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

Jenyu Chou, Yifei Cao and Patrick Minford

November 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

This working paper is produced for discussion purpose only. These working papers are expected to be published in
due course, in revised form, and should not be quoted or cited without the author’s written permission.
Cardiff Economics Working Papers are available online from:
http://econpapers.repec.org/paper/cdfwpaper/ and
business.cardiff.ac.uk/research/academic-sections/economics/working-papers
Enquiries: EconWP@cardiff.ac.uk
Evaluation and Indirect Inference

Estimation of Inattentive Feature in a New Keynesian Framework*

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

†School of Economics, University of Nottingham Ningbo China
‡Cardiff University Business School, Cardiff, UK

Abstract

We test the standard New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model with and without inattentive features, where inattentiveness is modeled in the form of sticky information and imperfect information data revision. All models are tested by Indirect Inference, 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 is used for model evaluation.

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

JEL Classification: 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. 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 expectations in determining aggregate outcomes of the macroeconomy, such as inflation dynamics and the business cycle, has been widely discussed. However, studies of how people form their expectations are relatively rare. Yet this topic is important as the most fundamental macroeconomic decisions are driven by people’s expectations of the future. In this paper, we test DSGE models with different degrees of rationality in expectations against the macro data for the postwar US. We begin by surveying the literature on fully attentive expectations or full-information rational expectation (FIRE) and then explore the weakness of this early expectation assumption. To address its weakness, another assumption deviating from full-information rationality has been proposed, namely inattentive expectations. We focus in this paper 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 assumption is imperfect information (IF) data revision (Aruoba, 2008; Vázquez et al., 2010, 2012; Casares and Vázquez, 2016). Both inattentiveness assumptions will be specified carefully in what follows.

Simon (1989) criticises the ‘unrealistic’ view of the FIRE hypothesis. He argues that economic agents, knowing all their problems, choices and possible results, can certainly choose the best solution from all alternatives through some reasonable calculation. However, 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 constraints stemming from the superior authority of government in terms of legislation or personal constraints originating from limited time and energy). Thus, economic agents have to seek coordination to achieve efficiency and profits. In other words, economic agents cannot simply reach the optimal solution but only reach the satisfying or ‘good enough’ solution. As a result, 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 they can reach their maximum profit position. However, due to people’s physical and intellectual capacity limitations, added to the uncertainties originating from the external environment, people understand and solve complex problems only in a restricted manner.

2) Under the assumption of FIRE, economic agents are able to collect all available information to make economic decisions. But this 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 amount and quality of information obtained by economic entities is limited, which implies that economic agents cannot reach their optimum. Similar points are made by Caballero (2010), Stiglitz (2011) and Coibion and Gorodnichenko (2012, 2015).

2 Three Competitors

To address these issues, inattentiveness in expectations has been proposed. Of different such approaches, the two most prominent are SI (Mankiw and Reis, 2007) and IF data revision (Casares and Vázquez, 2016). These assumptions will be explored in the current study. We will do so using a small-scale closed-economy dynamic stochastic general equilibrium (DSGE) model as our benchmark model.

Although full-information rationality has weaknesses, as just argued, the assumption of rationality need not be jettisoned nor do other types of irrational behavior need to be
introduced to help the model fit the data (Collard et al., 2009; Coibion and Gorodnichenko, 2012). Thus, in this study, we focus only on the information assumptions.

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 macroeconomic 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.1

The second objective is to discover 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 the competing models solely according to their ability to match the behaviour of the data. Although the Bayesian approach provides a simple way to compare the relative performance of different models, it does so not merely on data fit but also on the basis of priors which may well not be widely accepted; hence its rankings will not be treated as objective by those not sharing these priors. The indirect inference estimation (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.

To preview in outline what we do, we propose three competing models. First, a New Keynesian DSGE model for a small-scale closed economy under the FIRE assumption. The economy consists of three types of agents, namely, households, firms, and monetary

---

1Standard 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.
authorities. The baseline FIRE model, which has been largely applied in previous studies (Milani and Rajbhandari, 2012), is the standard Calvo model without any inattentive features. Next, we replace FIRE in the model by SI, and third by IF data revision, as in Casares and Vázquez (2016). Unlike some earlier work, we use a small-scale DSGE model, the basic three-equation model of the IS curve, Phillips curve and Taylor rule, instead of a medium-sized one. Adding additional features might be a useful step (Smets and Wouters, 2003, 2007). However, the greater complexity could lead to difficulty in assessing the differences between the information features of the models.

