Report on the First Deliverable of Work Package 6.1 (Partial Information) for the FP7-funded Grant MONFISPOL

Joseph Pearlman
London Metropolitan University

November 12, 2009

Part I

This report documents the work that has been undertaken so far for estimating and forecasting linear rational expectations models of the economy when agents and econometricians share the same imperfect (or partial) information.

1 Summary of State of the Art

There are two further parts of this report and a suite of programs. Part II entitled 'Partial Information Implementation in Dynare' is the backbone of the software, while Part III is a paper written using the newly developed software.

Part II first describes an algorithm for converting a linearized model of the economy expressed in Sims form

$$A_0 Y_{t+1,t} + A_1 Y_t = A_2 Y_{t-1} + B\varepsilon_t \tag{1}$$

with agents' and econometricians' measurements given by

$$m_t = LY_t + v_t \tag{2}$$

into a model expressed in Blanchard-Kahn form:

$$\begin{bmatrix} z_{t+1} \\ x_{t+1,t} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} \varepsilon_{t+1} \tag{3}$$

with agents' measurements given by

$$m_t = \begin{bmatrix} K_1 & K_2 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + v_t \tag{4}$$

Surprisingly (for the author) the algorithm for making this conversion is original, or at least has never before been documented.

The document then provides solutions to this, with representations of the impulse response functions, covariances, autocovariances and correlation matrix.

It also provides the formulae for calculating the likelihood function for a given parameter set using a variant of the Kalman filter suitable for the RE forward-looking case.

Part III, entitled 'Endogenous Persistence in an Estimated New Keynesian Model under Imperfect Information' is an application of the software to evaluate the importance of partial information for parameter estimation and impulse response matching for a US dataset of macro variables.

The suite of programs are based on programs that are currently available in Dynare, so that they run within Dynare in much the same way. The only change is that the *modfile* requires the additional commands:

options_.usePartInfo=1;

options_.use_k_order=0;

The programs have been subjected to rigorous testing and are almost ready for distribution. The tests are as follows:

1. Calculation of the likelihood function for a simple model with one backward-looking and one forward-looking variable. These match the calculation using a spreadsheet to 7 significant figures.

2. Testing of impulse response functions under partial information when in fact there is full information. These have been compared with impulse response functions generated using the standard Dynare package, and they are correct to 7 significant figures.

3. Comparison of the likelihood function when there is full information to the likelihood function generated by the standard Dynare package. These match to 7 significant figures.

The model for item 1 is the analytical example of Section 12 below. The model for items 2 and 3 above is given by

$$\pi_t = \frac{\beta}{1+\beta\gamma}E_t\pi_{t+1} + \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{(1-\beta\xi)(1-\xi)}{(1+\beta\gamma)\xi}mc_t + mstot_t \tag{5}$$

$$mc_t = mun_t - muc_t - a_t + (1-\alpha)n_t + \epsilon_{mc,t} \tag{6}$$

$$mun_t = \frac{1}{1-\frac{h_c}{\gamma_g}}(c_t - \frac{h_c}{\gamma_g}c_{t-1}) + \frac{w_d}{1-w_d}n_t + muc_t; \tag{7}$$

$$muc_t = \frac{(1-\rho)(1-\sigma)-1}{1-h_c}(c_t - \frac{h_c}{\gamma_g}c_{t-1}) - \frac{w_d\rho(1-\sigma)}{1-w_d}n_t \tag{8}$$

$$muc_{t+1} = muc_t - (r_t - \pi_{t+1}) + \epsilon_t \tag{9}$$

$$y_t = c_y c_t + (1 - c_y) g_t \tag{10}$$

$$n_t = (y_t - a_t)/\alpha \tag{11}$$

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)\theta_p \pi_{t+1} + \epsilon_{e,t} \tag{12}$$

$$g_t = \rho_g g_{t-1} + \varepsilon_{g,t} \tag{13}$$

$$a_t = \rho_a a_{t-1} + \varepsilon_{a,t} \tag{14}$$

$$ms_t = \rho_{ms} ms_{t-1} + \varepsilon_{ms,t} \tag{15}$$

$$mstot_t = ms_t + \varepsilon_{m,t} \tag{16}$$

Under partial information, for testing purposes, it was assumed that the measured variables are $\{\pi, y, r, g, a, ms, mstot\}$. Inspection of the equations shows that this is equivalent to having full information. The software is then used to calculate impulse response functions, and in addition the model is simulated, and the likelihood calculated. In both cases, there was a match between these and the model using standard Dynare software.

2 Further Work

There are still a number of items that need to be addressed before the software is ready for general distribution.

1. Incorporation of current best estimates $Y_{t,t}$ into the model. An example of this might be when firms set wages based on best current estimates of the technology shock, or when there is an interest rate rule based on best current estimates of the output gap. We are awaiting directions on this from the authors of Dynare.

2. Incorporation of more general measurements; for example observations of the term structure of interest rates requires measurements of the form $(r_t + r_{t+1,t} + .. + r_{t+T,t})/(T + 1)$. This requires work on the parser by the authors of Dynare.

3. Currently the partial information software only permits one lag and one lead. We shall be extending this to any number of lags and leads as in standard Dynare.

4. There is as yet no variance decomposition for forecasts, but this will be added.

5. At the moment the new software requires its own Matlab subdirectory. We intend to ensure that the software will sit in the same Matlab subdirectory as standard Dynare.

Part II

Partial Information Implementation in Dynare

3 Introduction

The aim of this document is to describe an algorithm for turning the state space setup of Dynare into one that is suitable for obtaining the partial information setup that conforms to that of Pearlman *et al.* (1986). The state space setup for Dynare is based on writing an RE system as:

$$A_0 Y_{t+1,t} + A_1 Y_t = A_2 Y_{t-1} + B u_t \tag{17}$$

where A_0 is not of full rank and u_t is a vector containing instruments w_t and shocks ε_t. Currently estimation within Dynare assumes that agents have full information about the system, so that a calculation is done which solves (17) under full information. The estimation step then assumes that econometricians have only a limited information set, and processes this via the Kalman filter to obtain the likelihood function for a given set of parameters. In reality, agents too have a partial information set (which may or may not coincide with that of the econometricians) given by

$$m_t = L Y_t + v_t \tag{18}$$

where typically there is no observation error ($v_t = 0$) and L picks out most of the economic variables, typically excluding capital stock, Tobin's q and shocks.

The Pearlman *et al.* (1986) setup is given by

$$\begin{bmatrix} z_{t+1} \\ x_{t+1,t} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} C \\ 0 \end{bmatrix} \varepsilon_{t+1} + \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} w_t \tag{19}$$

with agents' measurements given by

$$m_t = \begin{bmatrix} K_1 & K_2 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + v_t \tag{20}$$

and these can be solved together to yield a reduced-form system. This can then be processed via the Kalman filter to obtain the likelihood function as above.

The next section describes an algorithm for converting the state space (17), (18) under partial information to the form (19), (20).

4 Conversion to Pearlman *et al.* (1986) Setup

In order to reduce the amount of notation we impose a particular way of incorporating shocks into the system. Suppose a particular shock $\bar{m}_t$ affects an equation of the system, where $\bar{m}_{t+1} = \rho\bar{m}_t + \bar{u}_{t+1}$. Redefine $m_t = \bar{m}_{t+1}$, $u_t = \bar{u}_{t+1}$, so that now the relevant equation of the system is affected by m_{t-1}, and the law of motion of the shock is described within the matrices A_1, A_2, B. This makes no difference to the Kalman filter below or to system estimation, but means that for simulation purposes, a shock to u_t at time 0 will have an effect that is diminished by ρ compared with a shock to $\bar{u}_t$ at time 0.

To repeat, all shocks $\bar{m}_t$ to the system at time t are dated as though they were m_{t-1}.

The procedure for conversion to a form suitable for filtering is then as follows:

1. Obtain the singular value decomposition for matrix A_0: $A_0 = UDV^T$, where U, V are unitary matrices. Assuming that only the first m values of the diagonal matrix D are non-zero, we can rewrite this as $A_0 = U_1 D_1 V_1^T$, where U_1 are the first m columns of U, D_1 is the first $m \times m$ block of D and V_1^T are the first m rows of V^T.

