ESSAYS ON THE EVALUATION AND ESTIMATION OF THE INFORMATION FRICTION IN A DSGE MODEL

By

Jen-Yu Chou

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

Economics Section of Cardiff Business School
Cardiff University

September 2017

Abstract

Using US quarterly data (i.e., real-time data and survey data respectively) from 1969 to 2015 through two different estimation approaches (i.e., Bayesian estimation approach and indirect inference estimation approach) to investigate the empirical performance of the standard reduced-form New-Keynesian Dynamic Stochastic General Equilibrium (DSGE) model under the condition without (i.e., full-information rationality) and with inattentive features (i.e., sticky information and imperfect information data revision), we find some consistent results. Firstly, the model of sticky Information is detected to be the preferred model to fit the real-time data behavior. Secondly, the model with sticky information is the only one can generate delay response, which is matching the evidence observed in actual data and in line with most consequences from the previous studies. Thirdly, the imperfect information data revision model performs better when we substitute the real-time data with the survey data, through which we can deduce that the survey data contains extra information to help improve imperfect information data revision model’s performance. Three main contributions are made in this thesis. The first contribution is the estimation and comparison of different types of inattentive DSGE model (sticky information versus imperfect information data revision) for US small-closed economy through Bayesian approach using the US quarterly data (i.e., real-time data and survey data) representing the main macroeconomic time series from 1969 to 2015. What the second contribution is that through comparing different inattentive New-Keynesian DSGE models basing on the full structure (relative to the single equations competition), we inspect which way of inattentive expectation is closer to the way that people form their expectation in real economy. Besides, the thesis adopts Indirect Inference approach as the robust check methodology, which delivers a new way to assess inattentive macroeconomic models, which is the third contribution.

Acknowledgments

Firstly, I want to say appreciate to my first and second supervisors, Dr. Joshy Easaw and Professor Patrick Minford for providing me with guidance, encouragement and much more. I have benefited from every time when I have a conversation and meeting with each of them. Without their support, and kind sharing of their knowledge, I would be not able to reach this stage and finish this Ph.D. journey.

Especially thanks to my first supervisor Dr.Joshy Easaw again for his generosity and patient what support me to overcome the heavy pressure and encourage me forward to finally finish my PhD thesis.

I would need to say thank you for all the members of the department staff at Cardiff Business School who supported me generously in many ways. I would also like to say many thanks to all my dear friends and course-mate in Cardiff, Dr.Zhirong Ou, Yue Gai, Xinran Zhao, Xue Dong, Yao Yao ,Xingchen Li, Ruimin Li and many other friends and colleagues in China, Taiwan and the UK whose names are not mentioned here due to space consideration.

Lastly, I have to say thanks to my dear family, my dear father Mr. Hong-Jen Chou, my mother Mrs Yue-Wei Chou, my two dear brothers Juin-Yu Chou and Juin-Cheng Chou, and my partner Mr Yirui Su. Without their continuous support and encouragement, it is impossible for me to have this courage to take and overcome this challenge.

Table of Content

List of Tables

List of Figures

General Introduction

Background

Expectation is important for economic agents in terms of making economic decisions, because they have to face various cases requiring such decisions in real life. For example, how to balance their consumptions and savings and what price to set, etc. Even now, concerning how economic agents forming their expectations, scholars do not have a unified model to explain the process. The New-Keynesian framework, which was characterized by full-information rationality assumption and the 'extreme-sticky' prices, has been proposed to solve this issue in some resent works (e.g., Calvo, 1983). It revealed the essential factors to understand the dynamics of the real world, such as imperfect competition, price rigidities, but there were some arguments about its fail to explain some facts observed in actual data. For instance, as Jeff Fuhrer and George Moore (1995) argue, the monetary policy shock who has a delaying and gradual impact on inflation cannot be explained by the original New-Keynesian type model. Mankiw and Reis (2002) demonstrate that the postponed reaction to monetary shock on inflation cannot be produced without any information friction (inattentive feature) or the price indexed its counterfactual hypothesis.

Thus, two alternative expanded models based on the New-Keynesian framework emerged in recent decades to solve the problems which cannot be explained by the original New-Keynesian type model. Among them, the first is sticky information model which was defined by Mankiw and Reis (2002, 2007). According to their assumption of sticky information, there is a delay in the spreading procedure of the information of macro-economic conditions. The lagged spreading through the population may be caused from two aspects: the cost of re-optimized information and the cost of requiring information.

Due to its rapid growth in recent years, this approach was successfully applied to explain economic behavior. For instance, Reis (2006a, 2006b) asserts with the belief of inattentive expectation hypothesis that economic agents choose to use the updated information only when the expected benefit from the newly arrived information is higher than the cost of it. For example, we postulate assume that in period t there is a proportion of the people from the supply side will absorb the current-period information, and meanwhile the rest (proportion) of people keeps the opinion that they are preserving in period t-1 of period t+1's inflation rate. Thus, different from the full-information expectation, the current inflation depends on not only this period t's expectations of the future inflation but also the past expectation of the future inflation rate. The other one is the imperfect information (data revision) model (Woodford, 2001, 2003; Aruoba, 2008; Casares and Vázquez, 2016). The imperfect information agents refer to who constantly update their own information sets under the premise that never can they fully observe the real state. According to this, they form and renovate their beliefs regarding the underlying economic situations accompanied with the problem of signal extraction (Woodford, 2001, 2003). We take data revision as a solution of the signal extraction problem, which indicates that imperfect information agents through two ways of using data revision process to reduce noise and incorporating entire of the involved information to figure out the real situation of economy to reach the same goal which is forming their expectations. Thus, the current state not only depends on not only the final revised observations but also the initial released observations. The details of the data revision processes will be well stated in the Chapter 2. The way or definition of data revision process borrowing from Casares and Vázquez (2016) and Vázquez et al. (2010, 2012) will also be well clarified in chapter 2.

Inspired by the previous studies, three questions will be investigated in this thesis. First question is whether the inclusion of inattentive features can help the original

New-Keynesian DSGE model to replicate some important stylized[1] facts better. The second question is whether the inclusion can help to give a better overall performance. The third question is whether the different types of inattentive feature are distinctive in explaining the dynamics of the observed actual data. To discover the answers of the questions above, three rivals will be selected: the model with full-information rational expectation, the sticky information expectation model, and the imperfect information data revision expectation model. Each of them will be evaluated through two methodologies: Bayesian estimation method and Indirect Inference estimation method within this thesis.

Model Evaluation Methodology

The Bayesian Estimation Approach to Evaluate the New-Keynesian DSGE type models

Bayesian estimation has been implemented as a relatively 'strong' econometric estimation method by some recent studies (Geweke, 2006; An and Schorfheide, 2007). Where it is superior to the 'weak' econometric estimation methods should be the capability of embodying all the features and implications of the model in the estimating procedure, yet the 'weak' ones, for instance, the calibration methods are only can reproduce some chosen moments of the observed variables through simply assigned values to parameters.

Bayesian estimation method is catalogued under the group of 'strong' interpretation. To be more specific, through the comparison of the Bayesian estimation with the classical maximum likelihood estimation, it is easy to conclude that in the perspective

[1] The persistence property of output and inflation, and the delay effect of monetary policy shock on inflation. And such stylized facts are taken as serviceable norms what assistant to evaluate models. The observed hump-shaped response of inflation to monetary policy shock has been paid attention in these recent years. This is because the fact that this hump-shaped response is not only robust but also hard to be generated in a simple model. Most notably, the New Keynesian Phillips curve which is basing on the assumption that firms face expense to adjust price is not able to reproduce such a response without any information rigidities (Mankiw and Reis, 2002).

of the working step most of them are similar apart from the last few ones. To be specific, posterior density function is obtained by using Bayesian estimation approach given by the combination of the likelihood function and prior distributions of the model's parameters. Then this optimization of posterior can done concerning parameters of the model. What the most distinguished point between the two methods is that the classical maximum likelihood misses the steps of including the additional prior function to reweigh the likelihood function.

Two general reasons for using Bayesian estimation approach have been discussed frequently in recent studies (Schorfheide, 2000; An and Schorfheide, 2007). The use of prior information come from either the previous relevant studies or the reflection of researcher's subjective perception. So, this method directly builds a link between our study and previous studies. Besides, the Bayesian method can evaluate misspecified models according to the criteria of measurement which are the marginal likelihood and the Bayes' factor. The model's marginal likelihood which is connected to the density function of prediction directly which can be taken as an acceptable criterion to measure the level of overall model fit. The competing models selected by us will be estimated by using the real-time data of US 1969Q1-2015Q4 (survey of professional forecaster data will be used in robustness check).

Concerning the structural parameters and impulse response functions estimated through Bayesian estimation, it shows that the set of estimates for the structural parameters are plausible. For instance, the estimated price stickiness for US economy is considerable, which is in accordance with many previous studies (Smets and Wouters, 2007; Milani and Rajbhabdari, 2012). Besides, the impacts of the three main shocks on the US economy after analyzing are consistent with the existing studies quantitatively. For example, a positive monetary policy shock is along with a rise of nominal interest rate, a decline in output gap, and a decrease in inflation (Peersman and Smets, 2002). Moreover, the positive cost-push shock has positive impacts on inflation and nominal interest rate, but impact on output gap negatively.

Besides, a positive demand shock has a positive effect on the output gap.

From the perspective of the overall model fit estimated through Bayesian estimation using real-time data, the results show that the inclusion of inattentive features has significant effects on the mode’s ability in fitting macroeconomic time series. To be specific, the inclusion improves the model's ability to explain the real world, which is in line with the suggestions from most of the related literatures (Mankiw and Reis, 2002, 2007; Collard et al., 2009). Apart from that, we find that the model achieves its best fit under sticky information model through Bayesian estimation. By using diffuse prior distribution, different specifications of Taylor rule, and different periods of lag information in sticky information model (j=4, 6 and 8) in the robustness check, we draw a conclusion that none of them can change the ranking among the three rivals. Surprisingly, when we use survey of professional forecaster data instead of real-time data to evaluate the models’ performances, although the models with inattentive still be superior to the baseline model, the rank between the two inattentive models’ changes.

The Indirect Inference Approach to Evaluate the New-Keynesian DSGE type models

First Stage: Calibration-based Indirect Inference Test

Current studies attempt to formalize the test method to evaluate model’s performance in an absolute sense relying on Indirect Inference. The Indirect Inference as a testing method utilizes that the solution of the log-linearized DSGE model is able to be expressed by a restricted Vector-Autoregressive-Moving-Average (VARMA) model and can be closely expressed by a Vector Autoregressive (VAR) model. Indirect inference test can be understood as a process that through comparing the simulated-data-based unrestricted VAR estimates with the alternative actual-data-based unrestricted VAR estimates, after then we can confirm whether these two sets of the auxiliary models’ estimated parameters (i.e., VAR) are ‘close enough’ (i.e., each competing DSGE model is correctly specified)

While conducting Indirect Inference test, we employ Wald test on VAR estimates. In general, Indirect Inference testing procedure contains three general steps. The first step, to construct the errors implied by the actual data and one of the model of the previously estimation-based and calibration-based structural models. In the second step, the innovations of structural errors are bootstrapped to be employed to produce the pseudo data which are based on candidate model. After that, an auxiliary model (i.e., VAR) is fitted to each set of pseudo data and the sampling distribution of the coefficients of the auxiliary VAR model. In the final step, the Wald statistic is calculated to judge whether or not the functions of the parameters of the auxiliary VAR model estimated on the actual data lie within the confidence interval allusive by the sampling distribution. According to the results through Indirect Inference calibration-based test, none of the three competing models can pass the test. Comparing with the previous studies argue in the literature, the performance of the model with imperfect information data revision is much worse than that of the baseline model, which is contradict to the conclusion from the Bayesian approach.

Second Stage: Estimation-based Indirect Inference Test

The Indirect Inference has been long-standing applied (Gregory and Smith, 1991, 1993; Gourieroux and Monfort, 1993). As far as we concern, when Indirect Inference is applied for evaluating model, structural model's parameters are provided at the beginning. However, the nature of fixed calibrated parameters leads it an overly strong condition for testing models and contradistinguishing one model from another. Seeing the values of parameter of the candidate model could be estimated or calibrated within a permissible range throughout the theoretical structure of the model, it is probable for a rejected model with the presumptive set of parameters to pass the test when it with another set of parameters. To have a fair result of the testing, it is necessary for investigators to find a set of 'good' structural parameters. Thus, we estimate the models to get the optimal sets of parameters before the evaluating process.

The general working steps of Indirect Inference estimation-based test are summarized as follows and which are similarly and common to those mentioned in the previous studies (Le et al., 2011, 2013[2], 2016; Minford and Ou, 2013; Liu and Minford, 2014): Firstly, to select an auxiliary model (e.g., VAR) to estimate it based on the actual data to achieve the benchmark estimates. Secondly, give presumptive values to structural parameters which are needed to be estimated, after which the parameters will be used to create numerous pseudo samples of simulated data with the investigated theoretical model. Thirdly, to estimate the selected auxiliary model derived by the simulated data obtained from step two, which is done to produce the joint distribution of the selected estimates (from the first step) so that we can have the mean of this distribution. In the fourth step, we compute the Wald statistics and the transformed Wald statistics (normalized t-statistics)[3] to measure the distance between the benchmark estimates achieved in the first step and the mean of the estimates achieved in the third step. Finally, the second step to the fourth step will be duplicated until the minimum of Wald statistic is achieved.

It is obvious that the process of the second stage through Indirect Inference is similar to that of the first stage, apart from the last step. The reason for the distinction is the purpose of the second stage which aims not only to gauge the gap between the to-be-examined model and the actual data but also to narrow the gap by searching for an optimal set of parameters under the premise of the theoretical model being true.

[2] One advantage of Indirect Inference over the other method in terms of testing procedure, an alternative hypothesis suitable for testing of the specification of the model can be automatically generated by the unrestricted VAR model based on actual data, which leads us not have to specify different DSGE models as the alternative hypothesis. As a result, the identified VAR derived by the DSGE model is the only factor that required in this testing procedure.
[3] This function of Transformed Wald statistics (normalized t-statistic) is based on Wilson and Hilferty 1983's method of transforming Chi-square distribution into a standard normal distribution calculated.

After re-evaluating the three competing models through estimation-based test with the same US real-time data through the Indirect Inference method, we find that only the model with sticky information can pass the test meanwhile perform no worse than the baseline model. Additionally, among three competing models, only the model with Imperfect Information data revision fails to pass the test. However, when we use US survey data over the same period to evaluate each model, none of them can pass the test.

Overall, there are some consistent results through implementing two different estimation methods. Firstly, the sticky Information model is found to be the preferred model to fit the real-time data behavior which is examined in terms of Wald statistics (Wald percentile) and transformed Wald statistic. Secondly, the sticky information model is the only one can generate delay reaction to monetary policy shock and this is matching the observing evidence in actual data. Thirdly, the imperfect information data revision model performs better when we substitute the real-time data with the survey data, through which we can deduce that the survey data contains extra information to help improve imperfect information data revision model's performance.

Contributions

The main intention in this thesis is to evaluate the available original New-Keynesian reduced-form DSGE model with three different expectation assumptions respectively. The three models, which are taken into consideration, can be categorized into two groups: one is without inattentive features, the other one is including inattentive features. Within the first group which only has one model, we take it with full-information rationality expectation assumption to be the baseline. While in the other one, two inattentive expectation models are contained. They are the model of sticky information (Mankiw and Reis, 2002, 2007) and the model of imperfect information data revision (Vázquez et al., 2010, 2012, and Casares and Vázquez, 2016) respectively.[4]

We carry out the evaluation of the three competing models from three aspects: 1) to assess them through estimated impulse response function; 2) to compare the model-fit through Bayesian estimation approach, in which the relatively performance is determined by the log marginal likelihood or the Bayes factor; and 3) to use Indirect Inference as robust check method to see whether the candidate theoretical model can generate data close to reality. According to this point of view, the analysis in this thesis can be taken as the competing and selecting procedure of empirical models.

There are three main contributions of this thesis. The first contribution is the estimation and comparison of different types of inattentive DSGE model (sticky information versus imperfect information data revision) for US small-closed economy through Bayesian approach using the US quarterly data (i.e., real-time data and survey data) representing the main macroeconomic time series from 1969 to 2015. And the reason why we choose real-time data to estimate each model as Paloviita (2007b) asserts the significance of people's current knowledge and belief in leading their behavior in economic activities. As a result, in some cases, such as policy decision, if the economic relationships can be described potentially, so we can obtain a more precise research result. Se we use real-time data obtainable on the occasion instead of recently. Besides, another kind of data is used in our research in robustness check. Due to people's deficiency in predict the economy, we introduce the Survey of Professional Forecaster (SPF) data to simulate people's reliance on the experts. However, SPF data is not flawless. Its defect may be exposed when there is a big news which is opposite or averse to some experts' expectations, in which case experts may have intention to avoid significant changing of their predictions for maintaining their reputations. Overall, the two kinds of data selected lead us to find

[4] We use small-closed DSGE model instead medium-scale DSGE model different from Miguel Casares and Vázquez (2016); Vázquez et al. (2010, 2012) use the reduced-form model to study the data revision its impact on monetary policy and leave the rest economic agents without involving data revision issues.

the best way to describe people’s expectation formation. Through adopting these two kinds of data to evaluate each model which may provide us more accurate guidance to find the best way to describe people's expectation formation. What’s more, once we find the best way to describe how people form their expectation the government can affect real activity in ways that are correlated with that information (i.e., noisy revision information, sticky delayed information), this should greatly increase the credible range of conducting more stabilized policy.

The second contribution, through comparing the different inattentive New-Keynesian DSGE models basing on the full structure (relative to the single equations competition), we inspect which way of inattentive expectation is closer to the way that people form their expectation in real economy. The third contribution, the thesis is adopted Indirect Inference approach as the robust check methodology, which delivers a new way to assess inattentive macroeconomic models.

The outline of each chapter is demonstrated as follows. In Chapter 1, we survey the literatures on different New-Keynesian type DSGE models including the ones with and without inattentive feature. We also discuss the main findings from previous literatures. In Chapter 2 is the introduction of each competing model. Chapter 3 and Chapter 4 apply the two main analyses to examine three selected competing models respectively. In Chapter 3, we estimate reduced-form New-Keynesian type model without and with inattentive ingredient (sticky Information and imperfect Information data revision) through Bayesian estimation approach; Chapter 4 uses the Indirect Inference as the robust check method to test and estimate each competing model to re-examine the results obtained through Bayesian estimation approach. Chapter 5 contains the conclusion and discussion of further research direction.

Chapter 1

Whether Different Inattentive Features Matter for Economy Dynamics?

1.1 Introduction

The role of people's expectation in determining aggregate outcomes of the macro economy, such as inflation dynamics and the business cycle, has often been discussed and well established. However, the study involves how people form their expectation is relatively rare and less well studied. One recent study by Milani and Rajbhandari (2012) compares the full-information rationality New-Keynesian type model with the alternative models that deviate from the full-information rationality.[5] However, this topic is quite important for making the most fundamental macroeconomic decisions, such as the allocation of consumption or savings, how to set the appropriate price and so forth, some of which are underlying macroeconomic dynamics and driven by people's expectation of the future. In the following sections, we survey the literature focusing on the early assumption of fully attentive expectation or full-information rational expectation firstly and explore the weakness of this early expectation assumption. In order to remedy the weakness of full-information rational assumption, another assumption deviating from the full-information rationality has been proposed, which is so-called inattentive expectation assumption. In particular, we mainly focus on two types of inattentiveness, which are the most commonly discussed. The first is the model with sticky information expectation, and the assumption of sticky information is basing on the study proposed by Mankiw and Reis (2002, 2007). The second popular inattentiveness is imperfect information data revision (Aruoba, 2008; Vázquez et al., 2010, 2012; Casares and Vázquez, 2016). Both inattentive assumptions mentioned above will be well stated and discussed in later sections.

[5] Those models are set as being with the allowances of ‘news’ about future shocks, near-rational expectations, learning, and observed subjective expectations from surveys respectively.

1.2 Literature Survey of Classical New-Keynesian type Model without Inattentive Feature: Full-Information Rational Expectation

The full-information rational expectation hypothesis is the starting point of the traditional economic theory. However, a gap between this classical New-Keynesian full-information rational expectation (without any inattentive ingredient, i.e., Calvo, 1983) and the real world has been criticized for many economies. Simon (1989) criticizes the "unrealistic" view of the idea of full-information rational expectations. He argues that regarding the case of economic agents having known all of their problems, choices and possible results, the economic agents could certainly choose the best solution from all alternatives through some reasonable calculation. But in practice, such 'perfect situation' cannot be existent in real world. Besides, some unavoidable constraints always restrict economic agents from making good decisions (e.g., social constraint stemmed from the superior authority of government in terms of legislation or personal constraints originated from limited time and energy). Thus, economic agents have to seek coordination from the aspects of efficiency, profits and other factors. In other words, economic agents cannot simply reach the optimal solution but only reaching the self-satisfied or ‘good enough' solution. As a result, the full-information rational expectation can hardly be applied to explain economic problems.

On the other hand, the implicit hypothesis of full-information rational expectation is that the economic agents are homogeneous. But in real world, economic agents may form different expectations due to their different abilities in information acquisition, absorption, and procession. In other words, not all economic agents hold full information. To sum up, the unrealistic feature of early assumption of full information rational expectation can be showed from two aspects as follows:

1) The full-information rational expectation hypothesizes the economic agents having such full information that can do their best to reach the maximum profit. However, due to people's physical and intellectual capacity limitation, adding to the uncertainties originated from external environment, people are capable to understand and solve complex problems but in a restricted way.

2) Under the assumption of full information rationality, information is a kind of scarce resource that economic agents are willing to try their best to collect all available information to make economic decisions. Despite the desire to acquire information, it did not take the information costs (i.e., costs of accessing required information) into consideration. It is understandable that agents have to pay while collecting the information required for decision making. In practice, it is impossible to get and process information without the payment of time, money, or physical efforts. Due to these potential costs, the number and the quality of information obtained by the economic entities are limited, which lead to the fact that economic agents are impossible to reach the best situation.

To sum up, under the assumption of full-information rationality, economic agents are supposed to clear about the all relevant parameters' value, such as the distribution of shock, the correct structure of the economic model and so on. However, it is an unreasonable assumption in practice because economic agents cannot hold all the information needed to reach the equilibrium of the whole economy (Caballero, 2010). Particularly, when an economy undergoes a big structural transformation such as Great Recession, it will need never implanted policies (Stiglitz, 2011). The tune to full-information rationality hypothesis is favorable according to recent empirical work. Coibion and Gorodnichenko (2012, 2015) strongly deny the legitimacy of hypothesis of full-information rationality. Furthermore, in their paper published in 2012, they clarify that the reason of rejection to full-information rationality hypothesis is not the rationality but the assumption of full information.

1.3 Literature Survey of New-Keynesian type Model with Inattentive Feature: Sticky Information versus Imperfect Information Data Revision

To remedy the unrealistic aspect of the early full-information rationality assumption and deal with the well-known empirical weaknesses (i.e., the delay effect of monetary shock on inflation, persistent of output, and inflation observed in macro data), the New-Keynesian type model with the features deviated from the full-information expectation assumption appears as a modified version.[6] Thus, the inattentive expectation was proposed. As inattentive expectation has different approaches, the two most prominent of them are sticky information (Mankiw and Reis, 2007) and imperfect information data revision (Casares and Vázquez, 2016). These two assumptions will be applied in our research, being different from the sticky information model from Mankiw and Reis (2007) and the imperfect information data revision model from Casares and Vázquez (2016), we use the small-scale closed economy DSGE model instead of medium-scale DSGE to be in line with the baseline model selected.

Although there are weaknesses of the full-information rationality, as recent studies suggest that there is no need to abandon its assumption of rationality or to introduce other types of irrational behavior to help model fit data (Collard et al., 2009; Coibion

[6] There are also some literatures focusing on how to compensate the impractical aspects of the full-information expectation New-Keynesian type models through multiple ways (Rotemberg and Woodford, 1996; Gali and Gertler, 1999; Smets and Wouter, 2003, 2007). In these papers, the most attention is received and focus on real rigidities, such as habit persistence, capital or investment costs, capital utilization, and backwards-looking price setting schemes for the subset of the economic agents (Christiano et al., 2005; Collard et al., 2009). However, Dhyne et al. (2006) argues that backwards-looking price indexation setting scheme cannot support the empirical evidence. The European Central Bank Report pointed out that individual price changes its movement are not consistent with the movement of aggregate inflation. In explaining the observed situation, the idea of reducing controversy that encourages scholars to continue making efforts to resolve this issues in the past few years.

and Gorodnichenko, 2012). Thus, in this thesis, two major inattentive rational models, sticky information, and imperfect information data revision models, are used and compared, and meanwhile, rationality is assumed.

1.3.1 Literature Survey of Rational Expectation condition on Sticky Information

After the year of 2000, the problem of how economic agents forming their expectations of the aggregate economy begins to draw several scholars' attention. To address this issue, Carroll (2003), as one of the funders of this area, introduces the idea of "epidemiological expectations", in which the households form their inflation expectations by receiving the news reports that reflect views of professional forecasters, to explain the origin of the sticky information expectations. According to his study, the slowness of information diffusing through the entire population is due to people's inattentiveness to the arrived information[7].

Sticky information expectation which based on the idea of information slow diffuse through entire population is recommended in many studies. Being one related study of them, Mankiw et al. (2003) research the topic of how disagreement may appear among different agents' expectations of inflation. Their study is distinguished from other researches by finding the ubiquitous heterogeneity of different households' and professionals' inflation expectations. The heterogeneity was derived from different frequencies of the agents updating their information sets. Reis (2006a, 2006b) supports sticky information inattentive assumptions due to the cost of newly arrived information. He asserts that economic agents will only choose to obtain new arrival

[7] Some recent articles have based on Carroll (2003)'s studies to study the implications for monetary policy (Ball et al., 2005) and the dynamics of aggregate economy (Mankiw and Reis, 2007).

information if the expected benefit is higher than the information cost. Later, Mankiw and Reis (2007) develops and analyzes the medium-scale general equilibrium models for the US economy under sticky information assumption. They find that information stickiness exists in all markets throughout the quarterly data from the 1954 Q3 to 2006Q1. Moreover, the information stickiness is especially pronounced for consumers and workers in their study, the feature of information that being slowly disseminated in microeconomic data on price provides more credit to sticky information expectation (Klenow and Wills, 2007; Knotek and Edward, 2010). Mitchell and Pearce (2015) provide direct evidence of sticky information through examining the frequency of revision forecasts for individual professional forecasters. They find that the forecasters do not revise their forecasts usually, which is consistent with the sticky information hypothesis. In most cases, these literatures support sticky information assumptions.

1.3.2 Literature Survey of Rational Expectation condition on Imperfect Information Data Revision

Another strand about people's negligence deviates from the full-information rationality assumption is imperfect information data revision. On the perspective of microeconomic area, imperfect information refers to asymmetric information which is a common characteristic of the imperfect market. However, in macroeconomic area imperfect information means that economic agents are struggling to figure out the actual state of economy. In detail, the definition of imperfect information in terms of microeconomics implies that consumers can be easily fooled by the supply and price. However, under the environment of macroeconomics, imperfect information implies that economic agents involve signal-extraction problem (data revision issue). To be specific, economic agents are disturbed by noises and demand to filter useful signal or information from disturbing noises in observed actual data. The essence of the imperfect information is the inattentive behaviour that the economic agents can constantly update their beliefs, but suffering from the noises, which results the fact that the economic agents cannot fully observe the real state of economy (cannot be fully attentive). Hence, they renew their beliefs about the fundamentals of economy via signal extraction or data revision process to reduce noise.

Imperfect information expectation is recommended in many studies. Woodford (2001, 2003) integrates the idea of people's limited capacity in processing information, imperfect common knowledge, and the monopoly pricing competition to explain the persistence impulse response to real variables. Schorfheide (2005) who allow monetary authority to hold imperfect information (imperfect common knowledge) about the inflation target by modelling economic agent to learn and understand the fluctuating values over time. Although the model under imperfect information catches important periods like the early 1980s' disinflation better, the model under perfect information fits real economic data better. Additionally, Collard et al. (2009) demonstrate that the new Keynesian model under imperfect information environment could produce considerable inertia on an empirically reasonable level.[8]

---

[8] In the study by Levine et al. (2012), regarding the fact that people may not have all information of all state variables and all impacts on the economy, researchers establish a complete structural DSGE model in which the economic agents need to solve the signal-extraction problem to derive the values of state variables and impacting shocks, but such model is mainly governed by habit formation and adaptive learning. Therefore, the endogenous persistence impulse response generated from the model under the assumption of imperfect information the impulse response function generated by the model is close to the real situation. At the same time, they showed an example of analysis of the model under the assumption of imperfect information which fits the economic data well without introducing real rigidities (e.g., habit formation) or indexation price. The setup of our models does not have any interruptions of other features (i.e., habit formation) to check how model itself can reproduce the observed stylized facts.

1.3.3 Differentiate Inattentive Features: Sticky Information versus Imperfect Information Data Revision

The introduction of first inattentiveness is sticky information in Section 1.3.1 which emphasizes the recurring cost of collecting the latest information during making economic decisions, which may lead people updating the information reluctantly for the expense (i.e., cost of processing information) can be higher than its interest. Imperfect information data revision as the second inattentiveness is introduced in Section 1.3.2. It stresses the existence of the noises that influence people’s decision by not reflecting the real state of the economy. Therefore, people via signal-extraction process or data revision process to reduce noises to figure out the real state of economy. Moreover, the model of imperfect information is based on the assumption of economic agents' limitation of information processing, so economic agents’ decisions are determined by the information merely obtained through their information processing channel or communication channel (Sims, 2003).

It may be enquired that why we care about the inattentive feature -- imperfect information data revision. Diebold and Rudebusch (1991) give an example as the best answer from all analytical data revision papers. They explain that the major US economic indicators are doing well in forecasting the recession ex-post only because it is made to explain the past. Its tracking record in real-time, on the contrary, is very poor. Two reasons are given to this contrast. One is that the initial announced data may appear to be very different from the latest announced data. The other one is that the methodology of index changes as time goes by after the real-time indicator failing to forecast the recession. Beyond that, there is another example that easily to be understood to demonstrate this issue. Assuming we use the simple Taylor rule as a monetary policy to remain the level of inflation invariable, when the output gap is negative, the interest rate should decrease. Should th[e interestxxx rate increase, the case would be opposite (i.e., positive output gap). Evidenced by the same token, if the central bank holds economic growth data which is exaggerated before the recession, it would lead to the delay in adjusting interest rate to lessen inflationary pressure after the economic downturn. This example endorses the importance of inattentive feature.

Although a large quantity of literature has suggested to incorporate inattentive features into models to explain the real world, some issues still have not been well discussed. To supplement the areas that omitted by previous papers, our research focuses on verifying the three topics: 1). Do these two inattentive features matter in economic dynamics response; 2) If they are, what are the distinctions between them; 3). Which one can give a better explanation.

1.4 Conclusion and Objectives

In the literatures mentioned above, there are three relevant models which can be divided into two groups, i.e. with and without inattentive ingredients. One of them is the classical 'attentive' expectation model, which is New-Keynesian type model with full-information rationality hypothesis. The second is sticky information model. The third is imperfect information data revision model. Three objectives will be reached through comparing the three models under different conditions.

The first objective of this thesis is to verify whether incorporating inattentive features into the popular reduced-form New-Keynesian model can perform better in replicating the empirical persistence found in macro-economic data than the full-information rationality alternative. The way to measure the performance of the model is to check its ability to generate persistent and delayed responses on output (output gap) and inflation to monetary policy (e.g., Christiano et al., 2005). Moreover, the model simulations will be carried out through Dynare 4.4.3 software.[9]

The second objective is to compare which expectation type model explains the US economy in the best way by using quarterly real-time data (survey of professional forecaster data will be used in robustness check). The process is implemented through Bayesian estimation approach. Through the comparison of Bayes Factor and the comparison of the log marginal likelihood of three competing models, the overall performance according to three rivals under different assumptions (i.e., fully attentive expectation versus inattentive expectation) can be compared and ranked relatively. The first advantage of using Bayesian estimation is that the application of priors which provides a chance to take the previous relevant studies into consideration and it facilitates to reduce identification issues in evaluating DSGE models.[10] The second advantage of Bayesian estimation is that Bayes factor provides an effortless way to evaluate model's relative performance.

The third objective is to use indirect inference to re-evaluate each competing model and make model comparison in an absolute way. Although the Bayesian factor provides a simple way to compare the relative performance of different models, it cannot be used to evaluate model's performance in an absolute way due to its limitation of judging that whether a to-be-examined model itself has a satisfactory performance that can be verified by the actual data. The method of distinguishing indirect inference estimation (estimation-based indirect inference test) from the Bayesian estimation method is to generate a data descriptor that indirectly evaluates the theoretical model by using a completely independent auxiliary model, e.g. VAR.

---

[9] From http://vermandel.fr/dsge-dynare-model-matlab-codes/, provide standard DSGE Models Dynare code, include the simple dynamic three-equation New Keynesian Model.

[10] Due to the structural interpretation of the parameters in DSGE models, sensible proper priors are usually available. These priors may be purely subjective or could reflect data from other sources (e.g., the estimates of structural parameters produced in macroeconometric studies and the estimates based on training sample of macroeconomic data). As the prior information given, Bayesian researchers do not need to worry about the identification issue. However, if a parameter is not identified, the data-based learning about it may be absent and its posterior only gives the reflection of prior information.

The intention of implementing estimation-based indirect inference test is to discover the optimal set of parameters about the actual data in the context of the model to make a fair model comparison.

Chapter 2
Introduction and Establishment of Three Competing Models

2.1 Introduction

The inclusion of inattentive features into macroeconomic model has become an active area of recent research. Carroll (2003) finds that the public's prediction is lags behind the prediction of professionals' through adopting survey of inflation expectation data. The study of Mankiw et al. (2003) their study shows that the disagreement of inflation expectations from survey data is matching the idea of sticky information. Furthermore, regarding to the recent work proposed by Dräger et al. (2013) they found that the impact of information friction on prediction errors at the individual level which provides support for imperfect information assumption (i.e., the economic agents suffer from noisy disturbance).

It is worth noting that, our study is not the first one to make a comparison between alternative expectation models and the full-information rationality type model. For instance, Milani and Rajbhandari (2012), who evaluated the alternatives (e.g., these alternatives include allowed "news" shocks, adaptive learning and observed survey expectations) deviate from fully-information rationality assumption in small-scale New-Keynesian DSGE model. Moreover, they have shown that the econometric characteristics of the model are susceptible to the different formations of expectation. Then our study can be understood as an analysis contributing to the selection of empirical models, which considers inattentive expectation type model as alternatives comparing with the baseline with full-information rationality.

2.2 The Introduction of Three Competing Models

The overview of each of the attentive and inattentive models will be specified as follows. The derivation of each model has been shown in Support Annex and the

Appendix B of Chapter 2. The three competing modes is a reduced-form New-Keynesian type DSGE model for a small-scale closed economy. Three types of agents are constituted the small-scale closed economy which are households, firms, and monetary authorities. The baseline model has been largely applied in previous studies (Milani and Rajbhandari, 2012) is the standard Calvo model without any inattentive features. In terms of the two other rivals, one is the model characterized by sticky information which has been discussed in Mankiw and Reis (2007), and the other one is the model characterized imperfect information data revision which has been constructed by Casares and Vázquez (2016). Being different from those two inattentive expectation model settings we are using the small-scale DSGE model instead of medium-sized DSGE model. Adding additional features might be a useful step (Smets and Wouter, 2003, 2007). However, it may also cause some fundamental issues to blur our main focus. Precisely, when each model being inserted with inclusion of some more new features taken into account, it may potentially distract some attention from the original focus to those new considered features, which leads to the difficulty of assessing the differences between the two inattentiveness (i.e., sticky information and imperfect information data revision). As well as the differences between the baseline model and the models with inattentive features, due to considering so many features.

2.2.1 Reduced-Form New-Keynesian Model without Inattentive Feature: Full-Information Rationality

The derivation of the classical small-closed New-Keynesian model is quite standard in the literature (Woodford, 2003). Here we present a more traditional version of the micro-foundation under the assumption of full-information rationality,[11] The details of

[11] The full-information rationality assumption type model applied in this thesis is chosen without indexation to past inflation and habit formation in consumers' preference, since the

the derivation have been presented in Supporting Annex at the end of this thesis. And the baseline model is as follows:

IS equation： (2.1)

PC equation: (2.2)

Interest rate smoothed Taylor Rule: (2.3)

We have seen from the above presented baseline model, it can be indicated that the aggregate economy under reduced-form New- Keynesian type model with full-information rationality which can be characterized by the dynamics of three main economic variables (i.e., output gap, inflation, and interest rate). The represents output gap, which is a gap between actual output and potential output (i.e., is the output under flexible price economy). The coefficient represents the elasticity of the intertemporal substitution. The new Keynesian Phillips Curve (PC curve) derived under the full-Information rationality assumption is equivalent to the current inflation driven by the expectation of future inflation , current output gap , and the supply shock . The coefficient stands for the time discount factor and is the combined parameter.[12] Interest rate equation that follows the simple 'interest-rate smoothed' Taylor rule (1993). Monetary policy makers set the interest rate basing on simple Taylor rule. The interest rate is driven by the current inflation and current output gap .

premise of indexation has been shown to be not consistent with the microeconomic evidence on price set (Nakamura and Steinsson, 2008). The evidence regarding agents' habit formation is less obvious, but it seems difficult to find supportive evidence through households' consumption data (Dynan, 2000)

[12] Where the composite parameter =0.15 has been taken as fixed and less than one which it implies strategic complementary, to keep it as fixed and less than 1 and in line with the suggestion from previous literature (Woodford, 2001, 2003; Ball et al, 2005). Besides, Woodford (2003) surveys and discusses the existing literature at length and concludes that firms pricing decision should be strategic complements rather than strategic substitutes to allow for potential inflation inertia. And this has been tested in some recent works, for instance, Coibion (2006) these authors when which produce inconsistent results with the actual data.

2.2.2 Reduced-Form New-Keynesian Model with Inattentive Features: Sticky Information and Imperfect Information Data Revision

Before the introduction of the selected inattentive expectation models, we need to clarify the assumption concerning two inattentive expectations respectively. Regarding to the assumption of sticky information, the economic agents update their information sets infrequently due to information costs which reference to the idea offered by Mankiw and Reis (2002, 2007). Distinguished from the conception of sticky information, the conception of imperfect information data revision is that economic agents suffer from the noises, thus they continuously revise their information to extract the useful signals (Aruoba, 2008; Casares and Vázquez, 2016; Vázquez et al., 2010, 2012). In other words, the two different inattentive features can be taken as two distinct information arrivals. One of the principal purposes of this thesis is verifying whether different inattentive features matter in explaining economic dynamics. Furthermore, under the premise of confirming the determinacy of inattentive features, we will explore which feature can explain the US economy better from 1969 to 2015[13].

2.2.2.1 The Model with Sticky Information

The first inattentive feature to be introduced is the sticky information which assumes that some of economic agents use the old information rather than the current arrived information to make the economic decision and form their expectations. Since the cost of previously used information has been paid, there is no extra payment required for reusing old information, which is the way to reduce information costs. The main idea of the sticky information model is that when making economic decisions, due to the cost of acquiring newly arrived information as well as the cost of re-optimization,

[13] In order to construct the revised data in imperfect information data revision model, the sample period actually cover from 1969Q1 to 2016Q4.

only a small percentage of people are willing to use current arrived information to adjust their plans. On the other hand, the rest of people will still use the old information and old plan. The model with sticky information is presented as follows:

IS equation: (2.4)

PC equation: ——— + (2.5)

Interest rate smoothed Taylor Rule: (2.6)

Thus, according to the model with sticky information presented above, the two parameters δ and λ are the shares of updating households and the share of updating firms respectively in any given period (for example if there is no information stickiness of firms then λ=1). To compare with the economic agents in the full-information rational expectation model without inattentive feature, the economic agents are assumed under the premise of sticky information economy update their information sets with certain rate and regarding households and firms respectively (Mankiw and Reis, 2002, 2007; Reis, 2006a, 2006b, 2009). Reis (2006a, 2006b) gives more deep-seated micro-foundations for model features sticky information. The early classical New-Keynesian type model assumes of full-information rationality, which is the case of a pure forward-looking-expectation Phillips curve. However, under sticky information environment, the inclusion of inattentiveness leads to deviation from full-information rationality. The economic agents under this circumstance use the outdated information to form their expectation. Therefore, it yields the Philips curve (PC curve) not only depends on the current expectation but also the past expectation about the future, which is caused by information spreading slowly through the entire population of the economy (Mankiw and Reis, 2002)[14]. When looking into the previous empirical literature, several papers are aiming at comparing Phillips curve derived

[14] Being differentiated from the sticky information PC model of Mankiw and Reis (2002), the current inflation in our New Keynesian three-equation model is determined by both the current expectation and the past expectation of the future inflation rate. In contrast, the current inflation in Mankiw and Reis' model is inferred from flexible price assumption.

under the assumption of full-information rationality and alternative under the sticky-information assumption (Mankiw and Reis, 2002; Coibion and Gorodnichenko, 2012, 2015). However, in this thesis, regarding to the empirical evidence, we are more interested in the simple reduced-form New-Keynesian DSGE type models, rather than that based on single equation (Easaw, et al., 2014; Coibion and Gorodnichenko, 2015). Estimation of comprehensive DSGE models through introducing inattentive feature exists, but there is only a small quantity of papers. The recent papers on this aspect set a benchmark of neo-classical model with flexible prices and introduce sticky information regarding various economic decisions (i.e., consumption balancing, price setting, and wage setting) (Reis, 2009). To the best of our knowledge, no one has compared DSGE models under different inattentive conditions (i.e., sticky information assumption versus imperfect information data revision assumption).[15] So here one of our main emphasizes is to use the model with sticky information to compare with the alternative inattentive expectation model (i.e., the model with imperfect information data revision) to examine which inattentive expectation model can give the better explanation for US economy in around recent five decades (sample period US quarterly data from 1969 to 2015).

Comparing with the baseline model, it is more challengeable to solve the model with sticky information. Since it involves infinity lagged expectation what leads to the question of how we can approximate the model with sticky information in the DSGE equilibrium configuration. Firstly, from the angle of sticky-information model setting, we can see that the proportion of lagged expectations diminish geometrically meaning that the impact on economic agents' expectation derived from the current state is far greater than that of previous periods. Consequently, the expectations that are formed very far from the present situation might not influence current inflation or output gap

[15] From an empirical point of view, for instance, Smets and Wouters (2007) may consider that a more satisfying specification may take into account some frictions. However, in this thesis, we would like to keep it simple, since one of the main questions we would like to focus is to differentiate different inattentive feature and to see whether different inattentive feature matters for dynamics of the economy.

due to the minimal weight (i.e., may approximate to zero) attached to them. Thus, we set j=4 (which meaning the incorporation of lag information up to 4 periods) as the benchmark, the longer period such as j=6 and 8 have been taken in robust check section.[16]

2.2.2.2 The Model with: Imperfect Information Data Revision

For the extend model with imperfect information data revision process both real-time data and revised data has been used, the suggestion comes from the previous studies (Casares and Vázquez,2016; Vázquez et al., 2010). Before introducing imperfect information data revision model, firstly we need to know what is real-time data, for example, if we analyze the economic agent's decision using the data available to us today, we will make an incorrect inference about their economic decision-making. If we look at the time that economic agents made their economic decisions, we are engaging in real-time analysis or taking the data revision seriously into consideration. The model with imperfect information data revision is presented as follows:

IS equation: (2.7)

PC: (2.8)

Interest rate smoothed Taylor Rule: (2.9)

Where and . Data revision is potentially critical in both theoretically and empirically way, although many economic researchers have made an inappropriate assumption about the data available to economic agents at that point. The applied assumption of data is that they are available immediately, yet the reality those data are announced with a few lags. Furthermore, the data

[16] The result in Travandt (2007), by setting maximum j=19, the convergence of the recursive equilibrium law of motion can be achieved for sticky information Phillips Curve model. However, in our selecting sticky information model enter competition is using fewer periods j and which is sufficient to reach convergence.

revision, in general, has been thought either not exist or small, but in real situation data revision may have a significant and big influence on empirical results and which is particularly the case of some variables that are defined conceptually. For instance, such as output gap, where the economic agents when they are making decisions, take this kind of variables know without any doubt. In a real case, such variable as output gap often fluctuates over time. Thus, in this imperfect information model, the data revision has been taken into consideration to see how it has affected New-Keynesian type macroeconomic model as well as empirical results

Moreover, what does data revision look like is followed by the suggestion from Casares and Vazquez (2016) and has been well specified in Appendix B to Chapter 2. Apart from the point as mentioned earlier, another two points should be clarified: 1) under imperfect information data revision hypothesis, the information of the economy its real state matters, for instance, firms' price-setting decision depends on the expectation of marginal revenue and the future nominal marginal costs. Thus, depends on the future aggregate price level. 2) information friction or inattentive feature underlined across this thesis to be taken seriously, such inattentive assumption needs to be reasonable. Where the nominal interest rates made through professional monetary authority are fully observable without noise disturb, and the observation of output gap and inflation are influenced by noises, in other words, both variables involving data revision processes. Collard and Dellas (2010), they argued that, as the data revision process reveals, very few aggregate variables can be observed accurate and correct. Such that, under the assumption of imperfect information, when firms make the price-setting decision cannot fully observe its information, on the other hand, households when to make consumption decision cannot fully observe the state to support them to make consumption plan. Such that, both price (inflation) and consumption (output) can only observe with some random noises. From the above three-equation model where and have been taken as the observed variable realized at time t they are the real-time data. And and are the final revised variables and which are stated as followings.

(2.10)

(2.11)

And we follow by the argument of Aruoba (2008) that many US aggregate time-series (e.g., inflation and output) their revisions are not rational forecast errors and supposed to be connected to their initial realised variables and . Thus, following his argument, we presume that final revision process of US output gap and inflation are defined as follows,

(2.12)

(2.13)

These revision processes allow for the existence of non-zero correlation between final true variables (i.e., output gap and inflation) and their initial realised variables. Besides, the existence of persistence revision processes. In particularly, the shocks of revision processes, and , both are the AR (1) processes. The two data revision processes assumed aim to offer a simple framework to approximate the 'true' revision processes, and to examine whether the deviation of the way we use for assumption to the well-behaved revision processes (i.e., white noise) assumption, influences the estimation of policy and behavioural parameters

.

For simplicity, we assume that revisions process is linear, following Casares and Vazquez (2016), since our estimated model is a linearized-reduced form version of a small-scale closed New Keynesian model. However, noteworthy, Corradi, et al. (2009) finds the evidence which supports that there is a nonlinear relation between data revisions and variables, which can be an interesting further research in the future. In benchmark competing process, we assumed that the final revisions are reached after 3 quarters, namely s=3 when solving the imperfect information data revision models.

Worth noting that, there are existing studies to contradistinguish the DSGE model with full-information rationality with the alternative DSGE model with sticky information. For example, Paustian and Pytlarczyk (2006) evaluates DSGE model for euro area based on Smets and Wouter's (2003) model through Bayesian estimation approach, and their main finding is that, and Calvo full-information rationality type model overwhelmingly dominants the model with sticky information regarding the posterior odds ratio. Trabandt (2007), use the full-specified DSGE model under the sticky-information assumption and compare it to the Calvo full-information rationality type model, and with allowance for the dynamic inflation indexation (e.g., Christiano et al., 2005), and found that both do equally well. Meanwhile, studies aim to compare the full-information rationality type model with the alternative with Imperfect Information Data Revision also existing. (Paloviita, 2007b, 2008[17]; Vazquez et al., 2010; Casare and Vazquez, 2016[18]), and they provide that the employ of real-time-data variables improves the empirical behave of the classical New-Keynesian model, moreover relax the full-information rationality expectation tentative generates a remarkable distinction for the parameter of the New-Keynesian model.

2.3 Conclusion

For each model with and without inattentive feature, first, it has assumed AR (1) process for all disturbances to each structural equation to capture omitted variables.

---

[17] Paloviita (2007b, 2008) uses the European panel data and apply GMM system estimation to investigate the empirical performance of the standard three-equation New-Keynesian reduced-form model under different information assumption, compare the full-information rational expectation with measured expectation through using revised (final) data, but in their used three-equation without no systematic error. Their estimation results provide evidence that incorporate data revision make the significant difference for parameters, particularly for monetary policy.

[18] Vazquez et al. (2010, 2012), based on three-equation framework to incorporate data revision issue into monetary authority, on contrary we assume monetary authority leave without data revision issues, but economic agents (households and firms) through data revision process to reduce noise to in order to figure out the real state of economy; Casare and Vazquez (2016) to incorporate the data revision into Smets and Wouter's medium-scale type model.

Besides, the frequency of each variable is quarterly, and each variable is demeaned variable, detrend data will be applied. Note that these three models have different information friction constraint, therefore having different IS and Phillips Curve (PC), and therefore may influence monetary policy. After then, by comparing their data fit ability (i.e., log marginal likelihood and Bayes' Factor), one should be able to say whether the suggestion of incorporating inattentive feature from previous literature can provide a better explanation for US economy relatively. Moreover, further explore whether different inattentive feature matters to explain economy dynamics.

Various macro-econometric methods are applied to do model estimation and comparison. The first applied analyzing method is Bayesian estimation approach, which is used to evaluate each model's performance through using US quarterly data in Chapter 3. One of the most significant strengths of Bayesian estimation method is that it provides a solution to find the relatively 'best' model, which can be done with the assistance of a model's marginal likelihood which is directly relevant to the model's prediction ability. Thus, the models for forecasting and policy analysis can be verified by the benchmark of the performance of prediction. Meanwhile, another criterion to verify the relatively 'best' model is the Bayes factor. Different prior distributions and different types of observations are used for robustness check.

Appendix A to Chapter 2

Table 2A-1 Reduced form for each economy to be estimated

	+
	( )

Appendix B to Chapter 2

B1. Sticky Information Model Derivation

B1.1 Sticky Information Model: IS Curve

Now we assume economic agents, households under the sticky information economy use the outdated information from all past period up to t to form their forecast, and in aggregate level not all of them use the updated information to form their forecast, then we have the following IS equation. Where denotes the share of updating households.

(B1.1)

B1.2 Sticky Information Model: Phillips Curve (PC)

Similarly, for firms, also subject to sticky information, and in aggregate level they are using not all of them use the update information to form their forecast, firms use the outdated information up to time t to form their forecast then we have the following PC equation, where denotes the share of updating firms.

$$\frac{\quad}{\quad} +$$

$$\frac{\quad}{\quad} + \qquad \text{(B1.2)}$$

From above we can see the current inflation thus depends on the current output gap as well as on current and past expectation of the future inflation rate.

B2. Imperfect Information Data Revision Model Derivation

The derivation of Imperfect Information Data Revision Model is following the deriving procedure and assumption explanation are following by Aruoba (2008), and Vázquez et al. (2010, 2012) and Casares & Vázquez (2016). First, let us consider the following identities regarding revised data related to cyclical of output gap and inflation, and which is the combination of the initial announcement and the final revisions. Which can be interpreted in the sense of noise, and have been taken as the observed variable realised at time t they are the real-time data. And and are the final revised variables and which are defined as follows.

(B2.1)

(B2.2)

And we follow by the argument of Aruoba (2008) that many US aggregate time-series (e.g., inflation and output) their revisions are not rational forecast errors and supposed to be connected to their initial realised variables and . Thus, following his argument, we presume that final revision process of US output gap and inflation are defined as follows,

(B2.3)

(B2.4)

These revision processes allow for the existence of non-zero correlation between final true variables (i.e., output gap and inflation) and their initial realised variables. Besides, the existence of persistence revision processes. In particularly, the shocks of revision processes, and , both are the AR (1) processes. The two data revision processes assumed aim to offer a simple framework to approximate the 'true' revision processes, and to examine whether the deviation of the way we use for assumption to the well-behaved revision processes (i.e., white noise) assumption, influences the estimation of policy and behavioural parameters. Therefore, from above defined equation we can get,

(B2.5)

(B2.6)

Furthermore, notice that final revision process of output gap and inflation also imply the identities' equations that,

(B2.7)

(B2.8)

(B2.9)

(B2.10)

B2.1 Imperfect Information Model: IS Curve

Use the imperfect information data revision assumption, to distinguish from the baseline Full-Information Rational Expectation model, here we can get the IS equation where households involve data revision issues, these imperfect-information type people react to expected revised values of inflation and output gap,

(B2.11)

And then we use the identity equation and to substitute out and from above to get the imperfect information IS equation as follows,

(B2.12)

B2.2 Imperfect Information Model: Phillips Curve

Firms involve data revision issue (noise disturbance) we can get the imperfect information Phillips curve,

$$+$$

(B2.13)

Similarly, we use the identity equation to substitute out from above to get,

(B2.14)

Meanwhile, the monetary policy assumed perfect observed and live without data revision issue,

(B2.15)

Where the final revision and their data be constructed as demeaned observables between the first released and the more latest released as follows,

(B2.16)

(B2.17)

So, here for analysis we choose for s=3 to construct the observations of final revision and ,

(B2.18)

(B2.19)

Therefore, we can also construct the observation of revised data and .

Note that as argued by Croushore (2011), if we look at the US data, which will show us that s is neither constant with the passage of time nor across variables. One may need to check whether the alternative of s will significant influence Imperfect information data revision its model performance. Here we choose s=3, as the data released in 2016Q1 and as the data released in 2016Q3 to construct the revision process corresponding to sample period from 1969Q1 up 2015Q4. For the simplicity of the analyzing procedure, the number of periods after which without more revisions, except benchmark revisions, and which is represented by s and to be constant.

Chapter 3
Estimate New-Keynesian Type Models with Inattentive Feature through Bayesian Approach

3.1 Introduction

In chapter 3, we focus on estimation and comparison basing on the reduced-form New-Keynesian DSGE model which was restricted by different inattentive expectation assumptions, such as sticky information expectation and imperfect information data revision expectation for the US economy over period 1969-2015 using a Bayesian approach. The three aims of this chapter are to explore which expectation model can reproduce the dynamics behavior of the US real-time data best through Bayesian estimation approach (survey data also used as the alternative type of observations in robustness check), to verify whether incorporating inattentive features can improve the model's performance, and to discuss how different inattentive ingredients influence the dynamics of the economy which can be checked from estimated Impulse Response Functions.

