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

April 7, 2023

Abstract

We test the standard New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model with and without inattentiveness features, where inattentiveness is modelled either in the form of sticky information (SI) or imperfect information data revision (IF). All models are estimated and tested by Indirect Inference. The estimated model with sticky information somewhat outperformed the estimated NK model with full information and rational expectations (FIRE), though both passed the test easily; while the model with imperfect information data revision was strongly rejected.

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

JEL Classification:E12,E52,C52

*We are grateful for the constructive comments from editor and anonymous referee(s).We are also thanks for the 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 central 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 expectation are relatively rare. In this paper, we test DSGE models with different degrees of rationality in expectations against the macro data for the postwar US. Our starting point is a model of fully attentive expectations or full-information rational expectation (FIRE). We then explore the potential weakness of this early expectation assumption and to address it, examine two main assumptions that deviate from full-information rationality through inattentiveness. We focus in this paper on the two types of inattentiveness most commonly discussed. The first is sticky information (SI), developed by Mankiw and Reis (2002, 2007). The second is imperfect information (IF) data revision (Aruoba, 2008; Vázquez et al., 2010, 2012; Casares and Vázquez, 2016). We will carefully define both in what follows.

Simon (1989) criticises the FIRE hypothesis sa being unrealistic. He argues that economic agents, knowing all their problems, choices, and possible results, can certainly choose the best solution from all alternatives through reasonable calculation. However, in practice, such a 'perfect situation' cannot exist in the real world. Some unavoidable constraints always prevent economic agents from making good decisions (e.g., social and regulative constraints or the personal constraints of limited time and energy). Thus, economic agents have to seek coordination to achieve efficiency and profits. In other words, economic agents cannot simply reach their optimal solution but only reach a satisficing or 'good enough' solution. As a result, the FIRE cannot be used to explain economic decisions.

At the same time, 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 information acquisition, absorption, and procession abilities. Similar points are made by Caballero (2010), Stiglitz (2011) and Coibion and Gorodnichenko

(2012, 2015).

2 Three Competitors

To address these issues, the alternative theory of expectations inattentiveness has been proposed. Among the different approaches to inattentiveness, the two most prominent are SI (Mankiw and Reis, 2007) and IF data revision (Casares and Vázquez, 2016). These assumptions are what we explore 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 does not need to be jettisoned, nor do we need to introduce irrational behaviour to help the model fit the data (Collard et al., 2009; Coibion and Gorodnichenko, 2012). In this study, we focus only on the information assumptions.

Our first objective is to discover which expectations model best explains the US economy's behaviour in quarterly postwar data since the 1970s. We use indirect inference to evaluate the competing models solely on their ability to match the data behaviour. 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 based on priors that may well not be widely accepted; hence its rankings will not be treated as objective by those not sharing these priors. The indirect inference estimator locates the DSGE model whose simulated behaviour most closely matches the data behaviour as represented by an auxiliary model such as a vector autoregression (VAR), whose role is simply descriptive: the VAR is a natural choice since it is the reduced form of a DSGE model. Thus our estimate is the DSGE model whose simulated reduced form is closest to the reduced form of the unknown model generating the data. The estimation-based indirect inference test is implemented to discover the optimal set of parameters of the actual data in the context of the model to make a fair model comparison.

Our second objective of this study is to examine how incorporating inattentive features into the popular reduced-form New Keynesian model affects the model impulse responses to key macro shocks, including their degree of persistence and delay, compared with those of the standard FIRE model.

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 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 with SI, and third with 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 (PC) and Taylor rule (TR), 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, which is our focus here.

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

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 summary outline. The aggregate economy in our 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, the difference between actual and potential output (i.e., output under a flexible price economy). Coefficient σ represents the elasticity of 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.[1] The interest rate equation follows the simple 'interest rate smoothed' Taylor rule (Taylor, 1993). 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 $\lambda = 1$) (Mankiw and Reis, 2002, 2007; Reis, 2006a,b, 2009). Reis (2006a,b) provides deeper micro-foundations for the model features of SI. Under an SI environment, the inclusion of inattentiveness leads to deviations 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 on past expectations about the future. This is caused by information spreading slowly through the entire population of the economy (Mankiw and Reis, 2002).[2] The models are estimated for the US economy over five decades (from 1969 to 2015).