2. Multiply (17) by $D_1^{-1}U_1^T$, which yields

$$V_1^T Y_{t+1,t} + D_1^{-1}U_1^T A_1 Y_t = D_1^{-1}U_1^T A_2 Y_{t-1} + D_1^{-1}U_1^T B u_t \tag{21}$$

   Now define $x_t = V_1^T Y_t$, $s_t = V_2^T Y_t$, and use the fact that $I = VV^T = V_1 V_1^T + V_2 V_2^T$ to rewrite this as:

$$x_{t+1,t} + D_1^{-1}U_1^T A_1(V_1 x_t + V_2 s_t) = D_1^{-1}U_1^T A_2(V_1 x_{t-1} + V_2 s_{t-1}) + D_1^{-1}U_1^T B u_t \tag{22}$$

3. Multiply (17) by U_2^T which yields

$$U_2^T A_1 Y_t = U_2^T A_2 Y_{t-1} + U_2^T B u_t \tag{23}$$

   which can be rewritten as

$$U_2^T A_1(V_1 x_t + V_2 s_t) = U_2^T A_2(V_1 x_{t-1} + V_2 s_{t-1}) + U_2^T B u_t \tag{24}$$

4. Typically $U_2^T A_1 V_2$ is invertible, which means that we can rewrite (22) and (24) as

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ F & 0 & I \end{bmatrix} \begin{bmatrix} s_t \\ x_t \\ x_{t+1,t} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} & -G_{13} \\ 0 & 0 & I \\ G_{31} & G_{32} & -G_{33} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ x_{t-1} \\ x_t \end{bmatrix} + \begin{bmatrix} H_1 \\ 0 \\ H_3 \end{bmatrix} u_t \tag{25}$$

   where

$$G_{11} = (U_2^T A_1 V_1)^{-1} U_2^T A_2 V_2 \quad G_{12} = (U_2^T A_1 V_1)^{-1} U_2^T A_2 V_1 \quad G_{13} = (U_2^T A_1 V_1)^{-1} U_2^T A_1 V_2 \tag{26}$$

$$G_{21} = D_1^{-1}U_1^T A_2 V_2 \quad G_{22} = D_1^{-1}U_1^T A_2 V_1 \quad G_{23} = D_1^{-1}U_1^T A_1 V_2 \tag{27}$$

$$H_1 = (U_2^T A_1 V_1)^{-1} U_2^T B \quad H_3 = D_1^{-1}U_1^T B \quad F = D_1^{-1}U_1^T A_1 V_2 \tag{28}$$

which can be further rewritten as

$$\begin{bmatrix} s_t \\ x_t \\ x_{t+1,t} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} & -G_{13} \\ 0 & 0 & I \\ G_{31} - FG_{11} & G_{32} - FG_{12} & -G_{33} + FG_{13} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ x_{t-1} \\ x_t \end{bmatrix} + \begin{bmatrix} H_1 \\ 0 \\ H_3 - FH_1 \end{bmatrix} u_t \tag{29}$$

5. The measurements $m_t = MY_t + v_t$ can be written in terms of the states as $m_t = M(V_1 x_t + V_2 s_t) + v_t$. To write the system in a form which corresponds to that of Pearlman *et al.* (1986) we need to write the measurements in terms of the forward-looking variables x_t and in terms of the backward-looking variables s_{t-1}, x_{t-1}. We do this by substituting for s_t from (29); but this introduces a term in u_t into the expression, and Pearlman *et al.* (1986) assume that shock terms in the dynamics and in the measurements are uncorrelated with one another. To remedy this, we incorporate ε_t into the predetermined variables, but we can retain w_t as it stands.

   Defining

$$\begin{bmatrix} H_1 \\ 0 \\ H_3 - FH_1 \end{bmatrix} u_t = \begin{bmatrix} P_1 \\ 0 \\ P_3 \end{bmatrix} \varepsilon_t + \begin{bmatrix} N_1 \\ 0 \\ N_3 \end{bmatrix} w_t \tag{30}$$

   we may rewrite the dynamics and measurement equations in the form:

$$\begin{bmatrix} \varepsilon_{t+1} \\ s_t \\ x_t \\ x_{t+1,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ P_1 & G_{11} & G_{12} & -G_{13} \\ 0 & 0 & 0 & I \\ P_3 & G_{31} - FG_{11} & G_{32} - FG_{12} & -G_{33} + FG_{13} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ s_{t-1} \\ x_{t-1} \\ x_t \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & N_1 \\ 0 & 0 \\ 0 & N_3 \end{bmatrix} \begin{bmatrix} \varepsilon_{t+1} \\ w_t \end{bmatrix} \tag{31}$$

$$m_t = \begin{bmatrix} LV_2P_1 & LV_2G_{11} & LV_2G_{12} & LV_1 - LV_2G_{13} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ s_{t-1} \\ x_{t-1} \\ x_t \end{bmatrix} + LV_2N_1 w_t + v_t \tag{32}$$

Thus the setup is as required, with the vector of predetermined variables given by $[\varepsilon_t' \ s_{t-1}' \ x_{t-1}']'$, and the vector of jump variables given by x_t. Note that there is an issue not covered by Pearlman (1992), namely that the instrument w_t is part of the measurement equation; if we assume that the instruments are observed, then there is no problem to modify the theory.

There is also a minor issue that the states of the system are not readily identifiable, as they will be linear combinations of the identifiable variables, which may make debugging of errors more problematic.

5 Passing the Model to ACES

The model setup in this form is passed from Dynare to ACES where it is in **Form 2**:

$$\begin{bmatrix} z_{t+1} \\ E_t x_{t+1} \end{bmatrix} = A \begin{bmatrix} z_t \\ x_t \end{bmatrix} + Dw_t + \begin{bmatrix} C \\ 0 \end{bmatrix} u_{t+1} \tag{33}$$

$$Y_t + E2 \begin{bmatrix} z_t \\ x_t \end{bmatrix} + E5w_t = 0 \tag{34}$$

where w_t are the instruments and u_t are the shocks. Note that in ACES notation, $B = I, AB = 0, E1 = I, E4 = 0, E3 = 0$.

For the partial information setup we also require the measurements (18):

$$m_t = LY_t + v_t \tag{35}$$

N.B. There is one difference here, namely that there is a shock v_t to the measurement Y_t. This shock v_t could also be incorporated into the state vector, by having an additional predetermined variable v_{t+1}. Also Y_t plays a different role here from what it usually does in ACES. In ACES, it represents static relationships that are included in the dynamics, whereas here Y_t represents what is observed by agents and policymakers.

The matrices above then correspond to those of the previous section via:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ P_1 & G_{11} & G_{12} & -G_{13} \\ 0 & 0 & 0 & I \\ P_3 & G_{31} - FG_{11} & G_{32} - FG_{12} & -G_{33} + FG_{13} \end{bmatrix} \quad D = \begin{bmatrix} 0 \\ N_1 \\ 0 \\ N_3 \end{bmatrix} \quad C = \begin{bmatrix} I \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{36}$$

$$E2 = -\begin{bmatrix} V_2H_1 & V_2G_{11} & V_2G_{12} & V_1 - V_2G_{13} \end{bmatrix} \qquad E5 = -V_2N_1 \tag{37}$$

6 Impulse Response Functions

Full Information Case:

It is easy to see that the impulse response functions can be calculated from

$$z_{t+1} = (A_{11} - A_{12}N)z_t + Cu_{t+1} \qquad x_t = -Nz_t \qquad Y_t = -E2 \begin{bmatrix} z_t \\ x_t \end{bmatrix} \tag{38}$$

where

$$\begin{bmatrix} N & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \Lambda^U \begin{bmatrix} N & I \end{bmatrix} \tag{39}$$

Partial Information Case: First rewrite m_t as:

$$m_t = \left[\begin{array}{cc} K_1 & K_2 \end{array}\right] \left[\begin{array}{c} z_t \\ x_t \end{array}\right] + v_t \tag{40}$$

