View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Discussion Papers in Economics

ENDOGENOUS PERSISTENCE IN AN ESTIMATED DSGE MODEL UNDER IMPERFECT INFORMATION

By

Paul Levine

(University of Surrey)

Joseph Pearlman

(London Metropolitan University)

George Perendia

(London Metropolitan University)

&

Bo Yang

(University of Surrey)

DP 03/10

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk

ISSN: 1749-5075

Endogenous Persistence in an Estimated DSGE Model under Imperfect Information*

Paul Levine
University of Surrey

Joseph Pearlman
London Metropolitan University

George Perendia
London Metropolitan University

Bo Yang
University of Surrey

February 2010

Abstract

We provide a tool for estimating DSGE models by Bayesian Maximum-likelihood methods under very general information assumptions. This framework is applied to a New Keynesian model where we compare the standard approach, that assumes an informational asymmetry between private agents and the econometrician, with an assumption of informational symmetry. For the former, private agents observe all state variables including shocks, whereas the econometrician uses only data for output, inflation and interest rates. For the latter both agents have the same imperfect information set and this corresponds to what we term the '*informational consistency principle*'. We first assume rational expectations and then generalize the model to allow some households and firms to form expectations adaptively. We find that in terms of model posterior probabilities, impulse responses, second moments and autocorrelations, the assumption of informational symmetry by rational agents significantly improves the model fit. We also find qualified empirical support for the heterogenous expectations model.

JEL Classification: C11, C52, E12, E32.

Keywords: Imperfect Information, DSGE Model, Rational versus Adaptive Expectations, Bayesian Estimation

*To be presented at the 2010 Royal Economic Society Conference, March 29 – 31, University of Surrey. Earlier versions of this paper were presented at the Conference "Robust Monetary Rules for the Open Economy" at the University of Surrey, September 20-21, 2007; the CEF 2008 Annual Conference in Paris, June 26–28, 2008; a seminar at the Bank of England and at the annual CDMA Conference, 2009, University of St Andrews, 2–4 September; We acknowledge financial support for this research from the ESRC, project no. RES-000-23-1126 and from the EU Framework Programme 7 project MONFISPOL. Comments from participants at all these events are gratefully acknowledged, particularly those from the discussant at the Surrey Conference, Michel Juillard.

Contents

1 Introduction

A large recent literature has relaxed the extreme information assumptions of standard rational expectations in what are now referred to as Dynamic Stochastic General Equilibrium (DSGE) models. There are many approaches on offer ranging from those that stay within the conventional rational expectations paradigm to behavioural alternatives. In the former category are a number of refinements that assume that agents are not able to perfectly observe states that define the economy. Thus Pearlman *et al.* (1986) propose a general framework for introducing information limitations at the point agents form expectations. Pearlman (1992), Svensson and Woodford (2003) and Svensson and Woodford (2001) use this framework to study optimal monetary policy. Collard and Dellas (2004), Collard and Dellas (2006) (discussed below) investigate empirical issues associated with imperfect information. The 'Rational Inattention' literature that includes Mankiw and Reis (2002), Sims (2005), Adam (2007), Luo and Young (2009), Luo (2006)) and Reis (2009) fits into this agenda, the basic idea being that agents can process information subject to a constraint placing an upper bound on the information flow. The literature cited up to now all assumes homogeneous agents with a common information set, or a simple form of aggregation across staggered information up-dating; the examination of diverse agents with diverse information sets goes back to Townsend (1983) and has been recently developed by Woodford (2003) and Pearlman and Sargent (2003).

A relaxation of the rational expectations assumption itself is provided by the statistical rational learning literature pioneered by Evans and Honkapohja (2000) and adopted in a estimated macro-model by Milani (2007). This introduces a specific form of bounded rationality in which utility-maximizing agents make forecasts in each period based on standard econometric techniques such as least squares. In many cases this converges to a rational expectations equilibrium. A more drastic 'behavioural' alternative that limits the cognitive ability of agents still further is proposed by the 'animal spirits' approach of DeGrauwe (2009) where agents choose between, and learn about alternative simple forecasting rules.

At the same time the formal estimation of DSGE models by Bayesian methods has become standard.[1] However, as Levine *et al.* (2007) first pointed out, most of this DSGE estimation makes asymmetric information assumptions where perfect information about current shocks and other macroeconomic variables is available to the economic agents, but not to the econometricians. Although perfect information on idiosyncratic shocks may be available to economic agents, it is implausible to assume that they have full information on economy-wide shocks. It therefore makes sense to address empirically alternative information assumptions to assess whether parameter estimates are consistent across these assumptions and whether these alternatives lead to a better model fit.

In this paper we present two models: The first stays within the conventional rational expectations framework, but relaxes the extreme perfect information assumptions for the

[1]See Fernandez-Villaverde (2009) for a comprehensive and accessible review.

private sector. In a basic New Keynesian (NK) macro-model we make the assumption that either agents are better informed than the econometricians (the standard asymmetric information case in the estimation literature) or that they both have only the same imperfect information available, and that there is informational symmetry. We utilize the solution in the latter case, obtained for a completely general linear rational expectations model by Pearlman *et al.* (1986). The second model introduces heterogeneous expectations and encompasses the first. Proportions of households and firms form either rational or adaptive expectations. This captures the spirit of the simple learning rules in DeGrauwe (2009) and of the statistical learning literature whilst enabling the model to be expressed in a linear form. This, in turn, is essential for the Kalman-filter techniques we employ to solve the model under imperfect information. We elaborate on this point in section 4.

The symmetric information assumption is the informational counterpart to the "*cognitive consistency principle*" proposed in Evans and Honkapohja (2009) which holds that economic agents should be assumed to be "about as smart as, but no smarter than good economists". Whilst we make greater cognitive demands on rational agents, the formation of rational (model-consistent) expectations, our assumption that agents have no more information than the economist who constructs and estimates the model amounts to what we term the *informational consistency principle (ICP)*.[2] Certainly the ICP seems plausible – a central question is whether it adds realism to our model in practice by improving its empirical performance.

The possibility that imperfect information in NK models improves the empirical fit has been examined by Collard and Dellas (2004) and Collard and Dellas (2006), although an earlier assessment of the effects of imperfect information for an IS-LM model dates back to Minford and Peel (1983). They show that with imperfect information about output and the technology shock, or with misperceived money, the effect on inflation and output of a monetary shock is the hump-shaped one displayed empirically in VAR estimation. With perfect information, the hump-shaped effect is not in evidence in simulations of the NK model. Collard and Dellas (2006) in particular are able to reproduce this without resorting to lagged price indexation. The purpose of our paper is to investigate this issue formally within the Bayesian-maximum likelihood estimation framework examining model fit in terms of model posterior probabilities, impulse responses, second moments, autocorrelations and a comparison with a DSGE-VAR.[3]

[2]We are grateful to George Evans for pointing this out this analogy to us.

[3]Since writing an earlier version of this paper, we came across Collard *et al.* (2009) which carries out an exercise using the solution method of Pearlman *et al.* (1986) and Levine *et al.* (2007) that is similar to the analysis of our rational expectations model in some respects. We examine a more general behavioural model that encompasses that with rational expectations. A further distinguishing feature of our work is that our model validation alongside the marginal likelihood comparison is more comprehensive. But most importantly, whereas Collard *et al.* (2009) conclude that marginal likelihood differences between symmetric and asymmetric information assumptions are "rather small", we find very significant differences that are supported by our comparisons of second moments with those of the data and model impulse responses with that of a DSGE-VAR. This suggests that the importance of imperfect information for understanding business cycles may be underestimated by these authors.

The rest of the paper is organized as follows. Section 2 sets out the rational expectations model. Section 3 generalizes the model to a behavioural one where certain proportions of households and firms form rational and adaptive expectations. Section 4 sets out the solution method (summarizing Pearlman *et al.* (1986)) and pays particular attention to a technical but important issue of log-linearization. We also show that our framework encompasses the rational inattention approach of Sims (2005), Adam (2007) and Luo and Young (2009) as a special case. Sections 5 provides an analytical solution for a simplified version of our model that demonstrates how imperfect information gives rise to endogenous persistence in the sense that it is not solely driven by exogenous shocks. Sections 6 and 7 set out and discuss the results of our Bayesian estimation. Section 8 concludes.

2 The Rational Expectations Model

We utilize a fairly standard NK model with a Taylor-type interest rate rule. The simplicity of our model facilitates the separate examination of different sources of persistence in the model. First, the model in its most general form has external habit in consumption habit and price indexing. These are part of the model, albeit ad hoc in the case of indexing, and therefore endogenous. Persistent exogenous shocks to demand, technology and the price mark-up classify as exogenous persistence. A key feature of the model is a further endogenous source of persistence that arises when agents have imperfect information and learn about the state of the economy using Kalman-filter updating.

The full model in non-linear form is as follows

$$
\begin{aligned}
1 &= \beta(1+R_t)E_t\left[\frac{MU^C_{t+1}}{MU^C_t\Pi_{t+1}}\right] && (1)\\
\frac{W_t}{P_t} &= -\frac{1}{(1-\frac{1}{\eta})}\frac{MU^L_t}{MU^C_t} && (2)\\
MC_t &= \frac{W_t}{\alpha A_t P_t L_t^{\alpha-1}} && (3)\\
H_t-\xi\beta E_t[\tilde{\Pi}_{t+1}^{\zeta-1}H_{t+1}] &= Y_t MU^C_t && (4)\\
J_t-\xi\beta E_t[\tilde{\Pi}_{t+1}^{\zeta}J_{t+1}] &= \frac{1}{1-\frac{1}{\zeta}}MC_t MS_t Y_t MU^C_t && (5)\\
Y_t &= \frac{A_t L_t^\alpha}{\Delta_t} && (6)\\
\Delta_t &\equiv \frac{1}{n}\sum_{j=1}^{n}(P_t(j)/P_t)^{-\zeta} && (7)\\
1 &= \xi\tilde{\Pi}_t^{\zeta-1}+(1-\xi)\left(\frac{J_t}{H_t}\right)^{1-\zeta} && (8)\\
\tilde{\Pi}_t &\equiv \frac{\Pi_t}{\Pi_{t-1}^{\gamma}} && (9)
\end{aligned}
$$

$$Y_t = C_t + G_t \tag{10}$$

Equation (1) is the familiar Euler equation with β the discount factor, $1 + R_t$ the gross nominal interest rate, MU_t^C the marginal utility of consumption and $\Pi \equiv \frac{P_t}{P_{t-1}}$ the gross inflation rate, with P_t the price level. The operator $E_t[\cdot]$ denotes rational expectations conditional upon a general information set (see section 4). In (2) the real wage, $\frac{W_t}{P_t}$ is a mark-up on the marginal rate of substitution between leisure and consumption. MU_t^L is the marginal utility of labour supply L_t. Equation (3) defines the marginal cost. Equations (4) to (9) describe Calvo pricing with $1 - \xi$ equal to the probability of a monopolistically competitive firm re-optimizing its price, indexing by an amount γ with an exogenous mark-up shock MS_t. They are derived from the optimal price-setting first-order condition for a firm j setting a new optimized price $P_t^0(j)$ given by

$$P_t^0(j)E_t\left[\sum_{k=0}^{\infty}\xi^k D_{t,t+k}Y_{t+k}(j)\left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma}\right] = \frac{\kappa}{(1-1/\zeta)}E_t\left[\sum_{k=0}^{\infty}\xi^k D_{t,t+k}P_{t+k}MC_{t+k}Y_{t+k}(j)\right] \tag{11}$$

where the stochastic discount factor $D_{t,t+k} = \beta^k\frac{MU_{t+k}^C/P_{t+k}}{MU_t^C/P_t}$, and demand for firm j's output, $Y_{t+k}(j)$, is given by

$$Y_{t+k}(j) = \left(\frac{P_t^0(j)}{P_{t+k}}\right)^{-\zeta} Y_{t+k} \tag{12}$$

In equilibrium all firms that have the chance to reset prices choose the same price $P_t^0(j) = P_t^0$ and $\frac{P_t^0}{P_t} = \frac{J_t}{H_t}$ is the real optimized price in (7) and (8).

Equation (6) is the production function with labour the only variable input into production and the technology shock A_t exogenous. Price dispersion Δ_t, defined by (7), can be shown for large n, the number of firms, to be given by

$$\Delta_t = \xi\tilde{\Pi}_t^{\zeta}\Delta_{t-1} + (1-\xi)\left(\frac{J_t}{H_t}\right)^{-\zeta} \tag{13}$$

Finally (10), where C_t denotes consumption, describes output equilibrium, with an exogenous government spending demand shock G_t. To close the model we assume a current inflation based Taylor-type interest-rule

$$\log(1+R_t) = \rho_r\log(1+R_{n,t-1}) + (1-\rho_r)\left(\theta_\pi\log\frac{\Pi_t}{\Pi} + \log(\frac{1}{\beta}) + \theta_y\log\frac{Y_t}{Y}\right) + \epsilon_{e,t} \tag{14}$$

where $\epsilon_{e,t}$ is a monetary policy shock.[4]

The following form of the single period utility for household r is a non-separable function

[4]Note the Taylor rule feeds back on output relative to its steady state rather than the output gap so we avoid making excessive informational demands on the central bank when implementing this rule.

of consumption and labour effort that is consistent with a balanced growth steady state:

$$U_t = \frac{\left[(C_t(r) - h_C C_{t-1})^{1-\varrho}(1 - L_t(r))^{\varrho}\right]^{1-\sigma}}{1-\sigma} \tag{15}$$

where $h_C C_{t-1}$ is external habit. In equilibrium $C_t(r) = C_t$ and differentiating we have

$$\begin{aligned} MU_t^C &= (1-\varrho)(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)-1}(1-L_t)^{\varrho(1-\sigma)} && (16) \\ MU_t^L &= -(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)}\varrho(1-L_t)^{\varrho(1-\sigma)-1} && (17) \end{aligned}$$

Shocks A_t, G_t are assumed to follow AR(1) processes. Thus we have

$$\begin{aligned} \log\frac{A_{t+1}}{A} &= \rho_a \log\frac{A_t}{A} + \epsilon_{a,t+1} && (18) \\ \log\frac{G_{t+1}}{G} &= \rho_g \log\frac{G_t}{G} + \epsilon_{g,t+1} && (19) \end{aligned}$$

where A, G denote the non-stochastic balanced growth values or paths of the variables A_t, G_t. $\epsilon_{e,t}$, $\epsilon_{a,t}$ and $\epsilon_{g,t}$ are i.i.d. with mean zero and variances $\sigma^2_{\epsilon_e}$, $\sigma^2_{\epsilon_a}$ and $\sigma^2_{\epsilon_g}$ respectively. $\epsilon_{e,t}$ is assumed to be white noise. Following Smets and Wouters (2007) and others in the literature, we decompose the price mark-up shock into persistent and transient component: $MS_t = MS_{per,t}MS_{tran,t}$ where

$$\begin{aligned} \log\frac{MS_{per,t+1}}{MS_{per}} &= \rho_{ms}\log\frac{MS_{per,t}}{MS_{per}} + \epsilon_{msper,t+1} && (20) \\ \log\frac{MS_{tra,t+1}}{MS_{tra}} &= \epsilon_{mstra,t+1} && (21) \end{aligned}$$

This results in MS_t being an ARMA(1,1) process. We can normalize $A = 1$ and put $MS = MS_{per} = MS_{tra} = 1$ in the steady state. $\epsilon_{mstra,t}$, is also assumed to be i.i.d. with mean zero and variance $\sigma^2_{\epsilon_{mstra}}$. The innovations are assumed to have zero contemporaneous correlation. This completes the model. The equilibrium is described by 14 equations, (1)–(10) and (14)–(17) defining 13 endogenous variables Π_t $\tilde{\Pi}_t$ C_t Y_t Δ_t R_t MC_t MU_t^C U_t MU_t^L L_t H_t J_t and $\frac{W_t}{P_t}$. There are 4 shocks in the system: A_t, G_t, MS_t and $\epsilon_{e,t}$.