Table 1: Model Setting

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Model Summarised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 (FIRE):</td>
<td>IS: ( x_t = E_t x_{t+1} - \sigma(\tilde{r}<em>t - E_t \pi</em>{t+1}) + g_t )</td>
</tr>
<tr>
<td></td>
<td>PC: ( \pi_t = \beta E_t \pi_{t+1} + \gamma((1-\alpha)(1-\alpha\beta)/\alpha)x_t + u_t )</td>
</tr>
<tr>
<td></td>
<td>TR: ( \tilde{r}<em>t = \rho \tilde{r}</em>{t-1} + (1-\rho)[\chi x_t + \chi_x \pi_t] + v_t )</td>
</tr>
<tr>
<td>Model 2 (SI):</td>
<td>IS: ( x_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j E_{t-j} x_{t+1} - \sigma(\tilde{r}<em>t - \pi</em>{t+1}) + g_t )</td>
</tr>
<tr>
<td></td>
<td>PC: ( \pi_t = \beta \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j} \pi_{t+1} + \gamma((1-\alpha)(1-\alpha\beta)/\alpha)x_t + u_t )</td>
</tr>
<tr>
<td></td>
<td>TR: ( \tilde{r}<em>t = \rho \tilde{r}</em>{t-1} + (1-\rho)[\chi x_t + \chi_x \pi_t] + v_t )</td>
</tr>
<tr>
<td>Model 3 (IF):</td>
<td>IS: ( x_t = E_t (x_{t+1} + E_{t+2} \pi_{t+1}^2) - \sigma[\tilde{r}<em>t - E_t (\pi</em>{t+1} + E_{t+2} \pi_{t+1}^2)] + g_t )</td>
</tr>
<tr>
<td></td>
<td>PC: ( \pi_t = \beta E_t \pi_{t+1} + \gamma((1-\alpha)(1-\alpha\beta)/\alpha)x_t + u_t )</td>
</tr>
<tr>
<td></td>
<td>TR: ( \tilde{r}<em>t = \rho \tilde{r}</em>{t-1} + (1-\rho)[\chi x_t + \chi_x \pi_t] + v_t )</td>
</tr>
</tbody>
</table>

Note: FIRE: Full-Information Rational Expectation Model; SI: Sticky Information Expectation Model; IF: Imperfect Information Data Revision Model.

Table 1 sets out the three models in the summary outline. The presented three competitors above indicate that the aggregate economy in the 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 \( \sigma \) represents the elasticity of the intertemporal substitution. The new Keynesian PC derived under the FIRE is equivalent to the current inflation \( \pi_t \) 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 $\beta$ represents the time discount factor, and $\gamma$ is the combined parameter.\(^2\) The interest rate equation follows the simple ‘interest rate smoothed’ Taylor rule (Taylor, 1993). The interest rate $\bar{r}_t$ is driven by the current inflation $\pi_t$ and current output gap $x_t$.

Thus, on the basis of the model with SI, the two parameters $\delta$ and $\lambda$ are the shares of updating households and of updating firms in any given period, respectively (for example, if no SI of firms exists, then $\lambda = 1$) (Mankiw and Reis, 2002, 2007; Reis, 2006a,b, 2009). Reis (2006a,b) provides deeper micro-foundations for model features SI. Under an SI environment, the inclusion of inattentiveness leads to deviation from full-information rationality. The economic agents in these circumstances use outdated information to form their expectations. Therefore, the Phillips Curve depends not only on current expectations but also past expectations about the future, caused by information spreading slowly through the entire population of the economy (Mankiw and Reis, 2002).\(^3\) The models are estimated 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 FIRE model, the model with SI is more challenging to solve. Given that SI involves infinitely lagged expectations, we consider how we can approximate the model with SI in the DSGE equilibrium framework. Firstly, from the angle of the SI model setting, the proportion of lagged expectations diminishes geometrically. In other words, the effect on economic agents’ expectations 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

\(^2\)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 one and in line with the suggestion from the previous literature (Woodford, 2001; Ball et al., 2005). Woodford (2011) surveys and discusses the existing literature at length and concludes that firms’ pricing decisions 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.

\(^3\)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 assumptions.
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 information up to four periods); longer periods, such as \( j = 6 \) and 8, are considered in the robustness check.\(^4\)

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’ decisions 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 engaged 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, in general, data revision has been thought to either not exist or is small, but in reality, data revision may significantly influence empirical results, particularly in some variables defined conceptually. For instance, when economic agents make decisions for the output (or output gap), they take this variable without any doubt. In an actual stat, such a variable as the output gap often fluctuates over time. Thus, data revision is considered in the IF model to see how it affects the NK macroeconomic model and the empirical results.