Bayesian estimation approach evaluates different kinds of model by comparing the marginal likelihood of them in a reasonable way. The natural parameters with respect to chosen applied models and the stochastic processes, manage the structural shocks derived by three key quarterly macro data in US economy: output gap (use the real GDP, and output gap is the difference of log of real GDP and log of potential GDP), inflation (log of implicit price deflator) and nominal interest rate (effective federal funds rate). We follow Bayesian estimation approach to evaluate each competing model through three stages. Firstly, through integrating the prior information of the parameters and the likelihood of the data, we can have the log of posterior function, by computing the maximum of which the mode of the posterior distribution can be reached. Secondly, to implement MH algorithm which enables us to obtain a full picture of the posterior distribution and the evaluation of the model's marginal likelihood. Finally, the comparison of three various models in terms of models' performances: full-information rationality expectation model, sticky information model and imperfect information data revision model are analyzed in the result.

The findings presented in this chapter indicate that the US three main economic quarterly real-time data forcefully prefers the model under the assumption of sticky information. Moreover, then through Bayesian estimation approach, we find that the specification with the sticky information outperforms other versions according to marginal likelihood and formal criterion Bayes Factor. Furthermore, the estimated parameters have reasonable values that agree with those typically analyzed in the literature. The model with imperfect information data revision ranks as second outperform model. The baseline under full-information rationality hypothesis type model performs worse than either of inattentive assumption models. We interpret these findings through Bayesian estimating approach as suggesting that incorporating inattentive feature is needed for the New-Keynesian rational expectation model to be a better monetary business cycle model. Besides, different inattentiveness does have impact on the three aspects that used to explain economic dynamics. The three aspects are estimated posterior distribution, estimated impulse response function, and significant different values of log marginal likelihood.

The rest sections of Chapter 3 are structured as follows. Section 3.2 contains the involved literatures about DSGE model estimated through the Bayesian method. Section 3.3 contains the description of the Bayesian estimation approach applied in this chapter. Following that description, Section 3.4 includes the explanation of the data and priors' estimation. Section 3.5 analyzes the assessments of estimation results and comparison results will be showed in section 3.6 and 3.7. Finally, Section 3.8 summarizes this chapter.

3.2 Related Literature of Estimating DSGE Model through Bayesian Approach

Bayesian estimation approach has often been applied to estimate DSGE model in recent years. Within large-scale of recent literature, some of them have been paid significant attention, for instance, Schorfheide (2000) uses Bayesian approach to contradistinguish the model fit of two rival DSGE models of consumption. Lubik and Schorfheide (2005) studies whether small-open economies' central banks are in response to exchange rate volatility. Smets and Wouters (2003) evaluate European countries through using Bayesian estimation method. Rabanal and Rubio-Ramirez (2005) evaluate four various competing New-Keynesian type models with nominal rigidities though Bayesian estimation approach by comparing the model fit.

There are some papers studying the similar topic as ours by applying Bayesian estimation approach as well. For instance, Mankiw and Reis (2007) evaluates model through Bayesian estimation approach to check the influence of sticky information on macroeconomic dynamics and policy base on the general-equilibrium framework. In Collard et al.'s (2009) paper, the possibility that through introduction imperfect information in New-Keynesian type models improves the model fit has been evaluated through Bayesian estimation method. Milani and Rajbhandari (2012) evaluate a variety of expectations formation models through Bayesian estimation approach, and their study shows that when the assumption of full-information rationality has been relaxed and adjusted, the significant shift of the posterior distributions of the structural parameters exist. Levine et al. (2012) compare perfect information with the alternative imperfect information by applying Bayesian estimation approach and finds that information is an essential factor for estimation. Besides, in Levine et al.'s (2012) paper, the results show that the New-Keynesian type model under the imperfect information assumption fits the observed autocorrelation of the data. Whereas, the model under the assumption of perfect information results in a poor model fit. Thus, the analysis in this thesis can be thought as the empirical model competing and selecting exercise by comparing the New-Keynesian DSGE type model under full-information rationality assumption with the alternatives under inattentive expectation assumption (i.e., sticky information and imperfect information data revision).

3.3 Bayesian Estimation Methodology

The reason why Bayesian technology has become increasingly popular in recent studies

For the question of why Bayesian technology has become increasingly popular in recent studies, there are three reasons given and repeatedly been stated by researchers (Geweke 1999; Fernandez-Villaverde and Rubio-Ramirez, 2004; An and Schorfheide, 2007). Firstly, the use of priors allows the previous both macroeconomic and microeconomic researches to be taken into consideration, which offers a way to connect to the previous useful literature. Secondly, the Bayesian estimation approach can give us valuable and stable results under the circumstance of the sample data is comparably small. The Bayesian estimation approach offering a way to assess a model with fundamental misspecification is the third reason. Because what can be accomplished by using the models' marginal likelihood or by using formal criterion Bayes factor. Bayesian Economists may argue that the DSGE model is an approximately/comparative specific version of modelling reality for there is no one hundred percent true model. Therefore, the Bayesian estimation approach is consistent with the argument that no model can be used to describe the real world correctly. Hence, the Bayesian method is used to study the DSGE models which agree with the beliefs held by many macroeconomic researchers.

The classical Maximum Likelihood Estimation approach was argued as a relatively weak estimation method since this approach has been proved to be feasible only for relatively small size systems but not appropriate to be employed to estimate large-scale type models. For instance, Canova (2009, pp.432) says that: ‘One crucial but often neglected condition needed for a methodology to deliver sensible estimates and meaningful inference is the one of identifiability: the objective function must have a unique minimum and should display ‘enough' curvature in all relevant dimensions.' If a model has a lot of parameters required to be estimated, we will meet trouble to achieve the correct information about the estimated parameters from the data. Thus, two problems are exposed 1) likelihood may produce estimation results that are not reflecting the information which is held by researchers, namely, the likelihood peaks in odds area; and 2) the parameters with being given various values result in same joint distribution for the data observations, namely, likelihood without enough curvature (i.e., within a large subset of parameters its likelihood is flat). However, the two problems mentioned above can be avoided or at least reduced by ‘reweighting’ likelihood function through using Bayesian approach after introducing the prior distributions to yield a function with sufficient curvature, therefore, can yield a function with sufficient curvature. From this aspect, the Bayesian estimation method is more capable of dealing with identifying problems.[19] Additionally, Bayesian estimation approach enables one to take advantage of the prior information from the fore literatures, either the reflection of the subjective view of the investigators by a particular prior probability density function of the parameter. However, the classical maximum likelihood cannot take even the prior information with the most non-controversy.

.

Moreover, Bayesian estimation approach can minimize the problem which usually caused by using classical maximum likelihood estimate. By using classical maximum likelihood, the overall estimation process is not very insensitive to each parameter its

[19] Detailed discussion of identification problems in DSGE models see from Canova (2009).

estimated value, which means that if the observed data give poor supports to one or more parameters of the estimated model, will result in a bad estimating result. Instead of embodying a specific value of each estimated parameters, Bayesian estimation approach allows the estimated parameters to follow a distribution that encompasses possible estimates.

3.3.1 The Application of Basic Rules of Bayesian Econometrics

3.3.1.1 Bayes' rule

The basic rule of estimating model through Bayesian approach is the Bayes' rule. For example, suppose there are two events , and , but the probability of event given by event depends on both the relationship between event and event and the prior probability of each event occurrence, which case can be incorporated in the Bayesian estimation as follows:

The formulation above is the prior probability of event , which contains no information about event . is the conditional probability of event conditional on event , which is the posterior probability and which is derived from the specified value of . is the conditional probability of event conditional on event , which is also called the likelihood. is the prior probability of event .

3.3.1.2 Application of Bayes' rule

In the application of Bayes' rule, event is the counterpart of the model's parameter needed to estimate and event is the counterpart of the observable actual data . The posterior density of the parameters can be obtained through combining the likelihood function, prior, and the marginal probability. The parameter is a random variable under the premise of Bayesian econometrics. Bayes' rule is implemented through substituting the corresponding factors:

Where is the prior containing no actual-data information available about the parameters of the model . is the likelihood function which is the density of the observed actual data being conditional on the parameter of the model . is the posterior whose function is summarizing all the information about the parameter of the model after observing the data. The expectation of the posterior can give a point estimate after being calculated.[20] Besides, is the probability of new evidence under all cases usually normalized and constant.

3.3.2 The General Working Steps of Bayesian Estimation

Basically, Bayesian estimation can be taken as a link between priors and maximum likelihood function. The maximum likelihood approach enters through the estimation processes based on confronting the model with the data. The likelihood function implies the probability of observing a given set of data. The priors can be taken as a

[20] The highest posterior density interval (HPDI) gives the smallest credible interval : (i.e., interval estimates).

tool to re-weight the likelihood function to insure more significance to specific regions of the subspace of parameter., Likelihood functions and priors as two components are combined through Bayes' rule to construct posterior distribution. The general working steps of Bayesian estimation are summarizing as follows.

1) To formulate our understanding about the situation: Firstly, to define a distribution model which expresses the qualitative aspects of our understanding of the situation. This model will contain some unknown parameters, which can be treated as random variables. Secondly, to determine a prior probability distribution, which represents our subjective beliefs and uncertainties about the unknown parameters before observing the data.

2) To collect observed data.

3) To get posterior knowledge about our updated beliefs by the means of calculating the posterior probability distribution

Finding the posterior distribution of the models' parameters, if the posterior distribution is ordinary, such as Chi Squared distribution or Normal distribution, it indicates available for us to find the analytical solution. Accordingly, Monte Carlo integration can be carried out to get the estimation of posterior straightly. However, if , namely the posterior distribution is not ordinary, then we are not able to achieve the random draws straightly from the posterior. As solutions to this issue, two ways have been provided to get well approximated posterior density: 1) To implement Independent Draw Approach: draws from the posterior distribution are independent of each other. e.g., importance sampling and acceptance sampling[21]; and 2) To implement Dependent Draw

[21] Importance sampling is a technique for estimating properties can be used under particular situation, e.g the distribution possessing samples generated from another distribution; Acceptance sampling uses statistical sampling to determine whether to accept or reject candidates. It has been a common quality control technique.

Approach: Draws from the posterior distribution are dependent on the previous draws (Marko chain Monte Carlo (MCMC) algorithm) e.g., Metropolis-Hasting Algorithm

In this chapter, we use the Metropolis-Hasting algorithm as a sub-method of MCMC sampling method to get the model parameters' posterior distributions. To be specific, the concept of the Metropolis-Hasting algorithm refers to produce a Markov-Chain that exhibits a sequence of feasible estimates of parameter by exploring the entire domain of the parameter space. Hereafter, the simulated posterior distribution is constructed through using the frequencies related to each value of estimated parameters to construct histogram accordingly mimic posterior distribution. To take the process above, the Metropolis-Hasting algorithm first specifies a 'candidate' distribution, from which some parameter estimates can be drawn. An 'acceptance-rejection' rule is given to determine which several generated estimates are retained. The Metropolis-Hasting algorithm is used to pick qualified 'candidates' distribution under the general regularity conditions. Thus, asymptotically normal posterior distribution can be obtained. Then, by employing the Hessian and mode which achieved from the maximization of the posterior kernel to determine the mean and variance.[22]Summarized Metropolis-Hasting algorithm working steps

Metropolis-Hasting algorithm takes draws from convenient candidate generating density. Let indicate a draw taken from this density which we denote as . The Metropolis-Hasting algorithm always takes as follows:

1) Choose the starting value,

2) Take a candidate draw, from the candidate generating density, which denotes the density of depends on .

[22] The variance is constructed as an inverse of the Hessian multiplied by a scale factor which determines the acceptance ratio.

3) Calculate the ratio of acceptance , which provides the ratio of acceptance of as a draw from the posterior. The ratio of adoption tends to shift the chain from the low-posterior-probability region to the higher-posterior-probability regions.

Set with ratio of acceptance , and set with the ratio of non-acceptance . These new estimates will be reserved with probability and rejected with probability , which facilitates us to search the whole area of the posteriors' distribution. We are not supposed to simply remove the candidates immediately for the reason that a smaller value of the posterior kernel might give us a chance not to be trapped around a local maximum, enabling us to reach the global maximum. What forms the center of the distribution is the highly favorable values which are represented by various appearances, while the tails of the distribution are constituted by the less favorable values.

4) Repeat steps 2, 3 and 4 S times

5) To take the average of the S draws ,……

To assure that the influence of initial values has disappeared and draws converged to posterior, the initial draws are abandoned, and the rest draws are kept as those from the posterior for the estimate of .

Overall, the Metropolis-Hastings algorithm, as an approach classified from MCMC algorithm, is applied to simulate posterior distribution in the cases when we meet trouble to carry out directly sampling. Therefore, Metropolis-Hastings algorithm is chosen to get sequences of random samples from the probability distribution. The current draw of Metropolis-Hastings algorithm always depends on the previous draw of that to get a candidate's probability density function which is represented as a chi-square distribution or a normal distribution. The variance of the 'candidate' distribution, especially the scale factor of it, plays a key role of the procedure. Since the Markov chain relies on the rate of acceptance in visiting the entire distribution to get the global maximum. In practice, a too small variance leads to an obstacle of scanning all the maximum to get the global one, while a too high variance result in a difficulty in finding the global maximum due to taking too much time for visiting the tails of distribution. As a result, a proper j-scale is important to determine the variance of 'candidate' distribution. It is suggested to set j-scale to make the acceptance probability within 20% to 40%. Besides, basing on the principle that the more the 'buckets' iterated in distribution, the less each of them takes the weight, which leads the histogram more complying with the desired theoretical distribution. Thus, the level of smoothness of histogram is increased correspondently with the amount of iterations. In this way, we can get approximated posterior distributions[23].

3.3.3 Model Competing and Selecting through Bayesian Approach: Marginal likelihood, Bayes' Factor, and Posterior Odds Ratio

[23] Approximated posterior distributions featured by the location (i.e., mode and mean) and dispersion (i.e., probability intervals and standard deviation). This methodology offers not only point estimate of the structural parameters but also a measure of the uncertainty around these estimates.

3.3.3.1 Marginal Likelihood of the Model

The key for modeling comparisons through Bayesian approach is the function called marginal likelihood .In detailed, donates the observables, and denotes objective estimated model. Hence, the marginal likelihood function is the counterpart of the density of the observed data while being given objective estimated model yet unconditional on the models' parameters. Through integrating out the parameters of the model we can work out the marginal likelihood function which are presented as follows:

where and are the prior probability and the likelihood function respectively. To solve the function involving multidimensional integration which is not tractable analytically, it requires to repeat to gain numerical approximations. The two most commonly applied methods are the Harmonic mean estimates and Laplace approximation estimates. The former solution utilizes the Metropolis-Hastings runs to simulate the marginal likelihood to take the simple average of the values which obtained from the simulation. The latter solution assumes that the Gaussian distribution can be adopted to approximate the posterior kernel.[24] From the two methods mentioned above, the former method is preferable, because it does not assume the posterior kernel as any formation of function (e.g., Gaussian distribution). If the assumption or restriction is incorrect, it may result in inaccurate results. Although the latter method consumes relatively less time on computation since it only requires some numerical calculations of posterior model and requirement of Hessian matrix, it has a restriction on functional formation of the posterior kernel, which may issue in inaccurate results. In general, the marginal likelihood is a natural way to measure

[24] Gaussian distribution can be used to approximate the posterior kernel to assesses its integral at the mode and variance.

model's unconditional overall performance through representing the overall likelihood.

3.3.3.2 Bayes' Factor

One criterion used to measure model's relatively performance is the Bayes factor which takes the responsibility of verifying ability of models in empirical uses. This Bayes factor is the simple ratio of marginal likelihoods between any two models, such as, and . The Bayes Factor is given as,

In this way, the Bayesian factor is a rule that based on models' fit to sample data ( ) to compare two relative models.

3.3.3.3 Posterior Odds Ratio

Another criterion is the Posterior Odds Ratio which is a more completed tool can be applied to measure model's relatively performance. The construction of Posterior Odds Ratio considers the case of two rival models and .If one assigns prior probabilities to each model and after data being observed to name them and respectively. After then, applying Bayes theorem with them, we can work out and in the same way as we apply to obtain the posterior distribution of parameters.

is the unconditional density of the sample data. Since, does not rely on either parameter of the model or specification of the model, we can substitute into to get,

Similarly, for model ,

Then the Posterior Odds Ratio of model versus model can be represented by the function as follows:

The posterior odds ratio can be regarded as a measurement of the relative performance of two competing models not only based on their model fit to the same sample data (i.e.,), but also on their beliefs concerning the probability of belief of each model (i.e., prior ratio,) . If one knows nothing about which one is more aggregable, the equal weight has been assigned to each model (i.e.,). Herein, there is no different between the Posterior Odds Ratio and the Bayes' Factor. The optimal determination through Posterior Odds Ratio criterion is to choose the one gains the highest posterior support, for instance, if we will select model , yet if we will select model .

3.4 Sample Data and Priors

3.4.1 Whole Sample Data

To estimate the parameters of the three competing models through Bayesian approach, three main macroeconomic variables from year of 1969 to 2015 of US economy are used (i.e., real GDP, GDP deflator, and the nominal interest rate). Sample data and their descriptions are presented and defined in Appendix C of Chapter 3.

3.4.2 Priors

In applying Bayesian estimation approach, the incorporation of prior distribution plays an essential role in estimating DSGE models. The specification of prior distributions (i.e., probability density function of the parameter) is where the Bayesian estimation process begins. The selection of prior's distributions can be made basing on several norms. For example, some most common applied distributions are as follows, restricting the parameters to be positive through inverse gamma distribution, restricting the parameters between 0 and 1 through the beta distribution, and restricting the parameters without any bound through the normal distribution respectively. Besides, the values of priors can be presumed from either previous studies or the investigators' subjective views. The ways of selecting the value of prior have to be in line with the analyses of the context of the model, which means the construction of each priors includes non-sample-data information in the estimation. In other words, priors constitute extra independent information on the model's parameters.

Moreover, selecting values of parameter facilitate to define its distribution which contains the measure of location (e.g., mean and mode) and dispersion (e.g., probability intervals and standard deviation). To this end, the parameters are usually divided into two groups. To be specific, the first group refers to the parameters with relatively strong prior beliefs about (e.g., involve the core structural parameters of the model); the second group refers to parameters with relatively weak prior beliefs about, (e.g., involves the parameters that used to characterize the structural shocks). In the former group, the priors of the parameters are based on the survey of existence of the empirical evidence as well as their implications for macroeconomic dynamics. Although we can adopt the parameters in the latter group based on surveying the previous literatures, in order to constrain the prior distribution within a considerable scope of parameter values the strategy of setting priors needs to be reasonable with proper density that derives from sufficient supports. Since the priors are created from normal standard densities, its computation is quite straightforward.

Most of parameters' prior distribution are chosen from previous literature within a reasonable range for explaining US economy. For instance, the price stickiness which is represented as whose value is 0.6 has been used in many empirical studies (Blinder et al, 1998; Nakamura and Steinsson, 2008; Milani and Rajbhandari, 2012). Additionally, the values of sticky-information parameters and what are 0.5 both are borrowed from Mankiw and Reis (2007)[25]. Moreover, the values of the parameters, regarding imperfect-information data-revision, and are set with mean 0 under the circumstance of allowing large standard deviation from the reference of Casares and Vazquez (2016). Meanwhile, some of the parameters' priors are very strict, and are set fixed before the exercise. Taking the time discount factor and the strategic complementary parameter as examples, they are fixed as 0.99 and 0.15 respectively.[26] We have little knowledge regarding the process that describes the

[25] The value and , both centered at 0.5, implying average information update every two quarters.

[26] As noted by Keen (2007), this is not a completely innocuous assumption, since the hump-shaped behavior of inflation in Mankiw and Reis (2002) disappears if price-setting decisions

forcing variables, so we impose a beta distribution which is centered at 0.5 for the AR coefficients to guarantee the stationary shock process. An inverse gamma distribution is used to restrict the volatility of shock to guarantee its positive value with the mean of 0.33 for the demand shock, 0.33 is assigned cost-push shock, and assign 0.25 to policy shock respectively (Milani and Rajbhandari, 2012). The same strategy is applied for the standard deviation of the revision shocks in imperfect information data revision model with the mean value 0.25 and relative higher volatility 4 to capture uncertainties. We assign 1 to the mean value of the intertemporal elasticity of substitution as the implication of log utility in consumption (Gali and Gertler, 2002; Gali et al., 2003; Meyer-Gohde, 2010), while we set wide standard deviation of as 0.5 in order to restrict the fluctuation in a reasonable range based on previous studies.

Concerning the priors of Taylor rule being borrowed from the previous common selection (Smets and Wouters 2003, 2007; Meyer-Gohde, 2010), assigning 1.5 as mean value to the reaction to the inflation, the 0.25 as its standard deviation, and follows normal distribution. At the same time, the same distribution is applied to restrict the reaction to output gap, yet with the different mean value 0.12.and different standard deviation 0.05 respectively. The lagged interest rate its coefficient, is also restricted by the same distribution, but assign 0.75 to its mean value and 0.1 to its standard error respectively to describe the persistent property of the policy rule. The specifications of priors[27] and the estimated mean values of posterior of the rival models' parameters as well as shock processes are well presented in Appendix A of Chapter 3.

---

are strategic substitutes. The defense of the assumption of strategic complementary in price-setting decision can also see Woodford (2003, chapter 3)

[27] Those specification of priors include distribution types, mean and the standard deviation.

3.5 Way to assess Bayesian Estimation Results

The crucial point of using Bayesian approach to estimate DSGE type model is how to assess the result of Bayesian estimation results. We assess the Bayesian estimation results from checking the estimation diagnosis and results as follows.

Firstly, if the MCMC numerical procedures performs well, the inspection of the estimated parameters' mode as well as standard deviation estimates can be convincible correspondently. Namely, the estimate results of parameters should be satisfied on the perspectives of both statistics and economic theory. To check whether the estimated results are plausible, we compare them with previous research works and evidence from micro data.

Secondly, if the estimated results are regarded as sensible ones then those estimates can be taken as favorable starting values for the Metropolis-Hastings algorithm and, thus, its properties of convergence can be examined as the main source of feedback to hold confidence or may indicate problems of estimation results. To reach convergence, we should take many individual runs, each of which performs sufficient number of draws with a different starting value of Metropolis-Hastings simulations. If convergence is obtained, meanwhile the optimizer is not trapped within strange region over the parameters' subspace, we may get following scenarios: results within each iteration of different runs being similar, or results between different runs being close. If convergence cannot be achieved, the issue can be attributed to insufficient support from priors or a deficiency of Metropolis-Hastings iterations in quantity.

Thirdly, as An and Schorfheide (2007, pp.127) said: 'A direct comparisons of priors and posteriors can often provide valuable insight about the extent to which data provide information about parameters of interest.' To check the simulated posterior distribution is essential and can be taken from following aspects: 1) posterior distributions should be an approximated normal distribution; and 2） the prior and posterior should be neither extremely similar nor extremely dissimilar. To be specific, if they are too different from each other which may indicate that prior gives a poor restriction on the sample data. However, if they are too close, the estimated results may largely be guided by the priors and rarely rely on the selected sample data. If a sufficient tight prior distribution is appointed, the informative posterior distribution can still be achieved even that we cannot identify the estimated parameters by the selected sample data. That is the case when the sufficient tight prior distribution has been set, and the posteriors will show well-behaved due to the fact that prior has been chosen within a specific region of the parameter space. It is definitely that the prior would have been selected to preclude the illogical areas of the parameter space on perspectives of statistics and theory. At the same time, prior should also be chosen wisely and uninformatively within a reasonable range to prevent selected sample data from being silent and drawing deceived conclusion. In another word, the movement from the prior to the posterior can be considered as a sign that there is a tension between priors and selected sample data. If prior distribution and posterior are no different with a given parameter, we can draw a conclusion that the estimated results about estimate parameters largely depends on the prior, while the selected sample data is silent on that estimated parameter. In such case, adjusting both the distribution and the dispersion of the prior may be a useful step to check identification problems and offer a clear answer that whether the selected data strongly supports estimated parameters. Additionally, the estimation of the structural shocks need to be checked concerning its reasonable magnitudes and frequencies of innovations. Finally, the sensitivity check can be made regarding apply different reasonable priors or apply different sample data.

3.6 Estimation Results through Bayesian Method

The three competing models are solved and estimated with the Dynare 4.4.3. The methodology we implied to get posterior distribution is a Metropolis-Hastings algorithm which generates 20,000 draws with the acceptance rate within 20% and 40%. The estimated sample data are selected quarterly from the US starting from 1969Q1 to 2015Q4 (survey of professional forecaster data has been used in robust checking section over the same periods).

3.6.1 Assessment results of Bayesian Estimation

The diagnosis for the sampling algorithm, the Metropolis-Hastings is shown in Appendix A of Chapter 3. The information of three aspects, namely the analyzed mean of parameters (interval), the variance of parameters (m2) and third moment of parameters (m3), are concluded in three graphs, each of which represents convergence measures in detailed. The two distinct lines in the graphs shows the results within chains and between chains respectively. To reach reasonable estimated results, both lines concerning each of the three measured aspects must be steady and convergent to one another. What can be seen from the graphs that overall convergence is approached. In terms of prior densities, we use gamma (G), inverse gamma (IG), beta (B), and normal (N) distributions. The prior and posterior distributions have been presented in Appendix A of chapter 3.

The significant different between prior distributions and posterior distributions are not exist, which can be checked graphically. Also, on the perspective of most parameters, the prior and posterior distributions are not extremely close which implies that the observables provide extra information for most parameters of the estimated models, which indicates that the presumed priors are not the only factor that influence the estimates. The estimated posterior is an approximated normal distribution whose shape consistent with the Bayesian estimation its asymptotic properties.

3.6.2 Summary of Posterior Estimates

The construction of the posterior distribution under Bayesian econometrics can be achieved by combing the prior distribution together with the likelihood function by using Kalman filter mechanism. After accomplishing Kalman recursion and evaluation and maximization got the log likelihood function and log prior density, Chris Sim's csminwel is applied to approach the estimated posterior.[28] Afterwards, the posterior distribution can be achieved through running 20,000 draws by Metropolis-Hastings algorithm with optimal acceptance rate (i.e., between 20% and 40%). From the 20,000 draws, the initial 20% are discarded and the rest are kept to eliminate any dependence of chain from its steady state.

The Table 3-1 gives the estimated posterior distribution of the parameters for each group concerning reduced-form New-Keynesian DSGE model concerning with and without inattentiveness. Incorporating inattentive feature into modelling expectation seriously affects the estimation results of the parameters. For instance, although the estimated intertemporal elasticity of substitution (i.e.,) is lower than prior's value in all three competing models no matter with or without inattentive feature, it varies significantly. In detail, the estimated of the model without inattentive feature is 0.0225. From another perspective, the values of estimated of the SI model is around five times higher than that of one without an inattentive feature. A relatively higher intertemporal substitution implies that large changes in consumption are not

[28] Chris Sim's csminwel is a minimization routine and carry out to minimize the negative likelihood.

very costly to consumers through Euler equation. On another face, if is low, the motivation of the consumption smoothness will be very strong, which is caused by the fact that the consumers will be reluctant to save but consuming a lot relative to the former case.

Regarding to imperfect information data revision model, the economic agents involve signal extraction (data revision) process to understand the real state of the economy. Thus, the value of σ is estimated to be 0.0899 which is four times larger than the one estimated in the baseline model. Additionally, the estimated AR coefficients of imperfect information data revision model, especially the AR coefficients of demand shock and cost-push shock, shift to relatively lower value comparing with that of baseline model. In terms of the estimated parameters (i.e., the reaction toward inflation and the reaction toward the output gap) in monetary policy function, the values are estimated to be not very different under the three models of the estimating results.

Most of estimation results presented in Table 3-1 are remarkably consistent with the previous studies. We find that the reaction towards the inflation is not far away from the presumed prior 1.5 under the three models. The reaction towards the output gap is also not volatile under different expectation assumptions (i.e., varies between 0.1848 to 0.1974). Moreover, the estimated result of shows reasonably high degree of interest-rate smoothness (i.e., varies between 0.8801 to 0.9002) under different expectation assumptions as well. However, higher policy coefficients overall and some structural parameters shift a lot (i.e., varies between around 0.02 to 0.1). The estimated coefficients of AR processes of shocks which reflect the existence of substantial degree of persistence in the data. The highly persistence performance are captured by the high degree autocorrelation in demand shock which is estimated above 0.6 in all three models. The autocorrelation in cost-push shock is estimated around 0.7 of both baseline model and sticky information model. However, regarding imperfect information data revision model, the estimated is quite low (i.e., is estimated to be 0.3657). Moreover, compared to and , the coefficient of monetary policy shock is estimated relatively small, which is around 0.2 to 0.3 regarding three models.

The estimation results illustrated above concerning the estimated posterior mean are not meant to show that one specified model is superior to the other models. By comparing the variation between estimated posterior results under the two different situations (i.e. with and without inattentive feature), we can check the sensitivity of the results. Furthermore, through evaluating the posterior results under the models with two different inattentive expectation assumptions, it is available to check the sensitivity of them. The necessity of checking sensitivity of variation concerning the models with different inattentive expectation assumptions is derived from the case that it is usually ignored by the previous studies.

TABLE 3-1 SUMMARY ESTIMATION RESULTS OF DIFFERENT EXPECTATION FORMATION

Prior distribution				Posterior distributions		
Params.	Distr.	Mean	S.D .	FIRE	SI (j=4)	IF
	G	1	0.5	0.0225	0.1092	0.0899
	B	0.6	0.05	0.7257	0.6340	0.7389
	B	0.75	0.1	0.8834	0.9002	0.8801
	N	1.5	0.25	1.3891	1.3735	1.0884
	N	0.12	0.05	0.1974	0.1848	0.1962
	B	0.5	0.15	0.7995	0.8139	0.6186
	B	0.5	0.15	0.6948	0.6940	0.3657
	B	0.5	0.15	0.3094	0.2986	0.2235
	IG	0.33	1	0.1564	0.5548	0.2710
	IG	0.33	1	0.0878	0.2446	0.1551
	IG	0.25	1	0.2301	0.2294	0.2181
	N	0	2	-	-	1.8500
	N	0	2	-	-	1.1198
	B	0.5	0.2	-	-	0.7252
	B	0.5	0.2	-	-	0.8535
	IG	0.25	4	-	-	0.3270
	IG	0.25	4	-	-	0.0808
	B	0.5	0.2	-	0.3084	-
	B	0.5	0.2	-	0.2362	-

3.6.3 Models Comparison

3.6.3.1 Model Fit

Table 3-2 shows that the marginal likelihoods of three rivals concerning the different expectation assumptions (i.e., with and without inattentive features), along with the corresponding formal criterion the definition of Bayes factor is the simple ratio of marginal likelihoods between any two models where we take the model with full-information rationality as null hypothesis. Geweke's Harmonic mean is applied to calculate the marginal likelihoods of each case[29]. Comparing the values of marginal likelihood is a standard way of Bayesian approach to know which model fit the data best. The model under the conventional assumption without any inattentive feature produces the lowest value of model fit. Maintaining rationality but extending to include inattentive ingredients, the models' performances are improved. Particularly, the model with sticky information expectation achieves the best model fit among the three competing models.

The implementation of the sticky-information model requests a predicting horizon (i.e., truncation point j), however, there is no clear approach to select the value of truncation point j. If the short forecasting horizon, namely small value of j, is supposed to be two or three quarters which are comparably short periods, it would lead to the misperception of the distribution of agents regarding to updating their information relative to the distribution given by theoretical model. On the other hand, a long forecasting horizon will include too much forecast errors, which tend to form bias to reduce the estimated share of updating agents (i.e., and) (Khan and Zhu, 2002).

[29] There two common methods for computing marginal likelihood concerning Bayesian method, one is so-called Laplace approximation, which assumed that posterior kernel can be approximated by a Gaussian distribution and evaluates its integral at the mode and variance obtained with the numerical maximization of the posterior. The second method is so-called Geweke's Harmonic mean estimator uses MH runs to simulate the marginal likelihood, and then simply use the average of these simulate values.

Balancing the reduction of forecast error and the frequency of updating information theoretically, we set j=4 as our starting point[30], and the alternative j=6, and 8 also have been taken into consideration as the choices of robustness check.[31]

Guided by Jeffreys (1961), we have a way to evaluate the preponderance of the evidence in the light of a selective model concerning the model in the null hypothesis to interpret it into the comparable superiority of model. The detail of guidelines is presented in Table 3-3. Basing on his guidelines, the Bayes factors' values in Table 3-2 show that 'decisive' evidence for both models with inattentive expectation assumptions against the baseline model with full-information rational expectation assumption. Moreover, between two models with respect to different inattentive expectation assumptions, we take the imperfect information data revision model as the null hypothesis. Through Bayes factor, it implies that the model with sticky information shows the 'decisive' evidence as a preferable choice (Bayes factor).

However, one obvious limitation of this comparing approach is that by using this method the conclusion of the evaluation of model fit can only be drawn relatively. Thus, the best estimated model would still be deficient (potentially misspecified) in catching the essential dynamic in our selected sample data. The model's performance is assessed in an absolute way of one model against data, the indirect inference has been chosen as the robust check approach to re-examine model's performance, and this will be conducted in Chapter 4.

[30] Kiley (2007, p112) compares the sticky prices and sticky information empirically and noted that, 'in practice., the longest information lag is truncated as four quarters.'
[31] Paustian & Pytlarczyk (2006), they have examined the sticky-information with different truncation point j=12, 24 respectively, and they find that the sticky information its model fit is not sensitive regarding increasing the maximum lag for outdated information and almost does not change.

TABLE 3-2 MODEL FIT COMPARISON

Model	Log Marginal Likelihood	Bayes Factor relative to the FIRE
FIRE model	-267.05	
SI model (j=4)	-247.36	
IF model	-254.08	
Note: (1) Sample period: 1969Q1-2015Q4 US macro data; (2) FIRE represent Full-Information Rational Expectation Model; SI represent Sticky Information Expectation Model; IF represents Imperfect Information Data Revision Model.		

TABLE 3-3 JEFFREY'S GUIDELINES FOR INTERPRETING BAYES FACTOR[32]

Bayes Factor	Interpretation
1 to 3.2	Not worth more than a bare mention Evidence
3.2 to 10	Substantial Evidence
10 to 100	Strong Evidence
100	Decisive Evidence

3.6.3.2 Estimated Impulse Response Functions (IRFs)

Our selected models are mostly consistent with a large number of literatures with respect to New-Keynesian three-equation model. This section mainly concerns with the appearance of the distinguishing features when we introduce the inattentive features (i.e., sticky information and imperfect information data revision) into the model. Previous study results regarding the introduction the inattentive ingredients into DSGE model (Mankiw and Reis, 2002, 2007; Collard et al., 2009) find that the monetary policy shock has a tendency to produce more delay impact under inattentive expectation economy relative than that under the economy without any inattentive features. The information costs are one of the interpretations behind this (Mankiw and Reis, 2002, 2007), The information costs are consisted of two aspects, one of which is the monetary costs (e.g., payment need to be made to acquire updated information and receive the professional interpretation from a financial advisor). The timing cost is the other aspect (e.g., time of obtaining, processing, and interpreting updated

[32] The use of Bayes Factor to compare models was first suggested by Jeffrey's (1961), who suggest that the following rule of thumb for interpreting Bayes factor.

information) (Begg and Imperato, 2001; Reis, 2006a, 2006b). Thus, due to the information costs, some of the economic agents will chose to use the already-paid old information, which generates the delay response. The other interpretation is that people sustain noisy disturbances so that they need time to filter useful information through data revision process (Casares and Vázquez, 2016).

Accordingly, our main focus in this section is to check how the embrace of an inattentive feature in the model affects the macroeconomic model. Particularly, the delay impact of a monetary policy shock upon the main macro variables (i.e., the delay effects of inflation and output gap) will be verified. Afterwards, the estimated impulse response function results will be shown graphically to illustrate major macro variables of the positive monetary policy impact under two different inattentive hypotheses as well as the baseline without inattentiveness respectively.

As Figures 3-1 shows, the models with sticky information can produce a persistence and a delay reaction of inflation, which is mostly in line with the suggestion from the previous studies (Mankiw and Reis, 2002). However, neither the model without any inattentive features nor the model with imperfect information can accomplish the goal. The results regarding the model with imperfect information data revision, unexpectedly, are different from the suggestions from previous studies (Collard et al., 2009). Additionally, the estimated impulse response functions generated under the model that assumes households and firms involving data revision issues are quite similar with those generated from the baseline model. Overall, the effect of the positive monetary policy shock gives a raise to the nominal interest rate in three competing models.

We turn to examine the IRFs with respect to the model featuring sticky price under full-information rationality assumption. Basing on Euler Equation, there will be a negative power on the demand of households' consumption which leads to holding off consumption, if the nominal interest rate increases along with the raise of real interest rate. The case exactly complies with our estimated results concerning model with full-information rationality assumption. Since the economic activity is directed by demand, the results of decreasing demand lead to a drop of firms' production. At the same time, deflation is generated by a reduction in economic activity demand. As time passes, the economy recovers, in the light of the Taylor rule, since the reductions in both demand and the inflation rates cause a reduction in the nominal interest rate after early period. The two alternative competing models are quite similar to the baseline model in terms of the IRFs of the positive monetary policy shock to main variables quantitatively. Exclusively to the model with sticky information, the positive impact of monetary policy can produce a persistence and gradual response of inflation.

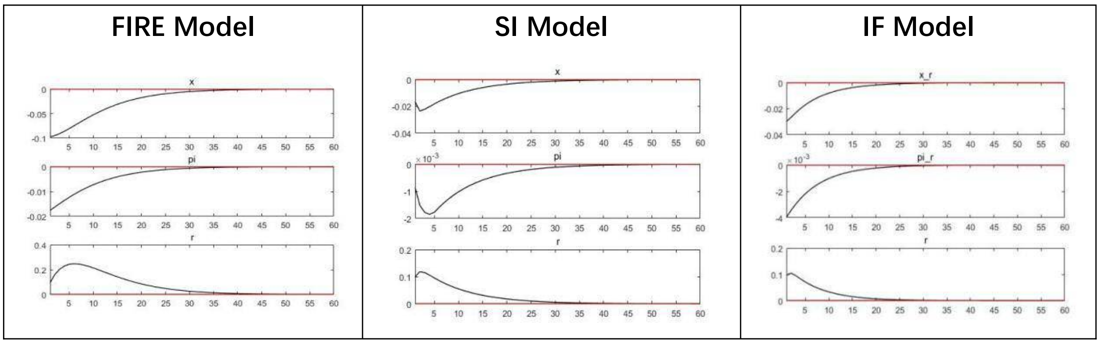


Figure 3-1 Estimated Impulse Response Function of One Unit Positive Policy Shock to Main Variable (x=output gap, pi=inflation, r=interest rate)

After then we turn to examine the effects of the positive demand shock to three main variables under three competing models through estimated impulse response function. The estimated impulse response function has been shown in Figure 3-2. We can see that the positive demand shock, in general, has a relatively long effect on interest rate since this variable converges after around 30 periods. Meanwhile, the demand shock has a relatively significant impact upon the output gap. Two long-run effect converges require 20 periods concerning FIRE model and SI model. However, it only takes 9 periods of convergence under IF model. In general, the demand shock impact inflation positively and converges quickly comparing to the effect on nominal interest rate under the three competing models. Under imperfect information data revision model, people's uncertainty of data revision at initial stage leads to small effects on inflation and output gap. But the turning point appears at the fifth period when people have strong enough confidences on their expectations after reducing the uncertainty. So, inflation and output gap under imperfect information may perform better at bringing about an efficient response and rapid convergence than those under full-information or sticky-information environments.

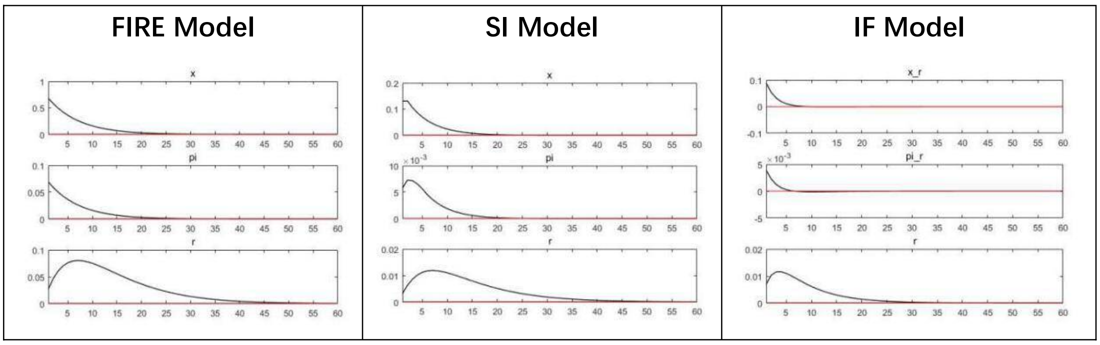


Figure 3-2 Estimated Impulse Response Function of One Unit Positive Demand Shock to Main Variable (x=output gap, pi=inflation, r=interest rate)

The positive cost-push shock impact inflation and interest rate positively regarding three competing models which is presented in Figure 3-3. But the positive cost-push shock leads to different consequences under different models. To be specific, the effect triggered by it under the baseline model is negative, while the those of SI model and IF model are almost null at the initial point on output gap. This distinction may be caused by the fact that people's inattentive behavior to some degree lessen the effect of cost-push shock, which is presented in Figure 3-3. The economic agents under imperfect information assumption environment cannot observe the real state. So, people reduce noise through data revision process and only take actions in reaction to their expected revised data. From the estimated impulse response function, it can be indicated that when people form their expectation through imperfect information data revision, the impact of the supply shock on inflation happens in short term. On the other hand, the models with sticky information will generate more persistence effects on output gap and require relatively longer time for converging. Furthermore, in aggregate level, the variables under the economic agents involving data revision issues converge more quickly than those under the baseline model and the sticky information model.

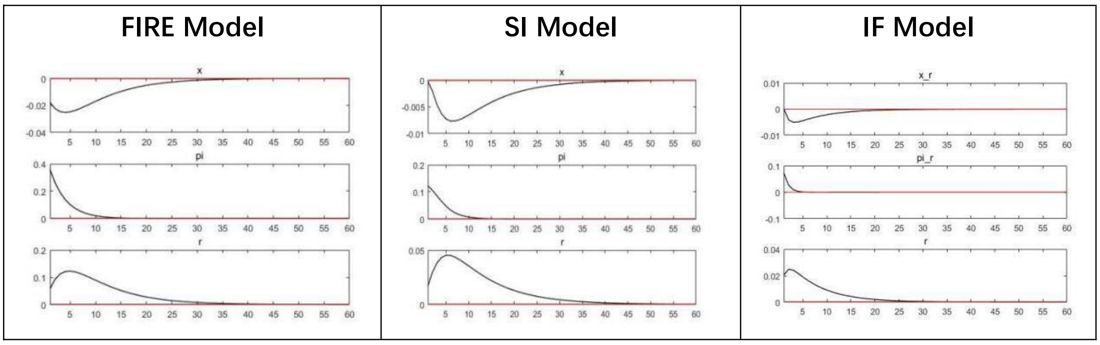


Figure 3-3 Estimated Impulse Response Function of One Unit Positive Cost-Push Shock to Main Variable (x=output gap, pi=inflation, r=interest rate)

To sum up, the model with sticky information expectation has strong abilities of generating more persistence and reproducing delay responses. However, the model with imperfect information data revision expectation cannot attain this goal. However, the fail to reproduce the delay response should not be taken as the reason to judge the model’s invalidity. The result may be caused by two key factors missing in our estimated inattentive expectation models. The origins are wage rigidities and the inclusion of capital variable utilization (Christiano et al., 2005). However, even if the model with sticky information can produce a persistence and delay impulse response, we still cannot confirm whether the model itself can indeed be used to explain the real world. So, it is still necessary to conduct indirect inference evaluation to examine the ability of model in an absolute way.

3.7 Robustness Check

3.7.1 Robustness to the Different Prior Distribution

As presented in Table 3-4, in this section we set to be 0.75 instead of 0.6 after adjusting the mean value of presumed prior of the degree of price stickiness (i.e., α) higher but still being one of the common options applied in many studies (Eichenbaum and Fisher, 2004; Woodford, 2003). It may be worth repeating the analysis with relative flatter prior, namely uninformative prior (i.e., the prior is assumed to follow uniform distribution instead of beta distribution used in starting comparison). The parameter depends on uniform distribution which is assumed within a fixed range of values (i.e., between 0 and 1). The estimated results in Table 3-4 show that the ranking of three competing models is the same as that reported before although different degree of tightness of priors leading different performance in each model. It facilitates us to remove the concerns that our estimation results may seriously be driven by the presumed distribution of the priors and give no chance to let the data speak.

Interestingly, although the models with inattentive feature still are superior to the baseline model in the light of model fit, the distance between sticky information model and imperfect information data revision model is narrowed down, which shows no evidence that the model of sticky information precede model of imperfect information data revision (i.e., when the imperfect information is taken as null hypothesis, the Bayes factor is approximately 1.42). It pushes us to re-examine the model's ability in an absolute way through indirect inference method.

TABLE 3-4 MODEL FIT COMPARISON

Model	Log Marginal Likelihood (Benchmark Priors)	Log Marginal Likelihood (Using Diffuse Prior)
FIRE model (baseline)	-267.05	-261.31
SI model (j=4)	-247.36	-246.75
IF model	-254.08	-247.10
Note: (1) Sample period: 1969Q1-2015Q4 US macro data; (2) FIRE represent Full-Information Rational Expectation Model; SI represent Sticky Information Expectation Model; IF represents Imperfect Information Data Revision Model.		

3.7.2 Robustness to the Different Truncation Point j of sticky information model

The empirical performance concerning sticky information model requires a forecasting horizon (i.e., truncation point j) to be taken into consideration. We set j=4 as our starting point, at the same time, we set relatively longer forecasting horizons j=6 and 8 respectively.[33] The estimation result shows that the truncation point has a so small influence on model fit of SI model that will not disturb the original rank.

TABLE 3-5 SENSITIVITY CHECK OF STICKY INFORMATION MODEL[34]

Model	Log Marginal Likelihood (benchmark priors)
SI model (j=4)	-247.36
SI model (j=6)	-247.27
SI model (j=8)	-247.11

3.7.3 Robustness to the specification of Taylor rule

Concerning that different specifications of the monetary policy rule may influence our estimation results, we re-estimate each model with two other specifications of Taylor rule (Smets and Wouter, 2003, 2007; Woodford, 2003). To be specific, one is the 'more complex Taylor rule' which includes the change of output gap and the change of inflation in monetary authority reaction function whose parameters are represented

[33] Khan & Zhu (2002). Estimates of the sticky-information Philips curve for the united states, Canada, and the United Kingdom. Bank of Canada; Paustian & Pytlarczyk (2006), examine the sticky-information with different truncation point j=12, 24 respectively, and they find that increasing the maximum lag for outdated information sets from j=12 to j=24 the fit of sticky information almost does not change

[34] Since the computation time grows rapidly as j increased, when faced such computational burdens, the attractive choice of truncating may just include a few lagged expectations. Estimating sticky information model with a higher but fixed j might be fairly accurate for some combined parameter, however the only combined is and somehow have been fixed as suggestion from previous studies; and too high j will also unnecessary burden the computations, such that I only consider j up to 8 and take j=4 as starter join model competition.

as and . We set the mean values and standard deviations equal to 0.12 and 0,05 respectively for both parameters and . The settings are in line with the previous studies (Smets and Wouter, 2003,2007) and enable the priors to follow the normal distribution. The other one is the 'less complex Taylor rule' (Woodford, 2003), which has been used in robust check as well and been suggested as a good description without interest rate smooth of the Fed's monetary policy between 1987 to 1992. Moreover, in this case: and have been asserted as good approximations to characterize the US policy (Woodford, 2003). Both alternative specifications have been presented in Table 3-6.

TABLE 3-6 ALTERNATIVE SPECIFICATION OF TAYLOR RULE

More complex Taylor Rule (e.g., Smets and Wouter, 2003, 2007)
Less complex Taylor Rule (e.g., Woodford, 2003)

The estimation results have been checked in Table 3-7, through which we can see that after introducing 'less complex Taylor rule' into the three-equation New-Keynesian framework, each of three competing models gains a worse model performance, which can be checked through the log marginal likelihood. But these results may not be surprising since it is too simple to closely match the optimal policy in the context of an economic model. But, the ranking among three competing models is fixed even though 'less complex Taylor rule' is introduced. But, on the contrary, while we are using the 'more complex Taylor rule' (Smets and Wouter, 2003, 2007), the performances of all the three models are improved.

In general, we can draw two conclusions under the situations regardless which specification of Taylor rule is adopted. The first is that the model with inattentive feature outperforms the baseline model without any inattentive feature. The second is that the ranking among three is identical to the previous results.[35]

TABLE 3-7 MODEL FIT COMPARISON

Taylor Rule / Model	Log Marginal Likelihood (benchmark Taylor Rule)	Log Marginal Likelihood (more complex Taylor rule)	Log Marginal Likelihood (less complex Taylor rule)
FIRE model (baseline)	-267.05	-260.47	-344.33
SI model (j=4)	-247.36	-238.24.	-256.46
IF model	-254.08	-250.80	-310.11

3.7.4 Robustness to alternative data resource: survey of professional forecaster data of output gap and inflation

To make our research more rigorous, the survey of professional forecaster data is chosen by us as a different type of data resource in robust check. Since this kind of data reflects the views of a few of the highly informed economic agents. The data is regarded as a standard so conservative that is available to assesses potential deviation from full-information rational expectations. As Ormeño and Molnár (2015) assert, survey data of inflation contributes to the way of modelling private expectations by providing useful information that macro data do not have. In this section, we extend to examine each model by using a different type of sample data (i.e., survey data). The estimation results obtained by using survey data is summarized in Table 3-8. The estimation results regarding the imperfect information data revision model performs best among three competing models. The gap of log marginal likelihoods of the model with imperfect information data revision model and

35 Of course, there are various monetary policy rule suggested in the previous studies, here we just choose two to do robustness check, the further research may necessary to consider more different monetary policy rules detailed and carefully.

that with full-information rationality is 19.64, which can be interpreted as Bayes factor (when we take the baseline model as null hypothesis). Similarly, the gap of log marginal likelihoods of the model with imperfect information data revision model and that with sticky information (j=4) is 6.68, which can be interpreted as Bayes factor (when we take the model with sticky information as null hypothesis). Regardless of different types of data resource in the estimation process, the gaps of log marginal likelihood of the model with imperfect information data revision and that with sticky information are quite similar (i.e., the gap is around 6.68 when estimated using survey data; the gap is 6.72 when without survey observations).

TABLE 3- 8 MODEL FIT COMPARISON (WITH SURVEY DATA)

Model	(1)	(2)	(1) - (2)
FIRE model (baseline)	-36.08	-267.05	230.97
SI model (j=4)	-23.12	-247.36	224.24
IF model	-16.44	-254.08	237.64
Note: (1) is the marginal likelihood estimated with survey data; (2) is the marginal likelihood estimated with US real-time data.			

Furthermore, when the survey data are introduced as observables, the performance of each model improves a lot. The number of log marginal likelihood increases a lot in three competing models, which demonstrates that there is an additional information in survey data to lift the performance of each model. However, Whatever type of resource we using to peruse the estimation result, the model with inattentive expectation are always superior to the baseline model in terms of model fit. However, under the same premise, the ranking of sticky information model and imperfect information is switched, which may because the extra information contained in survey data is in favor of model with imperfect information data revision. In terms of the three competing models in different types of data, we compare the estimation results with survey data (presented in Table 3-9) to the results with real-time data (presented in Table 3-1). It shows that most estimated values of the common parameters do not have significant difference. However, some differences exist. For example, the AR coefficients of cost-push shock and monetary policy shock are higher than those presented in Table 3-1. Besides, the estimated share of updating consumers is much lower than that estimated by using real-time data. While the estimated share of updating firms is relatively larger than that estimated by using real-time data

TABLE 3-9 SUMMARY ESTIMATION RESULTS OF DIFFERENT EXPECTATION FORMATION (WITH SURVEY DATA)[36]

Prior distribution				Posterior distributions (mean)		
Params.	Distr.	Mean	S.D	FIRE	SI (j=4)	IF
	G	1	0.5	0.0159	0.1344	0.0371
	B	0.6	0.05	0.6519	0.6277	0.6543
	B	0.75	0.1	0.8857	0.9164	0.9219
	N	1.5	0.25	1.4669	1.4146	1.3836
	N	0.12	0.05	0.1236	0.1214	0.1243
	B	0.5	0.15	0.5681	0.5983	0.4922
	B	0.5	0.15	0.6928	0.7033	0.4483
	B	0.5	0.15	0.3473	0.3234	0.3110
	IG	0.33	1	0.1158	0.2487	0.2446
	IG	0.33	1	0.0759	0.2106	0.1552
	IG	0.25	1	0.2384	0.2367	0.2414
	N	0	2	-	-	1.9627
	N	0	2	-	-	1.5134
	B	0.5	0.2	-	-	0.5612
	B	0.5	0.2	-	-	0.7457
	IG	0.25	4	-	-	0.2190
	IG	0.25	4	-	-	0.1132
	B	0.5	0.2	-	0.4474	-
	B	0.5	0.2	-	0.0916	-
Log marginal likelihood				-36.08	-23.11	-16.44
Bayes Factor relative to the FIRE				1		

It is noteworthy that survey data has been used to identify expectation mechanisms in recent studies. For instance, Carroll (2003) finds that the public's prediction is lags behind the prediction of professionals' through adopting survey of inflation

[36] The posterior estimated value of is quite different from prior mean which may due to the selected prior is suitable for final revised data but not suitable for real-time data or SPF data. The results of robustness check with revised data is given in Appendix D to Chapter 3.

expectation data. Being distinguished from the previous literature, Easaw and Golinelli (2010) investigate whether different agents or groups that make up the population have various information absorbing rates. Rather than treating economic agents as homogeneous type agents or groups and through using the UK survey data, and they find that homogeneous agents or group can be distinguished by their information absorbing rate respectively. Easaw and Golinelli (2014) establish a new structure (i.e., people can form their expectation multi-period) but basing on single equation method with the focus of inattentiveness (i.e., sticky information and imperfect information) using survey-based data for the US and UK. A more recent work involves using survey data to examine the model with deviation of full-information rationality, for instance, Del Negro and Eusepi (2011) study whether or not a DSGE model with imperfect information while keep rational expectation assumption can reproduce series of expected inflation that match the survey inflation data. Aruoba and Schorfheide (2011) apply inflation forecasts survey data in their observations as extra information which is able to be employed to indicate the time-varying Fed's Inflation Target. After endogenizing survey expectation in a standard DSGE model, Fuhrer (2017) asserts that most persistent in aggregate data is better due to slow-moving expectations but not habits, indexation or autocorrelated structural shocks.

3.7.5 Robustness to Different Detrend Method

Another problem which has been discussed widely in the DSGE literature is how to detrend real variables, particularly, the methodology to obtain the potential output for constructing the output gap. Most of the studies tend to use the statistical detrending method (e.g., HP filter, band-pass filter etc.). Alternatively, we can use the theory-derived potential output (i.e., the output solved under flexible price assumed economy) to construct output gap.