The log-linearization[5] of the model about the non-stochastic steady state is given by

$$\begin{aligned} y_t &= c_y c_t + (1-c_y)g_t \quad \text{where } c_y = \frac{C}{Y} && (22) \\ E_t mu^C_{t+1} &= mu^C_t - (r_t - E_t\pi_{t+1}) && (23) \\ \pi_t &= \frac{\beta}{1+\beta\gamma}E_t\pi_{t+1} + \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{(1-\beta\xi)(1-\xi)}{(1+\beta\gamma)\xi}(mc_t + ms_t) && (24) \end{aligned}$$

[5]Lower case variables are defined as $x_t = \log\frac{X_t}{X}$. r_t and π_t are log-deviations of *gross* rates. The validity of this log-linear procedure for general information sets is discussed in the next section.

where marginal utility, mu_t^C, and marginal costs, mc_t, are defined by

$$
\begin{aligned}
mu_t^C &= \frac{(1-\varrho)(1-\sigma)-1}{1-h_C}(c_t - h_C c_{t-1}) - \frac{\varrho(1-\sigma)L}{1-L} l_t \\
mc_t &= w_t - p_t - a_t + (1-\alpha) l_t \\
w_t - p_t &= mu_t^L - mu_t^C \\
y_t &= a_t + \alpha l_t \\
mu_t^L &= \frac{1}{1-h_C}(c_t - h_C c_{t-1}) + \frac{L}{1-L} l_t + mu_t^C
\end{aligned}
$$

Equations (22) and (23) constitute the micro-founded 'IS Curve' and demand side for the model, given the monetary instrument. According to (23) solved forward in time, the marginal utility of consumption is the sum of all future expected real interest rates. (24) is the 'NK Philips Curve', the supply side of our model. In the absence of indexing it says that the inflation rate is the discounted sum of all future expected marginal costs. Note that price dispersion, Δ_t, disappears up to a first order approximation and therefore does not enter the linear dynamics. Finally, shock processes and the Taylor rule are given by

$$
\begin{aligned}
g_{t+1} &= \rho_g g_t + \epsilon_{g,t+1} \\
a_{t+1} &= \rho_a a_t + \epsilon_{a,t+1} \\
msper_{t+1} &= \rho_{ms} msper_t + \epsilon_{msper,t+1} \\
ms_t &= msper_t + \epsilon_{mstra,t} \\
r_t &= \rho_r r_{t-1} + (1-\rho_r)[\theta_\pi \pi_t + \theta_y y_t] + \epsilon_{e,t}
\end{aligned}
$$

Bayesian estimation is based on the rational expectations solution of the log-linear model. The conventional approach assumes that the private sector has perfect information of the entire state vector mu_t^C, π_t, π_{t-1}, c_{t-1}, c_{t-1}^* and, crucially, current shocks $msper_t$, ms_t, a_t. These are extreme information assumptions and exceed the data observations on three data sets y_t, π_t and r_t that we subsequently use to estimate the model. If the private sector can only observe these data series (we refer to this as symmetric information) we must turn from a solution under *perfect* information on the part of the private sector (later referred to as *asymmetric information – AI* since the private sector's information set exceeds that of the econometrician) to one under *imperfect information – II.*

3 The Behavioural Model

We now assume that proportions λ_h and λ_f of households and firms respectively form rational expectations as before, but now the remaining agents form adaptive expectations. Our assumption of adaptive expectations differs from statistical learning as in Milani (2007), and from limiting the private sector to simple rules ("heuristics"), but allowing a degree of rationality through a selection process that evaluates their performance, as in DeGrauwe

(2009). These alternatives would be of interest in our imperfect information environment, but as yet our techniques cannot be applied to these cases (see section 4 for more discussion).

Consider first households. In principle those forming rational expectations could make different consumption and labour supply decisions. However we can avoid this complication by making the standard assumption of perfect insurance to equalize consumption decisions across the two types. Since they face the same real wage it follows from equating the marginal rate of substitution between consumption and leisure that labour supply decisions are the same too. What now changes for households is that expectations are composites of rational ($E_t^r[\cdot]$) and adaptive ($E_t^a[\cdot]$) with weights λ_h and $1-\lambda_h$ respectively. First order conditions for the representative household are now given by

$$1 = \beta(1+R_t)\left(\lambda_h E_t^r\left[\frac{MU_{t+1}^C}{MU_t^C\Pi_{t+1}}\right] + (1-\lambda_h)E_t^a\left[\frac{MU_{t+1}^C}{MU_t^C\Pi_{t+1}}\right]\right) \tag{25}$$

$$\frac{W_t}{P_t} = -\frac{1}{(1-\frac{1}{\eta})}\frac{MU_t^L}{MU_t^C} \tag{26}$$

Turning to firms, the optimal price-setting equation for *any* firm setting new prices is given as before by (11). However this choice is different for firms forming adaptive and rational expectations.[6] It is convenient to adopt the setup of in which we have perfectly competitive wholesale firms who produce a homogeneous good, which is bought by retail firms who differentiate the product at a fixed cost; then the real marginal cost of the wholesale firms, namely $MC_t = \frac{W_t}{\alpha A_t P_t L_t^{\alpha-1}}$ will be passed on to the retail firms, so that this is the same for all retail firms, be they rational or adaptive. Thus in (11) the only thing that differs for rational or adaptive firms is the way they form expectations. It follows that the RHS of the equations in J, H below are dependent on the economy-wide values of Y, MU^C, MC. Assume that the proportion of rational firms in the economy is λ_f. Then price setting and output equilibrium corresponding to (1)– (10) before are given by

$$\begin{aligned}
H_t^r - \xi\beta E_t^r[\tilde{\Pi}_{t+1}^{\zeta-1}H_{t+1}^r] &= Y_t MU_t^C \\
J_t^r - \xi\beta E_t^r[\tilde{\Pi}_{t+1}^{\zeta}J_{t+1}^r] &= \frac{1}{1-\frac{1}{\zeta}}MC_t MS_t Y_t MU_t^C \\
H_t^a - \xi\beta E_t^a[\tilde{\Pi}_{t+1}^{\zeta-1}H_{t+1}^a] &= Y_t MU_t^C \\
J_t^a - \xi\beta E_t^a[\tilde{\Pi}_{t+1}^{\zeta}J_{t+1}^a] &= \frac{1}{1-\frac{1}{\zeta}}MC_t MS_t Y_t MU_t^C \\
1 &= \xi\tilde{\Pi}_t^{\zeta-1} + (1-\xi)\lambda_f\left(\frac{J_t^r}{H_t^r}\right)^{1-\zeta} + (1-\xi)(1-\lambda_f)\left(\frac{J_t^a}{H_t^a}\right)^{1-\zeta} \\
\tilde{\Pi}_t &\equiv \frac{\Pi_t}{\Pi_{t-1}^{\gamma}} \\
MC_t &= \frac{W_t}{\alpha A_t P_t L_t^{\alpha-1}}
\end{aligned}$$

[6]DeGrauwe (2009) also has a model with composite expectations but fails to incorporate this aspect.

$$
\begin{aligned}
Y_t &= \frac{A_t L_t^{\alpha}}{\Delta_t} \\
\Delta_t &= \xi \tilde{\Pi}_t^{\zeta} \Delta_{t-1} + (1-\xi)\lambda_f \left(\frac{J_t^r}{H_t^r}\right)^{-\zeta} + (1-\xi)(1-\lambda_f)\left(\frac{J_t^a}{H_t^a}\right)^{-\zeta} \\
Y_t &= C_t + G_t
\end{aligned}
$$

Now we have a separate set of inflation dynamic for rational and adaptive firms with aggregate indexation-modified inflation $\tilde{\Pi}_t$ given by a weighted sum of optimized real prices. Similarly price dispersion, Δ_t, is now a weighted sum of contributions from the two types of firms.

Now define real optimized prices $Q_t^r \equiv J_t^r/H_t^r$, $Q_t^a \equiv J_t^a/H_t^a$ and put $\bar{Q}_t^r \equiv Q_t^r \tilde{\Pi}_t$ and $\bar{Q}_t^a \equiv Q_t^a \tilde{\Pi}_t$. We further distinguish between adaptive expectations of inflation by households and firms, $E_{h,t}^a[\pi_{t+1}]$ and $E_{f,t}^a[\pi_{t+1}]$ respectively. We can eliminate the latter and after some manipulation (see Appendix A) show that the log-linearization is as follows:

$$
\begin{aligned}
y_t &= c_y c_t + (1-c_y) g_t && (27) \\
\lambda_h E_t^r mu_{t+1}^C + (1-\lambda_h) E_{h,t}^a mu_{t+1}^C &= mu_t^C + \lambda_h E_t^r \pi_{t+1} + (1-\lambda_h) E_{h,t}^a \pi_{t+1} && (28) \\
\bar{q}_t^r - \xi\beta E_t^r \bar{q}_{t+1}^r &= \pi_t - \gamma\pi_{t-1} + (1-\beta\xi)(mc_t + ms_t) \\
&= \bar{q}_t^a - \xi\beta E_t^a \bar{q}_{t+1}^a && (29) \\
\pi_t - \gamma\pi_{t-1} &= (1-\xi)(\lambda_f \bar{q}_t^r + (1-\lambda_f)\bar{q}_t^a) && (30) \\
E_t^r \bar{q}_{t+1}^r - \xi\beta E_t^r \bar{q}_{t+2}^r &= (1-\xi\beta\mu_1) E_t^r \bar{q}_{t+1}^a - \xi\beta(1-\mu_1) E_{f,t}^a \bar{q}_{t+1}^a && (31)
\end{aligned}
$$

where mu_t^C, mc_t, shock processes and the Taylor rule are exactly as for the rational expectations model.[7] Adaptive expectations are given by

$$
\begin{aligned}
E_{f,t}^a \bar{q}_{t+1}^a &= \mu_1 \bar{q}_t^a + (1-\mu_1) E_{f,t-1}^a \bar{q}_t^a \\
E_{h,t}^a u_{c,t+1}^a &= \mu_2 mu_t^C + (1-\mu_2) E_{h,t-1}^a u_{c,t}^a \\
E_{h,t}^a \pi_{t+1}^a &= \mu_3 \pi_t + (1-\mu_3) E_{h,t-1}^a \pi_t^a
\end{aligned}
$$

Equations (27) and (28) constitute the 'IS' curve with composite expectations by the households. Equations (29) – (31) define the two NK Phillips Curves for rational and adaptive firms. (30) now gives aggregate inflation and (31) is the defining equation for $E_t^r \bar{q}_{t+1}^a$. Note that when all agents are rational, i.e., $\lambda_h = \lambda_f = 1$, then $\pi_t - \gamma\pi_{t-1} = (1-\xi)\bar{q}_t^r$ and we get back to the previous NK model.

[7] Note that for estimation purposes the coefficient on ms_t in (24) and (29) has been normalised to 1.

4 General Solution with Imperfect Information

Both RE and behavioural models are a special case of the following general setup in non-linear form

$$\begin{aligned} Z_{t+1} &= J(Z_t, E_tZ_t, X_t, E_tX_t) + \nu\sigma\epsilon_{t+1} && (32) \\ E_tX_{t+1} &= K(Z_t, E_tZ_t, X_t, E_tX_t) && (33) \end{aligned}$$

where Z_t, X_t are $(n-m)\times 1$ and $m \times 1$ vectors of backward and forward-looking variables, respectively, and ϵ_t is a $\ell \times 1$ shock variable, ν is an $(n-m)\times \ell$ matrix and σ is a small scalar. In log-linearized form with $z_t \equiv \log \frac{Z_t}{Z}$ where Z is the possibly trended steady state and $x_t \equiv \log \frac{X_t}{X}$. the state-space representation is

$$\begin{bmatrix} z_{t+1} \\ E_tx_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + B \begin{bmatrix} E_tz_t \\ E_tx_t \end{bmatrix} + \begin{bmatrix} u_{t+1} \\ 0 \end{bmatrix} \qquad (34)$$

where z_t, x_t are vectors of backward and forward-looking variables, respectively, and u_t is a shock variable; a more general setup allows for shocks to the equations involving expectations. In addition we assume that agents all make the same observations at time t, which are given by

$$\begin{aligned} W_t &= m(Z_t, E_tZ_t, X_t, E_tX_t) + \mu\sigma\epsilon_t && (35) \\ w_t &= \begin{bmatrix} M_1 & M_2 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + L \begin{bmatrix} E_tz_t \\ E_tx_t \end{bmatrix} + v_t && (36) \end{aligned}$$

in non-linear and linear forms respectively, where $\mu\sigma\epsilon_t$ and v_t represents measurement errors. Given the fact that expectations of forward-looking variables depend on the information set, it is hardly surprising that the absence of full information will impact on the path of the system.

In order to simplify the exposition we assume terms in E_tZ_t and E_tX_t do not appear in the set-up so that in the linearized form $B = L = 0$. Full details of the solution for the general setup are provided in Pearlman *et al.* (1986).[8]

4.1 Linear Approximation about the Non-Stochastic Steady State

Before proceeding to the rational expectations solution, we need to pose a basic question: is (34) linearized about the deterministic steady state, where expectations are conditional on any information set, a correct general form of the first-order approximation to the non-linear model above? In other words, up to a first order approximation, are the expected values of all variables in the non-linear model equal to their deterministic steady state values?

[8]Our model reduces to this form if we assume a pure inflation targeting rule with $\theta_y = 0$ in (14). In fact we find our empirical results to change very little with this simplification.

We draw upon and generalize the results of Schmitt-Grohe and Uribe (2004) on approximating non-linear RE models, Pearlman *et al.* (1986) PI solutions of linear RE models, and extended Kalman filter approximations for non-linear models. The latter is different from the standard engineering literature in which the Kalman filter is re-linearized at every stage (see Appendix B). However if the system is always close to the equilibrium, then there is no advantage to be gained from this, and we keep the linearization about the equilibrium. It is important to emphasize that there are no theorems to show that the extended Kalman filter guarantees convergence to the true nonlinear filter even for first-order deviations from the steady state. However it is a technique that is widely used, and the empirical evidence in its favour is good. The proofs in the appendix are therefore subject to the assumption that the approximation to the nonlinear filter is good to first order.

We now prove the following which establishes our requirement for the first order approximation:

Theorem

We look for a RE solution to to the non-linear model (32) and (33) under imperfect information which involves the innovations process variable $\tilde{Z}_t \equiv Z_t - E_{t-1}Z_t$:

$$X_t = g(Z_t, \tilde{Z}_t, \sigma)\,; \qquad Z_{t+1} = h(Z_t, \tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}\,; \qquad \tilde{Z}_{t+1} = f(\tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}$$

where σ is small. Then we have that $g_\sigma = h_\sigma = 0$.

Proof: See Appendix A.

This is the most important part of the generalization of Schmitt-Grohe and Uribe (2004), and the remainder represents a linearized version of Pearlman *et al.* (1986).

4.2 Solution Procedure

First assume *perfect information.* Following Blanchard and Kahn (1980), it is well-known that, there is then a saddle path satisfying:

$$x_t + Nz_t = 0 \quad \text{where} \quad \begin{bmatrix} N & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \Lambda^U \begin{bmatrix} N & I \end{bmatrix}$$

where Λ^U has unstable eigenvalues.

