Cardiff Economics Working Papers

Working Paper No. E2022/2

Evaluation and Indirect Inference Estimation of Inattentive Features in a New Keynesian Framework

Jenyu Chou, Yifei Cao, and Patrick Minford

January 2022

ISSN 1749-6010

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

Evaluation and Indirect Inference Estimation of Inattentive Features in a New Keynesian Framework*

Jenyu Chou†, **Yifei Cao**†, and **Patrick Minford**‡

† School of Economics, University of Nottingham Ningbo China

‡ Economic Section, Cardiff University Business School, Cardiff, UK

December 2021

Abstract

We test the standard New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model under the condition with and without inattentive features, where inattentiveness is modelled in the form of sticky information and imperfect information data revision. All models are tested with the Indirect Inference method, and our test result based on real-time data suggests that the model with sticky information passes the test and consistently outperforms the baseline NK model with full information and rational expectation, while the model with imperfect information data revision fails to pass the test. Furthermore, we show that none of the models passes the test when Survey of Professional Forecaster data are used for model evaluation. Overall, our findings provide important implications on the modelling of expectation formation in the DSGE framework.

Keywords: Inattentive expectation, New Keynesian, DSGE, Indirect Inference

JEL Classficiation:E12,E52,C52

*This paper has benefited from comments and suggestions from the seminar participants at the Cardiff University Business School. Any remaining errors are the sole responsibility of the authors. Any remaining errors are the sole responsibility of the authors. E-mail addresses: jen-yu.chou@nottingham.edu.cn (J. Chou), yifei.cao@nottingham.edu.cn (Y. Cao), minfordp@cardiff.ac.uk (P. Minford).

1 Introduction

The role of people's expectation in determining aggregate outcomes of the macro economy, such as inflation dynamics and business cycle, has often been discussed and well-established. However, studies involving how people form their expectation are relatively rare. One recent study by Milani and Rajbhandari (2012) compare a full-information rationality New-Keynesian-type model with alternative models that deviate from full-information rationality.[1] However, this topic is important for making the most fundamental macroeconomic decisions, such as the allocation of consumption or savings and setting an appropriate price, some of which are underlying macroeconomic dynamics and driven by people's expectation of the future. In the following sections, we initially survey the literature focusing on the early assumption of fully attentive expectation or full-information rational expectation (FIRE) and then explore the weakness of this early expectation assumption. To address the weakness of FIRE assumption, another assumption deviating from the full-information rationality has been proposed, namely, inattentive expectation assumption. Particularly, we mainly focus on two types of inattentiveness, which are the most commonly discussed. The first is the model with sticky information (SI), and the assumption of SI is borrowed from Mankiw and Reis (2002, 2007). The second popular inattentiveness is imperfect information (IF) data revision (Aruoba, 2008; Vázquez et al., 2010, 2012; Casares and Vázquez, 2016). Both inattentive assumptions will be well-stated and discussed in the later sections.

The FIRE hypothesis is the starting point of the traditional economic theory. However, a gap between this classical New Keynesian FIRE (Calvo, 1983) and the real world has been criticised for many economies. Simon (1989) criticises the 'unrealistic' view of the idea of FIRE. He argues that regarding the case of economic agents, having known all of their problems, choices and possible results, economic agents can certainly choose the best solution from all alternatives through some reasonable calculation. However,

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

in practice, such a 'perfect situation' cannot exist in the real world. Moreover, 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 reach the self-satisfied or 'good enough' solution. As a result, the FIRE can hardly be applied to explain economic problems.

On the other hand, the implicit hypothesis of FIRE is that economic agents are homogeneous. However, in the 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. In sum, the unrealistic features of the early assumption of FIRE can be shown from two aspects as follows:

1) The FIRE hypothesises that economic agents have 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 understand and solve complex problems but in a restricted manner.

2) Under the assumption of FIRE, information is a type 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 does not consider information costs (i.e. costs of accessing required information). Agents have to pay while collecting the information required for decision making. In practice, obtaining and processing information without the payment of time, money or physical effort is impossible. 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 cannot reach the best situation.

In sum, under the assumption of FIRE, economic agents are supposed have clear values of the relevant parameters (e.g. shock distribution and correct structure of economic model). However, it is an unreasonable assumption in practice because economic agents cannot hold all the information needed to reach the equilibrium of the entire economy (Caballero, 2010). Particularly, when an economy undergoes a remarkable structural transformation (e.g. the Great Recession), it will need new 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 the full-information rationality hypothesis. They also clarify that the reason for rejecting the full-information rationality hypothesis is not the rationality but the assumption of full information.

2 Three Competitors

To address the unrealistic aspect of the early FIRE and the well-known empirical weaknesses (i.e. the delay effect of monetary shock on inflation, persistent output and inflation observed in macro data), the New-Keynesian-type model with the features deviated from the FIRE appears as a modified version.[2] Thus, the inattentive expectation has been proposed. As inattentive expectation has different approaches, the two most prominent are SI (Mankiw and Reis, 2007) and IF data revision (Casares and Vázquez, 2016). These assumptions will be applied in the current study. Different from the model with SI from Mankiw and Reis (2007) and the model with IF data revision from (Casares and Vázquez, 2016), we use a small-scale closed-economy dynamic stochastic general equilibrium (DSGE) model instead of a medium-scale DSGE to be in line with the baseline model selected.

[2]Some studies have focused on how to compensate the impractical aspects of full-information expectation New-Keynesian-type models through multiple ways (Rotemberg and Woodford, 1996; Galı and Gertler, 1999; Smets and Wouters, 2003, 2007). In these studies, most attention is received on real rigidities, such as habit persistence, capital or investment costs, capital utilization and backwards-looking price setting schemes for the subset of 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. According to the European Central Bank Report, individual price movement are is not consistent with the movement of aggregate inflation. In explaining the observed situation, the idea of reducing controversy, which encourages scholars to continue making efforts to resolve these issues, has been adopted in the past few years.

Although the full-information rationality has weaknesses, as recent studies suggest, its assumption of rationality need not be ignored or other types of irrational behavior need not be introduced to help the model fit the data (Collard et al., 2009; Coibion and Gorodnichenko, 2012). Thus, in this study, two major inattentive rational models, SI and IF data revision models, are used and compared while assuming rationality.

In the above literature, three relevant models can be divided into two groups, namely, with and without inattentive features. The first model is the classical 'attentive' expectation model, which is a New-Keynesian-type model with full-information rationality hypothesis. The second model is the SI model. The third model is the IF data revision model. Three objectives will be reached by comparing the three models under different conditions.

The first objective of this study 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 performance of the model is measured by checking its ability to generate persistent and delayed responses on the output (output gap) and inflation to monetary policy (Christiano et al., 2005). Model simulations are performed using Dynare 4.4.3 software.[3]

The second objective is to compare which expectation-type model best explains the US economy by using quarterly real-time data (survey of professional forecaster (SPF) data are used for robustness check). The process is implemented through indirect inference to evaluate each competing model and make model comparison in an absolute manner. Although the Bayesian approach provides a simple way to compare the relative performance of different models, it cannot be used to evaluate the models' performance in an absolute manner due to its limitation of judging whether a to-be-examined model has a satisfactory performance, which can be verified by the actual data. The indirect inference estimation

[3]Standard DSGE models with Dynare code are provided in http://vermandel.fr/dsge-dynare-model-matlab-codes/, including the simple dynamic three-equation New Keynesian model.