The reduced-form solution is then given by:

$$\begin{aligned}
System: \quad z_{t+1} &= Fz_t + (A-F)\tilde{z}_t \\
&\quad + (F-A)PH^T(HPH^T+V)^{-1}(H\tilde{z}_t + v_t) + Cu_{t+1} && (41)\\
x_t &= -Nz_t + (N - A_{22}^{-1}A_{21})\tilde{z}_t \\
&\quad -(N - A_{22}^{-1}A_{21})PH^T(HPH^T+V)^{-1}(H\tilde{z}_t + v_t) && (42)\\
Innovations: \quad \tilde{z}_{t+1} &= A\tilde{z}_t - APH^T(HPH^T+V)^{-1}(H\tilde{z}_t+v_t) + Cu_{t+1} && (43)\\
Measurement: \quad m_t &= Ez_t + (H-E)\tilde{z}_t + v_t \\
&\quad -(H-E)PH^T(HPH^T+V)^{-1}(H\tilde{z}_t + v_t) \\
&= Ez_{t,t-1} + (EPH^T+V)(HPH^T+V)^{-1}(H\tilde{z}_t+v_t) && (44)
\end{aligned}$$

$$where \quad F = A_{11} - A_{12}N \quad A = A_{11} - A_{12}A_{22}^{-1}A_{21} \quad E = K_1 - K_2N \quad H = K_1 - K_2A_{22}^{-1}A_{21}$$

V is the covariance matrix of the measurement errors, and P is the solution of the Riccati equation given by

$$P = APA^T - APH^T(HPH^T+V)^{-1}HPA^T + CUC^T \tag{45}$$

and U is the covariance matrix of the shocks to the system.

Note that to obtain the impulse response for the underlying variables Y_t we use the relationship

$$Y_t = V_1x_t + V_2s_t \tag{46}$$

Noting that $s_t = [0 \;\; I \;\; 0]z_{t+1}$, it follows that we may write

$$Y_t = V_1x_t + \begin{bmatrix} 0 & V_2 & 0 \end{bmatrix}\Big(Fz_t + (A-F)\tilde{z}_t + (F-A)PH^T(HPH^T+V)^{-1}(H\tilde{z}_t+v_t)\Big) \tag{47}$$

or more simply

$$Y_t = \begin{bmatrix} 0 & V_2 & V_1 \end{bmatrix} z_{t+1} \tag{48}$$

6.1 Covariances and Autocovariances for the Partial Information Case

Pearlman *et al.* (1986) show that

$$cov\left[\begin{array}{c} \tilde{z}_t \\ z_t \end{array}\right] = \left[\begin{array}{cc} P & P \\ P & P+M \end{array}\right] \equiv P_0 \tag{49}$$

where$M$ satisfies

$$M = FMF^T + FPH^T(HPH^T + V)^{-1}HPF^T \tag{50}$$

If the dimension of the vector Y_t is n, define Ω_0 as the bottom right $n \times n$ matrix of $(P+M)$. Then it follows that

$$cov(Y_t) = \begin{bmatrix} V_2 & V_1 \end{bmatrix} \Omega_0 \begin{bmatrix} V_2^T \\ V_1^T \end{bmatrix} \equiv R_0 \tag{51}$$

To calculate the autocovariances, define

$$\Gamma = \begin{bmatrix} A(I - PH^T(HPH^T + V)^{-1}H) & 0 \\ (A - F)(I - PH^T(HPH^T + V)^{-1}H) & F \end{bmatrix} \tag{52}$$

Then the sequence of auto-covariance matrices of Y_t are defined as follows:

$$E\left(\begin{bmatrix} \tilde{z}_{t+k} \\ z_{t+k} \end{bmatrix}, \begin{bmatrix} \tilde{z}_t \\ z_t \end{bmatrix} \right) \equiv P_k = \Gamma^k P_0 = \Gamma P_{k-1} \tag{53}$$

Defining Ω_k as the bottom right $n \times n$ matrix of P_k, it follows that

$$cov(Y_{t+k}, Y_t) = E(Y_{t+k}Y_t^T) = \begin{bmatrix} V_2 & V_1 \end{bmatrix} \Omega_k \begin{bmatrix} V_2^T \\ V_1^T \end{bmatrix} \equiv R_k \tag{54}$$

These correspond to the matrices **gamma_y** defined at the bottom of page 41 of the Dynare User Guide. These are then use to generate **autocorr**, the autocorrelation functions of the variables. Thus the autocorrelation function of the ith element of Y is given by the sequence $\frac{(R_1)_{ii}}{(R_0)_{ii}}, \frac{(R_2)_{ii}}{(R_0)_{ii}}, \frac{(R_3)_{ii}}{(R_0)_{ii}}, ...$

In addition the **correlation matrix** of the Y_t variables is defined as

$$Corr = \Delta R_0 \Delta^T \text{ where } \Delta = diag(\surd(R_0)_{11}, \surd(R_0)_{22}, \surd(R_0)_{33}, ...) \tag{55}$$

7 Likelihood function calculation

Here we assume that there are no policy instruments w_t and that the system is saddlepath stable.

The Kalman filtering equation is given by

$$z_{t+1,t} = Fz_{t,t-1} + FP_tH^T(EP_tH^T + V)^{-1}e_t \tag{56}$$

where $e_t = m_t - Ez_{t,t-1}$

$$P_{t+1} = AP_tA^T + U - AP_tH^T(HP_tH^T + V)^{-1}HP_tA^T \tag{57}$$

the latter being a time-dependent Ricatti equation.

The period-t likelihood function is standard:

$$2lnL = -\sum lndet(cov(e_t) - \sum e_t^T(cov(e_t))^{-1}e_t \tag{58}$$

where

$$cov(e_t) = (EP_tH^T + V)(HP_tH^T + V)^{-1}(HP_tE^T + V) \tag{59}$$

Following Pearlman *et al.* (1986), the system is initialised at

$$z_{1,0} = 0 \qquad P_1 = P + M \tag{60}$$

where $P$is the steady state of the Riccati equation above, and M is the solution of the Lyapunov equation

$$M = FMF^T + FPH^T(HPH^T + V)^{-1}HPF^T \tag{61}$$

8 Extension to the case of Expectations of Current Variables

Suppose that expectations (or best estimates) of current variables are included in agents' decision-making and measurements. Then a general setup will be of the form

$$A_0Y_{t+1,t} + A_1Y_t = A_2Y_{t-1} + A_3Y_{t,t} + Bu_t \qquad m_t = LY_t + MY_{t,t} + v_t \tag{62}$$

To get this into Blanchard-Kahn format, we follow the same procedures as above with $Y_{t,t}$ as a member of the exogenous variables, and end up with a representation of the form

$$\begin{bmatrix} z_{t+1} \\ E_tx_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \begin{bmatrix} z_{t,t} \\ x_{t,t} \end{bmatrix} + \begin{bmatrix} C \\ 0 \end{bmatrix} \varepsilon_{t+1} \tag{63}$$

$$m_t = \begin{bmatrix} K_1 & K_2 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} R_1 & R_2 \end{bmatrix} \begin{bmatrix} z_{t,t} \\ x_{t,t} \end{bmatrix} + v_t \tag{64}$$

Then all the equations above for filtering, likelihood calculation, IRFs are identical, with the following altered definitions:

$$F = A_{11} + J_{11} - (A_{12} + J_{12})N \qquad E = K_1 + R_1 - (K_2 + R_2)N \tag{65}$$

$$\begin{bmatrix} N & I \end{bmatrix} \begin{bmatrix} A_{11} + J_{11} & A_{12} + J_{12} \\ A_{21} + J_{21} & A_{22} + J_{22} \end{bmatrix} = \Lambda^U \begin{bmatrix} N & I \end{bmatrix} \tag{66}$$

Part III

Endogenous Persistence in an Estimated DSGE Model under Imperfect Information

Presented at the annual CDMA Conference, 2009, University of St Andrews, 2–4 September.
Paul Levine, University of Surrey
Joseph Pearlman, London Metropolitan University
George Perendia, London Metropolitan University
Bo Yang, University of Surrey

Abstract

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 standard 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 only 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 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.

JEL Classification: C11, C52, E12, E32.