In the *imperfect information* case, following Pearlman *et al.* (1986), we use the Kalman filter updating given by

$$\begin{bmatrix} z_{t,t} \\ x_{t,t} \end{bmatrix} = \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix} + J \left[w_t - M \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix} \right]$$

where we denote $z_{t,t} \equiv E_t[z_t]$ etc. Thus the best estimator of the state vector at time $t-1$ is updated by multiple J of the innovation for the vector of observables $w_t - M \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix}$.

The matrix J is given by

$$J = \begin{bmatrix} PD^T \\ -NPD^T \end{bmatrix} \Gamma^{-1}$$

where $D \equiv M_1 - M_2 A_{22}^{-1} A_{21}$, $M \equiv [M_1 \; M_2]$ partitioned conformably with $\begin{bmatrix} z_t \\ x_t \end{bmatrix}$, $\Gamma \equiv EPD^T + V$ where $E \equiv M_1 - M_2 N$, $V = \text{cov}(v_t)$ is the covariance matrix of the measurement errors and P satisfies the Ricatti equation (40) below.

Using the Kalman filter, the solution as derived by Pearlman *et al.* (1986)[9] is given by the following processes describing the pre-determined and non-predetermined variables z_t and x_t and a process describing the innovations $\tilde{z}_t \equiv z_t - z_{t,t-1}$:

$$\begin{aligned}
\text{Predetermined}: \quad z_{t+1} &= Cz_t + (A - C)\tilde{z}_t + (C - A)PD^T(DPD^T + V)^{-1}(D\tilde{z}_t + v_t) \\
&+ u_{t+1} && (37) \\
\text{Non-predetermined}: \quad x_t &= -Nz_t + (N - A_{22}^{-1}A_{21})\tilde{z}_t && (38) \\
\text{Innovations}: \quad \tilde{z}_{t+1} &= A\tilde{z}_t - APD^T(DPD^T + V)^{-1}(D\tilde{z}_t + v_t) + u_{t+1} && (39)
\end{aligned}$$

$$\text{where} \quad C \equiv A_{11} - A_{12}N, \quad A \equiv A_{11} - A_{12}A_{22}^{-1}A_{21}, \quad D \equiv L_1 - L_2 A_{22}^{-1}A_{21}$$

and P is the solution of the Riccati equation given by

$$P = APA^T - APD^T(DPD^T + V)^{-1}DPA^T + U \tag{40}$$

where $U = \text{cov}(u_t)$ is the covariance matrix of the shocks to the system.

We can see that the solution procedure above is a generalization of the Blanchard-Kahn solution for perfect information by putting $\tilde{z}_t = v_t = 0$ to obtain

$$z_{t+1} = Cz_t + u_{t+1}; \quad x_t = -Nz_t \tag{41}$$

By comparing (41) with (37) we see that the determinacy of the system is independent of the information set. This is an important property that contrasts with the case where private agents use statistical learning to form forward expectations.

4.3 Statistical Learning

We now pose the question as to whether our framework can handle statistical learning. The work of Milani (2007) assumes that non-rational agents form expectations on the basis of having estimated the relationship between forward and backward-looking variables (including shocks) using discounted least squares. However he does this in the context of a very simple model in which there is only one representative agent, with all variables

[9] A less general solution procedure for linear models with imperfect information is provided by Lungu *et al.* (2008) with an application to a small open economy model, which they also extend to a non-linear version.

observable, and as a consequence all shocks are observable with a lag of one period; hence one can apply e-stability results to show that there is convergence of the system to an equilibrium, which coincides with the rational expectations equilibrium. Our model is more complex, requiring inferences to be made about the shocks, so that the filtered values are obtained rather than the true values. In general Bullard and Eusepi (2009), among others, show that there can be convergence to a system that exhibits indeterminacy. Any theory to account for this in estimation under imperfect information is currently non-existent, so we have taken a line of least resistance and assume non-rational agents form expectations adaptively.

4.4 Rational Inattention

On the theme of rational inattention, the fact that the dynamics of z_t depend on the dynamics of $\tilde{z}_t$ is equivalent to the result of Luo and Young (2009). For a simple stochastic growth model with rational inattention, they show that the dynamics of capital in their model, k_t, depends on $\hat{k}_t$ where the latter is last period's expected value of k_t, which in our notation would be $k_t - \tilde{k}_t$. it is also interesting to note that when there is only one predetermined variable in the system (as in Adam (2007) and Luo and Young (2009)), and it is observed with measurement error, then there is a one-to-one relationship between the variance of this error and the information channel capacity, the latter measuring the inverse of the degree of rational inattention. This is because if k_t has a normal distribution, then the difference in entropy at time t before and after a noisy measurement of k_t is a function[10] of $\sigma_k^2/(p_k + \sigma_k^2)$, where $p_k = var_{t-1}k_t$ and σ_k^2 is the variance of the noise. Thus if σ_k^2 is defined, then after solving the Riccati equation above, one can evaluate the capacity of the channel. Conversely, when the capacity is given, one can evaluate p_k/σ_k^2, followed by p_k from the Riccati equation, which then implies σ_k^2. When there are several predetermined variables, with noisy observations made on only one, then there is still a one-to-one relationship; thus if $k_t = h^T z_t$, then the difference in entropy is $\sigma_k^2/(h^T Ph + \sigma_k^2)$ where $P = var_{t-1}z_t$.

Thus our general framework with measurement error encompasses the rational inattention literature that assumes a single predetermined variable and relies on information channel capacity. However when more than one variable is observed with error, then the variance of the shock to measurements is a square matrix whose number of elements are obviously larger than the single parameter that represents the channel capacity. Thus we may consider estimating the capacity when there is one variable that is measured, but this does not easily generalise to the case when when there is more than one measurement per time period.

[10] For a Gaussian process the variance conditioned by the latest measurement is given by $p_k - p_k^2/(p_k+\sigma_k^2) = p_k\sigma_k^2/(p_k+\sigma_k^2)$, so that the defined value of the channel capacity is given by $\frac{1}{2}(log(p_k) - log(p_k\sigma_k^2/(p_k+\sigma_k^2))) = -\frac{1}{2}log(\sigma_k^2/(p_k+\sigma_k^2))$.

5 Analytical Example

To demonstrate the imperfect information solution procedure and the possible implications for endogenous persistence we consider a special case of our model without habit or indexation:

$$\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\beta\xi)(1-\xi)}{\xi} mc_t + ms_t$$

which for convenience we write as $E_t\pi_{t+1} = \frac{1}{\beta}\pi_t + x_t + w_t$ where $x_t \equiv \frac{(1-\beta\xi)(1-\xi)}{\beta\xi} mc_t$ and $w_t \sim N(0, \sigma_w^2)$ is now our transient shock to the mark-up. We now assume that x_t follows an exogenous AR(1) process

$$x_{t+1} = \rho x_t + \varepsilon_{t+1} \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

For our purposes this is most easily set up in the form

$$\begin{bmatrix} w_{t+1} \\ x_{t+1} \\ E_t[\pi_{t+1}] \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \rho & 0 \\ 1 & 1 & \alpha \end{bmatrix} \begin{bmatrix} w_t \\ x_t \\ \pi_t \end{bmatrix} + \begin{bmatrix} w_{t+1} \\ \varepsilon_{t+1} \\ 0 \end{bmatrix}$$

where $\alpha \equiv \frac{1}{\beta}$.

Under perfect information agents (somehow) observe the entire state vector consisting of the mark-up shock, the marginal cost and inflation. $[w_t \; x_t \; \pi_t]'$. We compare this with imperfect information where agents observe only inflation π_t with no measurement error. Then from our general solution procedure in section 4, the following matrices are defined

$$A = C = \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix}; \quad N = \begin{bmatrix} \frac{1}{\alpha} & \frac{1}{\alpha-\rho} \end{bmatrix} = -E; \quad D = \begin{bmatrix} -\frac{1}{\alpha} & -\frac{1}{\alpha} \end{bmatrix}; \quad U = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix}; \quad V = 0$$

It follows from (40) that

$$P = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & p \end{bmatrix} \qquad \text{where } p = \frac{\rho^2 p \sigma_w^2}{\sigma_w^2 + p} + \sigma_\varepsilon^2$$

From (37) it follows that the innovations are given by

$$\begin{bmatrix} \tilde{w}_{t+1} \\ \tilde{x}_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \tilde{w}_t \\ \tilde{x}_t \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \sigma_w^2 \\ p \end{bmatrix} \frac{(\tilde{w}_t + \tilde{\pi}_t)}{(p + \sigma_w^2)} + \begin{bmatrix} w_{t+1} \\ \epsilon_{t+1} \end{bmatrix}$$

Noting that $N - A_{22}^{-1}A_{21} = \begin{bmatrix} 0 & \frac{\rho}{\alpha(\alpha-\rho)} \end{bmatrix}$, it follows that the solution is given by

$$\begin{aligned} x_t &= \rho x_{t-1} + \varepsilon_t \\ \tilde{x}_t &= \frac{\rho}{\sigma_w^2 + p}(\sigma_w^2 \tilde{x}_{t-1} - p w_{t-1}) + \varepsilon_t \end{aligned}$$

$$\pi_t = -\frac{1}{\alpha}\left(1 + \frac{\rho\sigma_w^2 p}{(\alpha-\rho)(\sigma_w^2+p)}\right) w_t - \frac{1}{\alpha-\rho} x_t + \frac{\rho\sigma_w^2}{\alpha(\alpha-\rho)(\sigma_w^2+p)}\tilde{x}_t \quad (42)$$

Figure 1 in Appendix E illustrates the solution for $\beta = 0.99$, $\rho = 0.9$, $\sigma_\epsilon = 1$ and $\sigma_w^2 = 0, 1, 2$. The figure shows an impulse response to the mark-up, $x_0 = 1$. Under perfect information $\sigma_w^2 = 0$ and inflation is given by $\pi = -\frac{1}{\alpha-\rho}x_t$ with $x_t = \rho x_{t-1}$, $x_0 = 1$. Inflation jumps immediately to -9.1 but then proceeds to return to zero driven by the exogenous process for x_t. With imperfect information (II) the last term in (42) associated with the innovation introduces *endogenous persistence* arising from the rational learning of the private sector about this unobserved shock using Kalman updating. The inflation trajectory is now *hump-shaped* and the deviation from the v-shaped perfect information path increases as the variance of the transient shock σ_w^2 increases.

6 Bayesian Estimation

In the same year that Blanchard and Kahn (1980) provide a general solution for a linear model under RE in the state space form, Sims (1980) suggests the use of Bayesian methods for solving multivariate systems. This leads to the development of Bayesian VAR (BVAR) models (Doan *et al.* (1984)), and, during the 1980s, the extensive development and application of Kalman filtering-based state space systems methods in statistics and economics (Aoki (1987), Harvey (1989)).

Modern DSGE methods further enhance this Kalman filtering based Bayesian VAR state space model with Monte-Carlo Markov Chain (MCMC) optimising, stochastic simulation and importance-sampling (Metropolis-Hastings (MH) or Gibbs) algorithms. The aim of this enhancement is to provide the optimised estimates of the expected values of the currently unobserved, or the expected future values of the variables and of the relational parameters together with their posterior probability density distributions (Geweke (1999)). It has been shown that DSGE estimates are generally superior, especially for the longer-term predictive estimation than the VAR (but not BVAR) estimates (Smets and Wouters (2007)), and particularly in data-rich conditions (Boivin and Giannoni (2005)).

The crucial aspect is that agents in DSGE models are forward-looking. As a consequence, any expectations that are formed are dependent on the agents' information set. Thus unlike a backward-looking engineering system, the information set available will affect the path of a DSGE system.

The Bayesian approach uses the Kalman filter to combine the prior distributions for the individual parameters with the likelihood function to form the posterior density. This posterior density can then be obtained by optimizing with respect to the model parameters through the use of the Monte-Carlo Markov Chain sampling methods. Four variants of our linearized model are estimated using the Dynare software (Juillard (2003)), which has been extended by the paper's authors to allow for imperfect information on the part of the private sector.

In the process of parameter estimation, the mode of the posterior is first estimated using Chris Sim's csminwel after the models' log-prior densities and log-likelihood functions are obtained by running the Kalman recursion and are evaluated and maximized. Then a sample from the posterior distribution is obtained with the Metropolis-Hasting algorithm using the inverse Hessian at the estimated posterior mode as the covariance matrix of the jumping distribution. The scale used for the jumping distribution in the MH is set in order to allow a good acceptance rate (20%-40%). A number of parallel Markov chains of 100000 runs each are run for the MH in order to ensure the chains converge. The first 25% of iterations (initial burn-in period) are discarded in order to remove any dependence of the chain from its starting values.

6.1 Data, Priors and Model Identifiability

To estimate the system, we use three macro-economic observables at quarterly frequency for the US: real GDP, the GDP deflator and the nominal interest rate. Since the variables in the model are measured as deviations from a constant steady state, the time series are simply de-trended against a linear trend in order to obtain approximately stationary data. Following Smets and Wouters (2003), all variables are treated as deviations around the sample mean. Real variables are measured in logarithmic deviations from linear trends, in percentage points, while inflation (the GDP deflator) and the nominal interest rate are detrended by the same linear trend in inflation and converted to quarterly rates. The estimation results are based on a sample from 1970:1 to 2004:4.

The values of priors are taken from Levin *et al.* (2006) and Smets and Wouters (2007). Table 6 in Appendix D provides an overview of the priors used for each model variant described below. In general, inverse gamma distributions are used as priors when non-negativity constraints are necessary, and beta distributions for fractions or probabilities. Normal distributions are used when more informative priors seem to be necessary. We use the same prior means as in previous studies and allow for larger standard deviations, i.e. less informative priors, in particular for the habit parameter and price indexation. The priors on α, ξ are the exceptions and based on Smets and Wouters (2007) with smaller standard deviations. Also, for the parameters γ, h_C, ξ and ϱ we centre the prior density in the middle of the unit interval. The priors related to the process for the price mark-up shock are taken from Smets and Wouters (2007). The priors for $\mu_1, \mu_2, \mu_3, \lambda_h, \lambda_f$ are also assumed beta distributed with means 0.5 and standard deviations 0.2. Three of the structural parameters are kept fixed in the estimation procedure. These calibrated parameters are $\beta = 0.99$; $L = 0.4$, $c_y = 0.6$.

As emphasized by Canova and Sala (2009), it is necessary to confront the question of parameter identifiability in any DSGE model before taking the model to the data. Model/parameters identification is a prerequisite for the informativeness of different estimators, and their effectiveness when one uses the models to address policy questions and sources of identification failure could be marginalization (from the model structure), or lack of information (from the data).

Before estimating our models, we carry out a simple experiment examining parameter identifiability in our most general composite-expectations model. In this experiment, using the log-linearized solutions as the data generating process, we generate artificial data sets of length T = 5000 for all the observable variables from the DSGE model. To limit the influence of the initial conditions, we discard the first 100 observations. In particular, we simulate the data by imposing the prior means to the parameters. We then re-estimate the model on the artificial data sets using the standard Maximum Likelihood (ML) method and ask whether the ML estimates recover the DSGE model's priors. Convergence of the ML procedure then implies that the likelihood surface is not flat, suggesting there may be no identifiability problem.

One advantage of this technique is that it is completely independent of the nature or the size of the data used in estimation so that we can detect potential identification failures which are inherent in the model structure. The simulation and estimation results are then compared with the prior distributions and reported in Table 6. In the table, we measure the bias as the absolute value of the difference between the prior mean and the ML estimate for each parameter. We see that the bias is not markedly greater than one standard deviation in all cases, and much smaller in many cases indicating a 90% confidence interval for the true model. Overall the identification check suggests that identifiability in our DSGE model is generally very strong for much of the parameter space.