(estimation-based indirect inference test) is distinguished from the Bayesian estimation method by generating a data descriptor that indirectly evaluates the theoretical model by using a completely independent auxiliary model (e.g. vector autoregression (VAR). The estimation-based indirect inference test is implemented in discovering the optimal set of parameters of the actual data in the context of the model to make a fair model comparison.

The overview of each of the attentive and inattentive models is specified as follows. The three competing modes is a reduced-form New Keynesian-type DSGE model for a small-scale closed economy. The economy consists of three types of agents, namely, households, firms and monetary authorities. The baseline model, which 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 competitors, one is the model characterised by SI, which has been discussed in Mankiw and Reis (2007), and the other is the model characterised by IF data revision, which has been constructed by Casares and Vázquez (2016). Different from the two inattentive expectation model settings, we use a small-scale DSGE model instead of a medium-sized one. Adding additional features might be a useful step (Smets and Wouters, 2003, 2007). However, it may also cause some fundamental issues that can blur our main focus. Specifically, when a model includes additional new features, it may potentially be distracted from its original focus, which leads to the difficulty in assessing the differences between the two inattentiveness (i.e. SI and IF data revision). Another difference between the baseline model and the models with inattentive features is the consideration of many features.

[Insert Table 1]

The presented three competitors above indicate that the aggregate economy under the reduced-form New Keynesian framework can be characterised by the dynamics of three main economic variables (i.e. output gap, inflation and interest rate). x_t represents the output gap, which is a difference between the actual and potential outputs (i.e. output under a flexible price economy). Coefficient σ represents the elasticity of the intertemporal substitution. The new Keynesian PC derived under the FIRE is equivalent to the current inflation π_t driven by the expectation of future inflation $E_t\pi_{t+1}$, the current output gap x_t and the supply shock u_t. Coefficient β. represents the time discount factor, and γ is the combined parameter.[4] The interest rate equation follows the simple 'interest rate smoothed' Taylor rule (Taylor, 1993). Monetary policy makers set the interest rate based on the simple Taylor rule. The interest rate $\tilde{r_t}$ is driven by the current inflation π_t and current output gap x_t.

Thus, on the basis of the model with SI, the two parameters δ and λ are the shares of updating households and of updating firms in any given period, respectively (for example, if no SI of firms exist, then $\lambda = 1$). To compare with the economic agents in the baseline model without inattentive features, we assume that the economic agents under the premise of SI economy update their information sets with certain rate δ and λ regarding households and firms, respectively (Mankiw and Reis, 2002, 2007; Reis, 2006a,b, 2009). Reis (2006a,b) provides deeper micro-foundations for model features SI. The early classical New-Keynesian-type model is based on the assumption of FIRE, which is the case of delivering a pure forward-looking-expectation PC. However, under an SI 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, yielding the PC 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) .[5] When looking

[4]In $\gamma \equiv \chi + \sigma^{-1}$, the composite parameter $\gamma = 0.15$ has been taken as fixed and less than one, which implies strategic complementarity, to keep it as fixed and less than 1 and in line with the suggestion from previous literature (Woodford, 2001; Ball et al., 2005). Woodford (2003) surveys and discusses the existing literature at length and concludes that firms' pricing decision should be strategically complementary rather than strategic substitutes to allow for potential inflation inertia. This approach has been tested in some recent works; for example, Coibion et al. (2006) claims that when $\gamma > 1$, inconsistent results are produced with the actual data.

[5]Different from the SI PC model of (Mankiw and Reis, 2002), the current inflation in our New Keynesian three-equation model is determined by the current expectation and the past expectation of the future inflation rate. By contrast, the current inflation in Mankiw and Reis' model is inferred from flexible price assumption.

into the previous empirical literature, several studies have aimed to compare the PC derived under the assumption of FIRE and alternative under the SI assumption (Mankiw and Reis, 2002; Coibion and Gorodnichenko, 2012, 2015). However, in the present study, in terms of empirical evidence, we are more interested in the simple reduced-form New-Keynesian-type DSGE models rather than that based on a single equation (Easaw and Golinelli, 2010; Coibion and Gorodnichenko, 2015). Estimation of comprehensive DSGE models by introducing inattentive feature exists, but only few papers have conducted it. Recent papers on this aspect have set a benchmark of neoclassical model with flexible prices and introduced SI 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. SI assumption versus IF data revision assumption).[6] Thus, in this study, our main emphasis is to use the model with SI to compare with the alternative inattentive expectation model (i.e. the model with IF data revision) to examine which inattentive expectation model can yield better explanation for the US economy in the recent five decades (the sample period of the US quarterly data is from 1969 to 2015).

In comparison with the baseline model, the model with SI is more challengeable to solve. Given that SI involves infinitely lagged expectations, we question how we can approximate the model with SI in the DSGE equilibrium framework. Firstly, from the angle of SI model setting, the proportion of lagged expectations diminish geometrically. In other words, the effect on economic agents' expectation derived from the current state is far greater than that of previous periods. Consequently, the expectations that are formed extremely far from the present situation might not influence the current inflation or output gap due to the minimal weight (i.e. may be approximately zero) attached to them. Thus, we set $j = 4$ as the benchmark, which indicates the incorporation of lag

[6]From an empirical point of view, Smets and Wouters (2007) claim that a more satisfying specification may consider some frictions. However, in this study, we aim to keep it simple, because one of the main questions we would like to focus on differentiating the various inattentive features and seeing whether different inattentive features matter for the dynamics of the economy.

information up to four periods); longer periods, such as $j = 6$ and 8, are considered in the robustness check.[7]

Real-time and revised data are used for the extended model with IF data revision, as suggested by previous studies (Casares and Vázquez, 2016; Vázquez et al., 2010). Before introducing the IF data revision model, we must initially know what the real-time data are. For example, if we analyse the economic agents' decision using the data available to us today, then we will make an incorrect inference about their economic decision making. If we look at the time that economic agents made their economic decisions, then we are engaging in real-time analysis or taking considering the data revision seriously.

Data revision is potentially critical theoretically and empirically, although many economic researchers have made an inappropriate assumption about the data available to economic agents at each point in time. The applied data assumption is that they are available immediately, yet the reality is that those data are announced with a few lags. Furthermore, data revision, in general, has been thought either not to exist or is small, but in reality, data revision may have a significant influence on empirical results, which is particularly the case of some variables that are defined conceptually. For instance, for output gap, when economic agents are making decisions, they take this variable without any doubt. In a real stat, such variable as the output gap often fluctuates over time. Thus, in the IF model, data revision is considered to see how it affects the New-Keynesian-type macroeconomic model and the empirical results.