Keywords: Imperfect Information, DSGE Model, Bayesian Estimation

9 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 are on offer 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 (2001) and Svensson and Woodford (2003) 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 is that agents can process information subject to a constraint that places an upper bound on the information flow. These studies assume 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 have been recently developed by Woodford (2003) and Pearlman and Sargent (2003).

A more drastic deviation from rational expectations is provided by the statistical rational learning literature pioneered by Evans and Honkapohja (2000). 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. All these refinements contrast with the drastic behavioural alternative offered by the very recent 'Animal Spirits' approach (Akerlof and Shiller (2009), DeGrauwe (2009)), the latter paper proposing a radical alternative to a standard New Keynesian model with rational expectations.

At the same time the formal estimation of DSGE models by Bayesian methods has become standard.[1] However, as Levine *et al.* (2007a) first point out, most of this DSGE estimation makes asymmetric information assumptions where perfect information about past shocks 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 in order to assess whether parameter estimates are consistent across these assumptions and whether these alternatives lead to a better model fit.

In this paper we stay within the conventional rational expectations framework but relax the extreme perfect information assumptions for the private sector. We focus on a fairly standard New Keynesian (NK) macro-model, and make the assumption that either agents

[1] See Fernandez-Villaverde (2009) for an excellent review.

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 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 good economists. Whilst we make greater cognitive demands on both economic agents and the economists in sticking with rational (model-consistent) expectations, our assumption that agents have no more information than the economist constructing and estimating 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. With perfect information, the hump-shaped effect is not in evidence in simulations of the NK model. Collard and Dellas (2006) in particular is able to reproduce this without resorting to lagged price indexation. The implication of these examples is that since VAR estimates of macroeconomic aggregates, when simulated, lead to hump-shaped responses, it is easier to get a closer fit to the data when we assume that information about some or all of these aggregates is imperfect. 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 comparison with a DSGE-VAR.[3]

The rest of the paper is organized as follows. Section 2 sets out the model. Section

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

[3] Since writing and presenting this paper we came across Collard *et al.* (2009) which carries out a similar exercise using the solution method of Pearlman *et al.* (1986). We provide more details of the methodology and a formal comparison with the rational inattention approach of Sims (2005). 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.

3 sets out the solution method (summarizing Pearlman *et al.* (1986)) and pays particular attention to the issue of log-linearization. Sections 4 provides an analytical solution for a simplified version of our model. Sections 5 and 6 set out and discuss the results of our Bayesian estimation. Section 7 concludes.

10 The Model

We utilize a simple 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, and the shocks in particular, using Kalman-filter updating.

The full model in non-linear form is as follows

$$
\begin{aligned}
1 &= \beta(1+R_t)E_t\left[\frac{MU_{t+1}^C}{MU_t^C\Pi_{t+1}}\right] && (67)\\
\frac{W_t}{P_t} &= -\frac{1}{(1-\frac{1}{\eta})}\frac{MU_t^L}{MU_t^C} && (68)\\
Y_t &= A_tL_t^\alpha && (69)\\
MC_t &= \frac{W_t}{\alpha A_tP_tL_t^{\alpha-1}} && (70)\\
H_t-\xi\beta E_t[\tilde{\Pi}_{t+1}^{\zeta-1}H_{t+1}] &= Y_tMU_t^C && (71)\\
J_t-\xi\beta E_t[\tilde{\Pi}_{t+1}^{\zeta}J_{t+1}] &= \frac{\zeta}{\zeta-1}MC_tMS_tY_tMU_t^C && (72)\\
1 &= \xi\tilde{\Pi}_t^{\zeta-1}+(1-\xi)\left(\frac{J_t}{H_t}\right)^{1-\zeta} && (73)\\
\tilde{\Pi}_t &\equiv \frac{\Pi_t}{\Pi_{t-1}^\gamma} && (74)\\
Y_t &= C_t+G_t && (75)
\end{aligned}
$$

Equation (67) 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 expectations conditional upon a general information set (see next section). In (68) 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 the marginal utility of labour supply L_t. Equation (69) is the production function with labour the only variable input into production and the technology shock A_t exogenous. Equation (70) defines the marginal cost. Equations (71) to (74) describe Calvo pricing with $1-\xi$ equal to the probability of a monopolistically competitive firm re-optimizing its price, indexing by an amount γ and an exogenous mark-up shock MS_t (see Levine *et al.* (2007b)). Finally (75), where C_t denotes consumption, describes output equilibrium, with an exogenous government spending demand shock G_t.

To close the model we assume a general one-period ahead Inflation Forecast Based Taylor-type interest-rule

$$\log(1+R_t) = \rho_r \log(1+R_{n,t-1})+(1-\rho_r)\left(\theta_\pi E_t\left[\log\frac{\Pi_{t+1}}{\Pi}\right]+\log(\frac{1}{\beta})+\theta_y E_t\left[\log\frac{Y_t}{Y_t^*}\right]\right)+\epsilon_{e,t} \tag{76}$$

where Y_t^* is the flexi-price natural rate of output and $\epsilon_{e,t}$ is a monetary policy shock.

The form of the single period utility for household r is a non-separable function of consumption and labour effort that is consistent with a balanced growth steady state is:

$$U \equiv \frac{\left[(C_t(r)-h_C C_{t-1})^{1-\varrho}(1-L_t(r))^{\varrho}\right]^{1-\sigma}}{1-\sigma} \tag{77}$$

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

$$MU_t^C = (1-\varrho)(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)-1}(1-L_t)^{\varrho(1-\sigma)} \tag{78}$$

$$MU_t^L = -(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)}\varrho(1-L_t)^{\varrho(1-\sigma)-1} \tag{79}$$

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

$$\log\frac{A_{t+1}}{A} = \rho_a \log\frac{A_t}{A} + \epsilon_{a,t+1} \tag{80}$$

$$\log\frac{G_{t+1}}{G} = \rho_g \log\frac{G_t}{G} + \epsilon_{g,t+1} \tag{81}$$

where X denotes the non-stochastic balanced growth value or path of the variable X_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) we decompose the price mark-up shock into persistent and transient component: $MS_t = MS_{per,t}MS_{tran,t}$ where

$$\log\frac{MS_{per,t+1}}{MS_{per}} = \rho_{ms}\log\frac{MS_{per,t}}{MS_{per}} + \epsilon_{msper,t+1} \tag{82}$$

$$\log\frac{MS_{tra,t+1}}{MS_{tra}} = \epsilon_{mstra,t+1} \tag{83}$$

This results in MS_t being an ARMA(1,1) process.

Note that 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 13 equations, (67)–(75) and (76)–(79) defining 13 endogenous variables Π_t $\tilde{\Pi}_t$ C_t Y_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 natural rate Y_t^* is determined by the same set of equations with $MC_t = 0$.

The log-linearization[4] of the model about the non-stochastic steady state is given by

$$\begin{aligned}
\pi_t &= \frac{\beta}{1+\beta\gamma} E_t\pi_{t+1} + \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{(1-\beta\xi)(1-\xi)}{(1+\beta\gamma)\xi} mc_t + ms_t \\
w_t - p_t &= mu_t^L - mu_t^C
\end{aligned}$$

$$\begin{aligned}
y_t &= a_t + \alpha l_t \\
mc_t &= w_t - p_t - a_t + (1-\alpha) l_t \\
E_t mu_{t+1}^C &= mu_t^C - (r_t - E_t\pi_{t+1}) \\
mu_t^C &= \frac{(1-\varrho)(1-\sigma)-1}{1-h_C}(c_t - h_C c_{t-1}) - \frac{\varrho(1-\sigma)L}{1-L} l_t \\
mu_t^L &= \frac{1}{1-h_C}(c_t - h_C c_{t-1}) + \frac{L}{1-L} l_t + mu_t^C \\
y_t &= c_y c_t + (1-c_y) g_t \quad \text{where } c_y = \frac{C}{Y} \\
g_{t+1} &= \rho_g g_t + \epsilon_{g,t+1} \\
a_{t+1} &= \rho_a a_t + \epsilon_{a,t+1} \\
msper_{t+1} &= \rho_{ms} msper_t + \epsilon_{msper,t+1} \\
ms_t &= msper_t + \epsilon_{mstra,t}
\end{aligned}$$