Moreover, as specified in the Appendix, we follow the suggestion by Casares and Vázquez (2016) for the data revision. In addition to the above discussion, 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

\(^4\)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.
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 features highlighted in this study must be taken seriously; such inattentive assumption needs to be reasonable. The nominal interest rates made through professional monetary authorities are fully observable without noise disturbance, and noises influence the output gap and inflation. 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. Under the assumption of IF, firms cannot fully observe their 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. Price (inflation) and consumption (output) can only be observed with random noises. The above three-equation models are the real-time data, where \( x^*_t \) and \( \pi^*_t \) are taken as the observed variables realised at time \( t \). \( x_t \) and \( \pi_t \) 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 detrended data are applied. The three models have different information friction constraints, thereby having other IS and PC influencing 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 economic dynamics.
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 for filtering models divides the tested models into qualified and unqualified groups. 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) use 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 standard 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 regarding our research topic are available using 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 indicates that the ignorance of the data revision process may not result in a severe drawback in analysing a monetary policy based on an NK framework. Our assumption substitutes the subjects which involve IF data revision issues with households and firms instead of the monetary authority. Moreover, the agents 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 macro dimen-
sions after estimating it through indirect inference.\(^5\) However, we are more interested in a full structural model than a single-equation model.

This paper evaluates each model, focusing on its overall dynamic properties in connecting with the actual data by adopting indirect inference as the new evaluation method. While applying indirect inference to evaluate an existing structural model, two factors are inevitable in 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 an indirect inference test based on comparing the observed actual data with the data simulated from the theoretical model with the assistance of an auxiliary model. In this study, VAR, a stochastic process model used to capture the linear inter-dependencies amongst multiple time series, is selected as the auxiliary model for two reasons. The structural model can always be represented as a restricted VARMA (i.e. Vector Auto-regression Moving-Average), 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 covariance 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, which is derived from a function of these parameters, can be regarded as two criteria of the testing model to measure the distance to 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. The theoretical model cannot significantly reproduce the actual data if the consequence is rejected. Conversely, the accepted result implies that the data generated from the theoretical model do not significantly differ from the actual observed data.

\(^5\)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.
3.1 Wald Test Statistics

The Wald testing process can generally be summarised into three general steps as follows. Firstly, the model’s observed actual data and calibrated/estimated parameters derive the structural errors. The errors can be constructed under two different circumstances. The structural errors can be backed directly by the structural equations and the actual data when the structural model possesses no expectation terms. Under the situation that the 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. The fitted values are computed from a VAR (1), 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\(^6\) 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 to direct the dynamics and variances of the data.

\(^6\)We can obtain estimate distribution by estimating the auxiliary VAR model on each pseudo sample. The dynamic properties are captured by VAR estimates, whereas the variance of the main variables can capture the volatility properties. For the individual estimates, the confidence interval (95\%) is calculated directly from their bootstrapped distribution.
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 are computed from the VAR(1) coefficients and the three variances of the three main variables. Therefore, the Wald test statistic determines 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 expressed 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-covariance matrix of the distribution \(G_T(\alpha_T) - G_S(\bar{\alpha}_S(\theta))\). \(\alpha_T\) and \(\alpha_S(\theta)\) are the actual and simulated data sets, respectively. \(\theta\) 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 proceeds by considering the percentile of the Wald distribution - 100 minuses this percentile is the p-value of the model. 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 the bootstrapping method.

### 3.2 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 can rank these models by their p-value or equivalently the TM statistic derived from it. The TM provides a way to examine how poorly 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 and the less probable the model is. The TM distance is defined as follows.

\[
TM \text{ distance(normalised } t - \text{ statistic}) = \frac{(\sqrt{2W_{S_a}} - \sqrt{2p})}{(\sqrt{2W_{S_{95\%}} - \sqrt{2p}})} \times 1.645 \tag{2}
\]

This function of TM distance is based on Wilson and Hilferty (1931) method of transforming Chi-square distribution into a standard normal distribution calculated. Herein, the TM distance is the transformation of the Wald test statistics. \(W_{S_a}\) is the Mahalanobis distance (value of Wald statistics) using the actual data, \(W_{S_{95\%}}\) 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 the crystal (in the 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 on a global scale instead of being trapped in the 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.

