

# Imperfect Information, Optimal Monetary Policy and the Informational Consistency Principle\*

PAUL LEVINE  
UNIVERSITY OF SURREY

JOSEPH PEARLMAN  
LONDON METROPOLITAN UNIVERSITY

BO YANG  
UNIVERSITY OF SURREY

FEBRUARY 21, 2011

## Abstract

This paper examines the implications of imperfect information (II) for optimal monetary policy with a consistent set of informational assumptions for the modeller and the private sector – the ‘Informational Consistency Principle’ (ICP). We use an estimated simple NK model from Levine *et al.* (2010), where the imposition of the ICP significantly improves the fit of the model to US data. The questions we then address are first, what are the welfare costs associated with the private sector possessing only II of the state variables; second, what are the implications of II for the gains from commitment; third, how does II affect the form of optimized Taylor rules and finally how do interest rate zero-lower-bound considerations impact on optimal policy with II.

**JEL Classification:** C11, C52, E12, E32.

**Keywords:** Imperfect Information, DSGE Model, Optimal Monetary Policy, Bayesian Estimation, zero lower bound.

---

\*An earlier version of this paper was presented at a CIMS workshop on February 16, 2011 at the University of Surrey; this version has benefited from the comments of participants. We acknowledge financial support for this research from the ESRC, project no. RES-000-23-1126 and from the EU Framework Programme 7 project MONFISPOL.

# CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Model</b>	<b>2</b>
<b>3</b>	<b>Bayesian Estimation</b>	<b>4</b>
3.1	Data and Priors . . . . .	5
3.2	Estimation Results . . . . .	6
<b>4</b>	<b>The General Set-Up</b>	<b>9</b>
4.1	State-Space Representation and Information Sets . . . . .	9
4.2	LQ Approximation of the Optimization Problem . . . . .	10
4.3	Imposition of the ZLB Constraint . . . . .	11
<b>5</b>	<b>Optimal Policy Under Perfect Information</b>	<b>12</b>
5.1	The Optimal Policy with Commitment . . . . .	12
5.1.1	Implementation . . . . .	14
5.1.2	Optimal Policy from a Timeless Perspective . . . . .	14
5.2	The Dynamic Programming Discretionary Policy . . . . .	14
5.3	Optimized Simple Rules . . . . .	15
5.4	The Stochastic Case . . . . .	16
<b>6</b>	<b>Optimal Policy Under Imperfect Information</b>	<b>17</b>
<b>7</b>	<b>Optimal Monetary Policy in the NK Model: Results</b>	<b>19</b>
<b>8</b>	<b>Conclusions</b>	<b>21</b>
<b>A</b>	<b>Linearization of RE Model</b>	<b>24</b>
<b>B</b>	<b>Priors and Posterior Estimates</b>	<b>25</b>

# 1 INTRODUCTION

The formal estimation of DSGE models by Bayesian methods has now become standard.<sup>1</sup> However, as Levine *et al.* (2007) first pointed out, in the standard approach there is an implicit asymmetric informational assumption that needs to be critically examined: whereas perfect information about current shocks and other macroeconomic variables is available to the economic agents, it is not to the econometricians. By contrast, in Levine *et al.* (2007) and Levine *et al.* (2010) a symmetric information assumption is adopted. This can be thought of as the informational counterpart to the “*cognitive consistency principle*” proposed in Evans and Honkapohja (2009) which holds that economic agents should be assumed to be “about as smart as, but no smarter than good economists”. The assumption that agents have no more information than the economist who constructs and estimates the model on behalf of the policymaker, amounts to what we term the *informational consistency principle (ICP)*. Certainly the ICP seems plausible and in fact Levine *et al.* (2010) shows that this informational assumption improves the empirical performance of a standard NK model.<sup>2</sup>

The focus of our paper here is on the implications of imperfect information (II) for optimal monetary policy in a model estimated assuming the ICP. The questions we pose are first, what are the welfare costs associated with the private sector possesses only II of the state variables; second, what are the implications of II for the gains from commitment; third, how does II affect the form of optimized Taylor rules and finally how do interest rate zero-lower-bound considerations impact on optimal policy with II.

A sizeable literature now exists on this subject - a by no means exhaustive selection of contributions include: Cukierman and Meltzer (1986), Pearlman (1992), Svensson and Woodford (2001), Svensson and Woodford (2003), Faust and Svensson (2001), Faust and Svensson (2002) Aoki (2003), Aoki (2006) and and (Melecky *et al.* (2008)).<sup>3</sup> However, as far as we are aware, it is the first paper to study the latter in a estimated DSGE model with informational consistency at both the estimation and policy design stages of the exercise.

The rest of the paper is organized as follows. Section 2 describes the standard NK model used for the policy analysis and section 3 describes the estimation by Bayesian methods drawing upon Levine *et al.* (2010). Section 4 sets out the general framework for calculating optimal policy. Section 5 assumes perfect information for both the private sector and the policymaker setting out solution procedures for optimal policy, first assuming an ability to commit, second assuming no commitment mechanism is available and the central bank exercises discretion and third, assuming policy conducted in the form of a simple interest rate, Taylor-type rule. A novel feature of treatment is the consideration the zero

---

<sup>1</sup>See Fernandez-Villaverde (2009) for a comprehensive and accessible review.

<sup>2</sup>The possibility that imperfect information in NK models improves the empirical fit has also been examined by Collard and Dellas (2004), Collard and Dellas (2006), Collard *et al.* (2009), although an earlier assessment of the effects of imperfect information for an IS-LM model dates back to Minford and Peel (1983)

<sup>3</sup>Section provides a taxonomy of the various assumed information structures assumed in these papers.

lower bound in the design of policy rules. In section 5 we assume perfect information - both the central bank and the private sector (but not the econometrician) observe the full state vector describing the model model dynamics. Section 6 relaxes this assumption and considers rules that correspond to the ICP adopted at the estimation side. Section 7 provides an application to our estimated DSGE model. Section 8 concludes.

## 2 THE MODEL

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

The full model in non-linear form is as follows

$$1 = \beta(1 + R_t)E_t \left[ \frac{MU_{t+1}^C}{MU_t^C \Pi_{t+1}} \right] \quad (1)$$

$$\frac{W_t}{P_t} = -\frac{1}{(1 - \frac{1}{\eta})} \frac{MU_t^L}{MU_t^C} \quad (2)$$

$$MC_t = \frac{W_t}{A_t P_t} \quad (3)$$

$$H_t - \xi \beta E_t [\tilde{\Pi}_{t+1}^{\zeta-1} H_{t+1}] = Y_t MU_t^C \quad (4)$$

$$J_t - \xi \beta E_t [\tilde{\Pi}_{t+1}^{\zeta} J_{t+1}] = \frac{1}{1 - \frac{1}{\zeta}} MC_t MS_t Y_t MU_t^C \quad (5)$$

$$Y_t = \frac{A_t L_t}{\Delta_t} \text{ where } \Delta_t \equiv \frac{1}{n} \sum_{j=1}^n (P_t(j)/P_t)^{-\zeta} \quad (6)$$

$$1 = \xi \tilde{\Pi}_t^{\zeta-1} + (1 - \xi) \left( \frac{J_t}{H_t} \right)^{1-\zeta} \text{ where } \tilde{\Pi}_t \equiv \frac{\Pi_t}{\Pi_{t-1}^\gamma} \quad (7)$$

$$Y_t = C_t + G_t \quad (8)$$

Equation (1) is the familiar Euler equation with  $\beta$  the discount factor,  $1 + R_t$  the gross nominal interest rate,  $MU_t^C$  the marginal utility of consumption and  $\Pi \equiv \frac{P_t}{P_{t-1}}$  the gross inflation rate, with  $P_t$  the price level. The operator  $E_t[\cdot]$  denotes rational expectations conditional upon a general information set (see section 4). In (2) the real wage,  $\frac{W_t}{P_t}$  is a mark-up on the marginal rate of substitution between leisure and consumption.  $MU_t^L$  is the marginal utility of labour supply  $L_t$ . Equation (3) defines the marginal cost. Equations

(4) to (7) describe Calvo pricing with  $1 - \xi$  equal to the probability of a monopolistically competitive firm re-optimizing its price, indexing by an amount  $\gamma$  with an exogenous mark-up shock  $MS_t$ . They are derived from the optimal price-setting first-order condition for a firm  $j$  setting a new optimized price  $P_t^0(j)$  given by

$$P_t^0(j)E_t \left[ \sum_{k=0}^{\infty} \xi^k D_{t,t+k} Y_{t+k}(j) \left( \frac{P_{t+k-1}}{P_{t-1}} \right)^\gamma \right] = \frac{\kappa}{(1 - 1/\zeta)} E_t \left[ \sum_{k=0}^{\infty} \xi^k D_{t,t+k} P_{t+k} MC_{t+k} Y_{t+k}(j) \right] \quad (9)$$

where the stochastic discount factor  $D_{t,t+k} = \beta^k \frac{MU_{t+k}^C/P_{t+k}}{MU_t^C/P_t}$ ,  $MS_t$  is a mark-up shock common to all firms and demand for firm  $j$ 's output,  $Y_{t+k}(j)$ , is given by

$$Y_{t+k}(j) = \left( \frac{P_t^0(j)}{P_{t+k}} \right)^{-\zeta} Y_{t+k} \quad (10)$$

All of these nonlinear equations depend in part on expectations of future variables. How these expectations are formed depends on individual agents, and these may be rational or adaptive, which are the possibilities that we consider here, or may be formed on the basis of least squares learning.