In this thesis, the HP trend is the approximation of the potential output, which is used to construct the output gap. HP filter is a methodology of statistics that can be used to extract the trend after filtering the actual GDP data as the estimates of potential output. The HP filter is a convenient way to get potential output since it only needs the actual output data. However, HP filter is not impeccable because it does not utilize fully of the information from other economic time series data to direct the estimates of potential output. The absence of economic theory forces it to generate the potential output through a technique instead of a model. As a result, it is not a favored method to model the actual potential output. Another suggested method from the previous literature uses a linear detrending to get the output gap, However, this is not a suitable method, either. Since the potential output has a great chance of being non-linear, which can be proofed by the function derived from the model with flexible price assumption driven by technology shock. Although we have no idea about what a technology shock is, the probability that it is non-linear is very high.[37] However, the aim of this thesis is studying the empirical implications through model comparison. A more detailed study of using different detrending methods to obtain potential output for constructing output gap is surely warranted which can be remained for a future research.

3.8 Conclusion

The previous macroeconomic theory is basing on the assumption which full-information rationality restricts the consumers and households to form their expectations. The conclusions drawn by the empirical studies of macroeconomics also depend on such validity of full-information rationality hypothesis. In this chapter,

[37] In recent years, some contributions have made through using DSGE models to estimate potential and the output gap (Vetlov et al, 2011). These models have more visible micro foundations and are very attractive. Despite that, these are difficult to interpret and still a challenge for policy makers to apply in formulation of policies.

the consequences of the inclusion of inattentive expectations (inattentive features) in a popular small-scale reduced-form New-Keynesian DSGE model have been evaluated. What’s more, the econometric features of the model have been shown that are not insensitive to the introduction of inattentive ingredients through Bayesian estimation approach. The sensitive analysis focusing on comparing different inattentive features largely lack in the previous study.

The empirical evidence shown in this chapter implies that, firstly, the results of Bayesian estimation indicate that the modelling of incorporating inattentive feature has significant influence on the capability of the model in fitting macro-economic time series. Secondly, these are essential to be studied more critically in estimated. The limitation worth to mention here is that the model we have chosen to make comparison may be misspecified.[38] The source of misspecification not only due to the application of linear approximation solution but also the truth that DSGE model is an abstract of the real world.

It is necessary to study estimated DSGE type model in a critical way while different inattentive features incorporate it. However, there may be misspecifications as leaks in the models we have chosen during making comparisons.[39] The misspecification can be originated from that the DSGE model is not perfect to copy the ‘real model’. Thirdly, among three competing models, model with sticky information expectation wins the best model performance through Bayesian estimation using real-time data. In addition, the results show that the model with an inattentive feature improves the model's ability to explain the real world, which is in line with most consequences from the previous studies (Mankiw and Reis, 2002, 2007; Collard et al., 2009). However, between the two inattentive models, only the model with sticky information expectation can generate persistence and delay response, which has been checked

[38] Gourieroux, Monfort and Renault (1993) fully account for the fact that DSGE models are misspecified.

[39] Due to the nature of approximated linear solution methods, the DSGE model may encounter the loss of some components during the process of solving.

through estimated impulse response functions. Finally, the robustness checks with real-time data for our concerns regarding using different prior distribution (diffuse priors), different specifications of Taylor rule, and different truncation points in sticky information model (=4,6 and 8) draw a conclusion that none can switch the ranking position of three competing models. However, when we use survey data[40] to re-examine each competing model, although the model with inattentive features still outperform the baseline model, the ranking between sticky information and imperfect information data revision model changes. This contradict result may due to different types of sample data containing different information to favor different inattentive expectations.

Practically, there is no absolute best way to select an econometric method to estimate and evaluate one model. The Bayesian estimation approach by introducing a prior has its advantages, meanwhile, the most challenging factor is prior as well since its distribution needs to be determined or limited before carrying out estimation. Also, how to choose priors' distribution before implementing estimation is still a controversial point in recent studies (Fernández-Villaverde, 2010). Besides, it is crucial to note that Bayesian estimation method can only check model’s relative ability. Thus, it is still essential to check model's absolute ability in order to make fair comparison, which will be conducted in the following chapter through Indirect Inference approach.

[40] https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/ is where the survey of professional come from.

Appendix A to Chapter 3

Prior Interpretation

TABLE 3A-1 PRIORS MEAN OF PARAMETERS

Common Structural parameter		
	Elasticity of intertemporal substitution	1
	Sticky price degree	0.6
	Strategic complementary	0.15
Common Taylor Rule in three models		
	Degree of partially adjustment in Taylor rule	0.75
	Coefficient of inflation on Taylor rule	1.5
	Coefficient of output gap in Taylor rule	0.12
Common Forcing Variables in three models		
	AR coefficient of demand shock	0.5
	AR coefficient of cost-push shock	0.5
	AR coefficient of policy shock	0.5
	Standard deviation of demand shock	0.33
	Standard deviation of cost-push shock	0.33
	Standard deviation of policy shock	0.25
Note: The priors of parameter are mostly chosen from previous literatures, i.e., Miliani and Rajbhandari (2012), and Smets and Wounter (2003, 2007).		

TABLE 3A-2 PRIORS MEAN OF PARAMETERS

Imperfect Information model		
	output coefficient in output revision process	0
	inflation coefficient in inflation revision process	0
	AR term of shock in final revision process of x	0.5
	AR term of shock in final revision process of	0.5
	SD of measurement error of x	0.25
	SD of measurement error of	0.25
Sticky Information model		
	Share of updating firms (Mankiw and Reis,2007)	0.5
	Share of updating consumer (Mankiw and Reis,2007)	0.5
Note: The priors of parameter for SI model are chosen from previous literatures, from Mankiw and Reis (2007).and for IF model the priors of parameters borrow from Casares and Vazquez (2016).		

Appendix B to Chapter 3

Estimates without Survey Data

TABLE 3B-1 PARAMETERS ESTIMATE OF FULL-INFORMATION RATIONALITY

Prior distribution				Posterior distributions			
Params.	Distr.	Mean	S.D .	Mode	Mean	90% HPDIs/ Bayesian confidence bands	
	G	1	0.5	0.0167	0.0225	0.0051	0.0395
	B	0.6	0.05	0.7285	0.7257	0.6796	0.7733
	B	0.75	0.1	0.8913	0.8834	0.8473	0.9209
	N	1.5	0.25	1.3923	1.3891	1.0148	1.7447
	N	0.12	0.05	0.1940	0.1974	0.1197	0.2769
	B	0.5	0.15	0.8072	0.7995	0.7556	0.8452
	B	0.5	0.15	0.7015	0.6948	0.6352	0.7530
	B	0.5	0.15	0.2958	0.3094	0.1974	0.4257
	IG	0.33	1	0.1473	0.1564	0.1183	0.1943
	IG	0.33	1	0.0849	0.0878	0.0693	0.1049
	IG	0.25	1	0.2271	0.2301	0.2102	0.2494
Log marginal likelihood				-267.05			

Note:

- Metropolis-Hastings algorithm is applied to solve posterior distributions. 20000 draws with acceptance rate between 20% and 40%. and we discard the initial 20% of MH draw and keep 16000 draws.
- For the prior densities, we used beta (B), gamma (G), normal (N), and inverse gamma (IG) distributions.

Figure 3B-1 Full-Information Rational Expectation Multivariate MH Convergence Diagnosis

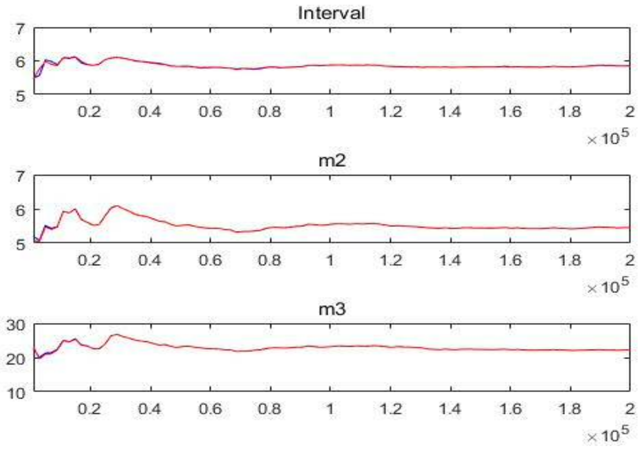
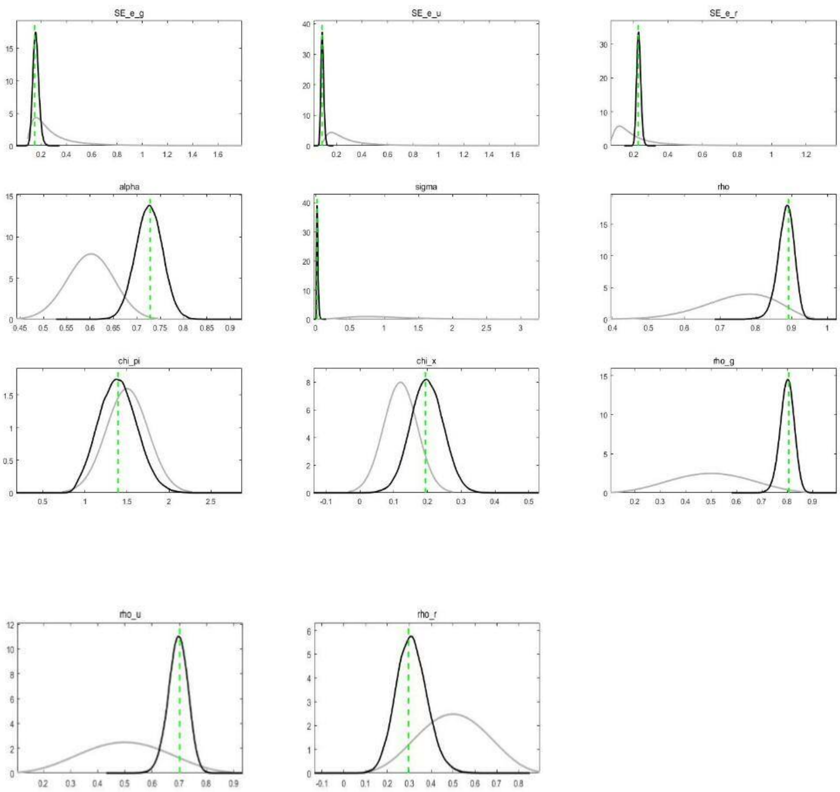


Figure 3B-2 Estimated Parameters Distribution of Full-Information Rationality



(Note: Black line: posterior distribution; green line: posterior mean)

Figure 3B-3 Full-Information Rational Expectation Smoothed Variables[41]

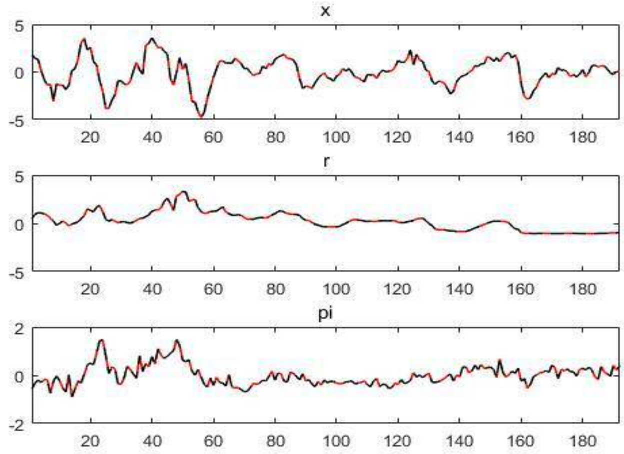
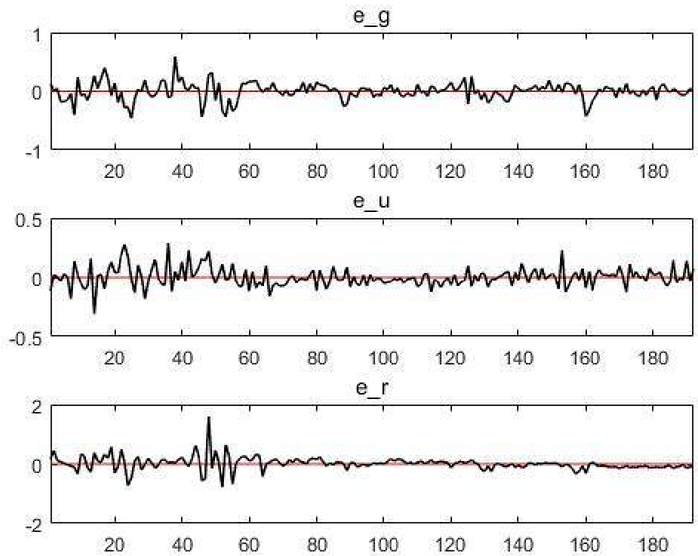


Figure 3B-4 Full-Information Rational Expectation Smoothed Shocks



[41] Dotted black line depicts the actually observed data, while the red line depicts the estimate of the smoothed variables ('best guess for the observed variables given observations') derived from Kalman smother at the posterior mode or posterior mean.

TABLE 3B-2 PARAMETERS ESTIMATE OF STICKY INFORMATION （J=4）

Prior distribution				Posterior distributions of SI (j=4)			
Params.	Distr.	Mean	S.D	Mode	Mean	90% HPDIs/ Bayesian confidence bands	
	G	1	0.5	0.0817	0.1092	0.0245	0.1894
	B	0.6	0.05	0.6314	0.6340	0.5685	0.6991
	B	0.75	0.1	0.9046	0.9002	0.8629	0.9372
	N	1.5	0.25	1.3863	1.3735	0.9735	1.7517
	N	0.12	0.05	0.1847	0.1848	0.1063	0.2646
	B	0.5	0.15	0.8101	0.8139	0.7558	0.8755
	B	0.5	0.15	0.7047	0.6940	0.6092	0.7785
	B	0.5	0.15	0.2848	0.2986	0.1891	0.4109
	IG	0.33	1	0.5252	0.5548	0.4500	0.6594
	IG	0.33	1	0.2455	0.2446	0.2159	0.2726
	IG	0.25	1	0.2265	0.2294	0.2105	0.2490
	B	0.5	0.25	0.1014	0.3084	0.0083	0.9264
	B	0.5	0.25	0.2612	0.2362	0.1257	0.3478
Log marginal likelihood				-247.36			

Note:

- Metropolis-Hastings algorithm is applied to solve posterior distributions. 20000 draws with acceptance rate between 20% and 40%. and we discard the initial 20% of MH draw and keep 16000 draws.
- For the prior densities, we used beta (B), gamma (G), normal (N), and inverse gamma (IG) distributions.

Figure 3B-5 Sticky Information (j=4) Multivariate MH Convergence Diagnosis

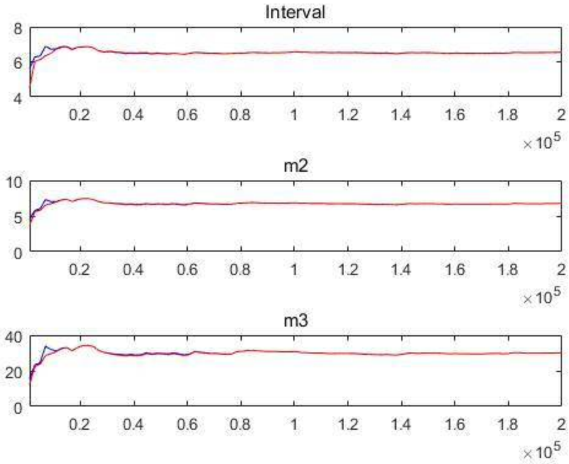
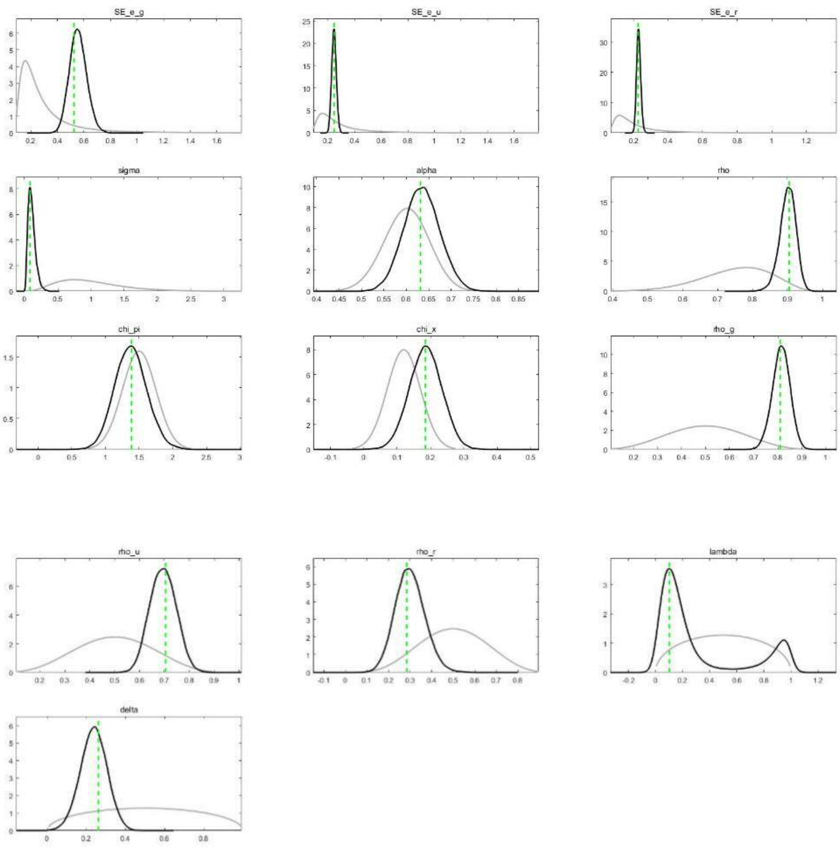


Figure 3B-6 Estimated Parameters Distribution of Sticky Information (j=4)



(Note: Black line: posterior distribution; green line: posterior mean)

Figure 3B-7 Sticky Information (j=4) Smoothed Variables

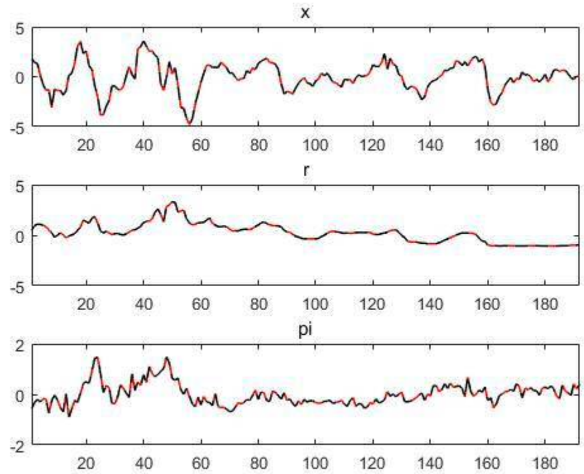


Figure 3B-8 Sticky Information (j=4) Smoothed Shocks

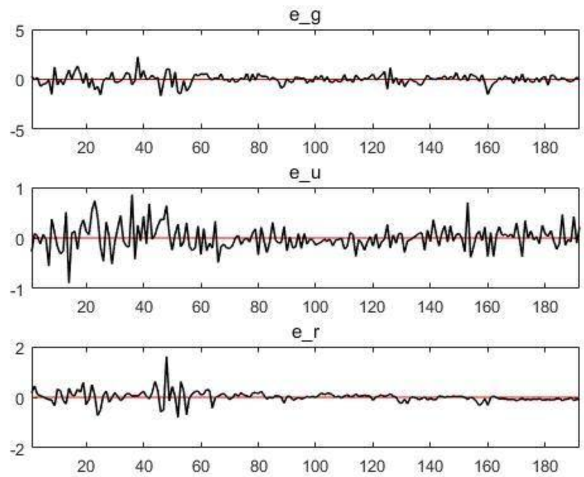


TABLE 3B-3 PARAMETERS ESTIMATE OF IMPERFECT INFORMATION DATA REVISION

Prior distribution				Posterior distribution			
Params.	Distr.	Mean	S.D	Mode	Mean	90% HPDIs/ Bayesian confidence bands	
	G	1	0.5	0.0535	0.0899	0.0152	0.1636
	B	0.6	0.05	0.7298	0.7389	0.6831	0.7960
	B	0.75	0.1	0.8698	0.8801	0.8420	0.9178
	N	1.5	0.25	1.0182	1.0884	0.8409	1.3467
	N	0.12	0.05	0.2031	0.1962	0.1311	0.2608
	N	0	2	1.2492	1.8500	0.5268	3.0401
	N	0	2	0.9908	1.1198	0.5698	1.6884
	B	0.5	0.2	0.8310	0.7252	0.4802	0.8929
	B	0.5	0.2	0.8585	0.8535	0.8118	0.8923
	B	0.5	0.15	0.4986	0.6186	0.3313	0.8647
	B	0.5	0.15	0.3678	0.3657	0.1720	0.5549
	B	0.5	0.15	0.2172	0.2235	0.1209	0.3205
	IG	0.33	1	0.1426	0.2710	0.0891	0.4923
	IG	0.33	1	0.1447	0.1551	0.0827	0.2212
	IG	0.25	1	0.2147	0.2181	0.2001	0.2363
	IG	0.25	4	0.2912	0.3270	0.0633	0.5879
	IG	0.25	4	0.0726	0.0808	0.0513	0.1089
Log marginal likelihood				-254.08			

Note:

- Metropolis-Hastings algorithm is applied to solve posterior distributions. 20000 draws with acceptance rate between 20% and 40%. and we discard the initial 20% of MH draw and keep 16000 draws.
- For the prior densities, we used beta (B), gamma (G), normal (N), and inverse gamma (IG) distributions.

Figure 3B-9 Imperfect Information Multivariate MH Convergence Diagnosis s=3

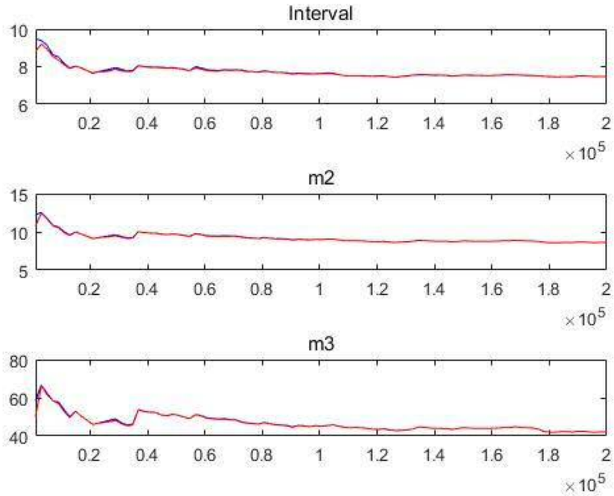
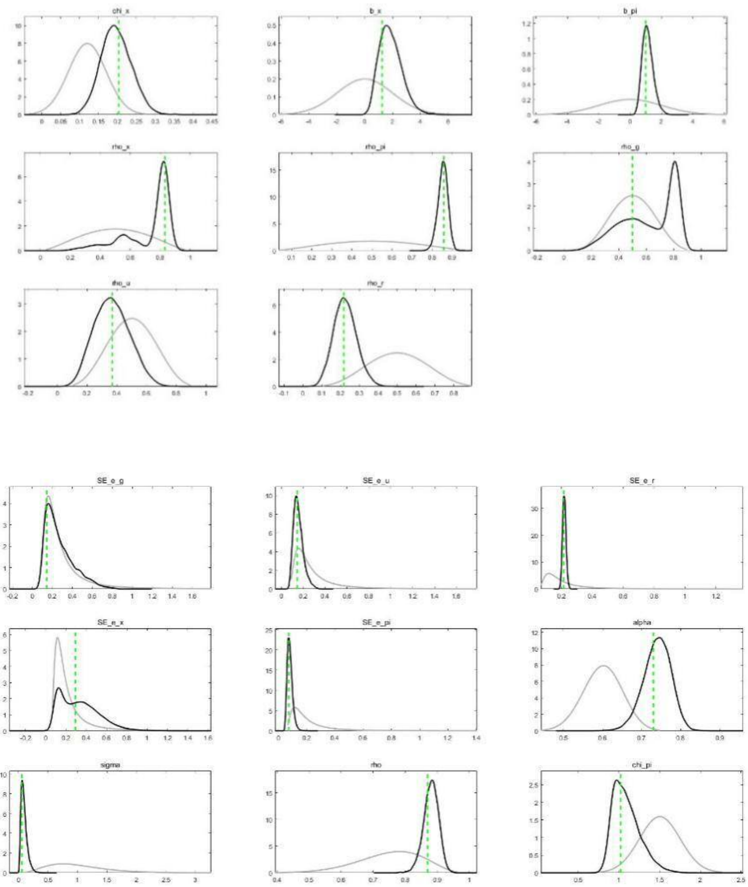


Figure 3B-10 Estimated Parameters Distribution of Imperfect Information Model s=3



(Note: Black line: posterior distribution; green line: posterior mean)

Figure 3B-11 Imperfect Information Data Revision Smoothed Variables s=3

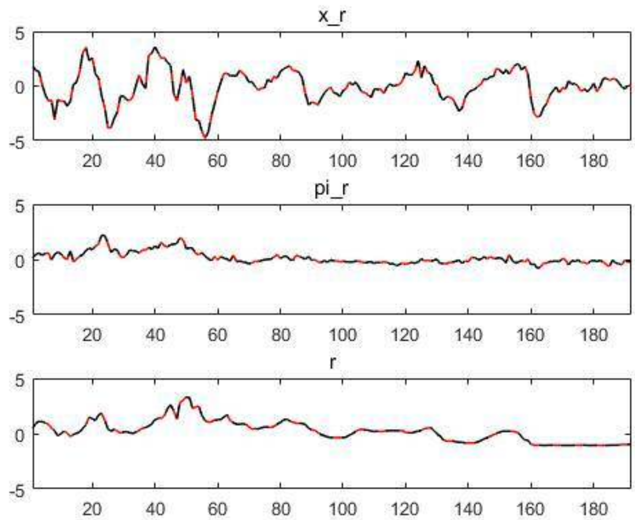
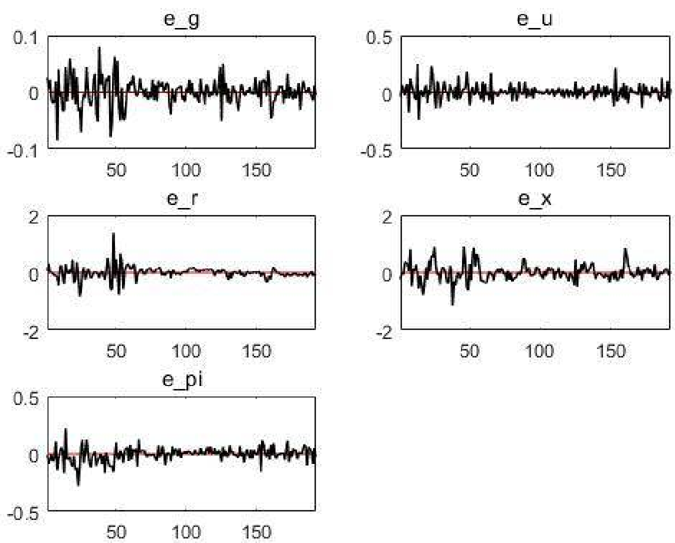


Figure 3B-12 Imperfect Information Data Revision Smoothed Shocks s=3



Appendix C to Chapter 3

All data are of a quarterly frequency and are seasonally adjusted. And all the series are demeaned before estimation.

United States Data Source

1) Effective Federal Funds Rate=FEDFUNDS, the federal funds rate is divided it by four to express it in quarterly rates. The observable is matched to the variable , where ————

2) The real-time data[42] from Real-time data set for macroeconomists are collected from Federal Reserve Bank of Philadelphia, the real-time Real GDP=ROUTPUT initial released in 2016Q1 (i.e., which only release real-time Real GDP up to time 2015Q4), then the quarterly real-time GDP is the deviation of the natural logarithm of total real-time GDP, potential output is from its HP filter. For imperfect information model to construct the revised observables corresponding to output gap up to time 2015Q4, the real-time data released after one period (2016Q1) as well as the real-time data of GDP released after three periods also applied (2016Q3).

3) For the real-time Implicit Price Deflator=P. Index level initial released in 2016Q1 (i.e., which only release real-time Implicit Price Deflator up to 2015Q4), seasonally adjusted, also from the real-time data set from Federal Reserve Bank of Philadelphia, the series is demeaned. The real-time inflation .Similarly, to construct the revised observables correspond to

[42] https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files is where the real-time data set from.

inflation up to time 2015Q4, the real-time data of Implicit Price Deflator released after one period and the data released after three periods also be used.

4) The survey data using in robust check section is the median of Survey of Professional Forecaster one quarter ahead forecast of GDP deflator and real GDP. In imperfect information data revision model, both one-quarter ahead and four-quarter ahead forecast has been used to construct the final revised observables.

Appendix D to Chapter 3

TABLE 3D-1 SUMMARY ESTIMATION RESULTS OF DIFFERENT EXPECTATION FORMATION (WITH FRED REVISED DATA)

Prior distribution				Posterior distributions (mean)		
Params.	Distr.	Mean	S.D	FIRE	SI (j=4)	IF
	G	1	0.5	0.0471	0.1350	0.1380
	B	0.6	0.05	0.7151	0.6238	0.7283
	B	0.75	0.1	0.8136	0.8256	0.8646
	N	1.5	0.25	1.1768	1.1392	1.1293
	N	0.12	0.05	0.2178	0.2081	0.1921
	B	0.5	0.15	0.8021	0.8114	0.5284
	B	0.5	0.15	0.6873	0.6850	0.4118
	B	0.5	0.15	0.2724	0.2705	0.2268
	IG	0.33	1	0.1596	0.5316	0.2013
	IG	0.33	1	0.0898	0.2393	0.1447
	IG	0.25	1	0.2214	0.2209	0.2205
	N	0	2	-	-	1.8349
	N	0	2	-	-	0.8741
	B	0.5	0.2	-	-	0.8157
	B	0.5	0.2	-	-	0.7099
	IG	0.25	4	-	-	0.4205
	IG	0.25	4	-	-	0.1192
	B	0.5	0.2	-	0.1376	-
	B	0.5	0.2	-	0.2642	-
Log marginal likelihood				-249.13	-232.39	-245.01
Bayes Factor relative to the FIRE				1		

Chapter 4
Testing and Estimating New-Keynesian Type Models with Inattentive Feature through Indirect Inference Approach

4.1 Introduction

This Chapter proposes an approach to evaluate reduced-form New-Keynesian DSGE models which are basing on indirect inference and this method is employed to the previous chapter's three competing models concerning the cases of attentive expectation (i.e., the model with full-information rational expectation) and inattentive expectation (i.e., the sticky information expectation model and the imperfect information data revision expectation model). The approach commonly used by pervious economists to solve the problem which has been existing for a long time to asses a calibrated and estimated DSGE model is simply comparing the features of simulated data and those of true data, in which the sample data are stimulated by calibrated and estimated DSGE model. In this chapter, we choose the indirect inference which can be divided into two stages as a more rigorous approach to evaluate DSGE models.

In the first stage, we will implement indirect inference as a calibrated-based testing method to test each competing model by given initial presumptive parameters. The content of the test can be understood as that through comparing the unrestricted VAR estimates (derived from the simulation data) with the alternative unrestricted VAR estimates (derived from the actual data), we can confirm whether these two groups of parameters' estimates of the auxiliary model are 'close enough' (i.e., each competing DSGE model is correctly specified). If the result shows one model is correctly specified, then the distance of the unrestricted VAR estimates and the alternative unrestricted VAR estimates should be minimized. In other words, the assumed model and the 'real model' will not be far away from to each other. The apparent strength of the indirect inference test method is that it is unnecessary to specify each competing model as the alternative hypothesis. However, we need to identify the auxiliary VAR what is generated by each competing model.

In the second stage, we will implement the estimation-based indirect inference test. In this stage, the Indirect Inference is not just used to gauge the 'distance' between the theory and the reality through using the auxiliary model but also finding a set of parameters to minimize such distance. The extra searching step the distinction between this stage and the last stage.

In this chapter, we have three main purposes. The first purpose is to take indirect inference as a calibration-based test approach. We intent to evaluate the already estimated model (in the previous chapter through Bayesian estimation approach). The focus of our test is to detect whether the data which are simulated from the three competing structural models can explain the actual data. The evaluation of three rival models are done through an indirect inference test which is basing on comparing Wald statistics that concentrate on the total capability of the model to fit the overall dynamic behavior of the actual data.

The second purpose, being distinguished from the first purpose, is to use the indirect inference as the estimation-based test approach whose duty includes exploring the optimal set of structural parameters which enables the model to copy the trajectory of the behavior of actual data to the maximum extent. In the second stage, indirect inference testing process will introduce the optimal searching procedure. To be more specific, the nature of being fixed of the model parameters is an overly strong condition for testing and contradistinguishing models. Seeing the parameter values of the candidate model could be estimated or calibrated within a permissible scope throughout the theoretical structure of the model, it is probable for a rejected model with the presumptive set of parameters to pass the test when it with another set of parameters. To have a fair result of the testing, it is necessary for investigators to find a set of 'good' structural parameters. Thus, we estimate the models to get the optimal sets of parameters before the evaluating process. The third purpose is to reach absolute performances of the models for comparing, which can be realized through the introduction of the distributions of the two groups of estimated parameters of the auxiliary models. From this point of view, it is the most significant difference comparing to the evaluation of Bayesian estimation.

To sum up, despite the conclusion of the 'best' model found through Bayesian estimation approach, it is necessary for us to verify the result through other method to give credit of the 'best' model.

4.2 Description of Indirect Inference as Evaluation Method

In this chapter, Indirect Inference is applied for measuring how close the three models are to real world. The principle of this method is basing on the idea that through comparing the moments of simulated data and actual data, a model can be measured in an absolute way in a framework that contains an auxiliary model. Two characteristics of this method make it superior to other solutions. Firstly, a statistical threshold given for filtering models divides the tested models into two groups of qualified and unqualified. Secondly, it enables us to evaluate the distance statistically in the middle of the theoretical models (simulated data) and the real world (actual data).

The approach of Indirect Inference already applied diffusely in the field of estimation by scholars (Gregory and Smith, 1991, 1993; Gallant and Tauchen, 1996; Keane and Smith, 2003; Minford, Theodoridis et al., 2009). For instance, in the year of 2011, Le et al. applied the same method to evaluate the model of the US economy which was constructed by Smets and Wouter (2007) and ultimately obtained a rejected consequence on the testing. In this thesis, our evaluation will take the common procedure of indirect inference evaluation for reference from previous studies (Le et al., 2011, 2016; Minford and Ou, 2013; Liu and Minford, 2014; Minford et al., 2015)

It is worth noting that there are two most relevant papers regarding to our research topic through using indirect inference method. One is published by Vázquez et al. (2010, 2012) who assess the importance of data revisions on the estimated monetary policy rule. The estimation conducted through indirect inference finds that the ignorance of the data revision process may not result in a serious drawback in analyzing monetary policy based on New-Keynesian framework. Our assumption substitutes the subjects who involve imperfect information data revision issue with households and firms instead of monetary authority. Meanwhile the subjects can perfectly observe monetary policy. The other related paper is published by Knotek and Edward (2010) who investigates a single equation model incorporating both sticky price and sticky information and detect that such a model can match the real world in both dimensions of micro and macro after estimating it through indirect inference.[43] However, we are more interested in full-structural model rather than single equation model.

The complete estimation of three competing models through Bayesian approach has been done in Chapter 3. With the consequence that competing models with inattentive features are preferable, in this chapter, we will turn to re-evaluate each model focusing on its overall dynamic properties in connecting with the actual data by adopting Indirect Inference as the new evaluation method.

While we are applying Indirect Inference to evaluate an existing structural model, two factors are inevitable in the process of stimulating the data from theoretical model. One is the parameters of theoretical model and the other one is the distribution of the errors. We evaluate the theoretical model through Indirect Inference test which is based on the contradistinction of the actual data with the simulated data obtained from the theoretical model with the assistance of auxiliary model. In this chapter, VAR (i.e. Vector auto-regression), which is a stochastic process model used to capture the

[43] Knotek and Edward (2010) finds that when the empirical Phillips curve is embodied with sticky prices and sticky information, its ability tends to be improved to match the macro data.

linear interdependencies among multiple time series. is selected as the auxiliary model.

There are two reasons for us to choose VAR as the auxiliary model. Firstly, the structural model can always be manifested as a restricted VARMA (i.e. Vector Auto-regression Moving-Average), which is close to a VAR representation. Secondly, VAR can reflect two properties of the data. They are the relation of variance-covariance among the variables through the co-variance matrix of the VAR disturbances, and the dynamic behaviour of the data via the dynamics and the impulse response functions of the VAR. The Wald statistic, which is derived by the distributions of these functions of the parameters of VAR, and TM distance (normalized t-statistics), which are derived from a function of these parameters can be regarded as two criteria of the testing model to measure the distance to the reality. From the consequence of the testing model regarding the two criteria, we can judge whether the hypothesis, which assumes the testing model is correctly specified, is accepted or rejected. If the consequence shows rejected, it implies that the theoretical model cannot reproduce the actual data significantly. While the consequence of being non-rejected implies the data generated from the theoretical model not different from the actual observed data significantly.

Wald Test Statistics

In general, the Wald testing process can be summarized into three general steps as follows. Firstly, to derive the structural errors by using the observed actual data and parameters calibrated or estimated in the model. There are two ways to construct the errors under two different circumstances. When the structural model possesses no expectation terms, the structural errors can be backed up straight from the structural equations and the actual data. While under the situation that structural equation includes the computation of expectations, the method used is the robust instrument variables estimation[44]. Therefore, the expected future variables of output gap and inflation are approximated by the fitted values of VAR (1), which are the linear combinations of the lagged three main variables. Secondly, the structural errors are bootstrapped to be employed to produce the pseudo data which are based on candidate theoretical model. After that, an auxiliary VAR model is fitted to each set of pseudo data and the sampling distribution of the coefficients of the auxiliary VAR model are achieved from these estimates of the auxiliary model. Thirdly, the Wald statistic is calculated to judge whether or not the functions of the parameters of the auxiliary VAR model estimated on the actual data lie within the confidence interval implied by this sampling distribution[45]of the coefficients of the auxiliary time series model (Minford et al., 2015; Fan et al., 2016).

The test is through comparing the performance of the overall capacity of the model with the dynamics performance of actual data to determine whether the hypothesis is qualified. The process of comparison is available through checking if coefficients of the actual-data-based VAR lie in the acceptable range of the theoretical model's implied joint distribution. By the means of that, we can even inspect the model's capability of directing the dynamics and variances of the data.

In this chapter, VAR (1) is used as the auxiliary model by us and is treated as the descriptors of the actual data for three main macro variables (i.e., output gap, inflation, and interest rate). The Wald statistics is calculated from the VAR (1) coefficients and the variances of the three main economic variables. Therefore, the Wald test statics is a criterion to determine whether the observed dynamics and volatility of the selected three main variables are interpreted by the simulated joint distribution of

[44] Robust instrument variables estimation is suggested by McCallum (1976) and Wickens (1982), in which the lagged endogenous data are set as instruments, and the fitted values are computed from a VAR (1) what is used as the auxiliary model during evaluation procedure as well.

[45] By estimating the auxiliary model VAR on each pseudo sample, we can have the distribution of the estimates. The dynamics properties are captured by VAR estimates, while the volatility properties can be captured by the variance of the main variables. For the individual estimates, the confidence interval (95%) is calculated directly from their bootstrapped distribution.

these at a given 95% confidence level. The Wald statistics is formulated by,

The equation above is a function of the gap between and . is the vector of VAR estimates of the selected US data descriptors. is the arithmetic mean of the N estimated vector of VAR estimates derived from bootstrap simulations. is the variance and covariance matrix of the distribution . In addition, and are the actual data sets and simulated data sets respectively. is the vector of the parameters of the theoretical model. Then we can check the positions of Wald test statistics within the distribution generated by model.

Indirect Inference can be proceeded by comparing the percentile of the Wald distribution. In detailed, for a 5% significant level, a percentile above 95% would not lie outside the non-rejection area. The distribution of as well as the Wald statistics are obtained through bootstrapping method.

Transformed Mahalanobis Distance (Normalized t-statistics)

The TM statistic is used in the situation which we are hardly able to distinguish the models' relative performances. For instance, there are two or more specified models rejected simultaneously by Wald test statistics, we have to use the TM statistic to rank these models after comparison. Additionally, the TM provides a way to examine how bad the model is by observing how far it deviates away from 1.645. The bigger the number is, the worse the model fit. The Transformed Mahalanobis (TM) distance is defined as follows.

Herein, the TM distance is the transformation of the Wald test statistics.[46] Where
is the Mahalanobis distance (value of Wald statistics) using the actual data,
is the 95% critical Mahalanobis distance from simulated data (is the value of the Wald statistics falling at 95th percentile of the bootstrap distribution), and p is the number of parameters concerned or defined as degree of freedom respectively.

4.3 Indirect Inference Estimation Results

4.3.1 First Stage: Results of Calibration-based Indirect Inference Test

The testing steps presented above will be employed to test three rival models by using the US real-time quarterly data from 1969 to 2015 (the survey data will be used in robust check section). The data (variables) has been well defined in chapter 3 Appendix in the previous chapter. VAR (1) is taken as the auxiliary model in this chapter and the estimation of VAR is implemented with three economic observables: output gap, inflation, and nominal interest rate, which are quarterly observables.

First, we test the New-Keynesian three-equation models for both cases with and without inattentive features. The baseline model under full-information rationality assumption with an interest-rate smooth Taylor rule. Moreover, all three errors are presumed to follow AR (1) processes, which is in line with the previous studies. We evaluate the three competing models basing on the actual errors derived from estimation on the actual data. Moreover, it requires an estimation of the model's structural errors which are the residuals in each equation of the structural model given by the actual data and the expected variables in that equation. The residuals of

[46] This function of Transformed Mahalanobis distance (normalized t-statistic) is based on Wilson and Hilferty (1983)'s method of transforming Chi-square distribution into a standard normal distribution calculated.

demand, cost-push, and monetary policy are estimated respectively. There are two extra AR (1) processes corresponding to final data revision processes in the model under imperfect information data revision assumption,

After evaluating each competing model, we can assess each mechanism of them. The model with full-information rationality assumption has been argued failure to generate delay response by the previous studies. Thus, two alternative approaches have been proposed by recent studies in order to remove such a fail. The two approaches are modeling with sticky information, and modelling with imperfect information data revision respectively. Regarding to the former approach, the economic agents adjust their decisions with delaying behavior and such delay behavior is generated by the information costs. On the other hand, concerning the latter approach, the economic agents adjust their decisions with delaying behavior due to data revision issue. The two explanations, which have been suggested from the previous studies to remedy the weakness in the baseline model, are selected to be examined in this thesis. Indirect inference test (full Wald test) is employed in this chapter to check each competing model's overall data dynamic performance in an absolute way. The transformed Mahalanobis distance (normalized t-statistics) is also used to measure how similar the to-be-examined model is to the real world.

4.3.1.1 Calibration Parameters (Initial-Presumptive parameters)

The overview of all structural parameters along with their initially presumptive values are presented in Table 4-1 (Part 1) and most of the values are identical to the prior means which have been used in the previous chapter.

TABLE 4-1 (PART 1) STARTING CALIBRATION STRUCTURAL PARAMETER VALUE

Parameters	Definition	Values
Common Parameters		
	Time discount factor (fixed)	0.99
	Price stickiness	0.6
	Elasticity of intertemporal substitution	1
	strategic complementary parameter (fixed)	0.15
	Degree of partially adjustment in Taylor rule	0.75
	Coefficient of inflation on Taylor rule	1.5
	Coefficient of output gap in Taylor rule	0.12
SI Expectation Model		
	Share of updating firms (Mankiw & Reis, 2007)	0.5
	Share of updating consumer (Mankiw & Reis, 2007)	0.5
IF Expectation Model[47]		
	output coefficient in output revision process	0.5
	inflation coefficient in inflation revision process	0.5

Table 4-1 (Part 2) Starting Calibration Parameter Value of AR Coefficients[48]

FIRE Model		
	AR coefficient of demand shock	0.90
	AR coefficient of cost-push shock	0.79
	AR coefficient of policy shock	0.59
SI Expectation Model		
	AR coefficient of demand shock	0.89
	AR coefficient of cost-push shock	0.79
	AR coefficient of policy shock	0.64
IF Expectation Model		
	AR coefficient of demand shock	0.67
	AR coefficient of cost-push shock	0.56
	AR coefficient of policy shock	0.30
	AR term of shock in final revision process of x	0.41
	AR term of shock in final revision process of	0.61

The AR coefficients' parameters and correspondent values shown in Table 4-1 (Part 2) are achieved from the sample estimation of US real-time data for each applied

47 The initial null hypothesis is that , meaning not well-behaved revision processes.

48 The AR coefficients of the structural errors implied by the models, all of them are sample estimated base on the real-time data.

model.[49] The presumptive parameters' values (calibration value) are largely in line with the mean values of priors we adopted in Chapter 3.

4.3.1.2 Comparison through Calibration-Based Testing

The model cannot be bootstrapped without the solution of the structural error which can be reached if the observed actual data and presumptive parameters are given. As a rule, the times of bootstrapping is normally set as 1000. Following by this step, the test statistics are reached through examining the distribution of simulated pseudo samples. The main focus of this section is to shows that testing results by using presumptive parameters. The implementation of the overall model performance test is realized through the combination of dynamics parameters and volatility parameters.

TABLE 4-2 COMPARISON TM DISTANCE BY USING CALIBRATION PARAMETER

Model	Full Wald percentile %	TM by using Calibration Parameter
FIRE Model	100	4.1538
SI (j=4) Model	99.4	2.7338
IF Model	100	28.5625
Note: Above results, VAR (1) has been used as auxiliary model.		

The calibration-based testing results are shown in Table 4-2. Since the full Wald percentiles of the model all above 95, the models are implied to be not fit for dynamic properties of the actual data as they are not falling within the non-rejection area (i.e., 95 % confident interval). The smaller the value of Wald percentile is, the better level of fit its model reaches. The values of TM distance, which are higher than the norm of 1.645, imply the same trend that the Wald percentile tell us by displaying the extent

[49] It is doubtable that OLS is a biased estimator of the auxiliary model, due to the presence of lagged endogenous variables as regressors. However, it should not influence the power of test as the identical auxiliary model and estimators are applied for depicting the simulated data and the actual data. In other words, the same bias is translated into each model. Such that, in fact indirect inference is used to test whether the model-based OLS-estimated auxiliary model would generate the actual-data-based OLS-estimated auxiliary model.

of the gap. Basing on the initial presumptive error properties of US reduced-form New-Keynesian DSGE type models, the three competing models do not fit the actual data accurately. However, according to the result of the initial calibration-based test by using the presumptive parameters, it shows that the model with sticky information wins the best performance, which can be assessed through comparing the TM distance (normalized t-statistics). Surprisingly, what is contrary to the result obtained by Bayesian estimation approach that the model with imperfect information data revision is far worse than its rivals. However, the contradiction of the calibration-based testing result may be due to our initial presumptive parameters which is not the best options to closely copy the overall dynamic properties of the actual data. Thus, indirect inference will be conducted later as an estimation method aiming to search the 'best' set of parameters for the three competing models respectively.

4.3.1.3 Robustness Check

1) *Higher-Order VAR as Auxiliary Model*

We have implement the calibration-based indirect inference test towards our three competing models. Suffering from the same problems of robustness with the previous scholars, we realize that the choice of auxiliary model may influence the testing results, since it plays a role of independent intermediary to evaluate the gap between the theoretical model to the reality. To mitigate our concerns about the interference regarding auxiliary, the higher-order VAR as auxiliary models are introduced in robustness check. We apply the estimates of the coefficient matrix and the volatility of the data as the descriptors selected in the auxiliary model.

The results shown by the higher-order VAR auxiliary model robustness check, the result of VAR (1) is robust, which gives an answer to why a VAR (1) could approximate the DSGE models. Higher-order auxiliary VAR model, which was for model evaluation originally, is used by us out of detecting whether the ranking of three competing models is robust. We achieve this transformation of aim by the fact that using a VAR as the auxiliary model with higher order contributes to the strictness of the model evaluation because it requires more details of the data to be fitted. Although, in general, applying higher-order VAR as auxiliary model gives worse result for each competing model (i.e., since it requires data to match more specific characteristics) and which could be a way to develop the evaluation when the difference of models' performances is not obvious through less order auxiliary model, VAR (1).

In summary, in this section we mainly focus on checking whether the rank of the three competing models will change when a higher-order VAR model is used. The results show that the model with sticky information expectation under more stringent condition still outperform the alternative models.

TABLE 4-3 MODEL PERFORMANCE UNDER DIFFERENT AUXILIARY MODELS

Competing model	FIRE	SI (j=4)	IF	FIRE	SI (j=4)	IF
Auxiliary model	VAR (2)			VAR (3)		
TM Distance (Full Wald percentile %)	27.0317 (100)	10.3896 (100)	37.7490 (100)	29.4240 (100)	14.1153 (100)	45.0990 (100)

2) *Using Alternative data resource: survey of professional forecaster data of output gap and inflation*

Concerning different types of data resource may provide extra information in favor of different model, we implement the same testing procedure in this section by using survey of professional forecaster data. There are two noteworthy things in the testing results. The first one is that the survey data does not provide extra useful information to improve models' performance excepting for the model with imperfect information data revision. The TM distance of imperfect information data revision model decreases from 28.5625 to 15.7632, which indicates that the distance between theoretical model and 'real-world model' has been narrowed down. The second thing is that although none of them can pass the test, it implies that the survey data contains some extra information that help us to make further distinguish between the baseline model and sticky information model. However, such calibration-based testing result may due to the initial presumptive parameters may be not the best options to describe the survey data, so that it is essential to carry out indirect inference estimation in the next stage to search the 'best' collection of parameters which can be applied to maximum degree narrow down the gap between theoretical model and the reality to make a fair comparison for models.

TABLE 4-4 COMPARISON TM DISTANCE BY USING CALIBRATION PARAMETERS (WITH SURVEY DATA)

Model	Full Wald percentile %	TM by using Calibration Parameter
FIRE Model	100	17.9522
SI (j=4) Model	100	4.1554
IF Model	100	15.7632
Note: Above results, VAR (1) has been used as auxiliary model.		

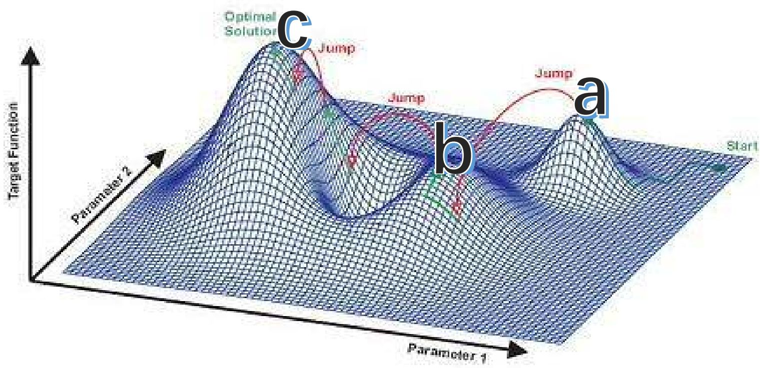
4.3.2 Second Stage: Results of Estimation-based Indirect Inference Test

Since the initial presumptive parameters used in the first stage may not lead to the optimal results, we decide to try to find out another collection of parameters to interpret the way of the data's generation in this chapter. If there not exists such a group of parameters which enables the model to pass indirect inference test, the model will be judged as rejected. As the aim of the second stage, searching the parameter set leading the model to replicate the real world as well as possible which can be defined as 'estimation-based Indirect Inference test', through which the chance of being accepted to the testing model will be maximized.

Essentially, this stage gives a way to solve the problem of parameter uncertainty. In practice, we can reduce the parameter uncertainty in a direct way by checking the Wald statistic derived from the group of parameters for the model. In detailed, the more the Wald statistic decreases, the better the parameter set performs. Herein, an effective algorithm basing on Simulate Annealing (SA) is introduced to search the optimal parameter set by starting from an extensive scope around the initial values along with random jumps around the space. With SA algorithm, we can have the lowest value of full Wald statistic for three rival models.

The SA algorithm refers to a stochastic optimization based on Monte-Carlo iterative solution strategy. The principle is inspired by the annealing process of metal heating and cooling through which the temperature of the object will be controlled to increase the size of the metal's crystals and reduce its defects. By mimicking the mechanism, the SA searches for the probabilities with lower energy to minimize the defects of crystal (resemble that of the steps of minimizing Wald statistics in estimation process of indirect inference). It tries to find the optimal parameter set repeatedly until the system reaches a minimum value of Wald statistics, or until a given computation budget has been exhausted. Since the principle of accepting a less optimal consequence temporarily, SA can reach the optimal consequence in a global scale instead of being trapped in local optimum. For example, according to Figure 4-1, SA mechanism allows one to search over the whole apace starting from the initial state (in an indirect inference estimation process, a current state is equivalent to the group of structural parameters) and jump to nearby local optimal 'a' and continue to search toward global optimal 'c'. The less optimums in between 'a' and 'c', taking 'b' as an example, will be accepted as a 'springboard' which one can jump to in order to jump and search the other space to reach the global optimum.

Figure 4-1 Simulated Annealing (SA) Avoiding Getting Stuck in Local Optimum[50]



Overall, in the application of indirect inference estimation, SA is used to seek the optimal set of parameters, which will facilitate to discover lowering Wald statistic until the computation budget used up. To carry out the numerical iterations to minimize the Wald statistics, the initial values of the parameters of structural models are required. Here, the starting values are the values of the presumptive parameters, such presumptive parameters are plausible and from the previous studies, meanwhile, we permit the parameters to seek around -0.5 to +0.5 of their starting values under estimation.

To implement estimation-based Indirect Inference test, the VAR (1) needs to be used continuously as the auxiliary model to give a reference substance for the estimated models to those of the calibrated models. The VAR (1) are used as descriptors of the coefficient matrix and the variance of the data. Just like the previous testing exercise.

In the first stage, the structural parameters were assigned by the initial presumptive values but those are selected in line with the commonly accepted values from the previous studies. Being distinguished from the first stage, the second stage uses indirect inference as estimating method to re-assess the three competing models on

50 Figure Source: http://www.frankfurt-consulting.de/img/SimAnn.jpg

their grooves based on actual data, which means that the restriction of initial presumptive parameters has been released.

