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

September 5, 2011

Abstract

This paper examines the implications of imperfect information for optimal monetary policy with a consistent set of informational assumptions for the modeller and the private sector. 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*. We use an estimated simple NK model from Levine *et al.* (2010), where the assumption of symmetric imperfect information significantly improves the fit of the model to US data. The policy questions we then 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 to assess the welfare costs of imperfect information under commitment, discretion and simple Taylor-type rules. Our main results are: limiting information to only lagged macro-variables has significant implications both for welfare and for the form of the simple rule. In the unconstrained exercise without ZLB concerns, the gains from commitment are very small, but variances of the nominal interest rate indicate that the ZLB needs to be addressed. Then the picture changes drastically and the welfare gains from commitment are large. A price-level rule mimics the optimal commitment rule best and we observe a 'tying one's hands' effect in which under discretion there are welfare gains from only observing lagged rather than current output and inflation.

JEL Classification: C11, C52, E12, E32.

Keywords: Imperfect Information, DSGE Model, Optimal Monetary Policy, Bayesian Estimation

---

*To be presented at the MONFISPOL final Conference at Goethe University, September 19 - 20, 2011. The paper has also been presented at the CDMA Conference "Expectations in Dynamic Macroeconomic Models" at St Andrews University, August 31 - September 2, 2011; the 17th International Conference on Computing in Economics and Finance, San Francisco, June 29 - July 1, 2011 and the European Monetary Forum, University of York, March 4 - 5, 2011. Comments by participants at these events are gratefully acknowledged, as are those by seminar participants at Glasgow University and the University of Surrey. We also acknowledge financial support from ESRC project RES-062-23-2451 and from the EU Framework Programme 7 project MONFISPOL. **File: Optpol15_Frankfurt.tex**

Contents

1 Introduction

The formal estimation of DSGE models by Bayesian methods has now become standard.[1] 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. This underlying informational assumption corresponds to the second category above. 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.[2]

The focus of our paper here is on the implications of imperfect information for optimal monetary policy. 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.

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).[3] 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. Section 3 sets out the general solution procedure for solving such a model under imperfect information given a particular (and usually sub-optimal) policy rule. Section 4 describes the estimation by Bayesian methods drawing upon Levine *et al.* (2010). Section 5 sets out the general framework for calculating optimal policy. Section 6 turns to optimal policy assuming perfect information for both the private sector and the policymaker, 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 treat-

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

[2] 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)

[3] Section provides a taxonomy of the various assumed information structures assumed in these papers.

ment is the consideration the zero lower bound in the design of policy rules. In section 6 both sets of agents, the central bank and the private sector observed the full state vector describing the model model dynamics. Section 7 relaxes this assumption by allowing two forms of symmetric imperfect information and considers rules that correspond to the ICP adopted at the estimation side. Section 8 provides an application of optimal policy under perfect and imperfect information using our estimated DSGE model. Section 9 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

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

Equation (1) is the familiar Euler equation with β the discount factor, $1+R_t$ the gross nominal interest rate, MU^C_t 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^L_t 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 γ with an exogenous mark-up shock MS_t. They are derived from the optimal price-setting first-order condition for a firm j setting a new optimized price $P_t^0(j)$ given by

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

where the stochastic discount factor $D_{t,t+k}=\beta^k\frac{MU^C_{t+k}/P_{t+k}}{MU^C_t/P_t}$, 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 (**??**) and (7).

Equation (6) is the production function with labour the only variable input into production and the technology shock A_t exogenous. Price dispersion Δ_t, defined by (**??**), 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

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

where $\Pi_{tar,t}$ is a time-varying inflation target following an AR(1) process, (**??**), and $\epsilon_{e,t}$ is a monetary policy shock.[4] The following form of the single period utility for household r is a non-separable function of consumption and labour effort that is consistent with a balanced growth steady state:

$$U_t=\frac{\left[(C_t(r)-h_C C_{t-1})^{1-\varrho}(1-L_t(r))^{\varrho}\right]^{1-\sigma}}{1-\sigma} \quad (13)$$

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

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:

$$MU_t^C = (1-\varrho)(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)-1}(1-L_t)^{\varrho(1-\sigma)} \quad (14)$$
$$MU_t^L = -(C_t - h_C C_{t-1})^{(1-\varrho)(1-\sigma)}\varrho(1-L_t)^{\varrho(1-\sigma)-1} \quad (15)$$

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 Π_t $\tilde{\Pi}_t$ C_t Y_t Δ_t R_t MC_t MU_t^C U_t MU_t^L L_t H_t J_t and $\frac{W_t}{P_t}$. There are 6 shocks in the system: A_t, G_t, $MS_{per,t}$, $MS_{tra,t}$, $\Pi_{tar,t}$ and $\epsilon_{e,t}$.

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

3 General Solution with Imperfect Information

The model with a particular and not necessarily optimal rule is a special case of the following general setup in non-linear form

$$Z_{t+1} = J(Z_t, E_t Z_t, X_t, E_t X_t) + \nu\sigma_\epsilon\epsilon_{t+1} \quad (16)$$
$$E_t X_{t+1} = K(Z_t, E_t Z_t, X_t, E_t X_t) \quad (17)$$

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_ϵ is a small

[5] Lower case variables are defined as $x_t = \log\frac{X_t}{X}$. r_t and π_t are log-deviations of *gross* rates. Details of the linearization are provided in Appendix A.

scalar. In log-linearized form the state-space representation is

$$\begin{bmatrix} z_{t+1} \\ E_t x_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + B \begin{bmatrix} E_t z_t \\ E_t x_t \end{bmatrix} + \begin{bmatrix} u_{t+1} \\ 0 \end{bmatrix} \tag{18}$$

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, in non-linear and linear forms respectively, by

$$M_t^{obs} = m(Z_t, E_t Z_t, X_t, E_t X_t) + \mu\sigma_\epsilon \epsilon_t \tag{19}$$

$$m_t = \begin{bmatrix} M_1 & M_2 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} L_1 & L_2 \end{bmatrix} \begin{bmatrix} E_t z_t \\ E_t x_t \end{bmatrix} + v_t \tag{20}$$

where $\mu\sigma_\epsilon\epsilon_t$ 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 $B = L = 0$.[6] Full details of the solution for the general setup are provided in PCL.

In our model expressed in linearized form (see Appendix A) we consider two forms of imperfect information for the private sector consistent with the ICP

$$\textbf{Information Set I}: \quad m_t = \begin{bmatrix} y_t \\ \pi_t \\ r_t \end{bmatrix} \tag{21}$$

$$\textbf{Information Set II}: \quad m_t = \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \\ r_t \end{bmatrix} \tag{22}$$

This contrasts with the information set under perfect information which consists of all the state variables including the shock processes a_t, g_t, etc.