In equilibrium all firms that have the chance to reset prices choose the same price  $P_t^0(j) = P_t^0$  and  $\frac{P_t^0}{P_t} = \frac{J_t}{H_t}$  is the real optimized price in (9).

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

$$\Delta_t = \xi \tilde{\Pi}_t^\zeta \Delta_{t-1} + (1 - \xi) \left( \frac{J_t}{H_t} \right)^{-\zeta} \quad (11)$$

Finally (8), where  $C_t$  denotes consumption, describes output equilibrium, with an exogenous government spending demand shock  $G_t$ . To close the model we assume a current inflation based Taylor-type interest-rule

$$\log(1 + R_t) = \rho_r \log(1 + R_{n,t-1}) + (1 - \rho_r) \left( \theta_\pi \log \frac{\Pi_t}{\Pi} + \log\left(\frac{1}{\beta}\right) + \theta_y \log \frac{Y_t}{Y} \right) + \epsilon_{e,t} \quad (12)$$

$$\log(1 + R_{n,t}) = \rho_r \log(1 + R_{n,t-1}) + (1 - \rho_r) \theta E_t \left[ \log \frac{\Pi_{t+1}}{\Pi_{tar,t+1}} \right] + (1 - \rho_r) \log\left(\frac{1}{\beta}\right) + \epsilon_{e,t} \quad (13)$$

$$\log \frac{\Pi_{tar,t+1}}{\Pi} = \rho_\pi \log \frac{\Pi_{tar,t}}{\Pi} + \epsilon_{\pi,t+1} \quad (14)$$

where  $\Pi_{tar,t}$  is a time-varying inflation target following an AR(1) process, (14), and  $\epsilon_{e,t}$  is a monetary policy shock.<sup>4</sup> The following form of the single period utility for household  $r$  is a

<sup>4</sup>Note the Taylor rule feeds back on output relative to its steady state rather than the output gap so we

non-separable function of consumption and labour effort that is consistent with a balanced growth steady state:

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

where  $h_C C_{t-1}$  is external habit. In equilibrium  $C_t(r) = C_t$  and marginal utilities  $MU_t^C$  and  $MU_t^L$  are obtained by differentiation:

$$MU_t^C = (1 - \varrho)(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)-1} (1 - L_t)^\varrho (1-\sigma) \quad (16)$$

$$MU_t^L = -(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)} \varrho (1 - L_t)^{\varrho(1-\sigma)-1} \quad (17)$$

Shocks  $A_t = Ae^{a_t}$ ,  $G_t = Ge^{g_t}$ ,  $\Pi_{tar,t}$  are assumed to follow log-normal AR(1) processes, where  $A, G$  denote the non-stochastic balanced growth values or paths of the variables  $A_t, G_t$ . Following Smets and Wouters (2007) and others in the literature, we decompose the price mark-up shock into persistent and transient components:  $MS_t = MS_{per} e^{msper_t} MS_{tra} e^{\varepsilon_{mstra,t}}$  where  $msper_t$  is an AR(1) process, which results in  $MS_t$  being an ARMA(1,1) process. We can normalize  $A = 1$  and put  $MS = MS_{per} = MS_{tra} = 1$  in the steady state. The innovations are assumed to have zero contemporaneous correlation. This completes the model. The equilibrium is described by 14 equations, (1)–(8), (12) and the expressions for  $MU_t^C$  and  $MU_t^L$ , defining 13 endogenous variables  $\Pi_t, \tilde{\Pi}_t, C_t, Y_t, \Delta_t, R_t, MC_t, MU_t^C, U_t, MU_t^L, L_t, H_t, J_t$  and  $\frac{W_t}{P_t}$ . There are 6 shocks in the system:  $A_t, G_t, MS_{per,t}, MS_{tra,t}, \Pi_{tar,t}$  and  $\varepsilon_{e,t}$ .

Bayesian estimation is based on the rational expectations solution of the log-linear model.<sup>5</sup> The conventional approach assumes that the private sector has perfect information of the entire state vector  $mu_t^C, \pi_t, \pi_{t-1}, c_{t-1}$ , and, crucially, current shocks  $msper_t, ms_t, a_t$ . These are extreme information assumptions and exceed the data observations on three data sets  $y_t, \pi_t$  and  $r_t$  that we subsequently use to estimate the model. If the private sector can only observe these data series (we refer to this as symmetric information) we must turn from a solution under *perfect* information on the part of the private sector (later referred to as *asymmetric information* – *AI* since the private sector’s information set exceeds that of the econometrician) to one under *imperfect information* – *II*.

### 3 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)

---

avoid making excessive informational demands on the central bank when implementing this rule.

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

models (Doan *et al.* (1984)), and, during the 1980s, the extensive development and application of Kalman filtering-based state space systems methods in statistics and economics (Aoki (1987), Harvey (1989)).

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

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

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

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

### 3.1 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. As a robustness check we also ran estimations using an output series detrending output with

a linear-quadratic trend. 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 1981:1 to 2006:4.

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

### 3.2 ESTIMATION RESULTS

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

First it is interesting to note that the parameter estimates are fairly consistent across the information assumptions despite the fact that these alternatives lead to a considerably better model fit based on the corresponding posterior marginal data densities. 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  $\gamma$  imply that inflation is intrinsically not very persistent in the relevant model specifications. The posterior mean estimates for the Calvo price-setting parameter,  $\xi$ , obtained from Model GH across all the information sets imply an average price contract duration of about 3 – 4 (quarters compared with the prior of 2 quarters) similar to the findings of Christiano *et al.* (2005), Levin *et al.* (2006) and Smets and Wouters (2007). The external habit parameter is estimated to be around 90% 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.

In Table 1 we report the posterior marginal data density from the estimation which is computed using the Geweke (1999) modified harmonic-mean estimator. The marginal data density can be interpreted as maximum log-likelihood values, penalized for the model dimensionality, and adjusted for the effect of the prior distribution (Chang *et al.* (2002)). Appendix E compares these results obtained with linear trend with those where output is detrended using a linear-quadratic trend. In fact the results change very little, so we continue to use linear detrending. Whichever model variant has the highest marginal data density attains the best relative model fit. The values for imperfect information with measurement error are virtually identical to those without measurement error, so we have excluded them from the table.

Model	AI	II	IIME
H	-92.85	-90.90	-92.18
G	-103.77	-102.03	-99.79
GH	-96.95	-96.62	-94.74
Z	-99.48	-96.48	-97.14

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

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

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

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

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

where  $p_i(\theta|m_i)$  is the prior density for model  $m_i$ , and  $L(y|m_i)$  is the data density for model  $m_i$  given parameter vector  $\theta$ . 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):

$$PO_{i,j} = \frac{p(m_i|y)}{p(m_j|y)} = \frac{L(y|m_i)p(m_i)}{L(y|m_j)p(m_j)} \quad (18)$$

$$BF_{i,j} = \frac{L(y|m_i)}{L(y|m_j)} = \frac{\exp(LL(y|m_i))}{\exp(LL(y|m_j))} \quad (19)$$

in terms of the log-likelihoods. Components (18) and (19) provide a framework for comparing alternative and potentially misspecified models based on their marginal likelihood. Such comparisons are important in the assessment of rival models.

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

Pr(H, II)=0.688
Pr(H, IIME)=0.1913
Pr(H, AI)=0.0979
Pr(GH, IIME)=0.0148
Pr(Z, II)=0.0026
Pr(GH, II)=0.0023
Pr(GH, AI)=0.0016
Pr(Z, IIME)=0.0013
Remaining prob. are almost zero

TABLE 2: Model Probabilities Across Model Variants and Information Sets

Tables 1 and 2 reveal that a combination of Model H and with information set II outperforms the same with information set AI by a Bayes factor of approximately 7. For all models  $II \succ AI$  in terms of LL. This is a striking result; although informational consistency in intuitively appealing there is no inevitability that models that assume this will perform better in LL terms than the traditional assumption of AI. By the same token introducing measurement error into the private sector's observations (information set IIME) is not bound to improve performance and indeed we see that the IIME case does not uniformly improve LL performance except for models G and GH where we do see  $IIME \succ II \succ AI$ .

Our model comparison analysis contains two other important results. First, uniformly across all information sets indexation does not improve the model fit, but the existence of habit is crucial. The poor performance of indexation is in a sense encouraging as this feature of the NK is ad hoc and vulnerable to the Lucas critique. The existence of habit by contrast is a plausible formulation of utility that addresses issues examined in the happiness

literature.<sup>6</sup> Second, the II as compared with AI specification leads to significantly better fit for Model Z, and a better improvement than for the other three model variants. Model Z we recall is the model with zero persistence mechanisms. Its substantial improvement of performance on introducing II on the part of the private sector confirms our earlier analytical results that show how II introduces endogenous persistence. But where other persistence mechanisms habit and indexation exist in models H and GH these to some extent overshadow the improvement brought by II.

## 4 THE GENERAL SET-UP

This section describes the general set-up that applies irrespective of the informational assumptions.