We may expect that, by indirect inference estimation or simulated-annealing estimation, estimated version of the three competing models would behave no worse than that found in the first stage. Seeing that when we take calibration values as the initial presumptive ones to assign the structural parameters, the SA mechanism will begin to explore from these initial presumptive values to substitute for them with 'better' values based on the actual data if only a minimum Wald statistic can be discovered. The process will be terminated when the Wald statistic can no longer be reduced, which implies that we have discovered the 'best' estimates of the structural parameters. The Simulated Annealing method, which facilitates to adjust the initial presumptive values, is helpful for the models to pass the test.

4.3.2.1 Estimation-Based Indirect Inference Testing Results: Full Information Rationality Assumption Model

The Simulated-Annealing-estimation-based test as well as the Bayesian-estimation-based test with respect to the three competing models for US economy are presented in Table 4-5, Table 4-6 and Table 4-7 respectively. The numbers in the column regarding the indirect inference estimation are obtained through SA estimation method. The scope of the value of parameters during SA exploring is limited within plus or minus 50% of the presumptive values of coefficients.

The main idea of indirect inference as an assessment methodology is to test the existing model to detect whether the structural parameters are capable to generate the actual data. However, if these initial presumptive parameters cannot be used to explain the generating process of the actual data, another set of parameters may be somewhere existed and can be applied to explain how the actual data is generated. If the model with initial presumptive parameters already fall within the non-rejection scope, it is still necessary to explore another group of parameters that can narrow the gap in the middle of the theoretical model and the reality, which leads to better testing results. The ‘best’ set of the structural models’ parameters are those to the maximum degree to shorten the distance between theoretical model and the reality.

In the second stage, we aim to explore the ‘best’ collection of parameters throughout the entire parameter space by the implementation of Indirect Inference without changing the signs of parameters as an estimation-based test approach. The minimized value of the distance (Mahalanobis distance) is captured for each competitor over the US sample periods through a Simulated Annealing algorithm. The ‘best’ collection of parameters that can furthest shorten the distance between the theory and the reality will be used for our estimation-based test. Using these optimal sets of parameters to compare models can reduce the unfairness in model comparisons.

Table 4-5 displays the estimation results of the model with full-information rationality assumption (FIRE Model). Overall, the estimated values of parameters of the FIRE model through indirect inference estimation of are not significantly far away from those obtained by Bayesian estimation. However, some distinguished cases exist. Particularly, the estimated value of the elasticity of intertemporal substitution is 0.5180, which is quite higher than that obtained from the Bayesian estimation. Besides, the same trend can be found in the value of price stickiness versus that of Bayesian estimation. Subsequently, examining the estimates of the major behavioral parameters of FIRE model, we toward to examine the parameters of the monetary policy function, which are based on standard interest-rate smoothed Taylor rule (1993). Regarding to the estimated coefficients of monetary policy, excepting which is increased less than 8%, the other two (i.e., and) both increase around 35% comparing to their estimated values achieved from Bayesian estimation. Within the system, all the three stationary shocks are quite highly persistent and two of them, excepting for the AR coefficient of monetary policy which is increased above 60% than that obtained through Bayesian estimation, are similar to the Bayesian estimated results.

In detailed, through SA estimation, the estimated value of is 1.5079 which is slightly higher than that obtained by Bayesian estimation. The two estimates regarding different estimation methods are both close to the initial calibration value (i.e., 1.5). The estimated value of the reaction to output gap is 0.1439 which is lower than that obtained by Bayesian estimation, which indicates that the monetary policy does not seem to react forcefully to the output gap level. Moreover, the parameter of interest rate smoothness which is estimated to be 0.6580 and lower than that obtained through Bayesian estimation. However, it is not far away from the initial presumptive value (i.e., 0.75). Besides. the AR coefficients regarding to the three exogeneous stationary shocks which are demand shock, cost-push shock and monetary policy shock are estimated to be very persistent, which are 0.8587,0.7318 and 0.8155 respectively.

Furthermore, the test statistic implies a Wald percentile of 64.8, so the FIRE model is not rejected at the 5% significant level. In practice, the Wald statistic is within the non-rejection region of the bootstrap distribution. Overall, many of the estimates obtained through SA estimation have shifted away from the estimates obtained through Bayesian estimation for a distance (e.g., the elasticity of intertemporal substitution is increased around 97% higher than the Bayesian estimated value what is 0.0225. The SA estimated value of price stickiness is around 25% higher than the counterpart of Bayesian approach). It is indicated in Table 4-5 that the model estimated with SA estimates performs better than the model estimated with Bayesian estimates in fitting the actual data. The reported Wald percentile has gain the significant reduction comparing with the one obtained through using Bayesian estimates. The full Wald statistics implies that the FIRE model with SA estimates fall within the non-rejection area, meaning that the model cannot be rejected at a chance of 95%. Furthermore, the model with Bayesian estimates performance worse than the model with the initial presumptive parameters (calibration parameters).

TABLE 4-5 ESTIMATES OF FIRE MODEL

Parameters	Starting Calibration	Bayesian Estimates	SA Estimates
	1	0.0225	0.5180
	0.6	0.7257	0.9677
	0.75	0.8834	0.6580
	1.5	1.3891	1.5079
	0.12	0.1974	0.1439
	0.86	0.7995	0.8587
	0.73	0.6948	0.7318
	0.82	0.3094	0.8155
Full Wald %	100	100	64.8
TM (normalize t-statistic)	4.1538	26.0498	0.6587

4.3.2.2 Estimation-Based Indirect Inference Testing Results: Sticky Information Expectation Model

Table 4-6 displays the estimation results of the model with sticky information (SI model). Overall, most estimates through SA estimation are higher than those obtained from Bayesian estimation, excepting that the estimate of interest rate smoothed parameter is 0.7672 which is a little bit lower than that obtained through Bayesian estimation. The reaction parameter of output gap in monetary policy is estimated to be around 13%, which is lower than that in Bayesian estimates as well but being not quite far from its initial presumptive value. However, some SA estimates are higher than the Bayesian estimates, particularly the AR coefficient of monetary policy which is two times higher than that obtained through Bayesian estimation.

Furthermore, the test statistic indicates a Wald percentile of 53.10, so the SI model cannot be rejected at the 5% significant level, meaning that Wald statistic is well included in non-rejection region of the bootstrap distribution. Additionally, many SA estimates are somehow different from the estimates achieved by Bayesian estimation. For instance, the elasticity of intertemporal substitution is seven times higher than the Bayesian estimated value 0.1092. As well as the SA estimated share of updating firms whose estimate is 0.4504, it is about 1.5 times larger than that (i.e., 0.3084) obtained through Bayesian estimates but closer to the counterpart (i.e., 0.657) in empirical studies (Reis, 2009). Besides, the share of updating consumers is estimated 2 times larger than that obtained though Bayesian approach.

TABLE 4-6 ESTIMATES OF SI MODEL (J=4)

Parameters	Starting Calibration	Bayesian Estimates	SA Estimates
	1	0.1092	0.9050
	0.6	0.6340	0.5542
	0.75	0.9002	0.7672
	1.5	1.3735	1.6266
	0.12	0.1848	0.1299
	0.89	0.8139	0.8842
	0.79	0.6490	0.6421
	0.64	0.2986	0.7351
	0.5	0.3084	0.4504
	0.5	0.2362	0.5138
Full Wald %	99.4	54.00	53.10
TM (normalize t-statistic)	2.7338	-0.2072	0.1092

4.3.2.3 Estimation-Based Indirect Inference Testing Results: Imperfect Information Data Revision Expectation Model

In Table 4-7, in general, although none of the three cases concerning calibration-based model test, Bayesian-estimated-based model test and SA-estimated-based model test, can pass the test, the model with Bayesian estimates gives the worst result which can be inspected through TM distance (normalized t-statistics). The most significant difference between the SA-estimated-based model test and the Bayesian-estimated-based model test is that the estimated value of coefficient of the former test, being closer to its initial presumptive value, is ten times larger than the value obtained through the latter test.

TABLE 4-7 ESTIMATES OF IF DATA REVISION MODEL

Parameters	Starting Calibration	Bayesian Estimates	SA Estimates
	1	0.0899	0.8639
	0.6	0.7389	0.5623
	0.75	0.8801	0.6495
	1.5	1.0884	1.3342
	0.12	0.1962	0.1131
	0.5	1.8500	0.4404
	0.5	1.1198	0.4683
	0.67	0.6186	0.6292
	0.56	0.3657	0.5083
	0.30	0.2235	0.2718
	0.42	0.7252	0.3443
	0.61	0.8535	0.5099
Full Wald %	100	100	100
TM (normalize t-statistic)	28.5625	94.6459	20.3812

4.3.2.4 Comparison through Estimation-based Test

4.3.2.4.1 TM Distance Comparison

Overall, due to the norm of 1.645 as a threshold of judging the succeed of pass, only the models whose absolute values of TM Distance are below 1.645 can be qualified being 'good enough' models. According to Table 4-8, the SI Model can pass Bayesian-estimated-based test and SA-estimated-based test with a fail in calibration-based test, while the FIRE Model and the IF Model can pass 1 and 0 test respectively. We can drop a conclusion that the SI Model is superior to the other ones in terms of overall model fit.

The assessment of model is more precise by using the SA estimates from the point of view of actual data. Since the AR coefficients in SA estimates are estimated basing on the structural errors which use the actual observed data and parameters estimated in the model. The SA Estimation, in which the initial presumptive parameters are replaced by the optimal ones for re-test leading to higher passing possibility for the competing models, does not allow the IF model to pass. In general, the results of SA-estimation-based testing are better than the results of initial calibration-based testing as expected. This improvement can be attributed to the application of SA estimation approach what explores all the potential parameters over wild space to discover the best fit.

TABLE 4-8 COMPARISON TM DISTANCE (NORMALIZED T-STATISTICS)

Model	Starting Calibration	Bayesian Estimates	SA Estimates
FIRE Model	4.1538	26.0498	0.6587
SI (j=4) Model	2.7338	-0.2072	0.1092
IF Model	28.5625	94.6459	20.3812

4.3.2.4.2 Estimated Impulse Response Functions (IRFs)

In this section, the estimated impulse response functions have been used as the main tools to explore each competing model's behavior under all three shocks (i.e., demand shock, cost-push shock and monetary policy shock).

IRFs of Monetary Policy Shock

Figure 4-2 displays the estimated impulse response of the three main variables (i.e., output gap, inflation, and interest rate) to the monetary policy shock of three competing models respectively. In general, under the estimated monetary policy reaction function, the responses of the same variable under different models are quantitively similar. To be specific, nominal interest rate increases, but output gap and inflation decrease with respect to the three competing models. As shown in Figure 4-2, throughout the impacts of monetary policy shock on inflation and output gap, the hump-shaped response only appears under the SI model. Regarding to the period of convergence, the convergences of three main variables under FIRE model (the baseline model) and SI model are around 18 periods, but under IF model (i.e., the model with imperfect information data revision) they converge faster. Surprisingly, under the model with imperfect information data revision, the impact of monetary policy shock not only fails to generate the hump-shape response on inflation and output gap, but also weakens the delay response on interest rate.

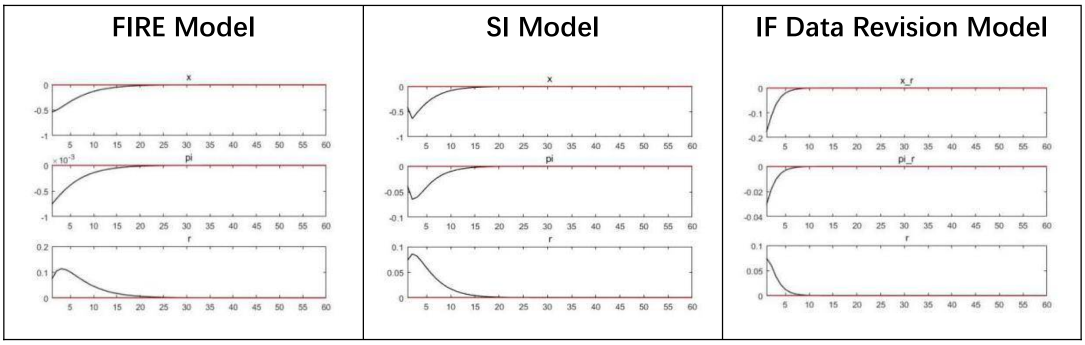


Figure 4-2 Estimated Impulse Response Function of One Unit Positive Policy Shock to Main Variables (x=output gap, pi=inflation, r=nominal interest rate)

IRFs of Demand Shock

Figure 4-3 presents the estimated impulse response functions of the three main variables to demand shock regarding the three rivals. Overall, the positive demand shock has a positive effect on three main variables. Besides, the effect last for a long time (i.e., around 20 periods more) under FIRE model and SI model. However, the effects on three main variables are relatively short with respect to the IF model. Furthermore, the demand shock has a persistent impact on inflation and output gap under SI model, which does not appear under the other two competing models.

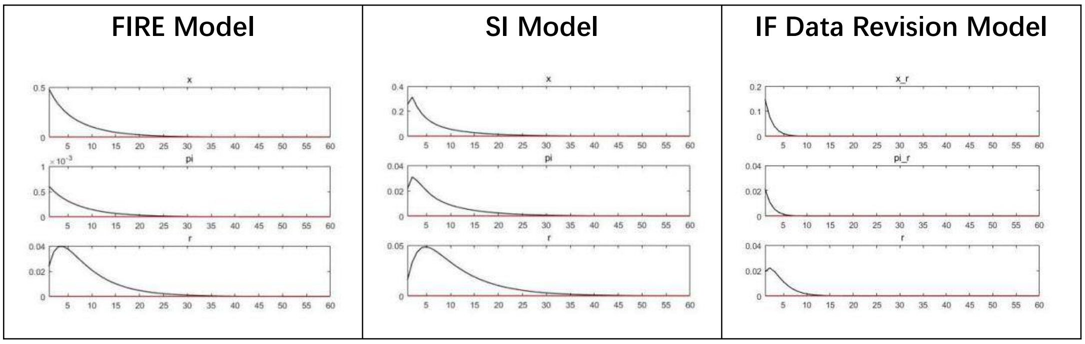


Figure 4-3 Estimated Impulse Response Function of One Unit Positive Demand Shock to Main Variable (x=output gap, pi=inflation, r=nominal interest rate)

IRFs of Cost-Push Shock

Figure 4-4 shows the behavior of three main variables in response to the positive cost-push shock with respect to three competitors. In general, all three competing models generate similar dynamics quantitatively. In detailed, both the inflation and interest rate are affected positively by the positive cost-push shock which delivers a negative effect on output gap. Additionally, the cost-push shock has the largest effect at initial point under FIRE model on three main variables. Meanwhile, it has a moderate effect at initial point under SI model and a minimal effect under IF model in terms of periods return to steady state.

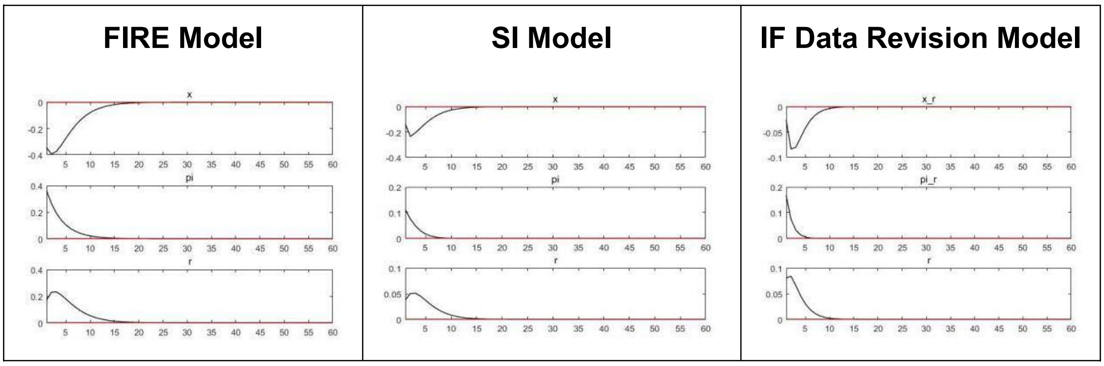


Figure 4-4 Estimated Impulse Response Function of One Unit Positive Cost-Push Shock to Main Variable (x=output gap, pi=inflation, r=nominal interest rate)

To sum up, the estimated IRFs are not very different from those obtained by Bayesian estimation. The SI model has strong abilities of generating more persistence and reproducing delay responses to monetary policy shock. However, the IF model still cannot achieve this goal.

4.3.2.5 Robustness check

1) Higher-Order Auxiliary Models

In this section, as the same in Section 4.3.1.3, we need to check that whether the rank among the three competing models in terms of higher-order auxiliary models is robust with the optimal set of parameters. We chose a VAR (1) as the auxiliary model in which the selected descriptors are equivalent to the estimates of its coefficients matrix and data variance incorporated in the indirect inference estimation procedure. As stated earlier, there are two factors, which are the model required to fit and the its extent of fit, that decide which option we should choose as the auxiliary model from a higher-order VAR model and other multiple types of time series models. When the higher-order auxiliary model VAR (2) and VAR (3) have been applied, the results show that although none can pass the test, the models' performances still can be compared. According to Table 4-9, the leading position of SI model in terms of overall dynamic properties over the competitors has not been switched when we choose higher-order VAR (i.e., VAR (2) or VAR (3)) instead of VAR (1) as auxiliary model.

Overall, the results of TM statistics in Table 4-9 indicates that raising VAR's order would make the acceptance of all the three estimated models weaker due to the greater burden placed on them. Comparing the results of TM statistics from Table 4-9 and Table 4-8, we can draw three conclusions. Firstly, it is obvious that when we use lower order VAR (i.e., VAR (1)) as the auxiliary model, all three competing models are less rejected. Secondly, the SI model is always less rejected than the competitors, which indicates that the SI model is preferred from the angle of model's overall performance regardless of the auxiliary VAR models' order. Thirdly, the ranking of three competing models is identical to the previous regardless of different choices of auxiliary models (i.e., VAR (2) or VAR (3)) through SA estimation among three rivals. So, VAR (1) can be an accepted auxiliary model to mimic the theoretical models.

TABLE 4-9 MODEL PERFORMANCE UNDER DIFFERENT AUXILIARY MODELS

Competing model	FIRE	SI (j=4)	IF	FIRE	SI (j=4)	IF
DATA SAMPLE: WITHOUT SURVEY DATA						
Auxiliary model	VAR (2)			VAR (3)		
TM Distance (Full Wald %)	8.1734 (100)	7.4455 (100)	32.1638 (100)	11.7022 (100)	9.1573 (100)	47.4983 (100)

2) *Different Truncation Point j of Sticky Information Model*

In this section, as the same as in Section 3.7, we need to check the robustness of different truncation point j in SI model but through the indirect inference approach. We have selected alternatives =6 and 8 to imply them into robust check procedure. According to Table 4-10, we receive the same suggestion as the one provided by Bayesian estimation approach that incorporating more lagged information into SI model has merely influence on its model performance after checking the TM distance (normalized t-statistics). Furthermore, the ranking among three rivals is identical as the previous ranking no matter which value of truncation point j (i.e., =6 and 8) in SI model is applied.

TABLE 4-10 SENSITIVITY CHECK BY USING MINIMIZING COEFFICIENT VALUES FOR SI MODEL

Model	TM by using SA Estimated Parameter
FIRE model	0.6587
SI model (j=4)	0.1092
SI model (j=6)	-0.2796
SI model (j=8)	-0.3518
IF model	20.3812

3) *Using alternative data resource: survey of professional forecaster data of output gap and inflation*

The estimation result by using Survey of Professional Forecaster Data (survey data) is presented in Table 4-12. The results obtained through Bayesian estimation approach show that the performance of IF model is far more superior to its rivals'. However, through indirect inference estimation, it shows that the full ability of IF model is far inferior to its competitors'. When each model is estimated by using survey data instead of real-time data, none of them can pass the test. In addition, it becomes more difficult to tell which one from FIRE expectation model and SI expectation model can give the better replication of the full dynamics of the actual observables (i.e., survey data) better. However, SI model performs at least no worse than the baseline when SPF data has been used.

TABLE 4-11 STARTING CALIBRATION PARAMETER VALUE OF AR COEFFICIENTS[51]

FIRE Model		
	AR coefficient of demand shock	0.94
	AR coefficient of cost-push shock	0.75
	AR coefficient of policy shock	0.56
SI Expectation Model		
	AR coefficient of demand shock	0.93
	AR coefficient of cost-push shock	0.74
	AR coefficient of policy shock	0.56
IF Expectation Model		
	AR coefficient of demand shock	0.70
	AR coefficient of cost-push shock	0.54
	AR coefficient of policy shock	0.29
	AR term of shock in final revision process of x	0.39
	AR term of shock in final revision process of	0.59

TABLE 4-12 COMPARISON TM BY USING MINIMIZING COEFFICIENT VALUES (WITH SURVEY DATA)

Model	SA Estimation Parameter
FIRE Model	5.6900
SI (j=4) Model	5.2699
IF Model	12.4718

[51] The AR coefficients of the structural errors implied by the models, all of them are sample estimated base on survey of professional forecaster data.

4.5 Conclusion

In this chapter, we use indirect inference as a testing method (i.e., calibration-based testing method) at starting stage and take the same approach as an estimation method (i.e., estimation-based testing method) in the next stage. We aim to contradistinguish the performance of the simulated-data-based estimated auxiliary model, with the performance of actual-data-based estimated auxiliary model through indirect inference test method.

We implement indirect inference methodology to test the three competing models regarding its dynamic performance for US economic real-time quarterly data from 1969 to 2015 (also use the other type of sample data, i.e. survey of the professional forecaster data, over the same period in robustness check). We compared three versions of model and found that none of them can fit the actual data through the initial calibrated-based test. Surprisingly, the imperfect information has the worst performance among the three models, which is contradicted to the results obtained by Bayesian estimation approach. However, the calibration-based testing results obtained by Indirect Inference approach shows that the model with sticky information expectation performs best among three competitors.

In the second stage, Indirect inference has been applied as estimation approach to both types of expectation models: with and without inattentiveness which were investigated in chapter 3. The comparisons of each competing models through Bayesian-estimated-based test and SA-estimated-based (Indirect Inference) test have been conducted respectively. The results indicate that the performance of each competing model with SA (indirect inference) estimates (i.e., best fitting parameters) has been improved, when compared with the results of the calibration-based test from the first stage.

Four achievements can be reflected through the results of indirect inference estimation. Firstly, regardless of two different estimation methods (i.e., Bayesian estimation and Indirect Inference estimation) by using the real-time data, the model with sticky information expectation is all the way preferred among the three competitors. Secondly, when we tried to find a robust superior model in terms of dynamic performance by changing the conditions, such as auxiliary model, truncation point in SI model, and type of data resource, we found that the model with sticky information expectation still the best choice to fit the US economy, Thirdly, the impacts of the structural shocks on US economy have been analyzed by the estimated impulse response functions. In general, these impacts are not significant different from the previous studies quantitatively, as well as those estimated through Bayesian estimation in chapter 3. For instance, a positive demand shock result in a raise in output gap, inflation, and interest rate. A positive monetary policy shock impact interest rate positively but creates a decrease in both output gap and inflation. Fourthly, unexpectedly, the model features imperfect information data revision fails to pass the test and gain the worst performance, which is contradict to not only the result obtained through Bayesian approach but also the suggestions from previous studies.

Overall, although Bayesian estimation approach is an effective practical tool to inspect model's performance by taking prior information about the macro economy into consideration, the prior is restricted while being applied because prior distribution need to be determined before entering estimation process. Besides, the model's performance obtained by Bayesian estimation are showed in a relative way that impossible to evaluate their absolute abilities. Thus, the method of indirect inference used in this chapter is an advanced tool to re-estimate each competing model in an 'unrestricted' way by exploring all the potential sets of parameters which can be accepted by models. In addition, the independent VAR has been used as an auxiliary model which offers a way to examine each model in an absolute sense. Besides, the optimal set of parameters can be discovered through SA mechanism for each competing model, to mitigate the unfairness in model comparisons.

While we were replacing the real-time data with survey data to apply them in estimation procedure, we found that the performances of models were increased excepting the cases of FIRE model and SI model through Indirect Inference. This contraction indicates that the survey data may contain useful information to improve the imperfect information data revision model's performance.

Appendix A to Chapter 4

TABLE 4A-1 ADF TEST RESULTS OF THE REVISED VARIABLES

Variables	Option	Critical value	t-statistics	Inference
	None	-1.942013	-5.411552 (-6.62896)	stationary (stationary)
	None	-1.942013	-3.242983 (-3.280844)	stationary (stationary)
Note: the number in the bracket is tested by using SPF revised data; outside the bracket is tested by using real-time revised data.				

TABLE 4A-2 ADF TEST RESULTS OF THE SURVEY OF PROFESSIONAL FORECASTER VARIABLES

Variables	Option	Critical value	t-statistics	Inference
	None	-1.942013	-7.191524	stationary
	None	-1.942013	-5.285229	stationary
	None	-1.942013	-5.145850	stationary
	None	-1.942013	-13.82232	stationary
Note: Here, and are the SPF data which denote that use survey conducted at time t and release in next period; and similar for and .				

TABLE 4A-3 ADF TEST RESULTS OF REAL TIME VARIABLES

Variables	Option	Critical value	t-statistics	Inference
	None	-1.942013	-4.19852	stationary
	None	-1.942013	-7.128462	stationary
	None	-1.942013	-2.332022	stationary
	None	-1.942013	-2.344756	stationary
Note: Here, and are the real-time data t released after one period; and are the real-time data t release after three periods.				

Appendix B to Chapter 4

TABLE 4B-1 MINIMIZING COEFFICIENT VALUES FOR FIRE MODEL (WITH SURVEY DATA)

Parameters	SA Estimates
	0.0275
	0.6286
	0.7476
	1.7401
	0.0749
	0.7759
	0.6537
	0.2772
Full Wald %	100
TM (normalize t-statistic)	5.6900

TABLE 4B-2 MINIMIZING COEFFICIENT VALUES FOR SI (J=4) MODEL (WITH SURVEY DATA)

Parameters	SA Estimates
	0.9878
	0.5713
	0.7180
	1.5641
	0.1238
	0.7696
	0.6570
	0.5505
	0.5179
	0.4849
Full Wald %	100
TM (normalize t-statistic)	5.2699

TABLE 4B-3 MINIMIZING COEFFICIENT VALUES FOR IF MODEL (WITH SURVEY DATA)

Parameters	SA Estimates
	0.4386
	0.8435
	0.5477
	1.4304
	0.1292
	0.4655
	0.4399
	0.6968
	0.5394
	0.2788
	0.3977
	0.5777
Full Wald %	100
TM (normalize t-statistic)	12.4718

Chapter 5
General Conclusion and Further Research Direction

5.1 Some Valuable Summarizes of The Thesis

Through comparing the models with inattentive expectation, we have a flexible way to explain which inattentive feature can give a better explanation of the US economy. To be specific, basing on the most commonly used stylized New-Keynesian model, we successfully incorporate the inattentive expectation assumption into the model out of the existence of the cost for acquiring and processing the updated information, or the data revision issues. In the sticky information assumption, the agents are slowly incorporating information about macroeconomic conditions (i.e., output, inflation, and interest rate). For another, in the assumption of data revision, economic agents cannot observe the true state because of noises. These noises are originated from people's imperfect knowledge about the real economy.

This research arises from the two inattentive assumptions above which are suggested from the two proposals in previous commonly discussed literature - one is sticky information expectation (Mankiw and Reis, 2002, 2007); the other one is imperfect information data revision expectation (Casares and Vazquez, 2016; Arouba, 2008). These studies all share the same goal of remedying deficiencies in the classical full-information expectation type models.

The deviation from full-information rationality after incorporating inattentive feature should be significant in solving issues of macroeconomics (Akerlof, 2002; Sargent, 1993). For example, after incorporating inattentive expectation, they find that many problems arising from the New-Keynesian model under full-information rationality assumption can be solved. Firstly, it can solve the problem of New-Keynesian full-information Phillips curve which leads nonsensically counterfactual forecasts about the impacts of monetary policy due to lack of any source of inflation inertia. Secondly, the counterfactual evidence regarding disinflations resulting in booms rather than recessions (Ball, 1994) can be removed which is argued by Mankiw and Reis (2002).

Thirdly, it removes the inability of full-information New-Keynesian type model that offer the explanation to the question why monetary policy shock has a delayed and gradual impact on inflation (Mankiw and Reis, 2002). Thus, such inattentive behaviour assumption considering what role people act in terms of behavioural economics is 'satisficer' rather than full-information rational maximiser (Simon,1989).

However, these approaches incorporating inattentive features do have their weaknesses. They are not successful in explaining why people not apply diffusely obtainable information about real economy into their economic decision making. However, people may easily find out what the information, such as interest rate, published by central bank, but it is hard to interpret the meanings behind the numbers for people lacking professional knowledge. As a result, the real problem is not get access to information but dealing with it. Unluckily, economics does not hole the instruments to model imperfect information dealing process. The methods proposed by Woodford (2003), Ball (2000) and Mankiw and Reis (2002, 2007) are none of the hope that a model of imperfect information procurement may take as a rough replacement. Despite the weaknesses of incorporating inattentive ingredients, its characteristic of explaining inflation inertia leads the model more complying with the situation of real world.

The alternative inattentive expectation models are applied in this thesis to compare with the baseline model. The selects are two-specific reduced-form three-equation DSGE model with inattentive feature. The sticky-information model as the first select which is based on the idea that while people forming their expectation, they are restricted by the cost of processing and acquiring the current information from using the latest information (Mankiw and Reis, 2002, 2007). The imperfect information model as the second select can reduce noise through data revision process (Casares and Vazquez, 2016; Arouba, 2008). One of the most significant motivations of the data revision comes from that there is a remarkably deep output gap misperception during the great inflation of the 1970s. This misperception can be coming down to the mis-measurement of actual output. Such mis-measurement, which is present in almost macroeconomic series, is a quantitatively substantial source of misperceptions (Collard and Dellas, 2010).

.

Concerning the results through two estimation methods, there are some part coincident. Firstly, the model of sticky Information is detected to be the most favorable model in the light of fitting the real-time data behavior. Secondly, the model with sticky information is the only one can generate delay response, which is in line with the evidence observed in actual data. Thirdly, the imperfect information data revision model with the survey data has better performance than that with the real-time data. The gap of the model with different conditions indicates that the survey data contains extra information to help improve imperfect information data revision model's performance.

However, there are some conflicts between the two estimation methods. In detailed, through the Bayesian estimation approach by using survey data the model with imperfect information data revision wins the best position among three competing models, but such result is not robust under alternative estimation methodology (i.e., Indirect Inference). The conflicts may be stemmed from the following reasons. Firstly, due to the unobserved potential output, the traditional measures of the output gap are probably burdened with error. The mismeasurement of the true output gap could influence the ability of each selected competing model (Lown and Rich, 1997). Secondly, different estimation methodologies may potentially lead to different conclusions. However, it is obvious that there is no absolute optimal way to choose a macro econometric method to estimate and evaluate models. Different estimation methodologies have their strengths and weaknesses. For instance, the Bayesian estimation approach is superior on the aspect of incorporating priors linking to the previous studies, but it is deficient for the same aspect because these priors have been put 'restrictions' before estimation. Besides, how to set prior distribution before estimation is still a disputable issue. Moreover, Bayesian estimation only offers a way to obtain model's relative performance by comparison, which cannot examine a model's absolute ability individually. Thus, we decide to use an 'unrestricted' estimation and evaluation method, indirect inference, to estimate different models as a robust check approach. It may be doubted that there is no model of any sort is qualified enough to simulate the 'real world' for its complexity. However, as asserted by Friedman (1953): 'Complete realism is clearly unattainable, and the question whether a theory is realistic enough can be settled only by seeing whether it yields predictions that are good enough for the purpose in hand or are better than predictions from alternative theories'. Thus, a qualified model should not be assessed by 'literal truth', but by 'if it is true'. He gives the perfect competition as an example to demonstrate his idea. Although the perfect competition never actually exists, it predicts the industries' highly competitive behaviour. Thus, even there is no model perfect match the reality, we still test its own ability to what extent can be used to explain the real world. That is why the indirect inference is chosen as the robust evaluating method in this thesis.

5.2 Further Research

In this thesis, we estimate and test New-Keynesian reduced-form type models with respect to two different expectation assumptions--with and without inattentiveness--by using US macro-economic data (survey of professional forecaster data have been adopted in robust check section). In choosing inattentive models for comparing, many options are left by us, but they can be developed in future work in the following ways.

Firstly, we only consider inattentive expectation with small-closed economy. Future work could be conduct through empirically evaluating small-open economy by incorporating exchange rate, import and export to develop more complicated models for comparison. Secondly, we can investigate mix-inattentive model (Dräger, 2016) to compare with the single-inattentive model. This process could also be applied into both close and open economies. Thirdly, the robust check in this thesis regarding to different specification of monetary policy shows that although the rank among three competing models do not switch, with respect to different monetary policy specifications, each model’s performance changes significantly. Thus, further research can take the inattentive expectation as the base structure model but with different monetary policy to examine whether the monetary authority does a good job over recent decades, which can also be carried out through both Bayesian and indirect inference approach.

Supporting Annex

Full-Information Rationality Assumption Model Micro-foundations and Derivations (Baseline Model)

The main derivation is following the common deriving procedure in New Keynesian literature (e.g., Walsh, 2003; Menz and Vogel, 2009).

Full-Information Rational Expectation Model: IS Curve

Representative households are assumed to consume a composite of differentiated foods by monopolistically competitive firms that make up of a continuum of measure. The composite consumption that enters that utility function in each period is:

(A.1)

Where is the price elasticity of demand for good i. The cost minimization process of representative households implies that demand for good i is,

(A.2)

Where is the price of good i and is the aggregate price in period t. Each household maximizes the following discounted sum of future expected utility functions

(A.3)

Where stands for the time discount factor, while and denote the elasticities of inter-temporal substitution and the inverse of the elasticity of labour supply respectively. Subject to the period budget constraint

(A.4)

Each household derives utility from consumption and disutility from hours of labor supplied . In the budget constraint, stands for nominal bond holdings, denotes the aggregate price level, the real wage, the nominal interest rate, is the real term of dividend distributions, and is the real term of net transfer or taxes. The utility maximization problem can be described using the Lagrangean function as follows:

(A.5)

First order conditions imply,

(A.6)

(A.7)

(A.8)

And then we can get,

(A.9)

(A.10)

After log-linearization equation (A.9) around a zero-inflation steady state, where , and denote the percentage deviation from steady state.

(A.11)

And log-linearizing the resource constraint is [52],

(A.12)

Then the output gap is defined as the difference between actual output and potential output, where the potential output is the output under flexible price. The potential output can be solved approximately use the log difference of actual output from its HP trend.

(A.13)

Furthermore, here use the output gap rewrite the above log-linearizing Euler equation,

(A.14)

Where demand shock is an exogenous shock driven by exogeneous productivity shocks.

Full-Information Rationality Assumption Model: Phillips Curve

As explained in this small-closed economy the representative agent's households' own firms. Under monopolistically competitive environment each firm has production function. And the production function, in line with the standard NK model, I assume a Cobb-Douglas production with constant return to scale

(A.15)

[52] Follow by Walsh (2003), we also assume , then

Where i denote the firm; [53]is the technology. Under Calvo (1983) contract, each frim re-optimizes its price in every period with probability (1-) and keep its price fixed to the previously set price with probability , have to keep these remain due to menu cost. However, for simplicity, the nominal wage in the labour market are presumed to be fully flexible. And then where we have used the expressions for the product's demand curve,

(A.16)

So, in each period firms producing differentiated goods but processing identical price strategy would set individual prices , subject to the production constraint , the Calvo contract resetting probability is 1- and the demand curve , to maximize the discounted real profits. Then here we let denotes the real marginal cost to each firms' production, and solve the firms' cost minimization problem, we can solve,

(A.17)

(A.18)

Using the Lagrange

(A.19)

Solve the first order condition we get the firms' real marginal costs

(A.20)

[53] with , where is the iid productivity shock.

Such that each firm maximizes the expected discounted sum of future profits to choose an individual

(A.21)

Where —— is the discount factor, indicating the ratio of marginal utilities of consumption between periods. Then using the demand curve, we can rewrite the firm's maximization problem,

(A.22)

Then the first order condition of firms' maximized equation with respect to individual price implies

(A.23)

Log-linearization of the firm's maximized problem's first order condition, around zero inflation steady state yields the optimal reset price for each firm as follows:

(A.24)

The aggregate price level in each period given the Calvo contract can be written as the weighted average of this up-to-date reset prices and the unchanged, with the weights being the reset probability, and its opposite, respectively, and is this process each individual frim have the same price strategy,

(A.25)

Then log-linearized above equation we can solve

(A.26)

(A.27)

Use equation here we can get

$$—————$$ (A.28)

And since the log linearized of real marginal cost is,

(A.29)

Combine with the log-linearized of ——— — (which have been solved from first order condition from household side), then we get

$$-$$ (A.30)

Then we can have also solved the real marginal cost as following,

$$- \quad =$$ (A.31)[54]

Then we can get the new Keynesian Phillips Curve is,

$$—————$$ (A.32)

[54] The interpretation of is follow by Woodford (2001) as the strategic complementarity between different pricing decisions of different suppliers. Woodford suggest is an empirically plausible value for the US.

Follow by many authors simple adds an additive cost-push-shock after having derived the Philips curve in the standard way.

$$\underline{\qquad\qquad} \quad + \qquad (A.33)$$

Government and Monetary Policy (Taylor rule)

Finally, equation (A.34) is the interest rate smoothing rule with a lagged interest rate that has been added into the classic form that is developed by Taylor (1993) to obtain smoothing behaviour.

$$(A.34)$$

Where is the degree of partially adjustment, is the monetary policy shock. All disturbances , ,and are AR(1) processes with AR coefficients , , and ,

$$(A.35)$$

$$(A.36)$$

$$(A.37)$$

Reference

Akerlof, George A. (2002). "Behavioral Macroeconomics and Macroeconomic Behavior." American Economic Review, 92 (3), 411-433

An, S., & Schorfheide, F. (2007). Bayesian analysis of DSGE models. Econometric reviews, 26(2-4), 113-172.

Aruoba, S. B. (2008). Data revisions are not well behaved. Journal of Money, Credit and Banking, 40(2-3), 319-340.

Aruoba, B., & Schorfheide, F. (2011). Sticky prices versus monetary frictions: An estimation of policy trade-offs. American Economic Journal: Macroeconomics, 3(1), 60-90.

Ball, L. (1994). Credible disinflation with staggered price-setting. The American Economic Review, 84(1), 282-289.

Blinder, A., Canetti, E. R., Lebow, D. E., & Rudd, J. B. (1998). Asking about prices: a new approach to understanding price stickiness. Russell Sage Foundation.

Ball, L. (2000). Near-rationality and inflation in two monetary regimes (No. w7988). National Bureau of Economic Research.

Begg, D. K., & Imperato, I. (2001). The rationality of information gathering: monopoly. The Manchester School, 69(3), 237-252.

Ball, L., Mankiw, N.,G., and R. Reis, (2005), "Monetary Policy for Inattentive Economies," Journal of Monetary Economics 52(4), 703-725.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. Journal of Monetary Economics, 12(3), 383-398.

Clarida, R., Gali, J., & Gertler, M. (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. The Quarterly Journal of Economics, 115(1), 147-180.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. the Quarterly Journal of Economics, 118(1), 269-298.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy, 113(1), 1-45.

Coibion, Olivier, (2006). “Inflation Inertia in Sticky-Information Models,” Contributions to Macroeconomics 6(1), Article 1.

Corradi, Valentina, Andres Fernandez, and Norman R. Swanson (2009). “Information in the Revision Process of Real-Time Datasets.” Journal of Business and Economic Statistics 27, 455–467.

Collard, F., H. Dellas and F. Smets (2009), “Imperfect information and the business cycle”, Journal of Monetary Economics, 56 (1), 38-56

Canova, F., & Sala, L. (2009). Back to square one: Identification issues in DSGE models. Journal of Monetary Economics, 56(4), 431-449.

Collard, F., & Dellas, H. (2010). Monetary misperceptions, output, and inflation dynamics. Journal of Money, Credit and Banking, 42(2-3), 483-502.

Caballero, R. (2010), 'Macroeconomics after the Crisis: Time to Deal with the Pretense of Knowledge Syndrome', Journal of Economic Perspectives, 24(4), 85-102.

Coibion, O. (2010). Testing the sticky information Phillips curve. The Review of Economics and Statistics, 92(1), 87-101.

Croushore, D. (2011). Frontiers of real-time data analysis. Journal of economic literature, 49(1), 72-100.

Coibion, O. and Y. Gorodnichenko (2012). What Can Survey Forecasts Tell Us About Information Rigidities? Journal of Political Economy 120 (1), 116–159.

Coibion, O., & Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. The American Economic Review, 105(8), 2644-2678.

Casares, M., & Vázquez, J. (2016). Data Revisions in the Estimation of DSGE models. Macroeconomic Dynamics, 20(7), 1683-1716.

Diebold, F. X., & Rudebusch, G. D. (1991). Forecasting output with the composite leading index: A real-time analysis. Journal of the American Statistical Association, 86(415), 603-610.

Dynan, Karen (2000) “Habit Formation in Consumer Preferences: Evidence from Panel Data,” American Economic Review, 90, 391-406

Dhyne, E., L. J. Alvarez, H. L. Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, ´ P. Lunnemann, F. Rumler, and J. Vilmunen (2006): “Price Setting in the Euro Area and the United States: Some Facts From Individual Consumer Price Data,” Journal of Economic Perspectives, 20(2), 171–192.

Del Negro, M., & Eusepi, S. (2011). Fitting observed inflation expectations. Journal of Economic Dynamics and control, 35(12), 2105-2131.

Dräger, L., (2011). Endogenous persistence with recursive inattentiveness. Working Paper 285. KOF Swiss Economic Institute.
-(2016), Recursive inattentiveness with heterogeneous expectations. Macroeconomic Dynamics, 20(4), 1073-1100.

Dräger. L, Michael J. Lamla b (2013), Updating inflation expectations: Evidence from micro-data L, Economics Letters 117, 807–810

Eichenbaum, M., & Fisher, J. D. (2004). Evaluating the Calvo model of sticky prices (No. w10617). National Bureau of Economic Research.

Joshy, E., & Golinelli, R. (2010). Households Forming Inflation Expectations: Active and Passive Absorption Rates. The BE Journal of Macroeconomics, 10(1), 1-32.

Friedman, M. (1953). Essays in positive economics. University of Chicago Press.

Friedman, B. M. (1979). Optimal expectations and the extreme information assumptions of ‘rational expectations’ macromodels. Journal of Monetary Economics, 5(1), 23-41.

Fuhrer, Jeff and George Moore (1995). “Inflation Persistence.” Quarterly Journal of Economics, 110, 127-160

Fernández-Villaverde, J., & Rubio-Ramı rez, J. F. (2004). Comparing dynamic equilibrium models to data: a Bayesian approach. Journal of Econometrics, 123(1), 153-187.

Fernández-Villaverde, J. (2010). The econometrics of DSGE models. SERIES: Journal of the Spanish Economic Association, 1(1), 3-49.

Fan, J., Minford, P., & Ou, Z. (2016). The role of fiscal policy in Britain's Great Inflation. Economic Modelling, 58, 203-218.

Fuhrer, J. (2017). Expectations as a source of macroeconomic persistence: Evidence from survey expectations in a dynamic macro model. Journal of Monetary Economics, 86, 22-35.

Gregory, A. and Smith, G. (1991). Calibration as testing: Inference in simulated macro models. Journal of Business and Economic Statistics, 9 (3), 293-303.

Gregory, A. and Smith, G. (1993). Calibration in macroeconomics, in: Maddala, G.(Ed.), Handbook of Statistics,11, Elsevier, St. Louis, Mo., 703-719

Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect inference. Journal of Applied Econometrics, 8, 85-118.

Gallant, A. R., & Tauchen, G. (1996). Which moments to match? Econometric Theory, 12(4), 657-681.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: inference, development, and communication. Econometric reviews, 18(1), 1-73.

Gali., J. and M. Gertler, (1999), "Inflation Dynamics: A Structural Econometric Analysis," Journal of Monetary Economics, 44, 195-222.

Galí, J. (2002). New perspectives on monetary policy, inflation, and the business cycle (No. w8767). National Bureau of Economic Research.

Gali J, López-Salido J D, Vallés J. (2003) Technology shocks and monetary policy: assessing the Fed's performance[J]. Journal of Monetary Economics, 50(4), 723-743.

Geweke, J., & Whiteman, C. (2006). Bayesian forecasting. Handbook of economic forecasting, 1, 3-80.

II, K., & Edward, S. (2010). A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment. Journal of Money, Credit and Banking, 42(8), 1543-1564.

Jeffreys, H. (1961). Theory of probability (3rd ed.). New York, NY: Oxford University Press.

Jones, B. D. (2002). Bounded rationality and public policy: Herbert A. Simon and the decisional foundation of collective choice. Policy Sciences, 35(3), 269-284.

Joshy Easaw and Roberto Golinelli (2014): "Inflation Expectations and the Two Forms of Inattentiveness," Cardiff Economic Working paper

Khan, H., & Zhu, Z. (2002). Estimates of the Sticky-Information Phillips Curve for the United States, Canada, and the United Kingdom (No. 02-19). Bank of Canada.
-(2006), Estimates of the sticky-information Phillips curve for the United States. Journal of Money, Credit, and Banking, 38(1), 195-207.

Keane, M., & Smith, A. A. (2003). Generalized indirect inference for discrete choice models. Yale University.

Keen, B. D. (2007). Sticky Price and Sticky Information Price-Setting Models: What is the difference?, Economic Inquiry, 45(4), 770-786.

Kiley, M. T. (2007). A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting. Journal of Money, Credit and Banking, 39(s1), 101-125.

Klenow, Peter J., and Jonathan L. Willis. (2007). “Sticky Information and Sticky Prices.” Journal of Monetary Economics, 54 (10), 79–99.

Korenok, O., & Swanson, N. R. (2007). How sticky is sticky enough? A distributional and impulse response analysis of new Keynesian DSGE models. Journal of Money, Credit and Banking, 39(6), 1481-1508.

Knotek, Edward S. (2010). “A Tale of Two Rigidities: Sticky Prices in a Sticky Information Environment.” J. Money, Credit and Banking 42 (12), 1543– 64.

Lown, Cara S. and Robert Rich. (1997). “Is there an inflation puzzle?” Federal Reserve Bank of New York Economics Policy Review, 3(4), 51–69.

Lubik, T., & Schorfheide, F. (2005). A Bayesian look at new open economy macroeconomics. NBER macroeconomics annual, 20, 313-366.

Le, V. P. M., Meenagh, D., Minford, P., & Wickens, M. (2010). Two orthogonal continents? testing a two-country DSGE model of the US and the EU using indirect inference. Open Economies Review, 21(1), 23-44.

Le, V. P. M., Meenagh, D., Minford, P., & Wickens, M. (2011). How much nominal rigidity is there in the US economy? Testing a New Keynesian DSGE Model using indirect inference. Journal of Economic Dynamics and Control, 35(12), 2078-2104.

Le, V. P. M., Meenagh, D., Minford, P., & Ou, Z. (2013). What causes banking crises? An empirical investigation for the world economy. Open Economies Review, 24(4), 581-611.

Levine, P., Pearlman, J., Perendia, G., & Yang, B. (2012). Endogenous persistence in an estimated DSGE model under imperfect information. The Economic Journal, 122(565), 1287-1312.

Liu, C., & Minford, P. (2014). Comparing behavioral and rational expectations for the US post-war economy. Economic Modelling, 43, 407-415.

Le, V. P. M., Meenagh, D., Minford, P., Wickens, M., & Xu, Y. (2016). Testing macro models by indirect inference: a survey for users. Open Economies Review, 27(1), 1-38.

McCallum, B. T. (1976). Rational expectations and the natural rate hypothesis: some consistent estimates. Econometrica: Journal of the Econometric Society, 44(1), 43-52.

Mankiw and Reis (2002). STICKY INFORMATION VERSUS STICKY PRICES: A PROPOSAL TO REPLACE THE NEW KEYNESIAN PHILLIPS CURVE. The Quarterly Journal of Economics, 117(4), 1295-1328.

Mankiw, N. Gregory, Ricardo Reis and Justin Wolfers (2003). "Disagreement about Inflation Expectations." NBER Macroeconomics Annual, 18, 209-248.

Mankiw, N. Gregory, and Ricardo Reis (2006). "Pervasive Stickiness." American Economic Review, 96(2), 164-169.

Mankiw, N. G., & Reis, R. (2007). Sticky information in general equilibrium. Journal of the European Economic Association, 5(2-3), 603-613.

Meenagh, D., Minford, P., & Wickens, M. (2008). Testing a DSGE model of the US using indirect inference. Cardiff University. Cardiff Economics Working Papers 2008/11.
-(2009), Testing a DSGE model of the EU using indirect inference. Open Economies Review, 20(4), 435-471.

Menz, J. O., & Vogel, L. (2009). A Detailed Derivation of the Sticky Price and Sticky Information New Keynesian DSGE Model (No. 2/2009). DEP (Socioeconomics) Discussion Papers, Macroeconomics and Finance Series.

Minford, P., Theodoridis, K., & Meenagh, D. (2009). Testing a model of the UK by the method of indirect inference. Open Economies Review, 20(2), 265-291.

Meyer-Gohde, A. (2010). Linear rational-expectations models with lagged expectations: A synthetic method. Journal of Economic Dynamics and Control, 34(5), 984-1002.

Milani, F. (2012). The modeling of expectations in empirical DSGE models: A survey. In DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments (pp. 3-38). Emerald Group Publishing Limited.

Milani, F., & Rajbhandari, A. (2012). Expectation formation and monetary DSGE models: Beyond the rational expectations paradigm. In DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments (pp.253-288). Emerald Group Publishing Limited.

Minford, P., & Ou, Z. (2013). Taylor Rule or optimal timeless policy? Reconsidering the Fed's behavior since 1982. Economic Modelling, 32, 113-123.

Minford, P., Ou, Z., & Wickens, M. (2015). Revisiting the Great Moderation: policy or luck?. Open Economies Review, 26(2), 197-223.

Mitchell, K., & Pearce, D. (2015). Direct Evidence on Sticky Information from the Revision Behavior of Professional Forecasters (No. 66172). University Library of Munich, Germany.
-(2017). Direct Evidence on Sticky Information from the Revision Behavior of Professional Forecasters. Southern Economic Journal, 84(2), 637-653.

Nakamura, E., & Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. The Quarterly Journal of Economics, 123(4), 1415-1464.

Ormeño, A., & Molnár, K. (2015). Using survey data of inflation expectations in the estimation of learning and rational expectations models. Journal of Money, Credit and Banking, 47(4), 673-699.

Peersman, G., & Smets, F. (2002). Are the effects of monetary policy in the euro area greater in recessions than in booms?. In Conference on Transmission Mechanism of Monetary Policy (pp. 28-48). Cambridge University Press.

Paustian, M., & Pytlarczyk, E. (2006). Sticky contracts or sticky information? Evidence from an estimated Euro area DSGE model. Manuscript, Bowling Green State University.

Paloviita, M (2007b) Estimating a small DSGE model under rational and measured expectations: some comparisons. Bank of Finland Research Discussion Papers, No. 14/2007.
- (2008). Comparing alternative Phillips curve specifications: European results with survey-based expectations. Applied Economics, 40(17), 2259-2270.

Rotemberg, J. J., & Woodford, M. (1996). Real-business-cycle models and the forecastable movements in output, hours, and consumption. The American Economic Review, 86 (1), 71-89.

Rabanal, P., & Rubio-Ramírez, J. F. (2005). Comparing New Keynesian models of the business cycle: A Bayesian approach. Journal of Monetary Economics, 52(6), 1151-1166.

Reis, R. (2006a). Inattentive consumers. Journal of monetary Economics, 53(8), 1761-1800.

Reis, R. (2006b). Inattentive producers. The Review of Economic Studies, 73(3), 793-821.

Reis, Ricardo (2009). “Optimal Monetary Policy Rules in an Estimated Sticky-Information Model.” American Economic Journal: Macroeconomics, 1(2), 1-28.

Simon, H. A. (1989). Models of Thought, Volume II. New Haven: Yale University Press.

Sargent, T. J. (1993). Bounded rationality in macroeconomics: The Arne Ryde memorial lectures. OUP Catalogue.

Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. Journal of Applied Econometrics, 15(6), 645-670.

Sims, C. A. (2003). Implications of rational inattention. Journal of monetary Economics, 50(3), 665-690.

Smets, F., & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. Journal of the European economic association, 1(5), 1123-1175.

Schorfheide, F. (2005). Learning and monetary policy shifts. Review of Economic dynamics, 8(2), 392-419.

Smets, F., & Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. The American Economic Review, 97(3), 586-606.

Stiglitz, J. E. (2011). Rethinking macroeconomics: What failed, and how to repair it. Journal of the European Economic Association, 9(4), 591-645.

Taylor, John B. (1993). "Discretion versus policy rules in practice." Carnegie-Rochester Conference Series on Public Policy 39, 195-214.

Trabandt, Mathias (2004). “Sticky Information vs. Sticky Prices: A Horse Race in a DSGE Framework.” Mimeo, Humboldt University Berlin.
- (2005). “Sticky Information vs. Sticky Prices: A Horse Race In A DSGE Framework.” Manuscript.
-(2007). Sticky information vs. sticky prices: A horse race in a DSGE framework.

Wickens, M. R. (1982). The efficient estimation of econometric models with rational expectations. The Review of Economic Studies, 49(1), 55-67.

Wilson, E. B., & Hilferty, M. M. (1931). The distribution of chi-square. Proceedings of the National Academy of Sciences, 17(12), 684-688.

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy (No. w8673). National Bureau of Economic Research.
-(2003), Imperfect Common Knowledge and the effects of monetary policy, Knowledge, Information, and Expectation In Modern Macroeconomics: In Honor of Edmund S. Phelps. (pp.5)

Woodford, M., (2003), Interest and Prices: Foundations of a Theory of Monetary Policy, Princeton University Press, Princeton, NJ.

Walsh, C, E (2003), Monetary Theory and Policy (3rd), Cambridge and London: MIT Press.

Vázquez, J., María-Dolores, R., & Londoño, J. M. (2010). Data Revisions and the Monetary Policy Rule: An analysis based on an extension of the basic New Keynesian model. The University of the Basque Country, Mimeo.
- (2012). The effect of data revisions on the basic New Keynesian model. International Review of Economics & Finance, 24, 235-249.

Vetlov, I., Hlédik, T., Jonsson, M., Kucsera, H., & Pisani, M. (2011). Potential output in DSGE models (No. 1351). ECB Working Paper.