3.1 Linear Solution Procedure

Now we turn to the solution for (18) and (20). First assume *perfect information*. Following Blanchard and Kahn (1980), it is well-known that there is then a saddle path satisfying:

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

[6] In fact our model is of this simplified form.

where Λ^U has unstable eigenvalues. In the *imperfect information* case, following PCL, we use the Kalman filter updating given by

$$\begin{bmatrix} z_{t,t} \\ x_{t,t} \end{bmatrix} = \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix} + J \left[m_t - \begin{bmatrix} M_1 + L_1 & M_2 + L_2 \end{bmatrix} \begin{bmatrix} z_{t,t-1} \\ x_{t,t-1} \end{bmatrix} \right]$$

where we denote $z_{t,t} \equiv E_t[z_t]$ etc. Thus the best estimator of the state vector at time $t-1$ is updated by multiple J of the error in the predicted value of the measurement. The matrix J is given by

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

where $D \equiv M_1 - M_2 A_{22}^{-1} A_{21}$, $M \equiv [M_1 \;\; M_2]$ is partitioned conformably with $\begin{bmatrix} z_t \\ x_t \end{bmatrix}$, $\Gamma \equiv EPD^T + V$ where $E \equiv M_1 + L_1 - (M_2 + L_2)N$, $V = \text{cov}(v_t)$ is the covariance matrix of the measurement errors and P satisfies the Ricatti equation (26) below.

Using the Kalman filter, the solution as derived by PCL[7] 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 $\tilde{z}_t \equiv z_t - z_{t,t-1}$:

$$\begin{aligned} \text{Predetermined}: \quad z_{t+1} &= Cz_t + (A - C)\tilde{z}_t + (C - A)PD^T(DPD^T + V)^{-1}(D\tilde{z}_t + v_t) \\ &+ u_{t+1} && (23) \\ \text{Non-predetermined}: \quad x_t &= -Nz_t + (N - A_{22}^{-1}A_{21})\tilde{z}_t && (24) \\ \text{Innovations}: \quad \tilde{z}_{t+1} &= A\tilde{z}_t - APD^T(DPD^T + V)^{-1}(D\tilde{z}_t + v_t) + u_{t+1} && (25) \end{aligned}$$

where

$$C \equiv A_{11} - A_{12}N, \quad A \equiv A_{11} - A_{12}A_{22}^{-1}A_{21}, \quad D \equiv M_1 - M_2 A_{22}^{-1} A_{21}$$

and P is the solution of the Riccati equation given by

$$P = APA^T - APD^T(DPD^T + V)^{-1}DPA^T + U \tag{26}$$

where $U \equiv \text{cov}(u_t)$ is the covariance matrix of the shocks to the system. The measurement m_t can now be expressed as

$$m_t = Ez_t + (D - E)\tilde{z}_t + v_t - (D - E)PD^T(DPD^T + V)^{-1}(D\tilde{z}_t + v_t) \tag{27}$$

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

$$z_{t+1} = Cz_t + u_{t+1}\,; \quad x_t = -Nz_t \tag{28}$$

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

By comparing (28) with (23) we see that the determinacy of the system is independent of the information set. This is an important property that contrasts with the case where private agents use statistical learning to form forward expectations.[8]

3.2 The Filtering and Likelihood Calculations

To evaluate the likelihood for a given set of parameters (prior to multiplying by their prior probabilities), the econometrician takes the equations (23), (25) and (27) as given, and for the case when measurement error shocks are zero[9] evaluates the the Kalman filtering equations for $z_{t,t-1}$ and $Z_t \equiv \text{cov}(z_t)$:

$$z_{t+1,t} = Cz_{t,t-1} + CZ_tD^T(DPD^T)^{-1}e_t \qquad e_t \equiv w_t - Ez_{t,t-1}$$

$$Z_{t+1} = CZ_tC^T + PD^T(DPD^T)^{-1}DP - CZ_tE^T(EZ_tE^T)^{-1}EZ_tC^T \tag{29}$$

the latter being a time-dependent Riccati equation.

The period-t likelihood function is standard:

$$2\ln L = -\sum \ln\det(\text{cov}(e_t) - \sum e_t^T(\text{cov}(e_t))^{-1}e_t$$

where from $e_t \equiv w_t - Ez_{t,t-1}$ and (27)

$$\text{cov}(e_t) = EZ_tE$$

The system is initialized at $z_{1,0} = Z_0 = 0$ to obtain Z_1 from (29).

3.3 When Can Perfect Information be Inferred?

We now pose the question: under what conditions do the RE solutions under perfect and imperfect information actually differ? By observing a subset of outcomes can agents actually infer the full state vector, including shocks?

To answer this basic question we first explore the possibility of representing the solution to the model under imperfect information as a VAR.[10] First define $s_t \equiv \begin{bmatrix} z_t \\ \tilde{z}_t \end{bmatrix}$ and $\epsilon_t \equiv \begin{bmatrix} u_t \\ v_{t-1} \end{bmatrix}$ and

$$m_t = \begin{bmatrix} \tilde{M}_1 & \tilde{M}_2 \end{bmatrix} \begin{bmatrix} s_t \\ x_t \end{bmatrix} + v_t \tag{30}$$

[8]Our imperfect information framework encompasses the rational inattention approach of Sims (2005), Adam (2007) and Luo and Young (2009) as a special case. See Levine *et al.* (2010).

[9]It is straightforward to generalize to the the case for V non-zero.

[10]This section essentially generalizes Fernandez-Villaverde *et al.* (2007) to the case of imperfect information.

Then the solution set out in the previous section can be written as

$$s_{t+1} = \tilde{A}s_t + \tilde{B}\epsilon_{t+1} \quad (31)$$

$$x_t = -\tilde{N}s_t \quad (32)$$

where $\tilde{A}$, $\tilde{B}$ and $\tilde{N}$ are functions of A, B, C, N, P, D, U and V. Hence

$$m_{t+1} = (\tilde{M}_1 - \tilde{M}_2\tilde{N})(\tilde{A}s_t + \tilde{B}\epsilon_{t+1}) + v_{t+1} \equiv \tilde{C}s_t + \tilde{D}\epsilon_{t+1} \quad (33)$$

Suppose that *the number of shocks=the number of observed variables.* With at least one shock this can only be true if there is no measurement error; so we also put $v_t = 0$. With this assumption D is square. Suppose first that it is invertible. Then we can write

$$\epsilon_{t+1} = \tilde{D}^{-1}(m_{t+1} - \tilde{C}s_t)$$

Substituting into (31) we then have

$$[I - (\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C})L]s_{t+1} = \tilde{B}\tilde{D}^{-1}m_{t+1}$$

Iterating we arrive at

$$s_t = \sum_{j=0}^{\infty}[\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C}]^j\tilde{B}\tilde{D}^{-1}m_{t-j} \quad (34)$$

$$m_{t+1} = \tilde{C}\sum_{j=0}^{\infty}[\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C}]^j\tilde{B}\tilde{D}^{-1}w_{t-j} + \tilde{D}\epsilon_{t+1} \quad (35)$$

Then provided matrix $[\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C}]$ has stable eigenvalues, the summations converge.[11] Then (35) is an *infinite* VAR representation of the solution to our DSGE model. Furthermore, from (34), observations on the history of m_t imply that s_t is observed. This is consistent with our full information RE assumption. Thus we have the result that *if agents observe m_t without measurement error and if the number of shocks = the number of observations, then by observing the latter agents can infer the full state vector if $\tilde{D}$ is invertible. Imperfect information is equivalent to complete information in this special case.*

Under what conditions would $\tilde{D}$ be singular? An obvious case is under imperfect information case II where some variables are observed only with a lag. Then the current shocks cannot influence these observed variables so some of rows (two in this case) are zero