### 4.1 STATE-SPACE REPRESENTATION AND INFORMATION SETS

The non-linear DSGE model is linearized about a deterministic balanced growth path resulting in a state-space representation

$$\begin{bmatrix} \mathbf{z}_{t+1} \\ E_t \mathbf{x}_{t+1} \end{bmatrix} = A^1 \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} + A^2 \begin{bmatrix} E_t \mathbf{z}_t \\ E_t \mathbf{x}_t \end{bmatrix} + B \mathbf{w}_t + \begin{bmatrix} \mathbf{u}_{t+1} \\ 0 \end{bmatrix} \quad (20)$$

where  $\mathbf{z}_t, \mathbf{x}_t$  are vectors of backward and forward-looking variables, respectively,  $\mathbf{w}_t$  is a vector of policy variables, and  $\mathbf{u}_t$  is a i.d. zero mean shock variable with covariance matrix  $\Sigma_u$ ; a more general setup allows for shocks to the equations involving expectations. In addition for the imperfect information case, we assume that agents all make the same observations at time  $t$ , which are given by

$$\mathbf{m}_t = M^1 \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} + M^2 \begin{bmatrix} E_t \mathbf{z}_t \\ E_t \mathbf{x}_t \end{bmatrix} + L \mathbf{w}_t + \mathbf{v}_t \quad (21)$$

where  $\mathbf{v}_t$  is a i.d. zero mean shock variable with covariance matrix  $\Sigma_v$ , representing measurement errors.

Define target variables  $\mathbf{s}_t$  by

$$\mathbf{s}_t = J \mathbf{y}_t + H \mathbf{w}_t \quad (22)$$

Then the policy-maker's loss function at time  $t$  by

$$\Omega_t = \frac{1}{2} \sum_{\tau=0}^{\infty} \beta^\tau [\mathbf{s}_{t+\tau}^T Q_1 \mathbf{s}_{t+\tau} + \mathbf{w}_{t+\tau}^T Q_2 \mathbf{w}_{t+\tau}] \quad (23)$$

---

<sup>6</sup>In particular the ‘‘Easterin paradox’’, Easterlin (2003). See also Layard (2006) and Choudhary *et al.* (2011) for the role of external habit in the explanation of the paradox.

which we can rewrite as

$$\Omega_t = \frac{1}{2} \sum_{i=0}^{\infty} \beta^i [y_{t+i}^T Q y_{t+i} + 2y_{t+i}^T U w_{t+i} + w_{t+i}^T R w_{t+i}] \quad (24)$$

where  $Q = J^T Q_1 M$ ,  $U = J^T Q_1 H$ ,  $R = Q_2 + H^T Q_1 H$ ,  $Q_1$  and  $Q_2$  are symmetric and non-negative definite,  $R$  is required to be positive definite and  $\beta \in (0, 1)$  is discount factor.

For the literature described in the introduction, rational expectations are formed assuming the following information sets:

1. For perfect information the private sector and policymaker/modeller have the following information set:

$$I_t = \{z_\tau, x_\tau\}, \tau \leq t; A^1, A^2, B, \Sigma_u, [Q, U, R, \beta] \text{ or the monetary rule}$$

2. For symmetric imperfect information (see Pearlman (1992), Svensson and Woodford (2003) and for Bayesian estimation, Levine *et al.* (2010)):

$$I_t = \{m_\tau\}, \tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v, [Q, U, R, \beta] \text{ or the monetary rule.}$$

3. For the first category of asymmetric imperfect information (see Svensson and Woodford (2001), Aoki (2003), Aoki (2006) and standard Bayesian estimation):

$$I_t^{ps} = I_t = \{z_\tau, x_\tau\}, \tau \leq t; A^1, A^2, B, \Sigma_u, [Q, U, R, \beta] \text{ or the monetary rule for the private sector and}$$

$$I_t^{pol} = \{m_\tau\}, \tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v, [Q, U, R, \beta] \text{ or the monetary rule for the policymaker.}$$

4. For the second category of asymmetric imperfect information (see Cukierman and Meltzer (1986), Faust and Svensson (2001), Faust and Svensson (2002)) and (Melecky *et al.* (2008)):

$$I_t^{pol} = \{m_\tau\}, \tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v, [Q, U, R, \beta] \text{ or the monetary rule for the policymaker sector and}$$

$$I_t^{ps} = \{m_\tau\}, \tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v \text{ for the private sector.}$$

In the rest of the paper we confine ourselves to information set 1 for perfect information and information set 2 for imperfect information. Information set 3 is incompatible with the ICP. Information set 4 is however compatible and is needed to address the issue of optimal ambiguity. This is beyond the scope of this paper.

## 4.2 LQ APPROXIMATION OF THE OPTIMIZATION PROBLEM

In our models there are a number of distortions that result in the steady state output being below (or possibly above) the social optimum. We cannot assume that these distortions are small in the steady state and use the ‘small distortions’ (Woodford (2003)). Instead, our computations use the large distortions approximation to this welfare function as described in Levine *et al.* (2008a).

To work out the welfare in terms of a consumption equivalent percentage increase, expanding  $U(C, L)$  as a Taylor series, a 1% permanent increase in consumption of 1 per cent yields a first-order welfare increase  $U_C C \times 0.01$ . Since standard deviations are expressed in terms of percentages, the welfare loss terms which are proportional to the covariance matrix (and pre-multiplied by 1/2) are of order  $10^{-4}$ . The losses reported in the paper in the subsections that follow are scaled by a factor  $1 - \beta$ . Letting  $\Delta\Omega$  be these losses relative to the optimal policy, then  $c_e = \Delta\Omega \times 0.01\%$ .

### 4.3 IMPOSITION OF THE ZLB CONSTRAINT

We can modify welfare criterion so as to approximately impose an interest rate zero lower bound (ZLB) so that this event hardly ever occurs. Let  $L_t$  be our quadratic approximation to the single-period loss function. As in Woodford (2003), chapter 6, the ZLB constraint is implemented by modifying the single period welfare loss to  $L_t + w_r r_{n,t}^2$ . Then following Levine *et al.* (2008b), the policymaker's optimization problem is to choose  $w_r$  and the unconditional distribution for  $r_{n,t}$  (characterized by the steady state variance) shifted to the right about a new non-zero steady state inflation rate and a higher nominal interest rate, such that the probability,  $p$ , of the interest rate hitting the lower bound is very low.<sup>7</sup> This is implemented by calibrating the weight  $w_r$  for each of our policy rules so that  $z_0(p)\sigma_r < R_n$  where  $z_0(p)$  is the critical value of a standard normally distributed variable  $Z$  such that  $\text{prob}(Z \leq z_0) = p$ ,  $R_n = \frac{1}{\beta} - 1 + \pi^* \equiv R_n(\pi^*)$  is the steady state nominal interest rate (assuming zero growth),  $\sigma_r^2 = \text{var}(r_n)$  is the unconditional variance and  $\pi^*$  is the new steady state inflation rate. Given  $\sigma_r$  the steady state positive inflation rate that will ensure  $r_{n,t} \geq 0$  with probability  $1 - p$  is given by<sup>8</sup>

$$\pi^* = \max[z_0(p)\sigma_r - R_n(0) \times 100, 0] \quad (25)$$

In our linear-quadratic framework we can write the intertemporal expected welfare loss at time  $t = 0$  as the sum of stochastic and deterministic components,  $\Omega_0 = \tilde{\Omega}_0 + \bar{\Omega}_0$ . Note that  $\bar{\Omega}_0$  incorporates in principle the new steady state values of all the variables; however the NK Phillips curve being almost vertical, the main extra term comes from the quadratic inflation terms in the loss function. By increasing  $w_r$  we can lower  $\sigma_r$  thereby decreasing  $\pi^*$  and reducing the deterministic component, but at the expense of increasing the stochastic component of the welfare loss. By exploiting this trade-off, we then arrive at the optimal policy that, in the vicinity of the steady state, imposes the ZLB constraint,  $r_t \geq 0$  with probability  $1 - p$ .

<sup>7</sup>The idea that the ZLB should be avoided by choosing a long-run inflation rate so as to increase the corresponding long-run interest rate and make room for an active interest rate rule at all times has been put forward recently by Blanchard *et al.* (2010).

<sup>8</sup>If the inefficiency of the steady-state output is negligible, then  $\pi^* \geq 0$  is a credible new steady state inflation rate. Note that in our LQ framework, the zero interest rate bound is very occasionally hit in which case the interest rate is allowed to become negative.

## 5 OPTIMAL POLICY UNDER PERFECT INFORMATION

Under perfect information,  $\begin{bmatrix} E_t z_t \\ E_t x_t \end{bmatrix} = \begin{bmatrix} z_t \\ x_t \end{bmatrix}$ . Let  $A \equiv A^1 + A^2$  and first consider the purely deterministic problem with a model then in state-space form:

$$\begin{bmatrix} z_{t+1} \\ x_{t+1,t}^e \end{bmatrix} = A \begin{bmatrix} z_t \\ x_t \end{bmatrix} + B w_t \quad (26)$$

where  $z_t$  is an  $(n - m) \times 1$  vector of predetermined variables including non-stationary processed,  $z_0$  is given,  $w_t$  is a vector of policy variables,  $x_t$  is an  $m \times 1$  vector of non-predetermined variables and  $x_{t+1,t}^e$  denotes rational (model consistent) expectations of  $x_{t+1}$  formed at time  $t$ . Then  $x_{t+1,t}^e = x_{t+1}$  and letting  $y_t^T = [z_t^T \ x_t^T]$  (26) becomes

$$y_{t+1} = A y_t + B w_t \quad (27)$$

The procedures for evaluating the three policy rules are outlined in the rest of this appendix (or Currie and Levine (1993) for a more detailed treatment).