6.2 The Rational Expectations Model

We consider 4 model variants: GH ($\gamma, h_C > 0$), G ($h_C = 0$), H ($\gamma = 0$) and Z (zero persistence or $\gamma = h_C = 0$). Then for each model variant we examine three information sets: first we make the assumption that private agents are better informed than the econometricians (the standard asymmetric information case in the estimation literature) – the Asymmetric Information (AI) case. Then we examine two symmetric information sets for both econometrician and private agents: Imperfect Information without measurement error on the three observables r_t, π_t, y_t (II) and measurement error on *two* observables π_t, y_t (IIME). This gives 12 sets of results. First Table 7 in Appendix D reports the parameter estimates using Bayesian methods. It summarizes posterior means of the studied parameters and 90% confidence intervals for the four model specifications across the three information sets, AI, II and IIME, as well as the posterior model odds. Overall, the parameter estimates are plausible and reasonably robust across model and information specifications. The results are generally similar to those of Levin *et al.* (2006) and Smets and Wouters (2007) for the US, thus allowing us to conduct relevant empirical comparisons.

First it is interesting to note that the parameter estimates are fairly consistent across the information assumptions despite the fact that these alternatives lead to a considerably better model fit based on the corresponding posterior marginal data densities. On the other hand, the point estimates are relatively less robust across different model specifications, particularly for the Calvo price parameter and those in relation to the policy rule and process of mark-up shock.

Focusing on the parameters characterising the degree of price stickiness and the existence of real rigidities, we find that the price indexation parameters are estimated to be smaller than assumed in the prior distribution (in line with those reported by Smets and Wouters (2007)). The estimates of γ imply that inflation is intrinsically not very persistent in the relevant model specifications (the weight on lagged inflation in the Phillips curve is 0.27 implied by Model GH when assuming perfect information). If we assume an imperfect information set on GH, the model estimates that inflation is sightly more persistent as the weight becomes 0.33. The posterior mean estimates for the Calve price-setting parameter, ξ, obtained from Model GH across all the information sets imply an average price contract duration of about five quarters, similar to the findings of Christiano *et al.* (2005), Levin *et al.* (2006) and Smets and Wouters (2007). The external habit parameter is estimated to be around 80% of past consumption, which is somewhat higher than the estimates reported in Christiano *et al.* (2005), although this turns out to be a very robust outcome of the estimated models. The point estimates of h_C obtained from the imperfect information version seems to be slightly closer to the plausible values.

In Table 1 we report the posterior marginal data density from the estimation which is computed using the Geweke (1999) modified harmonic-mean estimator. The marginal data density can be interpreted as maximum log-likelihood values, penalized for the model dimensionality, and adjusted for the effect of the prior distribution (Chang *et al.* (2002)). Whichever model variant has the highest marginal data density attains the best relative model fit.

Model	AI	II	IIME
H	-238.20	-230.89	-231.37
G	-245.30	-239.15	-238.40
GH	-239.59	-230.95	-230.52
Z	-244.37	-242.04	-239.21

TABLE 1: Marginal Log-likelihood Values Across Model Variants and Information Sets

The model posterior probabilities are constructed as follows. Let $p_i(\theta|m_i)$ represent the *prior* distribution of the parameter vector $\theta \in \Theta$ for some model $m_i \in M$ and let $L(y|\theta, m_i)$ denote the likelihood function for the observed data $y \in Y$ conditional on the model and the parameter vector. Then the joint posterior distribution of θ for model m_i combines the likelihood function with the prior distribution:

$$p_i(\theta|y, m_i) \propto L(y|\theta, m_i)\, p_i(\theta|m_i)$$

Bayesian inference also allows a framework for comparing alternative and potentially misspecified models based on their marginal likelihood. For a given model $m_i \in M$ and common dataset, the latter is obtained by integrating out vector θ,

$$L(y|m_i) = \int_{\Theta} L(y|\theta, m_i)\, p(\theta|m_i)\, d\theta$$

where $p_i(\theta|m_i)$ is the prior density for model m_i, and $L(y|m_i)$ is the data density for model m_i given parameter vector θ. To compare models (say, m_i and m_j) we calculate the posterior odds ratio which is the ratio of their posterior model probabilities (or Bayes Factor when the prior odds ratio, $\frac{p(m_i)}{p(m_j)}$, is set to unity):

$$PO_{i,j} = \frac{p(m_i|y)}{p(m_j|y)} = \frac{L(y|m_i)p(m_i)}{L(y|m_j)p(m_j)} \tag{43}$$

$$BF_{i,j} = \frac{L(y|m_i)}{L(y|m_j)} = \frac{\exp(LL(y|m_i))}{\exp(LL(y|m_j))} \tag{44}$$

in terms of the log-likelihoods. Components (43) and (44) provide a framework for comparing alternative and potentially misspecified models based on their marginal likelihood. Such comparisons are important in the assessment of rival models.

Given Bayes factors we can compute the model probabilities $p_1, p_2, \cdots p_n$ for n models. Since $\sum_{i=1}^{n} p_i = 1$ we have that $\frac{1}{p_1} = \sum_{i=2}^{n} BF_{i,1}$, from which p_1 is obtained. Then $p_i = p_1 BF(i,1)$ gives the remaining model probabilities. These are reported in Table 2 where we denote the probability of variant G, information assumption II say, by Pr(G, II) etc.

Pr(GH, IIME)=0.3610
Pr(H, II)=0.2494
Pr(GH, II)=0.2348
Pr(H, IIME)=0.1543
Pr(H, AI)=0.0002
Pr(G, IIME)=0.0001
Pr(G, II)=0.0001
Remaining prob. are almost zero

Table 2: Model Probabilities Across Model Variants and Information Sets

Tables 1 and 2 reveal that a combination of Model GH and with information set IIME outperforms its 11 rivals with a posterior probability of 36%. However, the differences in log marginal likelihood or the posterior odds ratio are not substantive between Models GH and H under either IIME or II. For example, the log marginal likelihood difference between Model GH under IIME and Model H under II is 0.43. As suggested by Kass and Raftery (1995), in order to choose the former over later, we need a prior probability over Model GH under IIME 1.54 ($\approx e^{0.43}$) times larger than our prior probability over Model H under II. This factor is believed to be small and therefore we are unable to conclude that Model GH under IIME outperforms Model H under II. Equivalently, in Bayesian model comparison, a posterior Bayes factor needs to be at least 3 for there to be a positive evidence favouring Model m_i over m_j.

Our analysis of the model comparison contains several important results. First, price indexation does not improve the model fit, but the existence of habit is crucial as the results clearly suggest that incorporating habit persistence in consumption in the US model imparts greater inertia to the model, and improves the fit (relatively). Second, the II (or IIME) specification leads to significantly better fit for all model variants. Third, we find substantial evidence that the combinations of Models GH/H and IIME/II are far superior to any other combinations in terms of the ability to explain the data highlighting the importance of the underlying model persistence mechanisms and informational symmetry.

The focus on various alternative specifications seeks to address some of the concerns with Bayesian model comparisons pointed out by Sims (2003). By estimating a large number of model variants, this method intends to complete the space of competing models and to compute posterior odds that take into consideration other (seemingly irrelevant) aspects of the specification. One obvious limitation of this methodology is that the assessment of model fit is only relative to its other rivals with different restrictions. The outperforming model in the space of competing models may still be poor (potentially misspecified) in capturing the important dynamics in the data. To further evaluate the absolute performance of one particular model (or information assumption) against data, it is necessary to compare the model's implied characteristics with those of the actual data (or the VAR model).

6.3 The Behavioural Model

We consider the same four model variants and information sets as for the behavioural model. Table 8 in Appendix D reports the posterior estimates. The estimated policy coefficients are fairly robust across specifications and are reasonably consistent with those using the rational expectations model. The estimates of γ imply that inflation is less persistent compared to the previous model. The results from the behavioural model also show that the price stickiness parameter is estimated to be larger than assumed in the prior distribution. This implies that there is a considerable degree of price stickiness. Although this high degree of nominal stickiness in price is implausible and far from our priors, it is in line with the findings by Smets and Wouters (2005) and others.[11] Similar to the rational expectations model, the estimates of risk-aversion parameter are close to the prior assumption, indicating that the intertemporal elasticity of substitution (equal to $1/\sigma$ is less than one which is plausible as suggested in much of RBC literature.

Note that the estimates of the 5 parameters associated with adaptive expectations across all the specifications are statistically different from zero according to the 90% interval

[11] It reflects the low slope in the Phillips curve in output-inflation space as revealed by the data. This can be reconciled with a plausible degree of price stickiness without significantly changing the rest of the parameter estimates by either introducing Kimball preferences, as in Smets and Wouters (2007), or state-contingent price contracts, as in Gertler and Leahy (2008). To keep the core model relatively simple, we have chosen not to go down this route.

suggesting that they are playing important roles in the US economy. In particular, looking at Model H under the II assumption, the estimated λ_h and λ_f suggest that a share of around 70% of households form rational expectations while only around 20% firms are rational. But the sensitivity of restrictions is strong in these estimates as the estimated values of μ_2 and λ_h are smaller in the absence of habit persistence. Nevertheless, most of the estimated adaptive expectation parameters are tight estimates based the percentiles obtained from the estimation suggesting that they are statistically reliable.

Table 3 reports the posterior marginal data density obtained from estimating the behavioural model and Table 4 shows the probability ranking. We find that Model H with imperfect information set II has the highest marginal data density and attains the best relative model fit. This again suggests that incorporating the consumption habit seems to offer significant improvement in terms of the model fit to US data. The degree of consumption habit is also high and statistically significant using the behavioural model suggesting, in principle, the empirical relevance of the parameter. Also as found in the previous section, the data shows no support for the price indexation. Finally, in terms of using different information assumptions, we find that the II and IIME specifications do not lead to significantly better fit for all model variants. As one would expect, the ability of information symmetry to improve the model fit fall sharply when only 20% of firms turn out to form adaptive expectations in our estimation. There is evidence that the combinations of Model H and all three information assumptions are far superior to any other combinations in terms of the ability to explain the data, again highlighting the importance of the underlying habit persistence mechanisms.

Model	AI	II	IIME
H	-227.14	-226.55	-227.27
G	-242.87	-242.98	-243.15
GH	-230.87	-230.79	-230.72
Z	-239.27	-238.55	-239.12

Table 3: Marginal Log-likelihood Values Across Model Variants and Information Sets (The Behavioural Model)

6.4 Impulse Response Analysis

This subsection investigates the importance of shocks to the endogenous variables of interests by analyzing the impulse responses to the structural shocks in the models. As an alternative way of validating the model performance, we also compare the estimated DSGE model and an identified VAR model in terms of matching their impulse responses. To focus the presentation, this exercise is only performed for Model GH (the 'best' rational expectations model), Model Z (with zero persistence) and Model_COM_H (the 'best' behavioural model) across different information sets AI and II. The aim is to investigate the impact of

Pr(H, II)=0.4798
Pr(H, AI)=0.2660
Pr(H, IIME)=0.2335
Pr(GH, IIME)=0.0074
Pr(GH, II)=0.0069
Pr(GH, AI)=0.0064
Remaining prob. are zero

TABLE 4: Model Probabilities Across Model Variants and Information Sets (The Behavioural Model)

changing information assumptions in terms of the impulse response dynamics and to what extent the simplest NK model Z can be brought closer to the data by introducing imperfect information.

The estimated model impulse response functions (IRFs) can be directly related from the state space representation of the above economic model. To tackle the degree of freedom problem of the VAR models, a solution is to improve these by tilting them towards the values implied by the DSGE parameters. The latter impose a prior on the VAR, yielding the so-called DSGE-VAR approach proposed by Del Negro and Schorfheide (2004).

In general, their method implements the DSGE model prior by generating dummy observations from the DSGE model, and adding them to the actual data and leads to an estimation of the VAR based on a mixed sample of artificial and actual observations. The ratio of dummy over actual observations (called the hyperparameter λ) controls the variance and therefore the weight of the DSGE prior relative to the sample. If λ is small the prior is diffuse. For extreme values of this parameter (0 or ∞) either an unrestricted VAR or the DSGE model is estimated. The empirical performance of a DSGE-VAR will depend on the tightness of the DSGE prior. Details on the algorithm used to implement this DSGE-VAR are to be found in Del Negro and Schorfheide (2004) and Del Negro *et al.* (2005).

We fit our VAR to the same data set used to estimate the DSGE model. We consider a VAR with 4 lags.[12] We use a data-driven procedure to determine the tightness of the prior endogenously based on the marginal data density. Our choice of the optimal λ is 0.5 and this is found by comparing different VAR models using the estimates of the marginal data density. In particular, we iterate over a grid that contains the values of $\lambda = [0; 0.25; 0.5; 0.75; 1; 2; 5; \infty]$ and we find that the DSGE-VAR(4) with $\lambda = 0.5$ has the highest posterior probability.[13] This implies that the mixed sample that is used to estimate the VAR has slightly lower weight on the DSGE model (artificial observations) than on the VAR (actual observations).

Figure 2 in Appendix E depicts the mean responses corresponding to a positive one

[12]The choice of the lag length maximizes the marginal data density associated with the DSGE-VAR($\hat{\lambda}$).

[13]Alternatively, one can simply find the ‘optimal’ $\hat{\lambda}$ by estimating the parameter λ as one of the deep parameters.

standard deviation shock. The endogenous variables of interest are the observables in the estimation and each response is for a 10 period (2.5 years) horizon. All DSGE impulse responses are computed simulating the vector of DSGE model parameters at the posterior mean values reported in Table 5. The impulse responses for VAR(4) are obtained using the DSGE-VAR identification procedure. Overall, we find that the sign and magnitude of the DSGE and VAR impulse responses are quite similar implying that the DSGE model seems to mimic the VAR model in, at least, some dimensions. This confirms that the estimated DSGE model under both AI and II seems to be able to capture the main features of the US data. The overall impact of the model dynamics can be broadly described using the estimated impulse responses.

In response to an exogenous policy tightening, our model GH under asymmetric information (AI) predicts a decline in output that dies out within a few years, a gradual decrease in the inflation rate over several periods following a hump shaped response and a rise in the nominal interest rate. These findings are robust across many empirical studies and can be viewed as evidence of sizeable and persistent real effects of monetary policy shock captured by our model GH. When we assume informational symmetry, results for the DSGE model responses change dramatically. In particular, the imperfect information (II) specification produces a large hump-shaped decline in output (the peak effect occurs roughly over one year after the shock) and a gradual and lagged response in inflation when consumption habit and indexation are present. The larger decline and sluggish response of output to the policy shock in the II model show the evidence of endogenous persistence that is driven by informational symmetry. It is noteworthy that model GH succeeds in accounting for the inertial responses of inflation and output. Model Z without any persistence mechanisms fails to replicate the observed hump-shaped IRF for inflation under both information sets. It is interesting to note that our behavioural model H under both information sets manages to simulate a much larger hump-shaped decline in output (the peak effect occurs roughly over 5 years following the policy change) and does a better job at mimicking the response generated by the data. This shows further evidence of some endogenous persistence.