Moreover, we follow the suggestion by Casares and Vázquez (2016) for the data revision, as specified in the Appendix. In addition to the above discussion, another two points should be clarified. 1) Under the IF data revision hypothesis, the information about the real state of the economy matters. For example, firms' price-setting decision depends on the expectation of marginal revenue and the future nominal marginal costs. Thus, it depends on the future aggregate price level. 2) The information friction or inattentive

[7] From the result of Trabandt (2007), by setting maximum $j = 19$, the convergence of the recursive equilibrium law of motion can be achieved for the SI PC model. However, the SI model uses fewer periods j, which cannot sufficiently reach convergence.

features highlighted in this study must be taken seriously; such inattentive assumption needs to be reasonable. The nominal interest rates made through professional monetary authority are fully observable without noise disturbance, and the observation of output gap and inflation are influenced by noises. In other words, both variables are involved in the data revision. Collard and Dellas (2010) argue that few aggregate variables can be observed accurately , as the data revision process reveals. That is, under the assumption of IF, firms cannot fully observe its information when making a price-setting decision; similarly, when households make a consumption decision, they cannot fully observe the state to support them to make a consumption plan. Particularly, price (inflation) and consumption (output) can only be observed with some random noises. From the above three-equation models, where x_t^r and π_t^r are taken as the observed variables realised at time t, they are the real-time data. x_t and π_t are the final revised variables.

For each model with and without inattentive feature, the AR (1) process is assumed for all the disturbances to each structural equation to capture omitted variables. In addition, the frequency of each variable is quarterly, each variable is a demeaned variable, and detrend data are applied. The three models have different information friction constraints, thereby having different IS and PC, which may influence the monetary policy. Then, by comparing their model performance (i.e. transformed Mahalanobis (TM) distance), whether the suggestion of incorporating inattentive features from the previous literature can provide a better explanation for the US economy relatively should be determined. Moreover, whether different inattentive features matter should be explored to explain the economy dynamics.

3 Estimation through Indirect Inference

In this study, indirect inference is applied to measure how close the three models are to the real world. The principle of this method is based on the idea that a model can be measured in an absolute manner in a framework that contains an auxiliary model by comparing the moments of simulated and actual data. 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, qualified and unqualified. Secondly, it enables us to measure the distance statistically between the theoretical models (model-simulated data) and the real world (actual data).

The approach of indirect inference has been applied widely in the field of estimation (Gregory and Smith, 1991; Gallant and Tauchen, 1996; Keane and Smith, 2003; Minford et al., 2009). For instance, Le et al. (2011) apply the same method to evaluate the model of the US economy, which is constructed by Smets and Wouters (2007), and ultimately obtain a rejected consequence on the testing. In the present work, we take the common procedure of indirect inference evaluation for reference from previous studies Le et al. (2011); Liu and Minford (2014); Minford et al. (2015).

Notably, two relevant papers are available regarding our research topic that uses the 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. Estimation conducted through indirect inference indicate that the ignorance of the data revision process may not result in a serious drawback in analysing a monetary policy based on a New Keynesian framework. Our assumption substitutes the subjects who involve IF data revision issue with households and firms instead of monetary authority. Moreover, the subjects can perfectly observe the monetary policy. The other related paper is published by KNOTEK II (2010), who investigate a single-equation model incorporating sticky price and SI. They find that such a model can match the real world in micro and macrodimensions after estimating it through indirect inference.[8] However, we are more interested in a full structural model rather than a single-equation model.

In this study, we evaluate each model, focusing on its overall dynamic properties in connecting with the actual data by adopting indirect inference as the new evaluation

[8]KNOTEK II (2010) find that when the empirical PC is embodied with sticky prices and SI; its ability tends to be improved to match the macro data.

method. While applying indirect inference to evaluate an existing structural model, two factors are inevitable in the process of stimulating the data from the theoretical model. One is the parameters of the theoretical model, and the other is the distribution of errors. We evaluate the theoretical model through indirect inference test, which is based on the comparison of the observed actual data with the data simulated from the theoretical model with the assistance of an auxiliary model. In this study, VAR, which is a stochastic process model used to capture the linear inter-dependencies amongst multiple time series, is selected as the auxiliary model for two reasons. Firstly, the structural model can always be represented 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 amongst the variables through the co-variance matrix of the VAR disturbances, and the dynamic behavior 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, which is 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, then the theoretical model cannot reproduce the actual data significantly. Conversely, the consequence of being accepted implies that the data generated from the theoretical model do not significantly differ from the actual observed data.

Wald Test Statistics

In general, the Wald testing process can be summarised into three general steps as follows. Firstly, the observed actual data and parameters calibrated or estimated in the model are used to derive the structural errors. The errors can be constructed under two different circumstances. When the structural model possesses no expectation terms, the structural errors can be backed directly from the structural equations and the actual data. Under the situation that structural equation includes the computation of expectations, the method used is the robust instrument variables estimation suggested by Wickens (1982), in which the lagged endogenous data are set as instruments, and the fitted values are computed from a VAR (1), which is also used as the auxiliary model during the evaluation procedure. 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 used to produce the pseudo data that are based on the candidate theoretical model. An auxiliary VAR model is then fitted to each set of pseudo data, and the sampling distribution of the coefficients of the auxiliary VAR model is achieved from these estimates of the auxiliary model. Thirdly, the Wald statistic is computed to determine whether the functions of the parameters of the auxiliary VAR model estimated on the actual data lie in the confidence interval implied by this sampling distribution[9] of the coefficients of the auxiliary time series model (Minford et al., 2015).

The test is conducted by comparing the performance of the overall capacity of the model with the dynamic performance of the actual data to determine whether the hypothesis is qualified. The comparison is performed by checking if the coefficients based on the actual data-based VAR lie in the acceptable range of the theoretical model's implied joint distribution. Then, we can examine the model's ability of directing the dynamics and variances of the data.

In this study, VAR (1) is used as the auxiliary model and is treated as the descriptors of the actual data for the three main macro variables (i.e. output gap, inflation and interest rate). The Wald statistics is computed from the VAR (1) coefficients and the

[9]We can obtain the estimate distribution by estimating the auxiliary VAR model on each pseudo sample. The dynamic properties are captured by VAR estimates, whereas 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.

three variances of the three main variables. Therefore, the Wald test statics is a criterion to determine whether the observed dynamics and volatility of the selected three main variables are explained by their simulated joint distribution at a given confidence level (95%). The Wald statistics can be expressedde as follows.

$$Wald\ test\ statistics = [G_T(\alpha_T) - G_S(\bar{\alpha}_S(\theta)]' W[G_T(\alpha_T) - G_S(\bar{\alpha}_S(\theta)] \tag{1}$$

The equation above is a function of the gap between $G_S(\bar{\alpha}_S(\theta))$ and $G_T(\alpha_T)$. $G_T(\alpha_T)$ is the vector of VAR estimates of the selected US data descriptors. $G_S(\bar{\alpha}_S(\theta))$ is the arithmetic mean of the N estimated vector of VAR estimates derived from bootstrap simulations. W is the variance and covariance matrix of the distribution $G_T(\alpha_T) - G_S(\bar{\alpha}_S(\theta))$. α_T and $\alpha_S(\theta)$ are the actual and simulated data sets, respectively. θ 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 the model.

Indirect inference can be proceeded by comparing the percentile of the Wald distribution. Specifically, for a 5% significant level, a percentile above 95% will not lie outside the non-rejection area. The distribution of $G_T(\alpha_T) - G_S(\bar{\alpha}_S(\theta))$ and the Wald statistics are obtained through bootstrapping method.