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …………………………………………………………… (candidate)
Date ………………….…………….………

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ………(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed …………………………………………..…………… (candidate)
Date …………………………….……………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ………………………………………….………….…… (candidate)
Date …………………….……………………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …………………………………………………..….…….. (candidate)
Date …………………………………………

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed ……………………………………………..……… (candidate)
Date …………………………………….………

Abstract

Using US quarterly data (i.e., real-time data and survey data respectively) from 1969 to 2015 through two different estimation approaches (i.e., Bayesian estimation approach and indirect inference estimation approach) to investigate the empirical performance of the standard reduced-form New-Keynesian Dynamic Stochastic General Equilibrium (DSGE) model under the condition without (i.e., full-information rationality) and with inattentive features (i.e., sticky information and imperfect information data revision), we find some consistent results. Firstly, the model of sticky Information is detected to be the preferred model to fit the real-time data behavior. Secondly, the model with sticky information is the only one can generate delay response, which is matching the evidence observed in actual data and in line with most consequences from the previous studies. Thirdly, the imperfect information data revision model performs better when we substitute the real-time data with the survey data, through which we can deduce that the survey data contains extra information to help improve imperfect information data revision model's performance. Three main contributions are made in this thesis. The first contribution is the estimation and comparison of different types of inattentive DSGE model (sticky information versus imperfect information data revision) for US small-closed economy through Bayesian approach using the US quarterly data (i.e., real-time data and survey data) representing the main macroeconomic time series from 1969 to 2015. What the second contribution is that through comparing different inattentive New-Keynesian DSGE models basing on the full structure (relative to the single equations competition), we inspect which way of inattentive expectation is closer to the way that people form their expectation in real economy. Besides, the thesis adopts Indirect Inference approach as the robust check methodology, which delivers a new way to assess inattentive macroeconomic models, which is the third contribution.

Acknowledgments

Firstly, I want to say appreciate to my first and second supervisors, Dr. Joshy Easaw and Professor Patrick Minford for providing me with guidance, encouragement and much more. I have benefited from every time when I have a conversation and meeting with each of them. Without their support, and kind sharing of their knowledge, I would be not able to reach this stage and finish this Ph.D. journey.

Especially thanks to my first supervisor Dr.Joshy Easaw again for his generosity and patient what support me to overcome the heavy pressure and encourage me forward to finally finish my PhD thesis.

I would need to say thank you for all the members of the department staff at Cardiff Business School who supported me generously in many ways. I would also like to say many thanks to all my dear friends and course-mate in Cardiff, Dr.Zhirong Ou, Yue Gai, Xinran Zhao, Xue Dong, Yao Yao ,Xingchen Li, Ruimin Li and many other friends and colleagues in China, Taiwan and the UK whose names are not mentioned here due to space consideration.

Lastly, I have to say thanks to my dear family, my dear father Mr. Hong-Jen Chou, my mother Mrs Yue-Wei Chou, my two dear brothers Juin-Yu Chou and Juin-Cheng Chou, and my partner Mr Yirui Su. Without their continuous support and encouragement, it is impossible for me to have this courage to take and overcome this challenge.

Table of Content

List of Tables

List of Figures

General Introduction

Background

Expectation is important for economic agents in terms of making economic decisions, because they have to face various cases requiring such decisions in real life. For example, how to balance their consumptions and savings and what price to set, etc. Even now, concerning how economic agents forming their expectations, scholars do not have a unified model to explain the process. The New-Keynesian framework, which was characterized by full-information rationality assumption and the ‘extreme-sticky' prices, has been proposed to solve this issue in some resent works (e.g., Calvo, 1983). It revealed the essential factors to understand the dynamics of the real world, such as imperfect competition, price rigidities, but there were some arguments about its fail to explain some facts observed in actual data. For instance, as Jeff Fuhrer and George Moore (1995) argue, the monetary policy shock who has a delaying and gradual impact on inflation cannot be explained by the original New-Keynesian type model. Mankiw and Reis (2002) demonstrate that the postponed reaction to monetary shock on inflation cannot be produced without any information friction (inattentive feature) or the price indexed its counterfactual hypothesis.

Thus, two alternative expanded models based on the New-Keynesian framework emerged in recent decades to solve the problems which cannot be explained by the original New-Keynesian type model. Among them, the first is sticky information model which was defined by Mankiw and Reis (2002, 2007). According to their assumption of sticky information, there is a delay in the spreading procedure of the information of macro-economic conditions. The lagged spreading through the population may be caused from two aspects: the cost of re-optimized information and the cost of requiring information.

Due to its rapid growth in recent years, this approach was successfully applied to explain economic behavior. For instance, Reis (2006a, 2006b) asserts with the belief of inattentive expectation hypothesis that economic agents choose to use the updated information only when the expected benefit from the newly arrived information is higher than the cost of it. For example, we postulate assume that in period t there is a proportion of the people from the supply side will absorb the current-period information, and meanwhile the rest (proportion) of people keeps the opinion that they are preserving in period t-1 of period t+1's inflation rate. Thus, different from the full-information expectation, the current inflation depends on not only this period t's expectations of the future inflation but also the past expectation of the future inflation rate. The other one is the imperfect information (data revision) model (Woodford, 2001, 2003; Aruoba, 2008; Casares and Vázquez, 2016). The imperfect information agents refer to who constantly update their own information sets under the premise that never can they fully observe the real state. According to this, they form and renovate their beliefs regarding the underlying economic situations accompanied with the problem of signal extraction (Woodford, 2001, 2003). We take data revision as a solution of the signal extraction problem, which indicates that imperfect information agents through two ways of using data revision process to reduce noise and incorporating entire of the involved information to figure out the real situation of economy to reach the same goal which is forming their expectations. Thus, the current state not only depends on not only the final revised observations but also the initial released observations. The details of the data revision processes will be well stated in the Chapter 2. The way or definition of data revision process borrowing from Casares and Vázquez (2016) and Vázquez et al. (2010, 2012) will also be well clarified in chapter 2.

Inspired by the previous studies, three questions will be investigated in this thesis. First question is whether the inclusion of inattentive features can help the original

New-Keynesian DSGE model to replicate some important stylized[1] facts better. The second question is whether the inclusion can help to give a better overall performance. The third question is whether the different types of inattentive feature are distinctive in explaining the dynamics of the observed actual data. To discover the answers of the questions above, three rivals will be selected: the model with full-information rational expectation, the sticky information expectation model, and the imperfect information data revision expectation model. Each of them will be evaluated through two methodologies: Bayesian estimation method and Indirect Inference estimation method within this thesis.

Model Evaluation Methodology

The Bayesian Estimation Approach to Evaluate the New-Keynesian DSGE type models

Bayesian estimation has been implemented as a relatively 'strong' econometric estimation method by some recent studies (Geweke, 2006; An and Schorfheide, 2007). Where it is superior to the 'weak' econometric estimation methods should be the capability of embodying all the features and implications of the model in the estimating procedure, yet the 'weak' ones, for instance, the calibration methods are only can reproduce some chosen moments of the observed variables through simply assigned values to parameters.

Bayesian estimation method is catalogued under the group of 'strong' interpretation. To be more specific, through the comparison of the Bayesian estimation with the classical maximum likelihood estimation, it is easy to conclude that in the perspective

[1] The persistence property of output and inflation, and the delay effect of monetary policy shock on inflation. And such stylized facts are taken as serviceable norms what assistant to evaluate models. The observed hump-shaped response of inflation to monetary policy shock has been paid attention in these recent years. This is because the fact that this hump-shaped response is not only robust but also hard to be generated in a simple model. Most notably, the New Keynesian Phillips curve which is basing on the assumption that firms face expense to adjust price is not able to reproduce such a response without any information rigidities (Mankiw and Reis, 2002).

of the working step most of them are similar apart from the last few ones. To be specific, posterior density function is obtained by using Bayesian estimation approach given by the combination of the likelihood function and prior distributions of the model's parameters. Then this optimization of posterior can done concerning parameters of the model. What the most distinguished point between the two methods is that the classical maximum likelihood misses the steps of including the additional prior function to reweigh the likelihood function.

Two general reasons for using Bayesian estimation approach have been discussed frequently in recent studies (Schorfheide, 2000; An and Schorfheide, 2007). The use of prior information come from either the previous relevant studies or the reflection of researcher's subjective perception. So, this method directly builds a link between our study and previous studies. Besides, the Bayesian method can evaluate misspecified models according to the criteria of measurement which are the marginal likelihood and the Bayes' factor. The model's marginal likelihood which is connected to the density function of prediction directly which can be taken as an acceptable criterion to measure the level of overall model fit. The competing models selected by us will be estimated by using the real-time data of US 1969Q1-2015Q4 (survey of professional forecaster data will be used in robustness check).

Concerning the structural parameters and impulse response functions estimated through Bayesian estimation, it shows that the set of estimates for the structural parameters are plausible. For instance, the estimated price stickiness for US economy is considerable, which is in accordance with many previous studies (Smets and Wouters, 2007; Milani and Rajbhabdari, 2012). Besides, the impacts of the three main shocks on the US economy after analyzing are consistent with the existing studies quantitatively. For example, a positive monetary policy shock is along with a rise of nominal interest rate, a decline in output gap, and a decrease in inflation (Peersman and Smets, 2002). Moreover, the positive cost-push shock has positive impacts on inflation and nominal interest rate, but impact on output gap negatively.

Besides, a positive demand shock has a positive effect on the output gap.

From the perspective of the overall model fit estimated through Bayesian estimation using real-time data, the results show that the inclusion of inattentive features has significant effects on the mode’s ability in fitting macroeconomic time series. To be specific, the inclusion improves the model's ability to explain the real world, which is in line with the suggestions from most of the related literatures (Mankiw and Reis, 2002, 2007; Collard et al., 2009). Apart from that, we find that the model achieves its best fit under sticky information model through Bayesian estimation. By using diffuse prior distribution, different specifications of Taylor rule, and different periods of lag information in sticky information model (j=4, 6 and 8) in the robustness check, we draw a conclusion that none of them can change the ranking among the three rivals. Surprisingly, when we use survey of professional forecaster data instead of real-time data to evaluate the models’ performances, although the models with inattentive still be superior to the baseline model, the rank between the two inattentive models’ changes.

The Indirect Inference Approach to Evaluate the New-Keynesian DSGE type models

First Stage: Calibration-based Indirect Inference Test

Current studies attempt to formalize the test method to evaluate model’s performance in an absolute sense relying on Indirect Inference. The Indirect Inference as a testing method utilizes that the solution of the log-linearized DSGE model is able to be expressed by a restricted Vector-Autoregressive-Moving-Average (VARMA) model and can be closely expressed by a Vector Autoregressive (VAR) model. Indirect inference test can be understood as a process that through comparing the simulated-data-based unrestricted VAR estimates with the alternative actual-data-based unrestricted VAR estimates, after then we can confirm whether these two sets of the auxiliary models’ estimated parameters (i.e., VAR) are ‘close enough’ (i.e., each competing DSGE model is correctly specified)

While conducting Indirect Inference test, we employ Wald test on VAR estimates. In general, Indirect Inference testing procedure contains three general steps. The first step, to construct the errors implied by the actual data and one of the model of the previously estimation-based and calibration-based structural models. In the second step, the innovations of structural errors are bootstrapped to be employed to produce the pseudo data which are based on candidate model. After that, an auxiliary model (i.e., VAR) is fitted to each set of pseudo data and the sampling distribution of the coefficients of the auxiliary VAR model. In the final step, the Wald statistic is calculated to judge whether or not the functions of the parameters of the auxiliary VAR model estimated on the actual data lie within the confidence interval allusive by the sampling distribution. According to the results through Indirect Inference calibration-based test, none of the three competing models can pass the test. Comparing with the previous studies argue in the literature, the performance of the model with imperfect information data revision is much worse than that of the baseline model, which is contradict to the conclusion from the Bayesian approach.

Second Stage: Estimation-based Indirect Inference Test

The Indirect Inference has been long-standing applied (Gregory and Smith, 1991, 1993; Gourieroux and Monfort, 1993). As far as we concern, when Indirect Inference is applied for evaluating model, structural model's parameters are provided at the beginning. However, the nature of fixed calibrated parameters leads it an overly strong condition for testing models and contradistinguishing one model from another. Seeing the values of parameter of the candidate model could be estimated or calibrated within a permissible range throughout the theoretical structure of the model, it is probable for a rejected model with the presumptive set of parameters to pass the test when it with another set of parameters. To have a fair result of the testing, it is necessary for investigators to find a set of 'good' structural parameters. Thus, we estimate the models to get the optimal sets of parameters before the evaluating process.

The general working steps of Indirect Inference estimation-based test are summarized as follows and which are similarly and common to those mentioned in the previous studies (Le et al., 2011, 2013[2], 2016; Minford and Ou, 2013; Liu and Minford, 2014): Firstly, to select an auxiliary model (e.g., VAR) to estimate it based on the actual data to achieve the benchmark estimates. Secondly, give presumptive values to structural parameters which are needed to be estimated, after which the parameters will be used to create numerous pseudo samples of simulated data with the investigated theoretical model. Thirdly, to estimate the selected auxiliary model derived by the simulated data obtained from step two, which is done to produce the joint distribution of the selected estimates (from the first step) so that we can have the mean of this distribution. In the fourth step, we compute the Wald statistics and the transformed Wald statistics (normalized t-statistics)[3] to measure the distance between the benchmark estimates achieved in the first step and the mean of the estimates achieved in the third step. Finally, the second step to the fourth step will be duplicated until the minimum of Wald statistic is achieved.

It is obvious that the process of the second stage through Indirect Inference is similar to that of the first stage, apart from the last step. The reason for the distinction is the purpose of the second stage which aims not only to gauge the gap between the to-be-examined model and the actual data but also to narrow the gap by searching for an optimal set of parameters under the premise of the theoretical model being true.

[2] One advantage of Indirect Inference over the other method in terms of testing procedure, an alternative hypothesis suitable for testing of the specification of the model can be automatically generated by the unrestricted VAR model based on actual data, which leads us not have to specify different DSGE models as the alternative hypothesis. As a result, the identified VAR derived by the DSGE model is the only factor that required in this testing procedure.

[3] This function of Transformed Wald statistics (normalized t-statistic) is based on Wilson and Hilferty 1983's method of transforming Chi-square distribution into a standard normal distribution calculated.

After re-evaluating the three competing models through estimation-based test with the same US real-time data through the Indirect Inference method, we find that only the model with sticky information can pass the test meanwhile perform no worse than the baseline model. Additionally, among three competing models, only the model with Imperfect Information data revision fails to pass the test. However, when we use US survey data over the same period to evaluate each model, none of them can pass the test.

Overall, there are some consistent results through implementing two different estimation methods. Firstly, the sticky Information model is found to be the preferred model to fit the real-time data behavior which is examined in terms of Wald statistics (Wald percentile) and transformed Wald statistic. Secondly, the sticky information model is the only one can generate delay reaction to monetary policy shock and this is matching the observing evidence in actual data. Thirdly, the imperfect information data revision model performs better when we substitute the real-time data with the survey data, through which we can deduce that the survey data contains extra information to help improve imperfect information data revision model's performance.

Contributions

The main intention in this thesis is to evaluate the available original New-Keynesian reduced-form DSGE model with three different expectation assumptions respectively. The three models, which are taken into consideration, can be categorized into two groups: one is without inattentive features, the other one is including inattentive features. Within the first group which only has one model, we take it with full-information rationality expectation assumption to be the baseline. While in the other one, two inattentive expectation models are contained. They are the model of sticky information (Mankiw and Reis, 2002, 2007) and the model of imperfect information data revision (Vázquez et al., 2010, 2012, and Casares and Vázquez, 2016) respectively.[4]

We carry out the evaluation of the three competing models from three aspects: 1) to assess them through estimated impulse response function; 2) to compare the model-fit through Bayesian estimation approach, in which the relatively performance is determined by the log marginal likelihood or the Bayes factor; and 3) to use Indirect Inference as robust check method to see whether the candidate theoretical model can generate data close to reality. According to this point of view, the analysis in this thesis can be taken as the competing and selecting procedure of empirical models.

There are three main contributions of this thesis. The first contribution is the estimation and comparison of different types of inattentive DSGE model (sticky information versus imperfect information data revision) for US small-closed economy through Bayesian approach using the US quarterly data (i.e., real-time data and survey data) representing the main macroeconomic time series from 1969 to 2015. And the reason why we choose real-time data to estimate each model as Paloviita (2007b) asserts the significance of people's current knowledge and belief in leading their behavior in economic activities. As a result, in some cases, such as policy decision, if the economic relationships can be described potentially, so we can obtain a more precise research result. Se we use real-time data obtainable on the occasion instead of recently. Besides, another kind of data is used in our research in robustness check. Due to people's deficiency in predict the economy, we introduce the Survey of Professional Forecaster (SPF) data to simulate people's reliance on the experts. However, SPF data is not flawless. Its defect may be exposed when there is a big news which is opposite or averse to some experts' expectations, in which case experts may have intention to avoid significant changing of their predictions for maintaining their reputations. Overall, the two kinds of data selected lead us to find

[4] We use small-closed DSGE model instead medium-scale DSGE model different from Miguel Casares and Vázquez (2016); Vázquez et al. (2010, 2012) use the reduced-form model to study the data revision its impact on monetary policy and leave the rest economic agents without involving data revision issues.

the best way to describe people’s expectation formation. Through adopting these two kinds of data to evaluate each model which may provide us more accurate guidance to find the best way to describe people's expectation formation. What’s more, once we find the best way to describe how people form their expectation the government can affect real activity in ways that are correlated with that information (i.e., noisy revision information, sticky delayed information), this should greatly increase the credible range of conducting more stabilized policy.

The second contribution, through comparing the different inattentive New-Keynesian DSGE models basing on the full structure (relative to the single equations competition), we inspect which way of inattentive expectation is closer to the way that people form their expectation in real economy. The third contribution, the thesis is adopted Indirect Inference approach as the robust check methodology, which delivers a new way to assess inattentive macroeconomic models.

The outline of each chapter is demonstrated as follows. In Chapter 1, we survey the literatures on different New-Keynesian type DSGE models including the ones with and without inattentive feature. We also discuss the main findings from previous literatures. In Chapter 2 is the introduction of each competing model. Chapter 3 and Chapter 4 apply the two main analyses to examine three selected competing models respectively. In Chapter 3, we estimate reduced-form New-Keynesian type model without and with inattentive ingredient (sticky Information and imperfect Information data revision) through Bayesian estimation approach; Chapter 4 uses the Indirect Inference as the robust check method to test and estimate each competing model to re-examine the results obtained through Bayesian estimation approach. Chapter 5 contains the conclusion and discussion of further research direction.

Chapter 1

Whether Different Inattentive Features Matter for Economy Dynamics?

1.1 Introduction

The role of people's expectation in determining aggregate outcomes of the macro economy, such as inflation dynamics and the business cycle, has often been discussed and well established. However, the study involves how people form their expectation is relatively rare and less well studied. One recent study by Milani and Rajbhandari (2012) compares the full-information rationality New-Keynesian type model with the alternative models that deviate from the full-information rationality.[5] However, this topic is quite important for making the most fundamental macroeconomic decisions, such as the allocation of consumption or savings, how to set the appropriate price and so forth, some of which are underlying macroeconomic dynamics and driven by people's expectation of the future. In the following sections, we survey the literature focusing on the early assumption of fully attentive expectation or full-information rational expectation firstly and explore the weakness of this early expectation assumption. In order to remedy the weakness of full-information rational assumption, another assumption deviating from the full-information rationality has been proposed, which is so-called inattentive expectation assumption. In particular, we mainly focus on two types of inattentiveness, which are the most commonly discussed. The first is the model with sticky information expectation, and the assumption of sticky information is basing on the study proposed by Mankiw and Reis (2002, 2007). The second popular inattentiveness is imperfect information data revision (Aruoba, 2008; Vázquez et al., 2010, 2012; Casares and Vázquez, 2016). Both inattentive assumptions mentioned above will be well stated and discussed in later sections.

[5] Those models are set as being with the allowances of 'news' about future shocks, near-rational expectations, learning, and observed subjective expectations from surveys respectively.

1.2 Literature Survey of Classical New-Keynesian type Model without Inattentive Feature: Full-Information Rational Expectation

The full-information rational expectation hypothesis is the starting point of the traditional economic theory. However, a gap between this classical New-Keynesian full-information rational expectation (without any inattentive ingredient, i.e., Calvo, 1983) and the real world has been criticized for many economies. Simon (1989) criticizes the "unrealistic" view of the idea of full-information rational expectations. He argues that regarding the case of economic agents having known all of their problems, choices and possible results, the economic agents could certainly choose the best solution from all alternatives through some reasonable calculation. But in practice, such 'perfect situation' cannot be existent in real world. Besides, some unavoidable constraints always restrict economic agents from making good decisions (e.g., social constraint stemmed from the superior authority of government in terms of legislation or personal constraints originated from limited time and energy). Thus, economic agents have to seek coordination from the aspects of efficiency, profits and other factors. In other words, economic agents cannot simply reach the optimal solution but only reaching the self-satisfied or ‘good enough' solution. As a result, the full-information rational expectation can hardly be applied to explain economic problems.

On the other hand, the implicit hypothesis of full-information rational expectation is that the economic agents are homogeneous. But in real world, economic agents may form different expectations due to their different abilities in information acquisition, absorption, and procession. In other words, not all economic agents hold full information. To sum up, the unrealistic feature of early assumption of full information rational expectation can be showed from two aspects as follows:

1) The full-information rational expectation hypothesizes the economic agents having such full information that can do their best to reach the maximum profit. However, due to people's physical and intellectual capacity limitation, adding to the uncertainties originated from external environment, people are capable to understand and solve complex problems but in a restricted way.

2) Under the assumption of full information rationality, information is a kind of scarce resource that economic agents are willing to try their best to collect all available information to make economic decisions. Despite the desire to acquire information, it did not take the information costs (i.e., costs of accessing required information) into consideration. It is understandable that agents have to pay while collecting the information required for decision making. In practice, it is impossible to get and process information without the payment of time, money, or physical efforts. Due to these potential costs, the number and the quality of information obtained by the economic entities are limited, which lead to the fact that economic agents are impossible to reach the best situation.

To sum up, under the assumption of full-information rationality, economic agents are supposed to clear about the all relevant parameters' value, such as the distribution of shock, the correct structure of the economic model and so on. However, it is an unreasonable assumption in practice because economic agents cannot hold all the information needed to reach the equilibrium of the whole economy (Caballero, 2010). Particularly, when an economy undergoes a big structural transformation such as Great Recession, it will need never implanted policies (Stiglitz, 2011). The tune to full-information rationality hypothesis is favorable according to recent empirical work. Coibion and Gorodnichenko (2012, 2015) strongly deny the legitimacy of hypothesis of full-information rationality. Furthermore, in their paper published in 2012, they clarify that the reason of rejection to full-information rationality hypothesis is not the rationality but the assumption of full information.

1.3 Literature Survey of New-Keynesian type Model with Inattentive Feature: Sticky Information versus Imperfect Information Data Revision

To remedy the unrealistic aspect of the early full-information rationality assumption and deal with the well-known empirical weaknesses (i.e., the delay effect of monetary shock on inflation, persistent of output, and inflation observed in macro data), the New-Keynesian type model with the features deviated from the full-information expectation assumption appears as a modified version.[6] Thus, the inattentive expectation was proposed. As inattentive expectation has different approaches, the two most prominent of them are sticky information (Mankiw and Reis, 2007) and imperfect information data revision (Casares and Vázquez, 2016). These two assumptions will be applied in our research, being different from the sticky information model from Mankiw and Reis (2007) and the imperfect information data revision model from Casares and Vázquez (2016), we use the small-scale closed economy DSGE model instead of medium-scale DSGE to be in line with the baseline model selected.

Although there are weaknesses of the full-information rationality, as recent studies suggest that there is no need to abandon its assumption of rationality or to introduce other types of irrational behavior to help model fit data (Collard et al., 2009; Coibion

[6] There are also some literatures focusing on how to compensate the impractical aspects of the full-information expectation New-Keynesian type models through multiple ways (Rotemberg and Woodford, 1996; Gali and Gertler, 1999; Smets and Wouter, 2003, 2007). In these papers, the most attention is received and focus on real rigidities, such as habit persistence, capital or investment costs, capital utilization, and backwards-looking price setting schemes for the subset of the economic agents (Christiano et al., 2005; Collard et al., 2009). However, Dhyne et al. (2006) argues that backwards-looking price indexation setting scheme cannot support the empirical evidence. The European Central Bank Report pointed out that individual price changes its movement are not consistent with the movement of aggregate inflation. In explaining the observed situation, the idea of reducing controversy that encourages scholars to continue making efforts to resolve this issues in the past few years.

and Gorodnichenko, 2012). Thus, in this thesis, two major inattentive rational models, sticky information, and imperfect information data revision models, are used and compared, and meanwhile, rationality is assumed.

1.3.1 Literature Survey of Rational Expectation condition on Sticky Information

After the year of 2000, the problem of how economic agents forming their expectations of the aggregate economy begins to draw several scholars' attention. To address this issue, Carroll (2003), as one of the funders of this area, introduces the idea of "epidemiological expectations", in which the households form their inflation expectations by receiving the news reports that reflect views of professional forecasters, to explain the origin of the sticky information expectations. According to his study, the slowness of information diffusing through the entire population is due to people's inattentiveness to the arrived information[7].

Sticky information expectation which based on the idea of information slow diffuse through entire population is recommended in many studies. Being one related study of them, Mankiw et al. (2003) research the topic of how disagreement may appear among different agents' expectations of inflation. Their study is distinguished from other researches by finding the ubiquitous heterogeneity of different households' and professionals' inflation expectations. The heterogeneity was derived from different frequencies of the agents updating their information sets. Reis (2006a, 2006b) supports sticky information inattentive assumptions due to the cost of newly arrived information. He asserts that economic agents will only choose to obtain new arrival

[7] Some recent articles have based on Carroll (2003)'s studies to study the implications for monetary policy (Ball et al., 2005) and the dynamics of aggregate economy (Mankiw and Reis, 2007).

information if the expected benefit is higher than the information cost. Later, Mankiw and Reis (2007) develops and analyzes the medium-scale general equilibrium models for the US economy under sticky information assumption. They find that information stickiness exists in all markets throughout the quarterly data from the 1954 Q3 to 2006Q1. Moreover, the information stickiness is especially pronounced for consumers and workers in their study, the feature of information that being slowly disseminated in microeconomic data on price provides more credit to sticky information expectation (Klenow and Wills, 2007; Knotek and Edward, 2010). Mitchell and Pearce (2015) provide direct evidence of sticky information through examining the frequency of revision forecasts for individual professional forecasters. They find that the forecasters do not revise their forecasts usually, which is consistent with the sticky information hypothesis. In most cases, these literatures support sticky information assumptions.

1.3.2 Literature Survey of Rational Expectation condition on Imperfect Information Data Revision

Another strand about people's negligence deviates from the full-information rationality assumption is imperfect information data revision. On the perspective of microeconomic area, imperfect information refers to asymmetric information which is a common characteristic of the imperfect market. However, in macroeconomic area imperfect information means that economic agents are struggling to figure out the actual state of economy. In detail, the definition of imperfect information in terms of microeconomics implies that consumers can be easily fooled by the supply and price. However, under the environment of macroeconomics, imperfect information implies that economic agents involve signal-extraction problem (data revision issue). To be specific, economic agents are disturbed by noises and demand to filter useful signal or information from disturbing noises in observed actual data. The essence of the imperfect information is the inattentive behaviour that the economic agents can constantly update their beliefs, but suffering from the noises, which results the fact that the economic agents cannot fully observe the real state of economy (cannot be fully attentive). Hence, they renew their beliefs about the fundamentals of economy via signal extraction or data revision process to reduce noise.

Imperfect information expectation is recommended in many studies. Woodford (2001, 2003) integrates the idea of people's limited capacity in processing information, imperfect common knowledge, and the monopoly pricing competition to explain the persistence impulse response to real variables. Schorfheide (2005) who allow monetary authority to hold imperfect information (imperfect common knowledge) about the inflation target by modelling economic agent to learn and understand the fluctuating values over time. Although the model under imperfect information catches important periods like the early 1980s' disinflation better, the model under perfect information fits real economic data better. Additionally, Collard et al. (2009) demonstrate that the new Keynesian model under imperfect information environment could produce considerable inertia on an empirically reasonable level.[8]

[8] In the study by Levine et al. (2012), regarding the fact that people may not have all information of all state variables and all impacts on the economy, researchers establish a complete structural DSGE model in which the economic agents need to solve the signal-extraction problem to derive the values of state variables and impacting shocks, but such model is mainly governed by habit formation and adaptive learning. Therefore, the endogenous persistence impulse response generated from the model under the assumption of imperfect information the impulse response function generated by the model is close to the real situation. At the same time, they showed an example of analysis of the model under the assumption of imperfect information which fits the economic data well without introducing real rigidities (e.g., habit formation) or indexation price. The setup of our models does not have any interruptions of other features (i.e., habit formation) to check how model itself can reproduce the observed stylized facts.

1.3.3 Differentiate Inattentive Features: Sticky Information versus Imperfect Information Data Revision

The introduction of first inattentiveness is sticky information in Section 1.3.1 which emphasizes the recurring cost of collecting the latest information during making economic decisions, which may lead people updating the information reluctantly for the expense (i.e., cost of processing information) can be higher than its interest. Imperfect information data revision as the second inattentiveness is introduced in Section 1.3.2. It stresses the existence of the noises that influence people's decision by not reflecting the real state of the economy. Therefore, people via signal-extraction process or data revision process to reduce noises to figure out the real state of economy. Moreover, the model of imperfect information is based on the assumption of economic agents' limitation of information processing, so economic agents' decisions are determined by the information merely obtained through their information processing channel or communication channel (Sims, 2003).

It may be enquired that why we care about the inattentive feature -- imperfect information data revision. Diebold and Rudebusch (1991) give an example as the best answer from all analytical data revision papers. They explain that the major US economic indicators are doing well in forecasting the recession ex-post only because it is made to explain the past. Its tracking record in real-time, on the contrary, is very poor. Two reasons are given to this contrast. One is that the initial announced data may appear to be very different from the latest announced data. The other one is that the methodology of index changes as time goes by after the real-time indicator failing to forecast the recession. Beyond that, there is another example that easily to be understood to demonstrate this issue. Assuming we use the simple Taylor rule as a monetary policy to remain the level of inflation invariable, when the output gap is negative, the interest rate should decrease. Should th[e interestxxx rate increase, the case would be opposite (i.e., positive output gap). Evidenced by the same token, if the central bank holds economic growth data which is exaggerated before the recession, it would lead to the delay in adjusting interest rate to lessen inflationary pressure after the economic downturn. This example endorses the importance of inattentive feature.

Although a large quantity of literature has suggested to incorporate inattentive features into models to explain the real world, some issues still have not been well discussed. To supplement the areas that omitted by previous papers, our research focuses on verifying the three topics: 1). Do these two inattentive features matter in economic dynamics response; 2) If they are, what are the distinctions between them; 3). Which one can give a better explanation.

1.4 Conclusion and Objectives

In the literatures mentioned above, there are three relevant models which can be divided into two groups, i.e. with and without inattentive ingredients. One of them is the classical ‘attentive' expectation model, which is New-Keynesian type model with full-information rationality hypothesis. The second is sticky information model. The third is imperfect information data revision model. Three objectives will be reached through comparing the three models under different conditions.

The first objective of this thesis is to verify whether incorporating inattentive features into the popular reduced-form New-Keynesian model can perform better in replicating the empirical persistence found in macro-economic data than the full-information rationality alternative. The way to measure the performance of the model is to check its ability to generate persistent and delayed responses on output (output gap) and inflation to monetary policy (e.g., Christiano et al., 2005). Moreover, the model simulations will be carried out through Dynare 4.4.3 software.[9]

The second objective is to compare which expectation type model explains the US economy in the best way by using quarterly real-time data (survey of professional forecaster data will be used in robustness check). The process is implemented through Bayesian estimation approach. Through the comparison of Bayes Factor and the comparison of the log marginal likelihood of three competing models, the overall performance according to three rivals under different assumptions (i.e., fully attentive expectation versus inattentive expectation) can be compared and ranked relatively. The first advantage of using Bayesian estimation is that the application of priors which provides a chance to take the previous relevant studies into consideration and it facilitates to reduce identification issues in evaluating DSGE models.[10] The second advantage of Bayesian estimation is that Bayes factor provides an effortless way to evaluate model's relative performance.

The third objective is to use indirect inference to re-evaluate each competing model and make model comparison in an absolute way. Although the Bayesian factor provides a simple way to compare the relative performance of different models, it cannot be used to evaluate model's performance in an absolute way due to its limitation of judging that whether a to-be-examined model itself has a satisfactory performance that can be verified by the actual data. The method of distinguishing indirect inference estimation (estimation-based indirect inference test) from the Bayesian estimation method is to generate a data descriptor that indirectly evaluates the theoretical model by using a completely independent auxiliary model, e.g. VAR.

---

[9] From http://vermandel.fr/dsge-dynare-model-matlab-codes/, provide standard DSGE Models Dynare code, include the simple dynamic three-equation New Keynesian Model.

[10] Due to the structural interpretation of the parameters in DSGE models, sensible proper priors are usually available. These priors may be purely subjective or could reflect data from other sources (e.g., the estimates of structural parameters produced in macroeconometric studies and the estimates based on training sample of macroeconomic data). As the prior information given, Bayesian researchers do not need to worry about the identification issue. However, if a parameter is not identified, the data-based learning about it may be absent and its posterior only gives the reflection of prior information.

The intention of implementing estimation-based indirect inference test is to discover the optimal set of parameters about the actual data in the context of the model to make a fair model comparison.

Chapter 2
Introduction and Establishment of Three Competing Models

2.1 Introduction

The inclusion of inattentive features into macroeconomic model has become an active area of recent research. Carroll (2003) finds that the public's prediction is lags behind the prediction of professionals' through adopting survey of inflation expectation data. The study of Mankiw et al. (2003) their study shows that the disagreement of inflation expectations from survey data is matching the idea of sticky information. Furthermore, regarding to the recent work proposed by Dräger et al. (2013) they found that the impact of information friction on prediction errors at the individual level which provides support for imperfect information assumption (i.e., the economic agents suffer from noisy disturbance).

It is worth noting that, our study is not the first one to make a comparison between alternative expectation models and the full-information rationality type model. For instance, Milani and Rajbhandari (2012), who evaluated the alternatives (e.g., these alternatives include allowed "news" shocks, adaptive learning and observed survey expectations) deviate from fully-information rationality assumption in small-scale New-Keynesian DSGE model. Moreover, they have shown that the econometric characteristics of the model are susceptible to the different formations of expectation. Then our study can be understood as an analysis contributing to the selection of empirical models, which considers inattentive expectation type model as alternatives comparing with the baseline with full-information rationality.

2.2 The Introduction of Three Competing Models

The overview of each of the attentive and inattentive models will be specified as follows. The derivation of each model has been shown in Support Annex and the

Appendix B of Chapter 2. The three competing modes is a reduced-form New-Keynesian type DSGE model for a small-scale closed economy. Three types of agents are constituted the small-scale closed economy which are households, firms, and monetary authorities. The baseline model has been largely applied in previous studies (Milani and Rajbhandari, 2012) is the standard Calvo model without any inattentive features. In terms of the two other rivals, one is the model characterized by sticky information which has been discussed in Mankiw and Reis (2007), and the other one is the model characterized imperfect information data revision which has been constructed by Casares and Vázquez (2016). Being different from those two inattentive expectation model settings we are using the small-scale DSGE model instead of medium-sized DSGE model. Adding additional features might be a useful step (Smets and Wouter, 2003, 2007). However, it may also cause some fundamental issues to blur our main focus. Precisely, when each model being inserted with inclusion of some more new features taken into account, it may potentially distract some attention from the original focus to those new considered features, which leads to the difficulty of assessing the differences between the two inattentiveness (i.e., sticky information and imperfect information data revision). As well as the differences between the baseline model and the models with inattentive features, due to considering so many features.

2.2.1 Reduced-Form New-Keynesian Model without Inattentive Feature: Full-Information Rationality

The derivation of the classical small-closed New-Keynesian model is quite standard in the literature (Woodford, 2003). Here we present a more traditional version of the micro-foundation under the assumption of full-information rationality,[11] The details of

[11] The full-information rationality assumption type model applied in this thesis is chosen without indexation to past inflation and habit formation in consumers' preference, since the

the derivation have been presented in Supporting Annex at the end of this thesis. And the baseline model is as follows:

IS equation : (2.1)

PC equation: ———— (2.2)

Interest rate smoothed Taylor Rule: (2.3)

We have seen from the above presented baseline model, it can be indicated that the aggregate economy under reduced-form New- Keynesian type model with full-information rationality which can be characterized by the dynamics of three main economic variables (i.e., output gap, inflation, and interest rate). The represents output gap, which is a gap between actual output and potential output (i.e., is the output under flexible price economy). The coefficient represents the elasticity of the intertemporal substitution. The new Keynesian Phillips Curve (PC curve) derived under the full-Information rationality assumption is equivalent to the current inflation driven by the expectation of future inflation , current output gap , and the supply shock . The coefficient stands for the time discount factor and is the combined parameter.[12] Interest rate equation that follows the simple 'interest-rate smoothed' Taylor rule (1993). Monetary policy makers set the interest rate basing on simple Taylor rule. The interest rate is driven by the current inflation and current output gap .

---

premise of indexation has been shown to be not consistent with the microeconomic evidence on price set (Nakamura and Steinsson, 2008). The evidence regarding agents' habit formation is less obvious, but it seems difficult to find supportive evidence through households' consumption data (Dynan, 2000)

[12] Where the composite parameter =0.15 has been taken as fixed and less than one which it implies strategic complementary, to keep it as fixed and less than 1 and in line with the suggestion from previous literature (Woodford, 2001, 2003; Ball et al, 2005). Besides, Woodford (2003) surveys and discusses the existing literature at length and concludes that firms pricing decision should be strategic complements rather than strategic substitutes to allow for potential inflation inertia. And this has been tested in some recent works, for instance, Coibion (2006) these authors when which produce inconsistent results with the actual data.

2.2.2 Reduced-Form New-Keynesian Model with Inattentive Features: Sticky Information and Imperfect Information Data Revision

Before the introduction of the selected inattentive expectation models, we need to clarify the assumption concerning two inattentive expectations respectively. Regarding to the assumption of sticky information, the economic agents update their information sets infrequently due to information costs which reference to the idea offered by Mankiw and Reis (2002, 2007). Distinguished from the conception of sticky information, the conception of imperfect information data revision is that economic agents suffer from the noises, thus they continuously revise their information to extract the useful signals (Aruoba, 2008; Casares and Vázquez, 2016; Vázquez et al., 2010, 2012). In other words, the two different inattentive features can be taken as two distinct information arrivals. One of the principal purposes of this thesis is verifying whether different inattentive features matter in explaining economic dynamics. Furthermore, under the premise of confirming the determinacy of inattentive features, we will explore which feature can explain the US economy better from 1969 to 2015[13].

2.2.2.1 The Model with Sticky Information

The first inattentive feature to be introduced is the sticky information which assumes that some of economic agents use the old information rather than the current arrived information to make the economic decision and form their expectations. Since the cost of previously used information has been paid, there is no extra payment required for reusing old information, which is the way to reduce information costs. The main idea of the sticky information model is that when making economic decisions, due to the cost of acquiring newly arrived information as well as the cost of re-optimization,

[13] In order to construct the revised data in imperfect information data revision model, the sample period actually cover from 1969Q1 to 2016Q4.

only a small percentage of people are willing to use current arrived information to adjust their plans. On the other hand, the rest of people will still use the old information and old plan. The model with sticky information is presented as follows:

IS equation: (2.4)

PC equation: ———— + (2.5)

Interest rate smoothed Taylor Rule: (2.6)

Thus, according to the model with sticky information presented above, the two parameters δ and λ are the shares of updating households and the share of updating firms respectively in any given period (for example if there is no information stickiness of firms then λ=1). To compare with the economic agents in the full-information rational expectation model without inattentive feature, the economic agents are assumed under the premise of sticky information economy update their information sets with certain rate and regarding households and firms respectively (Mankiw and Reis, 2002, 2007; Reis, 2006a, 2006b, 2009). Reis (2006a, 2006b) gives more deep-seated micro-foundations for model features sticky information. The early classical New-Keynesian type model assumes of full-information rationality, which is the case of a pure forward-looking-expectation Phillips curve. However, under sticky information environment, the inclusion of inattentiveness leads to deviation from full-information rationality. The economic agents under this circumstance use the outdated information to form their expectation. Therefore, it yields the Philips curve (PC curve) not only depends on the current expectation but also the past expectation about the future, which is caused by information spreading slowly through the entire population of the economy (Mankiw and Reis, 2002)[14]. When looking into the previous empirical literature, several papers are aiming at comparing Phillips curve derived

[14] Being differentiated from the sticky information PC model of Mankiw and Reis (2002), the current inflation in our New Keynesian three-equation model is determined by both the current expectation and the past expectation of the future inflation rate. In contrast, the current inflation in Mankiw and Reis' model is inferred from flexible price assumption.

under the assumption of full-information rationality and alternative under the sticky-information assumption (Mankiw and Reis, 2002; Coibion and Gorodnichenko, 2012, 2015). However, in this thesis, regarding to the empirical evidence, we are more interested in the simple reduced-form New-Keynesian DSGE type models, rather than that based on single equation (Easaw, et al., 2014; Coibion and Gorodnichenko, 2015). Estimation of comprehensive DSGE models through introducing inattentive feature exists, but there is only a small quantity of papers. The recent papers on this aspect set a benchmark of neo-classical model with flexible prices and introduce sticky information regarding various economic decisions (i.e., consumption balancing, price setting, and wage setting) (Reis, 2009). To the best of our knowledge, no one has compared DSGE models under different inattentive conditions (i.e., sticky information assumption versus imperfect information data revision assumption).[15] So here one of our main emphasizes is to use the model with sticky information to compare with the alternative inattentive expectation model (i.e., the model with imperfect information data revision) to examine which inattentive expectation model can give the better explanation for US economy in around recent five decades (sample period US quarterly data from 1969 to 2015).

Comparing with the baseline model, it is more challengeable to solve the model with sticky information. Since it involves infinity lagged expectation what leads to the question of how we can approximate the model with sticky information in the DSGE equilibrium configuration. Firstly, from the angle of sticky-information model setting, we can see that the proportion of lagged expectations diminish geometrically meaning that the impact on economic agents' expectation derived from the current state is far greater than that of previous periods. Consequently, the expectations that are formed very far from the present situation might not influence current inflation or output gap

[15] From an empirical point of view, for instance, Smets and Wouters (2007) may consider that a more satisfying specification may take into account some frictions. However, in this thesis, we would like to keep it simple, since one of the main questions we would like to focus is to differentiate different inattentive feature and to see whether different inattentive feature matters for dynamics of the economy.

due to the minimal weight (i.e., may approximate to zero) attached to them. Thus, we set j=4 (which meaning the incorporation of lag information up to 4 periods) as the benchmark, the longer period such as j=6 and 8 have been taken in robust check section.[16]

2.2.2.2 The Model with: Imperfect Information Data Revision

For the extend model with imperfect information data revision process both real-time data and revised data has been used, the suggestion comes from the previous studies (Casares and Vázquez,2016; Vázquez et al., 2010). Before introducing imperfect information data revision model, firstly we need to know what is real-time data, for example, if we analyze the economic agent's decision using the data available to us today, we will make an incorrect inference about their economic decision-making. If we look at the time that economic agents made their economic decisions, we are engaging in real-time analysis or taking the data revision seriously into consideration. The model with imperfect information data revision is presented as follows:

IS equation: (2.7)

PC: (2.8)

Interest rate smoothed Taylor Rule: (2.9)

Where and . Data revision is potentially critical in both theoretically and empirically way, although many economic researchers have made an inappropriate assumption about the data available to economic agents at that point. The applied assumption of data is that they are available immediately, yet the reality those data are announced with a few lags. Furthermore, the data

[16] The result in Travandt (2007), by setting maximum j=19, the convergence of the recursive equilibrium law of motion can be achieved for sticky information Phillips Curve model. However, in our selecting sticky information model enter competition is using fewer periods j and which is sufficient to reach convergence.

revision, in general, has been thought either not exist or small, but in real situation data revision may have a significant and big influence on empirical results and which is particularly the case of some variables that are defined conceptually. For instance, such as output gap, where the economic agents when they are making decisions, take this kind of variables know without any doubt. In a real case, such variable as output gap often fluctuates over time. Thus, in this imperfect information model, the data revision has been taken into consideration to see how it has affected New-Keynesian type macroeconomic model as well as empirical results

Moreover, what does data revision look like is followed by the suggestion from Casares and Vazquez (2016) and has been well specified in Appendix B to Chapter 2. Apart from the point as mentioned earlier, another two points should be clarified: 1) under imperfect information data revision hypothesis, the information of the economy its real state matters, for instance, firms' price-setting decision depends on the expectation of marginal revenue and the future nominal marginal costs. Thus, depends on the future aggregate price level. 2) information friction or inattentive feature underlined across this thesis to be taken seriously, such inattentive assumption needs to be reasonable. Where the nominal interest rates made through professional monetary authority are fully observable without noise disturb, and the observation of output gap and inflation are influenced by noises, in other words, both variables involving data revision processes. Collard and Dellas (2010), they argued that, as the data revision process reveals, very few aggregate variables can be observed accurate and correct. Such that, under the assumption of imperfect information, when firms make the price-setting decision cannot fully observe its information, on the other hand, households when to make consumption decision cannot fully observe the state to support them to make consumption plan. Such that, both price (inflation) and consumption (output) can only observe with some random noises. From the above three-equation model where and have been taken as the observed variable realized at time t they are the real-time data. And and are the final revised variables and which are stated as followings.

(2.10)

(2.11)

And we follow by the argument of Aruoba (2008) that many US aggregate time-series (e.g., inflation and output) their revisions are not rational forecast errors and supposed to be connected to their initial realised variables and . Thus, following his argument, we presume that final revision process of US output gap and inflation are defined as follows,

(2.12)

(2.13)

These revision processes allow for the existence of non-zero correlation between final true variables (i.e., output gap and inflation) and their initial realised variables. Besides, the existence of persistence revision processes. In particularly, the shocks of revision processes, and , both are the AR (1) processes. The two data revision processes assumed aim to offer a simple framework to approximate the 'true' revision processes, and to examine whether the deviation of the way we use for assumption to the well-behaved revision processes (i.e., white noise) assumption, influences the estimation of policy and behavioural parameters

.

For simplicity, we assume that revisions process is linear, following Casares and Vazquez (2016), since our estimated model is a linearized-reduced form version of a small-scale closed New Keynesian model. However, noteworthy, Corradi, et al. (2009) finds the evidence which supports that there is a nonlinear relation between data revisions and variables, which can be an interesting further research in the future. In benchmark competing process, we assumed that the final revisions are reached after 3 quarters, namely s=3 when solving the imperfect information data revision models.

Worth noting that, there are existing studies to contradistinguish the DSGE model with full-information rationality with the alternative DSGE model with sticky information. For example, Paustian and Pytlarczyk (2006) evaluates DSGE model for euro area based on Smets and Wouter's (2003) model through Bayesian estimation approach, and their main finding is that, and Calvo full-information rationality type model overwhelmingly dominants the model with sticky information regarding the posterior odds ratio. Trabandt (2007), use the full-specified DSGE model under the sticky-information assumption and compare it to the Calvo full-information rationality type model, and with allowance for the dynamic inflation indexation (e.g., Christiano et al., 2005), and found that both do equally well. Meanwhile, studies aim to compare the full-information rationality type model with the alternative with Imperfect Information Data Revision also existing. (Paloviita, 2007b, 2008[17]; Vazquez et al., 2010; Casare and Vazquez, 2016[18]), and they provide that the employ of real-time-data variables improves the empirical behave of the classical New-Keynesian model, moreover relax the full-information rationality expectation tentative generates a remarkable distinction for the parameter of the New-Keynesian model.

2.3 Conclusion

For each model with and without inattentive feature, first, it has assumed AR (1) process for all disturbances to each structural equation to capture omitted variables.

---

[17] Paloviita (2007b, 2008) uses the European panel data and apply GMM system estimation to investigate the empirical performance of the standard three-equation New-Keynesian reduced-form model under different information assumption, compare the full-information rational expectation with measured expectation through using revised (final) data, but in their used three-equation without no systematic error. Their estimation results provide evidence that incorporate data revision make the significant difference for parameters, particularly for monetary policy.

[18] Vazquez et al. (2010, 2012), based on three-equation framework to incorporate data revision issue into monetary authority, on contrary we assume monetary authority leave without data revision issues, but economic agents (households and firms) through data revision process to reduce noise to in order to figure out the real state of economy; Casare and Vazquez (2016) to incorporate the data revision into Smets and Wouter's medium-scale type model.

Besides, the frequency of each variable is quarterly, and each variable is demeaned variable, detrend data will be applied. Note that these three models have different information friction constraint, therefore having different IS and Phillips Curve (PC), and therefore may influence monetary policy. After then, by comparing their data fit ability (i.e., log marginal likelihood and Bayes' Factor), one should be able to say whether the suggestion of incorporating inattentive feature from previous literature can provide a better explanation for US economy relatively. Moreover, further explore whether different inattentive feature matters to explain economy dynamics.

Various macro-econometric methods are applied to do model estimation and comparison. The first applied analyzing method is Bayesian estimation approach, which is used to evaluate each model's performance through using US quarterly data in Chapter 3. One of the most significant strengths of Bayesian estimation method is that it provides a solution to find the relatively 'best' model, which can be done with the assistance of a model's marginal likelihood which is directly relevant to the model's prediction ability. Thus, the models for forecasting and policy analysis can be verified by the benchmark of the performance of prediction. Meanwhile, another criterion to verify the relatively 'best' model is the Bayes factor. Different prior distributions and different types of observations are used for robustness check.

Appendix A to Chapter 2

Table 2A-1 Reduced form for each economy to be estimated

	$+$
	( )

Appendix B to Chapter 2

B1. Sticky Information Model Derivation

B1.1 Sticky Information Model: IS Curve

Now we assume economic agents, households under the sticky information economy use the outdated information from all past period up to t to form their forecast, and in aggregate level not all of them use the updated information to form their forecast, then we have the following IS equation. Where denotes the share of updating households.

(B1.1)

B1.2 Sticky Information Model: Phillips Curve (PC)

Similarly, for firms, also subject to sticky information, and in aggregate level they are using not all of them use the update information to form their forecast, firms use the outdated information up to time t to form their forecast then we have the following PC equation, where denotes the share of updating firms.

$$+$$

$$+ \qquad \text{(B1.2)}$$

From above we can see the current inflation thus depends on the current output gap as well as on current and past expectation of the future inflation rate.

B2. Imperfect Information Data Revision Model Derivation

The derivation of Imperfect Information Data Revision Model is following the deriving procedure and assumption explanation are following by Aruoba (2008), and Vázquez et al. (2010, 2012) and Casares & Vázquez (2016). First, let us consider the following identities regarding revised data related to cyclical of output gap and inflation, and which is the combination of the initial announcement and the final revisions. Which can be interpreted in the sense of noise, and have been taken as the observed variable realised at time t they are the real-time data. And and are the final revised variables and which are defined as follows.

(B2.1)

(B2.2)

And we follow by the argument of Aruoba (2008) that many US aggregate time-series (e.g., inflation and output) their revisions are not rational forecast errors and supposed to be connected to their initial realised variables and . Thus, following his argument, we presume that final revision process of US output gap and inflation are defined as follows,

(B2.3)

(B2.4)

These revision processes allow for the existence of non-zero correlation between final true variables (i.e., output gap and inflation) and their initial realised variables. Besides, the existence of persistence revision processes. In particularly, the shocks of revision processes, and , both are the AR (1) processes. The two data revision processes assumed aim to offer a simple framework to approximate the 'true' revision processes, and to examine whether the deviation of the way we use for assumption to the well-behaved revision processes (i.e., white noise) assumption, influences the estimation of policy and behavioural parameters. Therefore, from above defined equation we can get,

(B2.5)

(B2.6)

Furthermore, notice that final revision process of output gap and inflation also imply the identities' equations that,

(B2.7)

(B2.8)

(B2.9)

(B2.10)

B2.1 Imperfect Information Model: IS Curve

Use the imperfect information data revision assumption, to distinguish from the baseline Full-Information Rational Expectation model, here we can get the IS equation where households involve data revision issues, these imperfect-information type people react to expected revised values of inflation and output gap,

(B2.11)

And then we use the identity equation and to substitute out and from above to get the imperfect information IS equation as follows,

(B2.12)

B2.2 Imperfect Information Model: Phillips Curve

Firms involve data revision issue (noise disturbance) we can get the imperfect information Phillips curve,

$$+ \qquad (B2.13)$$

Similarly, we use the identity equation to substitute out from above to get,

$$(B2.14)$$

Meanwhile, the monetary policy assumed perfect observed and live without data revision issue,

$$(B2.15)$$

Where the final revision and their data be constructed as demeaned observables between the first released and the more latest released as follows,

$$(B2.16)$$

$$(B2.17)$$

So, here for analysis we choose for s=3 to construct the observations of final revision and ,

$$(B2.18)$$

$$(B2.19)$$

Therefore, we can also construct the observation of revised data and .

Note that as argued by Croushore (2011), if we look at the US data, which will show us that s is neither constant with the passage of time nor across variables. One may need to check whether the alternative of s will significant influence Imperfect information data revision its model performance. Here we choose s=3, as the data released in 2016Q1 and as the data released in 2016Q3 to construct the revision process corresponding to sample period from 1969Q1 up 2015Q4. For the simplicity of the analyzing procedure, the number of periods after which without more revisions, except benchmark revisions, and which is represented by s and to be constant.

Chapter 3
Estimate New-Keynesian Type Models with Inattentive Feature through Bayesian Approach

3.1 Introduction

In chapter 3, we focus on estimation and comparison basing on the reduced-form New-Keynesian DSGE model which was restricted by different inattentive expectation assumptions, such as sticky information expectation and imperfect information data revision expectation for the US economy over period 1969-2015 using a Bayesian approach. The three aims of this chapter are to explore which expectation model can reproduce the dynamics behavior of the US real-time data best through Bayesian estimation approach (survey data also used as the alternative type of observations in robustness check), to verify whether incorporating inattentive features can improve the model's performance, and to discuss how different inattentive ingredients influence the dynamics of the economy which can be checked from estimated Impulse Response Functions.

Bayesian estimation approach evaluates different kinds of model by comparing the marginal likelihood of them in a reasonable way. The natural parameters with respect to chosen applied models and the stochastic processes, manage the structural shocks derived by three key quarterly macro data in US economy: output gap (use the real GDP, and output gap is the difference of log of real GDP and log of potential GDP), inflation (log of implicit price deflator) and nominal interest rate (effective federal funds rate). We follow Bayesian estimation approach to evaluate each competing model through three stages. Firstly, through integrating the prior information of the parameters and the likelihood of the data, we can have the log of posterior function, by computing the maximum of which the mode of the posterior distribution can be reached. Secondly, to implement MH algorithm which enables us to obtain a full picture of the posterior distribution and the evaluation of the model's marginal likelihood. Finally, the comparison of three various models in terms of models' performances: full-information rationality expectation model, sticky information model and imperfect information data revision model are analyzed in the result.