Following a positive technological shock, inflation and the interest rate fall gradually as higher productivity shrinks labour demand, pushing marginal cost down on impact, lowers prices and interest rate and monetary policy does not respond strongly enough to offset the downward pressure on marginal cost. Again these responses are predicted by many empirical studies on DSGE models (e.g. Levin *et al.* (2006) and Smets and Wouters (2007)) and the estimated reactions from our models account for these behaviours. In particular, Model GH when assuming II does well at accounting for the dynamic response of the US output to a productivity shock and Model Z when assuming II does a better job, compared to its AI counterpart, at predicting the reactions of inflation and interest rate computed from the data following a shock in technology. It is also worth noting that with AI the DSGE model somewhat overstates the initial responses particularly in inflation. In general, we conclude that the model's overall performance with respect to a technology shock is improved with informational symmetry. In addition, the overall performance of the behavioural model in matching the dynamic responses in the data is satisfactory, in particular, outperforming the rational expectation Model GH in predicting the response of inflation after a productivity shock.

With respect to the remaining shocks/responses, our models do well at accounting for the responses of output and interest rate to a government spending shock and the response of inflation to the transient part of a price mark-up shock. The qualitative effects are similar and the information specification does not seem to make a significant impact. The response of inflation following the government spending shock is somewhat overstated by our DSGE model under either information assumption. In terms of the persistent mark-up shock (referred to as Mark-up (ms) in Figure 2), the II assumption helps improve the model's performance in reflecting the central projection, particularly of inflation and output. To be specific, II helps generate a better shape of IRF while Model GH under AI predicts that output is not affected very much. Moreover, a result that is worth emphasizing is that Model Z when assuming II does very well at projecting the most likely after-shock path of inflation. Changing the information assumption slightly improves the IRFs of interest rate.

The simulations from the behavioural Model H using both II and AI show that the behavioural model does better at projecting the most likely after-shock path of almost all the variables in response to the government spending and the two versions of mark-up shocks, getting closer to the data. The only exception is the interest rate response after a government spending shock. Overall, these results from the estimated posterior impulse responses, combined with the simulated IRF based on the simple calibrated example (Figure 1), imply that the presence of the II specification or the assumption of agents' adaptive behaviour improve the fit of the model.

7 Further Model Validation

The summary statistics such as first and second moments have been standard for researchers to use to validate models in the literature on DSGE models, especially in the RBC tradition. As the Bayes factors (or posterior model odds) are used to assess the relative fit amongst a number of competing models, the question of comparing the moments is whether the models correctly predict population moments, such as the variables' volatility or their correlation, i.e. to assess the absolute fit of a model to macroeconomic data. Following Schorfheide (2000), let $\hat{y}_T$ be a sample of observation of length T that one could have observed in the past or that one might observe in the future. One can derive the sampling distribution of $\hat{y}_T$ given the current state of knowledge using the Bayes theorem: $p(\hat{y}_T|y_T) = \int L(\hat{y}_T|\theta)p(\theta|y_T)d\theta$. Assume that $T(y_T)$ is a test quantity that reflects an aspect of the data (moment) that one wants to check, e.g. correlation between output and inflation or the output volatility. In order to assess whether the estimated model can replicate population moments, one sequentially generates draws from the posterior distribution, $p(\theta|y_T)$ and the predictive distribution $p(\hat{y}_T|y_T)$ so that the predictive $T(\hat{y}_T)$ can be computed.

7.1 Standard Moment Criteria

To assess the contributions of assuming different information sets ***and proportions of adaptive agents*** in our estimated model variants, we compute some selected second moments and present the results in this subsection. Table 5 presents the second moments implied by the above estimations and compares with those in the actual data. In particular, we compute these model implied statistics by simulating the models at the posterior means obtained from estimation. The models are simulated by using 10000 series with 10000 periods. The first 1000 observations are dropped to eliminate the possible effect of initial conditions and an HP filter is applied before computing the moments to eliminate the possible trends. The results of model's second moments are compared with the second moments in the actual data to evaluate model's empirical performance for the selected model variants.

	Standard Deviation		
Model	Output	Inflation	Interest rate
Data	4.99	0.62	0.74
Model GH_AI	5.01	0.71	0.94
Model Z_AI	3.01	0.81	1.06
Model GH_II	4.57	0.67	0.88
Model Z_II	2.67	0.50	0.80
Model_COM_H_AI	2.77	0.51	0.66
Model_COM_H_II	2.58	0.49	0.65
	Cross-correlation with Output		
Data	1.00	-0.22	-0.36
Model GH_AI	1.00	-0.50	-0.71
Model Z_AI	1.00	-0.51	-0.46
Model GH_II	1.00	-0.47	-0.69
Model Z_II	1.00	-0.16	-0.21
Model_COM_H_AI	1.00	-0.19	-0.42
Model_COM_H_II	1.00	-0.18	-0.41
	Autocorrelations (Order=1)		
Data	0.96	0.85	0.94
Model GH_AI	0.98	0.88	0.95
Model Z_AI	0.95	0.91	0.96
Model GH_II	0.98	0.87	0.94
Model Z_II	0.98	0.89	0.95
Model_COM_H_AI	0.84	0.95	0.91
Model_COM_H_II	0.84	0.94	0.91

Table 5: Selected Second Moments

In terms of the standard deviations, almost all the rational expectations models generate relative high volatility compared to the actual data (except for output). In line with the Bayesian model comparison, Model GH (assuming II and rational expectations) fits the data better in terms of implied volatility, getting closer to the data in this dimension. Overall, the estimated models are able to reproduce acceptable volatility for the main variables of the DSGE model. The inflation volatilities implied by the models are close to that of the data. All rational expectations models under investigation appear to match well the autocorrelations (order=1) of all the endogenous variables. Using the behavioural model, output is less autocorrelated while inflation seems to be more autocorrelated then those in the data at order 1. Table 5 also reports the cross-correlations of the 3 observable variables vis-a-vis output. The data report that the inflation rate and nominal interest rate are countercyclical. All model variants perform successfully in generating the negative contemporaneous inflation-output and interest rate-output correlations observed in the data.

The 'preferred' model, Model GH (assuming imperfect information and rational expectations), does a better job at matching the data volatilities and first order autocorrelations, suggesting that habit formation and informational symmetry help fitting the data in these dimensions. In addition, the abilities of Model Z in capturing the inflation and interest rate volatilities and the contemporaneous cross-correlations are improved quite significantly when assuming there is informational symmetry. Overall, Bayesian Maximum-likelihood based methods suggest that all the implications of each model for fitting the data are contained in their likelihood functions. In other words, the simulation results mainly show that, switching from AI and II delivers a better fit to most features of the actual data, as suggested by the data and likelihood criterion.

The behavioural models, in general, are able to capture the main features of the data in most dimensions and strengthen the argument that the presence of partial rationality is supported by the data. In particular, Model H assuming adaptive expectations performs very well in generating and matching the contemporaneous cross-correlations of output with both inflation and interest rate, outperforming all its rivals. Model H is also more successful in replicating the interest rate volatility captured by the data. However, the main shortcoming that the behavioural model faces is the difficulty of replicating the output volatility.

7.2 Unconditional Autocorrelations

To further illustrate how the estimated models capture the data statistics based on different information or behavioural assumptions, we plot the unconditional autocorrelations of the actual data and those of the endogenous variables generated by the model variants in Figure 3. In general, all rational expectations models match reasonably well the autocorrelations shown in the data within a shorter period horizon and our 'best' rational expectations model, Model GH under II, does a slightly better job at matching the autocorrelations compared to its AI counterpart. The data report that all variables are positively and very significantly autocorrelated over short horizons. At a lag of one quarter, all the estimated models are able to generate the observed autocorrelations as noted above (except for the output autocorrelograms simulated by the behavioural model H), but at higher lags, the model simulated autocorrelations under AI are greater (more persistence) than those of the sample for the interest rate and inflation for the rational expectations model, but display *less* persistence for the behavioral model. When it comes to matching the interest rate a similar story applies except that the persistence switch-over of the behavoural model does not occur.

Of particular interest is that, when assuming II, the implied autocorrelograms produced by Model Z fit extremely well the observed autocorrelations of interest rate and inflation while its AI counterpart generates much sluggishness and is less able to match the inflation autocorrelation observed in the data from the second lag onwards. Imperfect information can therefore do much to improve the empirical performance of the simplest NK model Z with no added persistence mechanisms, though overall it loses out in comparison with models with these features. The results in this exercise again generally show that the DSGE models under II perform better at capturing the main features of the US data, strengthening the argument that the presence of informational symmetry helps improve the model fit to data.

8 Conclusions

Our paper makes both a methodological and substantive contribution to the macroeconomic literature on imperfect information. The methodological contribution is the provision of a general tool for estimating DSGE models by Bayesian Maximum-likelihood methods under very general information assumptions on the part of private agents. Our substantive contribution is an application to a NK model where we compare the standard approach, that assumes an informational asymmetry between private agents and the econometrician, with as assumption of informational symmetry. For the former private agents observe all state variables including shocks, whereas the econometrician only uses only data for output, inflation and interest rates. For the latter both agents have the same imperfect information set. For the rational model, we find that in terms of model posterior probabilities, impulse responses, second moments and autocorrelations, the assumption of informational symmetry significantly improves the model fit to data.The behavioural model easily wins the marginal likelihood race, but this must be qualified by the poor fit of output volatility and is failure to capture observed persistence in output and the interest rate.

There are three other notable results. First, we study variants of our model which close down the two endogenous persistence mechanisms of habit in consumption and indexing in turn. We then pose the question of whether imperfect information can provide an alternative source of endogenous persistence as illustrated in our simple analytical model. Indeed we find this is the case: our Model Z with neither mechanism and with imperfect (symmetric) information fits the observed autocorrelation of the data of the interest rate and inflation extremely well, whereas the same model with perfect (asymmetric) information on the part of the private sectors results in a poor fit in this dimension. Second we study symmetric information with measurement error for the observed macroeconomic series and find this improves the fit still further, thought the increase in the model probability is not significant. Finally there is little to be gained from the indexation mechanism in terms of model fit, an encouraging result for our workhorse NK model as price-indexation is generally deemed to be an unsatisfactory ad hoc compromise feature of this genre.

There are a number of directions for future research. We have deliberately chosen to apply our methodology to a relatively simple NK model with only few frictions. Having demonstrated that information plays an important role for the estimation of this model, the next step would be to examine its implications for closed- and open-economy models with a range of frictions such as Smets and Wouters (2007) and Adolfson *et al.* (2007), respectively. Second as alluded to in the introduction there are other ways of modelling information limitations associated with the rational inattention literature. We have shown that our general framework with a single measurement error is equivalent to models in the rational inattention literature that assumes a single predetermined variable and rely on information channel capacity. However a formal comparison with the sticky information approach of Mankiw and Reis (2002) would be of some interest. Finally optimal policy needs to be examined making consistent information assumptions at the estimation and policy analysis stages. If imperfect information on the part of the private sector proves (as in our rational model) to be strongly supported empirically in a range of DSGE models with various frictions, this suggests that the imperfect information solution of optimal policy set out in Pearlman (1992) is appropriate.

References

Adam, K. (2007). Optimal Monetary Policy with Imperfect Common Knowledge . *Journal of Monetary Economics*, **54**(2), 267–301.

Adolfson, M., Laseen, S., Linde, J., and Villani, M. (2007). Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through. *Journal of International Economics*, **72**(2), 481–511.

Aoki, M. (1987). *State Space Modelling of Time Series*. Springer-Verlag.

Blanchard, O. and Kahn, C. (1980). The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, **48**, 1305–1313.

Boivin, J. and Giannoni, M. (2005). DSGE Models in a Data-Rich Environment. *mimeo*.

Bullard, J. and Eusepi, S. (2009). When Does Determinacy Imply Expectational Stability? St Louis Fed Working Paper 2008-007B.

Canova, F. and Sala, L. (2009). Back to Square One: Identification Issues in DSGE Models. *Journal of Monetary Economics*, **56**(4), 431–449.

Chang, Y., Gomes, J. F., and Schorfheide, F. (2002). Learning-by-Doing as a propagation Mechanism. *American Economic Review*, **92**(5), 1498–1520.

Christiano, L., Eichenbaum, M., and Evans, C. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, **113**, 1–45.

Collard, F. and Dellas, H. (2004). The New Keynesian Model with Imperfect Information and Learning. *mimeo, CNRS-GREMAQ*.

Collard, F. and Dellas, H. (2006). Misperceived Money and Inflation Dynamics. *mimeo, CNRS-GREMAQ*.

Collard, F., Dellas, H., and Smets, F. (2009). Imperfect Information and the Business Cycle. *Journal of Monetary Economics*. Forthcoming.

DeGrauwe, P. (2009). Animal Spirits and Monetary Policy. Mimeo, University of Leuven.

Del Negro, M. and Schorfheide, F. (2004). Priors from General Equilibrium Models for VARs. *International Economic Review*, **45**, 643–673.

Del Negro, M., Schorfheide, F., Smets, F., and Wouters, R. (2005). On the fit and forecasting performance of New-Keynesian models. ECB Working Paper No. 491.

Doan, J., Litterman, R., and Sims, C. (1984). Forecasting and Conditional Projection Using Realistic Prior Distributions. *NBER*.

Evans, G. W. and Honkapohja, S. (2000). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Evans, G. W. and Honkapohja, S. (2009). Learning and Macroeconomics. *Annual Review of Economics*, **1**, 421–449.

Fernandez-Villaverde, J. (2009). The Econometrics of DSGE Models. CEPR Discussion Paper No. 7159.

Gertler, M. and Leahy, J. (2008). A Phillips Curve with an Ss Foundation. *Journal of Political Economy*, **113**(3), 533 – 572.

Geweke, J. (1999). Computational Experiments and Reality. University of Minnesota and Federal Reserve Bank of Minneapolis.

Harvey, A. (1989). *Forecasting Structural Time Series and the Kalman Filter*. CUP.

Juillard, M. (2003). DYNARE: A Program for Solving Rational Expectations Models. CEPREMAP.

Julier, S. and Uhlmann, J. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, **92**, 401–422.

Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, **90**, 773–795.

Levin, A., Onatski, A., Williams, J. C., and Williams, N. (2006). Monetary Policy Under Uncertainty in Micro-Founded Macroeconomic Models. in M. Gertler and K. Rogoff (eds.), NBER Macroeconomics Annual, 2005, pp 229–387 .

Levine, P., Pearlman, J., and Perendia, G. (2007). Estimating DSGE Models under Partial Information. Department of Economics Discussion Papers 1607, Department of Economics, University of Surrey .

Lungu, L., Matthews, K., and Minford, A. (2008). Partial Current Information and Signal Extraction in a Rational Expectations Macroeconomic Model: A Computational Solution. *Economic Modelling*, **25**(2), 255–273.

Luo, Y. (2006). Consumption Dynamics under Information Processing. Mimeo, University of Hong Kong.

Luo, Y. and Young, E. R. (2009). Rational Inattention and Aggregate Fluctuations. *The B.E. Journal of Macroeconomics*, **9**(1). Article 14.

Mankiw, N. and Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics*, **117**(4).

Milani, F. (2007). Expectations, Learning and Macroeconomic Persistence. *Journal of Monetary Economics*, **54**(7), 2065–2082. Forthcoming.

Minford, A. and Peel, D. (1983). Some Implications of Partial Information Sets in Macroeeconomic Models Embodying Rational Expectations. *Manchester School*, **51**, 235–249.

Pearlman, J. G. (1992). Reputational and Non-Reputational Policies with Partial Information. *Journal of Economic Dynamics and Control*, **16**, 339–357.

Pearlman, J. G. and Sargent, T. J. (2003). Knowing the Forecasts of Others. *Review of Economic Dynamics*, page to appear.

Pearlman, J. G., Currie, D., and Levine, P. (1986). Rational Expectations Models with Private Information. *Economic Modelling*, **3**(2), 90–105.