### 5.1 THE OPTIMAL POLICY WITH COMMITMENT

Consider the policy-maker's *ex-ante* optimal policy at  $t = 0$ . This is found by minimizing  $\Omega_0$  given by (24) subject to (27) and (22) and given  $z_0$ . We proceed by defining the Hamiltonian

$$\mathcal{H}_t(y_t, y_{t+1}, \mu_{t+1}) = \frac{1}{2} \beta^t (y_t^T Q y_t + 2 y_t^T U w_t + w_t^T R w_t) + \mu_{t+1} (A y_t + B w_t - y_{t+1}) \quad (28)$$

where  $\mu_t$  is a row vector of costate variables. By standard Lagrange multiplier theory we minimize

$$\mathcal{L}_0(y_0, y_1, \dots, w_0, w_1, \dots, \mu_1, \mu_2, \dots) = \sum_{t=0}^{\infty} \mathcal{H}_t \quad (29)$$

with respect to the arguments of  $L_0$  (except  $z_0$  which is given). Then at the optimum,  $\mathcal{L}_0 = \Omega_0$ .

Redefining a new costate column vector  $\mathbf{p}_t = \beta^{-t} \mu_t^T$ , the first-order conditions lead to

$$w_t = -R^{-1}(\beta B^T \mathbf{p}_{t+1} + U^T y_t) \quad (30)$$

$$\beta A^T \mathbf{p}_{t+1} - \mathbf{p}_t = -(Q y_t + U w_t) \quad (31)$$

Substituting (30) into (27)) we arrive at the following system under control

$$\begin{bmatrix} I & \beta B R^{-1} B^T \\ 0 & \beta(A^T - U R^{-1} U^T) \end{bmatrix} \begin{bmatrix} y_{t+1} \\ \mathbf{p}_{t+1} \end{bmatrix} = \begin{bmatrix} A - B R^{-1} U^T & 0 \\ -(Q - U R^{-1} U^T) & I \end{bmatrix} \begin{bmatrix} y_t \\ \mathbf{p}_t \end{bmatrix} \quad (32)$$

To complete the solution we require  $2n$  boundary conditions for (32). Specifying  $z_0$

gives us  $n - m$  of these conditions. The remaining condition is the ‘transversality condition’

$$\lim_{t \rightarrow \infty} \mu_t^T = \lim_{t \rightarrow \infty} \beta^t \mathbf{p}_t = 0 \quad (33)$$

and the initial condition

$$\mathbf{p}_{20} = 0 \quad (34)$$

where  $\mathbf{p}_t^T = [\mathbf{p}_{1t}^T \mathbf{p}_{2t}^T]$  is partitioned so that  $\mathbf{p}_{1t}$  is of dimension  $(n - m) \times 1$ . Equation (22), (30), (32) together with the  $2n$  boundary conditions constitute the system under optimal control.

Solving the system under control leads to the following rule

$$\mathbf{w}_t = -F \begin{bmatrix} I & 0 \\ -N_{21} & -N_{22} \end{bmatrix} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{p}_{2t} \end{bmatrix} \equiv D \begin{bmatrix} \mathbf{z}_t \\ \mathbf{p}_{2t} \end{bmatrix} = -F \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_{2t} \end{bmatrix} \quad (35)$$

where

$$\begin{bmatrix} \mathbf{z}_{t+1} \\ \mathbf{p}_{2t+1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ S_{21} & S_{22} \end{bmatrix} G \begin{bmatrix} I & 0 \\ -N_{21} & -N_{22} \end{bmatrix} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{p}_{2t} \end{bmatrix} \equiv H \begin{bmatrix} \mathbf{z}_t \\ \mathbf{p}_{2t} \end{bmatrix} \quad (36)$$

$$N = \begin{bmatrix} S_{11} - S_{12}S_{22}^{-1}S_{21} & S_{12}S_{22}^{-1} \\ -S_{22}^{-1}S_{21} & S_{22}^{-1} \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \quad (37)$$

$$\mathbf{x}_t = - \begin{bmatrix} N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{p}_{2t} \end{bmatrix} \quad (38)$$

where  $F = -(R + B^T S B)^{-1}(B^T S A + U^T)$ ,  $G = A - B F$  and

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \quad (39)$$

partitioned so that  $S_{11}$  is  $(n - m) \times (n - m)$  and  $S_{22}$  is  $m \times m$  is the solution to the steady-state Riccati equation

$$S = Q - U F - F^T U^T + F^T R F + \beta(A - B F)^T S (A - B F) \quad (40)$$

The cost-to-go for the optimal policy (OP) at time  $t$  is

$$\Omega_t^{OP} = -\frac{1}{2}(\text{tr}(N_{11}Z_t) + \text{tr}(N_{22}\mathbf{p}_{2t}\mathbf{p}_{2t}^T)) \quad (41)$$

where  $Z_t = \mathbf{z}_t \mathbf{z}_t^T$ . To achieve optimality the policy-maker sets  $\mathbf{p}_{20} = 0$  at time  $t = 0$ . At time  $t > 0$  there exists a gain from renegeing by resetting  $\mathbf{p}_{2t} = 0$ . It can be shown that  $N_{11} < 0$  and  $N_{22} < 0$ .<sup>9</sup>, so the incentive to renege exists at all points along the trajectory of the optimal policy. This is the time-inconsistency problem.

<sup>9</sup>See Currie and Levine (1993), chapter 5.

### 5.1.1 IMPLEMENTATION

The rule may also be expressed in two other forms: First as

$$\mathbf{w}_t = D_1 \mathbf{z}_t + D_2 H_{21} \sum_{\tau=1}^t (H_{22})^{\tau-1} \mathbf{z}_{t-\tau} \quad (42)$$

where  $D = [D_1 \ D_2]$  is partitioned conformably with  $\mathbf{z}_t$  and  $\mathbf{p}_{2t}$ . The rule then consists of a feedback on the lagged predetermined variables with geometrically declining weights with lags extending back to time  $t = 0$ , the time of the formulation and announcement of the policy.

The final way of expressing the rule is express the process for  $\mathbf{w}_t$  in terms of the target variables only,  $\mathbf{s}_t$ , in the loss function. This in particular eliminates feedback from the exogenous processes in the vector  $\mathbf{z}_t$ . Since the rule does not require knowledge of these processes to design, Woodford (2003) refers to this as “robust” in describing it as the *Robust Optimal Explicit* rule.

### 5.1.2 OPTIMAL POLICY FROM A TIMELESS PERSPECTIVE

Noting from (38) that for the optimal policy we have  $\mathbf{x}_t = -N_{21}\mathbf{z}_t - N_{22}\mathbf{p}_{2t}$ , the optimal policy “from a timeless perspective” proposed by Woodford (2003) replaces the initial condition for optimality  $p_{20} = 0$  with

$$J\mathbf{x}_0 = -N_{21}\mathbf{z}_0 - N_{22}\mathbf{p}_{20} \quad (43)$$

where  $J$  is some  $1 \times m$  matrix. Typically in New Keynesian models the particular choice of condition is  $\pi_0 = 0$  thus avoiding any once-and-for-all initial surprise inflation. This initial condition applies only at  $t = 0$  and only affects the deterministic component of policy and not the stochastic, stabilization component.

## 5.2 THE DYNAMIC PROGRAMMING DISCRETIONARY POLICY

To evaluate the discretionary (time-consistent) policy we rewrite the cost-to-go  $\Omega_t$  given by (24) as

$$\Omega_t = \frac{1}{2} [y_t^T Q y_t + 2y_t^T U w_t + w_t^T R w_t + \beta \Omega_{t+1}] \quad (44)$$

The dynamic programming solution then seeks a stationary solution of the form  $w_t = -Fz_t$  in which  $\Omega_t$  is minimized at time  $t$  subject to (1) in the knowledge that a similar procedure will be used to minimize  $\Omega_{t+1}$  at time  $t + 1$ .