[11] This is an innocuous requirement, and in general would be imposed in order to ensure uniqueness. In fact, provided that A has stable eigenvalues, then the w_t process is stationary. There are a finite number of representations of w_t of the form (31) and (33), which are all factorizations of the spectral density for w_t given by $(D + C(I - AL)^{-1}BL)\Sigma(D' + B'(I - A'L^{-1})^{-1}C'L^{-1})$ where $\Sigma = \mathrm{var}(\varepsilon_t)$. Conventionally one chooses the values of B, C, D which ensure that the system is invertible, namely that ε can be expressed as a stable representation in w. For example, the following two representations of m_t have identical second moment properties: $m_t = \theta\varepsilon_t - \varepsilon_{t-1}$, $w_t = \varepsilon_t - \theta\varepsilon_{t-1}$, but only the latter is invertible for $-1 < \theta < 1$.

meaning $\tilde{D}$ is not invertible. In our model then, both these sufficient conditions for imperfect information collapsing to the perfect information case do not hold, so we can expect differences between the two cases.[12]

4 Bayesian Estimation

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

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

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

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

In the process of parameter estimation, the mode of the posterior is first estimated using Chris Sim's csminwel after the models' log-prior densities and log-likelihood functions are obtained by running the Kalman recursion and are evaluated and maximized. Then a sample from the posterior distribution is obtained with the Metropolis-Hasting algorithm using the inverse Hessian at the estimated posterior mode as the covariance matrix of the jumping distribution. The scale used for the jumping distribution in the MH is set in order

[12] In fact many NK DSGE models do have the property that the number of shocks equal the number of observable and the latter are current values without lags - for example Smets and Wouters (2003).

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.

4.1 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 8 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 ξ are the exception and based on Smets and Wouters (2007) with smaller standard deviations. Also, for the parameters γ, h_C, ξ and ϱ we centre the prior density in the middle of the unit interval. The priors related to the process for the price mark-up shock are taken from Smets and Wouters (2007). The priors for $\mu_1, \mu_2, \mu_3, \lambda_h, \lambda_f$ are also assumed beta distributed with means 0.5 and standard deviations 0.2. Three of the structural parameters are kept fixed in the estimation procedure. These calibrated parameters are $\beta = 0.99$; $L = 0.4$, $c_y = 0.6$.

4.2 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, π_t, y_t (II) and measurement error on *two* observables π_t, y_t (IIME).

This gives 12 sets of results. First Table 9 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 γ imply that inflation is intrinsically not very persistent in the relevant model specifications. 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 $3 - 4$ (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 θ for model m_i combines the likelihood function with the prior distribution:

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

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

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

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

$$PO_{i,j} = \frac{p(m_i|y)}{p(m_j|y)} = \frac{L(y|m_i)p(m_i)}{L(y|m_j)p(m_j)} \tag{36}$$

$$BF_{i,j} = \frac{L(y|m_i)}{L(y|m_j)} = \frac{\exp(LL(y|m_i))}{\exp(LL(y|m_j))} \tag{37}$$

in terms of the log-likelihoods. Components (36) and (37) provide a framework for comparing alternative and potentially misspecified models based on their marginal likelihood. Such comparisons are important in the assessment of rival models.

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

Pr(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.[13] 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.

5 The General Set-Up and Optimal Policy Problem

This section describes the general set-up that applies irrespective of the informational assumptions. Removing the estimated rule (12), for a given set of observed policy instruments w_t we now consider a linearized model in a general state-space form:

$$\begin{bmatrix} \mathsf{z}_{t+1} \\ E_t\mathsf{x}_{t+1} \end{bmatrix} = A^1 \begin{bmatrix} \mathsf{z}_t \\ \mathsf{x}_t \end{bmatrix} + A^2 \begin{bmatrix} E_t\mathsf{z}_t \\ E_t\mathsf{x}_t \end{bmatrix} + B\mathsf{w}_t + \begin{bmatrix} \mathsf{u}_{t+1} \\ 0 \end{bmatrix} \tag{38}$$

where $\mathsf{z}_t, \mathsf{x}_t$ are vectors of backward and forward-looking variables, respectively, w_t is a vector of policy variables, and u_t is a i.d. zero mean shock variable with covariance matrix Σ_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 still given by (20).

Define target variables s_t by

$$\mathsf{s}_t = J\mathsf{y}_t + H\mathsf{w}_t \tag{39}$$

[13] 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.

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

$$\Omega_t = \frac{1}{2}\sum_{\tau=0}^{\infty}\beta^t[\mathsf{s}_{t+\tau}^T Q_1 \mathsf{s}_{t+\tau} + \mathsf{w}_{t+\tau}^T Q_2 \mathsf{w}_{t+\tau}] \tag{40}$$

This could be an ad hoc loss function or a large distortions approximation to the household's utility as described in Levine *et al.* (2008a). Substituting (39) into (40) results in the following form of the loss function used subsequently in the paper

$$\Omega_t = \frac{1}{2}\sum_{i=0}^{\infty}\beta^t[\mathsf{y}_{t+\tau}^T Q \mathsf{y}_{t+\tau} + 2\mathsf{y}_{t+\tau}^T U \mathsf{w}_{t+\tau} + \mathsf{w}_{t+\tau}^T R \mathsf{w}_{t+\tau}] \tag{41}$$

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:
   $I_t = \{\mathsf{z}_\tau, \mathsf{x}_\tau\}$, $\tau \leq t; A^1, A^2, B, \Sigma_u, [Q, U, R, \beta]$ or the monetary rule

2. For symmetric imperfect information (see Pearlman (1992), Svensson and Woodford (2003) and for Bayesian estimation, Levine *et al.* (2010)):
   $I_t = \{\mathsf{m}_\tau\}$, $\tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v, [Q, U, R, \beta]$ 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):
   $I_t^{ps} = I_t = \{\mathsf{z}_\tau, \mathsf{x}_\tau\}$, $\tau \leq t; A^1, A^2, B, \Sigma_u, [Q, U, R, \beta]$ or the monetary rule for the private sector and
   $I_t^{pol} = \{\mathsf{m}_\tau\}$, $\tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v, [Q, U, R, \beta]$ 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)):
   $I_t^{pol} = \{\mathsf{m}_\tau\}$, $\tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v, [Q, U, R, \beta]$ or the monetary rule for the policymaker sector and
   $I_t^{ps} = \{\mathsf{m}_\tau\}$, $\tau \leq t; A^1, A^2, B, M^1, M^2, L, \Sigma_u, \Sigma_v$ 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. However this interesting case is beyond the scope of this paper.

6 Optimal Policy Under Perfect Information

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

$$\begin{bmatrix} \mathsf{z}_{t+1} \\ \mathsf{x}^e_{t+1,t} \end{bmatrix} = A \begin{bmatrix} \mathsf{z}_t \\ \mathsf{x}_t \end{bmatrix} + B\mathsf{w}_t \tag{42}$$

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 $\mathsf{x}^e_{t+1,t}$ denotes rational (model consistent) expectations of x_{t+1} formed at time t. Then $\mathsf{x}^e_{t+1,t} = \mathsf{x}_{t+1}$ and letting $\mathsf{y}_t^T = \begin{bmatrix} \mathsf{z}_t^T & \mathsf{x}_t^T \end{bmatrix}$ (42) becomes

$$\mathsf{y}_{t+1} = A\mathsf{y}_t + B\mathsf{w}_t \tag{43}$$

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

6.1 The Optimal Policy with Commitment

Consider the policy-maker's *ex-ante* optimal policy at $t = 0$. This is found by minimizing Ω_0 given by (41) subject to (43) and (39) and given z_0. We proceed by defining the Hamiltonian

$$\mathcal{H}_t(y_t, y_{t+1}, \mu_{t+1}) = \frac{1}{2}\beta^t(\mathsf{y}_t^T Q\mathsf{y}_t + 2\mathsf{y}_t^T U\mathsf{w}_t + \mathsf{w}_t^T R\mathsf{w}_t) + \mu_{t+1}(A\mathsf{y}_t + B\mathsf{w}_t - \mathsf{y}_{t+1}) \tag{44}$$