In comparison with the baseline FIRE model, the model with SI is more challenging

[1]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.

[2]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 and past expectations of the future inflation rates. By contrast, the current inflation in Mankiw and Reis' model is inferred from flexible price assumptions.

to solve. Given that SI involves infinitely lagged expectations, we consider how we can approximate the model with SI in the DSGE equilibrium framework. In the SI model, the proportion of lagged expectations diminishes geometrically. Hence the effect on economic agents' expectations derived from the current state is far more significant than that of previous periods. Consequently, the expectations that are formed exceptionally 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, indicating that lagged information up to four periods old is included); more extended lags, such as $j = 6$ and 8, are considered in the robustness check.[3]

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). For the IF data revision model, we must use real-time data. For example, if we were to analyse the economic agents' decisions using the data available to us today, then we would make incorrect inferences about their economic decision-making. We need to use the data at the time that economic agents made their economic decisions; this real-time analysis takes account of data revision.

Data revision is potentially critical theoretically and empirically, although this is not always properly recognised. Data is announced with lags. Furthermore, data revision may significantly influence empirical results, particularly in some variables defined conceptually. For instance, when economic agents make decisions about output (or the output gap), they will not know this variable without any doubt. A variable such 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.

As specified in the Appendix, we follow the suggestion by Casares and Vázquez (2016) for data revision. Collard and Dellas (2010) argue that few aggregate variables can be observed accurately, as the data revision process reveals. Thus price (inflation) and con-

[3] From the result of Trabandt (2007), by setting maximum $j = 19$, the 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.

sumption (output) can only be observed with random noise. In the IF three-equation model, x_t^r and π_t^r are taken as the observed variables realised at time t, while x_t and π_t are the final revised variables.

For each model with and without inattentive feature, an AR (1) process is assumed for all the disturbances to each structural equation to capture omitted variables. Each variable is demeaned and detrended.

3 Estimation through Indirect Inference

Three main shocks drive the stochastic dynamics: the monetary policy shock, the demand shock, and the cost-push shock. The models are estimated using three essential macroeconomic time series: real GDP, federal fund rates, and implicit price deflator. The HP filter is used to calculate the unobserved variable output gap (i.e., x_t).[4]

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 it to evaluate the model of the US economy, which is constructed by Smets and Wouters (2007) and ultimately reject the model. In the present work, we follow the standard procedure of indirect inference evaluation as set out in 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 assumes IF data revision issues occur for households

[4]In the DSGE literature, the theoretical concepts are not captured against specific data figures (e.g., GDP levels or inflation rate) but filtered data (e.g., HP filtered series). Filtering decomposes the data into a cyclical component and a trend component, and it is the cyclical component we usually feed into the model. By doing so, the analysis focuses on the business cycle frequencies, mainly because it is considered that DSGE models are better suited to explain short-term fluctuations rather than long-run growth

and firms rather than the monetary authority. Moreover, the agents can perfectly observe monetary policy. The other related paper is published by KNOTEK II (2010), who investigate a single-equation model incorporating sticky prices and SI. They find that such a model can match the real world in micro and macro dimensions after estimating it through indirect inference.[5] However, here we use a full structural model rather than a single-equation model.

We use a VAR as noted above, since this is the reduced form of the structural model. Specifically, we use a VAR(1) with our three variables; higher order VARs create excessive power under the indirect inference procedure, causing the undesirable rejection of models that are very close to the truth- as noted in Meenagh et al (Le et al., 2016; Meenagh et al., 2019).The Wald statistic measures the distance of the model from reality.

3.1 The Wald Test Statistic

The Wald testing process can be summarised by three steps.

First, the model's observed actual data and calibrated/estimated parameters imply the structural errors. These structural errors can be backed out directly from the structural equations and the actual data when the structural model possesses no expectation terms. When the structural equation includes expectations, the method used is the robust instrument variables estimation suggested by Wickens (1982), in which the lagged endogenous data are used as instruments. The fitted values are computed from the VAR (1) auxiliary model.