In the log-linearized model the natural rate of output, y_t^* say, is given by putting $mc_t = mc_t^* = 0$; i.e.,

$$\begin{aligned}
w_t^* - p_t^* &= mrs_t^* = \frac{1}{1-h_C}(c_t^* - h_C c_{t-1}^*) + \frac{L}{1-L} l_t^* \\
(1-\alpha) l_t^* &= a_t - (w_t^* - p_t^*) \\
y_t^* &= c_y c_t^* + (1-c_y) g_t \\
y_t^* &= a_t + \alpha l_t^* \\
o_t &= y_t - y_t^*
\end{aligned}$$

which defines the output gap, o_t, as a function of the shocks a_t and g_t. The log-linearized interest rate rule is then

$$r_t = \rho_r r_{t-1} + (1-\rho_r)[\theta_\pi E_t\pi_{t+1} + \theta_y E_t o_t] + \epsilon_{e,t} \tag{84}$$

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

[4]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.

of the entire state vector mu_t^C, π_t, π_{t-1}, c_{t-1}, c_{t-1}^* and crucially the output gap (so that $E_t o_t = o_t$ in (84)), current shocks $msper_t$, ms_t, a_t, and g_t. Since all the other macro-variables can be expressed in terms of the state variables it follows that the information set of the private sector must include these as well. 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*. This we now turn to.

11 General Rational Expectations Solution under Perfect and Imperfect Information

Our model is a special case of the following general setup in non-linear form

$$\begin{aligned} Z_{t+1} &= J(Z_t, E_t Z_t, X_t, E_t X_t) + \nu\sigma\epsilon_{t+1} && (85) \\ E_t X_{t+1} &= K(Z_t, E_t Z_t, X_t, E_t X_t) && (86) \end{aligned}$$

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

$$\begin{bmatrix} z_{t+1} \\ E_t x_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + B \begin{bmatrix} E_t z_t \\ E_t x_t \end{bmatrix} + \begin{bmatrix} u_{t+1} \\ 0 \end{bmatrix} \tag{87}$$

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

$$\begin{aligned} W_t &= m(Z_t, E_t Z_t, X_t, E_t X_t) + \mu\sigma\epsilon_t && (88) \\ w_t &= \begin{bmatrix} M_1 & M_2 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + L \begin{bmatrix} E_t z_t \\ E_t x_t \end{bmatrix} + v_t && (89) \end{aligned}$$

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 $E_t Z_t$ and $E_t X_t$ 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).[5]

11.1 Linear Approximation about the Non-Stochastic Steady State

Before proceeding to the rational expectations solution, we need to pose a basic question: is (87) 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?

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.

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 (85) and (86) under imperfect information which involves the innovations process variable $\tilde{Z}_t \equiv Z_t - E_{t-1}Z_t$:

$$X_t = g(Z_t, \tilde{Z}_t, \sigma)\,; \qquad Z_{t+1} = h(Z_t, \tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}\,; \qquad \tilde{Z}_{t+1} = f(\tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1} \quad (90)$$

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).

11.2 Rational Expectations Solution

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

$$x_t + Nz_t = 0 \quad \text{where} \quad \begin{bmatrix} N & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \Lambda^U \begin{bmatrix} N & I \end{bmatrix} \quad (91)$$

where Λ^U has unstable eigenvalues.

[5] Our model reduces to this form if we assume a pure inflation targeting rule with $\theta_y = 0$ in (76) and (84). In fact we find our empirical results to change very little with this simplification.

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

$$\begin{bmatrix} z_{t,t} \\ x_{t,t} \end{bmatrix} = \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix} + J \left[w_t - M \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix} \right] \tag{92}$$

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

$$J = \begin{bmatrix} PD^T \\ -NPD^T \end{bmatrix} \Gamma^{-1} \tag{93}$$

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 (97) below.

Using the Kalman filter, the solution as derived by Pearlman *et al.* (1986)[6] 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 $\tilde{z}_t \equiv z_t - z_{t,t-1}$:

$$\begin{aligned}
\text{Predetermined}: \quad z_{t+1} &= Cz_t + (A-C)\tilde{z}_t + (C-A)PD^T(DPD^T+V)^{-1}(D\tilde{z}_t + v_t) \\
&+ u_{t+1} &(94)\\
\text{Non-predetermined}: \quad x_t &= -Nz_t + (N - A_{22}^{-1}A_{21})\tilde{z}_t &(95)\\
\text{Innovations}: \quad \tilde{z}_{t+1} &= A\tilde{z}_t - APD^T(DPD^T+V)^{-1}(D\tilde{z}_t + v_t) + u_{t+1} &(96)
\end{aligned}$$

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

and P is the solution of the Riccati equation given by

$$P = APA^T - APD^T(DPD^T + V)^{-1}DPA^T + U \tag{97}$$

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

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 $k_t - \tilde{k}_t$.

On the theme of rational inattention, 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)),

[6] 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.

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 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[7] 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 $\sigma_k^2/(h^T Ph+\sigma_k^2)$. 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.

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

$$z_{t+1} = Cz_t + u_{t+1} \tag{98}$$

$$x_t = -Nz_t \tag{99}$$

12 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:

$$\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\beta\xi)(1-\xi)}{\xi} mc_t + ms_t \tag{100}$$

which for convenience we write as

$$E_t \pi_{t+1} = \frac{1}{\beta}\pi_t + x_t + w_t \tag{101}$$

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

$$x_{t+1} = \rho x_t + \varepsilon_{t+1} \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{102}$$

[7] 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 channel capacity is given by $\frac{1}{2}(log(p_k) - log(p_k\sigma_k^2/(p_k+\sigma_k^2)))$.

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

$$\begin{bmatrix} w_{t+1} \\ x_{t+1} \\ E_t[\pi_{t+1}] \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \rho & 0 \\ 1 & 1 & \alpha \end{bmatrix} \begin{bmatrix} w_t \\ x_t \\ \pi_t \end{bmatrix} + \begin{bmatrix} w_{t+1} \\ \varepsilon_{t+1} \\ 0 \end{bmatrix} \tag{103}$$

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. $\begin{bmatrix} w_t \\ x_t \\ \pi_t \end{bmatrix}$. We compare this with imperfect information where agents observe only inflation π_t. Then from our general solution procedure in section 3 the following matrices are defined:

$$A = C = \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \qquad N = \begin{bmatrix} \frac{1}{\alpha} & \frac{1}{\alpha - \rho} \end{bmatrix} = -E \qquad D = \begin{bmatrix} -\frac{1}{\alpha} & -\frac{1}{\alpha} \end{bmatrix} \qquad U = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} \tag{104}$$

It follows from (97) that

$$P = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & p \end{bmatrix} \qquad \text{where } p = \frac{\rho^2 p \sigma_w^2}{\sigma_w^2 + p} + \sigma_\varepsilon^2 \tag{105}$$

From (94) it follows that the innovations are given by

$$\begin{bmatrix} \tilde{w}_{t+1} \\ \tilde{x}_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \tilde{w}_t \\ \tilde{x}_t \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \sigma_w^2 \\ p \end{bmatrix} \frac{(\tilde{w}_t + \tilde{\pi}_t)}{(p + \sigma_w^2)} + \begin{bmatrix} w_{t+1} \\ \epsilon_{t+1} \end{bmatrix} \tag{106}$$

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

$$\begin{aligned} x_t &= \rho x_{t-1} + \varepsilon_t && (107) \\ \tilde{x}_t &= \frac{\rho}{\sigma_w^2 + p}(\sigma_w^2 \tilde{x}_{t-1} - p w_{t-1}) + \varepsilon_t && (108) \\ \pi_t &= -\frac{1}{\alpha}\left(1 + \frac{\rho \sigma_w^2 p}{(\alpha - \rho)(\sigma_w^2 + p)}\right) w_t - \frac{1}{\alpha - \rho} x_t + \frac{\rho \sigma_w^2}{\alpha(\alpha - \rho)(\sigma_w^2 + p)} \tilde{x}_t && (109) \end{aligned}$$

Figure 1 in Appendix D 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 (109) 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.