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

$$\mathcal{L}_0(y_0, y_1, \ldots, w_0, w_1, \ldots, \mu_1, \mu_2, \ldots) = \sum_{t=0}^{\infty} \mathcal{H}_t \tag{45}$$

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 $\mathsf{p}_t = \beta^{-t}\mu_t^T$, the first-order conditions lead to

$$\mathsf{w}_t = -R^{-1}(\beta B^T \mathsf{p}_{t+1} + U^T \mathsf{y}_t) \tag{46}$$

$$\beta A^T \mathsf{p}_{t+1} - \mathsf{p}_t = -(Q\mathsf{y}_t + U\mathsf{w}_t) \tag{47}$$

Substituting (46) into (43)) we arrive at the following system under control

$$\begin{bmatrix} I & \beta BR^{-1}B^T \\ 0 & \beta(A^T - UR^{-1}B^T) \end{bmatrix} \begin{bmatrix} \mathsf{y}_{t+1} \\ \mathsf{p}_{t+1} \end{bmatrix} = \begin{bmatrix} A - BR^{-1}U^T & 0 \\ -(Q - UR^{-1}U^T & I \end{bmatrix} \begin{bmatrix} \mathsf{y}_t \\ \mathsf{p}_t \end{bmatrix} \tag{48}$$

To complete the solution we require $2n$ boundary conditions for (48). Specifying z_0

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

$$\lim_{t\to\infty}\mu_t^T = \lim_{t\to\infty}\beta^t \mathsf{p}_t = 0 \tag{49}$$

and the initial condition

$$\mathsf{p}_{20} = 0 \tag{50}$$

where $\mathsf{p}_t^T = \left[\mathsf{p}_{1t}^T\, \mathsf{p}_{2t}^T\right]$ is partitioned so that p_{1t} is of dimension $(n-m)\times 1$. Equation (39), (46), (48) together with the $2n$ boundary conditions constitute the system under optimal control.

Solving the system under control leads to the following rule

$$\mathsf{w}_t = -F\begin{bmatrix} I & 0 \\ -N_{21} & -N_{22}\end{bmatrix}\begin{bmatrix}\mathsf{z}_t \\ \mathsf{p}_{2t}\end{bmatrix} \equiv D\begin{bmatrix}\mathsf{z}_t \\ \mathsf{p}_{2t}\end{bmatrix} = -F\begin{bmatrix}\mathsf{z}_t \\ \mathsf{x}_{2t}\end{bmatrix} \tag{51}$$

where

$$\begin{bmatrix}\mathsf{z}_{t+1} \\ \mathsf{p}_{2t+1}\end{bmatrix} = \begin{bmatrix} I & 0 \\ S_{21} & S_{22}\end{bmatrix} G \begin{bmatrix} I & 0 \\ -N_{21} & -N_{22}\end{bmatrix}\begin{bmatrix}\mathsf{z}_t \\ \mathsf{p}_{2t}\end{bmatrix} \equiv H\begin{bmatrix}\mathsf{z}_t \\ \mathsf{p}_{2t}\end{bmatrix} \tag{52}$$

$$N = \begin{bmatrix} S_{11} - S_{12}S_{22}^{-1}S_{21} & S_{12}S_{22}^{-1} \\ -S_{22}^{-1}S_{21} & S_{22}^{-1}\end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22}\end{bmatrix} \tag{53}$$

$$\mathsf{x}_t = -\begin{bmatrix} N_{21} & N_{22}\end{bmatrix}\begin{bmatrix}\mathsf{z}_t \\ \mathsf{p}_{2t}\end{bmatrix} \tag{54}$$

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

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22}\end{bmatrix} \tag{55}$$

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 Ricatti equation

$$S = Q - UF - F^T U^T + F^T RF + \beta(A - BF)^T S(A - BF) \tag{56}$$

The welfare loss for the optimal policy (OPT) at time t is

$$\Omega_t^{OPT} = -\frac{1}{2}(\mathrm{tr}(N_{11}Z_t) + \mathrm{tr}(N_{22}\mathsf{p}_{2t}\mathsf{p}_{2t}^T)) \tag{57}$$

where $Z_t = \mathsf{z}_t\mathsf{z}_t^T$. To achieve optimality the policy-maker sets $\mathsf{p}_{20} = 0$ at time $t = 0$.[14] At

[14] Noting from (54) that for the optimal policy we have $\mathsf{x}_t = -N_{21}\mathsf{z}_t - N_{22}\mathsf{p}_{2t}$, the optimal policy "from a timeless perspective" proposed by Woodford (2003) replaces the initial condition for optimality $p_{20} = 0$ with $J\mathsf{x}_0 = -N_{21}\mathsf{z}_0 - N_{22}\mathsf{p}_{20}$ 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.

time $t > 0$ there exists a gain from reneging by resetting $\mathsf{p}_{2t} = 0$. It can be shown that $N_{11} < 0$ and $N_{22} < 0$.[15], so the incentive to renege exists at all points along the trajectory of the optimal policy. This is the time-inconsistency problem.

6.2 The Dynamic Programming Discretionary Policy

To evaluate the discretionary (time-consistent) policy we rewrite the welfare loss Ω_t given by (41) as

$$\Omega_t = \frac{1}{2}[\mathsf{y}_t^T Q \mathsf{y}_t + 2\mathsf{y}_t^T U \mathsf{w}_t + \mathsf{w}_t^T R \mathsf{w}_t + \beta\Omega_{t+1}] \tag{58}$$

The dynamic programming solution then seeks a stationary solution of the form $\mathsf{w}_t = -\mathsf{F}\mathsf{z}_t$ in which Ω_t is minimized at time t subject to (1) in the knowledge that a similar procedure will be used to minimize Ω_{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

$$\mathsf{x}_{t+\tau} = -N_{t+1}\mathsf{z}_{t+\tau}, \;\; \tau \geq 1 \tag{59}$$

The loss at time t for the *ex ante* optimal policy was from (57) found to be a quadratic function of x_t and p_{2t}. We have seen that the inclusion of p_{2t} was the source of the time inconsistency in that case. We therefore seek a lower-order controller

$$\mathsf{w}_t = -F\mathsf{z}_t \tag{60}$$

with the welfare loss in z_t only. We then write $\Omega_{t+1} = \frac{1}{2}\mathsf{z}_{t+1}^T S_{t+1}^{TCT}\mathsf{z}_{t+1}$ in (58). This leads to the following iterative process for F_t

$$\mathsf{w}_t = -F_t\mathsf{z}_t \tag{61}$$

where

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

[15]See Currie and Levine (1993), chapter 5.

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

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^{TCT} starting with some initial values for N_t and S_t^{TCT}. If the process converges to stationary values, F, N and S say, then the time-consistent feedback rule is $\mathsf{w}_t = -F\mathsf{z}_t$ with loss at time t given by

$$
\Omega_t^{TCT} = \frac{1}{2}\mathsf{z}_t^T S^{TCT} \mathsf{z}_t = \frac{1}{2}\mathrm{tr}(S^{TCT} Z_t) \tag{62}
$$

6.3 Optimized Simple Rules

We now consider simple sub-optimal rules of the form

$$
\mathsf{w}_t = D\mathsf{y}_t = D \begin{bmatrix} \mathsf{z}_t \\ \mathsf{x}_t \end{bmatrix} \tag{63}
$$

where D is constrained to be sparse in some specified way. Rule (63) can be quite general. By augmenting the state vector in an appropriate way it can represent a PID (proportional-integral-derivative)controller.