The findings presented in this chapter indicate that the US three main economic quarterly real-time data forcefully prefers the model under the assumption of sticky information. Moreover, then through Bayesian estimation approach, we find that the specification with the sticky information outperforms other versions according to marginal likelihood and formal criterion Bayes Factor. Furthermore, the estimated parameters have reasonable values that agree with those typically analyzed in the literature. The model with imperfect information data revision ranks as second outperform model. The baseline under full-information rationality hypothesis type model performs worse than either of inattentive assumption models. We interpret these findings through Bayesian estimating approach as suggesting that incorporating inattentive feature is needed for the New-Keynesian rational expectation model to be a better monetary business cycle model. Besides, different inattentiveness does have impact on the three aspects that used to explain economic dynamics. The three aspects are estimated posterior distribution, estimated impulse response function, and significant different values of log marginal likelihood.

The rest sections of Chapter 3 are structured as follows. Section 3.2 contains the involved literatures about DSGE model estimated through the Bayesian method. Section 3.3 contains the description of the Bayesian estimation approach applied in this chapter. Following that description, Section 3.4 includes the explanation of the data and priors' estimation. Section 3.5 analyzes the assessments of estimation results and comparison results will be showed in section 3.6 and 3.7. Finally, Section 3.8 summarizes this chapter.

3.2 Related Literature of Estimating DSGE Model through Bayesian Approach

Bayesian estimation approach has often been applied to estimate DSGE model in recent years. Within large-scale of recent literature, some of them have been paid significant attention, for instance, Schorfheide (2000) uses Bayesian approach to contradistinguish the model fit of two rival DSGE models of consumption. Lubik and Schorfheide (2005) studies whether small-open economies' central banks are in response to exchange rate volatility. Smets and Wouters (2003) evaluate European countries through using Bayesian estimation method. Rabanal and Rubio-Ramirez (2005) evaluate four various competing New-Keynesian type models with nominal rigidities though Bayesian estimation approach by comparing the model fit.

There are some papers studying the similar topic as ours by applying Bayesian estimation approach as well. For instance, Mankiw and Reis (2007) evaluates model through Bayesian estimation approach to check the influence of sticky information on macroeconomic dynamics and policy base on the general-equilibrium framework. In Collard et al.'s (2009) paper, the possibility that through introduction imperfect information in New-Keynesian type models improves the model fit has been evaluated through Bayesian estimation method. Milani and Rajbhandari (2012) evaluate a variety of expectations formation models through Bayesian estimation approach, and their study shows that when the assumption of full-information rationality has been relaxed and adjusted, the significant shift of the posterior distributions of the structural parameters exist. Levine et al. (2012) compare perfect information with the alternative imperfect information by applying Bayesian estimation approach and finds that information is an essential factor for estimation. Besides, in Levine et al.'s (2012) paper, the results show that the New-Keynesian type model under the imperfect information assumption fits the observed autocorrelation of the data. Whereas, the model under the assumption of perfect information results in a poor model fit. Thus, the analysis in this thesis can be thought as the empirical model competing and selecting exercise by comparing the New-Keynesian DSGE type model under full-information rationality assumption with the alternatives under inattentive expectation assumption (i.e., sticky information and imperfect information data revision).

3.3 Bayesian Estimation Methodology

The reason why Bayesian technology has become increasingly popular in recent studies

For the question of why Bayesian technology has become increasingly popular in recent studies, there are three reasons given and repeatedly been stated by researchers (Geweke 1999; Fernandez-Villaverde and Rubio-Ramirez, 2004; An and Schorfheide, 2007). Firstly, the use of priors allows the previous both macroeconomic and microeconomic researches to be taken into consideration, which offers a way to connect to the previous useful literature. Secondly, the Bayesian estimation approach can give us valuable and stable results under the circumstance of the sample data is comparably small. The Bayesian estimation approach offering a way to assess a model with fundamental misspecification is the third reason. Because what can be accomplished by using the models' marginal likelihood or by using formal criterion Bayes factor. Bayesian Economists may argue that the DSGE model is an approximately/comparative specific version of modelling reality for there is no one hundred percent true model. Therefore, the Bayesian estimation approach is consistent with the argument that no model can be used to describe the real world correctly. Hence, the Bayesian method is used to study the DSGE models which agree with the beliefs held by many macroeconomic researchers.

The classical Maximum Likelihood Estimation approach was argued as a relatively weak estimation method since this approach has been proved to be feasible only for relatively small size systems but not appropriate to be employed to estimate large-scale type models. For instance, Canova (2009, pp.432) says that: 'One crucial but often neglected condition needed for a methodology to deliver sensible estimates and meaningful inference is the one of identifiability: the objective function must have a unique minimum and should display 'enough' curvature in all relevant dimensions.' If a model has a lot of parameters required to be estimated, we will meet trouble to achieve the correct information about the estimated parameters from the data. Thus, two problems are exposed 1) likelihood may produce estimation results that are not reflecting the information which is held by researchers, namely, the likelihood peaks in odds area; and 2) the parameters with being given various values result in same joint distribution for the data observations, namely, likelihood without enough curvature (i.e., within a large subset of parameters its likelihood is flat). However, the two problems mentioned above can be avoided or at least reduced by 'reweighting' likelihood function through using Bayesian approach after introducing the prior distributions to yield a function with sufficient curvature, therefore, can yield a function with sufficient curvature. From this aspect, the Bayesian estimation method is more capable of dealing with identifying problems.[19] Additionally, Bayesian estimation approach enables one to take advantage of the prior information from the fore literatures, either the reflection of the subjective view of the investigators by a particular prior probability density function of the parameter. However, the classical maximum likelihood cannot take even the prior information with the most non-controversy.

.

Moreover, Bayesian estimation approach can minimize the problem which usually caused by using classical maximum likelihood estimate. By using classical maximum likelihood, the overall estimation process is not very insensitive to each parameter its

[19] Detailed discussion of identification problems in DSGE models see from Canova (2009).

estimated value, which means that if the observed data give poor supports to one or more parameters of the estimated model, will result in a bad estimating result. Instead of embodying a specific value of each estimated parameters, Bayesian estimation approach allows the estimated parameters to follow a distribution that encompasses possible estimates.

3.3.1 The Application of Basic Rules of Bayesian Econometrics

3.3.1.1 Bayes' rule

The basic rule of estimating model through Bayesian approach is the Bayes' rule. For example, suppose there are two events , and , but the probability of event given by event depends on both the relationship between event and event and the prior probability of each event occurrence, which case can be incorporated in the Bayesian estimation as follows:

‘

The formulation above is the prior probability of event , which contains no information about event . is the conditional probability of event conditional on event , which is the posterior probability and which is derived from the specified value of . is the conditional probability of event conditional on event , which is also called the likelihood. is the prior probability of event .

3.3.1.2 Application of Bayes' rule

In the application of Bayes' rule, event is the counterpart of the model's parameter needed to estimate and event is the counterpart of the observable actual data . The posterior density of the parameters can be obtained through combining the likelihood function, prior, and the marginal probability. The parameter is a random variable under the premise of Bayesian econometrics. Bayes' rule is implemented through substituting the corresponding factors:

Where is the prior containing no actual-data information available about the parameters of the model . is the likelihood function which is the density of the observed actual data being conditional on the parameter of the model . is the posterior whose function is summarizing all the information about the parameter of the model after observing the data. The expectation of the posterior can give a point estimate after being calculated.[20] Besides, is the probability of new evidence under all cases usually normalized and constant.

3.3.2 The General Working Steps of Bayesian Estimation

Basically, Bayesian estimation can be taken as a link between priors and maximum likelihood function. The maximum likelihood approach enters through the estimation processes based on confronting the model with the data. The likelihood function implies the probability of observing a given set of data. The priors can be taken as a

[20] The highest posterior density interval (HPDI) gives the smallest credible interval : (i.e., interval estimates).

tool to re-weight the likelihood function to insure more significance to specific regions of the subspace of parameter., Likelihood functions and priors as two components are combined through Bayes' rule to construct posterior distribution. The general working steps of Bayesian estimation are summarizing as follows.

1) To formulate our understanding about the situation: Firstly, to define a distribution model which expresses the qualitative aspects of our understanding of the situation. This model will contain some unknown parameters, which can be treated as random variables. Secondly, to determine a prior probability distribution, which represents our subjective beliefs and uncertainties about the unknown parameters before observing the data.

2) To collect observed data.

3) To get posterior knowledge about our updated beliefs by the means of calculating the posterior probability distribution

Finding the posterior distribution of the models' parameters, if the posterior distribution is ordinary, such as Chi Squared distribution or Normal distribution, it indicates available for us to find the analytical solution. Accordingly, Monte Carlo integration can be carried out to get the estimation of posterior straightly. However, if , namely the posterior distribution is not ordinary, then we are not able to achieve the random draws straightly from the posterior. As solutions to this issue, two ways have been provided to get well approximated posterior density: 1) To implement Independent Draw Approach: draws from the posterior distribution are independent of each other. e.g., importance sampling and acceptance sampling[21]; and 2) To implement Dependent Draw

---

[21] Importance sampling is a technique for estimating properties can be used under particular situation, e.g the distribution possessing samples generated from another distribution; Acceptance sampling uses statistical sampling to determine whether to accept or reject candidates. It has been a common quality control technique.

Approach: Draws from the posterior distribution are dependent on the previous draws (Marko chain Monte Carlo (MCMC) algorithm) e.g., Metropolis-Hasting Algorithm

In this chapter, we use the Metropolis-Hasting algorithm as a sub-method of MCMC sampling method to get the model parameters' posterior distributions. To be specific, the concept of the Metropolis-Hasting algorithm refers to produce a Markov-Chain that exhibits a sequence of feasible estimates of parameter by exploring the entire domain of the parameter space. Hereafter, the simulated posterior distribution is constructed through using the frequencies related to each value of estimated parameters to construct histogram accordingly mimic posterior distribution. To take the process above, the Metropolis-Hasting algorithm first specifies a 'candidate' distribution, from which some parameter estimates can be drawn. An 'acceptance-rejection' rule is given to determine which several generated estimates are retained. The Metropolis-Hasting algorithm is used to pick qualified 'candidates' distribution under the general regularity conditions. Thus, asymptotically normal posterior distribution can be obtained. Then, by employing the Hessian and mode which achieved from the maximization of the posterior kernel to determine the mean and variance.[22]Summarized Metropolis-Hasting algorithm working steps

Metropolis-Hasting algorithm takes draws from convenient candidate generating density. Let indicate a draw taken from this density which we denote as . The Metropolis-Hasting algorithm always takes as follows:

1) Choose the starting value,

2) Take a candidate draw, from the candidate generating density, which denotes the density of depends on .

[22] The variance is constructed as an inverse of the Hessian multiplied by a scale factor which determines the acceptance ratio.

3) Calculate the ratio of acceptance , which provides the ratio of acceptance of as a draw from the posterior. The ratio of adoption tends to shift the chain from the low-posterior-probability region to the higher-posterior-probability regions.

Set with ratio of acceptance , and set with the ratio of non-acceptance . These new estimates will be reserved with probability and rejected with probability , which facilitates us to search the whole area of the posteriors' distribution. We are not supposed to simply remove the candidates immediately for the reason that a smaller value of the posterior kernel might give us a chance not to be trapped around a local maximum, enabling us to reach the global maximum. What forms the center of the distribution is the highly favorable values which are represented by various appearances, while the tails of the distribution are constituted by the less favorable values.

4) Repeat steps 2, 3 and 4 S times

5) To take the average of the S draws ,……

To assure that the influence of initial values has disappeared and draws converged to posterior, the initial draws are abandoned, and the rest draws are kept as those from the posterior for the estimate of .

Overall, the Metropolis-Hastings algorithm, as an approach classified from MCMC algorithm, is applied to simulate posterior distribution in the cases when we meet trouble to carry out directly sampling. Therefore, Metropolis-Hastings algorithm is chosen to get sequences of random samples from the probability distribution. The current draw of Metropolis-Hastings algorithm always depends on the previous draw of that to get a candidate's probability density function which is represented as a chi-square distribution or a normal distribution. The variance of the 'candidate' distribution, especially the scale factor of it, plays a key role of the procedure. Since the Markov chain relies on the rate of acceptance in visiting the entire distribution to get the global maximum. In practice, a too small variance leads to an obstacle of scanning all the maximum to get the global one, while a too high variance result in a difficulty in finding the global maximum due to taking too much time for visiting the tails of distribution. As a result, a proper j-scale is important to determine the variance of 'candidate' distribution. It is suggested to set j-scale to make the acceptance probability within 20% to 40%. Besides, basing on the principle that the more the 'buckets' iterated in distribution, the less each of them takes the weight, which leads the histogram more complying with the desired theoretical distribution. Thus, the level of smoothness of histogram is increased correspondently with the amount of iterations. In this way, we can get approximated posterior distributions[23].

3.3.3 Model Competing and Selecting through Bayesian Approach: Marginal likelihood, Bayes' Factor, and Posterior Odds Ratio

[23] Approximated posterior distributions featured by the location (i.e., mode and mean) and dispersion (i.e., probability intervals and standard deviation). This methodology offers not only point estimate of the structural parameters but also a measure of the uncertainty around these estimates.

3.3.3.1 Marginal Likelihood of the Model

The key for modeling comparisons through Bayesian approach is the function called marginal likelihood .In detailed, donates the observables, and denotes objective estimated model. Hence, the marginal likelihood function is the counterpart of the density of the observed data while being given objective estimated model yet unconditional on the models' parameters. Through integrating out the parameters of the model we can work out the marginal likelihood function which are presented as follows:

where and are the prior probability and the likelihood function respectively. To solve the function involving multidimensional integration which is not tractable analytically, it requires to repeat to gain numerical approximations. The two most commonly applied methods are the Harmonic mean estimates and Laplace approximation estimates. The former solution utilizes the Metropolis-Hastings runs to simulate the marginal likelihood to take the simple average of the values which obtained from the simulation. The latter solution assumes that the Gaussian distribution can be adopted to approximate the posterior kernel.[24] From the two methods mentioned above, the former method is preferable, because it does not assume the posterior kernel as any formation of function (e.g., Gaussian distribution). If the assumption or restriction is incorrect, it may result in inaccurate results. Although the latter method consumes relatively less time on computation since it only requires some numerical calculations of posterior model and requirement of Hessian matrix, it has a restriction on functional formation of the posterior kernel, which may issue in inaccurate results. In general, the marginal likelihood is a natural way to measure

[24] Gaussian distribution can be used to approximate the posterior kernel to assesses its integral at the mode and variance.

model's unconditional overall performance through representing the overall likelihood.

3.3.3.2 Bayes' Factor

One criterion used to measure model's relatively performance is the Bayes factor which takes the responsibility of verifying ability of models in empirical uses. This Bayes factor is the simple ratio of marginal likelihoods between any two models, such as, and . The Bayes Factor is given as,

In this way, the Bayesian factor is a rule that based on models' fit to sample data ( ) to compare two relative models.

3.3.3.3 Posterior Odds Ratio

Another criterion is the Posterior Odds Ratio which is a more completed tool can be applied to measure model's relatively performance. The construction of Posterior Odds Ratio considers the case of two rival models and .If one assigns prior probabilities to each model and after data being observed to name them and respectively. After then, applying Bayes theorem with them, we can work out and in the same way as we apply to obtain the posterior distribution of parameters.

is the unconditional density of the sample data. Since, does not rely on either parameter of the model or specification of the model, we can substitute into to get,

Similarly, for model ,

Then the Posterior Odds Ratio of model versus model can be represented by the function as follows:

The posterior odds ratio can be regarded as a measurement of the relative performance of two competing models not only based on their model fit to the same sample data (i.e.,), but also on their beliefs concerning the probability of belief of each model (i.e., prior ratio,) . If one knows nothing about which one is more aggregable, the equal weight has been assigned to each model (i.e.,). Herein, there is no different between the Posterior Odds Ratio and the Bayes' Factor. The optimal determination through Posterior Odds Ratio criterion is to choose the one gains the highest posterior support, for instance, if we will select model , yet if we will select model .

3.4 Sample Data and Priors

3.4.1 Whole Sample Data

To estimate the parameters of the three competing models through Bayesian approach, three main macroeconomic variables from year of 1969 to 2015 of US economy are used (i.e., real GDP, GDP deflator, and the nominal interest rate). Sample data and their descriptions are presented and defined in Appendix C of Chapter 3.

3.4.2 Priors

In applying Bayesian estimation approach, the incorporation of prior distribution plays an essential role in estimating DSGE models. The specification of prior distributions (i.e., probability density function of the parameter) is where the Bayesian estimation process begins. The selection of prior's distributions can be made basing on several norms. For example, some most common applied distributions are as follows, restricting the parameters to be positive through inverse gamma distribution, restricting the parameters between 0 and 1 through the beta distribution, and restricting the parameters without any bound through the normal distribution respectively. Besides, the values of priors can be presumed from either previous studies or the investigators' subjective views. The ways of selecting the value of prior have to be in line with the analyses of the context of the model, which means the construction of each priors includes non-sample-data information in the estimation. In other words, priors constitute extra independent information on the model's parameters.

Moreover, selecting values of parameter facilitate to define its distribution which contains the measure of location (e.g., mean and mode) and dispersion (e.g., probability intervals and standard deviation). To this end, the parameters are usually divided into two groups. To be specific, the first group refers to the parameters with relatively strong prior beliefs about (e.g., involve the core structural parameters of the model); the second group refers to parameters with relatively weak prior beliefs about, (e.g., involves the parameters that used to characterize the structural shocks). In the former group, the priors of the parameters are based on the survey of existence of the empirical evidence as well as their implications for macroeconomic dynamics. Although we can adopt the parameters in the latter group based on surveying the previous literatures, in order to constrain the prior distribution within a considerable scope of parameter values the strategy of setting priors needs to be reasonable with proper density that derives from sufficient supports. Since the priors are created from normal standard densities, its computation is quite straightforward.

Most of parameters' prior distribution are chosen from previous literature within a reasonable range for explaining US economy. For instance, the price stickiness which is represented as whose value is 0.6 has been used in many empirical studies (Blinder et al, 1998; Nakamura and Steinsson, 2008; Milani and Rajbhandari, 2012). Additionally, the values of sticky-information parameters and what are 0.5 both are borrowed from Mankiw and Reis (2007)[25]. Moreover, the values of the parameters, regarding imperfect-information data-revision, and are set with mean 0 under the circumstance of allowing large standard deviation from the reference of Casares and Vazquez (2016). Meanwhile, some of the parameters' priors are very strict, and are set fixed before the exercise. Taking the time discount factor and the strategic complementary parameter as examples, they are fixed as 0.99 and 0.15 respectively.[26] We have little knowledge regarding the process that describes the

---

25 The value and , both centered at 0.5, implying average information update every two quarters.

26 As noted by Keen (2007), this is not a completely innocuous assumption, since the hump-shaped behavior of inflation in Mankiw and Reis (2002) disappears if price-setting decisions

forcing variables, so we impose a beta distribution which is centered at 0.5 for the AR coefficients to guarantee the stationary shock process. An inverse gamma distribution is used to restrict the volatility of shock to guarantee its positive value with the mean of 0.33 for the demand shock, 0.33 is assigned cost-push shock, and assign 0.25 to policy shock respectively (Milani and Rajbhandari, 2012). The same strategy is applied for the standard deviation of the revision shocks in imperfect information data revision model with the mean value 0.25 and relative higher volatility 4 to capture uncertainties. We assign 1 to the mean value of the intertemporal elasticity of substitution as the implication of log utility in consumption (Gali and Gertler, 2002; Gali et al., 2003; Meyer-Gohde, 2010), while we set wide standard deviation of as 0.5 in order to restrict the fluctuation in a reasonable range based on previous studies.

Concerning the priors of Taylor rule being borrowed from the previous common selection (Smets and Wouters 2003, 2007; Meyer-Gohde, 2010), assigning 1.5 as mean value to the reaction to the inflation, the 0.25 as its standard deviation, and follows normal distribution. At the same time, the same distribution is applied to restrict the reaction to output gap, yet with the different mean value 0.12.and different standard deviation 0.05 respectively. The lagged interest rate its coefficient, is also restricted by the same distribution, but assign 0.75 to its mean value and 0.1 to its standard error respectively to describe the persistent property of the policy rule. The specifications of priors[27] and the estimated mean values of posterior of the rival models' parameters as well as shock processes are well presented in Appendix A of Chapter 3.

---

are strategic substitutes. The defense of the assumption of strategic complementary in price-setting decision can also see Woodford (2003, chapter 3)

[27] Those specification of priors include distribution types, mean and the standard deviation.

3.5 Way to assess Bayesian Estimation Results

The crucial point of using Bayesian approach to estimate DSGE type model is how to assess the result of Bayesian estimation results. We assess the Bayesian estimation results from checking the estimation diagnosis and results as follows.

Firstly, if the MCMC numerical procedures performs well, the inspection of the estimated parameters' mode as well as standard deviation estimates can be convincible correspondently. Namely, the estimate results of parameters should be satisfied on the perspectives of both statistics and economic theory. To check whether the estimated results are plausible, we compare them with previous research works and evidence from micro data.

Secondly, if the estimated results are regarded as sensible ones then those estimates can be taken as favorable starting values for the Metropolis-Hastings algorithm and, thus, its properties of convergence can be examined as the main source of feedback to hold confidence or may indicate problems of estimation results. To reach convergence, we should take many individual runs, each of which performs sufficient number of draws with a different starting value of Metropolis-Hastings simulations. If convergence is obtained, meanwhile the optimizer is not trapped within strange region over the parameters' subspace, we may get following scenarios: results within each iteration of different runs being similar, or results between different runs being close. If convergence cannot be achieved, the issue can be attributed to insufficient support from priors or a deficiency of Metropolis-Hastings iterations in quantity.

Thirdly, as An and Schorfheide (2007, pp.127) said: 'A direct comparisons of priors and posteriors can often provide valuable insight about the extent to which data provide information about parameters of interest.' To check the simulated posterior distribution is essential and can be taken from following aspects: 1) posterior distributions should be an approximated normal distribution; and 2）the prior and posterior should be neither extremely similar nor extremely dissimilar. To be specific, if they are too different from each other which may indicate that prior gives a poor restriction on the sample data. However, if they are too close, the estimated results may largely be guided by the priors and rarely rely on the selected sample data. If a sufficient tight prior distribution is appointed, the informative posterior distribution can still be achieved even that we cannot identify the estimated parameters by the selected sample data. That is the case when the sufficient tight prior distribution has been set, and the posteriors will show well-behaved due to the fact that prior has been chosen within a specific region of the parameter space. It is definitely that the prior would have been selected to preclude the illogical areas of the parameter space on perspectives of statistics and theory. At the same time, prior should also be chosen wisely and uninformatively within a reasonable range to prevent selected sample data from being silent and drawing deceived conclusion. In another word, the movement from the prior to the posterior can be considered as a sign that there is a tension between priors and selected sample data. If prior distribution and posterior are no different with a given parameter, we can draw a conclusion that the estimated results about estimate parameters largely depends on the prior, while the selected sample data is silent on that estimated parameter. In such case, adjusting both the distribution and the dispersion of the prior may be a useful step to check identification problems and offer a clear answer that whether the selected data strongly supports estimated parameters. Additionally, the estimation of the structural shocks need to be checked concerning its reasonable magnitudes and frequencies of innovations. Finally, the sensitivity check can be made regarding apply different reasonable priors or apply different sample data.

3.6 Estimation Results through Bayesian Method

The three competing models are solved and estimated with the Dynare 4.4.3. The methodology we implied to get posterior distribution is a Metropolis-Hastings algorithm which generates 20,000 draws with the acceptance rate within 20% and 40%. The estimated sample data are selected quarterly from the US starting from 1969Q1 to 2015Q4 (survey of professional forecaster data has been used in robust checking section over the same periods).

3.6.1 Assessment results of Bayesian Estimation

The diagnosis for the sampling algorithm, the Metropolis-Hastings is shown in Appendix A of Chapter 3. The information of three aspects, namely the analyzed mean of parameters (interval), the variance of parameters (m2) and third moment of parameters (m3), are concluded in three graphs, each of which represents convergence measures in detailed. The two distinct lines in the graphs shows the results within chains and between chains respectively. To reach reasonable estimated results, both lines concerning each of the three measured aspects must be steady and convergent to one another. What can be seen from the graphs that overall convergence is approached. In terms of prior densities, we use gamma (G), inverse gamma (IG), beta (B), and normal (N) distributions. The prior and posterior distributions have been presented in Appendix A of chapter 3.

The significant different between prior distributions and posterior distributions are not exist, which can be checked graphically. Also, on the perspective of most parameters, the prior and posterior distributions are not extremely close which implies that the observables provide extra information for most parameters of the estimated models, which indicates that the presumed priors are not the only factor that influence the estimates. The estimated posterior is an approximated normal distribution whose shape consistent with the Bayesian estimation its asymptotic properties.

3.6.2 Summary of Posterior Estimates

The construction of the posterior distribution under Bayesian econometrics can be achieved by combing the prior distribution together with the likelihood function by using Kalman filter mechanism. After accomplishing Kalman recursion and evaluation and maximization got the log likelihood function and log prior density, Chris Sim's csminwel is applied to approach the estimated posterior.[28] Afterwards, the posterior distribution can be achieved through running 20,000 draws by Metropolis-Hastings algorithm with optimal acceptance rate (i.e., between 20% and 40%). From the 20,000 draws, the initial 20% are discarded and the rest are kept to eliminate any dependence of chain from its steady state.

The Table 3-1 gives the estimated posterior distribution of the parameters for each group concerning reduced-form New-Keynesian DSGE model concerning with and without inattentiveness. Incorporating inattentive feature into modelling expectation seriously affects the estimation results of the parameters. For instance, although the estimated intertemporal elasticity of substitution (i.e.,) is lower than prior’s value in all three competing models no matter with or without inattentive feature, it varies significantly. In detail, the estimated of the model without inattentive feature is 0.0225. From another perspective, the values of estimated of the SI model is around five times higher than that of one without an inattentive feature. A relatively higher intertemporal substitution implies that large changes in consumption are not

[28] Chris Sim's csminwel is a minimization routine and carry out to minimize the negative likelihood.

very costly to consumers through Euler equation. On another face, if   is low, the motivation of the consumption smoothness will be very strong, which is caused by the fact that the consumers will be reluctant to save but consuming a lot relative to the former case.

Regarding to imperfect information data revision model, the economic agents involve signal extraction (data revision) process to understand the real state of the economy. Thus, the value of σ is estimated to be 0.0899 which is four times larger than the one estimated in the baseline model. Additionally, the estimated AR coefficients of imperfect information data revision model, especially the AR coefficients of demand shock and cost-push shock, shift to relatively lower value comparing with that of baseline model. In terms of the estimated parameters (i.e., the reaction toward inflation and the reaction toward the output gap) in monetary policy function, the values are estimated to be not very different under the three models of the estimating results.

Most of estimation results presented in Table 3-1 are remarkably consistent with the previous studies. We find that the reaction towards the inflation   is not far away from the presumed prior 1.5 under the three models. The reaction towards the output gap is also not volatile under different expectation assumptions (i.e.,   varies between 0.1848 to 0.1974). Moreover, the estimated result of   shows reasonably high degree of interest-rate smoothness (i.e.,   varies between 0.8801 to 0.9002) under different expectation assumptions as well. However, higher policy coefficients overall and some structural parameters shift a lot (i.e.,   varies between around 0.02 to 0.1). The estimated coefficients of AR processes of shocks which reflect the existence of substantial degree of persistence in the data. The highly persistence performance are captured by the high degree autocorrelation in demand shock   which is estimated above 0.6 in all three models. The autocorrelation in cost-push shock   is estimated around 0.7 of both baseline model and sticky information model. However, regarding imperfect information data revision model, the estimated is quite low (i.e., is estimated to be 0.3657). Moreover, compared to and , the coefficient of monetary policy shock is estimated relatively small, which is around 0.2 to 0.3 regarding three models.

The estimation results illustrated above concerning the estimated posterior mean are not meant to show that one specified model is superior to the other models. By comparing the variation between estimated posterior results under the two different situations (i.e. with and without inattentive feature), we can check the sensitivity of the results. Furthermore, through evaluating the posterior results under the models with two different inattentive expectation assumptions, it is available to check the sensitivity of them. The necessity of checking sensitivity of variation concerning the models with different inattentive expectation assumptions is derived from the case that it is usually ignored by the previous studies.

TABLE 3-1 SUMMARY ESTIMATION RESULTS OF DIFFERENT EXPECTATION FORMATION

Prior distribution				Posterior distributions		
Params.	Distr.	Mean	S.D .	FIRE	SI (j=4)	IF
	G	1	0.5	0.0225	0.1092	0.0899
	B	0.6	0.05	0.7257	0.6340	0.7389
	B	0.75	0.1	0.8834	0.9002	0.8801
	N	1.5	0.25	1.3891	1.3735	1.0884
	N	0.12	0.05	0.1974	0.1848	0.1962
	B	0.5	0.15	0.7995	0.8139	0.6186
	B	0.5	0.15	0.6948	0.6940	0.3657
	B	0.5	0.15	0.3094	0.2986	0.2235
	IG	0.33	1	0.1564	0.5548	0.2710
	IG	0.33	1	0.0878	0.2446	0.1551
	IG	0.25	1	0.2301	0.2294	0.2181
	N	0	2	-	-	1.8500
	N	0	2	-	-	1.1198
	B	0.5	0.2	-	-	0.7252
	B	0.5	0.2	-	-	0.8535
	IG	0.25	4	-	-	0.3270
	IG	0.25	4	-	-	0.0808
	B	0.5	0.2	-	0.3084	-
	B	0.5	0.2	-	0.2362	-

3.6.3 Models Comparison

3.6.3.1 Model Fit

Table 3-2 shows that the marginal likelihoods of three rivals concerning the different expectation assumptions (i.e., with and without inattentive features), along with the corresponding formal criterion the definition of Bayes factor is the simple ratio of marginal likelihoods between any two models where we take the model with full-information rationality as null hypothesis. Geweke's Harmonic mean is applied to calculate the marginal likelihoods of each case[29]. Comparing the values of marginal likelihood is a standard way of Bayesian approach to know which model fit the data best. The model under the conventional assumption without any inattentive feature produces the lowest value of model fit. Maintaining rationality but extending to include inattentive ingredients, the models' performances are improved. Particularly, the model with sticky information expectation achieves the best model fit among the three competing models.

The implementation of the sticky-information model requests a predicting horizon (i.e., truncation point j), however, there is no clear approach to select the value of truncation point j. If the short forecasting horizon, namely small value of j, is supposed to be two or three quarters which are comparably short periods, it would lead to the misperception of the distribution of agents regarding to updating their information relative to the distribution given by theoretical model. On the other hand, a long forecasting horizon will include too much forecast errors, which tend to form bias to reduce the estimated share of updating agents (i.e., and) (Khan and Zhu, 2002).

[29] There two common methods for computing marginal likelihood concerning Bayesian method, one is so-called Laplace approximation, which assumed that posterior kernel can be approximated by a Gaussian distribution and evaluates its integral at the mode and variance obtained with the numerical maximization of the posterior. The second method is so-called Geweke's Harmonic mean estimator uses MH runs to simulate the marginal likelihood, and then simply use the average of these simulate values.

Balancing the reduction of forecast error and the frequency of updating information theoretically, we set j=4 as our starting point[30], and the alternative j=6, and 8 also have been taken into consideration as the choices of robustness check.[31]

Guided by Jeffreys (1961), we have a way to evaluate the preponderance of the evidence in the light of a selective model concerning the model in the null hypothesis to interpret it into the comparable superiority of model. The detail of guidelines is presented in Table 3-3. Basing on his guidelines, the Bayes factors' values in Table 3-2 show that 'decisive' evidence for both models with inattentive expectation assumptions against the baseline model with full-information rational expectation assumption. Moreover, between two models with respect to different inattentive expectation assumptions, we take the imperfect information data revision model as the null hypothesis. Through Bayes factor, it implies that the model with sticky information shows the 'decisive' evidence as a preferable choice (Bayes factor).

However, one obvious limitation of this comparing approach is that by using this method the conclusion of the evaluation of model fit can only be drawn relatively. Thus, the best estimated model would still be deficient (potentially misspecified) in catching the essential dynamic in our selected sample data. The model's performance is assessed in an absolute way of one model against data, the indirect inference has been chosen as the robust check approach to re-examine model's performance, and this will be conducted in Chapter 4.

---

[30] Kiley (2007, p112) compares the sticky prices and sticky information empirically and noted that, 'in practice., the longest information lag is truncated as four quarters.'

[31] Paustian & Pytlarczyk (2006), they have examined the sticky-information with different truncation point j=12, 24 respectively, and they find that the sticky information its model fit is not sensitive regarding increasing the maximum lag for outdated information and almost does not change.

TABLE 3-2 MODEL FIT COMPARISON

Model	Log Marginal Likelihood	Bayes Factor relative to the FIRE
FIRE model	-267.05	
SI model (j=4)	-247.36	
IF model	-254.08	
Note: (1) Sample period: 1969Q1-2015Q4 US macro data; (2) FIRE represent Full-Information Rational Expectation Model; SI represent Sticky Information Expectation Model; IF represents Imperfect Information Data Revision Model.		

TABLE 3-3 JEFFREY'S GUIDELINES FOR INTERPRETING BAYES FACTOR[32]

Bayes Factor	Interpretation
1 to 3.2	Not worth more than a bare mention Evidence
3.2 to 10	Substantial Evidence
10 to 100	Strong Evidence
100	Decisive Evidence

3.6.3.2 Estimated Impulse Response Functions (IRFs)

Our selected models are mostly consistent with a large number of literatures with respect to New-Keynesian three-equation model. This section mainly concerns with the appearance of the distinguishing features when we introduce the inattentive features (i.e., sticky information and imperfect information data revision) into the model. Previous study results regarding the introduction the inattentive ingredients into DSGE model (Mankiw and Reis, 2002, 2007; Collard et al., 2009) find that the monetary policy shock has a tendency to produce more delay impact under inattentive expectation economy relative than that under the economy without any inattentive features. The information costs are one of the interpretations behind this (Mankiw and Reis, 2002, 2007), The information costs are consisted of two aspects, one of which is the monetary costs (e.g., payment need to be made to acquire updated information and receive the professional interpretation from a financial advisor). The timing cost is the other aspect (e.g., time of obtaining, processing, and interpreting updated

[32] The use of Bayes Factor to compare models was first suggested by Jeffrey's (1961), who suggest that the following rule of thumb for interpreting Bayes factor.

information) (Begg and Imperato, 2001; Reis, 2006a, 2006b). Thus, due to the information costs, some of the economic agents will chose to use the already-paid old information, which generates the delay response. The other interpretation is that people sustain noisy disturbances so that they need time to filter useful information through data revision process (Casares and Vázquez, 2016).

Accordingly, our main focus in this section is to check how the embrace of an inattentive feature in the model affects the macroeconomic model. Particularly, the delay impact of a monetary policy shock upon the main macro variables (i.e., the delay effects of inflation and output gap) will be verified. Afterwards, the estimated impulse response function results will be shown graphically to illustrate major macro variables of the positive monetary policy impact under two different inattentive hypotheses as well as the baseline without inattentiveness respectively.

As Figures 3-1 shows, the models with sticky information can produce a persistence and a delay reaction of inflation, which is mostly in line with the suggestion from the previous studies (Mankiw and Reis, 2002). However, neither the model without any inattentive features nor the model with imperfect information can accomplish the goal. The results regarding the model with imperfect information data revision, unexpectedly, are different from the suggestions from previous studies (Collard et al., 2009). Additionally, the estimated impulse response functions generated under the model that assumes households and firms involving data revision issues are quite similar with those generated from the baseline model. Overall, the effect of the positive monetary policy shock gives a raise to the nominal interest rate in three competing models.

We turn to examine the IRFs with respect to the model featuring sticky price under full-information rationality assumption. Basing on Euler Equation, there will be a negative power on the demand of households' consumption which leads to holding off consumption, if the nominal interest rate increases along with the raise of real interest rate. The case exactly complies with our estimated results concerning model with full-information rationality assumption. Since the economic activity is directed by demand, the results of decreasing demand lead to a drop of firms' production. At the same time, deflation is generated by a reduction in economic activity demand. As time passes, the economy recovers, in the light of the Taylor rule, since the reductions in both demand and the inflation rates cause a reduction in the nominal interest rate after early period. The two alternative competing models are quite similar to the baseline model in terms of the IRFs of the positive monetary policy shock to main variables quantitatively. Exclusively to the model with sticky information, the positive impact of monetary policy can produce a persistence and gradual response of inflation.

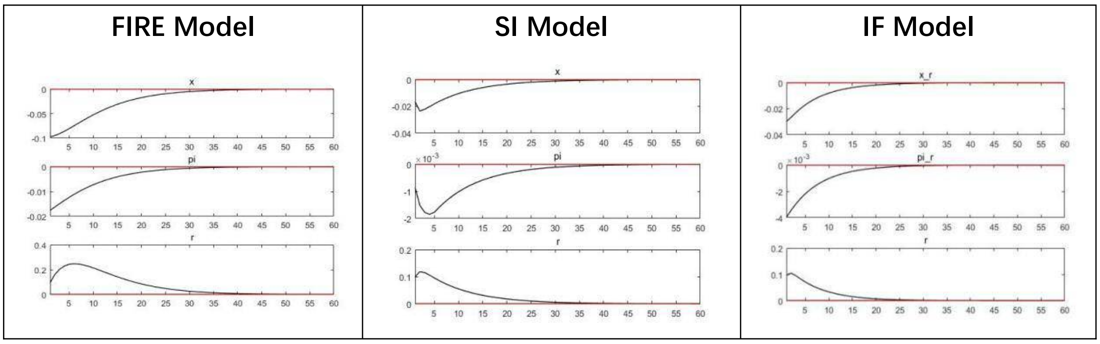


Figure 3-1 Estimated Impulse Response Function of One Unit Positive Policy Shock to Main Variable (x=output gap, pi=inflation, r=interest rate)

After then we turn to examine the effects of the positive demand shock to three main variables under three competing models through estimated impulse response function. The estimated impulse response function has been shown in Figure 3-2. We can see that the positive demand shock, in general, has a relatively long effect on interest rate since this variable converges after around 30 periods. Meanwhile, the demand shock has a relatively significant impact upon the output gap. Two long-run effect converges require 20 periods concerning FIRE model and SI model. However, it only takes 9 periods of convergence under IF model. In general, the demand shock impact inflation positively and converges quickly comparing to the effect on nominal interest rate under the three competing models. Under imperfect information data revision model, people's uncertainty of data revision at initial stage leads to small effects on inflation and output gap. But the turning point appears at the fifth period when people have strong enough confidences on their expectations after reducing the uncertainty. So, inflation and output gap under imperfect information may perform better at bringing about an efficient response and rapid convergence than those under full-information or sticky-information environments.

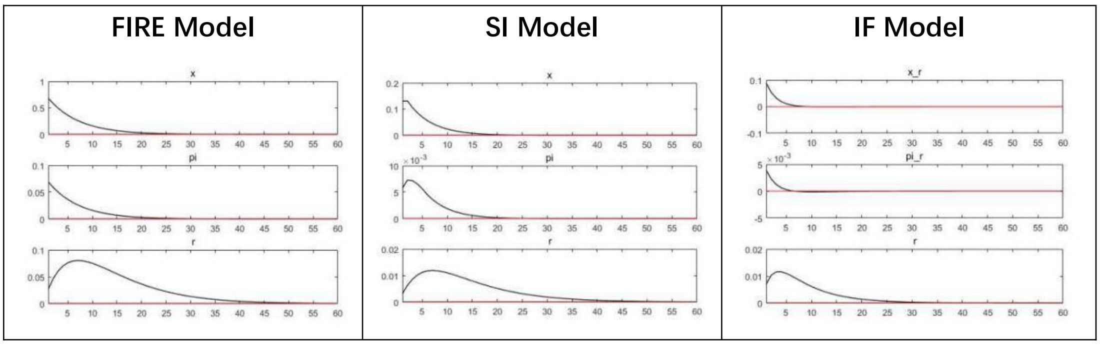


Figure 3-2 Estimated Impulse Response Function of One Unit Positive Demand Shock to Main Variable (x=output gap, pi=inflation, r=interest rate)

The positive cost-push shock impact inflation and interest rate positively regarding three competing models which is presented in Figure 3-3. But the positive cost-push shock leads to different consequences under different models. To be specific, the effect triggered by it under the baseline model is negative, while the those of SI model and IF model are almost null at the initial point on output gap. This distinction may be caused by the fact that people's inattentive behavior to some degree lessen the effect of cost-push shock, which is presented in Figure 3-3. The economic agents under imperfect information assumption environment cannot observe the real state. So, people reduce noise through data revision process and only take actions in reaction to their expected revised data. From the estimated impulse response function, it can be indicated that when people form their expectation through imperfect information data revision, the impact of the supply shock on inflation happens in short term. On the other hand, the models with sticky information will generate more persistence effects on output gap and require relatively longer time for converging. Furthermore, in aggregate level, the variables under the economic agents involving data revision issues converge more quickly than those under the baseline model and the sticky information model.

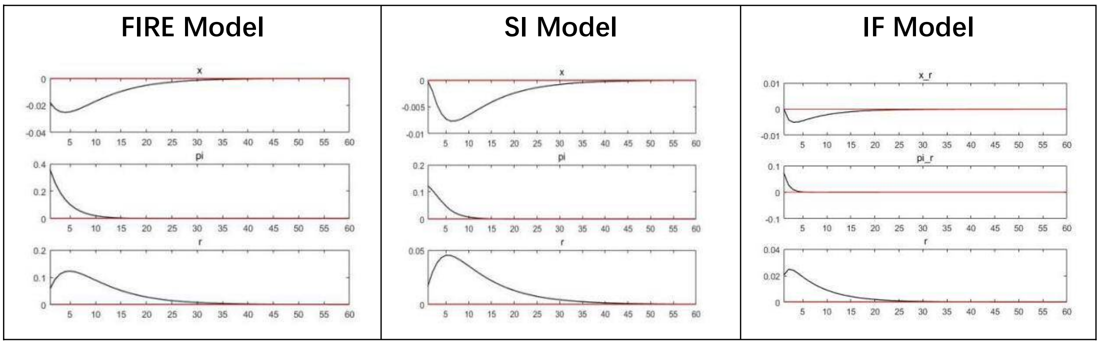


Figure 3-3 Estimated Impulse Response Function of One Unit Positive Cost-Push Shock to Main Variable (x=output gap, pi=inflation, r=interest rate)

To sum up, the model with sticky information expectation has strong abilities of generating more persistence and reproducing delay responses. However, the model with imperfect information data revision expectation cannot attain this goal. However, the fail to reproduce the delay response should not be taken as the reason to judge the model's invalidity. The result may be caused by two key factors missing in our estimated inattentive expectation models. The origins are wage rigidities and the inclusion of capital variable utilization (Christiano et al., 2005). However, even if the model with sticky information can produce a persistence and delay impulse response, we still cannot confirm whether the model itself can indeed be used to explain the real world. So, it is still necessary to conduct indirect inference evaluation to examine the ability of model in an absolute way.

3.7 Robustness Check

3.7.1 Robustness to the Different Prior Distribution

As presented in Table 3-4, in this section we set   to be 0.75 instead of 0.6 after adjusting the mean value of presumed prior of the degree of price stickiness (i.e., α) higher but still being one of the common options applied in many studies (Eichenbaum and Fisher, 2004; Woodford, 2003). It may be worth repeating the analysis with relative flatter prior, namely uninformative prior (i.e., the prior is assumed to follow uniform distribution instead of beta distribution used in starting comparison). The parameter depends on uniform distribution which is assumed within a fixed range of values (i.e., between 0 and 1). The estimated results in Table 3-4 show that the ranking of three competing models is the same as that reported before although different degree of tightness of priors leading different performance in each model. It facilitates us to remove the concerns that our estimation results may seriously be driven by the presumed distribution of the priors and give no chance to let the data speak.

Interestingly, although the models with inattentive feature still are superior to the baseline model in the light of model fit, the distance between sticky information model and imperfect information data revision model is narrowed down, which shows no evidence that the model of sticky information precede model of imperfect information data revision (i.e., when the imperfect information is taken as null hypothesis, the Bayes factor is approximately 1.42). It pushes us to re-examine the model's ability in an absolute way through indirect inference method.

TABLE 3-4 MODEL FIT COMPARISON

Model	Log Marginal Likelihood (Benchmark Priors)	Log Marginal Likelihood (Using Diffuse Prior)
FIRE model (baseline)	-267.05	-261.31
SI model (j=4)	-247.36	-246.75
IF model	-254.08	-247.10

Note: (1) Sample period: 1969Q1-2015Q4 US macro data; (2) FIRE represent Full-Information Rational Expectation Model; SI represent Sticky Information Expectation Model; IF represents Imperfect Information Data Revision Model.

3.7.2 Robustness to the Different Truncation Point j of sticky information model

The empirical performance concerning sticky information model requires a forecasting horizon (i.e., truncation point j) to be taken into consideration. We set j=4 as our starting point, at the same time, we set relatively longer forecasting horizons j=6 and 8 respectively.[33] The estimation result shows that the truncation point has a so small influence on model fit of SI model that will not disturb the original rank.

TABLE 3-5 SENSITIVITY CHECK OF STICKY INFORMATION MODEL[34]

Model	Log Marginal Likelihood (benchmark priors)
SI model (j=4)	-247.36
SI model (j=6)	-247.27
SI model (j=8)	-247.11

3.7.3 Robustness to the specification of Taylor rule

Concerning that different specifications of the monetary policy rule may influence our estimation results, we re-estimate each model with two other specifications of Taylor rule (Smets and Wouter, 2003, 2007; Woodford, 2003). To be specific, one is the 'more complex Taylor rule' which includes the change of output gap and the change of inflation in monetary authority reaction function whose parameters are represented

---

33 Khan & Zhu (2002). Estimates of the sticky-information Philips curve for the united states, Canada, and the United Kingdom. Bank of Canada; Paustian & Pytlarczyk (2006), examine the sticky-information with different truncation point j=12, 24 respectively, and they find that increasing the maximum lag for outdated information sets from j=12 to j=24 the fit of sticky information almost does not change

34 Since the computation time grows rapidly as j increased, when faced such computational burdens, the attractive choice of truncating may just include a few lagged expectations. Estimating sticky information model with a higher but fixed j might be fairly accurate for some combined parameter, however the only combined is and somehow have been fixed as suggestion from previous studies; and too high j will also unnecessary burden the computations, such that I only consider j up to 8 and take j=4 as starter join model competition.

as and . We set the mean values and standard deviations equal to 0.12 and 0,05 respectively for both parameters and . The settings are in line with the previous studies (Smets and Wouter, 2003,2007) and enable the priors to follow the normal distribution. The other one is the ‘less complex Taylor rule’ (Woodford, 2003), which has been used in robust check as well and been suggested as a good description without interest rate smooth of the Fed's monetary policy between 1987 to 1992. Moreover, in this case: and have been asserted as good approximations to characterize the US policy (Woodford, 2003). Both alternative specifications have been presented in Table 3-6.

TABLE 3-6 ALTERNATIVE SPECIFICATION OF TAYLOR RULE

More complex Taylor Rule (e.g., Smets and Wouter, 2003, 2007)
Less complex Taylor Rule (e.g., Woodford, 2003)

The estimation results have been checked in Table 3-7, through which we can see that after introducing ‘less complex Taylor rule’ into the three-equation New-Keynesian framework, each of three competing models gains a worse model performance, which can be checked through the log marginal likelihood. But these results may not be surprising since it is too simple to closely match the optimal policy in the context of an economic model. But, the ranking among three competing models is fixed even though ‘less complex Taylor rule’ is introduced. But, on the contrary, while we are using the ‘more complex Taylor rule’ (Smets and Wouter, 2003, 2007), the performances of all the three models are improved.

In general, we can draw two conclusions under the situations regardless which specification of Taylor rule is adopted. The first is that the model with inattentive feature outperforms the baseline model without any inattentive feature. The second is that the ranking among three is identical to the previous results.[35]

TABLE 3-7 MODEL FIT COMPARISON

Taylor Rule / Model	Log Marginal Likelihood (benchmark Taylor Rule)	Log Marginal Likelihood (more complex Taylor rule)	Log Marginal Likelihood (less complex Taylor rule)
FIRE model (baseline)	-267.05	-260.47	-344.33
SI model (j=4)	-247.36	-238.24.	-256.46
IF model	-254.08	-250.80	-310.11

3.7.4 Robustness to alternative data resource: survey of professional forecaster data of output gap and inflation

To make our research more rigorous, the survey of professional forecaster data is chosen by us as a different type of data resource in robust check. Since this kind of data reflects the views of a few of the highly informed economic agents. The data is regarded as a standard so conservative that is available to assesses potential deviation from full-information rational expectations. As Ormeño and Molnár (2015) assert, survey data of inflation contributes to the way of modelling private expectations by providing useful information that macro data do not have. In this section, we extend to examine each model by using a different type of sample data (i.e., survey data). The estimation results obtained by using survey data is summarized in Table 3-8. The estimation results regarding the imperfect information data revision model performs best among three competing models. The gap of log marginal likelihoods of the model with imperfect information data revision model and

35 Of course, there are various monetary policy rule suggested in the previous studies, here we just choose two to do robustness check, the further research may necessary to consider more different monetary policy rules detailed and carefully.

that with full-information rationality is 19.64, which can be interpreted as Bayes factor (when we take the baseline model as null hypothesis). Similarly, the gap of log marginal likelihoods of the model with imperfect information data revision model and that with sticky information (j=4) is 6.68, which can be interpreted as Bayes factor (when we take the model with sticky information as null hypothesis). Regardless of different types of data resource in the estimation process, the gaps of log marginal likelihood of the model with imperfect information data revision and that with sticky information are quite similar (i.e., the gap is around 6.68 when estimated using survey data; the gap is 6.72 when without survey observations).

TABLE 3- 8 MODEL FIT COMPARISON (WITH SURVEY DATA)

Model	(1)	(2)	(1) - (2)
FIRE model (baseline)	-36.08	-267.05	230.97
SI model (j=4)	-23.12	-247.36	224.24
IF model	-16.44	-254.08	237.64
Note: (1) is the marginal likelihood estimated with survey data; (2) is the marginal likelihood estimated with US real-time data.			

Furthermore, when the survey data are introduced as observables, the performance of each model improves a lot. The number of log marginal likelihood increases a lot in three competing models, which demonstrates that there is an additional information in survey data to lift the performance of each model. However, Whatever type of resource we using to peruse the estimation result, the model with inattentive expectation are always superior to the baseline model in terms of model fit. However, under the same premise, the ranking of sticky information model and imperfect information is switched, which may because the extra information contained in survey data is in favor of model with imperfect information data revision. In terms of the three competing models in different types of data, we compare the estimation results with survey data (presented in Table 3-9) to the results with real-time data (presented in Table 3-1). It shows that most estimated values of the common parameters do not have significant difference. However, some differences exist. For example, the AR coefficients of cost-push shock and monetary policy shock are higher than those presented in Table 3-1. Besides, the estimated share of updating consumers is much lower than that estimated by using real-time data. While the estimated share of updating firms is relatively larger than that estimated by using real-time data

TABLE 3-9 SUMMARY ESTIMATION RESULTS OF DIFFERENT EXPECTATION FORMATION (WITH SURVEY DATA)[36]

Prior distribution				Posterior distributions (mean)		
Params.	Distr.	Mean	S.D	FIRE	SI (j=4)	IF
	G	1	0.5	0.0159	0.1344	0.0371
	B	0.6	0.05	0.6519	0.6277	0.6543
	B	0.75	0.1	0.8857	0.9164	0.9219
	N	1.5	0.25	1.4669	1.4146	1.3836
	N	0.12	0.05	0.1236	0.1214	0.1243
	B	0.5	0.15	0.5681	0.5983	0.4922
	B	0.5	0.15	0.6928	0.7033	0.4483
	B	0.5	0.15	0.3473	0.3234	0.3110
	IG	0.33	1	0.1158	0.2487	0.2446
	IG	0.33	1	0.0759	0.2106	0.1552
	IG	0.25	1	0.2384	0.2367	0.2414
	N	0	2	-	-	1.9627
	N	0	2	-	-	1.5134
	B	0.5	0.2	-	-	0.5612
	B	0.5	0.2	-	-	0.7457
	IG	0.25	4	-	-	0.2190
	IG	0.25	4	-	-	0.1132
	B	0.5	0.2	-	0.4474	-
	B	0.5	0.2	-	0.0916	-
Log marginal likelihood				-36.08	-23.11	-16.44
Bayes Factor relative to the FIRE				1		

It is noteworthy that survey data has been used to identify expectation mechanisms in recent studies. For instance, Carroll (2003) finds that the public's prediction is lags behind the prediction of professionals' through adopting survey of inflation

[36] The posterior estimated value of is quite different from prior mean which may due to the selected prior is suitable for final revised data but not suitable for real-time data or SPF data. The results of robustness check with revised data is given in Appendix D to Chapter 3.

expectation data. Being distinguished from the previous literature, Easaw and Golinelli (2010) investigate whether different agents or groups that make up the population have various information absorbing rates. Rather than treating economic agents as homogeneous type agents or groups and through using the UK survey data, and they find that homogeneous agents or group can be distinguished by their information absorbing rate respectively. Easaw and Golinelli (2014) establish a new structure (i.e., people can form their expectation multi-period) but basing on single equation method with the focus of inattentiveness (i.e., sticky information and imperfect information) using survey-based data for the US and UK. A more recent work involves using survey data to examine the model with deviation of full-information rationality, for instance, Del Negro and Eusepi (2011) study whether or not a DSGE model with imperfect information while keep rational expectation assumption can reproduce series of expected inflation that match the survey inflation data. Aruoba and Schorfheide (2011) apply inflation forecasts survey data in their observations as extra information which is able to be employed to indicate the time-varying Fed's Inflation Target. After endogenizing survey expectation in a standard DSGE model, Fuhrer (2017) asserts that most persistent in aggregate data is better due to slow-moving expectations but not habits, indexation or autocorrelated structural shocks.

3.7.5 Robustness to Different Detrend Method

Another problem which has been discussed widely in the DSGE literature is how to detrend real variables, particularly, the methodology to obtain the potential output for constructing the output gap. Most of the studies tend to use the statistical detrending method (e.g., HP filter, band-pass filter etc.). Alternatively, we can use the theory-derived potential output (i.e., the output solved under flexible price assumed economy) to construct output gap.