13 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 the linearized model described below 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 acceptation rate (20%-40%). A number of parallel Markov chains of 100000 runs each are run for the MH in order to ensure the chains are converged. The first 25% of iterations (initial burn-in period) are discarded in order to remove any dependence of the chain from its starting values.

13.1 Data and Priors

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.* (2005) and Smets and Wouters (2007). Table 4 in Appendix C 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 center 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).

Three of the structural parameters are kept fixed in the estimation procedure. These calibrated arameters are $\beta = 0.99$; $L = 0.4$, $c_y = 0.6$ (Note the latter allows for investment with capital in the model, but held fixed).

13.2 Model Comparisons

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 symmetrical 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 in Table 1 we report the posterior marginal data density

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

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 with the highest marginal data density attains the best relative model fit.

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:

$$p_i(\theta|y, m_i) \propto L(y|\theta, m_i)\, p_i(\theta|m_i) \tag{110}$$

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,

$$L(y|m_i) = \int_\Theta L(y|\theta, m_i)\, p(\theta|m_i)\, d\theta \tag{111}$$

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):

$$PO_{i,j} = \frac{p(m_i|y)}{p(m_j|y)} = \frac{L(y|m_i)p(m_i)}{L(y|m_j)p(m_j)} \tag{112}$$

$$BF_{i,j} = \frac{L(y|m_i)}{L(y|m_j)} = \frac{\exp(LL(y|m_i))}{\exp(LL(y|m_j))} \tag{113}$$

in terms of the log-likelihoods. Components (112) and (113) 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 easily 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

$$\frac{1}{p_1} = \sum_{i=2}^{n} BF_{i,1} \tag{114}$$

from which p_1 is obtained. Then $p_i = p_1 BF(i, 1)$ gives the remaining model probabilities. These are in Table 2 reported as follows. Let the probability of model variant G, information assumption IIME say be Pr(G, IIME). Then the probability ranking is shown in Table 2.

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
Pr(Z, IIME)=0.0000
Pr(GH, AI)=0.0000
Pr(Z, II)=0.0000
Pr(Z, AI)=0.0000
Pr(G, AI)=0.0000

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 than 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 pitfall or limitation of this methodology is that the assessment of how fit a model 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).

13.3 Posterior Estimates

Table 5 in Appendix C 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.* (2005) 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.* (2005) 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.

As noted, it is not sufficient, at least at this stage, to establish the superiority of the II version over its perfect information counterpart against data based on the posterior estimates and data preferences. In what follows, we carry out a more general evaluation that is based on impulse responses, moment criteria and autocorrelations.

13.4 Impulse Response Analysis

This subsection investigates the importance of shocks to the endogenous variables of interests by analysing 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' model) and Model Z (contains zero persistence) across different information sets (i.e. AI and II). The aim is to investigate the impact of changing information assumptions in terms of the impulse response dynamics.

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, an alternative solution to improving VAR estimates by 'restricting' its parameter estimates is to tilt estimates toward a point in the parameter space. Careful construction of VAR prior (or restricting VARs) is crucial because matching impulse responses in the data and in the model requires the identification restrictions imposed on the VAR are consistent and compatible with the theoretical model. In order to estimate the identified IRFs from a VAR model, we follow the so-called DSGE-VAR approach proposed by Del Negro and Schorfheide (2004), where they use the DSGE model itself to construct a prior distribution for the VAR coefficients so that DSGE-VAR estimates are tilted toward DSGE model restriction and they find the resulting model can be useful for policy analysis.

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.[8] We use a data-driven procedure to determine the tightness of 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 $\lambda = [0; 0.25; 0.5; 0.75; 1; 2; 5; \infty]$, we find that $\lambda = 0.5$ has the highest posterior probability. Overall, DSGE-VAR(4) with $\lambda = 0.5$ has the highest posterior probability.[9] 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 D depicts the mean responses corresponding to a positive one 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 asymmetrical 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

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

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

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.

Following a positive technological shock, inflation and 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.* (2005) 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.

With respect to the remaining shocks, 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. Finally, the performance of changing the information assumptions in terms of the impulse response dynamics on output and interest rate to the transient component of the mark-up shock (Mark-up (m)) appears ambiguous to interpret. In this case, the effects are opposite to the data.

Overall, these results from the estimated posterior impulse responses, combined with the simulated IRF based on the simple calibrated example (Figure 1), mainly imply that the presence of the II specification in the US economy is capable of making an impact on producing different model-based dynamics. This further confirms the above findings that there is, quantitatively, substantial evidence in the data to support the assumption that both agents and econometricians have only the same imperfect information available.

14 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.

14.1 Standard Moment Criteria

To assess the contributions of assuming different information sets in our estimated model variants, we compute some selected second moments and present the results in this subsection. Table 3 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 (i.e. GH and Z).

In terms of the standard deviations, almost all the models generate relative high volatility compared to the actual data (except for output). In line with the Bayesian model comparison, Model GH (assuming imperfect information) fits the data better in terms of implied volatility, getting closer to the data in this dimension. Overall, the estimated mod-

Model	Output	Inflation	Interest rate
Standard Deviation			
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
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
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

TABLE 3: Selected Second Moments

els 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 models under investigation appear to match well the autocorrelations (order=1) of all the endogenous variables. Table 3 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), 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.

14.2 Unconditional Autocorrelations

To further illustrate how the estimated models capture the data statistics based on different information 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 models match reasonably well the autocorrelations shown in the data within a 10-period horizon and our 'best' 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, but at lags of 2 and 3 quarters, the model simulated autocorrelations under AI are greater than those of the sample for the interest rate and inflation. When it comes to matching the interest rate, all models do a better job, getting closer to the observed persistence.

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. Inflation in Model GH is less autocorrelated than in the data from lags of 2 quarters, but it becomes closer to the data towards the end of sample period. The results in this exercise again 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.

15 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.

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.

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 (symmetrical) information fits the observed autocorrelation of the data of the interest rate and inflation extremely well, whereas the same model with perfect (asymmetrical) information on the part of the private sectors results in a poor fit in this dimension. Second we study symmetrical 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.* (1983), 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 model) to be empirically supported 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 Poicy with Imperfect Common Knowledge. *Journal of Monetary Economics*, **54**(2), 267–361.

Adolfson, M., Laseen, S., Linde, J., and Villani, M. (1983). Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through. *Journal of International Economics*, **72**, 481–551.

Akerlof, G. A. and Shiller, R. J. (2009). *Animal Spirits*. Princeton University Press.

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*.

Chang, Y., Gomes, J. F., and Schorfheide, F. (2002). Learning-by-Doing as a propagation Mechanism. *American Economic Revie*, **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*, **112**, 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. Working Paper 491, ECB.

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.*

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.

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. (2005). Monetary Policy Under Uncertainty in Micro-Founded Macroeconomic Models. In M. Gertler and K. Rogoff, editors, *New Challenges for Monetary Policy*, pages 229–387. NBER Macroeconomics Annual,.

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

Levine, P., McAdam, P., and Pearlman, J. (2007b). Quantifying and Sustaining Welfare Gains from Monetary Commitment. ECB Working Paper No. 709 and *Journal of Monetary Economics*, forthcoming.

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. *B.E. Journal of Macroeconomic*, **9**.

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

Minford, A. and Peel, D. (1983). Some Implications of Partial Information Sets in Macroeeconomic Models Embodying Rational Expectations. *Manchester School*, **51**, 225–249.

Pearlman, J. (1992). Reputational and Non-reputational Policies under Partial Information. *Journal of Economic Dynamics and Control*, **16**(2), 339–357.

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

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

Reis, R. (2009). A Sticky Information General Equilibrium Model for Policy Analysis. *NBER Working Paper 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. (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 Varibles for Optimal Policy under Asymmetric Information. *Journal of Economic Dynamics and Control*, **28**(4), 661–690.