TM Distance (normalised t-statistics)

The TM statistic is used when the models' relative performance is difficult to distinguish. For instance, when two or more specified models are rejected simultaneously by Wald test statistics, we have to use the TM statistic to rank these models after comparison. In addition, the TM provides a way to examine how poor the model performs by observing how far it deviates away from 1.645. The larger the number is, the worse the model fit will be. The TM distance is defined as follows.

$$TM\ distance(normalised\ t-statistic) = \frac{(\sqrt{2WS_a} - \sqrt{2p})}{(\sqrt{2WS_{s95\%}} - \sqrt{2p})} * 1.645 \tag{2}$$

This function of TM distance is based on Wilson and Hilferty's (1983) method of transforming Chi-square distribution into a standard normal distribution calculated. Herein, the TM distance is the transformation of the Wald test statistics. WS_a is the Mahalanobis distance (value of Wald statistics) using the actual data, $WS_{s95\%}$ 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 degrees of freedom.

In practice, we can reduce the parameter uncertainty directly by checking the Wald statistic derived from the set of parameters for the model. Specifically, the more the Wald statistic decreases, the better the parameter set will perform. Herein, an effective algorithm based on simulated annealing (SA) is introduced to search the optimal parameter set by starting from a wide range around the initial values along with random jumps around the space. With this algorithm, we can have the minimum-value full Wald statistic for the three competing models.

The SA algorithm refers to a stochastic optimisation 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 minimise the defects of crystal (in indirect inference estimation procedure, which is similar to the step of minimising Wald statistics). It attempts to find the optimal parameter set repeatedly until the system reaches a minimum value of Wald statistics, or until a given computation budget is exhausted. Given 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.

Overall, in the application of indirect inference estimation, SA is used to seek the optimal set of parameters, which will facilitate the lowering of Wald statistic until the computation budget is used. Initial values of the structural model's parameters are required in performing the numerical iterations to minimise the Wald statistics. Here, the starting values are the values of the presumptive parameters. Such presumptive parameters are plausible and are based on previous studies. We also allow the parameters to seek around -0.5 to $+0.5$ of their starting values under estimation.

To implement the estimation-based indirect inference test, the VAR (1) needs to be used continuously as the auxiliary model to provide a reference substance for the estimated models to those of the calibrated models. The VAR (1) is used as descriptors of the coefficient matrix and the variance of the data. The SA mechanism will begin to explore from these initial presumptive values to substitute 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 SA method, which adjusts the initial presumptive values (calibrated values), is helpful for the models to pass the test.

Estimation-based Indirect Inference Testing Results: FIRE Model

The SA estimation-based test and the Bayesian estimation-based test with respect to the three competing models for US economy are presented in Tables 2 and 4. 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 exploration is limited within $\pm 50\%$ 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 can generate the actual data. However, if these initial presumptive parameters cannot be used to explain the generating process of the actual data, then another set of parameters may exist and can be applied to explain how the actual data are generated. If the model with initial presumptive parameters already fall within the non-rejection scope, then another set of parameters that can narrow the gap between the theoretical model and the reality should be explored, which leads to better testing results. The 'best' set of parameters for the structural model are those to the maximum degree to shorten the distance between the theoretical model and the reality.

In the indirect inference estimation stage, we aim to explore the 'best' collection of parameters throughout the entire parameter space by implementing indirect inference without changing the signs of the parameters as an estimation-based test approach. The minimised value of the Mahalanobis distance is captured for each competitor over the US sample periods through an SA 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 2 displays the estimation results of the FIRE model. Overall, the estimated values of parameters of the FIRE model through indirect inference estimation are not significantly far away from those obtained by Bayesian estimation. However, some distinguished cases exist. Particularly, the estimated value of the elasticity of inter-temporal substitution is 0.5180, which is quite higher than that obtained from the Bayesian estimation. Moreover, the same trend can be found in the value of price stickiness versus that of Bayesian estimation. After examining the estimates of the main behavioral parameters of the FIRE model, we examine the parameters of the monetary policy function, which are based on standard interest rate smoothed Taylor rule (1993). Regarding the estimated coefficients of monetary policy, except for χ_π, which is increased by $<8\%$, the other two (i.e. ρ and χ_x) increase around 35% compared with 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 detail, SA estimation indicates that the estimated value of χ_π is 1.5079, which is slightly higher than that obtained by Bayesian estimation. The two estimates regarding different estimation methods are close to the initial calibration value (i.e. 1.5). The estimated value of the reaction to the output gap χ_x is 0.1439, which is lower than that obtained by Bayesian estimation. Hence, the monetary policy does not appear to react strongly to the output gap level. Moreover, for the interest rate smoothness, the coefficient ρ on the lagged interest rate 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). Furthermore, the AR coefficients regarding the three exogenous stationary shocks (i.e. demand shock, cost-push shock and monetary policy shock) are estimated to be persistent (0.8587,0.7318 and 0.8155, respectively).

Furthermore, the test statistic implies a Wald percentile of 64.8. Thus, 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 inter-temporal substitution σ 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). Table 4-5 shows 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 gained significant reduction compared with the one obtained using Bayesian estimates. The full Wald statistics implies that the FIRE model with SA estimates fall within the non-rejection area; thus, the model cannot be rejected at a 95% chance. Furthermore, the model with Bayesian estimates perform worse than the model with the initial presumptive parameters (calibration parameters).

[Insert Table 2]

Estimation-based Indirect Inference Testing Results: SI Expectation Model

Table 3 displays the estimation results of the model with SI. Overall, most estimates through SA estimation are higher than those obtained from Bayesian estimation, except for the estimate of interest rate smoothed parameter ρ (0.7672), which is slightly lower than that obtained through Bayesian estimation. The reaction parameter of the output gap in monetary policy χ_x is estimated to be approximately 13%, which is lower than that in Bayesian estimates but 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 ρ_r, which is two times higher than that obtained through Bayesian estimation.

Furthermore, the test statistic indicates a Wald percentile of 53.10. Thus, the SI model cannot be rejected at the 5% significant level, which implies that the Wald statistic is well included in the non-rejection region of the bootstrap distribution. In addition, many SA estimates are somehow different from the estimates achieved by Bayesian estimation. For instance, the elasticity of inter-temporal substitution σ is seven times higher than the Bayesian estimated value (0.1092). Moreover, for the SA estimated share of updating firms λ 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). Moreover, the share of updating consumers δ is estimated to be two times larger than that obtained though Bayesian approach.

[Insert Table 3]

Estimation-based Indirect Inference Testing Results: IF Data Revision Expectation Model

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

[Insert Table 4]

4 Comparison through Estimation-based Test

TM Distance Comparison Overall, due to the norm of 1.645 as a threshold of judging the succeed of passing, only the models whose absolute values of TM distance are below 1.645 can be qualified as being 'good enough' models. As shown in Table 5, the SI model can pass the Bayesian and SA estimation-based test with a failure in calibration-based test, whereas the FIRE and IF models can pass 1 and 0 test, respectively. Therefore, the SI Model is superior to the other two in terms of overall model fit.

The model assessment is more precise by using the SA estimates from the point of view of the actual data. The AR coefficients in SA estimates are estimated based on the structural errors that 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 those of the initial calibration-based testing, as expected. This improvement can be attributed to the application of SA estimation approach, which explores all the potential parameters over wild space to discover the best fit.