Substituting (63) into (41) gives

$$
\Omega_t = \frac{1}{2}\sum_{i=0}^{\infty} \beta^t \mathsf{y}_{t+i}^T P_{t+i} \mathsf{y}_{t+i} \tag{64}
$$

where $P = Q + UD + D^T U^T + D^T RD$. The system under control (42), with w_t given by (63), has a rational expectations solution with $\mathsf{x}_t = -N\mathsf{z}_t$ where $N = N(D)$. Hence

$$
\mathsf{y}_t^T P \mathsf{y}_t = \mathsf{z}_t^T T \mathsf{z}_t \tag{65}
$$

where $T = P_{11} - N^T P_{21} - P_{12}N + N^T P_{22} N$, P is partitioned as for S in (55) onwards and

$$
\mathsf{z}_{t+1} = (G_{11} - G_{12}N)\mathsf{z}_t \tag{66}
$$

where $G = A + BD$ is partitioned as for P. Solving (66) we have

$$
\mathsf{z}_t = (G_{11} - G_{12}N)^t \mathsf{z}_0 \tag{67}
$$

Hence from (68), (65) and (67) we may write at time t

$$\Omega_t^{SIM} = \frac{1}{2} z_t^T V z_t = \frac{1}{2}\mathrm{tr}(V Z_t) \tag{68}$$

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

$$V^{LYA} = T + H^T V^{LYA} H \tag{69}$$

where $H = G_{11} - G_{12}N$. At time $t = 0$ the optimized simple rule is then found by minimizing Ω_0 given by (68) 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

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

6.4 The Stochastic Case

Consider the stochastic generalization of (42)

$$\begin{bmatrix} \mathsf{z}_{t+1} \\ \mathsf{x}^e_{t+1,t} \end{bmatrix} = A \begin{bmatrix} \mathsf{z}_t \\ \mathsf{x}_t \end{bmatrix} + B\mathsf{w}_t + \begin{bmatrix} \mathsf{u}_t \\ 0 \end{bmatrix} \tag{70}$$

where u_t is an $n \times 1$ vector of white noise disturbances independently distributed with $\mathrm{cov}(\mathsf{u}_t) = \Sigma$. 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 $\mathsf{z}_t^T X \mathsf{z}_t = \mathrm{tr}(X\mathsf{z}_t, Z_t^T)$ replaced with

$$E_t\left(\mathrm{tr}\left[X\left(z_t z_t^T + \sum_{i=1}^{\infty} \beta^t \mathsf{u}_{t+i}\mathsf{u}_{t+i}^T\right)\right]\right) = \mathrm{tr}\left[X\left(z_t^T z_t + \frac{\lambda}{1-\lambda}\Sigma\right)\right] \tag{71}$$

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

Thus for the optimal policy with commitment (57) becomes in the stochastic case

$$\Omega_t^{OPT} = -\frac{1}{2}\mathrm{tr}\left(N_{11}\left(Z_t + \frac{\beta}{1-\beta}\Sigma\right) + N_{22}\mathsf{p}_{2t}\mathsf{p}_{2t}^T\right) \tag{72}$$

For the time-consistent policy (62) becomes

$$\Omega_t^{TCT} = -\frac{1}{2}\mathrm{tr}\left(S\left(Z_t + \frac{\beta}{1-\beta}\Sigma\right)\right) \tag{73}$$

and for the simple rule, generalizing (68)

$$\Omega_t^{SIM} = -\frac{1}{2}\mathrm{tr}\left(V^{LYA}\left(Z_t + \frac{\beta}{1-\beta}\Sigma\right)\right) \tag{74}$$

The optimized simple rule is found at time $t = 0$ by minimizing Ω_0^{SIM} given by (74). Now we find that

$$D^* = D^* \left(\mathsf{z}_0 \mathsf{z}_0^T + \frac{\beta}{1-\beta}\Sigma \right) \tag{75}$$

or, in other words, the optimized rule depends both on the initial displacement z_0 and on the covariance matrix of disturbances Σ.

A very important feature of optimized simple rules is that unlike their optimal commitment or optimal discretionary counterparts they are *not certainty equivalent.* In fact if the rule is designed at time $t = 0$ then $D^* = f^* \left(Z_0 + \frac{\beta}{1-\beta}\Sigma \right)$ and so depends on the displacement z_0 at time $t = 0$ and on the covariance matrix of innovations $\Sigma = \mathrm{cov}(\epsilon_t)$. From non-certainty equivalence it follows that if the simple rule were to be re-designed at ant time $t > 0$, since the re-optimized D^* will then depend on Z_t the new rule will differ from that at $t = 0$. This feature is true in models with or without rational forward-looking behaviour and it implies that *simple rules are time-inconsistent even in non-RE models.*

7 Optimal Policy Under Imperfect Information

Here we assume that that there is a set of measurements as described above in section 3. The following is a summary of the solution provided by Pearlman (1992), with some details provided in Appendix C. It can be shown that the estimate for z_t at time t, denoted by $\mathsf{z}_{t,t}$ can be expressed in terms of the innovations process $\mathsf{z}_t - \mathsf{z}_{t,t-1}$ as

$$\mathsf{z}_{t,t} = \mathsf{z}_{t,t-1} + PD^T(DPD^T + V)^{-1}(D(\mathsf{z}_t - \mathsf{z}_{t,t-1}) + \mathsf{v}_t) \tag{76}$$

where $D = M_1^1 - M_2^1(A_{22}^1)^{-1}A_{21}^1$, $M^1 = [M_1^1, M_2^1]$, $M^2 = [M_1^2, M_2^2]$ partitioned conformably with $[\mathsf{z}_t^T, \mathsf{x}_t^T]^T$, and P is the solution of the Riccati equation describing the Kalman Filter

$$P = APA^T - APD^T(DPD^T + V)^{-1}DPA^T + \Sigma \tag{77}$$

where $A = A_{11}^1 - A_{12}^1(A_{22}^1)^{-1}A_{21}^1$. One can also show that $\mathsf{z}_t - \mathsf{z}_{t,t}$ and $\mathsf{z}_{t,t}$ are orthogonal in expectations. Note that this Riccati equation *is independent of policy.* We may then write the expected utility as

$$\begin{aligned} \frac{1}{2}E_t\Big[\sum_{i=0}^{\infty}\beta^t(\mathsf{y}_{t+\tau,t+\tau}^T Q \mathsf{y}_{t+\tau,t+\tau} \quad &+ \quad 2\mathsf{y}_{t+\tau,t+\tau}^T U \mathsf{w}_{t+\tau} + \mathsf{w}_{t+\tau}^T R \mathsf{w}_{t+\tau} \\ &+ \quad (\mathsf{y}_{t+\tau} - \mathsf{y}_{t+\tau,t+\tau})^T Q (\mathsf{y}_{t+\tau} - \mathsf{y}_{t+\tau,t+\tau}))\Big] \end{aligned} \tag{78}$$

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

$$\begin{bmatrix} \mathsf{z}_{t+1,t+1} \\ E_t \mathsf{x}_{t+1,t+1} \end{bmatrix} = (A^1 + A^2) \begin{bmatrix} \mathsf{z}_{t,t} \\ \mathsf{x}_{t,t} \end{bmatrix} + B\mathsf{w}_t + \begin{bmatrix} \mathsf{z}_{t+1,t+1} - \mathsf{z}_{t+1,t} \\ 0 \end{bmatrix} \tag{79}$$

which represents the expected dynamics of the system (where we note by the chain rule that $E_t \mathsf{x}_{t+1,t+1} \triangleq E_t[E_{t+1}x_{t+1}] = E_t x_{t+1}$. Note that $\mathrm{cov}(\mathsf{z}_{t+1,t+1} - \mathsf{z}_{t+1,t}) = PD^T(DPD^T + V)^{-1}DP$ and $\mathrm{cov}(\mathsf{z}_{t+1} - \mathsf{z}_{t+1,t+1}) = P - PD^T(DPD^T + V)^{-1}DP \triangleq \bar{P}$.