Suppose that the policy-maker at time  $t$  expects a private-sector response from  $t + 1$  onwards, determined by subsequent re-optimization, of the form

$$\mathbf{x}_{t+\tau} = -N_{t+1}\mathbf{z}_{t+\tau}, \quad \tau \geq 1 \quad (45)$$

The loss at time  $t$  for the *ex ante* optimal policy was from (41) found to be a quadratic function of  $\mathbf{x}_t$  and  $\mathbf{p}_{2t}$ . We have seen that the inclusion of  $\mathbf{p}_{2t}$  was the source of the time inconsistency in that case. We therefore seek a lower-order controller

$$\mathbf{w}_t = -F\mathbf{z}_t \quad (46)$$

with the cost-to-go quadratic in  $\mathbf{z}_t$  only. We then write  $\Omega_{t+1} = \frac{1}{2}\mathbf{z}_{t+1}^T S_{t+1} \mathbf{z}_{t+1}$  in (44). This leads to the following iterative process for  $F_t$

$$\mathbf{w}_t = -F_t \mathbf{z}_t \quad (47)$$

where

$$\begin{aligned} F_t &= (\bar{R}_t + \lambda \bar{B}_t^T S_{t+1} \bar{B}_t)^{-1} (\bar{U}_t^T + \beta \bar{B}_t^T S_{t+1} \bar{A}_t) \\ \bar{R}_t &= R + K_t^T Q_{22} K_t + U^{2T} K_t + K_t^T U^2 \\ K_t &= -(A_{22} + N_{t+1} A_{12})^{-1} (N_{t+1} B^1 + B^2) \\ \bar{B}_t &= B^1 + A_{12} K_t \\ \bar{U}_t &= U^1 + Q_{12} K_t + J_t^T U^2 + J_t^T Q_{22} J_t \\ \bar{J}_t &= -(A_{22} + N_{t+1} A_{12})^{-1} (N_{t+1} A_{11} + A_{12}) \end{aligned}$$

$$\begin{aligned} \bar{A}_t &= A_{11} + A_{12} J_t \\ S_t &= \bar{Q}_t - \bar{U}_t F_t - F_t^T \bar{U}^T + \bar{F}_t^T \bar{R}_t F_t + \beta (\bar{A}_t - \bar{B}_t F_t)^T S_{t+1} (\bar{A}_t - \bar{B}_t F_t) \\ \bar{Q}_t &= Q_{11} + J_t^T Q_{21} + Q_{12} J_t + J_t^T Q_{22} J_t \\ N_t &= -J_t + K_t F_t \end{aligned}$$

where  $B = \begin{bmatrix} B^1 \\ B^2 \end{bmatrix}$ ,  $U = \begin{bmatrix} U^1 \\ U^2 \end{bmatrix}$ ,  $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ , and  $Q$  similarly are partitioned conformably with the predetermined and non-predetermined components of the state vector.

The sequence above describes an iterative process for  $F_t$ ,  $N_t$ , and  $S_t$  starting with some initial values for  $N_t$  and  $S_t$ . If the process converges to stationary values,  $F$ ,  $N$  and  $S$  say, then the time-consistent feedback rule is  $w_t = -Fz_t$  with loss at time  $t$  given by

$$\Omega_t^{TC} = \frac{1}{2} \mathbf{z}_t^T S \mathbf{z}_t = \frac{1}{2} \text{tr}(S Z_t) \quad (48)$$

### 5.3 OPTIMIZED SIMPLE RULES

We now consider simple sub-optimal rules of the form

$$\mathbf{w}_t = D y_t = D \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} \quad (49)$$

where  $D$  is constrained to be sparse in some specified way. Rule (49) can be quite general. By augmenting the state vector in an appropriate way it can represent a PID (proportional-integral-derivative) controller.

Substituting (49) into (24) gives

$$\Omega_t = \frac{1}{2} \sum_{i=0}^{\infty} \beta^i y_{t+i}^T P_{t+i} y_{t+i} \quad (50)$$

where  $P = Q + UD + D^T U^T + D^T R D$ . The system under control (26), with  $w_t$  given by (49), has a rational expectations solution with  $x_t = -N z_t$  where  $N = N(D)$ . Hence

$$y_t^T P y_t = z_t^T T z_t \quad (51)$$

where  $T = P_{11} - N^T P_{21} - P_{12} N + N^T P_{22} N$ ,  $P$  is partitioned as for  $S$  in (39) onwards and

$$z_{t+1} = (G_{11} - G_{12} N) z_t \quad (52)$$

where  $G = A + BD$  is partitioned as for  $P$ . Solving (52) we have

$$z_t = (G_{11} - G_{12} N)^t z_0 \quad (53)$$

Hence from (54), (51) and (53) we may write at time  $t$

$$\Omega_t^{SIM} = \frac{1}{2} z_t^T V z_t = \frac{1}{2} \text{tr}(V Z_t) \quad (54)$$

where  $Z_t = z_t z_t^T$  and  $V$  satisfies the *Lyapunov* equation

$$V = T + H^T V H \quad (55)$$

where  $H = G_{11} - G_{12} N$ . At time  $t = 0$  the optimized simple rule is then found by minimizing  $\Omega_0$  given by (54) with respect to the non-zero elements of  $D$  given  $z_0$  using a standard numerical technique. An important feature of the result is that unlike the previous solution the optimal value of  $D$ ,  $D^*$  say, is not independent of  $z_0$ . That is to say

$$D^* = D^*(z_0)$$

#### 5.4 THE STOCHASTIC CASE

Consider the stochastic generalization of (26)

$$\begin{bmatrix} z_{t+1} \\ x_{t+1,t}^e \end{bmatrix} = A \begin{bmatrix} z_t \\ x_t \end{bmatrix} + B w_t + \begin{bmatrix} u_t \\ 0 \end{bmatrix} \quad (56)$$

where  $\mathbf{u}_t$  is an  $n \times 1$  vector of white noise disturbances independently distributed with  $\text{cov}(\mathbf{u}_t) = \Sigma$ . Then, it can be shown that certainty equivalence applies to all the policy rules apart from the simple rules (see Currie and Levine (1993)). The expected loss at time  $t$  is as before with quadratic terms of the form  $\mathbf{z}_t^T X \mathbf{z}_t = \text{tr}(X \mathbf{z}_t, \mathbf{z}_t^T)$  replaced with

$$E_t \left( \text{tr} \left[ X \left( \mathbf{z}_t \mathbf{z}_t^T + \sum_{i=1}^{\infty} \beta^i \mathbf{u}_{t+i} \mathbf{u}_{t+i}^T \right) \right] \right) = \text{tr} \left[ X \left( \mathbf{z}_t^T \mathbf{z}_t + \frac{\lambda}{1-\lambda} \Sigma \right) \right] \quad (57)$$

where  $E_t$  is the expectations operator with expectations formed at time  $t$ .

Thus for the optimal policy with commitment (41) becomes in the stochastic case

$$\Omega_t^{OP} = -\frac{1}{2} \text{tr} \left( N_{11} \left( Z_t + \frac{\beta}{1-\beta} \Sigma \right) + N_{22} \mathbf{p}_{2t} \mathbf{p}_{2t}^T \right) \quad (58)$$

For the time-consistent policy (48) becomes

$$\Omega_t^{TC} = -\frac{1}{2} \text{tr} \left( S \left( Z_t + \frac{\beta}{1-\beta} \Sigma \right) \right) \quad (59)$$

and for the simple rule, generalizing (54)

$$\Omega_t^{SIM} = -\frac{1}{2} \text{tr} \left( V \left( Z_t + \frac{\beta}{1-\beta} \Sigma \right) \right) \quad (60)$$

The optimized simple rule is found at time  $t = 0$  by minimizing  $\Omega_0^{SIM}$  given by (60). Now we find that

$$D^* = D^* \left( \mathbf{z}_0 \mathbf{z}_0^T + \frac{\beta}{1-\beta} \Sigma \right) \quad (61)$$

or, in other words, the optimized rule depends both on the initial displacement  $\mathbf{z}_0$  and on the covariance matrix of disturbances  $\Sigma$ .

## 6 OPTIMAL POLICY UNDER IMPERFECT INFORMATION

Here we assume that there is a set of measurements as described above. Pearlman (1992) shows that the estimate for  $z_t$  at time  $t$ , denoted by  $z_{t,t}$  is given in terms of its estimate in the previous period  $z_{t,t-1}$  via the updating equation

$$z_{t,t} = z_{t,t-1} + PD^T (EPD^T + V)^{-1} (m_t - Ez_{t,t-1}) \quad (62)$$

where  $D = M_1^1 - M_2^1 (A_{22}^1)^{-1} A_{21}^1$ ,  $E = M_1^1 + M_1^2 - (M_2^1 + M_2^2) N$ ,  $N$  represents the saddlepath relationship between  $x_{t,t-1}$  and  $z_{t,t-1}$  and  $P$  is the solution of the Riccati equation

$$P = APA^T - APD^T (DPD^T + V)^{-1} DPA^T + \Sigma \quad (63)$$

where  $A = A_{11}^1 - A_{12}^1(A_{22}^1)^{-1}A_{21}^1$ .  $z_{t,t}$  can also be written as

$$z_{t,t} = z_{t,t-1} + PD^T(DPD^T + V)^{-1}(D(z_t - Ez_{t,t-1}) + v_t) \quad (64)$$

and one can also show that  $z_t - z_{t,t}$  and  $z_{t,t}$  are orthogonal in expectations. Note that this relationship is independent of policy. We may then write the expected utility as

$$\frac{1}{2}E_t \left[ \sum_{i=0}^{\infty} \beta^i (y_{t+\tau,t+\tau}^T Q y_{t+\tau,t+\tau} + 2y_{t+\tau,t+\tau}^T U w_{t+\tau} + w_{t+\tau}^T R w_{t+\tau} + (y_{t+\tau} - y_{t+\tau,t+\tau})^T Q (y_{t+\tau} - y_{t+\tau,t+\tau})) \right] \quad (65)$$

where we note that  $w_{t+\tau}$  is dependent only on current and past  $y_{t+s,t+s}$ . This is minimized subject to the dynamics

$$\begin{bmatrix} z_{t+1,t+1} \\ E_t x_{t+1,t+1} \end{bmatrix} = (A^1 + A^2) \begin{bmatrix} z_{t,t} \\ x_{t,t} \end{bmatrix} + B w_t + \begin{bmatrix} z_{t+1,t+1} - z_{t+1,t} \\ 0 \end{bmatrix} \quad (66)$$

which represents the expected dynamics of the system. Note that  $cov(z_{t+1,t+1} - z_{t+1,t}) = PD^T(DPD^T + V)^{-1}DP$  and  $cov(z_{t+1,t+1} - z_{t+1,t+1}) = P - PD^T(DPD^T + V)^{-1}DP$ .