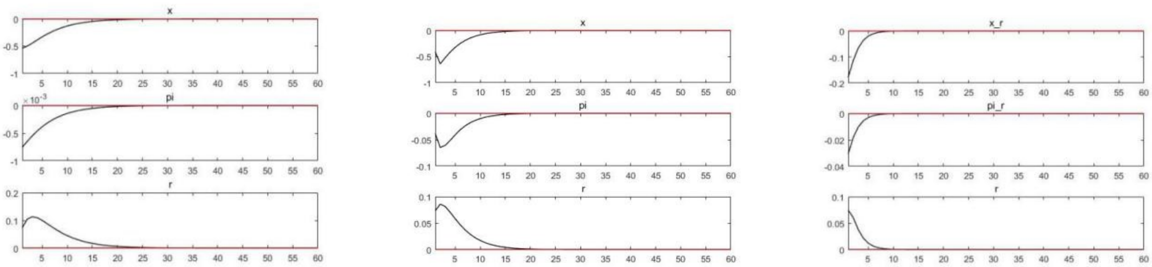
[Insert Table 5]

IRFs of Monetary Policy Shock

Figure 1 shows the estimated impulse response of the three main variables (i.e. output gap, inflation and interest rate) to the monetary policy shock of the three competing models. Generally, under the estimated monetary policy reaction function, the responses of the same variable under different models are quantitatively similar. Specifically, nominal interest rate increases, whereas the output gap and inflation decrease with respect to the three competing models. As shown in Figure 1, throughout the effect of monetary policy shock on inflation and output gap, the hump-shaped response only appears under the SI model. Regarding the period of convergence, the convergences of the three main variables under the FIRE model (the baseline model) and the SI model are around 18 periods; however, under the IF model (i.e. the model with IF data revision), they converge faster. Surprisingly, under the model with IF data revision, the effect 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 1: Estimated Impulse Response Function of One Unit of Positive Policy Shock to the Main Variables

FIRE Model | SI Model | IF Model



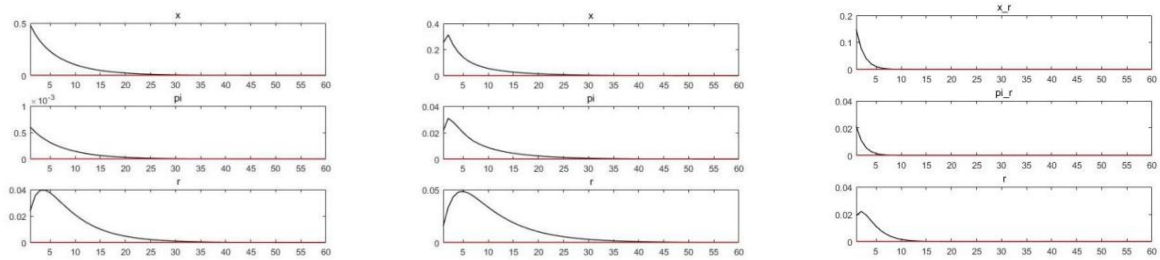
Notes: x indicates output gap, pi indicates inflation and r indicates interest rate.

IRFs of Demand Shock

Figure 2 presents the estimated impulse response functions of the three main variables to demand shock regarding the three competitors. 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 effect on inflation and output gap under SI model, which does not appear under the other two competing models.

Figure 2: Estimated Impulse Response Function of One-Unit Positive Demand Shock to the Main Variables

FIRE Model SI Model IF Model



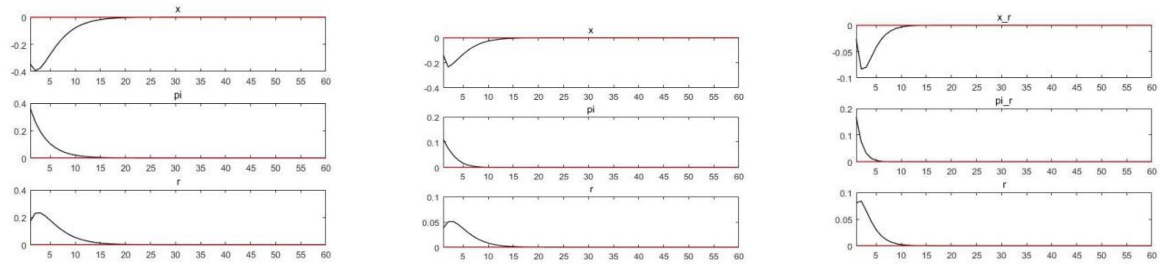
Notes: x indicates output gap, pi indicates inflation and r indicates interest rate.

IRFs of Cost-Push Shock

Figure 3 shows the behavior of the three main variables in response to the positive cost-push shock with respect to the three competitors. Generally, all the three competing models generate similar dynamics quantitatively. Specifically, the inflation and interest rate are affected positively by the positive cost-push shock, which delivers a negative effect on the output gap. In addition, the cost-push shock has the largest effect at the initial point under the FIRE model on three main variables. However, it has a moderate effect at the initial point under the SI model and a minimal effect under the IF model in terms of periods returning to a steady state.

Figure 3: Estimated Impulse Response Function of One-Unit Positive Cost-Push Shock to the Main Variables

FIRE Model SI Model IF Model



Notes: x indicates output gap, pi indicates inflation and r indicates interest rate.

5 Robustness Check

Higher-order Auxiliary Models

We need to check whether the rank amongst the three competing models in terms of higher-order auxiliary models is robust with the optimal set of parameters. We select VAR (1) as the auxiliary model in which the selected descriptors are equivalent to the estimates of its coefficient matrix and data variance incorporated in the indirect inference estimation procedure. As stated previously, two factors, which are the model required to fit and its extent of fit, decide which option we should choose as the auxiliary model from a VAR model with higher order and other types of time series models. When the higher-order auxiliary models, VAR (2) and VAR (3), have been applied, the results show that although none can pass the test, the models' performance can still be compared. As shown in Table 6, the leading position of the SI model in terms of overall dynamic properties over the competitors has not changed when a higher-order VAR (i.e. VAR (2) or VAR (3)) instead of VAR (1) is selected as the auxiliary model.

Overall, the results of TM statistics in Table 3-9 indicate that increasing the order of VAR will make the non-rejection of all the three estimated models even weaker due to the greater burden placed on them. By comparing the results of TM statistics from Tables 6 and 5, we can draw three conclusions. Firstly, 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 this model is preferred in terms of model's overall performance regardless of the auxiliary VAR models' order. Thirdly, the ranking of three competing models is identical to the previous studies regardless of different choices of auxiliary models (i.e. VAR (2) or VAR (3)) through SA estimation amongst three competitors. Thus, VAR (1) can be an accepted auxiliary model to mimic the theoretical models.

[Insert Table 6]

Different Truncation Points j of SI Model

In this section, we need to check the robustness of the different truncation points j in the SI model in an indirect inference approach. We select alternatives $j = 6$ and 8 to be used in the robustness check procedure. As shown in Table 7, we receive the same suggestion as the one provided by Bayesian estimation approach, that is, incorporating more lagged information into the SI model has influence on its model performance after checking the TM distance (normalised t-statistics). Furthermore, the ranking amongst the three competitors is identical as the previous ranking no matter the value of the truncation point j (i.e. $j = 6$ and 8) in the SI model.