The VAR(1) is used as descriptors of the coefficient matrix and the variance of the data, and continuously as the auxiliary model to provide a reference substance for the estimated models to those of the calibrated models for implementing the estimation-based indirect inference test. The SA method, which adjusts the initial presumptive values (calibrated values), is helpful for the models to pass the test. The SA mechanism will explore 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, implying that we have found the ‘best’ estimates of the structural parameters.

### 3.3 Estimation-based Indirect Inference Test: FIRE Model

The SA estimation-based and Bayesian estimation-based tests concerning the three competing models for the US economy are presented in Tables 2 and 4. The numbers in the column regarding the indirect inference estimation are obtained through the SA estimation method. The scope of the value of parameters during SA exploration is limited within ±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 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 falls within the non-rejection scope, then another set that can narrow the gap between the theoretical model and the reality should be explored, leading to better testing results.

The ‘best’ collection of parameters for the structural model are those to the maximum degree to shorten the distance between the theoretical model and 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. Our estimation-based test will use the best collection of parameters that can shorten the distance between the theory and reality. Using these optimal sets of parameters to compare models can reduce the unfairness in model comparisons. The minimised value of the Mahalanobis distance is captured for each competitor over the US sample periods through an SA algorithm.

After examining the estimates of the main behavioural parameters of the FIRE model, we explore the parameters of the monetary policy function, which are based on the standard interest rate smoothed Taylor rule (1993). 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 from those obtained by Bayesian analysis. However, some distinguished cases exist. Notably, the estimated value of the elasticity of inter-temporal substitution is 0.5180, which is relatively 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 analysis. Regarding the estimated coefficients of monetary policy, except for $\chi_x$, which is increased by <8%, the other two (i.e. $\rho$ and $\chi_e$) increase around 35% compared with their estimated values achieved from Bayesian estimation. Within the system, all three stationary shocks are quite highly persistent,
and two of them, except 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 $\chi_\pi$ is 1.5079, slightly higher than that obtained by Bayesian analysis. 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 $\chi_x$ 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 $\rho$ on the lagged interest rate is estimated to be 0.6580 and lower than that obtained through Bayesian estimation. However, it is not far 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 analysis for a distance (e.g. the elasticity of inter-temporal substitution $\sigma$ is increased around 97% higher than the Bayesian estimated value, which is 0.0225. The SA estimated value of price stickiness is around 25% higher than the counterpart of the Bayesian approach). Table 2 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 a significant reduction compared with the one obtained using Bayesian estimates. The full Wald statistics imply that the FIRE model with SA estimates falls within the non-rejection area; thus, the model cannot be rejected at a 95% chance. Furthermore, the model with Bayesian estimates performs worse than the model with the initial presumptive parameters (calibration parameters).
Table 2: Estimates of FIRE Model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Starting Calibration (Initial Values)</th>
<th>Bayesian Estimates</th>
<th>SA Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>1</td>
<td>0.0225</td>
<td>0.518</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.6</td>
<td>0.7257</td>
<td>0.9677</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.75</td>
<td>0.8834</td>
<td>0.658</td>
</tr>
<tr>
<td>$\chi_\pi$</td>
<td>1.5</td>
<td>1.3891</td>
<td>1.5079</td>
</tr>
<tr>
<td>$\chi_x$</td>
<td>0.12</td>
<td>0.1974</td>
<td>0.1439</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>0.86</td>
<td>0.7995</td>
<td>0.8587</td>
</tr>
<tr>
<td>$\rho_u$</td>
<td>0.73</td>
<td>0.6948</td>
<td>0.7318</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>0.82</td>
<td>0.3094</td>
<td>0.8155</td>
</tr>
<tr>
<td>Full Wald %</td>
<td>100</td>
<td>100</td>
<td>64.8</td>
</tr>
<tr>
<td>TM (Normalised t-Statistic)</td>
<td>4.1538</td>
<td>26.0498</td>
<td>0.6587</td>
</tr>
</tbody>
</table>

Note: The sample period of the US quarterly data is from 1969 to 2015.

3.4 Estimation-based Indirect Inference Test: 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 analysis, except for the estimate of interest rate smoothed parameter $\rho$ (0.7672), which is slightly lower than that obtained through Bayesian estimation. The reaction parameter of the output gap in monetary policy $\chi_x$ 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 $\rho_r$, 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 $\sigma$ is seven times higher than the Bayesian estimated value (0.1092). Moreover, for the SA estimated share of updating firms
$\lambda$ 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 $\delta$ is estimated to be two times larger than that obtained through the Bayesian approach.