Second, the structural errors are bootstrapped to produce the simulated data from the candidate theoretical model. An auxiliary VAR model is then fitted to each set of simulated data, and the sampling distribution of the coefficients of the auxiliary VAR model is computed from these estimates of the auxiliary model.

Third, the Wald statistic is computed to determine whether the parameters of the

[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.

auxiliary VAR model estimated on the actual data lie in the confidence interval implied by this distribution. (Minford et al., 2015).

The Wald statistic is computed from the VAR(1) coefficients and the three variances of the three main variables as follows.

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

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 $G_T(\alpha_T) - G_S(\bar{\alpha}_S(\theta))$. 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 proceeds by considering the percentile of the Wald distribution - 100 minus this percentile is the p-value of the model, which can be used to rank the models' ability to match the data behaviour. Specifically, for a 5% significant level, a percentile above 95% lies outside the rejection area. In the following section we explain the estimation algorithm and our ranking methods in more detail.

3.2 Estimation and assessing model distance from the data behaviour

The TM statistic, a normalised t-statistic, is used when the model's 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 5% rejection boundary. The larger the number is, the worse the model fit will be and the less probable the model. The TM distance is defined as follows.

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

This function of TM distance is based on Wilson and Hilferty (1931) method of transforming the Chi-square distribution of the Wald into a standard normal distribution. Herein, the TM distance is the transformation of the Wald test statistics. WS_a is the Mahalanobis distance (the value of the Wald statistic) using the actual data, $WS_{s95\%}$ is the 95% critical Mahalanobis distance from simulated data (is the value of the Wald statistics falling at the 95th percentile of the bootstrap distribution) and p is the number of auxiliary model coefficients, defined as the degrees of freedom.

In estimation, we search for the DSGE parameter set,θ, giving the closest match to the auxiliary model estimates from the data. The algorithm we use for this purpose is based on simulated annealing (SA). This searches for the optimal parameter set by starting from a wide range around the initial values along with random jumps around the space.

The SA algorithm refers to a stochastic optimisation based on a 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 the Wald statistic, or until a given computation budget is exhausted. Given the principle of accepting a less optimal consequence temporarily, SA can reach the optimal parameter set globally instead of being trapped in a local optimum.

Initial values of the structural model's parameters are required to start the search. Here, these starting values are the calibrated values of the parameters[6], based on previous

[6]The starting search points (i.e., starting calibration parameter) will not affect our indirect inference

studies. We allow the algorithm to search within plus or minus 50% of the parameter starting values.

4 Estimates

4.1 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.

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
P-Value %	0	0	0.3413

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

Table 2 displays the estimation results of the FIRE model; we show the original calibration, and the Bayesian estimates carried out in Chou (2018).

All three stationary shocks are quite highly persistent and similar to their Bayesian estimates, except for the monetary policy error, whose AR coefficient is 60% higher than the Bayesian estimate.

The Wald percentile is 64.8, implying a p-value of 0.352. Thus, the FIRE model is not

testing results but only the search time for the optimal set of parameters. More precisely, the initial presumptive parameters used in the first stage may not produce optimal results. Therefore, we need to try to find out another collection of parameters.

rejected at the 5% significance level. Overall, many of the estimates obtained through SA estimation have shifted away from the Bayesian estimates, under which the model is firmly rejected, with a zero p-value. While the estimated values of the parameters are generally not far from those obtained by Bayesian analysis, there are important exceptions. Thus, the estimated value of the elasticity of inter-temporal substitution is 0.5180, which is higher, as is the extent of price stickiness. The estimated coefficients of monetary policy are also higher:χ_π by <8%, the other two (i.e., ρ and χ_x) by around 35%.

4.2 Estimation-based Indirect Inference Test: SI Expectation Model

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
P-Value %	0.006	0.46	0.469

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

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 the interest rate smoothing parameter ρ (0.7672), which is slightly lower. The reaction parameter of the output gap in monetary policy χ_x is estimated as around 13% lower than its Bayesian estimates but not far from its initial calibrated value. However, some SA estimates are higher than the Bayesian estimates, particularly the AR coefficient of monetary policy ρ_r, two times higher than that obtained through Bayesian estimation.