In this thesis, the HP trend is the approximation of the potential output, which is used to construct the output gap. HP filter is a methodology of statistics that can be used to extract the trend after filtering the actual GDP data as the estimates of potential output. The HP filter is a convenient way to get potential output since it only needs the actual output data. However, HP filter is not impeccable because it does not utilize fully of the information from other economic time series data to direct the estimates of potential output. The absence of economic theory forces it to generate the potential output through a technique instead of a model. As a result, it is not a favored method to model the actual potential output. Another suggested method from the previous literature uses a linear detrending to get the output gap, However, this is not a suitable method, either. Since the potential output has a great chance of being non-linear, which can be proofed by the function derived from the model with flexible price assumption driven by technology shock. Although we have no idea about what a technology shock is, the probability that it is non-linear is very high.[37] However, the aim of this thesis is studying the empirical implications through model comparison. A more detailed study of using different detrending methods to obtain potential output for constructing output gap is surely warranted which can be remained for a future research.

3.8 Conclusion

The previous macroeconomic theory is basing on the assumption which full-information rationality restricts the consumers and households to form their expectations. The conclusions drawn by the empirical studies of macroeconomics also depend on such validity of full-information rationality hypothesis. In this chapter,

[37] In recent years, some contributions have made through using DSGE models to estimate potential and the output gap (Vetlov et al, 2011). These models have more visible micro foundations and are very attractive. Despite that, these are difficult to interpret and still a challenge for policy makers to apply in formulation of policies.

the consequences of the inclusion of inattentive expectations (inattentive features) in a popular small-scale reduced-form New-Keynesian DSGE model have been evaluated. What's more, the econometric features of the model have been shown that are not insensitive to the introduction of inattentive ingredients through Bayesian estimation approach. The sensitive analysis focusing on comparing different inattentive features largely lack in the previous study.

The empirical evidence shown in this chapter implies that, firstly, the results of Bayesian estimation indicate that the modelling of incorporating inattentive feature has significant influence on the capability of the model in fitting macro-economic time series. Secondly, these are essential to be studied more critically in estimated. The limitation worth to mention here is that the model we have chosen to make comparison may be misspecified.[38] The source of misspecification not only due to the application of linear approximation solution but also the truth that DSGE model is an abstract of the real world.

It is necessary to study estimated DSGE type model in a critical way while different inattentive features incorporate it. However, there may be misspecifications as leaks in the models we have chosen during making comparisons.[39] The misspecification can be originated from that the DSGE model is not perfect to copy the 'real model'. Thirdly, among three competing models, model with sticky information expectation wins the best model performance through Bayesian estimation using real-time data. In addition, the results show that the model with an inattentive feature improves the model's ability to explain the real world, which is in line with most consequences from the previous studies (Mankiw and Reis, 2002, 2007; Collard et al., 2009). However, between the two inattentive models, only the model with sticky information expectation can generate persistence and delay response, which has been checked

[38] Gourieroux, Monfort and Renault (1993) fully account for the fact that DSGE models are misspecified.
[39] Due to the nature of approximated linear solution methods, the DSGE model may encounter the loss of some components during the process of solving.

through estimated impulse response functions. Finally, the robustness checks with real-time data for our concerns regarding using different prior distribution (diffuse priors), different specifications of Taylor rule, and different truncation points in sticky information model (=4,6 and 8) draw a conclusion that none can switch the ranking position of three competing models. However, when we use survey data[40] to re-examine each competing model, although the model with inattentive features still outperform the baseline model, the ranking between sticky information and imperfect information data revision model changes. This contradict result may due to different types of sample data containing different information to favor different inattentive expectations.

Practically, there is no absolute best way to select an econometric method to estimate and evaluate one model. The Bayesian estimation approach by introducing a prior has its advantages, meanwhile, the most challenging factor is prior as well since its distribution needs to be determined or limited before carrying out estimation. Also, how to choose priors' distribution before implementing estimation is still a controversial point in recent studies (Fernández-Villaverde, 2010). Besides, it is crucial to note that Bayesian estimation method can only check model's relative ability. Thus, it is still essential to check model's absolute ability in order to make fair comparison, which will be conducted in the following chapter through Indirect Inference approach.

[40] https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/ is where the survey of professional come from.

Appendix A to Chapter 3

Prior Interpretation

TABLE 3A-1 PRIORS MEAN OF PARAMETERS

Common Structural parameter		
	Elasticity of intertemporal substitution	1
	Sticky price degree	0.6
	Strategic complementary	0.15
Common Taylor Rule in three models		
	Degree of partially adjustment in Taylor rule	0.75
	Coefficient of inflation on Taylor rule	1.5
	Coefficient of output gap in Taylor rule	0.12
Common Forcing Variables in three models		
	AR coefficient of demand shock	0.5
	AR coefficient of cost-push shock	0.5
	AR coefficient of policy shock	0.5
	Standard deviation of demand shock	0.33
	Standard deviation of cost-push shock	0.33
	Standard deviation of policy shock	0.25
Note: The priors of parameter are mostly chosen from previous literatures, i.e., Miliani and Rajbhandari (2012), and Smets and Wounter (2003, 2007).		

TABLE 3A-2 PRIORS MEAN OF PARAMETERS

Imperfect Information model		
	output coefficient in output revision process	0
	inflation coefficient in inflation revision process	0
	AR term of shock in final revision process of x	0.5
	AR term of shock in final revision process of	0.5
	SD of measurement error of x	0.25
	SD of measurement error of	0.25
Sticky Information model		
	Share of updating firms (Mankiw and Reis,2007)	0.5
	Share of updating consumer (Mankiw and Reis,2007)	0.5
Note: The priors of parameter for SI model are chosen from previous literatures, from Mankiw and Reis (2007).and for IF model the priors of parameters borrow from Casares and Vazquez (2016).		

Appendix B to Chapter 3

Estimates without Survey Data

TABLE 3B-1 PARAMETERS ESTIMATE OF FULL-INFORMATION RATIONALITY

Prior distribution				Posterior distributions			
Params.	Distr.	Mean	S.D .	Mode	Mean	90% HPDIs/ Bayesian confidence bands	
	G	1	0.5	0.0167	0.0225	0.0051	0.0395
	B	0.6	0.05	0.7285	0.7257	0.6796	0.7733
	B	0.75	0.1	0.8913	0.8834	0.8473	0.9209
	N	1.5	0.25	1.3923	1.3891	1.0148	1.7447
	N	0.12	0.05	0.1940	0.1974	0.1197	0.2769
	B	0.5	0.15	0.8072	0.7995	0.7556	0.8452
	B	0.5	0.15	0.7015	0.6948	0.6352	0.7530
	B	0.5	0.15	0.2958	0.3094	0.1974	0.4257
	IG	0.33	1	0.1473	0.1564	0.1183	0.1943
	IG	0.33	1	0.0849	0.0878	0.0693	0.1049
	IG	0.25	1	0.2271	0.2301	0.2102	0.2494
Log marginal likelihood				-267.05			

Note:

- Metropolis-Hastings algorithm is applied to solve posterior distributions. 20000 draws with acceptance rate between 20% and 40%. and we discard the initial 20% of MH draw and keep 16000 draws.
- For the prior densities, we used beta (B), gamma (G), normal (N), and inverse gamma (IG) distributions.

Figure 3B-1 Full-Information Rational Expectation Multivariate MH Convergence Diagnosis

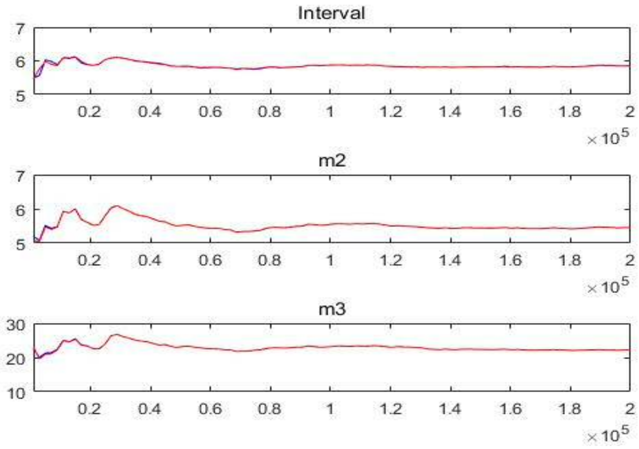
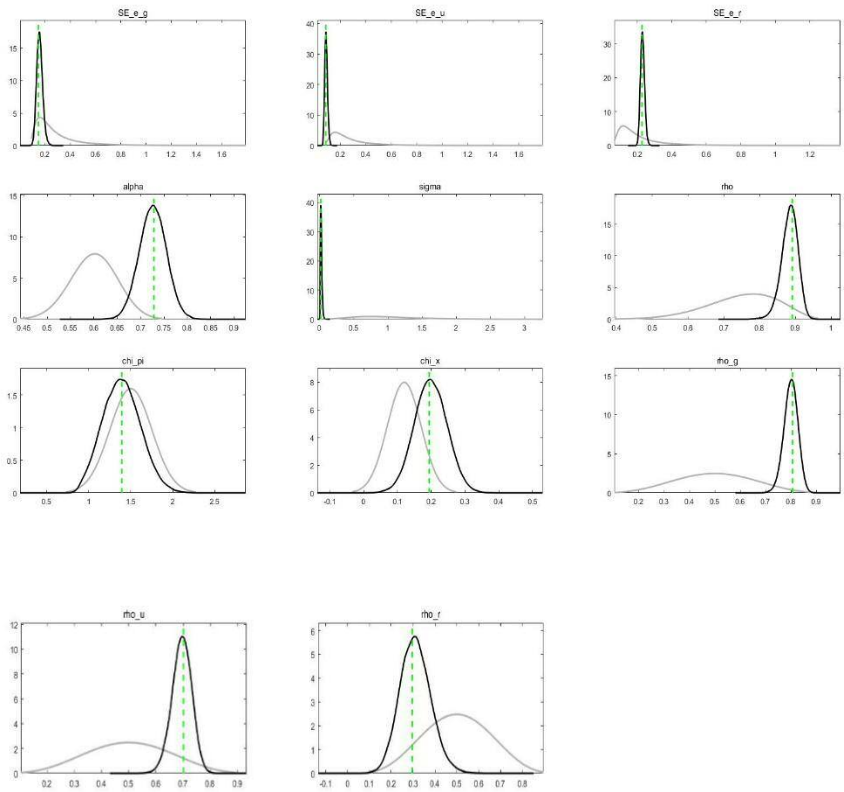


Figure 3B-2 Estimated Parameters Distribution of Full-Information Rationality



(Note: Black line: posterior distribution; green line: posterior mean)

Figure 3B-3 Full-Information Rational Expectation Smoothed Variables[41]

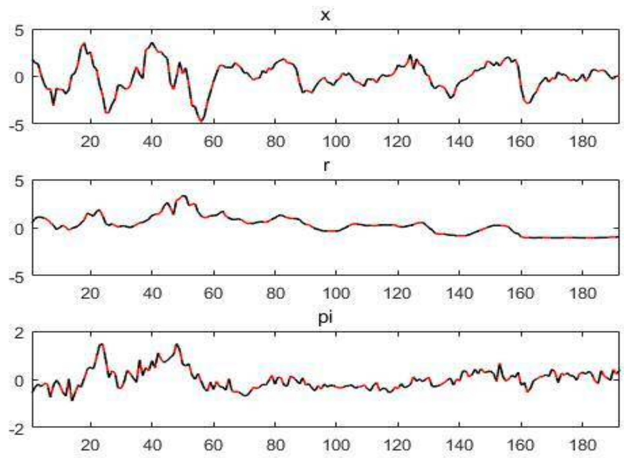
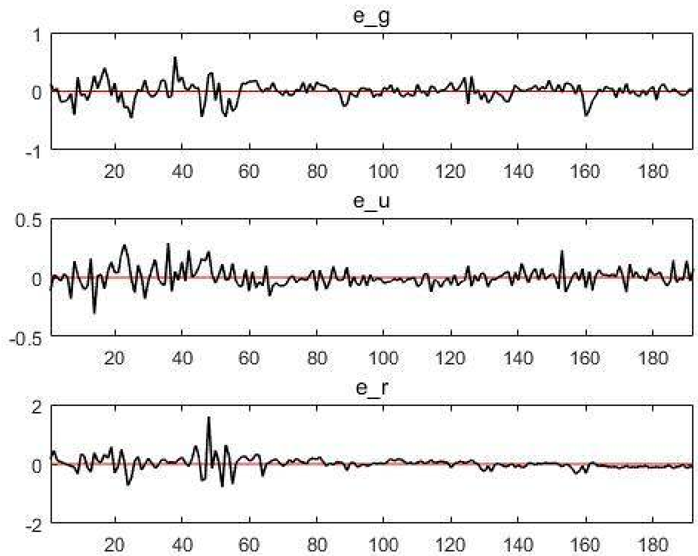


Figure 3B-4 Full-Information Rational Expectation Smoothed Shocks



[41] Dotted black line depicts the actually observed data, while the red line depicts the estimate of the smoothed variables ('best guess for the observed variables given observations') derived from Kalman smother at the posterior mode or posterior mean.

TABLE 3B-2 PARAMETERS ESTIMATE OF STICKY INFORMATION (J=4)

Prior distribution				Posterior distributions of SI (j=4)			
Params.	Distr.	Mean	S.D	Mode	Mean	90% HPDIs/ Bayesian confidence bands	
	G	1	0.5	0.0817	0.1092	0.0245	0.1894
	B	0.6	0.05	0.6314	0.6340	0.5685	0.6991
	B	0.75	0.1	0.9046	0.9002	0.8629	0.9372
	N	1.5	0.25	1.3863	1.3735	0.9735	1.7517
	N	0.12	0.05	0.1847	0.1848	0.1063	0.2646
	B	0.5	0.15	0.8101	0.8139	0.7558	0.8755
	B	0.5	0.15	0.7047	0.6940	0.6092	0.7785
	B	0.5	0.15	0.2848	0.2986	0.1891	0.4109
	IG	0.33	1	0.5252	0.5548	0.4500	0.6594
	IG	0.33	1	0.2455	0.2446	0.2159	0.2726
	IG	0.25	1	0.2265	0.2294	0.2105	0.2490
	B	0.5	0.25	0.1014	0.3084	0.0083	0.9264
	B	0.5	0.25	0.2612	0.2362	0.1257	0.3478
Log marginal likelihood				-247.36			

Note:

- Metropolis-Hastings algorithm is applied to solve posterior distributions. 20000 draws with acceptance rate between 20% and 40%. and we discard the initial 20% of MH draw and keep 16000 draws.
- For the prior densities, we used beta (B), gamma (G), normal (N), and inverse gamma (IG) distributions.

Figure 3B-5 Sticky Information (j=4) Multivariate MH Convergence Diagnosis

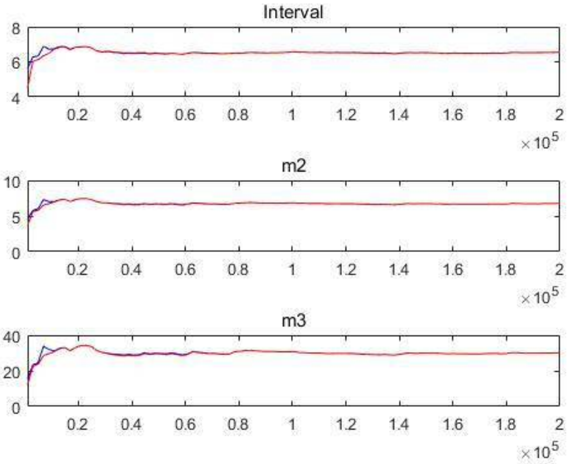
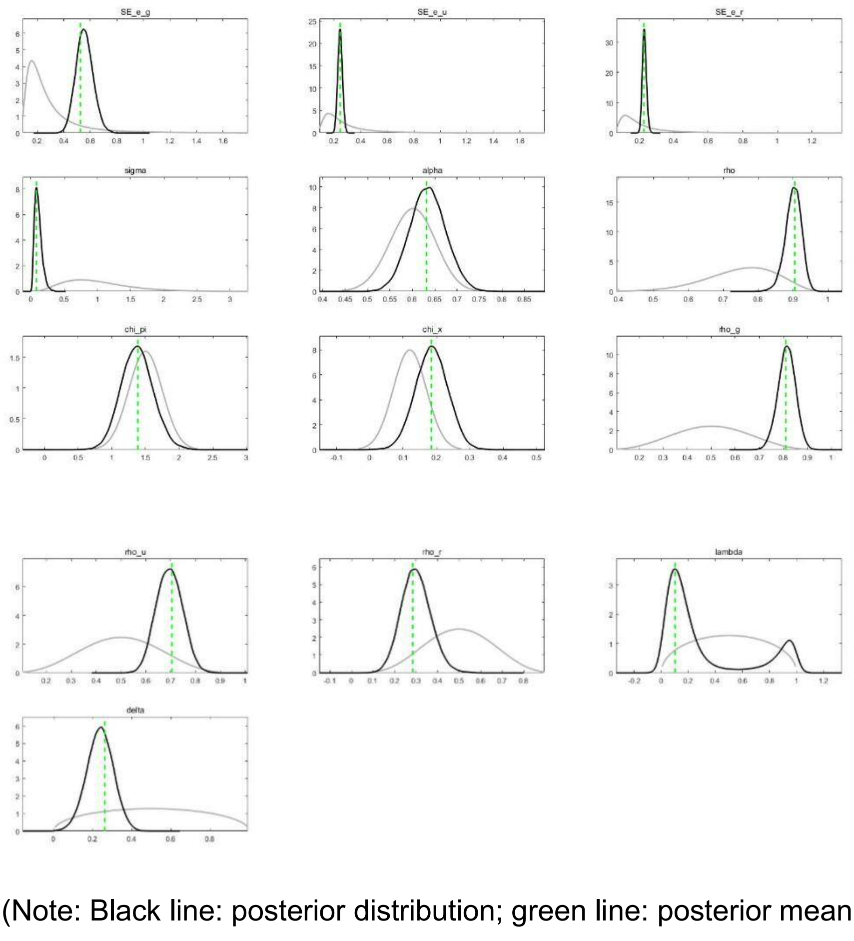


Figure 3B-6 Estimated Parameters Distribution of Sticky Information (j=4)



(Note: Black line: posterior distribution; green line: posterior mean)

Figure 3B-7 Sticky Information (j=4) Smoothed Variables

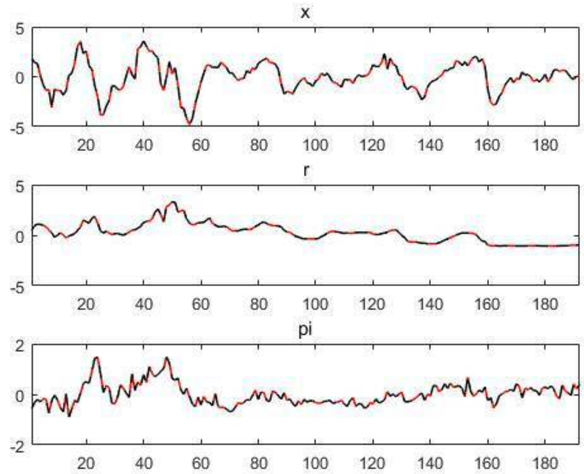


Figure 3B-8 Sticky Information (j=4) Smoothed Shocks

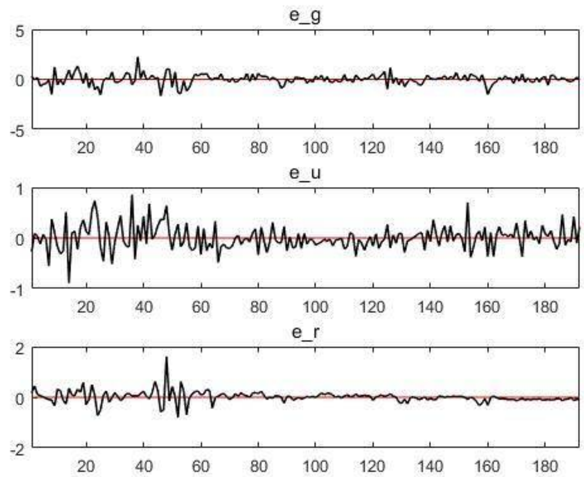


TABLE 3B-3 PARAMETERS ESTIMATE OF IMPERFECT INFORMATION DATA REVISION

Prior distribution				Posterior distribution			
Params.	Distr.	Mean	S.D	Mode	Mean	90% HPDIs/ Bayesian confidence bands	
	G	1	0.5	0.0535	0.0899	0.0152	0.1636
	B	0.6	0.05	0.7298	0.7389	0.6831	0.7960
	B	0.75	0.1	0.8698	0.8801	0.8420	0.9178
	N	1.5	0.25	1.0182	1.0884	0.8409	1.3467
	N	0.12	0.05	0.2031	0.1962	0.1311	0.2608
	N	0	2	1.2492	1.8500	0.5268	3.0401
	N	0	2	0.9908	1.1198	0.5698	1.6884
	B	0.5	0.2	0.8310	0.7252	0.4802	0.8929
	B	0.5	0.2	0.8585	0.8535	0.8118	0.8923
	B	0.5	0.15	0.4986	0.6186	0.3313	0.8647
	B	0.5	0.15	0.3678	0.3657	0.1720	0.5549
	B	0.5	0.15	0.2172	0.2235	0.1209	0.3205
	IG	0.33	1	0.1426	0.2710	0.0891	0.4923
	IG	0.33	1	0.1447	0.1551	0.0827	0.2212
	IG	0.25	1	0.2147	0.2181	0.2001	0.2363
	IG	0.25	4	0.2912	0.3270	0.0633	0.5879
	IG	0.25	4	0.0726	0.0808	0.0513	0.1089
Log marginal likelihood				-254.08			

Note:

- Metropolis-Hastings algorithm is applied to solve posterior distributions. 20000 draws with acceptance rate between 20% and 40%. and we discard the initial 20% of MH draw and keep 16000 draws.
- For the prior densities, we used beta (B), gamma (G), normal (N), and inverse gamma (IG) distributions.

Figure 3B-9 Imperfect Information Multivariate MH Convergence Diagnosis s=3

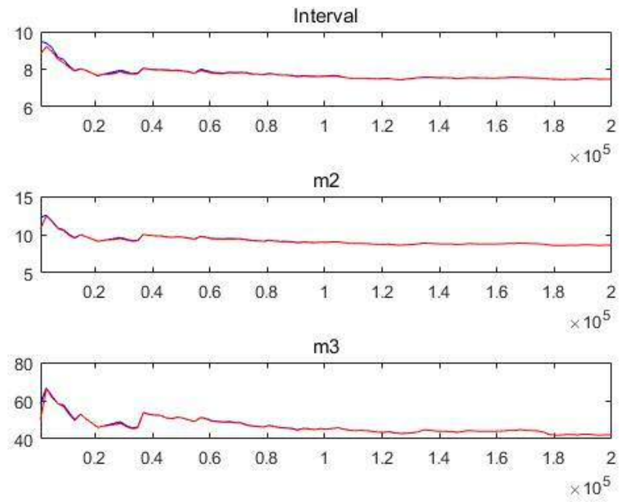
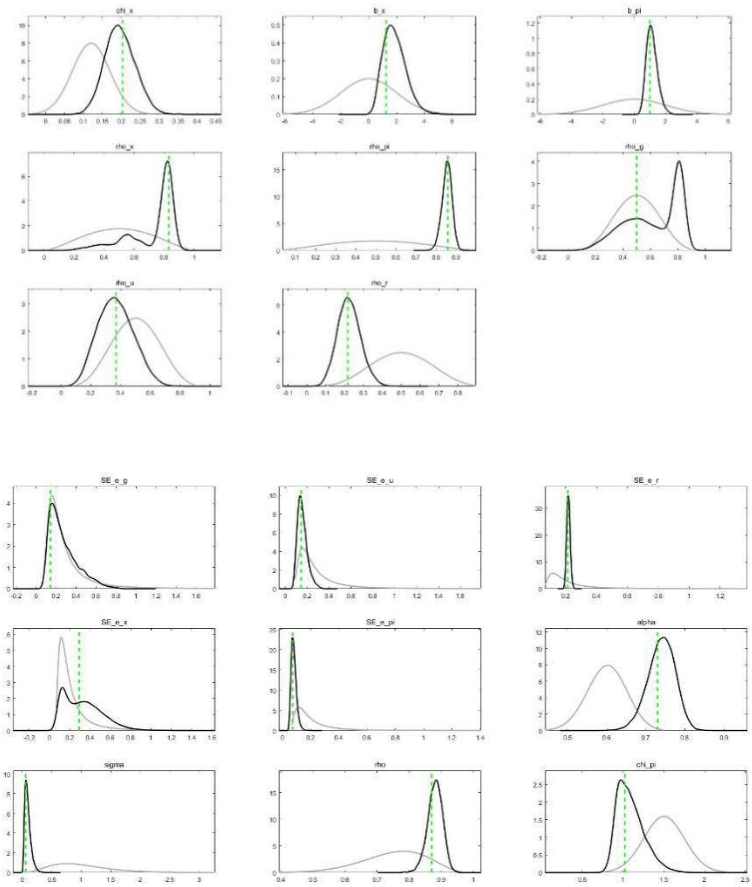


Figure 3B-10 Estimated Parameters Distribution of Imperfect Information Model s=3



(Note: Black line: posterior distribution; green line: posterior mean)

Figure 3B-11 Imperfect Information Data Revision Smoothed Variables s=3

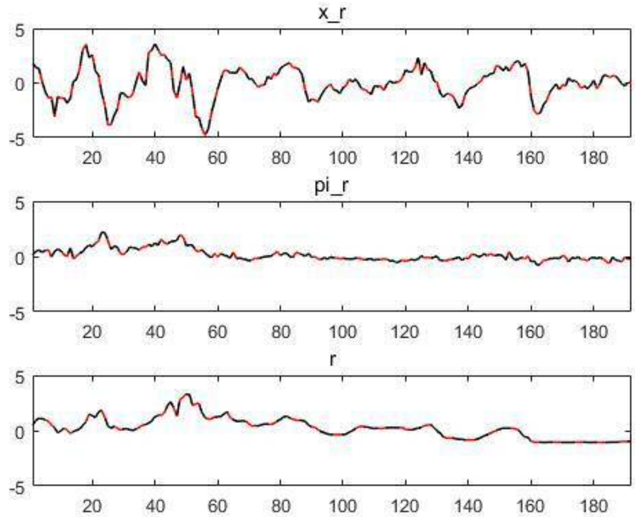
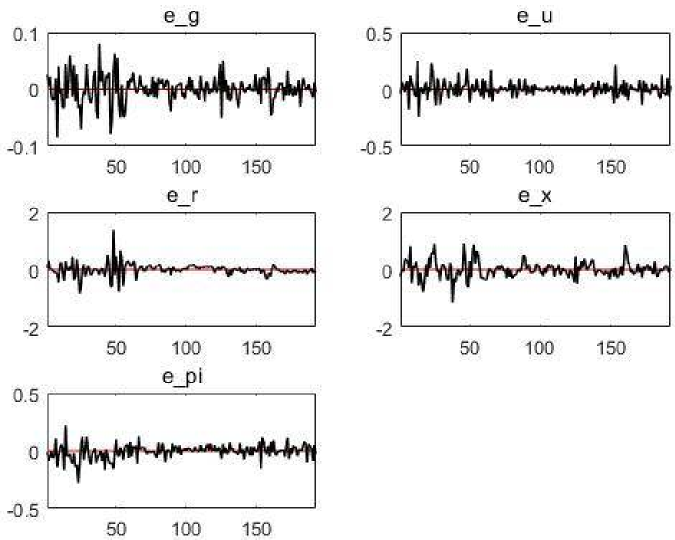


Figure 3B-12 Imperfect Information Data Revision Smoothed Shocks s=3



Appendix C to Chapter 3

All data are of a quarterly frequency and are seasonally adjusted. And all the series are demeaned before estimation.

United States Data Source

1) Effective Federal Funds Rate=FEDFUNDS, the federal funds rate is divided it by four to express it in quarterly rates. The observable is matched to the variable ,

   where

2) The real-time data[42] from Real-time data set for macroeconomists are collected from Federal Reserve Bank of Philadelphia, the real-time Real GDP=ROUTPUT initial released in 2016Q1 (i.e., which only release real-time Real GDP up to time 2015Q4), then the quarterly real-time GDP is the deviation of the natural logarithm of total real-time GDP, potential output is from its HP filter. For imperfect information model to construct the revised observables corresponding to output gap up to time 2015Q4, the real-time data released after one period (2016Q1) as well as the real-time data of GDP released after three periods also applied (2016Q3).

3) For the real-time Implicit Price Deflator=P. Index level initial released in 2016Q1 (i.e., which only release real-time Implicit Price Deflator up to 2015Q4), seasonally adjusted, also from the real-time data set from Federal Reserve Bank of Philadelphia, the series is demeaned. The real-time inflation .Similarly, to construct the revised observables correspond to

[42] https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files is where the real-time data set from.

inflation up to time 2015Q4, the real-time data of Implicit Price Deflator released after one period and the data released after three periods also be used.

4) The survey data using in robust check section is the median of Survey of Professional Forecaster one quarter ahead forecast of GDP deflator and real GDP. In imperfect information data revision model, both one-quarter ahead and four-quarter ahead forecast has been used to construct the final revised observables.

Appendix D to Chapter 3

TABLE 3D-1 SUMMARY ESTIMATION RESULTS OF DIFFERENT EXPECTATION FORMATION (WITH FRED REVISED DATA)

Prior distribution				Posterior distributions (mean)		
Params.	Distr.	Mean	S.D	FIRE	SI (j=4)	IF
	G	1	0.5	0.0471	0.1350	0.1380
	B	0.6	0.05	0.7151	0.6238	0.7283
	B	0.75	0.1	0.8136	0.8256	0.8646
	N	1.5	0.25	1.1768	1.1392	1.1293
	N	0.12	0.05	0.2178	0.2081	0.1921
	B	0.5	0.15	0.8021	0.8114	0.5284
	B	0.5	0.15	0.6873	0.6850	0.4118
	B	0.5	0.15	0.2724	0.2705	0.2268
	IG	0.33	1	0.1596	0.5316	0.2013
	IG	0.33	1	0.0898	0.2393	0.1447
	IG	0.25	1	0.2214	0.2209	0.2205
	N	0	2	-	-	1.8349
	N	0	2	-	-	0.8741
	B	0.5	0.2	-	-	0.8157
	B	0.5	0.2	-	-	0.7099
	IG	0.25	4	-	-	0.4205
	IG	0.25	4	-	-	0.1192
	B	0.5	0.2	-	0.1376	-
	B	0.5	0.2	-	0.2642	-
Log marginal likelihood				-249.13	-232.39	-245.01
Bayes Factor relative to the FIRE				1		

Chapter 4
Testing and Estimating New-Keynesian Type Models with Inattentive Feature through Indirect Inference Approach

4.1 Introduction

This Chapter proposes an approach to evaluate reduced-form New-Keynesian DSGE models which are basing on indirect inference and this method is employed to the previous chapter's three competing models concerning the cases of attentive expectation (i.e., the model with full-information rational expectation) and inattentive expectation (i.e., the sticky information expectation model and the imperfect information data revision expectation model). The approach commonly used by pervious economists to solve the problem which has been existing for a long time to asses a calibrated and estimated DSGE model is simply comparing the features of simulated data and those of true data, in which the sample data are stimulated by calibrated and estimated DSGE model. In this chapter, we choose the indirect inference which can be divided into two stages as a more rigorous approach to evaluate DSGE models.

In the first stage, we will implement indirect inference as a calibrated-based testing method to test each competing model by given initial presumptive parameters. The content of the test can be understood as that through comparing the unrestricted VAR estimates (derived from the simulation data) with the alternative unrestricted VAR estimates (derived from the actual data), we can confirm whether these two groups of parameters' estimates of the auxiliary model are 'close enough' (i.e., each competing DSGE model is correctly specified). If the result shows one model is correctly specified, then the distance of the unrestricted VAR estimates and the alternative unrestricted VAR estimates should be minimized. In other words, the assumed model and the 'real model' will not be far away from to each other. The apparent strength of the indirect inference test method is that it is unnecessary to specify each competing model as the alternative hypothesis. However, we need to identify the auxiliary VAR what is generated by each competing model.

In the second stage, we will implement the estimation-based indirect inference test. In this stage, the Indirect Inference is not just used to gauge the 'distance' between the theory and the reality through using the auxiliary model but also finding a set of parameters to minimize such distance. The extra searching step the distinction between this stage and the last stage.

In this chapter, we have three main purposes. The first purpose is to take indirect inference as a calibration-based test approach. We intent to evaluate the already estimated model (in the previous chapter through Bayesian estimation approach). The focus of our test is to detect whether the data which are simulated from the three competing structural models can explain the actual data. The evaluation of three rival models are done through an indirect inference test which is basing on comparing Wald statistics that concentrate on the total capability of the model to fit the overall dynamic behavior of the actual data.

The second purpose, being distinguished from the first purpose, is to use the indirect inference as the estimation-based test approach whose duty includes exploring the optimal set of structural parameters which enables the model to copy the trajectory of the behavior of actual data to the maximum extent. In the second stage, indirect inference testing process will introduce the optimal searching procedure. To be more specific, the nature of being fixed of the model parameters is an overly strong condition for testing and contradistinguishing models. Seeing the parameter values of the candidate model could be estimated or calibrated within a permissible scope throughout the theoretical structure of the model, it is probable for a rejected model with the presumptive set of parameters to pass the test when it with another set of parameters. To have a fair result of the testing, it is necessary for investigators to find a set of 'good' structural parameters. Thus, we estimate the models to get the optimal sets of parameters before the evaluating process. The third purpose is to reach absolute performances of the models for comparing, which can be realized through the introduction of the distributions of the two groups of estimated parameters of the auxiliary models. From this point of view, it is the most significant difference comparing to the evaluation of Bayesian estimation.

To sum up, despite the conclusion of the ‘best’ model found through Bayesian estimation approach, it is necessary for us to verify the result through other method to give credit of the ‘best’ model.

4.2 Description of Indirect Inference as Evaluation Method

In this chapter, Indirect Inference is applied for measuring how close the three models are to real world. The principle of this method is basing on the idea that through comparing the moments of simulated data and actual data, a model can be measured in an absolute way in a framework that contains an auxiliary model. Two characteristics of this method make it superior to other solutions. Firstly, a statistical threshold given for filtering models divides the tested models into two groups of qualified and unqualified. Secondly, it enables us to evaluate the distance statistically in the middle of the theoretical models (simulated data) and the real world (actual data).

The approach of Indirect Inference already applied diffusely in the field of estimation by scholars (Gregory and Smith, 1991, 1993; Gallant and Tauchen, 1996; Keane and Smith, 2003; Minford, Theodoridis et al., 2009). For instance, in the year of 2011, Le et al. applied the same method to evaluate the model of the US economy which was constructed by Smets and Wouter (2007) and ultimately obtained a rejected consequence on the testing. In this thesis, our evaluation will take the common procedure of indirect inference evaluation for reference from previous studies (Le et al., 2011, 2016; Minford and Ou, 2013; Liu and Minford, 2014; Minford et al., 2015)

It is worth noting that there are two most relevant papers regarding to our research topic through using indirect inference method. One is published by Vázquez et al. (2010, 2012) who assess the importance of data revisions on the estimated monetary policy rule. The estimation conducted through indirect inference finds that the ignorance of the data revision process may not result in a serious drawback in analyzing monetary policy based on New-Keynesian framework. Our assumption substitutes the subjects who involve imperfect information data revision issue with households and firms instead of monetary authority. Meanwhile the subjects can perfectly observe monetary policy. The other related paper is published by Knotek and Edward (2010) who investigates a single equation model incorporating both sticky price and sticky information and detect that such a model can match the real world in both dimensions of micro and macro after estimating it through indirect inference.[43] However, we are more interested in full-structural model rather than single equation model.

The complete estimation of three competing models through Bayesian approach has been done in Chapter 3. With the consequence that competing models with inattentive features are preferable, in this chapter, we will turn to re-evaluate each model focusing on its overall dynamic properties in connecting with the actual data by adopting Indirect Inference as the new evaluation method.

While we are applying Indirect Inference to evaluate an existing structural model, two factors are inevitable in the process of stimulating the data from theoretical model. One is the parameters of theoretical model and the other one is the distribution of the errors. We evaluate the theoretical model through Indirect Inference test which is based on the contradistinction of the actual data with the simulated data obtained from the theoretical model with the assistance of auxiliary model. In this chapter, VAR (i.e. Vector auto-regression), which is a stochastic process model used to capture the

[43] Knotek and Edward (2010) finds that when the empirical Phillips curve is embodied with sticky prices and sticky information, its ability tends to be improved to match the macro data.

linear interdependencies among multiple time series. is selected as the auxiliary model.

There are two reasons for us to choose VAR as the auxiliary model. Firstly, the structural model can always be manifested as a restricted VARMA (i.e. Vector Auto-regression Moving-Average), which is close to a VAR representation. Secondly, VAR can reflect two properties of the data. They are the relation of variance-covariance among the variables through the co-variance matrix of the VAR disturbances, and the dynamic behaviour of the data via the dynamics and the impulse response functions of the VAR. The Wald statistic, which is derived by the distributions of these functions of the parameters of VAR, and TM distance (normalized t-statistics), which are derived from a function of these parameters can be regarded as two criteria of the testing model to measure the distance to the reality. From the consequence of the testing model regarding the two criteria, we can judge whether the hypothesis, which assumes the testing model is correctly specified, is accepted or rejected. If the consequence shows rejected, it implies that the theoretical model cannot reproduce the actual data significantly. While the consequence of being non-rejected implies the data generated from the theoretical model not different from the actual observed data significantly.

Wald Test Statistics

In general, the Wald testing process can be summarized into three general steps as follows. Firstly, to derive the structural errors by using the observed actual data and parameters calibrated or estimated in the model. There are two ways to construct the errors under two different circumstances. When the structural model possesses no expectation terms, the structural errors can be backed up straight from the structural equations and the actual data. While under the situation that structural equation includes the computation of expectations, the method used is the robust instrument variables estimation[44]. Therefore, the expected future variables of output gap and inflation are approximated by the fitted values of VAR (1), which are the linear combinations of the lagged three main variables. Secondly, the structural errors are bootstrapped to be employed to produce the pseudo data which are based on candidate theoretical model. After that, an auxiliary VAR model is fitted to each set of pseudo data and the sampling distribution of the coefficients of the auxiliary VAR model are achieved from these estimates of the auxiliary model. Thirdly, the Wald statistic is calculated to judge whether or not the functions of the parameters of the auxiliary VAR model estimated on the actual data lie within the confidence interval implied by this sampling distribution[45]of the coefficients of the auxiliary time series model (Minford et al., 2015; Fan et al., 2016).

The test is through comparing the performance of the overall capacity of the model with the dynamics performance of actual data to determine whether the hypothesis is qualified. The process of comparison is available through checking if coefficients of the actual-data-based VAR lie in the acceptable range of the theoretical model's implied joint distribution. By the means of that, we can even inspect the model's capability of directing the dynamics and variances of the data.

In this chapter, VAR (1) is used as the auxiliary model by us and is treated as the descriptors of the actual data for three main macro variables (i.e., output gap, inflation, and interest rate). The Wald statistics is calculated from the VAR (1) coefficients and the variances of the three main economic variables. Therefore, the Wald test statics is a criterion to determine whether the observed dynamics and volatility of the selected three main variables are interpreted by the simulated joint distribution of

[44] Robust instrument variables estimation is suggested by McCallum (1976) and Wickens (1982), in which the lagged endogenous data are set as instruments, and the fitted values are computed from a VAR (1) what is used as the auxiliary model during evaluation procedure as well.

[45] By estimating the auxiliary model VAR on each pseudo sample, we can have the distribution of the estimates. The dynamics properties are captured by VAR estimates, while the volatility properties can be captured by the variance of the main variables. For the individual estimates, the confidence interval (95%) is calculated directly from their bootstrapped distribution.

these at a given 95% confidence level. The Wald statistics is formulated by,

The equation above is a function of the gap between and . is the vector of VAR estimates of the selected US data descriptors. is the arithmetic mean of the N estimated vector of VAR estimates derived from bootstrap simulations. is the variance and covariance matrix of the distribution . In addition, and are the actual data sets and simulated data sets respectively. is the vector of the parameters of the theoretical model. Then we can check the positions of Wald test statistics within the distribution generated by model.

Indirect Inference can be proceeded by comparing the percentile of the Wald distribution. In detailed, for a 5% significant level, a percentile above 95% would not lie outside the non-rejection area. The distribution of as well as the Wald statistics are obtained through bootstrapping method.

Transformed Mahalanobis Distance (Normalized t-statistics)

The TM statistic is used in the situation which we are hardly able to distinguish the models' relative performances. For instance, there are two or more specified models rejected simultaneously by Wald test statistics, we have to use the TM statistic to rank these models after comparison. Additionally, the TM provides a way to examine how bad the model is by observing how far it deviates away from 1.645. The bigger the number is, the worse the model fit. The Transformed Mahalanobis (TM) distance is defined as follows.

Herein, the TM distance is the transformation of the Wald test statistics.[46] Where is the Mahalanobis distance (value of Wald statistics) using the actual data, is the 95% critical Mahalanobis distance from simulated data (is the value of the Wald statistics falling at 95th percentile of the bootstrap distribution), and p is the number of parameters concerned or defined as degree of freedom respectively.

4.3 Indirect Inference Estimation Results

4.3.1 First Stage: Results of Calibration-based Indirect Inference Test

The testing steps presented above will be employed to test three rival models by using the US real-time quarterly data from 1969 to 2015 (the survey data will be used in robust check section). The data (variables) has been well defined in chapter 3 Appendix in the previous chapter. VAR (1) is taken as the auxiliary model in this chapter and the estimation of VAR is implemented with three economic observables: output gap, inflation, and nominal interest rate, which are quarterly observables.

First, we test the New-Keynesian three-equation models for both cases with and without inattentive features. The baseline model under full-information rationality assumption with an interest-rate smooth Taylor rule. Moreover, all three errors are presumed to follow AR (1) processes, which is in line with the previous studies. We evaluate the three competing models basing on the actual errors derived from estimation on the actual data. Moreover, it requires an estimation of the model's structural errors which are the residuals in each equation of the structural model given by the actual data and the expected variables in that equation. The residuals of

[46] This function of Transformed Mahalanobis distance (normalized t-statistic) is based on Wilson and Hilferty (1983)'s method of transforming Chi-square distribution into a standard normal distribution calculated.

demand, cost-push, and monetary policy are estimated respectively. There are two extra AR (1) processes corresponding to final data revision processes in the model under imperfect information data revision assumption,

After evaluating each competing model, we can assess each mechanism of them. The model with full-information rationality assumption has been argued failure to generate delay response by the previous studies. Thus, two alternative approaches have been proposed by recent studies in order to remove such a fail. The two approaches are modeling with sticky information, and modelling with imperfect information data revision respectively. Regarding to the former approach, the economic agents adjust their decisions with delaying behavior and such delay behavior is generated by the information costs. On the other hand, concerning the latter approach, the economic agents adjust their decisions with delaying behavior due to data revision issue. The two explanations, which have been suggested from the previous studies to remedy the weakness in the baseline model, are selected to be examined in this thesis. Indirect inference test (full Wald test) is employed in this chapter to check each competing model's overall data dynamic performance in an absolute way. The transformed Mahalanobis distance (normalized t-statistics) is also used to measure how similar the to-be-examined model is to the real world.

4.3.1.1 Calibration Parameters (Initial-Presumptive parameters)

The overview of all structural parameters along with their initially presumptive values are presented in Table 4-1 (Part 1) and most of the values are identical to the prior means which have been used in the previous chapter.

TABLE 4-1 (PART 1) STARTING CALIBRATION STRUCTURAL PARAMETER VALUE

Parameters	Definition	Values
	Common Parameters	
	Time discount factor (fixed)	0.99
	Price stickiness	0.6
	Elasticity of intertemporal substitution	1
	strategic complementary parameter (fixed)	0.15
	Degree of partially adjustment in Taylor rule	0.75
	Coefficient of inflation on Taylor rule	1.5
	Coefficient of output gap in Taylor rule	0.12
	SI Expectation Model	
	Share of updating firms (Mankiw & Reis, 2007)	0.5
	Share of updating consumer (Mankiw & Reis, 2007)	0.5
	IF Expectation Model[47]	
	output coefficient in output revision process	0.5
	inflation coefficient in inflation revision process	0.5

Table 4-1 (Part 2) Starting Calibration Parameter Value of AR Coefficients[48]

	FIRE Model	
	AR coefficient of demand shock	0.90
	AR coefficient of cost-push shock	0.79
	AR coefficient of policy shock	0.59
	SI Expectation Model	
	AR coefficient of demand shock	0.89
	AR coefficient of cost-push shock	0.79
	AR coefficient of policy shock	0.64
	IF Expectation Model	
	AR coefficient of demand shock	0.67
	AR coefficient of cost-push shock	0.56
	AR coefficient of policy shock	0.30
	AR term of shock in final revision process of x	0.41
	AR term of shock in final revision process of	0.61

The AR coefficients' parameters and correspondent values shown in Table 4-1 (Part 2) are achieved from the sample estimation of US real-time data for each applied

47 The initial null hypothesis is that , meaning not well-behaved revision processes.

48 The AR coefficients of the structural errors implied by the models, all of them are sample estimated base on the real-time data.

model.[49] The presumptive parameters' values (calibration value) are largely in line with the mean values of priors we adopted in Chapter 3.

4.3.1.2 Comparison through Calibration-Based Testing

The model cannot be bootstrapped without the solution of the structural error which can be reached if the observed actual data and presumptive parameters are given. As a rule, the times of bootstrapping is normally set as 1000. Following by this step, the test statistics are reached through examining the distribution of simulated pseudo samples. The main focus of this section is to shows that testing results by using presumptive parameters. The implementation of the overall model performance test is realized through the combination of dynamics parameters and volatility parameters.

TABLE 4-2 COMPARISON TM DISTANCE BY USING CALIBRATION PARAMETER

Model	Full Wald percentile %	TM by using Calibration Parameter
FIRE Model	100	4.1538
SI (j=4) Model	99.4	2.7338
IF Model	100	28.5625
Note: Above results, VAR (1) has been used as auxiliary model.		

The calibration-based testing results are shown in Table 4-2. Since the full Wald percentiles of the model all above 95, the models are implied to be not fit for dynamic properties of the actual data as they are not falling within the non-rejection area (i.e., 95 % confident interval). The smaller the value of Wald percentile is, the better level of fit its model reaches. The values of TM distance, which are higher than the norm of 1.645, imply the same trend that the Wald percentile tell us by displaying the extent

[49] It is doubtable that OLS is a biased estimator of the auxiliary model, due to the presence of lagged endogenous variables as regressors. However, it should not influence the power of test as the identical auxiliary model and estimators are applied for depicting the simulated data and the actual data. In other words, the same bias is translated into each model. Such that, in fact indirect inference is used to test whether the model-based OLS-estimated auxiliary model would generate the actual-data-based OLS-estimated auxiliary model.

of the gap. Basing on the initial presumptive error properties of US reduced-form New-Keynesian DSGE type models, the three competing models do not fit the actual data accurately. However, according to the result of the initial calibration-based test by using the presumptive parameters, it shows that the model with sticky information wins the best performance, which can be assessed through comparing the TM distance (normalized t-statistics). Surprisingly, what is contrary to the result obtained by Bayesian estimation approach that the model with imperfect information data revision is far worse than its rivals. However, the contradiction of the calibration-based testing result may be due to our initial presumptive parameters which is not the best options to closely copy the overall dynamic properties of the actual data. Thus, indirect inference will be conducted later as an estimation method aiming to search the 'best' set of parameters for the three competing models respectively.

4.3.1.3 Robustness Check

1) *Higher-Order VAR as Auxiliary Model*

We have implement the calibration-based indirect inference test towards our three competing models. Suffering from the same problems of robustness with the previous scholars, we realize that the choice of auxiliary model may influence the testing results, since it plays a role of independent intermediary to evaluate the gap between the theoretical model to the reality. To mitigate our concerns about the interference regarding auxiliary, the higher-order VAR as auxiliary models are introduced in robustness check. We apply the estimates of the coefficient matrix and the volatility of the data as the descriptors selected in the auxiliary model.

The results shown by the higher-order VAR auxiliary model robustness check, the result of VAR (1) is robust, which gives an answer to why a VAR (1) could approximate the DSGE models. Higher-order auxiliary VAR model, which was for model evaluation originally, is used by us out of detecting whether the ranking of three competing models is robust. We achieve this transformation of aim by the fact that using a VAR as the auxiliary model with higher order contributes to the strictness of the model evaluation because it requires more details of the data to be fitted. Although, in general, applying higher-order VAR as auxiliary model gives worse result for each competing model (i.e., since it requires data to match more specific characteristics) and which could be a way to develop the evaluation when the difference of models' performances is not obvious through less order auxiliary model, VAR (1).

In summary, in this section we mainly focus on checking whether the rank of the three competing models will change when a higher-order VAR model is used. The results show that the model with sticky information expectation under more stringent condition still outperform the alternative models.

TABLE 4-3 MODEL PERFORMANCE UNDER DIFFERENT AUXILIARY MODELS

Competing model	FIRE	SI (j=4)	IF	FIRE	SI (j=4)	IF
Auxiliary model	VAR (2)			VAR (3)		
TM Distance (Full Wald percentile %)	27.0317 (100)	10.3896 (100)	37.7490 (100)	29.4240 (100)	14.1153 (100)	45.0990 (100)

2) Using Alternative data resource: survey of professional forecaster data of output gap and inflation

Concerning different types of data resource may provide extra information in favor of different model, we implement the same testing procedure in this section by using survey of professional forecaster data. There are two noteworthy things in the testing results. The first one is that the survey data does not provide extra useful information to improve models' performance excepting for the model with imperfect information data revision. The TM distance of imperfect information data revision model decreases from 28.5625 to 15.7632, which indicates that the distance between theoretical model and 'real-world model' has been narrowed down. The second thing is that although none of them can pass the test, it implies that the survey data contains some extra information that help us to make further distinguish between the baseline model and sticky information model. However, such calibration-based testing result may due to the initial presumptive parameters may be not the best options to describe the survey data, so that it is essential to carry out indirect inference estimation in the next stage to search the ‘best’ collection of parameters which can be applied to maximum degree narrow down the gap between theoretical model and the reality to make a fair comparison for models.

TABLE 4-4 COMPARISON TM DISTANCE BY USING CALIBRATION PARAMETERS (WITH SURVEY DATA)

Model	Full Wald percentile %	TM by using Calibration Parameter
FIRE Model	100	17.9522
SI (j=4) Model	100	4.1554
IF Model	100	15.7632
Note: Above results, VAR (1) has been used as auxiliary model.		

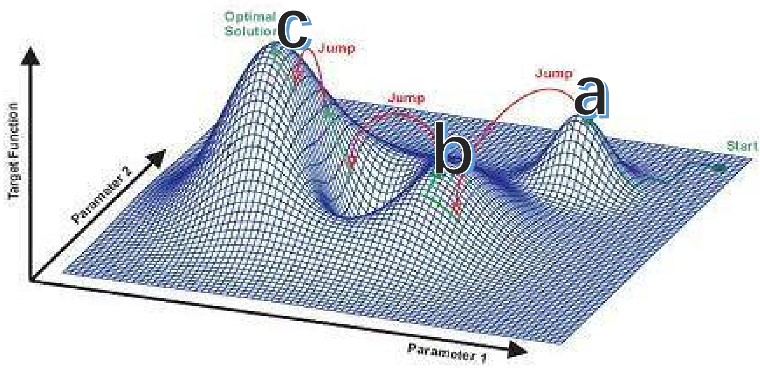
4.3.2 Second Stage: Results of Estimation-based Indirect Inference Test

Since the initial presumptive parameters used in the first stage may not lead to the optimal results, we decide to try to find out another collection of parameters to interpret the way of the data’s generation in this chapter. If there not exists such a group of parameters which enables the model to pass indirect inference test, the model will be judged as rejected. As the aim of the second stage, searching the parameter set leading the model to replicate the real world as well as possible which can be defined as ‘estimation-based Indirect Inference test’, through which the chance of being accepted to the testing model will be maximized.

Essentially, this stage gives a way to solve the problem of parameter uncertainty. In practice, we can reduce the parameter uncertainty in a direct way by checking the Wald statistic derived from the group of parameters for the model. In detailed, the more the Wald statistic decreases, the better the parameter set performs. Herein, an effective algorithm basing on Simulate Annealing (SA) is introduced to search the optimal parameter set by starting from an extensive scope around the initial values along with random jumps around the space. With SA algorithm, we can have the lowest value of full Wald statistic for three rival models.

The SA algorithm refers to a stochastic optimization based on Monte-Carlo iterative solution strategy. The principle is inspired by the annealing process of metal heating and cooling through which the temperature of the object will be controlled to increase the size of the metal's crystals and reduce its defects. By mimicking the mechanism, the SA searches for the probabilities with lower energy to minimize the defects of crystal (resemble that of the steps of minimizing Wald statistics in estimation process of indirect inference). It tries to find the optimal parameter set repeatedly until the system reaches a minimum value of Wald statistics, or until a given computation budget has been exhausted. Since the principle of accepting a less optimal consequence temporarily, SA can reach the optimal consequence in a global scale instead of being trapped in local optimum. For example, according to Figure 4-1, SA mechanism allows one to search over the whole apace starting from the initial state (in an indirect inference estimation process, a current state is equivalent to the group of structural parameters) and jump to nearby local optimal 'a' and continue to search toward global optimal 'c'. The less optimums in between 'a' and 'c', taking 'b' as an example, will be accepted as a 'springboard' which one can jump to in order to jump and search the other space to reach the global optimum.

Figure 4-1 Simulated Annealing (SA) Avoiding Getting Stuck in Local Optimum[50]



Overall, in the application of indirect inference estimation, SA is used to seek the optimal set of parameters, which will facilitate to discover lowering Wald statistic until the computation budget used up. To carry out the numerical iterations to minimize the Wald statistics, the initial values of the parameters of structural models are required. Here, the starting values are the values of the presumptive parameters, such presumptive parameters are plausible and from the previous studies, meanwhile, we permit the parameters to seek around -0.5 to +0.5 of their starting values under estimation.

To implement estimation-based Indirect Inference test, the VAR (1) needs to be used continuously as the auxiliary model to give a reference substance for the estimated models to those of the calibrated models. The VAR (1) are used as descriptors of the coefficient matrix and the variance of the data. Just like the previous testing exercise.

In the first stage, the structural parameters were assigned by the initial presumptive values but those are selected in line with the commonly accepted values from the previous studies. Being distinguished from the first stage, the second stage uses indirect inference as estimating method to re-assess the three competing models on

---

50 Figure Source: http://www.frankfurt-consulting.de/img/SimAnn.jpg

their grooves based on actual data, which means that the restriction of initial presumptive parameters has been released.