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

$$0 = A^1_{12}(\mathsf{z}_t - \mathsf{z}_{t,t}) + A^1_{22}(\mathsf{x}_t - \mathsf{x}_{t,t}) \tag{80}$$

Furthermore, as in Pearlman (1992) we can show 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

$$\begin{aligned} W^J =& \mathsf{z}^T_{0,0} S^J \mathsf{z}_{0,0} + \frac{\lambda}{1-\lambda}\mathrm{tr}\Big(S^J PD^T(DPD^T + V)^{-1}DP\Big) \\ &+ \frac{1}{1-\lambda}\mathrm{tr}\Big(Q_{11} - Q_{12}(A^1_{22})^{-1}A^1_{21} - (A^1_{21})^T(A^1_{22})^{-T}Q_{21} + (A^1_{21})^T(A^1_{22})^{-T}Q_{22}(A^1_{22})^{-1}A^1_{21}\Big)\bar{P} \end{aligned} \tag{81}$$

where $J =$OPT, TCT, SIM; the second term is the expected value of the first three terms of (78) under each of the rules, and the final term is independent of the policy rule, and is the expected value of the final term of (78), utilising (80). Also note that from the perfect information case in the previous subsection:

$$S^{OPT} = N_{11} \equiv S_{11} - S_{12}S^{-1}_{22}S_{21} \tag{82}$$

- S_{ij} are the partitions of S, the Ricatti matrix used to calculate the welfare loss under optimal policy with commitment.
- S^{TCT} is used to calculate the welfare loss in the time consistent solution algorithm.
- $S^{SIM} = V^{LYA}$ is calculated from the Lyapunov equation used to calculate the welfare under the optimized simple rule.

In the special case of perfect information, $M^1 = I$, $M_2 = v_t = V$ so that $D = E = I$. It follows that $\bar{P} = 0$ and the last term in (81) disappears. Moreover $P = \Sigma$, $\mathsf{z}_{0,0} = \mathsf{z}_0$ and (81) reduces to the welfare loss expressions obtained previously. Thus the effect of imperfect information is to introduce a new term into the welfare loss that depends only on the model's transmission of policy but is independent of that policy and to modify the first policy-dependent term by an effect that depends on the solution P to the Ricatti equation associated with the Kalman Filter.

8 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.

This section addresses all these questions. We examine two imperfect information sets.
Imperfect Information Set I: This consists of the current and past values of output, inflation and the interest rate In this scenario the private sector must infer the shocks from its observations.
Imperfect Information Set II: As for I but output and inflation are only observed with a lag, but the current interest rate is observed.

We considered simple inflation targeting rules that respond only to inflation.[16] The corresponding forms of the rules for the two information sets are

$$r_t = \rho_r r_{t-1} + \theta_\pi \pi_t \tag{83}$$

for perfect information and information set I, and either

$$r_t = \rho_r r_{t-1} + \theta_\pi E_t \pi_t \quad (\textbf{Form A}) \tag{84}$$

or

$$r_t = \rho_r r_{t-1} + \theta_\pi \pi_{t-1} \quad (\textbf{Form B}) \tag{85}$$

for information set II.

With this choice of Taylor rule the case where $\rho_r = 1$ is of particular interest as this then corresponds to a *price-level* rule. There has been a recent interest in the case for price-level rather than inflation stability. Gaspar *et al.* (2010) provide an excellent review of this literature. The basic difference between the two regimes in that under an inflation targeting mark-up shock leads to a commitment to use the interest rate to accommodate an increase in the inflation rate falling back to its steady state. By contrast a price-level rule commits to a inflation rate below its steady state after the same initial rise. Under inflation targeting one lets bygones be bygones allowing the price level to drift to a permanently different price-level path whereas price-level targeting restores the price level to its steady state path. The

[16] We also considered the following simple rules that responding to both inflation and output:

$$r_t = \rho_r r_{t-1} + \theta_\pi \pi_t + \theta_y y_t$$

for perfect information and information set I, and either

$$r_t = \rho_r r_{t-1} + \theta_\pi E_t \pi_t + \theta_y E_t y_t \quad (\textbf{Form A})$$

$$r_t = \rho_r r_{t-1} + \theta_\pi \pi_{t-1} + \theta_y y_{t-1} \quad (\textbf{Form B})$$

for information set II. However the results were very similar with a very small weight on output.

latter can lower inflation variance and be welfare enhancing because forward-looking price-setters anticipates that a current increase in the general price level will be undone giving them an incentive to moderate the current adjustment of its own price. In our results we will see if price-level targeting is indeed welfare optimal across different information assumptions.

8.1 Optimal Policy without Zero Lower Bound Considerations

Results are presented for a loss function that is formally a quadratic approximation about the steady state of the Lagrangian, and which represents the true approximation about the fully optimal solution. This welfare-based loss function has been obtained numerically.

A permanent drop in consumption of 0.1464% produces a welfare loss per period of 100. So from Table 3 we can see (a) that the gains from commitment are very small under any information set; (b) that imperfect information in the form of only observing a subset of current state variables (information set I) imposes only a tiny welfare loss, whereas if only lagged output and inflation are observed, the losses are significant, and are of the order of 0.11% consumption equivalent.

Simple rules are able to quite well replicate the welfare losses under the fully optimal solution, albeit with Simple Rule A requiring a large weight $\theta_\pi = 16$ on inflation, and $\rho_r = 1.0$ on lagged interest rates (virtually the same for all information sets). This derivative rule in inflation is equivalent to a price level rule. For Simple Rule B, with lagged information on inflation, the welfare loss is similar but achieved with weights $\theta_\pi = 2.26$ and $\rho_r = 0.6$, so the price level rule does not apply in this case and the II with only lagged observations of output and inflation has important implications for the form of the optimized rule.

Information	Information Set	Optimal	Time Consistent	Simple Rule A	Simple Rule B
Perfect	Full state vector	20.06	21.22	21.39	n. a.
Imperfect I	$I_t = [y_t, \pi_t, r_t]$	20.45	21.62	20.46	n. a.
Imperfect II	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	95.61	97.6	95.62	95.85

Table 3: Welfare Costs per period of Imperfect Information without ZLB Considerations

Information	Information Set	Simple Rule A $[\rho_r,\ \theta_\pi]$	Simple Rule B $[\rho_r,\ \theta_\pi]$
Perfect	Full state vector	[1, 16]	n. a.
Imperfect I	$I_t = [y_t, \pi_t, r_t]$	[1,16]	n. a.
Imperfect II	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	[1, 16.5]	[0.6, 2.26]

Table 4: Optimized Coefficients in Simple Rules without ZLB Considerations

Information	Information Set	Optimal	Time Cons	Simple Rule A	Simple Rule B
Perfect	Full state vector	1.31	2.39	1.48	n. a.
Imperfect I	$I_t = [y_t, \pi_t, r_t]$	1.32	2.32	1.35	n. a.
Imperfect II	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	2.42	4.16	2.47	2.34

Table 5: Interest Rate Variances

However Table 5 indicates the aggressive nature of these rules leads to high interest rate variances resulting in a ZLB problem for all the rules and information sets. This now needs to be addressed.