Furthermore, the test statistic indicates a Wald percentile of 53.10, and so a p-value of 0.469. Again, many SA estimates of the parameters differ from the Bayesian ones. 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 through the Bayesian approach. Nevertheless, the model under Bayesian estimation also matches the data behaviour reasonably well.

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

The table 4 shows that the IF model is strongly rejected, whether in calibrated form or on either Bayesian or indirect inference estimates. In all cases, the p-value is zero and the TM large.

5 Effects of shocks in the different models

The Table 5 below compares the data-matching performance of the three models. As can be seen, the SI and FIRE models both pass the matching test but the SI model is the more probable, getting closest to the auxiliary model behaviour. The IF model is strongly rejected.

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
P-Value %	0	0	0

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

Table 5: 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

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.

5.1 Model IRFs

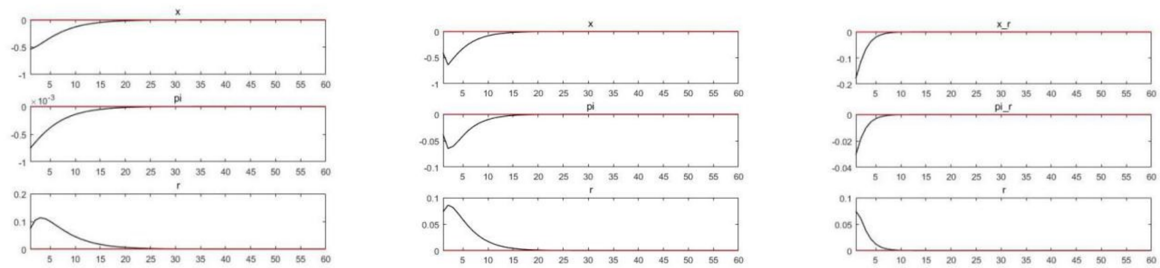
In this section we examine the model IRFs, noting that the IF model can largely be disregarded as it is strongly rejected.

5.1.1 IRFs of Monetary Policy Shock

Figure 1 shows the estimated impulse response to a restrictive monetary policy shock for the three models. For the FIRE and SI models, the size and duration of the effects are similar, but slightly more humped for SI. The IF effects are much smaller and disappear rapidly.

Figure 1: Estimated Impulse Response Function of One Unit of Positive Policy Shock to the Main Variables

FIRE Model SI Model IF Model



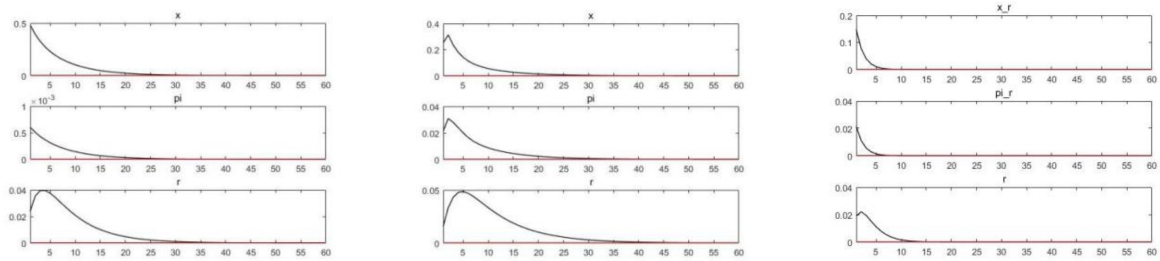
Notes: x indicates output gap, pi indicates inflation and r indicates interest rate.

5.1.2 IRFs of Demand Shock

Figure 2 presents the IRFs to the demand shock. The effect is biggest under the FIRE model and long drawn out under FIRE and SI; it is humped only under SI.