**Table 3: Estimates of SI Model**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Starting Calibration (Initial Values)</th>
<th>Bayesian Estimates</th>
<th>SA Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>1</td>
<td>0.1092</td>
<td>0.905</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.6</td>
<td>0.634</td>
<td>0.5542</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.75</td>
<td>0.9002</td>
<td>0.7672</td>
</tr>
<tr>
<td>$\chi_\pi$</td>
<td>1.5</td>
<td>1.3735</td>
<td>1.6266</td>
</tr>
<tr>
<td>$\chi_x$</td>
<td>0.12</td>
<td>0.1848</td>
<td>0.1299</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>0.89</td>
<td>0.8139</td>
<td>0.8842</td>
</tr>
<tr>
<td>$\rho_u$</td>
<td>0.79</td>
<td>0.649</td>
<td>0.6421</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>0.64</td>
<td>0.2986</td>
<td>0.7351</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.5</td>
<td>0.3084</td>
<td>0.4504</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.5</td>
<td>0.2362</td>
<td>0.5138</td>
</tr>
<tr>
<td>Full Wald %</td>
<td>99.4</td>
<td>54</td>
<td>53.1</td>
</tr>
<tr>
<td>TM (Normalised t-Statistic)</td>
<td>2.7338</td>
<td>-0.2072</td>
<td>0.1092</td>
</tr>
</tbody>
</table>

**Note:** The sample period of the US quarterly data is from 1969 to 2015.

### 3.5 Estimation-based Indirect Inference Test: IF Data Revision Expectation Model

Generally, although none of the three cases concerning the 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) in Table 4. The most significant difference between the SA and Bayesian estimation-based model tests is that the estimated value of coefficient $\sigma$ 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: Estimates of IF Data Revision Model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Starting Calibration (Initial Values)</th>
<th>Bayesian Estimates</th>
<th>SA Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>1</td>
<td>0.0899</td>
<td>0.8639</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.6</td>
<td>0.7389</td>
<td>0.5623</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.75</td>
<td>0.8801</td>
<td>0.6495</td>
</tr>
<tr>
<td>$\chi\pi$</td>
<td>1.5</td>
<td>1.0884</td>
<td>1.3342</td>
</tr>
<tr>
<td>$\chi_x$</td>
<td>0.12</td>
<td>0.1962</td>
<td>0.1131</td>
</tr>
<tr>
<td>$b_x$</td>
<td>0.5</td>
<td>1.85</td>
<td>0.4404</td>
</tr>
<tr>
<td>$b_{\pi}$</td>
<td>0.5</td>
<td>1.1198</td>
<td>0.4683</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>0.67</td>
<td>0.6186</td>
<td>0.6292</td>
</tr>
<tr>
<td>$\rho_u$</td>
<td>0.56</td>
<td>0.3657</td>
<td>0.5083</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>0.3</td>
<td>0.2235</td>
<td>0.2718</td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>0.42</td>
<td>0.7252</td>
<td>0.3443</td>
</tr>
<tr>
<td>$\rho_{\pi}$</td>
<td>0.61</td>
<td>0.8535</td>
<td>0.5099</td>
</tr>
<tr>
<td>Full Wald %</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>TM (Normalised t-Statistic)</td>
<td>28.5625</td>
<td>94.6459</td>
<td>20.3812</td>
</tr>
</tbody>
</table>

Note: The sample period of the US quarterly data is from 1969 to 2015.

4 Comparison through Estimation-based Test

4.1 TM Distance Comparison

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

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 a 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 those
of the initial calibration-based testing, as expected. This improvement can be attributed to applying the SA estimation approach, which explores all the potential parameters over wild space to discover the best fit.

Table 5: TM Distance (Normalised T-Statistics)

<table>
<thead>
<tr>
<th>Model</th>
<th>Starting Calibration</th>
<th>Bayesian Estimates</th>
<th>SA Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRE Model</td>
<td>4.1538</td>
<td>26.0498</td>
<td>0.6587</td>
</tr>
<tr>
<td>SI (j=4) Model</td>
<td>2.7338</td>
<td>-0.2072</td>
<td>0.1092</td>
</tr>
<tr>
<td>IF Model</td>
<td>28.5625</td>
<td>94.6459</td>
<td>20.3812</td>
</tr>
</tbody>
</table>

Note: (1) The sample period of the US quarterly data is from 1969 to 2015; (2) FIRE: Full-Information Rational Expectation Model; SI: Sticky Information Expectation Model; IF: Imperfect Information Data Revision Model.