8.2 Imposing an Interest Rate Zero Lower Bound Constraint

In the absence of a lower bound constraint on the nominal interest rate the policymaker's optimization problem is to minimize Ω_0 given by (41) subject to (43) and (39) and given z_0. If the variances of shocks are sufficiently large, this will lead to a large nominal interest rate variability and the possibility of the nominal interest rate becoming negative.

To rule out this possibility but remain within the tractable LQ framework, we follow Woodford (2003), chapter 6, and modify our interest-rate rules to approximately impose an interest rate ZLB so that this event hardly ever occurs. Our quadratic approximation to the single-period loss function can be written as $L_t = \mathsf{y}_t' Q \mathsf{y}_t$ where $\mathsf{y}_t' = [\mathsf{z}_t', \mathsf{x}_t']'$ and Q is a symmetric matrix. As in Woodford (2003), chapter 6, the ZLB constraint is implemented by modifying the single period welfare loss to $L_t + w_r r_t^2$.

Then following Levine *et al.* (2008b), the policymaker's optimization problem is to choose w_r and the unconditional distribution for R_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. 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 prob $(Z \leq z_0) = p$, $R_n^* = (1+\pi^*)R + \pi^*$ is the steady state nominal interest rate, R is the steady state real interest rate, $\sigma_r^2 = \text{var}(R_n)$ is the unconditional variance and π^* is the new steady state inflation rate. Given σ_r the steady state positive inflation rate that will ensure $R_t \geq 0$ with probability $1-p$ is given by

$$\pi^* = \max\left[\frac{z_0(p)\sigma_r - R}{1+R} \times 100, 0\right] \tag{86}$$

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 a contribution from $(\pi^*)^2$. By increasing w_r we can lower σ_r thereby decreasing π^* 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 $1-p$. Figure 1 illustrates shows the solution to the problem for optimal policy and perfect information. with $p = 0.0025$; ie., a probability of hitting the zero lower bound once every 400 quarters or 100 years.

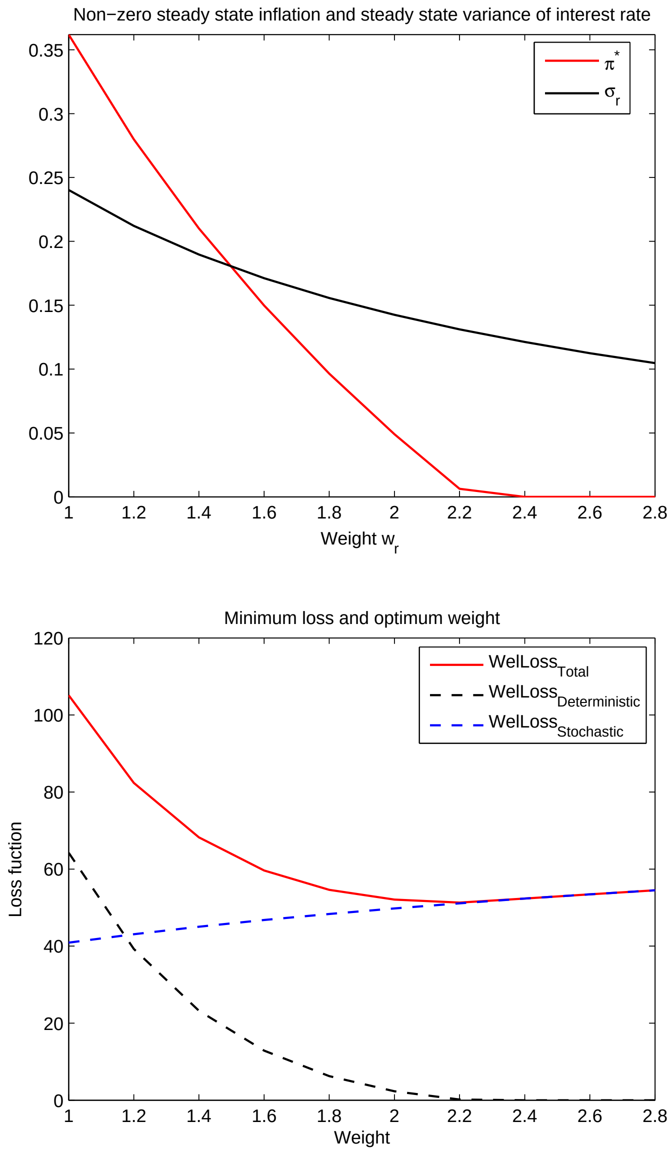

Figure 1: Imposition of ZLB for Optimal Policy and Perfect Information

Note that in our LQ framework, the zero interest rate bound is very occasionally hit. Then interest rate is allowed to become negative, possibly using a scheme proposed by Gesell (1934) and Keynes (1936). Our approach to the ZLB constraint (following Woodford

(2003))[17] in effect replaces it with a nominal interest rate variability constraint which ensures the ZLB is hardly ever hit. By contrast the work of a number of authors including Adam and Billi (2007), Coenen and Wieland (2003), Eggertsson and Woodford (2003) and Eggertsson (2006) study optimal monetary policy with commitment in the face of a non-linear constraint $i_t \geq 0$ which allows for frequent episodes of liquidity traps in the form of $i_t = 0$.

A problem with the procedure so far is that we shift the steady state to a new one with a higher inflation, but we continue to approximate the loss function and the dynamics about the original Ramsey steady state. We know from the work of Ascari and Ropele (2007a) and Ascari and Ropele (2007b) that the dynamic properties of the linearized model change significantly when the model is linearized about a non-zero inflation. This issue is addressed analytically in Coibion *et al.* (2011), but in a very simple NK model. We now propose a general solution and numerical procedure that can be used in any DSGE model.

1. Begin by defining a new parameter: p, the probability of hitting the ZLB, the weight w_r on the variance of the nominal net interest rate and a target steady state nominal interest rate $\hat{R}$.

2. Modify the single-period utility to $L_t = \Lambda_t - \frac{1}{2}w_r(R_t - \hat{R})^2$.

3. In the first iteration let w_r to be low to get through OPT, say $w_r = 0.001$ and $\hat{R} = \frac{1}{\beta} - 1$, the no-growth zero-inflation steady-state nominal interest rate corresponding to the standard Ramsey problem with no ZLB considerations.

4. Perform the LQ approximation of the Ramsey optimization problem with modified loss function L_t. For standard problems the steady state nominal net inflation rate $\pi^{Ramsey} = 0$ and $R^{Ramsey} = \frac{1}{\beta} - 1$.

