold financial sector’s equity as an asset and the devaluation of the financial sector’s asset may affect the balance sheet condition of the corporate sector. However, our model does not allow the corporate sector to hold banking sector’s equity as an asset (in our model, corporate sector is assumed to hold the asset fully in the form of physical capital). Thus, one of tasks is to expand to a model where banks and corporates hold each other’s assets.

6.4 Chapter 5: Impacts of Government Spending on Unemployment

6.4.1 Research Questions

Does government spending improve unemployment? Surprisingly, contradictory results are reported against this simple question. Chapter 5 examines the quantitative effect of government spending on unemployment in Japan. The question is quite simple: Does government spending improve unemployment, if so, how big is it?
6.4.2 Methods
First, as a channel for fiscal stimulus to create demand, we focus on non-wasteful government spending that government expenditure directly affects household utility and firm’s productivity. Second, we consider the so-called spending reversal rule which cuts government expenditure with the increase in debt outstanding. Third, we examine the quantitative effect of government spending on unemployment, relying not upon calibration but upon estimation. So, we adopt a simple way to introduce unemployment based on workers’ market power.

6.4.3 Results
There are four findings: (1) 1% increase of government consumption reduces unemployment by 0.18%. (2) 1% stimulus arising from government investment brings an improvement of unemployment by 0.06%. (3) Since the late 1990s and especially after the financial crisis of 2007-2008, the positive government consumption shock plays a significant role in the reduction of unemployment rate by more than 1.5%. (4) We also find the “crowd-in” channel where fiscal stimuli induce private consumption and investment to increase does not have much influence on unemployment variations.

6.4.4 Conclusions
Does government spending improve unemployment? Yes. How big is it? 1% increase of government consumption reduces unemployment by 0.18%. 1% stimulus arising from government investment brings an improvement of unemployment by 0.06%. But the channel is the traditional one. Complementarity between government consumption and private consumption is so small. Hence, the crowd in effect on government consumption does not work to improve unemployment.

6.4.5 Future Research
There are remaining issues as follows: First, several papers suggest different methods to introduce non-wasteful government spending under alternative specifications of utility and production functions. Thus, we should check the robustness of our result to ensure that the channels of non-wasteful fiscal expansions do not play important roles for variations of unemployment. Second, we introduce not only non-wasteful government spending but also tax rules. When we attempt to simulate the effect of a tax increase, we should estimate parameters, considering anticipated shocks, i.e. “news” shocks. Third, we should reconsider the validity of modeling unemployment based on the market power of workers. While our model predicts higher market power of workers will raise the unemployment rate, Christiano (2011) pointed out that the hypothesis is rejected in Japan. Thus, it should be necessary to model the labor market more richly, for instance, by embedding search and matching friction where firms have some bargaining power.
It is about eight years since I started studying the DSGE model. This doctoral thesis summarizes the results of this period. In the meantime, I received great influence and support from various researchers.

First of all, I would like to thank the co-authors. Hirokuni Iiboshi (Professor, Tokyo Metropolitan University) is a teacher of Bayesian statistics for me. It was my greatest fortune that he advised me on the concept, methodology and how to apply it. Shin-Ichi Nishiyama (Professor, Kobe University) is a teacher of the DSGE model for me. I still strongly remember a DSGE model's wonderful seminar he provided at the Cabinet Office, when he was still working at the Bank of Canada. Ryoichi Namba (Research fellow, Chubu Region Institute for Economic and Social Research) is the one who made the estimation work all night together. I was greatly influenced from him how to face research in honesty. Ryo Hasumi (Research fellow, Japan Center for Economic Research) is a genius. In addition to a deep understanding of both the theory and empirical aspects, he has extraordinary coding capability of whatever software, it was very awesome.

Next, I would like to thank for professors of the dissertation committee, Akiomi Kitagawa (Committee Chair, Professor of Tohoku University), Jun Nagayasu (Professor, Tohoku University) and Michio Suzuki (Adjunct Associate Professor, Hitotsubashi University) who provided beneficial comments for my dissertation. I would also like to thank Jiro Akita (Professor, Tohoku University) who gave valuable comments at the proposal of my doctoral thesis.