We may expect that, by indirect inference estimation or simulated-annealing estimation, estimated version of the three competing models would behave no worse than that found in the first stage. Seeing that when we take calibration values as the initial presumptive ones to assign the structural parameters, the SA mechanism will begin to explore from these initial presumptive values to substitute for them with 'better' values based on the actual data if only a minimum Wald statistic can be discovered. The process will be terminated when the Wald statistic can no longer be reduced, which implies that we have discovered the 'best' estimates of the structural parameters. The Simulated Annealing method, which facilitates to adjust the initial presumptive values, is helpful for the models to pass the test.

4.3.2.1 Estimation-Based Indirect Inference Testing Results: Full Information Rationality Assumption Model

The Simulated-Annealing-estimation-based test as well as the Bayesian-estimation-based test with respect to the three competing models for US economy are presented in Table 4-5, Table 4-6 and Table 4-7 respectively. The numbers in the column regarding the indirect inference estimation are obtained through SA estimation method. The scope of the value of parameters during SA exploring is limited within plus or minus 50% of the presumptive values of coefficients.

The main idea of indirect inference as an assessment methodology is to test the existing model to detect whether the structural parameters are capable to generate the actual data. However, if these initial presumptive parameters cannot be used to explain the generating process of the actual data, another set of parameters may be somewhere existed and can be applied to explain how the actual data is generated. If the model with initial presumptive parameters already fall within the non-rejection scope, it is still necessary to explore another group of parameters that can narrow the gap in the middle of the theoretical model and the reality, which leads to better testing results. The 'best' set of the structural models' parameters are those to the maximum degree to shorten the distance between theoretical model and the reality.

In the second stage, we aim to explore the 'best' collection of parameters throughout the entire parameter space by the implementation of Indirect Inference without changing the signs of parameters as an estimation-based test approach. The minimized value of the distance (Mahalanobis distance) is captured for each competitor over the US sample periods through a Simulated Annealing algorithm. The 'best' collection of parameters that can furthest shorten the distance between the theory and the reality will be used for our estimation-based test. Using these optimal sets of parameters to compare models can reduce the unfairness in model comparisons.

Table 4-5 displays the estimation results of the model with full-information rationality assumption (FIRE Model). Overall, the estimated values of parameters of the FIRE model through indirect inference estimation of are not significantly far away from those obtained by Bayesian estimation. However, some distinguished cases exist. Particularly, the estimated value of the elasticity of intertemporal substitution is 0.5180, which is quite higher than that obtained from the Bayesian estimation. Besides, the same trend can be found in the value of price stickiness versus that of Bayesian estimation. Subsequently, examining the estimates of the major behavioral parameters of FIRE model, we toward to examine the parameters of the monetary policy function, which are based on standard interest-rate smoothed Taylor rule (1993). Regarding to the estimated coefficients of monetary policy, excepting which is increased less than 8%, the other two (i.e., and) both increase around 35% comparing to their estimated values achieved from Bayesian estimation. Within the system, all the three stationary shocks are quite highly persistent and two of them, excepting for the AR coefficient of monetary policy which is increased above 60% than that obtained through Bayesian estimation, are similar to the Bayesian estimated results.

In detailed, through SA estimation, the estimated value of is 1.5079 which is slightly higher than that obtained by Bayesian estimation. The two estimates regarding different estimation methods are both close to the initial calibration value (i.e., 1.5). The estimated value of the reaction to output gap is 0.1439 which is lower than that obtained by Bayesian estimation, which indicates that the monetary policy does not seem to react forcefully to the output gap level. Moreover, the parameter of interest rate smoothness which is estimated to be 0.6580 and lower than that obtained through Bayesian estimation. However, it is not far away from the initial presumptive value (i.e., 0.75). Besides. the AR coefficients regarding to the three exogeneous stationary shocks which are demand shock, cost-push shock and monetary policy shock are estimated to be very persistent, which are 0.8587,0.7318 and 0.8155 respectively.

Furthermore, the test statistic implies a Wald percentile of 64.8, so the FIRE model is not rejected at the 5% significant level. In practice, the Wald statistic is within the non-rejection region of the bootstrap distribution. Overall, many of the estimates obtained through SA estimation have shifted away from the estimates obtained through Bayesian estimation for a distance (e.g., the elasticity of intertemporal substitution is increased around 97% higher than the Bayesian estimated value what is 0.0225. The SA estimated value of price stickiness is around 25% higher than the counterpart of Bayesian approach). It is indicated in Table 4-5 that the model estimated with SA estimates performs better than the model estimated with Bayesian estimates in fitting the actual data. The reported Wald percentile has gain the significant reduction comparing with the one obtained through using Bayesian estimates. The full Wald statistics implies that the FIRE model with SA estimates fall within the non-rejection area, meaning that the model cannot be rejected at a chance of 95%. Furthermore, the model with Bayesian estimates performance worse than the model with the initial presumptive parameters (calibration parameters).

TABLE 4-5 ESTIMATES OF FIRE MODEL

Parameters	Starting Calibration	Bayesian Estimates	SA Estimates
	1	0.0225	0.5180
	0.6	0.7257	0.9677
	0.75	0.8834	0.6580
	1.5	1.3891	1.5079
	0.12	0.1974	0.1439
	0.86	0.7995	0.8587
	0.73	0.6948	0.7318
	0.82	0.3094	0.8155
Full Wald %	100	100	64.8
TM (normalize t-statistic)	4.1538	26.0498	0.6587

4.3.2.2 Estimation-Based Indirect Inference Testing Results: Sticky Information Expectation Model

Table 4-6 displays the estimation results of the model with sticky information (SI model). Overall, most estimates through SA estimation are higher than those obtained from Bayesian estimation, excepting that the estimate of interest rate smoothed parameter is 0.7672 which is a little bit lower than that obtained through Bayesian estimation. The reaction parameter of output gap in monetary policy is estimated to be around 13%, which is lower than that in Bayesian estimates as well but being not quite far from its initial presumptive value. However, some SA estimates are higher than the Bayesian estimates, particularly the AR coefficient of monetary policy which is two times higher than that obtained through Bayesian estimation.

Furthermore, the test statistic indicates a Wald percentile of 53.10, so the SI model cannot be rejected at the 5% significant level, meaning that Wald statistic is well included in non-rejection region of the bootstrap distribution. Additionally, many SA estimates are somehow different from the estimates achieved by Bayesian estimation. For instance, the elasticity of intertemporal substitution is seven times higher than the Bayesian estimated value 0.1092. As well as the SA estimated share of updating firms whose estimate is 0.4504, it is about 1.5 times larger than that (i.e., 0.3084) obtained through Bayesian estimates but closer to the counterpart (i.e., 0.657) in empirical studies (Reis, 2009). Besides, the share of updating consumers is estimated 2 times larger than that obtained though Bayesian approach.

TABLE 4-6 ESTIMATES OF SI MODEL (J=4)

Parameters	Starting Calibration	Bayesian Estimates	SA Estimates
	1	0.1092	0.9050
	0.6	0.6340	0.5542
	0.75	0.9002	0.7672
	1.5	1.3735	1.6266
	0.12	0.1848	0.1299
	0.89	0.8139	0.8842
	0.79	0.6490	0.6421
	0.64	0.2986	0.7351
	0.5	0.3084	0.4504
	0.5	0.2362	0.5138
Full Wald %	99.4	54.00	53.10
TM (normalize t-statistic)	2.7338	-0.2072	0.1092

4.3.2.3 Estimation-Based Indirect Inference Testing Results: Imperfect Information Data Revision Expectation Model

In Table 4-7, in general, although none of the three cases concerning calibration-based model test, Bayesian-estimated-based model test and SA-estimated-based model test, can pass the test, the model with Bayesian estimates gives the worst result which can be inspected through TM distance (normalized t-statistics). The most significant difference between the SA-estimated-based model test and the Bayesian-estimated-based model test is that the estimated value of coefficient of the former test, being closer to its initial presumptive value, is ten times larger than the value obtained through the latter test.

TABLE 4-7 ESTIMATES OF IF DATA REVISION MODEL

Parameters	Starting Calibration	Bayesian Estimates	SA Estimates
	1	0.0899	0.8639
	0.6	0.7389	0.5623
	0.75	0.8801	0.6495
	1.5	1.0884	1.3342
	0.12	0.1962	0.1131
	0.5	1.8500	0.4404
	0.5	1.1198	0.4683
	0.67	0.6186	0.6292
	0.56	0.3657	0.5083
	0.30	0.2235	0.2718
	0.42	0.7252	0.3443
	0.61	0.8535	0.5099
Full Wald %	100	100	100
TM (normalize t-statistic)	28.5625	94.6459	20.3812

4.3.2.4 Comparison through Estimation-based Test

4.3.2.4.1 TM Distance Comparison

Overall, due to the norm of 1.645 as a threshold of judging the succeed of pass, only the models whose absolute values of TM Distance are below 1.645 can be qualified being 'good enough' models. According to Table 4-8, the SI Model can pass Bayesian-estimated-based test and SA-estimated-based test with a fail in calibration-based test, while the FIRE Model and the IF Model can pass 1 and 0 test respectively. We can drop a conclusion that the SI Model is superior to the other ones in terms of overall model fit.

The assessment of model is more precise by using the SA estimates from the point of view of actual data. Since the AR coefficients in SA estimates are estimated basing on the structural errors which use the actual observed data and parameters estimated in the model. The SA Estimation, in which the initial presumptive parameters are replaced by the optimal ones for re-test leading to higher passing possibility for the competing models, does not allow the IF model to pass. In general, the results of SA-estimation-based testing are better than the results of initial calibration-based testing as expected. This improvement can be attributed to the application of SA estimation approach what explores all the potential parameters over wild space to discover the best fit.

TABLE 4-8 COMPARISON TM DISTANCE (NORMALIZED T-STATISTICS)

Model	Starting Calibration	Bayesian Estimates	SA Estimates
FIRE Model	4.1538	26.0498	0.6587
SI (j=4) Model	2.7338	-0.2072	0.1092
IF Model	28.5625	94.6459	20.3812

4.3.2.4.2 Estimated Impulse Response Functions (IRFs)

In this section, the estimated impulse response functions have been used as the main tools to explore each competing model's behavior under all three shocks (i.e., demand shock, cost-push shock and monetary policy shock).

IRFs of Monetary Policy Shock

Figure 4-2 displays the estimated impulse response of the three main variables (i.e., output gap, inflation, and interest rate) to the monetary policy shock of three competing models respectively. In general, under the estimated monetary policy reaction function, the responses of the same variable under different models are quantitively similar. To be specific, nominal interest rate increases, but output gap and inflation decrease with respect to the three competing models. As shown in Figure 4-2, throughout the impacts of monetary policy shock on inflation and output gap, the hump-shaped response only appears under the SI model. Regarding to the period of convergence, the convergences of three main variables under FIRE model (the baseline model) and SI model are around 18 periods, but under IF model (i.e., the model with imperfect information data revision) they converge faster. Surprisingly, under the model with imperfect information data revision, the impact of monetary policy shock not only fails to generate the hump-shape response on inflation and output gap, but also weakens the delay response on interest rate.

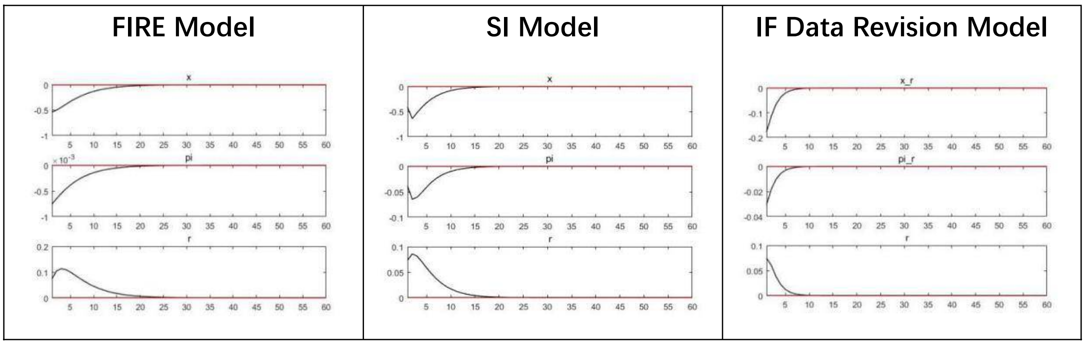


Figure 4-2 Estimated Impulse Response Function of One Unit Positive Policy Shock to Main Variables (x=output gap, pi=inflation, r=nominal interest rate)

IRFs of Demand Shock

Figure 4-3 presents the estimated impulse response functions of the three main variables to demand shock regarding the three rivals. Overall, the positive demand shock has a positive effect on three main variables. Besides, the effect last for a long time (i.e., around 20 periods more) under FIRE model and SI model. However, the effects on three main variables are relatively short with respect to the IF model. Furthermore, the demand shock has a persistent impact on inflation and output gap under SI model, which does not appear under the other two competing models.

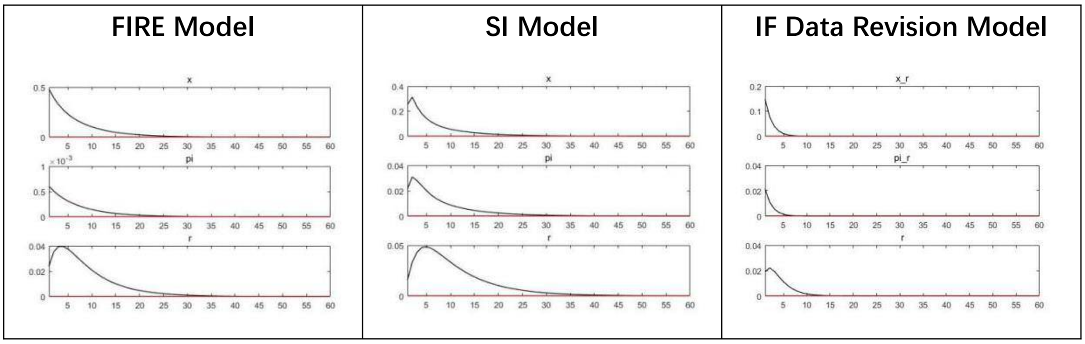


Figure 4-3 Estimated Impulse Response Function of One Unit Positive Demand Shock to Main Variable (x=output gap, pi=inflation, r=nominal interest rate)

IRFs of Cost-Push Shock

Figure 4-4 shows the behavior of three main variables in response to the positive cost-push shock with respect to three competitors. In general, all three competing models generate similar dynamics quantitatively. In detailed, both the inflation and interest rate are affected positively by the positive cost-push shock which delivers a negative effect on output gap. Additionally, the cost-push shock has the largest effect at initial point under FIRE model on three main variables. Meanwhile, it has a moderate effect at initial point under SI model and a minimal effect under IF model in terms of periods return to steady state.

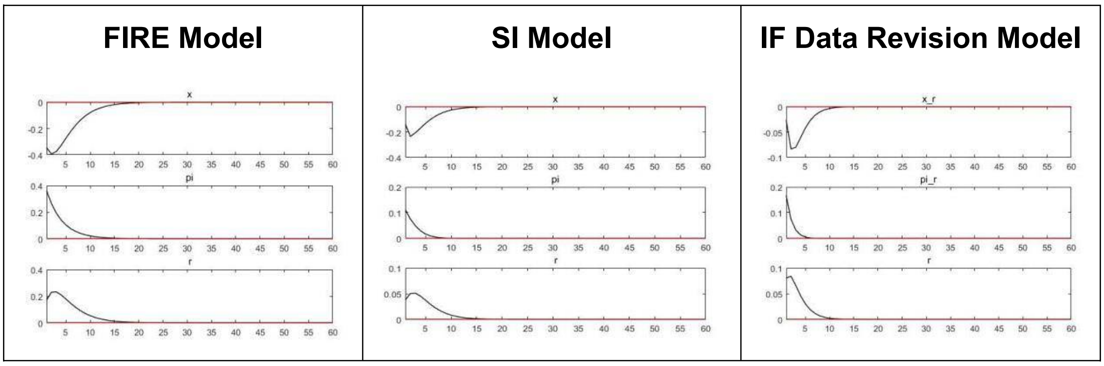


Figure 4-4 Estimated Impulse Response Function of One Unit Positive Cost-Push Shock to Main Variable (x=output gap, pi=inflation, r=nominal interest rate)

To sum up, the estimated IRFs are not very different from those obtained by Bayesian estimation. The SI model has strong abilities of generating more persistence and reproducing delay responses to monetary policy shock. However, the IF model still cannot achieve this goal.

4.3.2.5 Robustness check

1) Higher-Order Auxiliary Models

In this section, as the same in Section 4.3.1.3, we need to check that whether the rank among the three competing models in terms of higher-order auxiliary models is robust with the optimal set of parameters. We chose a VAR (1) as the auxiliary model in which the selected descriptors are equivalent to the estimates of its coefficients matrix and data variance incorporated in the indirect inference estimation procedure. As stated earlier, there are two factors, which are the model required to fit and the its extent of fit, that decide which option we should choose as the auxiliary model from a higher-order VAR model and other multiple types of time series models. When the higher-order auxiliary model VAR (2) and VAR (3) have been applied, the results show that although none can pass the test, the models' performances still can be compared. According to Table 4-9, the leading position of SI model in terms of overall dynamic properties over the competitors has not been switched when we choose higher-order VAR (i.e., VAR (2) or VAR (3)) instead of VAR (1) as auxiliary model.

Overall, the results of TM statistics in Table 4-9 indicates that raising VAR's order would make the acceptance of all the three estimated models weaker due to the greater burden placed on them. Comparing the results of TM statistics from Table 4-9 and Table 4-8, we can draw three conclusions. Firstly, it is obvious that when we use lower order VAR (i.e., VAR (1)) as the auxiliary model, all three competing models are less rejected. Secondly, the SI model is always less rejected than the competitors, which indicates that the SI model is preferred from the angle of model's overall performance regardless of the auxiliary VAR models' order. Thirdly, the ranking of three competing models is identical to the previous regardless of different choices of auxiliary models (i.e., VAR (2) or VAR (3)) through SA estimation among three rivals. So, VAR (1) can be an accepted auxiliary model to mimic the theoretical models.

TABLE 4-9 MODEL PERFORMANCE UNDER DIFFERENT AUXILIARY MODELS

Competing model	FIRE	SI (j=4)	IF	FIRE	SI (j=4)	IF
DATA SAMPLE: WITHOUT SURVEY DATA						
Auxiliary model	VAR (2)			VAR (3)		
TM Distance (Full Wald %)	8.1734 (100)	7.4455 (100)	32.1638 (100)	11.7022 (100)	9.1573 (100)	47.4983 (100)

2) *Different Truncation Point j of Sticky Information Model*

In this section, as the same as in Section 3.7, we need to check the robustness of different truncation point j in SI model but through the indirect inference approach. We have selected alternatives =6 and 8 to imply them into robust check procedure. According to Table 4-10, we receive the same suggestion as the one provided by Bayesian estimation approach that incorporating more lagged information into SI model has merely influence on its model performance after checking the TM distance (normalized t-statistics). Furthermore, the ranking among three rivals is identical as the previous ranking no matter which value of truncation point j (i.e., =6 and 8) in SI model is applied.

TABLE 4-10 SENSITIVITY CHECK BY USING MINIMIZING COEFFICIENT VALUES FOR SI MODEL

Model	TM by using SA Estimated Parameter
FIRE model	0.6587
SI model (j=4)	0.1092
SI model (j=6)	-0.2796
SI model (j=8)	-0.3518
IF model	20.3812

3) *Using alternative data resource: survey of professional forecaster data of output gap and inflation*

The estimation result by using Survey of Professional Forecaster Data (survey data) is presented in Table 4-12. The results obtained through Bayesian estimation approach show that the performance of IF model is far more superior to its rivals'. However, through indirect inference estimation, it shows that the full ability of IF model is far inferior to its competitors'. When each model is estimated by using survey data instead of real-time data, none of them can pass the test. In addition, it becomes more difficult to tell which one from FIRE expectation model and SI expectation model can give the better replication of the full dynamics of the actual observables (i.e., survey data) better. However, SI model performs at least no worse than the baseline when SPF data has been used.

TABLE 4-11 STARTING CALIBRATION PARAMETER VALUE OF AR COEFFICIENTS[51]

FIRE Model		
	AR coefficient of demand shock	0.94
	AR coefficient of cost-push shock	0.75
	AR coefficient of policy shock	0.56
SI Expectation Model		
	AR coefficient of demand shock	0.93
	AR coefficient of cost-push shock	0.74
	AR coefficient of policy shock	0.56
IF Expectation Model		
	AR coefficient of demand shock	0.70
	AR coefficient of cost-push shock	0.54
	AR coefficient of policy shock	0.29
	AR term of shock in final revision process of x	0.39
	AR term of shock in final revision process of	0.59

TABLE 4-12 COMPARISON TM BY USING MINIMIZING COEFFICIENT VALUES (WITH SURVEY DATA)

Model	SA Estimation Parameter
FIRE Model	5.6900
SI (j=4) Model	5.2699
IF Model	12.4718

51 The AR coefficients of the structural errors implied by the models, all of them are sample estimated base on survey of professional forecaster data.

4.5 Conclusion

In this chapter, we use indirect inference as a testing method (i.e., calibration-based testing method) at starting stage and take the same approach as an estimation method (i.e., estimation-based testing method) in the next stage. We aim to contradistinguish the performance of the simulated-data-based estimated auxiliary model, with the performance of actual-data-based estimated auxiliary model through indirect inference test method.

We implement indirect inference methodology to test the three competing models regarding its dynamic performance for US economic real-time quarterly data from 1969 to 2015 (also use the other type of sample data, i.e. survey of the professional forecaster data, over the same period in robustness check). We compared three versions of model and found that none of them can fit the actual data through the initial calibrated-based test. Surprisingly, the imperfect information has the worst performance among the three models, which is contradicted to the results obtained by Bayesian estimation approach. However, the calibration-based testing results obtained by Indirect Inference approach shows that the model with sticky information expectation performs best among three competitors.

In the second stage, Indirect inference has been applied as estimation approach to both types of expectation models: with and without inattentiveness which were investigated in chapter 3. The comparisons of each competing models through Bayesian-estimated-based test and SA-estimated-based (Indirect Inference) test have been conducted respectively. The results indicate that the performance of each competing model with SA (indirect inference) estimates (i.e., best fitting parameters) has been improved, when compared with the results of the calibration-based test from the first stage.

Four achievements can be reflected through the results of indirect inference estimation. Firstly, regardless of two different estimation methods (i.e., Bayesian estimation and Indirect Inference estimation) by using the real-time data, the model with sticky information expectation is all the way preferred among the three competitors. Secondly, when we tried to find a robust superior model in terms of dynamic performance by changing the conditions, such as auxiliary model, truncation point in SI model, and type of data resource, we found that the model with sticky information expectation still the best choice to fit the US economy, Thirdly, the impacts of the structural shocks on US economy have been analyzed by the estimated impulse response functions. In general, these impacts are not significant different from the previous studies quantitatively, as well as those estimated through Bayesian estimation in chapter 3. For instance, a positive demand shock result in a raise in output gap, inflation, and interest rate. A positive monetary policy shock impact interest rate positively but creates a decrease in both output gap and inflation. Fourthly, unexpectedly, the model features imperfect information data revision fails to pass the test and gain the worst performance, which is contradict to not only the result obtained through Bayesian approach but also the suggestions from previous studies.

Overall, although Bayesian estimation approach is an effective practical tool to inspect model's performance by taking prior information about the macro economy into consideration, the prior is restricted while being applied because prior distribution need to be determined before entering estimation process. Besides, the model's performance obtained by Bayesian estimation are showed in a relative way that impossible to evaluate their absolute abilities. Thus, the method of indirect inference used in this chapter is an advanced tool to re-estimate each competing model in an 'unrestricted' way by exploring all the potential sets of parameters which can be accepted by models. In addition, the independent VAR has been used as an auxiliary model which offers a way to examine each model in an absolute sense. Besides, the optimal set of parameters can be discovered through SA mechanism for each competing model, to mitigate the unfairness in model comparisons.

While we were replacing the real-time data with survey data to apply them in estimation procedure, we found that the performances of models were increased excepting the cases of FIRE model and SI model through Indirect Inference. This contraction indicates that the survey data may contain useful information to improve the imperfect information data revision model's performance.

Appendix A to Chapter 4

TABLE 4A-1 ADF TEST RESULTS OF THE REVISED VARIABLES

Variables	Option	Critical value	t-statistics	Inference
	None	-1.942013	-5.411552 (-6.62896)	stationary (stationary)
	None	-1.942013	-3.242983 (-3.280844)	stationary (stationary)
Note: the number in the bracket is tested by using SPF revised data; outside the bracket is tested by using real-time revised data.				

TABLE 4A-2 ADF TEST RESULTS OF THE SURVEY OF PROFESSIONAL FORECASTER VARIABLES

Variables	Option	Critical value	t-statistics	Inference
	None	-1.942013	-7.191524	stationary
	None	-1.942013	-5.285229	stationary
	None	-1.942013	-5.145850	stationary
	None	-1.942013	-13.82232	stationary
Note: Here, and are the SPF data which denote that use survey conducted at time t and release in next period; and similar for and .				

TABLE 4A-3 ADF TEST RESULTS OF REAL TIME VARIABLES

Variables	Option	Critical value	t-statistics	Inference
	None	-1.942013	-4.19852	stationary
	None	-1.942013	-7.128462	stationary
	None	-1.942013	-2.332022	stationary
	None	-1.942013	-2.344756	stationary
Note: Here, and are the real-time data t released after one period; and are the real-time data t release after three periods.				

Appendix B to Chapter 4

TABLE 4B-1 MINIMIZING COEFFICIENT VALUES FOR FIRE MODEL (WITH SURVEY DATA)

Parameters	SA Estimates
	0.0275
	0.6286
	0.7476
	1.7401
	0.0749
	0.7759
	0.6537
	0.2772
Full Wald %	100
TM (normalize t-statistic)	5.6900

TABLE 4B-2 MINIMIZING COEFFICIENT VALUES FOR SI (J=4) MODEL (WITH SURVEY DATA)

Parameters	SA Estimates
	0.9878
	0.5713
	0.7180
	1.5641
	0.1238
	0.7696
	0.6570
	0.5505
	0.5179
	0.4849
Full Wald %	100
TM (normalize t-statistic)	5.2699

TABLE 4B-3 MINIMIZING COEFFICIENT VALUES FOR IF MODEL (WITH SURVEY DATA)

Parameters	SA Estimates
	0.4386
	0.8435
	0.5477
	1.4304
	0.1292
	0.4655
	0.4399
	0.6968
	0.5394
	0.2788
	0.3977
	0.5777
Full Wald %	100
TM (normalize t-statistic)	12.4718

Chapter 5
General Conclusion and Further Research Direction

5.1 Some Valuable Summarizes of The Thesis

Through comparing the models with inattentive expectation, we have a flexible way to explain which inattentive feature can give a better explanation of the US economy. To be specific, basing on the most commonly used stylized New-Keynesian model, we successfully incorporate the inattentive expectation assumption into the model out of the existence of the cost for acquiring and processing the updated information, or the data revision issues. In the sticky information assumption, the agents are slowly incorporating information about macroeconomic conditions (i.e., output, inflation, and interest rate). For another, in the assumption of data revision, economic agents cannot observe the true state because of noises. These noises are originated from people's imperfect knowledge about the real economy.

This research arises from the two inattentive assumptions above which are suggested from the two proposals in previous commonly discussed literature - one is sticky information expectation (Mankiw and Reis, 2002, 2007); the other one is imperfect information data revision expectation (Casares and Vazquez, 2016; Arouba, 2008). These studies all share the same goal of remedying deficiencies in the classical full-information expectation type models.

The deviation from full-information rationality after incorporating inattentive feature should be significant in solving issues of macroeconomics (Akerlof, 2002; Sargent, 1993). For example, after incorporating inattentive expectation, they find that many problems arising from the New-Keynesian model under full-information rationality assumption can be solved. Firstly, it can solve the problem of New-Keynesian full-information Phillips curve which leads nonsensically counterfactual forecasts about the impacts of monetary policy due to lack of any source of inflation inertia. Secondly, the counterfactual evidence regarding disinflations resulting in booms rather than recessions (Ball, 1994) can be removed which is argued by Mankiw and Reis (2002).

Thirdly, it removes the inability of full-information New-Keynesian type model that offer the explanation to the question why monetary policy shock has a delayed and gradual impact on inflation (Mankiw and Reis, 2002). Thus, such inattentive behaviour assumption considering what role people act in terms of behavioural economics is ‘satisficer’ rather than full-information rational maximiser (Simon,1989).

However, these approaches incorporating inattentive features do have their weaknesses. They are not successful in explaining why people not apply diffusely obtainable information about real economy into their economic decision making. However, people may easily find out what the information, such as interest rate, published by central bank, but it is hard to interpret the meanings behind the numbers for people lacking professional knowledge. As a result, the real problem is not get access to information but dealing with it. Unluckily, economics does not hole the instruments to model imperfect information dealing process. The methods proposed by Woodford (2003), Ball (2000) and Mankiw and Reis (2002, 2007) are none of the hope that a model of imperfect information procurement may take as a rough replacement. Despite the weaknesses of incorporating inattentive ingredients, its characteristic of explaining inflation inertia leads the model more complying with the situation of real world.

The alternative inattentive expectation models are applied in this thesis to compare with the baseline model. The selects are two-specific reduced-form three-equation DSGE model with inattentive feature. The sticky-information model as the first select which is based on the idea that while people forming their expectation, they are restricted by the cost of processing and acquiring the current information from using the latest information (Mankiw and Reis, 2002, 2007). The imperfect information model as the second select can reduce noise through data revision process (Casares and Vazquez, 2016; Arouba, 2008). One of the most significant motivations of the data revision comes from that there is a remarkably deep output gap misperception during the great inflation of the 1970s. This misperception can be coming down to the mis-measurement of actual output. Such mis-measurement, which is present in almost macroeconomic series, is a quantitatively substantial source of misperceptions (Collard and Dellas, 2010).

.

Concerning the results through two estimation methods, there are some part coincident. Firstly, the model of sticky Information is detected to be the most favorable model in the light of fitting the real-time data behavior. Secondly, the model with sticky information is the only one can generate delay response, which is in line with the evidence observed in actual data. Thirdly, the imperfect information data revision model with the survey data has better performance than that with the real-time data. The gap of the model with different conditions indicates that the survey data contains extra information to help improve imperfect information data revision model's performance.

However, there are some conflicts between the two estimation methods. In detailed, through the Bayesian estimation approach by using survey data the model with imperfect information data revision wins the best position among three competing models, but such result is not robust under alternative estimation methodology (i.e., Indirect Inference). The conflicts may be stemmed from the following reasons. Firstly, due to the unobserved potential output, the traditional measures of the output gap are probably burdened with error. The mismeasurement of the true output gap could influence the ability of each selected competing model (Lown and Rich, 1997). Secondly, different estimation methodologies may potentially lead to different conclusions. However, it is obvious that there is no absolute optimal way to choose a macro econometric method to estimate and evaluate models. Different estimation methodologies have their strengths and weaknesses. For instance, the Bayesian estimation approach is superior on the aspect of incorporating priors linking to the previous studies, but it is deficient for the same aspect because these priors have been put 'restrictions' before estimation. Besides, how to set prior distribution before estimation is still a disputable issue. Moreover, Bayesian estimation only offers a way to obtain model's relative performance by comparison, which cannot examine a model's absolute ability individually. Thus, we decide to use an 'unrestricted' estimation and evaluation method, indirect inference, to estimate different models as a robust check approach. It may be doubted that there is no model of any sort is qualified enough to simulate the 'real world' for its complexity. However, as asserted by Friedman (1953): 'Complete realism is clearly unattainable, and the question whether a theory is realistic enough can be settled only by seeing whether it yields predictions that are good enough for the purpose in hand or are better than predictions from alternative theories'. Thus, a qualified model should not be assessed by 'literal truth', but by 'if it is true'. He gives the perfect competition as an example to demonstrate his idea. Although the perfect competition never actually exists, it predicts the industries' highly competitive behaviour. Thus, even there is no model perfect match the reality, we still test its own ability to what extent can be used to explain the real world. That is why the indirect inference is chosen as the robust evaluating method in this thesis.

5.2 Further Research

In this thesis, we estimate and test New-Keynesian reduced-form type models with respect to two different expectation assumptions--with and without inattentiveness--by using US macro-economic data (survey of professional forecaster data have been adopted in robust check section). In choosing inattentive models for comparing, many options are left by us, but they can be developed in future work in the following ways.

Firstly, we only consider inattentive expectation with small-closed economy. Future work could be conduct through empirically evaluating small-open economy by incorporating exchange rate, import and export to develop more complicated models for comparison. Secondly, we can investigate mix-inattentive model (Dräger, 2016) to compare with the single-inattentive model. This process could also be applied into both close and open economies. Thirdly, the robust check in this thesis regarding to different specification of monetary policy shows that although the rank among three competing models do not switch, with respect to different monetary policy specifications, each model’s performance changes significantly. Thus, further research can take the inattentive expectation as the base structure model but with different monetary policy to examine whether the monetary authority does a good job over recent decades, which can also be carried out through both Bayesian and indirect inference approach.

Supporting Annex

Full-Information Rationality Assumption Model Micro-foundations and Derivations (Baseline Model)

The main derivation is following the common deriving procedure in New Keynesian literature (e.g., Walsh, 2003; Menz and Vogel, 2009).

Full-Information Rational Expectation Model: IS Curve

Representative households are assumed to consume a composite of differentiated foods by monopolistically competitive firms that make up of a continuum of measure. The composite consumption that enters that utility function in each period is:

(A.1)

Where is the price elasticity of demand for good i. The cost minimization process of representative households implies that demand for good i is,

(A.2)

Where is the price of good i and is the aggregate price in period t. Each household maximizes the following discounted sum of future expected utility functions

(A.3)

Where stands for the time discount factor, while and denote the elasticities of inter-temporal substitution and the inverse of the elasticity of labour supply respectively. Subject to the period budget constraint

(A.4)

Each household derives utility from consumption and disutility from hours of labor supplied . In the budget constraint, stands for nominal bond holdings, denotes the aggregate price level, the real wage, the nominal interest rate, is the real term of dividend distributions, and is the real term of net transfer or taxes. The utility maximization problem can be described using the Lagrangean function as follows:

(A.5)

First order conditions imply,

(A.6)

(A.7)

(A.8)

And then we can get,

(A.9)

(A.10)

After log-linearization equation (A.9) around a zero-inflation steady state, where , and denote the percentage deviation from steady state.

(A.11)

And log-linearizing the resource constraint is [52],

(A.12)

Then the output gap is defined as the difference between actual output and potential output, where the potential output is the output under flexible price. The potential output can be solved approximately use the log difference of actual output from its HP trend.

(A.13)

Furthermore, here use the output gap rewrite the above log-linearizing Euler equation,

(A.14)

Where demand shock is an exogenous shock driven by exogeneous productivity shocks.

Full-Information Rationality Assumption Model: Phillips Curve

As explained in this small-closed economy the representative agent's households' own firms. Under monopolistically competitive environment each firm has production function. And the production function, in line with the standard NK model, I assume a Cobb-Douglas production with constant return to scale

(A.15)

[52] Follow by Walsh (2003), we also assume , then

Where i denote the firm; [53]is the technology. Under Calvo (1983) contract, each frim re-optimizes its price in every period with probability (1-) and keep its price fixed to the previously set price with probability , have to keep these remain due to menu cost. However, for simplicity, the nominal wage in the labour market are presumed to be fully flexible. And then where we have used the expressions for the product's demand curve,

(A.16)

So, in each period firms producing differentiated goods but processing identical price strategy would set individual prices , subject to the production constraint , the Calvo contract resetting probability is 1- and the demand curve , to maximize the discounted real profits. Then here we let denotes the real marginal cost to each firms' production, and solve the firms' cost minimization problem, we can solve,

(A.17)

(A.18)

Using the Lagrange

(A.19)

Solve the first order condition we get the firms' real marginal costs

(A.20)

[53] with , where is the iid productivity shock.

Such that each firm maximizes the expected discounted sum of future profits to choose an individual

(A.21)

Where is the discount factor, indicating the ratio of marginal utilities of consumption between periods. Then using the demand curve, we can rewrite the firm's maximization problem,

(A.22)

Then the first order condition of firms' maximized equation with respect to individual price implies

(A.23)

Log-linearization of the firm's maximized problem's first order condition, around zero inflation steady state yields the optimal reset price for each firm as follows:

(A.24)

The aggregate price level in each period given the Calvo contract can be written as the weighted average of this up-to-date reset prices and the unchanged, with the weights being the reset probability, and its opposite, respectively, and is this process each individual frim have the same price strategy,

(A.25)

Then log-linearized above equation we can solve

(A.26)

(A.27)

Use equation here we can get

(A.28)

And since the log linearized of real marginal cost is,

(A.29)

Combine with the log-linearized of (which have been solved from first order condition from household side), then we get

$$-$$ (A.30)

Then we can have also solved the real marginal cost as following,

$$- \quad =$$ (A.31)[54]

Then we can get the new Keynesian Phillips Curve is,

(A.32)

[54] The interpretation of is follow by Woodford (2001) as the strategic complementarity between different pricing decisions of different suppliers. Woodford suggest is an empirically plausible value for the US.

Follow by many authors simple adds an additive cost-push-shock after having derived the Philips curve in the standard way.

$$ \text{———} \quad + \qquad (A.33) $$

Government and Monetary Policy (Taylor rule)

Finally, equation (A.34) is the interest rate smoothing rule with a lagged interest rate that has been added into the classic form that is developed by Taylor (1993) to obtain smoothing behaviour.

$$ (A.34) $$

Where is the degree of partially adjustment, is the monetary policy shock. All disturbances , ,and are AR(1) processes with AR coefficients , , and ,

$$ (A.35) $$

$$ (A.36) $$

$$ (A.37) $$

Reference

Akerlof, George A. (2002). "Behavioral Macroeconomics and Macroeconomic Behavior." American Economic Review, 92 (3), 411-433

An, S., & Schorfheide, F. (2007). Bayesian analysis of DSGE models. Econometric reviews, 26(2-4), 113-172.

Aruoba, S. B. (2008). Data revisions are not well behaved. Journal of Money, Credit and Banking, 40(2-3), 319-340.

Aruoba, B., & Schorfheide, F. (2011). Sticky prices versus monetary frictions: An estimation of policy trade-offs. American Economic Journal: Macroeconomics, 3(1), 60-90.

Ball, L. (1994). Credible disinflation with staggered price-setting. The American Economic Review, 84(1), 282-289.

Blinder, A., Canetti, E. R., Lebow, D. E., & Rudd, J. B. (1998). Asking about prices: a new approach to understanding price stickiness. Russell Sage Foundation.

Ball, L. (2000). Near-rationality and inflation in two monetary regimes (No. w7988). National Bureau of Economic Research.

Begg, D. K., & Imperato, I. (2001). The rationality of information gathering: monopoly. The Manchester School, 69(3), 237-252.

Ball, L., Mankiw, N.,G., and R. Reis, (2005), "Monetary Policy for Inattentive Economies," Journal of Monetary Economics 52(4), 703-725.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. Journal of Monetary Economics, 12(3), 383-398.

Clarida, R., Gali, J., & Gertler, M. (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. The Quarterly Journal of Economics, 115(1), 147-180.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. the Quarterly Journal of Economics, 118(1), 269-298.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy, 113(1), 1-45.

Coibion, Olivier, (2006). “Inflation Inertia in Sticky-Information Models,” Contributions to Macroeconomics 6(1), Article 1.

Corradi, Valentina, Andres Fernandez, and Norman R. Swanson (2009). “Information in the Revision Process of Real-Time Datasets.” Journal of Business and Economic Statistics 27, 455–467.

Collard, F., H. Dellas and F. Smets (2009), “Imperfect information and the business cycle”, Journal of Monetary Economics, 56 (1), 38-56

Canova, F., & Sala, L. (2009). Back to square one: Identification issues in DSGE models. Journal of Monetary Economics, 56(4), 431-449.

Collard, F., & Dellas, H. (2010). Monetary misperceptions, output, and inflation dynamics. Journal of Money, Credit and Banking, 42(2-3), 483-502.

Caballero, R. (2010), 'Macroeconomics after the Crisis: Time to Deal with the Pretense of Knowledge Syndrome', Journal of Economic Perspectives, 24(4), 85-102.

Coibion, O. (2010). Testing the sticky information Phillips curve. The Review of Economics and Statistics, 92(1), 87-101.

Croushore, D. (2011). Frontiers of real-time data analysis. Journal of economic literature, 49(1), 72-100.

Coibion, O. and Y. Gorodnichenko (2012). What Can Survey Forecasts Tell Us About Information Rigidities? Journal of Political Economy 120 (1), 116–159.

Coibion, O., & Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. The American Economic Review, 105(8), 2644-2678.

Casares, M., & Vázquez, J. (2016). Data Revisions in the Estimation of DSGE models. Macroeconomic Dynamics, 20(7), 1683-1716.

Diebold, F. X., & Rudebusch, G. D. (1991). Forecasting output with the composite leading index: A real-time analysis. Journal of the American Statistical Association, 86(415), 603-610.

Dynan, Karen (2000) “Habit Formation in Consumer Preferences: Evidence from Panel Data,” American Economic Review, 90, 391-406

Dhyne, E., L. J. Alvarez, H. L. Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, ´ P. Lunnemann, F. Rumler, and J. Vilmunen (2006): "Price Setting in the Euro Area and the United States: Some Facts From Individual Consumer Price Data," Journal of Economic Perspectives, 20(2), 171–192.

Del Negro, M., & Eusepi, S. (2011). Fitting observed inflation expectations. Journal of Economic Dynamics and control, 35(12), 2105-2131.

Dräger, L., (2011). Endogenous persistence with recursive inattentiveness. Working Paper 285. KOF Swiss Economic Institute.
-(2016), Recursive inattentiveness with heterogeneous expectations. Macroeconomic Dynamics, 20(4), 1073-1100.

Dräger. L, Michael J. Lamla b (2013), Updating inflation expectations: Evidence from micro-data L, Economics Letters 117, 807–810

Eichenbaum, M., & Fisher, J. D. (2004). Evaluating the Calvo model of sticky prices (No. w10617). National Bureau of Economic Research.

Joshy, E., & Golinelli, R. (2010). Households Forming Inflation Expectations: Active and Passive Absorption Rates. The BE Journal of Macroeconomics, 10(1), 1-32.

Friedman, M. (1953). Essays in positive economics. University of Chicago Press.

Friedman, B. M. (1979). Optimal expectations and the extreme information assumptions of 'rational expectations' macromodels. Journal of Monetary Economics, 5(1), 23-41.

Fuhrer, Jeff and George Moore (1995). "Inflation Persistence." Quarterly Journal of Economics, 110, 127-160

Fernández-Villaverde, J., & Rubio-Ramı rez, J. F. (2004). Comparing dynamic equilibrium models to data: a Bayesian approach. Journal of Econometrics, 123(1), 153-187.

Fernández-Villaverde, J. (2010). The econometrics of DSGE models. SERIES: Journal of the Spanish Economic Association, 1(1), 3-49.

Fan, J., Minford, P., & Ou, Z. (2016). The role of fiscal policy in Britain's Great Inflation. Economic Modelling, 58, 203-218.

Fuhrer, J. (2017). Expectations as a source of macroeconomic persistence: Evidence from survey expectations in a dynamic macro model. Journal of Monetary Economics, 86, 22-35.

Gregory, A. and Smith, G. (1991). Calibration as testing: Inference in simulated macro models. Journal of Business and Economic Statistics, 9 (3), 293-303.

Gregory, A. and Smith, G. (1993). Calibration in macroeconomics, in: Maddala, G.(Ed.), Handbook of Statistics,11, Elsevier, St. Louis, Mo., 703-719

Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect inference. Journal of Applied Econometrics, 8, 85-118.

Gallant, A. R., & Tauchen, G. (1996). Which moments to match? Econometric Theory, 12(4), 657-681.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: inference, development, and communication. Econometric reviews, 18(1), 1-73.

Gali., J. and M. Gertler, (1999), "Inflation Dynamics: A Structural Econometric Analysis," Journal of Monetary Economics, 44, 195-222.

Galí, J. (2002). New perspectives on monetary policy, inflation, and the business cycle (No. w8767). National Bureau of Economic Research.

Gali J, López-Salido J D, Vallés J. (2003) Technology shocks and monetary policy: assessing the Fed's performance[J]. Journal of Monetary Economics, 50(4), 723-743.

Geweke, J., & Whiteman, C. (2006). Bayesian forecasting. Handbook of economic forecasting, 1, 3-80.

II, K., & Edward, S. (2010). A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment. Journal of Money, Credit and Banking, 42(8), 1543-1564.

Jeffreys, H. (1961). Theory of probability (3rd ed.). New York, NY: Oxford University Press.

Jones, B. D. (2002). Bounded rationality and public policy: Herbert A. Simon and the decisional foundation of collective choice. Policy Sciences, 35(3), 269-284.

Joshy Easaw and Roberto Golinelli (2014): "Inflation Expectations and the Two Forms of Inattentiveness," Cardiff Economic Working paper

Khan, H., & Zhu, Z. (2002). Estimates of the Sticky-Information Phillips Curve for the United States, Canada, and the United Kingdom (No. 02-19). Bank of Canada.
-(2006), Estimates of the sticky-information Phillips curve for the United States. Journal of Money, Credit, and Banking, 38(1), 195-207.

Keane, M., & Smith, A. A. (2003). Generalized indirect inference for discrete choice models. Yale University.

Keen, B. D. (2007). Sticky Price and Sticky Information Price-Setting Models: What is the difference?, Economic Inquiry, 45(4), 770-786.

Kiley, M. T. (2007). A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting. Journal of Money, Credit and Banking, 39(s1), 101-125.

Klenow, Peter J., and Jonathan L. Willis. (2007). "Sticky Information and Sticky Prices." Journal of Monetary Economics, 54 (10), 79–99.

Korenok, O., & Swanson, N. R. (2007). How sticky is sticky enough? A distributional and impulse response analysis of new Keynesian DSGE models. Journal of Money, Credit and Banking, 39(6), 1481-1508.

Knotek, Edward S. (2010). "A Tale of Two Rigidities: Sticky Prices in a Sticky Information Environment." J. Money, Credit and Banking 42 (12), 1543– 64.

Lown, Cara S. and Robert Rich. (1997). "Is there an inflation puzzle?" Federal Reserve Bank of New York Economics Policy Review, 3(4), 51–69.

Lubik, T., & Schorfheide, F. (2005). A Bayesian look at new open economy macroeconomics. NBER macroeconomics annual, 20, 313-366.

Le, V. P. M., Meenagh, D., Minford, P., & Wickens, M. (2010). Two orthogonal continents? testing a two-country DSGE model of the US and the EU using indirect inference. Open Economies Review, 21(1), 23-44.

Le, V. P. M., Meenagh, D., Minford, P., & Wickens, M. (2011). How much nominal rigidity is there in the US economy? Testing a New Keynesian DSGE Model using indirect inference. Journal of Economic Dynamics and Control, 35(12), 2078-2104.

Le, V. P. M., Meenagh, D., Minford, P., & Ou, Z. (2013). What causes banking crises? An empirical investigation for the world economy. Open Economies Review, 24(4), 581-611.

Levine, P., Pearlman, J., Perendia, G., & Yang, B. (2012). Endogenous persistence in an estimated DSGE model under imperfect information. The Economic Journal, 122(565), 1287-1312.

Liu, C., & Minford, P. (2014). Comparing behavioral and rational expectations for the US post-war economy. Economic Modelling, 43, 407-415.

Le, V. P. M., Meenagh, D., Minford, P., Wickens, M., & Xu, Y. (2016). Testing macro models by indirect inference: a survey for users. Open Economies Review, 27(1), 1-38.

McCallum, B. T. (1976). Rational expectations and the natural rate hypothesis: some consistent estimates. Econometrica: Journal of the Econometric Society, 44(1), 43-52.

Mankiw and Reis (2002). STICKY INFORMATION VERSUS STICKY PRICES: A PROPOSAL TO REPLACE THE NEW KEYNESIAN PHILLIPS CURVE. The Quarterly Journal of Economics, 117(4), 1295-1328.

Mankiw, N. Gregory, Ricardo Reis and Justin Wolfers (2003). "Disagreement about Inflation Expectations." NBER Macroeconomics Annual, 18, 209-248.

Mankiw, N. Gregory, and Ricardo Reis (2006). "Pervasive Stickiness." American Economic Review, 96(2), 164-169.

Mankiw, N. G., & Reis, R. (2007). Sticky information in general equilibrium. Journal of the European Economic Association, 5(2-3), 603-613.

Meenagh, D., Minford, P., & Wickens, M. (2008). Testing a DSGE model of the US using indirect inference. Cardiff University. Cardiff Economics Working Papers 2008/11.
-(2009), Testing a DSGE model of the EU using indirect inference. Open Economies Review, 20(4), 435-471.

Menz, J. O., & Vogel, L. (2009). A Detailed Derivation of the Sticky Price and Sticky Information New Keynesian DSGE Model (No. 2/2009). DEP (Socioeconomics) Discussion Papers, Macroeconomics and Finance Series.

Minford, P., Theodoridis, K., & Meenagh, D. (2009). Testing a model of the UK by the method of indirect inference. Open Economies Review, 20(2), 265-291.

Meyer-Gohde, A. (2010). Linear rational-expectations models with lagged expectations: A synthetic method. Journal of Economic Dynamics and Control, 34(5), 984-1002.

Milani, F. (2012). The modeling of expectations in empirical DSGE models: A survey. In DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments (pp. 3-38). Emerald Group Publishing Limited.

Milani, F., & Rajbhandari, A. (2012). Expectation formation and monetary DSGE models: Beyond the rational expectations paradigm. In DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments (pp.253-288). Emerald Group Publishing Limited.

Minford, P., & Ou, Z. (2013). Taylor Rule or optimal timeless policy? Reconsidering the Fed's behavior since 1982. Economic Modelling, 32, 113-123.

Minford, P., Ou, Z., & Wickens, M. (2015). Revisiting the Great Moderation: policy or luck?. Open Economies Review, 26(2), 197-223.

Mitchell, K., & Pearce, D. (2015). Direct Evidence on Sticky Information from the Revision Behavior of Professional Forecasters (No. 66172). University Library of Munich, Germany. -(2017). Direct Evidence on Sticky Information from the Revision Behavior of Professional Forecasters. Southern Economic Journal, 84(2), 637-653.

Nakamura, E., & Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. The Quarterly Journal of Economics, 123(4), 1415-1464.

Ormeño, A., & Molnár, K. (2015). Using survey data of inflation expectations in the estimation of learning and rational expectations models. Journal of Money, Credit and Banking, 47(4), 673-699.

Peersman, G., & Smets, F. (2002). Are the effects of monetary policy in the euro area greater in recessions than in booms?. In Conference on Transmission Mechanism of Monetary Policy (pp. 28-48). Cambridge University Press.

Paustian, M., & Pytlarczyk, E. (2006). Sticky contracts or sticky information? Evidence from an estimated Euro area DSGE model. Manuscript, Bowling Green State University.

Paloviita, M (2007b) Estimating a small DSGE model under rational and measured expectations: some comparisons. Bank of Finland Research Discussion Papers, No. 14/2007.
- (2008). Comparing alternative Phillips curve specifications: European results with survey-based expectations. Applied Economics, 40(17), 2259-2270.

Rotemberg, J. J., & Woodford, M. (1996). Real-business-cycle models and the forecastable movements in output, hours, and consumption. The American Economic Review, 86 (1), 71-89.

Rabanal, P., & Rubio-Ramírez, J. F. (2005). Comparing New Keynesian models of the business cycle: A Bayesian approach. Journal of Monetary Economics, 52(6), 1151-1166.

Reis, R. (2006a). Inattentive consumers. Journal of monetary Economics, 53(8), 1761-1800.

Reis, R. (2006b). Inattentive producers. The Review of Economic Studies, 73(3), 793-821.

Reis, Ricardo (2009). “Optimal Monetary Policy Rules in an Estimated Sticky-Information Model.” American Economic Journal: Macroeconomics, 1(2), 1-28.

Simon, H. A. (1989). Models of Thought, Volume II. New Haven: Yale University Press.

Sargent, T. J. (1993). Bounded rationality in macroeconomics: The Arne Ryde memorial lectures. OUP Catalogue.

Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. Journal of Applied Econometrics, 15(6), 645-670.

Sims, C. A. (2003). Implications of rational inattention. Journal of monetary Economics, 50(3), 665-690.

Smets, F., & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. Journal of the European economic association, 1(5), 1123-1175.

Schorfheide, F. (2005). Learning and monetary policy shifts. Review of Economic dynamics, 8(2), 392-419.

Smets, F., & Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. The American Economic Review, 97(3), 586-606.

Stiglitz, J. E. (2011). Rethinking macroeconomics: What failed, and how to repair it. Journal of the European Economic Association, 9(4), 591-645.

Taylor, John B. (1993). "Discretion versus policy rules in practice." Carnegie-Rochester Conference Series on Public Policy 39, 195-214.

Trabandt, Mathias (2004). “Sticky Information vs. Sticky Prices: A Horse Race in a DSGE Framework.” Mimeo, Humboldt University Berlin.
- (2005). “Sticky Information vs. Sticky Prices: A Horse Race In A DSGE Framework.” Manuscript.
-(2007). Sticky information vs. sticky prices: A horse race in a DSGE framework.

Wickens, M. R. (1982). The efficient estimation of econometric models with rational expectations. The Review of Economic Studies, 49(1), 55-67.

Wilson, E. B., & Hilferty, M. M. (1931). The distribution of chi-square. Proceedings of the National Academy of Sciences, 17(12), 684-688.

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy (No. w8673). National Bureau of Economic Research.
-(2003), Imperfect Common Knowledge and the effects of monetary policy, Knowledge, Information, and Expectation In Modern Macroeconomics: In Honor of Edmund S. Phelps. (pp.5)

Woodford, M., (2003), Interest and Prices: Foundations of a Theory of Monetary Policy, Princeton University Press, Princeton, NJ.

Walsh, C, E (2003), Monetary Theory and Policy (3rd), Cambridge and London: MIT Press.

Vázquez, J., María-Dolores, R., & Londoño, J. M. (2010). Data Revisions and the Monetary Policy Rule: An analysis based on an extension of the basic New Keynesian model. The University of the Basque Country, Mimeo.
- (2012). The effect of data revisions on the basic New Keynesian model. International Review of Economics & Finance, 24, 235-249.

Vetlov, I., Hlédik, T., Jonsson, M., Kucsera, H., & Pisani, M. (2011). Potential output in DSGE models (No. 1351). ECB Working Paper.