[Insert Table 7]

Using Alternative Data Resource: SPF Data of Output Gap and Inflation

The estimation result by using SPF data (survey data) is presented in Table 9. The results obtained through Bayesian estimation approach show that the performance of the IF model is far more superior to its competitors. However, through indirect inference estimation, the full ability of the IF model is far inferior to its competitors. When each model is estimated by using the survey data instead of a real-time data, none of them can pass the test. In addition, determining which one from the FIRE expectation model and the SI expectation model can yield the better replication of the full dynamics of the actual observable has become difficult. However, the SI model performs at least no worse than the baseline when the SPF data are used.

[Insert Table 8]

[Insert Table 9]

6 Conclusion

In this study, we use indirect inference as a testing method (i.e. calibration-based testing method) at the starting stage and take the same approach as an estimation method (i.e. estimation-based testing method) in the second stage. We aim to compare 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 the US economic real-time quarterly data from 1969 to 2015 [we also use other types of sample data (i.e. SPF data) over the same period in the robustness check]. We compare three versions of model and find that none of them can fit the actual data through the initial calibrated-based test.

In this study, indirect inference is applied as the estimation approach to both types of expectation models: with and without inattentiveness. The comparisons of each competing models through Bayesian and SA estimation-based (indirect inference) test are conducted. Four achievements can be reflected through the results of indirect inference estimation. Firstly, regardless of two estimation methods (i.e. Bayesian and indirect inference estimations), by using the real-time data, the model with the SI expectation is the preferred approach amongst the three competitors. Secondly, when we attempt to find a robust superior model in terms of dynamic performance by changing the conditions, such as auxiliary model, truncation point in the SI model and type of data resource, we find that the model with the SI expectation is still the best choice to fit the US economy, Thirdly, the effects of the structural shocks on the US economy are analysed by the estimated IRFs.

Although the Bayesian estimation approach is an effective practical tool in inspecting a model's performance by considering prior information about the macro economy, the prior is restricted while being applied because prior distribution needs to be determined before entering the estimation process. Moreover, the model's performance obtained by Bayesian estimation indicate that their absolute abilities are impossible to evaluate. Thus, the method of indirect inference used in this study is an advanced tool for re-estimating each competing model in an 'unrestricted' manner by exploring all the potential sets of parameters that can be accepted by the models. In addition, the independent VAR is used as an auxiliary model to offer a way to examine each model in an absolute sense. Moreover, the optimal set of parameters can be discovered through SA mechanism for each competing model to mitigate the unfairness in model comparisons.

While we are replacing the real-time data with survey data to apply them in estimation procedure, we find that the performance of the models increase, except for the cases of FIRE model and SI model, through indirect inference. This contradiction indicates that the survey data may contain useful information to improve the IF data revision model's performance.

References

Aruoba, S. B. (2008). Data revisions are not well behaved. *Journal of money, credit and banking 40*(2-3), 319–340.

Ball, L., N. G. Mankiw, and R. Reis (2005). Monetary policy for inattentive economies. *Journal of monetary economics 52*(4), 703–725.

Caballero, R. J. (2010). Macroeconomics after the crisis: Time to deal with the pretense-of-knowledge syndrome. *Journal of Economic Perspectives 24*(4), 85–102.

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

Casares, M. and J. Vázquez (2016). Data revisions in the estimation of dsge models. *Macroeconomic Dynamics 20*(7), 1683–1716.

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

Coibion, O. et al. (2006). Inflation inertia in sticky information models. *Contributions to Macroeconomics 6*(1), 1–29.

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. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

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

Collard, F., H. Dellas, and F. Smets (2009). Imperfect information and the business cycle. *Journal of Monetary Economics 56*, S38–S56.

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

Dhyne, E., L. J. Alvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lunnemann, F. Rumler, and J. Vilmunen (2006). Price changes in the euro area and the united states: Some facts from individual consumer price data. *Journal of Economic Perspectives 20*(2), 171–192.

Easaw, J. and R. Golinelli (2010). Households forming inflation expectations: Active and passive absorption rates. *The BE Journal of Macroeconomics 10*(1).

Galı, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics 44*(2), 195–222.

Gallant, A. R. and G. Tauchen (1996). Which moments to match? *Econometric theory 12*(4), 657–681.

Gregory, A. W. and G. W. Smith (1991). Calibration as testing: inference in simulated macroeconomic models. *Journal of Business & Economic Statistics 9*(3), 297–303.

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

KNOTEK II, E. S. (2010). A tale of two rigidities: Sticky prices in a sticky-information environment. *Journal of Money, Credit and Banking 42*(8), 1543–1564.

Le, V. P. M., D. Meenagh, P. Minford, and M. Wickens (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.

Liu, C. and P. Minford (2014). Comparing behavioural and rational expectations for the us post-war economy. *Economic Modelling 43*, 407–415.

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

Milani, F. and A. Rajbhandari (2012). Expectation formation and monetary dsge models: Beyond the rational expectations paradigm. *Advances in Econometrics 28*, 253.

Minford, P., Z. Ou, and M. Wickens (2015). Revisiting the great moderation: policy or luck? *Open Economies Review 26*(2), 197–223.

Minford, P., K. Theodoridis, and D. Meenagh (2009). Testing a model of the uk by the method of indirect inference. *Open Economies Review 20*(2), 265–291.

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, R. (2009). Optimal monetary policy rules in an estimated sticky-information model. *American Economic Journal: Macroeconomics 1*(2), 1–28.

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

Simon, H. A. (1989). Models of thought. volume ii (p. 291).

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

Smets, F. and R. Wouters (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *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, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, Volume 39, pp. 195–214. Elsevier.

Trabandt, M. (2007). Sticky information vs. sticky prices: A horse race in a dsge framework. *Riksbank Research Paper Series* (209).

Vázquez, J., R. María-Dolores, and J. M. Londoño (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.*

Vázquez, J., R. María-Dolores, and J. M. Londoño (2012). The effect of data revisions on the basic new keynesian model. *International Review of Economics & Finance 24*, 235–249.

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

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. Technical report, National Bureau of Economic Research.

Appendix A

Table A1: Common Parameters

Common Structural parameter		
σ	Elasticity of intertemporal substitution	1
α	Sticky price degree	0.6
γ	Strategic complementary	0.2
Common Taylor Rule in three models		
ρ	Degree of partially adjustment in Taylor rule	0.8
χ_π	Coefficient of inflation on Taylor rule	1.5
χ_x	Coefficient of output gap in Taylor rule	0.1
Common Forcing Variables in three models		
ρ_g	AR coefficient of demand shock	0.5
ρ_u	AR coefficient of cost-push shock	0.5
ρ_r	AR coefficient of policy shock	0.5
ρ_g	Standard deviation of demand shock	0.3
ρ_u	Standard deviation of cost-push shock	0.3
ρ_r	Standard deviation of policy shock	0.3

Note: The priors of parameter are mostly chosen from previous literature (Milani and Rajbhandari, 2012; Smets and Wouters, 2003, 2007)

Table A2: Inattentive Parameters

Imperfect Information model		
b_x	Output coefficient in output revision process	0
b_π	Inflation coefficient in inflation revision process	0
ρ_x	AR term of shock in final revision process of x	0.5
ρ_π	AR term of shock in final revision process of π	0.5
e_x	SD of measurement error of x	0.25
e_π	SD of measurement error of π	0.25
Sticky Information model		
λ	Share of updating firms	0.5
δ	Share of updating consumer	0.5

Note: The priors of parameter for SI model are chosen from Mankiw and Reis (2007) , and those for IF model are borrowed from Casares and Vázquez (2016) .