Taking time- $t$  expectations of the equation involving  $E_t x_{t+1}$  and subtracting from the original yields:

$$0 = A_{12}^1(z_t - z_{t,t}) + A_{22}^1(x_t - x_{t,t}) \quad (67)$$

Furthermore, since Pearlman (1992) shows that certainty equivalence holds for both the fully optimal and the time consistent solution, it is straightforward to show that expected welfare for each of the regimes is given by

$$\begin{aligned} W^J = & z_{0,0}^T S^J z_{0,0} + \frac{\lambda}{1-\lambda} tr(S^J PD^T(DPD^T + V)^{-1}DP) \\ & + \frac{1}{1-\lambda} tr(Q_{11} - Q_{12}(A_{22}^1)^{-1}A_{21}^1 - A_{21}^{1T}(A_{22}^1)^{-T}Q_{21} + A_{21}^{1T}(A_{22}^1)^{-T}Q_{22}(A_{22}^1)^{-1}A_{21}^1)\bar{P} \end{aligned} \quad (68)$$

where

$$S^R = S_{11} - S_{12}S_{22}^{-1}S_{21} \quad \bar{P} = P - PD^T(DPD^T + V)^{-1}DP \quad (69)$$

- $S_{ij}$  are the partitions of  $S^R$ , the Ricatti matrix.
- $S^{NR}$  is calculated from the standard time consistent solution algorithm.
- $S^{SIM}$  is calculated from the Lyapunov equation above.

The last term of (68) corresponds to the sum of the last term in (65) after utilizing (67).

## 7 OPTIMAL MONETARY POLICY IN THE NK MODEL: RESULTS

This section sets out numerical results for optimal policy under commitment, optimal discretionary (or time consistent) policy and for a optimized simple Taylor rule. The model is the estimated form of the best-fitting one, namely model H. For the first set of results we ignore ZLB considerations. The questions we pose are first, what are the welfare costs associated with the private sector possesses only imperfect information of the state variables; second, what are the implications of imperfect information for the gains from commitment and third, how does imperfect information affect the form of optimized Taylor rules.

Table 1 enables us to answer the first and second of these questions. We examine four imperfect information sets.

**Imperfect Information Set I:** This consists of the current and past values of the three data series used to estimate the model, output, inflation and the interest rate plus the inflation target. This represents the scenario in which a stochastic inflation target is announced in each period and believed.

**Imperfect Information Set II:** This consists of the current and past values of output, inflation and the interest rate but excludes the inflation target. This may be announced, but its absence from the information set can be interpreted as a lack of credibility. In this scenario the private sector must infer the target from its observations.

**Imperfect Information Set III:** As for I but output and inflation are only observed with a lag, but the current interest rate is observed.

**Imperfect Information Set IV:** As for III but excludes the inflation target.

The corresponding forms of the simple rules are

$$r_t = \rho r_{t-1} + \theta_\pi (\pi_t - \pi_t^{targ}) + \theta_y y_t \quad (70)$$

for perfect information and information sets I and II and

$$r_t = \rho r_{t-1} + \theta_\pi (\pi_{t-1} - \pi_t^{targ}) + \theta_y y_{t-1} \quad (\mathbf{Form\ A}) \quad (71)$$

or

$$r_t = \rho r_{t-1} + \theta_\pi E_t (\pi_t - \pi_t^{targ}) + \theta_y E_t y_t \quad (\mathbf{Form\ B}) \quad (72)$$

for information sets III and IV. Of course for III,  $E_t \pi_t^{targ} = \pi_t^{targ}$ .

Results are presented for an ad hoc form of the loss function

$$\Omega_0 = E_0 \left[ \frac{1}{2} \sum_{t=0}^{\infty} \beta^t [y_t^2 + b\pi_t^2] \right] \quad (73)$$

where  $b = \frac{\zeta\xi}{(1-\xi)(1-\beta\xi)\sigma^2}$ .<sup>10</sup>

From Table 3 we can see that imperfect information in the form of only observing a

<sup>10</sup>For model Z in fact this is the welfare-based loss function if we replace  $y_t$  with the output gap.

subset of the state variables (information sets I and II) and then only lagged observable variables see a steady welfare cost equivalent of a consumption equivalent loss of first 0.02% rising almost 0.04%. The gains from commitment are small as are the costs of simplicity as long as the inflation is observed (i.e., announced and believed).

Information	Information Set	Optimal	Time Consistent	Simple Rule A	Simple Rule B
Perfect	Full state vector	3.74	3.95	3.87	n. a.
Imperfect I	$I_t = [y_t, \pi_t, r_t, \pi_{tar,t}]$	3.75	3.95	3.89	n. a.
Imperfect II	$I_t = [y_t, \pi_t, r_t]$	5.66	5.86	5.80	n. a.
Imperfect III	$I_t = [y_{t-1}, \pi_{t-1}, r_t, \pi_{tar,t}]$	5.20	5.39	7.10	5.26
Imperfect IV	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	7.11	7.29	7.12	7.17

TABLE 3: Welfare Costs of Limited Information

Table 4 addresses the third question by setting out the optimized coefficients in the Taylor rule. We find that in the case of information set IV where the target is not credible and output and inflation are observed only with a lag, form B of the rule results in a Taylor with far more persistence than the others. Finally Table 5 indicates there is a ZLB problem especially for the perfect information set which need to be addressed using the approach of Levine *et al.* (2008b). Variances of the nominal interest rate and the probabilities of hitting the ZLB as described in section 4.3 indicate that there are ZLB issues for most cases. Interestingly these are lessened under II when observations of macro-economic variables are with a one-period lag. In this case optimal interest rate adjustment is more muted bringing down its volatility and the probability of reaching the ZLB.

Information	$[\rho, \theta_\pi, \theta_y]$
Perfect	0, 10, 0.08
Imperfect I	0, 10, 0.08
Imperfect II	0, 10, 0.08
Imperfect III (A)	0, 10, 0.08
Imperfect III (B)	0.8, 6.15, 0.03
Imperfect IV (A)	0, 10, 0.08
Imperfect IV (B)	0.8, 6.15, 0.03

TABLE 4: Optimized Simple Rules

Information	Information Set	Optimal	Time Cons	Simple Rule A	Simple Rule B
Perfect	Full state vector	0.918 (0.159)	0.725 (0.156)	0.270 (0.027)	n. a.
Imperfect I	$I_t = [y_t, \pi_t, r_t, \pi_{tar,t}]$	0.752 (0.125)	0.290 (0.031)	0.182 (0.01)	n. a.
Imperfect II	$I_t = [y_t, \pi_t, r_t]$	0.752 (0.125)	0.289 (0.031)	0.182 (0.01)	n. a.
Imperfect III	$I_t = [y_{t-1}, \pi_{t-1}, r_t, \pi_{tar,t}]$	0.253 (0.023)	0.041 (0.000)	0.131 (0.003)	0.082 (0.000)
Imperfect IV	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	0.252 (0.023)	0.041 (0.000)	0.131 (0.003)	0.081 (0.000)

TABLE 5: Interest Rate Variances (probability per quarter of hitting the ZLB in brackets).

## 8 CONCLUSIONS

This is the first paper to examine optimal policy with the ICP; i.e., in an estimated DSGE NK model where informational consistency is applied at both the estimation and policy stages. Preliminary results are encouraging; information assumptions have significant implications for both welfare and for the form of the simple rule. Future work will modify the optimal rules to enforce a low ZLB probability (as in Levine *et al.* (2008b)) and carry out the exercise using a welfare-based form of the loss function. Revisiting the issues raised in the context of a richer DSGE model that includes capital, sticky wages, search-match labour market frictions and financial friction will also be the subject of future research.

## REFERENCES

- Aoki, K. (2003). On the optimal monetary policy response to noisy indicators. *Journal of Monetary Economics*, **113**(3), 501 – 523.
- Aoki, K. (2006). Optimal commitment policy under noisy information. *Journal of Economic Dynamics and Control*, **30**(1), 81 – 109.
- 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.
- Blanchard, O., Giovanni, D., and Mauro, P. (2010). Rethinking Macroeconomic Policy. IMF Staff Position Note, SPN/10/03 .
- 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 Review*, **92**(5), 1498–1520.