Reis, R. (2009). A Sticky Information General Equilibrium Model for Policy Analysis. NBER WP No. 14732.

Schmitt-Grohe, S. and Uribe, M. (2004). Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. *Journal of Economic Dynamics and Control*, **28**, 755 –775.

Schorfheide, F. (2000). Loss Function-based Evaluation of DSGE Models. *Journal of Applied Econometrics*, **15**(6), 645–670.

Sims, C. (2003). The Role of Models and Probabilities in the Monetary Policy Process. *Brookings Paper on Economic Activity.*

Sims, C. (2005). Rational Inattention: A Research Agenda. Deutche Bundesbank, W.P. no. 34/2005.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, **48**, 1–48.

Smets, F. and Wouters, R. (2003). An estimated Stochastic Dynamic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association*, **1(5)**, 1123–1175.

Smets, F. and Wouters, R. (2005). Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE Approach. *Journal of Applied Econometrics*, **20**(2), 161–183.

Smets, F. and Wouters, R. (2007). Shocks and Frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, **97**(3), 586–606.

Svensson, L. E. O. and Woodford, M. (2001). Indicator variables for Optimal Policy under Asymmetric Information. *Journal of Economic Dynamics and Control*, **28**(4), 661–680.

Svensson, L. E. O. and Woodford, M. (2003). Indicator variables for Optimal Policy. *Journal of Monetary Economics*, **50**(3), 691–720.

Townsend, R. M. (1983). Forecasting the forecasts of others. *Journal of Political Economy*, **91**, 546–588.

Woodford, M. (2003). Imperfect Common Knowledge and the Effects of Monetary Policy. In P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, editors, *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S Phelps*. Princeton University Press.

Appendix

A Linearization of Behavioural Model

Log-linearizing the non-linear system as it stands gives

$$q_t^r = \xi\beta E_t^r[\pi_{t+1} + q_{t+1}^r] - \xi\beta\gamma\pi_t + (1-\beta\xi)(mc_t + ms_t) \tag{A.1}$$

$$q_t^a = \xi\beta E_{f,t}^a[\pi_{t+1} + q_{t+1}^a] - \xi\beta\gamma\pi_t + (1-\beta\xi)(mc_t + ms_t) \tag{A.2}$$

$$w_t - p_t = mu_t^L - mu_t^C \tag{A.3}$$

$$mc_t = w_t - p_t - a_t + (1-\alpha)l_t \tag{A.4}$$

$$\lambda_h E_t^r mu_{t+1}^C + (1-\lambda_h)E_{h,t}^a mu_{t+1}^C = mu_t^C - (r_t - \lambda_h E_t^r \pi_{t+1} - (1-\lambda_h)E_t^{h,a}\pi_{t+1}) \tag{A.5}$$

$$mu_t^C = \frac{(1-\varrho)(1-\sigma)-1}{1-h_C}(c_t - h_C c_{t-1}) - \frac{\varrho(1-\sigma)L}{1-L} l_t \tag{A.6}$$

$$mu_t^L = \frac{1}{1-h_C}(c_t - h_C c_{t-1}) + \frac{L}{1-L} l_t + mu_t^C \tag{A.7}$$

$$y_t = c_y c_t + (1-c_y)g_t \tag{A.8}$$

$$\xi(\pi_t - \gamma\pi_{t-1}) = (1-\xi)(\lambda_f q_t^r + (1-\lambda_f)q_t^a) \tag{A.9}$$

where we distinguish between adaptive expectations of inflation by households and firms, $E_{h,t}^a[\pi_{t+1}]$ and $E_{f,t}^a[\pi_{t+1}]$ respectively.. However the following obviates the need to define separate adaptive process for firms for both q_t^a and π_t. Define $\bar{q}_t^a = q_t^a + \pi_t - \gamma\pi_{t-1}, \bar{q}_t^r = q_t^r + \pi_t - \gamma\pi_{t-1}$. Then (A.1), (A.2) and (A.9) become

$$\bar{q}_t^r = \xi\beta E_t^r \bar{q}_{t+1}^r + \pi_t - \gamma\pi_{t-1} + (1-\beta\xi)(mc_t + ms_t) \tag{A.10}$$

$$\bar{q}_t^a = \xi\beta E_{f,t}^a \bar{q}_{t+1}^r + \pi_t - \gamma\pi_{t-1} + (1-\beta\xi)(mc_t + ms_t) \tag{A.11}$$

$$\pi_t - \gamma\pi_{t-1} = (1-\xi)(\lambda_f \bar{q}_t^r + (1-\lambda_f)\bar{q}_t^a) \tag{A.12}$$

Equation (A.12) represents the central cognitive requirement for the adaptive firms, and is central to deriving the remaining relationships of the model. We assume that these agents (1) observe the overall inflation rate π_t and since they choose their own optimal relative price q_t^a at time t they therefore observe $\bar{q}_t^a = q_t^a + \pi_t - \gamma\pi_{t-1}$, and (2) know the value of ξ as well as the proportion of adaptive firms $1-\lambda_f$. Armed with this information they can deduce the value of $\bar{q}_t^r$ from (A.12). Similarly rational firms can deduce the value of $\bar{q}_t^a$.

We also need to define adaptive processes for $E_{h,t}^a[\pi_{t+1}]$, $E_{h,t}u_{c,t+1}^a$ and $E_{f,t}\bar{q}_{t+1}^a$:

$$\begin{aligned} E_{f,t}\bar{q}_{t+1}^a &= \mu_1 \bar{q}_t^a + (1-\mu_1)E_{f,t-1}\bar{q}_t^a \\ E_{h,t}u_{c,t+1}^a &= \mu_2 mu_t^C + (1-\mu_2)E_{h,t-1}u_{c,t}^a \\ E_{h,t}\pi_{t+1}^a &= \mu_3 \pi_t + (1-\mu_3)E_{h,t-1}\pi_t^a \end{aligned}$$

Since adaptive and rational agents need to form their own (differing) expectations of future inflation we need to explain how expectations are formed of one another's expectations of future $\bar{q}_t^a$

and $\bar{q}^r_t$. We first note that (A.11) and (A.10) imply that

$$\bar{q}^r_t - \xi\beta E^r_t \bar{q}^r_{t+1} = \bar{q}^a_t - \xi\beta E^a_{f,t}\bar{q}^a_{t+1} \tag{A.13}$$

so that $E^r_t \bar{q}^r_{t+1}$ is known to adaptive firms. Therefore taking expectations $E^a_{f,t}$ of (A.13) and noting that $E^a_{f,t}\bar{q}^r_t = \bar{q}^r_t$ we deduce that

$$E^a_{f,t}[E^r_t \bar{q}^r_{t+1}] = E^r_t \bar{q}^r_{t+1} \tag{A.14}$$

We now make the assumption that adaptive agents do not think they are able to have better expectations of $E^a_{f,t}\bar{q}^r_{t+1}$ than do rational agents, meaning that they assume $E^a_{f,t}\bar{q}^r_{t+1} = E^a_{f,t}[E^r_t \bar{q}^r_{t+1}]$, and by (A.14) this implies $E^a_{f,t}\bar{q}^r_{t+1} = E^r_t \bar{q}^r_{t+1}$. Hence it follows that

$$E^a_{f,t}\pi_{t+1} - \gamma\pi_{t-1} = (1-\xi)(\lambda_f E^r_t \bar{q}^r_{t+1} + (1-\lambda_f)E^a_{f,t}\bar{q}^a_{t+1}) \tag{A.15}$$

Rational agents are aware that they can form superior estimates to adaptive agents of $\bar{q}^a_{t+1}$. Thus advancing (A.13) by one period and taking expectations yields

$$E^r_t \bar{q}^r_{t+1} - \xi\beta E^r_t \bar{q}^r_{t+2} = E^r_t \bar{q}^a_{t+1} - \xi\beta E^r_t[E^a_{f,t+1}\bar{q}^a_{t+2}] \tag{A.16}$$

But adaptive agents assume

$$E^a_{f,t+1}\bar{q}^a_{t+2} = \mu_1 \bar{q}^a_{t+1} + (1-\mu_1)E^a_{f,t}\bar{q}^a_{t+1} \tag{A.17}$$

When rational agents take expectations of this at time t, they know that $E^r_t[E^a_{f,t}\bar{q}^a_{t+1}] = E^a_{f,t}\bar{q}^a_{t+1}$. Hence (A.16) can be written as

$$E^r_t \bar{q}^r_{t+1} - \xi\beta E^r_t \bar{q}^r_{t+2} = (1-\xi\beta\mu_1)E^r_t \bar{q}^a_{f,t+1} - \xi\beta(1-\mu_1)E^a_t \bar{q}^a_{f,t+1}$$

which is the defining equation for $E^r_t \bar{q}^a_{t+1}$, which is substituted into

$$E^r_t \pi_{t+1} - \gamma\pi_{t-1} = (1-\xi)(\lambda_f E^r_t \bar{q}^r_{t+1} + (1-\lambda_f)E^r_t \bar{q}^a_{f,t+1})$$

found by taking RE of (A.15)

B Proof of Theorem

Assume a model of the form

$$Z_{t+1} = J(Z_t, X_t) + \eta\sigma\varepsilon_{t+1} \qquad E_t X_{t+1} = K(Z_t, X_t)$$

where σ is small, and with measurements

$$W_t = L(Z_t, X_t)$$

We shall assume that there is a solution to this which involves the innovations process variable $\tilde{Z}_t \equiv Z_t - E_{t-1}Z_t$:

$$X_t = g(Z_t, \tilde{Z}_t, \sigma) \qquad Z_{t+1} = h(Z_t, \tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1} \qquad \tilde{Z}_{t+1} = f(\tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}$$

Also assume that

$$E_t Z_t - E_t Z_{t-1} = E_t \tilde{Z}_t = i(\tilde{Z}_t)$$

Noting that $K(Z_t, X_t) = E_t K(Z_t, X_t)$ it follows that

$$K(Z_t, g(Z_t, \tilde{Z}_t, \sigma)) = E_t K(Z_t, g(Z_t, \tilde{Z}_t, \sigma))$$

and the 1st order approximation to this is

$$K_1 z_t + K_2 g_1 z_t + K_2 g_2 \tilde{z}_t = K_1(z_t - \tilde{z}_t + i_1 \tilde{z}_t) + K_2 g_1 (z_t - \tilde{z}_t + i_1 \tilde{z}_t) + K_2 g_2 i_1 \tilde{z}_t \tag{B.1}$$

Ultimately we shall be solving for the partial derivative values at the steady state of f, g, h, i, and in particular by equating terms in z_t and $\tilde{z}_t$ in (B.1) we obtain

$$K_1 + K_2 g_1 = g_1 h_1 \qquad K_1 + K_2 g_1 + K_2 g_2 = (K_1 + K_2 g_1 + K_2 g_2) i_1$$

In the 1-dimensional case it is clear that $i_1 = 1$, and if the dimension of W_t equals that of Z_t then $i_1 = I$, the identity matrix. Now consider the second non-linear equation:

$$\begin{aligned} K(Z_t, X_t) &= K(Z_t, g(Z_t, \tilde{Z}_t, \sigma)) = E_t g(Z_{t+1}, \tilde{Z}_{t+1}, \sigma) \\ &= E_t g(h(Z_t, \tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}, f(\tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}, \sigma) \end{aligned}$$

The first order approximation to this is

$$K_1 z_t + K_2 g_1 z_t + K_2 g_2 \tilde{z}_t + K_2 g_\sigma \sigma = g_1 h_1 (z_t - \tilde{z}_t + i_1 \tilde{z}_t) + g_1 h_2 i_1 \tilde{z}_t + g_2 f_1 i_1 \tilde{z}_t + g_1 h_\sigma \sigma + g_2 f_\sigma \sigma + g_\sigma \sigma$$

from which it follows that

$$K_1 + K_2 g_1 = g_1 h_1 \qquad K_2 g_2 = -g_1 h_1 + g_1 h_1 i_1 + g_1 h_2 i_1 + g_2 f_1 i_1 \qquad K_2 g_\sigma = g_1 h_\sigma + g_2 f_\sigma + g_\sigma$$

Equating the two Z_{t+1} equations implies

$$h(Z_t, \tilde{Z}_t, \sigma) = J(Z_t, X_t) = J(Z_t, g(Z_t, \tilde{z}_t, \sigma))$$

so that to first order

$$J_1 z_t + J_2 g_1 z_t + J_2 g_2 \tilde{z}_t + J_2 g_\sigma \sigma = h_1 z_t + h_2 \tilde{z}_t + h_\sigma \sigma$$

and hence

$$J_1 + J_2 g_1 = h_1 \qquad J_2 g_2 = h_2 \qquad J_2 g_\sigma = h_\sigma$$

Note that $K_1 + K_2 g_1 = g_1 h_1$ and $J_1 + J_2 g_1 = h_1$ are the standard saddlepath solutions for g_1 and h_1.

Finally we need an equation describing how $E_t z_t$ is calculated. Thus assume that $E_{t-1} z_t$ is known. We can use the extended Kalman filter,[14] but evaluated always around the steady state.

[14]The control theory literature provides numerous numerical studies of convergence of the extended Kalman filter. There appears to be no guarantee of convergence, so that the problem might possibly be exacerbated by the approximation chosen, but the vast majority of the studies show that the extended Kalman filter is very reliable. There are very few studies that compare the extended Kalman filter to the

The measurement is given by $W_t = L(Z_t, X_t) = L(Z_t, g(Z_t, \tilde{Z}_t, \sigma))$. It follows that

$$E_t z_t = E_{t-1} z_t + PH^T (HPH^T)^{-1} H \tilde{z}_t$$

where $H = L_1 + L_2 g_1 + L_2 g_2$. It follows that

$$i_1 = PH^T (HPH^T)^{-1} H$$

Having solved previously for g_1, h_1 we still have to solve for g_2, h_2, f_1, i_1 as well as for P. Note that the latter arises from

$$P = f_1 P f_1^T + \sigma^2 \eta \eta^T$$

In addition we require that the first-order approximation to $\tilde{Z}_{t+1}$ equation derived from the the Z_{t+1} equation should have the same first-order approximation as the $\tilde{Z}_{t+1}$ equation itself. This implies that

$$(h_1 + h_2)(I - i_1) = f_1 \qquad f_\sigma = 0$$

This implies that we need to solve for the $n_x + n_z$ unknowns g_σ, h_σ for which the remaining equations reduce to :

$$(K_2 - I) g_\sigma = g_1 h_\sigma \qquad J_2 g_\sigma = h_\sigma$$

Since K_2 is $n_x \times n_x$ and J_2 is $n_z \times n_x$, it follows that there are $n_x + n_z$ equations in (B) from which it follows that $g_\sigma = 0, h_\sigma = 0$.

Thus the Schmitt-Grohe and Uribe Theorem 1 applies to the case of partial information as well.

exact filter calculated numerically. However a more recent approach with considerably more accuracy is due to Julier and Uhlmann (2004).