Appendix B: Data Description

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

United States Data Source:

Effective Federal Funds Rate is indicated by FEDFUNDS, the federal funds rate is divided by four to express it in quarterly rates. The observable data are matched to the variable r_t, where $r_t = \frac{FEDFUNDS_t}{4}$.

The real-time data from the real-time data set for macroeconomists hosted by the Federal Reserve Bank of Philadelphia [10]. The real-time Real GDP is indicated by ROUTPUT which is initially released in 2016Q1 (i.e., which only contains real-time Real GDP up to time 2015Q4); the quarterly real-time GDP is the deviation of the natural logarithm of total real-time GDP. For the IF model to construct the revised observables corresponding to the output gap up to 2015Q4, the real-time data released after one period (2016Q1) and the real-time data of GDP released after three periods are also applied (2016Q3).

Real-time Implicit Price Deflator is indicated by P. The series is demeaned for the index level which is initially released in 2016Q1 (i.e., which only contains real-time Implicit Price Deflator up to 2015Q4), which is seasonally adjusted and is also from the real-time data set from Federal Reserve Bank of Philadelphia. The real-time inflation $\pi_t^r = (lnP_t - lnP_{t-1}) * 100$. Similarly, to construct the revised observables corresponding to inflation up to 2015Q4, the real-time data of the Implicit Price Deflator released after one period and the data released after three periods are also used.

The survey data used in the robust check section is the median of the Survey of Professional Forecaster one-quarter ahead forecasts of the GDP deflator and real GDP. In the IF data revision model, both one-quarter ahead and four-quarter ahead forecasts are used to construct the final revised observables.

[10]https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files

Appendix C: Model Derivation

IS Curve in the Sticky Information Model

Now, we assume that economic agents and households under the sticky information economy use the outdated information from all past periods up to t to form their forecast. In the aggregate level, not all of them use the updated information to form their forecasts, $E_t^{SI} = \delta \sum_{j=0}^{\infty}(1-\delta)^j E_{t-j}$. Thus, we have the following IS equation[11] :

$$x_t = \delta \sum_{j=0}^{\infty}(1-\delta)^j E_{t-j} x_{t+1} - \sigma(\tilde{r}_t - \pi_{t+1}) + g_t \tag{D1}$$

where δ denots the share of updating households.

Phillips Curve in the Sticky Information Model

Similarly, for firms that are also subject to sticky information, and because they do not all use the updated information to form their forecast at the aggregate level, firms must use the outdated information up to time t to form their forecast $E_t^{SI} = \lambda \sum_{j=0}^{\infty}(1-\lambda)^j E_{t_j}$. Then, we have the following PC equation[12]:

$$\pi_t = \beta\lambda \sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\pi_{t+1} + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t \tag{D2}$$

where λ denotes the share of the updating firms.

From above, we can see that the current inflation depends on the current output gap and on current and past expectations of the future inflation rate.

[11]Initially, this is $x_t = E_t^{SI} x_{t+1} - \sigma(\tilde{r}_t - E_t^{SI}\pi_{t+1}) + g_t$.
[12]Initially, this is $\pi_t = \beta E_t^{SI}\pi_{t+1} + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t$.

Imperfect Information Data Revision

The derivation of the imperfect information data revision model follows the deriving procedure and assumption explanation provided by Aruoba (2008), Vázquez et al. (2010), Vázquez et al. (2012) and Casares and Vázquez (2016). First, we consider the following identities regarding revised data related to the cyclical of output gap and inflation, which can also refer to the combination of the initial announcement and the final revisions. This can be interpreted in the sense of noise: x_t^r and π_t^r are taken as the observed variables realized at time t (they are the real-time data). In addition, x_t and π_t are the final revised variables, which are defined respectively as follows:

$$x_t \equiv x_t^r + v_t^x \tag{D3}$$

$$\pi_t \equiv \pi_t^r + v_t^\pi \tag{D4}$$

We also follow the argument of Aruoba (2008) that, for many US aggregate time-series (e.g., inflation and output), their revisions are not rational forecast errors and are supposed to be connected to their initial realized variables, x_t^r and π_t^r. Thus, following his argument, we presume that the final revision process of the US output gap and inflation are defined as follows:

$$v_t^x = b_x x_t^r + e_t^x \tag{D5}$$

$$v_t^\pi = b_\pi \pi_t^r + e_t^\pi \tag{D6}$$

These revision processes allow for the existence of non-zero correlation between final true variables (i.e., output gap and inflation) and their initial realized variables along with the existence of persistence revision processes. In particular, the shocks of the revision processes, e_t^x and e_t^π, are both AR (1) processes. The two data revision processes aim to offer a simple framework to approximate the "true" revision processes and examine whether the deviation in the way we use the assumption of well-behaved revision processes (i.e., white noise) influences the estimation of policy and behavioral parameters. Therefore, from the defined equation above, we can obtain the following:

$$x_t \equiv x_t^r + v_t^x = (1+b_x)x_t^r + e_t^x \tag{D7}$$

$$\pi_t \equiv \pi_t^r + v_t^\pi = (1+b_\pi)\pi_t^r + e_t^\pi \tag{D8}$$

Furthermore, notice that the final revision process of output gap and inflation also implies the identities' respective equations as follows:

$$v_t^x = E_{t+1}v_t^x + e_t^x = b_x x_t^r + e_t^x \tag{D9}$$

$$v_t^\pi = E_{t+1}v_t^\pi + e_t^\pi = b_\pi \pi_t^r + e_t^\pi \tag{D10}$$

$$E_{t+1}v_t^x = b_x e_t^r \tag{D11}$$

$$E_{t+1}v_t^\pi = b_\pi \pi_t^r \tag{D12}$$

IS Curve in the Imperfect Information Model

We use the imperfect information data revision assumption to distinguish the baseline FIRE model. We can obtain the IS equation below[13]:

$$x_t = E_t(x_{t+1}^r + E_{t+2}v_{t+1}^x) - \sigma[\tilde{r}_t - E_t(\pi_{t+1}^r + E_{t+2}v_{t+1}^\pi)] + g_t \tag{D13}$$

where households involve data revision issues, because these imperfect-information-type of people react to the expected revised values of inflation and output gap.