- Choudhary, A., P. Levine, P. McAdam, P., and Welz, P. (2011). The Happiness Puzzle: Analytical Aspects of the Easterlin Paradox. *Oxford Economic Papers*. Forthcoming.
- Christiano, L., Eichenbaum, M., and Evans, C. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, **113**, 1–45.
- Collard, F. and Dellas, H. (2004). The New Keynesian Model with Imperfect Information and Learning. *mimeo, CNRS-GREMAQ*.
- Collard, F. and Dellas, H. (2006). Misperceived Money and Inflation Dynamics. *mimeo, CNRS-GREMAQ*.
- Collard, F., Dellas, H., and Smets, F. (2009). Imperfect Information and the Business Cycle. *Journal of Monetary Economics*. Forthcoming.
- Cukierman, A. and Meltzer, A. H. (1986). A theory of ambiguity, credibility and inflation under discretion and asymmetric information. *Econometrica*, **54**, 1099–1128.
- Currie, D. and Levine, P. L. (1993). *Rules, Reputation and Macroeconomic Policy Coordination*. Cambridge University Press.
- Doan, J., Litterman, R., and Sims, C. (1984). Forecasting and Conditional Projection Using Realistic Prior Distributions. *NBER*.
- Easterlin, R. A. (2003). Explaining Happiness. *Proceedings of the National Academy of Sciences of the United States*, **100**(19), 11176–11183.
- Evans, G. W. and Honkapohja, S. (2009). Learning and Macroeconomics. *Annual Review of Economics*, **1**, 421–449.
- Faust, J. and Svensson, L. (2001). Transparency and credibility: monetary policy with unobservable goals. *International Economic Review*, **42**, 369–397.
- Faust, J. and Svensson, L. (2002). The equilibrium degree of transparency and control in monetary policy. *Journal of Money, Credit and Banking*, **34**(2), 520–539.
- 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.

- Layard, R. (2006). Happiness and Public Policy: A Challenge to the Profession. *Economic Journal*, **116**(510), C24–C33.
- Levin, A., Onatski, A., Williams, J. C., and Williams, N. (2006). Monetary Policy Under Uncertainty in Micro-Founded Macroeconomic Models. in M. Gertler and K. Rogoff (eds.), NBER Macroeconomics Annual, 2005, pp 229–387 .
- Levine, P., Pearlman, J., and Perendia, G. (2007). Estimating DSGE Models under Partial Information. Department of Economics Discussion Papers 1607, Department of Economics, University of Surrey .
- Levine, P., Pearlman, J., and Piersse, R. (2008a). Linear-Quadratic Approximation, Efficiency and Target-Implementability. *Journal of Economic Dynamics and Control*, **32**, 3315–3349.
- Levine, P., McAdam, P., and Pearlman, J. (2008b). Quantifying and Sustaining Welfare Gains from Monetary Commitment. *Journal of Monetary Economics*, **55**(7), 1253–1276.
- Levine, P., Pearlman, J., Perendia, G., and Yang, B. (2010). Endogenous Persistence in an Estimated DSGE Model under Imperfect Information. Presented at the 2010 Royal Economic Society Conference, March 29 – 31, University of Surrey and Department of Economics Discussion Papers 0310, Department of Economics, University of Surrey .
- Melecky, M., Rodriguez Palenzuela, D., and Soderstrom, U. (2008). Inflation Target Transparency and the Macroeconomy. MPRA Paper No. 10545.
- Minford, A. and Peel, D. (1983). Some Implications of Partial Information Sets in Macroeconomic Models Embodying Rational Expectations. *Manchester School*, **51**, 235–249.
- Pearlman, J. G. (1992). Reputational and Non-Reputational Policies with Partial Information. *Journal of Economic Dynamics and Control*, **16**, 339–357.
- 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 variables for Optimal Policy. *Journal of Monetary Economics*, **50**(3), 691–720.
- Svensson, L. E. O. and Woodford, M. (2003). Indicator variables for Optimal Policy under Asymmetric Information. *Journal of Economic Dynamics and Control*, **28**(4), 661–680.

Woodford, M. (2003). *Foundations of a Theory of Monetary Policy*. Princeton University Press.

## A LINEARIZATION OF RE MODEL

The log-linearization<sup>11</sup> of the model about the non-stochastic steady state zero-growth<sup>12</sup>, zero-inflation is given by

$$y_t = c_y c_t + (1 - c_y) g_t \quad \text{where } c_y = \frac{C}{Y} \quad (\text{A.1})$$

$$E_t mu_{t+1}^C = mu_t^C - (r_t - E_t \pi_{t+1}) \quad (\text{A.2})$$

$$\pi_t = \frac{\beta}{1 + \beta\gamma} E_t \pi_{t+1} + \frac{\gamma}{1 + \beta\gamma} \pi_{t-1} + \frac{(1 - \beta\xi)(1 - \xi)}{(1 + \beta\gamma)\xi} (mc_t + ms_t) \quad (\text{A.3})$$

where marginal utilities,  $mu_t^C$ ,  $mu_t^L$ , and marginal costs,  $mc_t$ , and output,  $y_t$  are defined by

$$mu_t^C = \frac{(1 - \varrho)(1 - \sigma) - 1}{1 - h_C} (c_t - h_C c_{t-1}) - \frac{\varrho(1 - \sigma)L}{1 - L} l_t \quad (\text{A.4})$$

$$mu_t^L = \frac{1}{1 - h_C} (c_t - h_C c_{t-1}) + \frac{L}{1 - L} l_t + mu_t^C \quad (\text{A.5})$$

$$w_t - p_t = mu_t^L - mu_t^C \quad (\text{A.6})$$

$$mc_t = w_t - p_t - a_t \quad (\text{A.7})$$

$$y_t = a_t + l_t \quad (\text{A.8})$$

Equations (A.1) and (A.2) constitute the micro-founded ‘IS Curve’ and demand side for the model, given the monetary instrument. According to (A.2) solved forward in time, the marginal utility of consumption is the sum of all future expected real interest rates. (A.3) is the ‘NK Philips Curve’, the supply side of our model. In the absence of indexing it says that the inflation rate is the discounted sum of all future expected marginal costs. Note that price dispersion,  $\Delta_t$ , disappears up to a first order approximation and therefore does not enter the linear dynamics. Finally, shock processes and the Taylor rule are given by

$$g_{t+1} = \rho_g g_t + \epsilon_{g,t+1}$$

$$a_{t+1} = \rho_a a_t + \epsilon_{a,t+1}$$

$$msper_{t+1} = \rho_{ms} msper_t + \epsilon_{msper,t+1}$$

$$ms_t = msper_t + \epsilon_{mstra,t}$$

$$\pi_{tar,t+1} = \rho_a \pi_{tar,t} + \epsilon_{tar,t+1}$$

$$r_t = \rho_r r_{t-1} + (1 - \rho_r) \theta (E_t \pi_{t+1} - \rho_{tar} \pi_{tar,t}) + \epsilon_{e,t}$$

$\epsilon_{e,t}$ ,  $\epsilon_{a,t}$ ,  $\epsilon_{g,t}$ ,  $\epsilon_{msper,t}$ ,  $\epsilon_{mstra,t}$  and  $\epsilon_{tar,t}$  are i.i.d. with mean zero and variances  $\sigma_{\epsilon_e}^2$ ,  $\sigma_{\epsilon_a}^2$ ,  $\sigma_{\epsilon_g}^2$ ,  $\sigma_{\epsilon_{msper}}^2$ ,  $\sigma_{\epsilon_{mstra}}^2$  and  $\sigma_{\epsilon_{tra}}^2$  respectively.

<sup>11</sup>Lower case variables are defined as  $x_t = \log \frac{X_t}{\bar{X}}$ .  $r_t$  and  $\pi_t$  are log-deviations of *gross* rates. The validity of this log-linear procedure for general information sets is discussed in the next section.

<sup>12</sup>With growth we simply replace  $\beta$  and  $h_C$  with  $\beta_g \equiv \beta(1 + g)^{(1-\varrho)(1-\sigma)-1}$  and  $h_{Cg} = \frac{h_C}{1+g}$ .

## B PRIORS AND POSTERIOR ESTIMATES

Parameter	Notation	Prior distribution		
		Density	Mean	
Risk aversion	$\sigma$	Normal	1.50	0.375
Price indexation	$\gamma$	Beta	0.50	0.15
Calvo prices	$\xi$	Beta	0.50	0.10
Consumption habit formation	$h_C$	Beta	0.70	0.10
Preference parameter	$\varrho$	Beta	0.50	0.20
Adaptive expectations				
Error adjustment - $E_{f,t}\bar{q}_{t+1}^a$	$\mu_1$	Beta	0.50	0.20
Error adjustment - $E_{h,t}u_{c,t+1}^a$	$\mu_2$	Beta	0.50	0.20
Error adjustment - $E_{h,t}[\pi_{t+1}]^a$	$\mu_3$	Beta	0.50	0.20
Proportion of rational households	$\lambda_h$	Beta	0.50	0.20
Proportion of rational firms	$\lambda_f$	Beta	0.50	0.20
Interest rate rule				
Inflation	$\theta_\pi$	Normal	1.50	0.25
Output	$\theta_y$	Normal	0.125	0.05
Interest rate smoothing	$\rho_r$	Beta	0.80	0.10
AR(1) coefficient				
Technology	$\rho_a$	Beta	0.85	0.10
Government spending	$\rho_g$	Beta	0.85	0.10
Price mark-up	$\rho_{ms}$	Beta	0.50	0.20
Inflation target	$\rho_{tar}$	Beta	0.85	0.10
Standard deviation of AR(1) innovations				
Technology	$sd(\epsilon_a)$	Inv. gamma	0.40	2.00
Government spending	$sd(\epsilon_g)$	Inv. gamma	1.50	2.00
Price mark-up	$sd(\epsilon_{ms})$	Inv. gamma	0.10	2.00
Inflation target	$sd(\epsilon_{tar})$	Inv. gamma	0.10	10.00
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 6: Prior Distributions