C Priors and Posterior Estimates

Parameter	Notation	Prior distribution			Identification check $^{\diamondsuit}$		
		Density	Mean	S.D/df	ML mode	S.D.	Bias *
Risk aversion	σ	Normal	2.00	0.50	2.17	0.39	0.17
Price indexation	γ	Beta	0.50	0.15	0.33	0.16	0.17
Calvo prices	ξ	Beta	0.50	0.10	0.53	0.06	0.02
Consumption habit formation	h_C	Beta	0.50	0.20	0.47	0.04	0.03
Preference parameter	ϱ	Beta	0.50	0.20	0.50	0.09	0.00
Labour share	α	Normal	0.80	0.10	0.78	0.24	0.02
Adaptive expectations							
Error adjustment - $E_{f,t}\bar{q}^a_{t+1}$	μ_1	Beta	0.50	0.20	0.62	0.19	0.12
Error adjustment - $E_{h,t}u^a_{c,t+1}$	μ_2	Beta	0.50	0.20	0.54	0.04	0.04
Error adjustment - $E^a_{h,t}[\pi_{t+1}]$	μ_3	Beta	0.50	0.20	0.53	0.02	0.03
Proportion of rational households	λ_h	Beta	0.50	0.20	0.46	0.04	0.04
Proportion of rational firms	λ_f	Beta	0.50	0.20	0.46	0.08	0.04
Interest rate rule							
Inflation	θ_π	Normal	2.00	0.50	2.09	0.13	0.09
Output	θ_y	Normal	0.125	0.05	0.124	0.03	0.001
Interest rate smoothing	ρ_r	Beta	0.80	0.10	0.80	0.01	0.00
AR(1) coefficient							
Technology	ρ_a	Beta	0.85	0.10	0.82	0.02	0.03
Government spending	ρ_g	Beta	0.85	0.10	0.83	0.02	0.02
Price mark-up	ρ_{ms}	Beta	0.50	0.20	0.37	0.16	0.13
Standard deviation of AR(1) innovations							
Technology	$sd(\epsilon_a)$	Inv. gamma	0.60	2.00	1.14 ✠	0.18	0.14
Government spending	$sd(\epsilon_g)$	Inv. gamma	1.67	2.00	0.97	0.04	0.03
Price mark-up	$sd(\epsilon_{ms})$	Inv. gamma	0.10	2.00	1.06	0.34	0.06
Standard deviation of I.I.D. shocks/mearsument errors							
Mark-up process	$sd(\epsilon_m)$	Inv. gamma	0.10	2.00	0.63	0.37	0.37
Monetary policy	$sd(\epsilon_e)$	Inv. gamma	0.10	2.00	1.02	0.02	0.02
Observation error (inflation)	$sd(\epsilon_\pi)$	Inv. gamma	0.10	2.00	-	-	-
Observation error (output)	$sd(\epsilon_y)$	Inv. gamma	0.10	2.00	-	-	-

Table 6: Prior Distributions and ML Estimation Based on Artificial Data

$^{\diamondsuit}$ We generated artificial data observations of length T=5000 by imposing the prior means to all of the parameters (except for the S.D. of the shocks). The results presented here are based on maximum likelihood estimates of Model GH (behavioural) for the T=5000 observations.
* Note that *Bias* is measured as the absolute value of the difference between the prior mean and the mean of ML estimates for each parameter.
✠ The artificial data are simulated assuming that the standard deviations of all the shocks are 1 instead of using their prior means.

	AI				II				IIME			
Parameter	Model GH	Model H	Model G	Model Z	Model GH	Model H	Model G	Model Z	Model GH	Model H	Model G	Model Z
σ	2.28 [1.51:3.01]	2.22 [1.44:3.03]	2.57 [1.92:3.25]	2.62 [1.93:3.24]	2.36 [1.60:3.07]	2.30 [1.57:3.06]	2.66 [1.99:3.29]	2.78 [2.07:3.42]	2.38 [1.64:3.09]	2.30 [1.53:3.04]	2.62 [2.00:3.26]	2.74 [2.01:3.39]
γ	0.38 [0.16:0.58]	-	0.34 [0.14:0.53]	-	0.43 [0.21:0.65]	-	0.39 [0.20:0.57]	-	0.49 [0.24:0.73]	-	0.47 [0.23:0.69]	-
ξ	0.82 [0.75:0.90]	0.85 [0.79:0.91]	0.60 [0.46:0.76]	0.67 [0.55:0.79]	0.83 [0.76:0.90]	0.85 [0.79:0.91]	0.64 [0.56:0.72]	0.70 [0.62:0.82]	0.83 [0.76:0.91]	0.86 [0.80:0.91]	0.67 [0.60:0.75]	0.71 [0.63:0.80]
h_C	0.84 [0.75:0.93]	0.86 [0.78:0.94]	-	-	0.80 [0.69:0.91]	0.84 [0.76:0.92]	-	-	0.81 [0.71:0.91]	0.84 [0.77:0.93]	-	-
ϱ	0.33 [0.07:0.58]	0.33 [0.06:0.59]	0.37 [0.10:0.63]	0.32 [0.08:0.53]	0.35 [0.08:0.61]	0.33 [0.08:0.58]	0.37 [0.12:0.65]	0.25 [0.04:0.47]	0.34 [0.07:0.59]	0.33 [0.07:0.58]	0.31 [0.07:0.53]	0.31 [0.07:0.54]
α	0.87 [0.72:1.00]	0.87 [0.72:1.02]	0.75 [0.62:0.88]	0.76 [0.61:0.89]	0.87 [0.73:1.01]	0.88 [0.73:1.03]	0.74 [0.61:0.87]	0.74 [0.58:0.89]]	0.86 [0.72:1.01]	0.87 [0.73:1.02]	0.73 [0.60:0.86]	0.72 [0.58:0.86]
Interest rate rule												
θ_π	1.58 [1.18:1.96]	1.55 [1.08:1.95]	2.97 [2.43:3.52]	2.84 [2.40:3.34]	1.57 [1.19:1.95]	1.48 [1.08:1.86]	2.87 [2.39:3.38]	2.73 [1.34:3.55]	1.57 [1.17:1.96]	1.48 [1.07:1.84]	2.90 [2.33:3.43]	2.83 [2.28:3.56]
θ_y	0.09 [0.00:0.17]	0.08 [-0.01:0.17]	0.23 [0.16:0.29]	0.24 [0.18:0.30]	0.08 [0.00:0.17]	0.08 [-0.01:0.17]	0.22 [0.16:0.28]	0.19 [0.08:0.29]	0.08 [0.00:0.17]	0.08 [-0.01:0.16]	0.23 [0.17:0.30]	0.21 [0.15:0.30]
ρ_r	0.80 [0.75:0.86]	0.81 [0.75:0.97]	0.58 [0.44:0.71]	0.53 [0.44:0.64]	0.80 [0.75:0.86]	0.81 [0.76:0.87]	0.52 [0.40:0.65]	0.54 [0.35:0.75]	0.81 [0.75:0.87]	0.81 [0.76:0.86]	0.55 [0.43:0.69]	0.46 [0.29:0.62]
AR(1) coefficient												
ρ_a	0.98 [0.97:0.99]	0.98 [0.96:0.99]	0.96 [0.95:0.98]	0.96 [0.95:0.98]	0.98 [0.97:0.99]	0.98 [0.97:0.99]	0.97 [0.95:0.99]	0.97 [0.94:0.99]	0.98 [0.97:0.99]	0.98 [0.97:0.99]	0.96 [0.95:0.98]	0.97 [0.95:0.98]
ρ_g	0.92 [0.87:0.97]	0.93 [0.87:0.98]	0.88 [0.84:0.93]	0.89 [0.85:0.94]	0.92 [0.86:0.97]	0.93 [0.88:0.98]	0.87 [0.82:0.92]	0.88 [0.82:0.95]]	0.91 [0.86:0.97]	0.93 [0.88:0.98]	0.86 [0.82:0.91]	0.88 [0.83:0.93]
ρ_{ms}	0.27 [0.04:0.47]	0.36 [0.05:0.65]	0.98 [0.96:0.99]	0.98 [0.96:0.99]	0.40 [0.10:0.69]	0.50 [0.19:0.80]	0.98 [0.97:0.99]	0.89 [0.53:0.99]	0.40 [0.11:0.69]	0.54 [0.19:0.83]	0.98 [0.97:0.99]	0.95 [0.96:0.99]
Standard deviation of AR(1) innovations												
$sd(\epsilon_a)$	1.39 [0.92:1.83]	1.62 [1.12:2.17]]	0.74 [0.58:0.89]	0.72 [0.57:0.86]	1.27 [0.85:1.70]	1.49 [1.02:2.00]	0.71 [0.57:0.84]	0.70 [0.53:0.88]	1.26 [0.87:1.65]	1.43 [0.94:1.98]	0.72 [0.59:0.86]	0.72 [0.57:0.88]
$sd(\epsilon_g)$	2.03 [1.80:2.56]	2.03 [1.80:2.25]	2.60 [2.07:3.09]	2.69 [2.19:3.17]	2.05 [1.83:2.28]	2.03 [1.81:2.25]	2.62 [2.12:3.07]	2.62 [1.98:3.17]]	2.05 [1.81:2.27]	2.02 [1.80:2.24]	2.71 [2.21:3.14]	2.65 [2.11:3.21]
$sd(\epsilon_{ms})$	0.07 [0.03:0.12]	0.07 [0.03:0.11]	0.23 [0.04:0.41]	0.17 [0.04:0.34]	0.07 [0.03:0.12]	0.06 [0.03:0.10]	0.11 [0.05:0.17]	0.14 [0.04:0.25]	0.06 [0.03:0.10]	0.06 [0.03:0.09]	0.09 [0.04:0.13]	0.11 [0.04:0.19]
Standard deviation of I.I.D. shocks/mearsument errors												
$sd(\epsilon_m)$	0.11 [0.04:0.17]	0.08 [0.03:0.13]	0.14 [0.05:0.20]	0.16 [0.09:0.24]	0.11 [0.04:0.16]	0.06 [0.03:0.10]	0.23 [0.18:0.27]	0.18 [0.03:0.26]	0.09 [0.03:0.14]	0.06 [0.03:0.09]	0.13 [0.03:0.23]	0.13 [0.03:0.24]
$sd(\epsilon_e)$	0.27 [0.24:0.30]	0.27 [0.24:0.29]	0.26 [0.22:0.30]	0.22 [0.17:0.26]	0.27 [0.24:0.30]	0.27 [0.24:0.30]	0.29 [0.25:0.33]	0.27 [0.22:0.31]]	0.27 [0.24:0.30]	0.27 [0.24:0.30]	0.28 [0.24:0.32]	0.25 [0.21:0.30]
$sd(\epsilon_\pi)$	-	-	-	-	-	-	-	-	0.09 [0.03:0.14]	0.06 [0.03:0.09]	0.15 [0.03:0.25]	0.14 [0.03:0.25]
$sd(\epsilon_y)$	-	-	-	-	-	-	-	-	0.07 [0.02:0.12]	0.07 [0.02:0.12]	0.06 [0.03:0.10]	0.07 [0.02:0.12]
Price contract length												
$\frac{1}{1-\xi}$	5.56	6.67	2.50	3.03	5.88	6.67	2.78	3.33	5.88	7.14	3.03	3.45
LL and posterior model odd												
LL	-239.59	-238.20	-245.30	-244.37	-230.95	-230.89	-239.15	-242.04	-230.52	-231.37	-238.40	-239.21
Prob.	0.00	0.00	0.00	0.00	0.23	0.25	0.00	0.00	0.36	0.15	0.00	0.00

Table 7: Bayesian Posterior Distributions[◇]

◇ Notes: we report posterior means and 90% probability intervals (in parentheses) based on the output of the Metropolis-Hastings Algorithm. Sample range: 1970:I to 2004:IV.