4.2 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, the nominal interest rate increases, whereas the output gap and inflation decrease concerning the three competing models. As shown in Figure 1, the hump-shaped response only appears under the SI model throughout the effect of monetary policy shock on inflation and output gap. 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, they converge faster under the IF data revision model. Surprisingly, under the model with IF data revision, the effect of monetary policy shock fails to generate the hump-shape response on inflation and output gap and weakens the delay response on the interest rate.
Figure 1: Estimated Impulse Response Function of One Unit of Positive Policy Shock to the Main Variables

<table>
<thead>
<tr>
<th>FIRE Model</th>
<th>SI Model</th>
<th>IF Model</th>
</tr>
</thead>
</table>

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

4.3 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 lasts for a long time (i.e. around 20 periods) under the FIRE and SI models. However, the effects on three main variables are relatively short for the IF data revision model. Furthermore, the demand shock has a persistent impact on inflation and output gap under the 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

<table>
<thead>
<tr>
<th>FIRE Model</th>
<th>SI Model</th>
<th>IF Model</th>
</tr>
</thead>
</table>

Notes: x indicates output gap, pi indicates inflation and r indicates interest rate.
4.4 IRFs of Cost-Push Shock

Figure 3 shows the behaviour of the three main variables in response to the positive cost-push shock concerning the three competitors. Generally, all three competing models generate similar dynamics quantitatively. Specifically, the inflation and interest rate are affected positively by the positive cost-push shock, which negatively affects the output gap. In addition, the cost-push shock has the most considerable effect at the initial point under the FIRE model on three main variables. However, it has a moderate impact 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](image)

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

5 Robustness Check

5.1 Higher-order Auxiliary Models

We need to check whether the rank amongst the three competing models in 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, 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 6 indicate that increasing the order of VAR will make the non-rejection of all 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. When using 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 the model’s overall performance regardless of the auxiliary VAR models’ order. Thirdly, the ranking of the 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.

<table>
<thead>
<tr>
<th>Table 6: Model Performance under Different Auxiliary Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competing Model</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>TM Distance</td>
</tr>
<tr>
<td>(Full Wald %)</td>
</tr>
</tbody>
</table>

Notes: FIRE: Full-Information Rational Expectation Model; SI: Sticky Information Expectation Model; IF: Imperfect Information Data Revision Model.
5.2 Different Truncation Points $j$ of SI Model

This section needs 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 $8$ for the robustness check procedure. As shown in Table 7, we receive the same suggestion as the one provided by the Bayesian estimation approach, that is, incorporating more lagged information into the SI model influences its model performance after checking the TM distance (normalised t-statistics). Furthermore, the ranking amongst the three competitors is identical to the previous ranking no matter the value of the truncation point $j$ (i.e. $j = 6$ and $8$) in the SI model.

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

<table>
<thead>
<tr>
<th>Model</th>
<th>TM Using SA Estimation Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRE Model</td>
<td>0.6587</td>
</tr>
<tr>
<td>SI Model ($j=4$)</td>
<td>0.1092</td>
</tr>
<tr>
<td>SI Model ($j=6$)</td>
<td>-0.2796</td>
</tr>
<tr>
<td>SI Model ($j=8$)</td>
<td>-0.3518</td>
</tr>
<tr>
<td>IF Model</td>
<td>20.3812</td>
</tr>
</tbody>
</table>

Note: FIRE: Full-Information Rational Expectation Model; SI: Sticky Information Expectation Model; IF: Imperfect Information Data Revision Model.

5.3 Alternative Data Resource

The results obtained through the Bayesian estimation approach show that the performance of the IF model is far superior to its competitors. However, through indirect inference estimation, the full ability of the IF model is far inferior to its competitors. The estimation result using SPF data (survey data) is presented in Table 9. None of them can pass the test when each model is estimated using the survey data instead of real-time data—in addition, determining which one from the FIRE expectation model and the SI expectation model can better replicate the full dynamics of the actual observable. However, the SI model performs worse than the baseline when the SPF data are used.
Table 8: Starting Calibration Parameter Value of AR Coefficients

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Interpretation</th>
<th>FIRE</th>
<th>SI</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_g$</td>
<td>AR coefficient of demand shock</td>
<td>0.94</td>
<td>0.93</td>
<td>0.7</td>
</tr>
<tr>
<td>$\rho_u$</td>
<td>AR coefficient of cost-push shock</td>
<td>0.75</td>
<td>0.74</td>
<td>0.54</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>AR coefficient of policy shock</td>
<td>0.56</td>
<td>0.56</td>
<td>0.29</td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>AR term of shock in final revision process of $x$</td>
<td>-</td>
<td>-</td>
<td>0.39</td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>AR term of shock in final revision process of $\pi$</td>
<td>-</td>
<td>-</td>
<td>0.59</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>Model</th>
<th>SA Estimation Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRE Model</td>
<td>5.6900</td>
</tr>
<tr>
<td>SI ($j=4$) Model</td>
<td>5.2699</td>
</tr>
<tr>
<td>IF Model</td>
<td>12.4718</td>
</tr>
</tbody>
</table>

Note: FIRE: Full-Information Rational Expectation Model; SI: Sticky Information Expectation Model; IF: Imperfect Information Data Revision Model.