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
$\sigma$	2.33 [1.84:2.80]	2.30 [1.82:2.80]	2.75 [1.32:3.17]	2.79 [2.35:3.21]	2.39 [1.92:2.86]	2.33 [1.84:2.81]	2.83 [1.40:3.26]	2.85 [2.40:3.29]	2.39 [1.92:2.86]	2.33 [1.84:2.81]	2.83 [1.40:3.26]	2.85 [2.40:3.29]
$\gamma$	0.26 [0.08:0.43]	-	0.27 [0.08:0.46]	-	0.22 [0.07:0.35]	-	0.23 [0.08:0.38]	-	0.22 [0.07:0.35]	-	0.23 [0.08:0.38]	-
$\xi$	0.72 [0.61:0.83]	0.76 [0.67:0.85]	0.64 [0.50:0.77]	0.68 [0.57:0.79]	0.73 [0.62:0.84]	0.75 [0.65:0.84]	0.67 [0.58:0.77]	0.69 [0.60:0.78]	0.73 [0.62:0.84]	0.75 [0.65:0.84]	0.67 [0.58:0.77]	0.69 [0.60:0.78]
$h_C$	0.90 [0.85:0.95]	0.90 [0.85:0.95]	-	-	0.89 [0.83:0.95]	0.89 [0.84:0.95]	-	-	0.89 [0.83:0.95]	0.89 [0.84:0.95]	-	-
$\varrho$	0.36 [0.07:0.62]	0.35 [0.07:0.62]	0.20 [0.03:0.36]	0.20 [0.03:0.35]	0.36 [0.08:0.64]	0.35 [0.05:0.62]	0.22 [0.03:0.38]	0.21 [0.04:0.37]	0.36 [0.08:0.64]	0.35 [0.05:0.62]	0.22 [0.03:0.38]	0.21 [0.04:0.37]
Interest rate rule												
$\theta_\pi$	1.95 [1.57:2.31]	1.90 [1.53:2.26]	2.00 [1.63:2.37]	1.99 [1.63:2.35]	1.89 [1.54:2.25]	1.87 [1.50:2.27]	1.92 [1.59:2.24]	1.93 [1.60:2.24]	1.89 [1.54:2.25]	1.87 [1.50:2.27]	1.92 [1.59:2.24]	1.93 [1.60:2.24]
$\theta_y$	0.12 [0.07:0.18]	0.12 [0.06:0.17]	0.13 [0.07:0.18]	0.14 [0.08:0.19]	0.12 [0.06:0.17]	0.12 [0.07:0.18]	0.12 [0.06:0.17]	0.13 [0.08:0.18]	0.12 [0.06:0.17]	0.12 [0.07:0.18]	0.12 [0.06:0.17]	0.13 [0.08:0.18]
$\rho_r$	0.67 [0.54:0.78]	0.63 [0.51:0.75]	0.60 [0.47:0.71]	0.58 [0.46:0.69]	0.63 [0.53:0.74]	0.60 [0.49:0.70]	0.54 [0.42:0.64]	0.51 [0.40:0.61]	0.63 [0.53:0.74]	0.60 [0.49:0.70]	0.54 [0.42:0.64]	0.51 [0.40:0.61]
AR(1) coefficient												
$\rho_a$	0.95 [0.93:0.98]	0.95 [0.93:0.98]	0.95 [0.93:0.98]	0.95 [0.93:0.98]	0.96 [0.93:0.99]	0.96 [0.93:0.99]	0.96 [0.94:0.99]	0.96 [0.94:0.99]	0.96 [0.93:0.99]	0.96 [0.93:0.99]	0.96 [0.94:0.99]	0.96 [0.94:0.99]
$\rho_g$	0.92 [0.89:0.95]	0.92 [0.89:0.95]	0.89 [0.86:0.92]	0.89 [0.86:0.92]	0.92 [0.88:0.95]	0.92 [0.89:0.95]	0.88 [0.85:0.91]	0.88 [0.85:0.91]	0.92 [0.88:0.95]	0.92 [0.89:0.95]	0.88 [0.85:0.91]	0.88 [0.85:0.91]
$\rho_{ms}$	0.39 [0.03:0.87]	0.42 [0.14:0.96]	0.39 [0.04:0.83]	0.41 [0.05:0.85]	0.39 [0.05:0.73]	0.48 [0.17:0.96]	0.38 [0.05:0.70]	0.41 [0.08:0.72]	0.39 [0.05:0.73]	0.48 [0.17:0.96]	0.38 [0.05:0.70]	0.41 [0.08:0.72]
$\rho_{targ}$	0.74 [0.54:0.93]	0.71 [0.54:0.90]	0.79 [0.58:0.96]	0.79 [0.93:0.94]	0.75 [0.57:0.90]	0.73 [0.55:0.89]	0.70 [0.48:0.91]	0.75 [0.59:0.90]	0.75 [0.57:0.90]	0.73 [0.55:0.89]	0.70 [0.48:0.91]	0.75 [0.59:0.90]
Standard deviation of AR(1) innovations												
$sd(\epsilon_a)$	0.48 [0.29:0.72]	0.48 [0.29:0.74]	0.41 [0.30:0.53]	0.41 [0.31:0.53]	0.50 [0.33:0.67]	0.47 [0.25:0.1]	0.41 [0.30:0.50]	0.41 [0.32:0.51]	0.50 [0.33:0.67]	0.47 [0.25:0.1]	0.41 [0.30:0.50]	0.41 [0.32:0.51]
$sd(\epsilon_g)$	1.79 [1.57:2.01]	1.79 [1.58:2.00]	2.32 [2.02:2.63]	2.31 [2.01:2.60]	1.80 [1.58:2.02]	1.80 [1.58:2.02]	2.23 [1.95:2.52]	2.26 [1.96:2.55]	1.80 [1.58:2.02]	1.80 [1.58:2.02]	2.23 [1.95:2.52]	2.26 [1.96:2.55]
$sd(\epsilon_{ms})$	0.05 [0.02:0.08]	0.05 [0.02:0.08]	0.05 [0.02:0.08]	0.05 [0.02:0.08]	0.05 [0.02:0.08]	0.05 [0.02:0.08]	0.05 [0.02:0.09]	0.06 [0.03:0.09]	0.05 [0.02:0.08]	0.05 [0.02:0.08]	0.05 [0.02:0.09]	0.06 [0.03:0.09]
$sd(\epsilon_{targ})$	0.18 [0.03:0.43]	0.23 [0.03:0.43]	0.13 [0.03:0.32]	0.12 [0.03:0.24]	0.18 [0.03:0.38]	0.21 [0.05:0.43]	0.22 [0.03:0.48]	0.14 [0.04:0.30]	0.18 [0.03:0.38]	0.21 [0.05:0.43]	0.22 [0.03:0.48]	0.14 [0.04:0.30]
Standard deviation of I.I.D. shocks/measurements errors												
$sd(\epsilon_m)$	0.09 [0.04:0.12]	0.07 [0.03:0.12]	0.08 [0.04:0.12]	0.08 [0.03:0.12]	0.07 [0.03:0.11]	0.06 [0.03:0.09]	0.09 [0.04:0.13]	0.08 [0.03:0.12]	0.07 [0.03:0.11]	0.06 [0.03:0.09]	0.09 [0.04:0.13]	0.08 [0.03:0.12]
$sd(\epsilon_e)$	0.14 [0.07:0.20]	0.12 [0.04:0.18]	0.16 [0.09:0.21]	0.14 [0.10:0.19]	0.16 [0.10:0.21]	0.14 [0.07:0.20]	0.14 [0.04:0.21]	0.15 [0.09:0.20]	0.16 [0.10:0.21]	0.14 [0.07:0.20]	0.14 [0.04:0.21]	0.15 [0.09:0.20]
$sd(\epsilon_\pi)$	-	-	-	-	-	-	-	-	-	-	-	-
$sd(\epsilon_y)$	-	-	-	-	-	-	-	-	-	-	-	-
Price contract length												
$\frac{1}{1-\xi}$	3.57	4.17	2.78	3.13	3.70	4.00	3.03	3.23	3.70	4.00	3.03	3.23
LL and posterior model odd												
LL	-96.95	-92.85	-103.77	-99.48	-96.62	-90.90	-102.03	-96.48	-96.62	-90.90	-102.03	-96.48
Prob.	0.002	0.124	0.000	0.000	0.003	0.868	0.000	0.003	0.003	0.868	0.000	0.003

TABLE 7: BAYESIAN POSTERIOR DISTRIBUTIONS  $\diamond$ 

$\diamond$  Notes: we report posterior means and 90% probability intervals (in parentheses) based on the output of the Metropolis-Hastings Algorithm. Sample range: 1981:I to 2006:IV.