	AI				II				IIME			
Parameter	Model GH	Model H	Model G	Model Z	Model GH	Model H	Model G	Model Z	Model GH	Model H	Model G	Model Z
σ	2.48 [1.74:3.32]	2.46 [1.70:3.17]	2.60 [1.90:3.31]	2.53 [1.80:3.25]	2.52 [1.76:3.26]	2.50 [1.77:3.22]	2.55 [1.84:3.24]	2.50 [1.81:3.18]	2.52 [1.81:3.25]	2.55 [1.83:3.28]	2.55 [1.82:3.30]	2.50 [1.83:3.15]
γ	0.29 [0.11:0.46]	-	0.27 [0.11:0.43]	-	0.29 [0.11:0.47]	-	0.24 [0.10:0.39]	-	0.30 [0.12:0.47]	-	0.24 [0.10:0.38]	-
ξ	0.91 [0.87:0.95]	0.89 [0.84:0.94]	0.87 [0.82:0.92]	0.86 [0.81:0.92]	0.91 [0.87:0.95]	0.88 [0.83:0.95]	0.88 [0.83:0.92]	0.87 [0.81:0.92]	0.90 [0.85:0.95]	0.88 [0.81:0.94]	0.87 [0.83:0.92]	0.87 [0.82:0.92]
h_C	0.87 [0.80:0.94]	0.88 [0.80:0.95]	-	-	0.88 [0.81:0.94]	0.87 [0.80:0.95]	-	-	0.87 [0.80:0.94]	0.87 [0.80:0.95]	-	-
ϱ	0.46 [0.12:0.76]	0.45 [0.12:0.76]	0.20 [0.04:0.36]	0.22 [0.03:0.39]	0.46 [0.14:0.77]	0.44 [0.12:0.73]	0.20 [0.03:0.36]	0.21 [0.05:0.38]	0.46 [0.15:0.76]	0.43 [0.11:0.72]	0.21 [0.04:0.37]	0.21 [0.04:0.38]
α	0.79 [0.63:0.95]	0.80 [0.64:0.97]	0.90 [0.74: 1.05]	0.89 [0.74:1.05]	0.79 [0.63:0.95]	0.79 [0.63:0.94]	0.91 [0.75:1.06]	0.90 [0.73:1.05]]	0.78 [0.63:0.94]	0.79 [0.63:0.93]	0.90 [0.75:1.06]	0.89 [0.74:1.06]
Adaptive expectations												
μ_1	0.30 [0.13:0.47]	0.29 [0.11:0.49]	0.25 [0.10:0.39]	0.29 [0.10:0.49]	0.29 [0.12:0.45]	0.32 [0.11:0.53]	0.28 [0.13:0.43]	0.30 [0.11:0.50]	0.28 [0.11:0.45]	0.30 [0.10:0.49]	0.29 [0.12:0.43]	0.32 [0.13:0.51]
μ_2	0.39 [0.18:0.59]	0.42 [0.20:0.62]	0.19 [0.01:0.36]	0.30 [0.07:0.51]	0.42 [0.19:0.68]	0.43 [0.21:0.65]	0.18 [0.01:0.36]	0.31 [0.07:0.52]	0.43 [0.18:0.70]	0.47 [0.21:0.76]	0.22 [0.02:0.42]	0.26 [0.06:0.43]
μ_3	0.47 [0.04:0.78]	0.46 [0.04:0.78]	0.23 [0.02:0.57]	0.12 [0.01:0.31]	0.49 [0.13:0.85]	0.45 [0.04:0.77]	0.23 [0.01:0.57]	0.16 [0.01:0.44]	0.48 [0.04:0.78]	0.46 [0.04:0.77]	0.18 [0.01:0.51]	0.18 [0.08:0.48]
λ_h	0.73 [0.56:0.89]	0.69 [0.51:0.87]	0.20 [0.03:0.36]	0.16 [0.02:0.29]	0.75 [0.60:0.93]	0.71 [0.54:0.96]	0.18 [0.02:0.34]	0.15 [0.02:0.30]	0.76 [0.60:0.93]	0.72 [0.54:0.93]	0.17 [0.02:0.33]	0.16 [0.01:0.29]
λ_f	0.17 [0.01:0.33]	0.18 [0.02:0.32]	0.14 [0.02:0.27]	0.16 [0.02:0.29]	0.18 [0.02:0.34]	0.23 [0.02:0.43]	0.16 [0.02:0.30]	0.18 [0.02:0.32]	0.20 [0.02:0.38]	0.24 [0.03:0.46]	0.15 [0.02:0.29]	0.18 [0.02:0.33]
Interest rate rule												
θ_π	1.92 [1.45:2.40]	1.89 [1.42:2.32]	1.71 [1.28:2.12]	1.69 [1.20:2.15]	1.95[1.46:2.42]	1.91 [1.43:2.40]	1.69 [1.26:2.09]	1.69 [1.20:2.09]	1.93 [1.46:2.37]	1.90 [1.43:2.37]	1.71 [1.28:2.13]	1.63 [1.16:2.06]
θ_y	0.11 [0.04:0.18]	0.11 [0.04:0.18]	0.14 [0.07:0.20]	0.13 [0.07:0.20]	0.11 [0.03:0.17]	0.11 [0.03:0.17]	0.14 [0.07:0.20]	0.14 [0.08:0.20]	0.10 [0.04:0.17]	0.10 [0.03:0.17]	0.14 [0.07:0.20]	0.14 [0.07:0.20]
ρ_r	0.89 [0.85:0.93]	0.88 [0.84:0.92]	0.89 [0.85:0.92]	0.88 [0.85:0.92]	0.89 [0.85:0.93]	0.88 [0.84:0.92]	0.89 [0.85:0.92]	0.88 [0.85:0.92]	0.88 [0.84:0.92]	0.88 [0.83:0.92]	0.89 [0.85:0.93]	0.88 [0.84:0.92]
AR(1) coefficient												
ρ_a	0.87 [0.73:0.99]	0.86 [0.70:0.99]	0.84 [0.70:0.99]	0.84 [0.70:0.98]	0.90 [0.76:0.99]	0.90 [0.7:0.99]	0.84 [0.70:0.98]	0.85 [0.72:0.97]	0.91 [0.79:0.99]	0.91 [0.78:0.99]	0.85 [0.91:0.99]	0.87 [0.76:0.98]
ρ_g	0.93 [0.89:0.98]	0.93 [0.88:0.98]	0.96 [0.94:0.99]	0.96 [0.94:0.99]	0.93 [0.88:0.98]	0.93 [0.88:0.98]	0.96 [0.94:0.99]	0.96 [0.94:0.99]]	0.93 [0.88:0.98]	0.93 [0.88:0.98]	0.96 [0.94:0.99]	0.96 [0.94:0.99]
ρ_{ms}	0.56 [0.19:0.91]	0.66 [0.40:0.93]	0.45 [0.12:0.77]	0.53 [0.26:0.81]	0.56 [0.20:0.90]	0.58 [0.24:0.91]	0.44 [0.13:0.76]	0.53 [0.26:0.81]	0.52 [0.15:0.89]	0.59 [0.24:0.93]	0.44 [0.14:0.76]	0.53 [0.26:0.81]
Standard deviation of AR(1) innovations												
$sd(\epsilon_a)$	0.94 [0.14:2.34]	0.71 [1.14:1.63]]	0.49 [0.15:0.83]	0.63 [0.13:1.30]	0.90 [0.16:1.89]	1.26 [1.14:2.53]	0.44 [0.15:0.75]	0.56 [0.15:1.15]	1.01 [0.16:1.88]	1.05 [0.18:1.84]	0.45 [0.15:0.77]	0.51 [0.16:0.89]
$sd(\epsilon_g)$	2.07 [1.84:2.31]	2.06 [1.83:2.28]	2.17 [1.91:3.42]	2.14 [1.86:2.42]	2.06 [1.84:2.29]	2.05 [1.82:2.28]	2.18 [1.91:2.45]	2.16 [1.89:2.43]]	2.07 [1.84:2.28]	2.05 [1.83:2.28]	2.17 [1.89:2.45]	2.15 [1.88:2.41]
$sd(\epsilon_{ms})$	0.08 [0.03:0.14]	0.15 [0.06:0.22]	0.08 [0.03:0.16]	0.14 [0.05:0.21]	0.08 [0.03:0.15]	0.12 [0.03:0.20]	0.08 [0.03:0.14]	0.13 [0.05:0.21]	0.08 [0.03:0.14]	0.13 [0.04:0.20]	0.08 [0.03:0.15]	0.13 [0.05:0.19]
Standard deviation of I.I.D. shocks/mearsument errors												
$sd(\epsilon_m)$	0.12 [0.07:0.17]	0.09 [0.03:0.14]	0.12 [0.06:0.17]	0.09 [0.03:0.15]	0.12 [0.06:0.17]	0.08 [0.03:0.13]	0.12 [0.06:0.17]	0.09 [0.03:0.14]	0.12 [0.07:0.17]	0.08 [0.0:0.14]	0.12 [0.05:0.17]	0.09 [0.03:0.14]
$sd(\epsilon_e)$	0.26 [0.23:0.29]	0.26 [0.23:0.29]	0.26 [0.23:0.28]	0.26 [0.23:0.28]	0.26 [0.23:0.28]	0.26 [0.23:0.29]	0.25 [0.23:0.28]	0.25 [0.23:0.28]]	0.26 [0.23:0.29]	0.26 [0.23:0.29]	0.25 [0.23:0.28]	0.25 [0.23:0.28]
$sd(\epsilon_y)$	-	-	-	-	-	-	-	-	0.001 [0.000:0.002]	0.001 [0.000:0.001]	0.001 [0.000:0.001]	0.001 [0.000:0.001]
Price contract length												
$\frac{1}{1-\xi}$	11.11	9.09	7.69	7.14	11.11	8.33	8.33	7.69	10.00	8.33	7.69	7.69
LL and posterior model odd												
LL	-230.87	-227.14	-242.87	-239.28	-230.79	-226.55	-242.98	-238.55	-230.72	-227.27	-243.15	-239.12
Prob.	0.006	0.266	0.000	0.000	0.007	0.480	0.000	0.000	0.007	0.2335	0.000	0.000

TABLE 8: BAYESIAN POSTERIOR DISTRIBUTIONS (THE BEHAVIOURAL MODEL)$^\diamondsuit$

$^\diamondsuit$ Notes: we report posterior means and 90% probability intervals (in parentheses) based on the output of the Metropolis-Hastings Algorithm. Sample range: 1970:I to 2004:IV.

D Figures

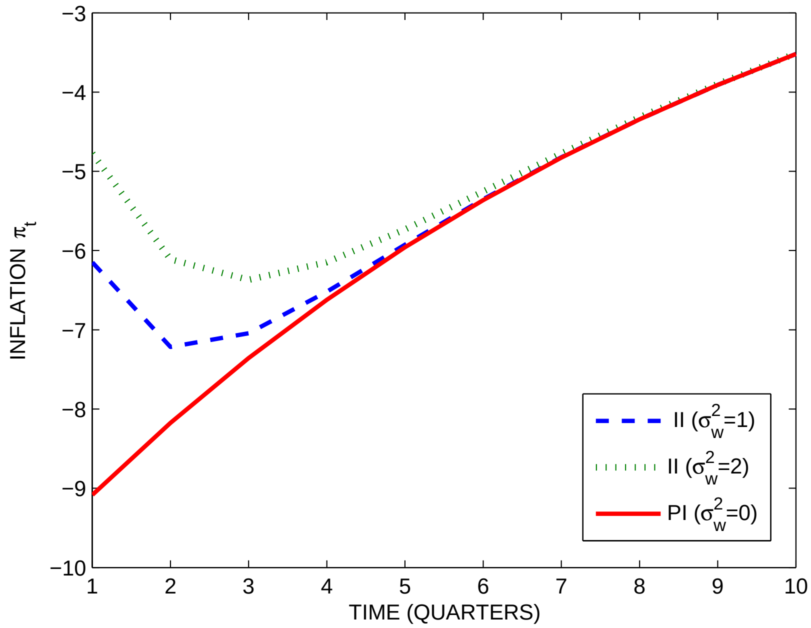

Figure 1: Inflation Dynamics under Perfect (PI) and Imperfect Information (II)

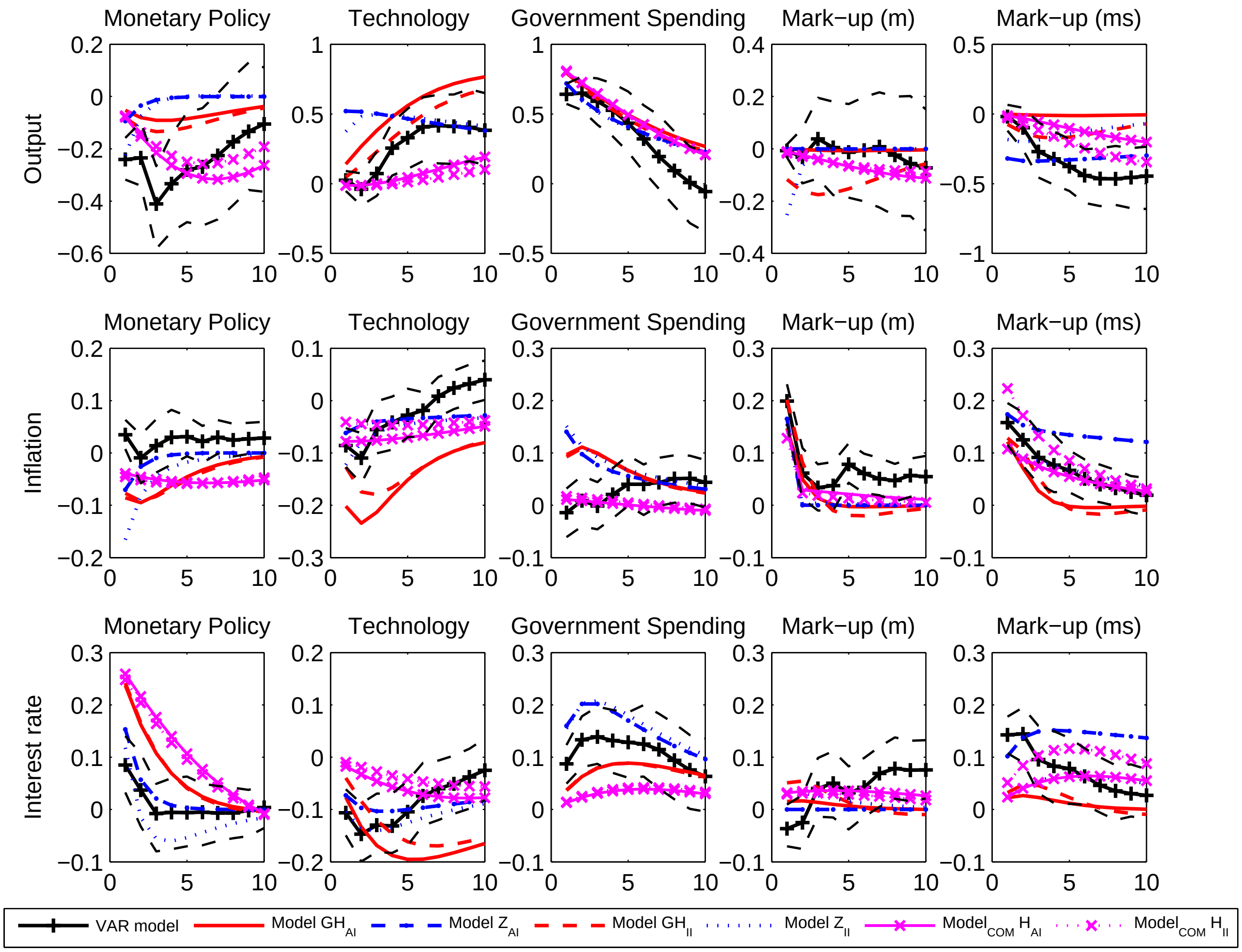

FIGURE 2: ESTIMATED IMPULSE RESPONSE FUNCTIONS - AI VS. II[◇]

[◇] Each panel plots the mean response corresponding a positive one standard deviation shock. Each response is for a 10 period horizon. All DSGE impulse responses are computed simulating the vector of DSGE model parameters at the posterior mean values reported in Table 5. The impulse responses for VAR(4) are obtained using the DSGE-VAR identification procedure described in the section 5.4. Mark-up(ms) and Mark-up(m) represent the price mark-up shocks (persistent and transient components respectively). The area in-between the black dashed lines covers the space between the first and ninth posterior deciles of the IRFs estimated by the VAR.

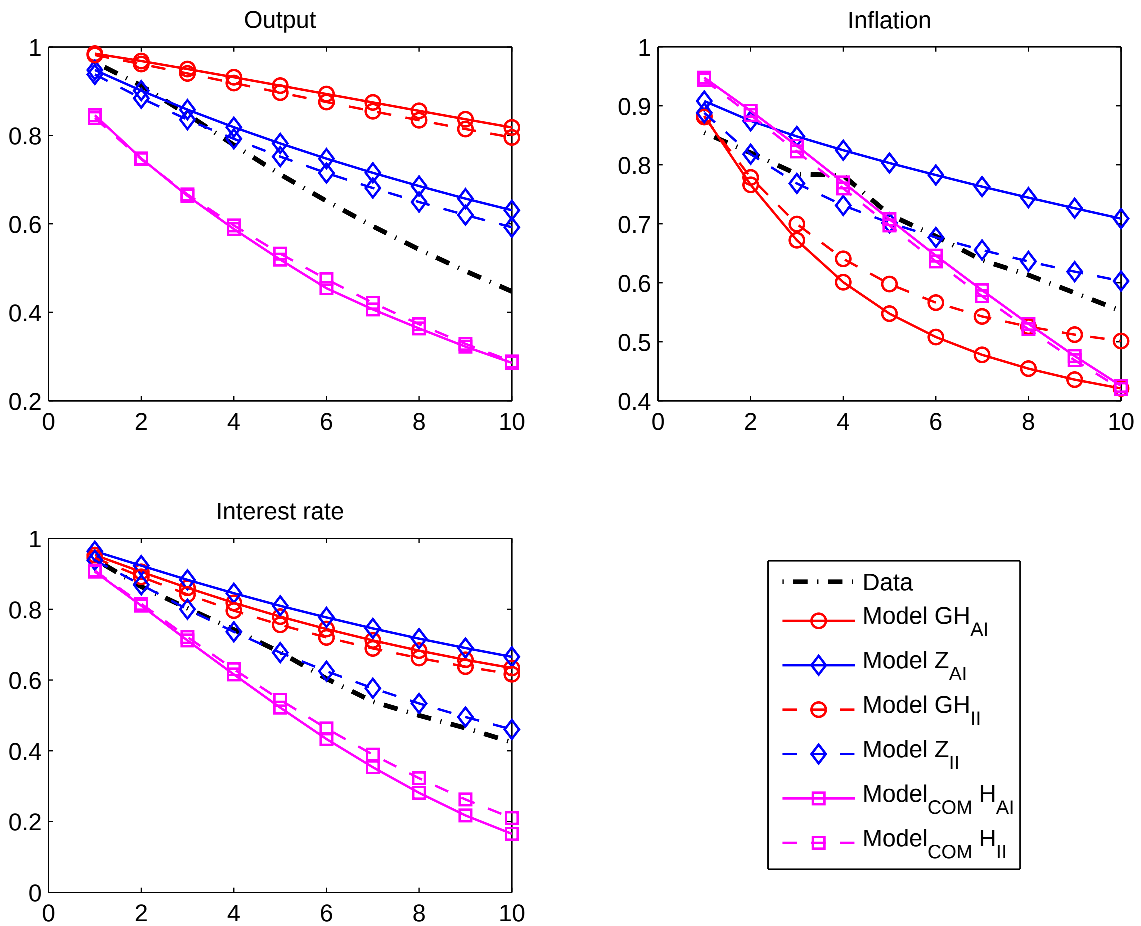

Figure 3: Autocorrelations of Observables in the Actual Data and in the Estimated Models