6 Conclusion

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

We compare three versions of the NK-DSGE model and find that none of them can fit the actual data through the initial calibrated-based test. We implement an 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. The comparisons of each competing model through Bayesian and SA estimation-based (indirect inference) tests are conducted. Three achievements can be reflected through the results of indirect inference estimation:
1) Regardless of two estimation methods (i.e. Bayesian and indirect inference estimations), the model with the SI expectation is the preferred approach amongst the three competitors by using real-time data.

2) We attempt to find a robust superior model in terms of dynamic performance by changing the 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.

3) The estimated IRFs analyse the effects of the structural shocks on the US economy.

Although the Bayesian estimation approach is an effective practical tool in inspecting a model’s performance by considering prior information about the macroeconomy, 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 indicates 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 the models can accept. 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 the 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 the estimation procedure, we find that the model’s performance increases through indirect inference, except for the cases of the FIRE model and SI model. This contradiction indicates that the survey data may contain useful information to improve the IF data revision model’s performance.
References


26


## Appendix A

### Table A1: Common Parameters

<table>
<thead>
<tr>
<th>Common Structural parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$ Elasticity of intertemporal substitution</td>
<td>1</td>
</tr>
<tr>
<td>$\alpha$ Sticky price degree</td>
<td>0.6</td>
</tr>
<tr>
<td>$\gamma$ Strategic complementary</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Common Taylor Rule in three models</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$ Degree of partially adjustment in Taylor rule</td>
<td>0.8</td>
</tr>
<tr>
<td>$\chi_\pi$ Coefficient of inflation on Taylor rule</td>
<td>1.5</td>
</tr>
<tr>
<td>$\chi_x$ Coefficient of output gap in Taylor rule</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Common Forcing Variables in three models</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_d$ AR coefficient of demand shock</td>
<td>0.5</td>
</tr>
<tr>
<td>$\rho_u$ AR coefficient of cost-push shock</td>
<td>0.5</td>
</tr>
<tr>
<td>$\rho_r$ AR coefficient of policy shock</td>
<td>0.5</td>
</tr>
<tr>
<td>$\rho_d$ Standard deviation of demand shock</td>
<td>0.3</td>
</tr>
<tr>
<td>$\rho_u$ Standard deviation of cost-push shock</td>
<td>0.3</td>
</tr>
<tr>
<td>$\rho_r$ Standard deviation of policy shock</td>
<td>0.3</td>
</tr>
</tbody>
</table>

**Note:** The priors of parameter are mostly chosen from previous literature (Milani and Rajbhandari, 2012; Smets and Wouters, 2003, 2007)
### Table A2: Inattentive Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_x$</td>
<td>Output coefficient in output revision process</td>
<td>0</td>
</tr>
<tr>
<td>$b_\pi$</td>
<td>Inflation coefficient in inflation revision process</td>
<td>0</td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>AR term of shock in final revision process of $x$</td>
<td>0.5</td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>AR term of shock in final revision process of $\pi$</td>
<td>0.5</td>
</tr>
<tr>
<td>$e_x$</td>
<td>SD of measurement error of $x$</td>
<td>0.25</td>
</tr>
<tr>
<td>$e_\pi$</td>
<td>SD of measurement error of $\pi$</td>
<td>0.25</td>
</tr>
</tbody>
</table>

**Imperfect Information model**

**Sticky Information model**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Share of updating firms</td>
<td>0.5</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Share of updating consumer</td>
<td>0.5</td>
</tr>
</tbody>
</table>

**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:

FEDFUNDS indicates an effective Federal Funds Rate. 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 \(^7\). 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 data revision 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 $\pi_t^r = (\ln P_t - \ln P_{t-1}) \times 100$. 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 of 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.

\(^7\)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\(^8\):

\[
x_t = \delta \sum_{j=0}^{\infty} (1 - \delta)^j E_{t-j} x_{t+1} - \sigma (\hat{r}_t - \pi_{t+1}) + g_t
\]

where \( \delta \) denotes 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\(^9\):

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

where \( \lambda \) denotes the share of the updating firms.