Experts of the DSGE model so far gave me numerous useful advice and guidance. In particular, from Ippei Fujiwara (Professor, Keio University and Australian National University), Kengo Nutahara (Professor, Senshu University) and Tatsuyoshi Okimoto (Associate Professor, Australian National University), I received very useful comments on the papers in this thesis at the conference and workshop. I would like to thank them for their special kindness.

In addition, many expert comments at various seminars have deepened my understanding of the DSGE model. I would like to thank for Akira Momota (Professor, Ritsumeikan University), Hiroki Arato (Associate Professor, Tokyo Metropolitan University), Junhyung Ko (Professor, Aoyama Gakuin University), Kazuki Hiraga (Associate Professor, Tokai University), Keiichi Morimoto (Associate Professor, Meisei University), Masaru Inaba (Professor, Kansai University), Masataka Suzuki (Associate Professor, Yokohama National University), Ryoji Ohdoi (Associate Professor, Tokyo Institute of Technology), Shiba Suzuki (Associate Professor, Seikei University), Shuhei Aoki (Associate Professor, Shinshu University), Takeo Hori (Professor, Tokyo Institute of Technology), Toshikane Watanabe (Professor, Hitotsubashi University), Yoshihiro Otsuka (Associate Professor, Tohoku Gakuin University) and Yoshiho Funashima (Associate Professor, Tohoku Gakuin University).

Furthermore, I would like to thank researchers I met at the Cabinet Office, Ayano Sato (Professor, Takasaki City University of Economics), Daisuke Nakamura (Cabinet Office, Consumer Affairs Agency), Fumihiko Suga (Assistant Professor, Kyushu University), Fumihira Nishizaki (Cabinet Office), Haruki Seitani (Cabinet Office), Junya Hamaaki (Associate Professor, Hosei University),
CHAPTER 6. CONCLUSION

Kohei Fukawa (Cabinet Office, OECD), Koichiro Iwamoto (Associate Professor, Aichi Toho University), Koiti Yano (Associate Professor, Komazawa University), Masaaki Maruyama (Cabinet Office), Masahiro Hori (Cabinet Office), Naoki Takayama (Cabinet Office), Takashi Sakuma (Professor, Nihon University), Takayuki Ogawa (Associate Professor, Osaka University of Economics), Takeshi Niizeki (Lecturer, Ehime University), Tetsuro Sakamaki (Cabinet Office, Ministry of Finance Policy Research Institute), Yasuharu Iwata (Cabinet Office, OECD), Yasushi Okada (the deceased, Research Fellow of Cabinet Office at that time) and Yasuyuki Iida (Associate Professor, Meiji University). And I would like to express my gratitude to Michel Juillard and Stephen Adjemian (CEPREMAP, France) who jointly held the DSGE workshop for seven years (ESRI-CEPREMAP Joint Workshop). In this workshop, it was very meaningful for me to exchange views on DSGE models as policy tools in policy institutions and researchers such as FED, ECB, BOC, IMF, BOJ and MOF.

Of course, I would like to thank my teachers of economics itself, Akihiko Kaneko (Associate Professor, Waseda University), Akira Hibiki (Professor, Tohoku University), Daisuke Shimizu (Professor, Gakushuin University), Hajime Takatsuka (Professor, Kagawa University), Hideki Konishi (Professor, Waseda University), Kuniyoshi Saito (Associate Professor, Meiji Gakuin University), Noriaki Matsushima (Professor, Osaka University), Shinichi Takeda (Associate Professor, Takushoku University), Takanori Adachi (Professor, Nagoya University), Takeshi Ebina (Associate Professor, Shinshu University), Takumi Naito (Professor, Waseda University), Tetsuya Shimane (Assistant Professor, Tokyo Institute of Technology), Toshihiro Matsumura (Professor, University of Tokyo), Ryuichi Tanaka (Professor, University of Tokyo), Yoichiro Higuchi (Professor, Tokyo Institute of Technology) and Yoshiyasu Ono (Professor, Osaka University).

Finally, I would like to thank my family for cheerfully supporting me over the years, father Ryuzo, mother Keiko, and my younger sister Sayaka.
Bibliography