We also use the identity equations $E_{t+2}v_{t+1}^x = b_x x_{t+1}^r$ and $E_{t+2}v_{t+1}^\pi = b_\pi \pi_{t+1}^r$ to substitute out $E_{t+2}v_{t+1}^x$ and $E_{t+2}v_{t+1}^\pi$ respectively, to obtain the imperfect information IS

[13] Initially, this is $x_t = E_t^{IF} x_{t+1} - \sigma(\tilde{r}_t - E_t^{IF}\pi_{t+1}) + g_t$.

equation below[14]:

$$x_t = (1 + b_x)E_t(x^r_{t+1}) - \sigma[\tilde{r}_t - (1 + b_\pi)E_t(\pi_{t+1^r})] + g_t \tag{D14}$$

Phillips Curve in the Imperfect Information Model

For firms with data revision issues (noise disturbance) we can obtain the imperfect information PC using the following equation:

$$\pi_t = \beta E_t(\pi^r_{t+1} + E_{t+2}v^\pi_{t+1}) + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t \tag{D15}$$

Similarly, we use the identity equation to substitute out $E_t v^\pi_{t+1}$ from the above equation to obtain [15]

$$\pi_t = \beta E^{IF}_t \pi_{t+1} + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t \tag{D16}$$

Meanwhile, the monetary policy assumed to be perfect is observed to have no data revision issue

$$\tilde{r}_t = \rho_r \tilde{r}_{t-1} + (1-\rho)[\chi_\pi x_t + \chi_x \pi_t] + v_t \tag{D17}$$

where the final revisions v^x_t and v^π_t their data can be constructed as demeaned observables between the first released $x^r_{t,t+1}$ and the latest released $x^r_{t,t+s}$ as follows:

$$v^x_t = (x^r_{t,t+1} - x^r_{t,t+s}) - M^{vx} \tag{D18}$$

$$v^\pi_t = (\pi^r_{t,t+1} - \pi^r_{t,t+s}) - M^{\pi x} \tag{D19}$$

Thus, for the analysis, we choose $s = 3$ to construct the observations of the final

[14]Initially, this is $x_t = E_t(x^r_{t+1} + b_x x^r_{t+1}) - \sigma[\tilde{r}_t - E_t(\pi^r_{t+1} + b_\pi \pi^r_{t+1})] + g_t$

[15]Initially, this is $\pi_t = (1 + b_\pi)\beta E_t(\pi^r_{t+1}) + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t$

revisions v_t^x and v_t^π :

$$v_t^x = (x_{t,t+1}^r - x_{t,t+3}^r) - M^{vx_3} \tag{D20}$$

$$v_t^\pi = (\pi_{t,t+1}^r - \pi_{t,t+3}^r) - M^{\pi x_3} \tag{D21}$$

Therefore, we can also construct the observations of the revised data x_t and π_t.

Note that, as argued by Croushore (2011), if we look at the US data, we can see 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 the performance of the imperfect information data revision. Here we choose $s = 3$, $x_{t,t+1}^r$ as the data released in 2016Q1, and $x_{t,t+3}^r$ as the data released in 2016Q3 to construct the revision process corresponding to the sample period from 1969Q1 up 2015Q4. For the simplicity of the analysis procedure, we consider the number of periods after which no more revisions can be done (except benchmark revisions, which is represented by s) and whether it is constant.

Table 1: Model Setting

Assumption	Model Summarised
Model 1 (FIRE):	IS: $x_t = E_t x_{t+1} - \sigma(\tilde{r}_t - E_t \pi_{t+1}) + g_t$
	PC: $\pi_t = \beta E_t \pi_{t+1} + \gamma((1-\alpha)(1-\alpha\beta)/\alpha)x_t + u_t$
	TR: $\tilde{r}_t = \rho_r \tilde{r}_{t-1} + (1-\rho)[\chi_\pi x_t + \chi_x \pi_t] + v_t$
Model 2 (SI):	IS: $x_t = \delta \sum_{j=0}^{\infty}(1-\delta)^j E_{t-j} x_{t+1} - \sigma(\tilde{r}_t - \pi_{t+1}) + g_t$
	PC: $\pi_t = \beta\lambda \sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\pi_{t+1} + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t$
	TR: $\tilde{r}_t = \rho_r \tilde{r}_{t-1} + (1-\rho)[\chi_\pi x_t + \chi_x \pi_t] + v_t$
Model 3 (IF):	IS: $x_t = E_t(x^r_{t+1} + E_{t+2} v^x_{t+1}) - \sigma[\tilde{r}_t - E_t(\pi^r_{t+1} + E_{t+2} v^\pi_{t+1})] + g_t$
	PC: $\pi_t = \beta E^{IF}_t \pi_{t+1} + \gamma(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha})x_t + u_t$
	TR: $\tilde{r}_t = \rho_r \tilde{r}_{t-1} + (1-\rho)[\chi_\pi x_t + \chi_x \pi_t] + v_t$

Table 2: Estimates of FIRE Model

Parameters	Starting Calibration (Initial Values)	Bayesian Estimates	SA Estimates
σ	1	0.0225	0.518
α	0.6	0.7257	0.9677
ρ	0.75	0.8834	0.658
χ_π	1.5	1.3891	1.5079
χ_x	0.12	0.1974	0.1439
ρ_g	0.86	0.7995	0.8587
ρ_u	0.73	0.6948	0.7318
ρ_r	0.82	0.3094	0.8155
Full Wald %	100	100	64.8
TM (Normalised t-Statistic)	4.1538	26.0498	0.6587

Table 3: Estimates of SI Model

Parameters	Starting Calibration (Initial Values)	Bayesian Estimates	SA Estimates
σ	1	0.1092	0.905
α	0.6	0.634	0.5542
ρ	0.75	0.9002	0.7672
χ_π	1.5	1.3735	1.6266
χ_x	0.12	0.1848	0.1299
ρ_g	0.89	0.8139	0.8842
ρ_u	0.79	0.649	0.6421
ρ_r	0.64	0.2986	0.7351
λ	0.5	0.3084	0.4504
δ	0.5	0.2362	0.5138
Full Wald %	99.4	54	53.1
TM (Normalised t-Statistic)	2.7338	-0.2072	0.1092

Table 4: Estimates of IF Data Revision Model

Parameters	Starting Calibration (Initial Values)	Bayesian Estimates	SA Estimates
σ	1	0.0899	0.8639
α	0.6	0.7389	0.5623
ρ	0.75	0.8801	0.6495
χ_π	1.5	1.0884	1.3342
χ_x	0.12	0.1962	0.1131
b_x	0.5	1.85	0.4404
b_π	0.5	1.1198	0.4683
ρ_g	0.67	0.6186	0.6292
ρ_u	0.56	0.3657	0.5083
ρ_r	0.3	0.2235	0.2718
ρ_x	0.42	0.7252	0.3443
ρ_π	0.61	0.8535	0.5099
Full Wald %	100	100	100
TM (Normalised t-Statistic)	28.5625	94.6459	20.3812

Table 5: Comparison TM Distance (Normalised 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

Table 6: 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)

Table 7: Sensitivity Check by using Minimising Coefficient Values for SI Model

Model	TM Using SA Estimation 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

Table 8: Starting Calibration Parameter Value of AR Coefficients

Parameters	Interpretation	FIRE	SI	IF
ρ_g	AR coefficient of demand shock	0.94	0.93	0.7
ρ_u	AR coefficient of cost-push shock	0.75	0.74	0.54
ρ_r	AR coefficient of policy shock	0.56	0.56	0.29
ρ_x	AR term of shock in final revision process of x	-	-	0.39
ρ_π	AR term of shock in final revision process of π	-	-	0.59

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

Table 9: Comparison of TM using Minimising Coefficient Values (with SPF Data)

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