From the 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.

\(^8\)Initially, this is \( x_t = E_{t}^{SI} x_{t+1} - \sigma (\hat{r}_t - E_{t}^{SI} \pi_{t+1}) + g_t \).

\(^9\)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 \( \pi_t^r \) are taken as the observed variables realized at time \( t \) (they are the real-time data). In addition, \( x_t \) and \( \pi_t \) are the final revised variables, which are defined respectively as follows:

\[
\begin{align*}
x_t & \equiv x_t^r + v_x^t \quad \text{(D3)} \\
\pi_t & \equiv \pi_t^r + v_\pi^t \quad \text{(D4)}
\end{align*}
\]

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 \( \pi_t^r \). Thus, following his argument, we presume that the final revision process of the US output gap and inflation is defined as follows:

\[
\begin{align*}
v_x^t & = b_x x_t^r + e_x^t \quad \text{(D5)} \\
v_\pi^t & = b_\pi \pi_t^r + e_\pi^t \quad \text{(D6)}
\end{align*}
\]

These revision processes allow for a 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, \( e_x^t \) and \( e_\pi^t \), 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 behavioural parameters. Therefore, from the defined equation above, we can obtain the following:

\[ x_t \equiv x_t^r + v_t^x = (1 + b_x)x_t^r + e_t^x \] (D7)

\[ \pi_t \equiv \pi_t^r + v_t^\pi = (1 + b_\pi)\pi_t^r + e_t^\pi \] (D8)

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

\[ v_t^x = E_{t+1}v_t^x + e_t^x = b_xx_t^r + e_t^x \] (D9)

\[ v_t^\pi = E_{t+1}v_t^\pi + e_t^\pi = b_\pi\pi_t^r + e_t^\pi \] (D10)

\[ E_{t+1}v_t^x = b_xx_t^r \] (D11)

\[ E_{t+1}v_t^\pi = b_\pi\pi_t^r \] (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\(^{10}\):

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

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

We also use the identity equations \( E_{t+2}v_{t+1}^x = b_xx_{t+1}^r \) 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

\(^{10}\)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\(^\text{11}\):

$$x_t = (1 + b_x)E_t(x_{t+1}^r) - \sigma[\tilde{r}_t - (1 + b_x)E_t(\pi_{t+1}^r)] + g_t \quad (D14)$$

**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:

$$\pi_t = \beta E_t(\pi_{t+1}^r + E_{t+2}v_{t+1}^\pi) + \gamma\left(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha}\right)x_t + u_t \quad (D15)$$

Similarly, we use the identity equation to substitute out $E_t v_{t+1}^\pi$ from the above equation to obtain \(^\text{12}\)

$$\pi_t = \beta E_{t}^{IF} \pi_{t+1} + \gamma\left(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha}\right)x_t + u_t \quad (D16)$$

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

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

where the final revisions $v_t^x$ and $v_t^\pi$ their data can be constructed as demeaned observables between the first released $x_{t,t+1}$ and the latest released $x_{t,t+s}$ as follows:

$$v_t^x = (x_{t,t+1}^r - x_{t,t+s}^r) - M^{xx} \quad (D18)$$

$$v_t^\pi = (\pi_{t,t+1}^r - \pi_{t,t+s}^r) - M^{\pi x} \quad (D19)$$

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

\(^{11}\)Initially, this is $x_t = E_t(x_{t+1}^r + b_x x_{t+1}^r) - \sigma[\tilde{r}_t - E_t(\pi_{t+1}^r + b_x \pi_{t+1}^r)] + g_t$

\(^{12}\)Initially, this is $\pi_t = (1 + b_x)\beta E_t(\pi_{t+1}^r) + \gamma\left(\frac{(1-\alpha)(1-\alpha\beta)}{\alpha}\right)x_t + u_t$
revisions $v_t^r$ and $v_t^\pi$:

\begin{align*}
  v_t^r &= (x_{t,t+1}^r - x_{t,t+3}^r) - M^{vx3} \\
  v_t^\pi &= (\pi_{t,t+1}^r - \pi_{t,t+3}^r) - M^{\pi x3}
\end{align*}

Therefore, we can also construct the observations of the revised data $x_t$ and $\pi_t$.

Note that, as argued by Croushore (2011), if we look at the US data, we can see that $s$ is neither constant over time nor across variables. One may need to check whether the alternative of $s$ will significantly 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 are represented by $s$) and whether it is constant.