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

FIRE Model SI Model IF Model



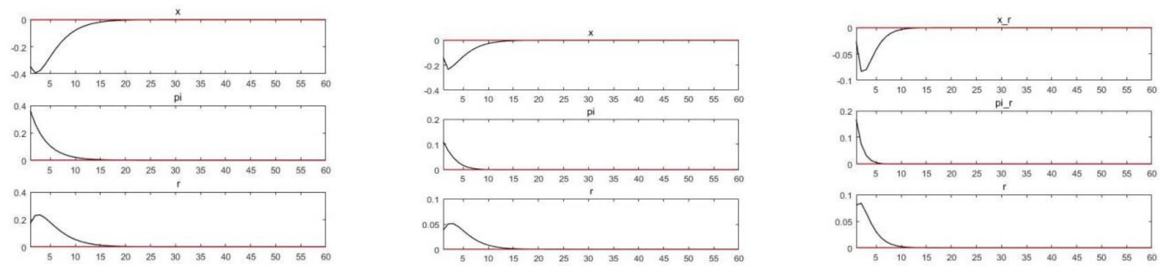
Notes: x indicates output gap, pi indicates inflation and r indicates interest rate.

5.1.3 IRFs of Cost-Push Shock

Figure 3 shows the IRFs for a positive cost-push shock concerning the three competitors. All three competing models generate similar dynamics. But the biggest effects occur under the FIRE model.

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.

5.2 Conclusions on model IRFs

What we see here is that the preferred SI model implies 'hump-like' responses for output, interest rates, and inflation, while the FIRE model implies relatively smoother responses.

The IF model IRFs imply quick spikes in response, falling off rapidly; these are implausible, no doubt explaining why the model fits the data behaviour so poorly.

6 Robustness Check:Different Truncation Points j of SI Model

Table 6: 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

Note: FIRE: Full-Information Rational Expectation Model; SI: Sticky Information Expectation Model; IF: Imperfect Information Data Revision 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 6, we receive the same suggestion as the one provided by the Bayesian estimation approach, 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.

7 Conclusion

In this paper we have tested the standard New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model with and without inattentive features, where inattentiveness is modelled either in the form of sticky information (SI) or imperfect information data revision (IF). All models are estimated and tested by Indirect Inference. The estimated model with sticky information somewhat outperformed the estimated NK model with full information and rational expectations (FIRE), though both passed the test easily; while the model with imperfect information data revision was strongly rejected.

The Bayesian version of the SI model also matched the data behaviour well, whereas the Bayesian versions of both the FIRE and IF models were strongly rejected. This suggests that the SI inattentiveness assumptions dominate the model's dynamic behaviour, sidelining the structural parameters.

The impulse responses of the estimated SI and FIRE models are quite similar in size and shape, though the SI ones show more delay and so are more hump-shaped, which may account for their somewhat closer match to the data behaviour. The IF responses are small, spiked and fall away rapidly, no doubt accounting for their complete failure to match the data behaviour.

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.

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

Chou, J.-Y. (2018). *Essays on the evaluation and estimation of the information friction in a DSGE model.* Ph. D. thesis, Cardiff University.

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.

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.

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

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.

Meenagh, D., P. Minford, M. Wickens, and Y. Xu (2019). Testing dsge models by indirect inference: a survey of recent findings. *Open Economies Review 30*(3), 593–620.

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.

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.

Wilson, E. B. and M. M. Hilferty (1931). The distribution of chi-square. *proceedings of the National Academy of Sciences of the United States of America 17*(12), 684.

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

Woodford, M. (2011). *Interest and prices: Foundations of a theory of monetary policy.* princeton university press.

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:

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 the time 2015Q4); the quarterly real-time GDP is the deviation of the natural the 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 the 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 was released after one period, and the data was released after three periods are also used.

[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 uses 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] :

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

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[9]:

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

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.

[8]Initially, this is $x_t = E_t^{SI}x_{t+1} - \sigma(\tilde{r}_t - E_t^{SI}\pi_{t+1}) + g_t$.
[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 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 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 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 behavioural parameters. Therefore, from the defined equation above, we can obtain the following:

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

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

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[10]:

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

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

[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[11]:

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

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 [12]

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

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

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

[12]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 over time nor across variables. One may need to check whether the alternative of s will significantly influence the imperfect information data revision's performance. 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 to 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.

Data Availability The data that support this study's findings are available from the corresponding author upon reasonable request.