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 Proof of Theorem

Assume a model of the form

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

where σ is small, and with measurements

$$W_t = L(Z_t, X_t) \tag{A.2}$$

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

$$X_t = g(Z_t, \tilde{Z}_t, \sigma) \qquad Z_{t+1} = h(Z_t, \tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1} \qquad \tilde{Z}_{t+1} = f(\tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1} \tag{A.3}$$

Also assume that

$$E_t Z_t - E_t Z_{t-1} = E_t \tilde{Z}_t = i(\tilde{Z}_t) \tag{A.4}$$

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

$$K(Z_t, g(Z_t, \tilde{Z}_t, \sigma)) = E_t K(Z_t, g(Z_t, \tilde{Z}_t, \sigma)) \tag{A.5}$$

and the 1st order approximation to this is

$$K_1 z_t + K_2 g_1 z_t + K_2 g_2 \tilde{z}_t = K_1(z_t - \tilde{z}_t + i_1\tilde{z}_t) + K_2 g_1(z_t - \tilde{z}_t + i_1\tilde{z}_t) + K_2 g_2 i_1 \tilde{z}_t \tag{A.6}$$

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 (A.6) we obtain

$$K_1 + K_2 g_1 = g_1 h_1 \qquad K_1 + K_2 g_1 + K_2 g_2 = (K_1 + K_2 g_1 + K_2 g_2) i_1 \tag{A.7}$$

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:

$$\begin{aligned} K(Z_t, X_t) &= K(Z_t, g(Z_t, \tilde{Z}_t, \sigma)) = E_t g(Z_{t+1}, \tilde{Z}_{t+1}, \sigma) \\ &= E_t g(h(Z_t, \tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}, f(\tilde{Z}_t, \sigma) + \eta\sigma\varepsilon_{t+1}, \sigma) \end{aligned} \tag{A.8}$$

The first order approximation to this is

$$K_1 z_t + K_2 g_1 z_t + K_2 g_2 \tilde{z}_t + K_2 g_\sigma \sigma = g_1 h_1(z_t - \tilde{z}_t + i_1\tilde{z}_t) + g_1 h_2 i_1 \tilde{z}_t + g_2 f_1 i_1 \tilde{z}_t + g_1 h_\sigma \sigma + g_2 f_\sigma \sigma + g_\sigma \sigma \tag{A.9}$$

from which it follows that

$$K_1 + K_2 g_1 = g_1 h_1 \qquad K_2 g_2 = -g_1 h_1 + g_1 h_1 i_1 + g_1 h_2 i_1 + g_2 f_1 i_1 \tag{A.10}$$

$$K_2 g_\sigma = g_1 h_\sigma + g_2 f_\sigma + g_\sigma \tag{A.11}$$

Equating the two Z_{t+1} equations implies

$$h(Z_t, \tilde{Z}_t, \sigma) = J(Z_t, X_t) = J(Z_t, g(Z_t, \tilde{z}_t, \sigma)) \tag{A.12}$$

so that to first order

$$J_1 z_t + J_2 g_1 z_t + J_2 g_2 \tilde{z}_t + J_2 g_\sigma \sigma = h_1 z_t + h_2 \tilde{z}_t + h_\sigma \sigma \tag{A.13}$$

and hence

$$J_1 + J_2 g_1 = h_1 \qquad J_2 g_2 = h_2 \qquad J_2 g_\sigma = h_\sigma \tag{A.14}$$

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, but evaluated always around the steady state. 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

$$E_t z_t = E_{t-1} z_t + PH^T (HPH^T)^{-1} H \tilde{z}_t \tag{A.15}$$

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

$$i_1 = PH^T (HPH^T)^{-1} H \tag{A.16}$$

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

$$P = f_1 P f_1^T + \sigma^2 \eta \eta^T \tag{A.17}$$

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

$$(h_1 + h_2)(I - i_1) = f_1 \qquad f_\sigma = 0 \tag{A.18}$$

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 :

$$(K_2 - I) g_\sigma = g_1 h_\sigma \qquad J_2 g_\sigma = h_\sigma \tag{A.19}$$

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 (A.19) 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.

B Background: Extended Kalman Filter

Consider the following system:

$$x_{t+1} = f(x_t) + w_t \qquad z_t = h(x_t) + v_t \tag{B.20}$$

Define the information set $I_t = \{z_t, z_{t-1}, z_{t-2}, ...\}$. Suppose we assume that

$$p(x_t|I_{t-1}) \sim N(x_t; \hat{x}_{t,t-1}, P_t) \tag{B.21}$$

It follows that

$$p(x_t \mid I_t) = \frac{N(x_t; \hat{x}_{t,t-1}, P_t) N(z_t; h(x_t), V)}{\int N(x_t; \hat{x}_{t,t-1}, P_t) N(z_t; h(x_t), V) dx_t} \tag{B.22}$$

Now assume that both x_t and $\hat{x}_{t,t-1}$ are near to the equilibrium level x. It follows that we may write a linear approximation

$$h(x_t) = h(x) + H(x_t - x) \tag{B.23}$$

where $H = h'(x)$. Hence

$$p(x_t \mid I_t) \cong N(x_t; x_{t,t}, P_{t,t}) \qquad x_{t,t} = x_{t,t-1} + P_t H'(HP_tH' + V)^{-1}(z_t - h(x) - Hx_{t,t-1}) \tag{B.24}$$

$$P_{t,t} = P_t - P_t H'(HP_tH' + V)^{-1} HP_t \tag{B.25}$$

These latter two formulae arise after performing the integration in (B.22) using (B.23). Note that the extended Kalman filter usually used in the literature linearized $h()$ about $x_{t,t-1}$. However,if we assume that the system is always close to its equilibrium value x, then there is little that is lost by linearizing about x.

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 exact filter calculated numerically.

C Priors and Posterior Estimates

Parameter	Notation	Prior distribution		
		Density	Mean	S.D/df
Risk aversion	σ	Normal	2.00	0.50
Price indexation	γ	Beta	0.50	0.15
Calvo prices	ξ	Beta	0.50	0.10
Consumption habit formation	h_C	Beta	0.50	0.20
Preference parameter	ϱ	Beta	0.50	0.20
Labour share	α	Normal	0.80	0.10
Interest rate rule				
Inflation	θ_π	Normal	2.00	0.50
Output gap	θ_y	Normal	0.125	0.05
Interest rate smoothing	ρ_r	Beta	0.80	0.10
AR(1) coefficient				
Technology	ρ_a	Beta	0.85	0.10
Government spending	ρ_g	Beta	0.85	0.10
Price mark-up	ρ_{ms}	Beta	0.50	0.20
Standard deviation of AR(1) innovations				
Technology	$sd(\epsilon_a)$	Inv. gamma	0.60	2.00
Government spending	$sd(\epsilon_g)$	Inv. gamma	1.67	2.00
Price mark-up	$sd(\epsilon_{ms})$	Inv. gamma	0.10	2.00
Standard deviation of I.I.D. shocks/mearsument errors				
Mark-up process	$sd(\epsilon_m)$	Inv. gamma	0.10	2.00
Monetary policy	$sd(\epsilon_e)$	Inv. gamma	0.10	2.00
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 4: Prior Distributions

	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 5: BAYESIAN POSTERIOR DISTRIBUTIONS$^\diamondsuit$

$\diamondsuit$ 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

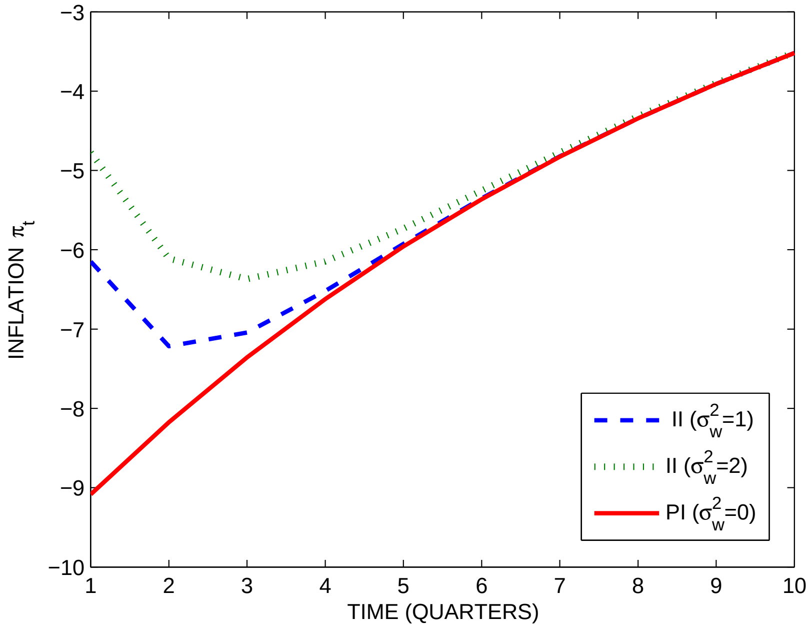

Figure 1: Inflation Dynamics under Perfect (PI) and Imperfect Information (II)

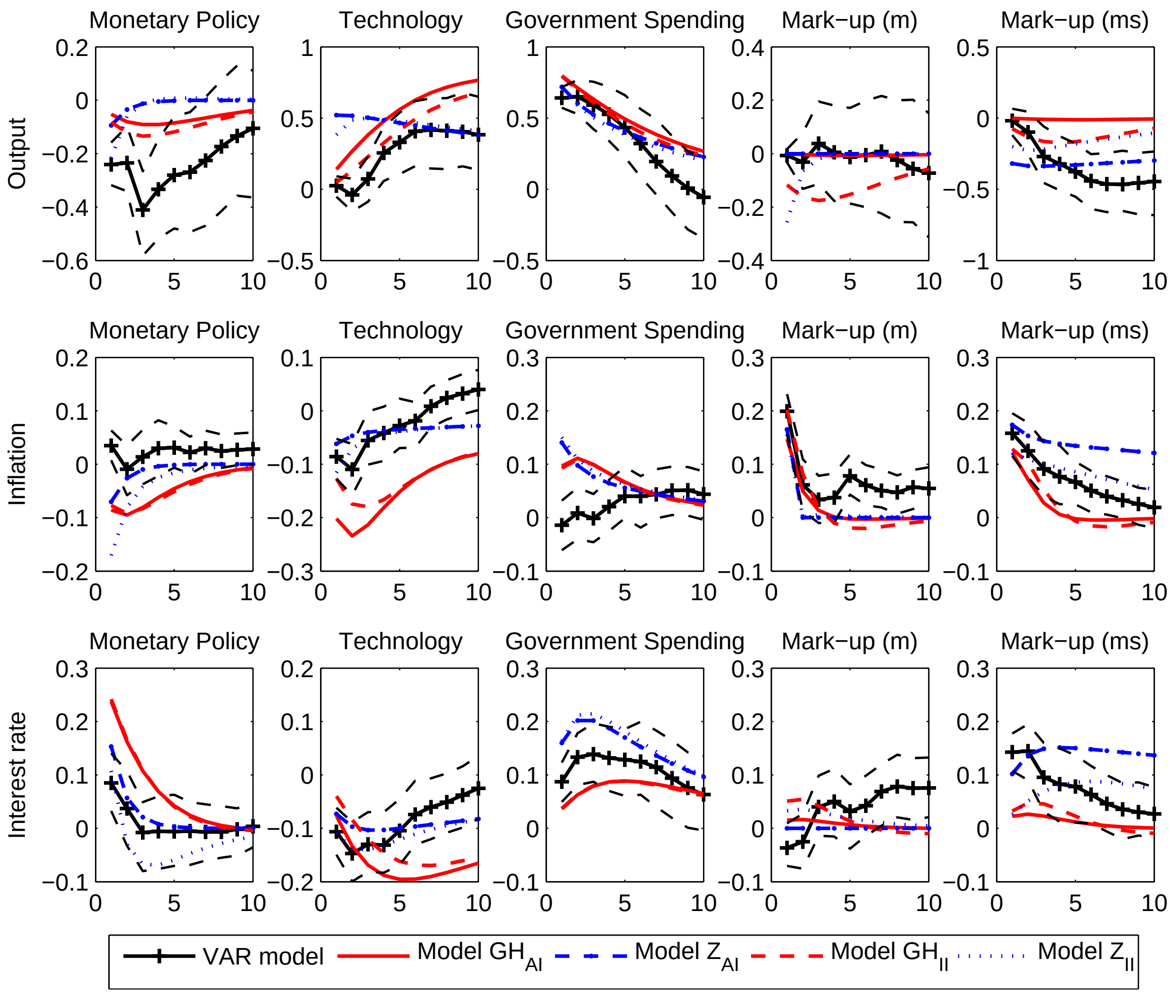

Figure 2: Estimated Impulse Response Functions - AI vs. II$^{\diamondsuit}$

$^{\diamondsuit}$ 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.

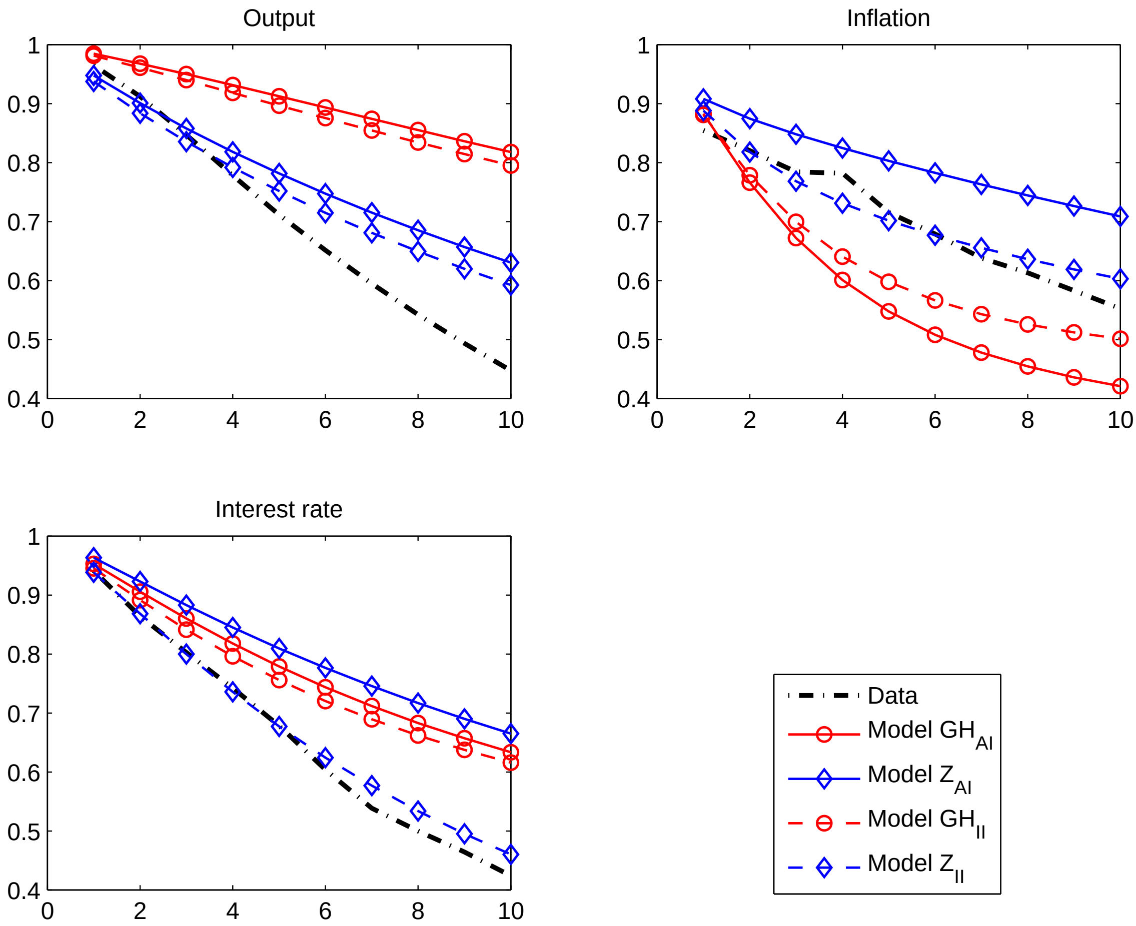

FIGURE 3: AUTOCORRELATIONS OF OBSERVABLES IN THE ACTUAL DATA AND IN THE ESTIMATED MODELS