5. Compute OPT or TCT or optimized simple rule SIM2 in as in the solution procedures above.

6. Extract $\sigma_r = \sigma_r(w_r)$.

7. Extract the minimized conditional (in the vicinity of the steady state, i.e. $z_0 = 0$ in ACES) stochastic loss function $\tilde{\Omega}_0(w_r)$

8. Compute $r^* = r^*(w_r)$ defined by $r^*(w_r) = \max\left[z_0(p)\sigma_r - R^{Ramsey} \times 100, 0\right]$, where in the first iteration $R^{Ramsey} = \frac{1}{\beta} - 1$ as noted above. This ensures that the ZLB is reached with a low probability p.

9. If $r^* < 0$, the ZLB constraint is not binding; if $r* > 0$ it is. Proceed in either case.

[17] We generalize the treatment of Woodford however by allowing the steady-state inflation rate to rise. Our policy prescription has recently been described as a “dual mandate” in which a central bank committed to a long-run inflation objective sufficiently high to avoid the ZLB constraint as well as a Taylor-type policy stabilization rule about such a rate - see Blanchard *et al.* (2010) and Gavin and Keen (2011).

10. Define $\pi^* = \pi^{Ramsey} + r^*$.

11. Compute the steady state $\bar{\Omega}_0(\pi^*)$ at the steady state of the model with a shifted new inflation rate π^*. Then compute $\Delta\bar{\Omega}_0(r^*(w_r)) \equiv \bar{\Omega}_0(\pi^*) - \bar{\Omega}_0(\pi^{Ramsey})$

12. Compute the actual total stochastic plus deterministic loss function that hits the ZLB with a low probability p

$$\Omega_0(w_r) = \tilde{\Omega}_0^{actual}(w_r) + \Delta\bar{\Omega}_0(r^*(w_r)) \tag{87}$$

13. A good approximation for $\tilde{\Omega}_0(w_r)^{actual}$ is $\tilde{\Omega}_0(w_r)^{actual} \simeq \tilde{\Omega}_0(w_r) - \frac{1}{2}w_r\sigma_r^2$ provided the welfare loss is multiplied by $1-\beta$.

14. Finally minimize $\Omega_0(w_r)$ with respect to w_r. This imposes the ZLB constraint as in Figure 1.

15. This is what we have currently in analysis. What now changes is to reset $\hat{R} = \frac{1}{\beta} - 1 + \alpha\pi^*$ where $\alpha \in (0,1]$ is a relaxation parameter to experiment with, i.e., $(\hat{R})^{new} = (\hat{R})^{old} + \alpha\pi^*$, $w_r^{new} = \text{argmin}\Omega_0(w_r)$ and return to the beginning. Iterate until $\pi^*(w_r) = 0$ and w_r is unchanged. In our experience with some appropriate choice of α this algorithm converges.

Information	Information Set	Optimal	Time Consis	Sim Rule A	Sim Rule B
Perfect (Wel Loss)	Full state vector	51.3 (0.08)	3652 (5.34)	106.4 (0.16)	n. a.
Perfect (Weight w_r)	Full state vector	2.2	0.75	5.5	n. a.
Imperf I (Wel Loss)	$I_t = [y_t, \pi_t, r_t]$	52.9 (0.08)	3873 (5.67)	101.4 (0.15)	n. a.
Imperf I ((Weight w_r)	$I_t = [y_t, \pi_t, r_t]$	2.3	0.7	5.0	n. a.
Imperf II (Wel Loss)	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	166.9 (0.24)	2761 (4.04)	275.2 (0.40)	188 (0.28)
Imperf II ((Weight w_r)	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	3.8	1.0	8.0	5.0

Table 6: Welfare Costs per period of Imperfect Information *with* ZLB Considerations. Consumption Equivalent Losses (%) in brackets. Prob of hitting ZLB=0.0025.

Information	Information Set	Simple Rule A $[\rho_r,\ \theta_\pi]$	Simple Rule B $[\rho_r,\ \theta_\pi]$
Perfect	Full state vector	[1, 0.45]	n. a.
Imperfect I	$I_t = [y_t, \pi_t, r_t]$	[1, 0.44]	n. a.
Imperfect II	$I_t = [y_{t-1}, \pi_{t-1}, r_t]$	[1, 0.3]	[1, 0.22]

Table 7: Optimized Coefficients in Simple Rules *with* ZLB Considerations

Using this algorithm we obtain the following results. Introducing the ZLB constraint drastically changes the relative welfare performance of commitment, simple rules and the

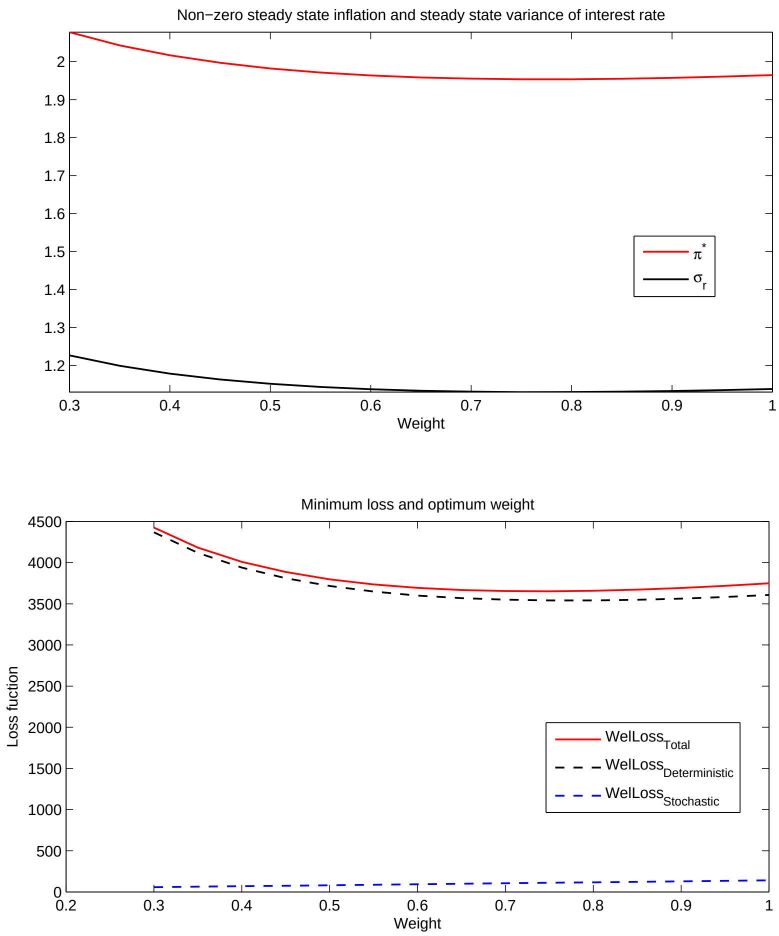

FIGURE 2: IMPOSITION OF ZLB FOR DISCRETION AND PERFECT INFORMATION

withdrawal of information. Now there are substantial gains from commitment of $4-5\%$ consumption equivalent. Simple rules are less able to mimic their optimal counterpart and the loss of information can impose a welfare loss of up to 0.24% consumption equivalent. The form of the optimized simple rules now becomes a price level rule for all cases highlighting an argument in their favour emphasized by Gaspar *et al.* (2010). The reason why the discretionary policy performs so badly with a ZLB constraint can be seen from Figure 2 that implements the first iteration of our algorithm as in Figure 1 for the PI case. Generally speaking it is now well understood that under discretion the policymaker lacks the leverage over private sector behaviour that is possible under commitment from say temporary loosening (or tightening) of monetary policy with promises to reverse this in the future. This in turn greatly inhibits the ability to reduce the unconditional variance of the nominal interest rate when it is penalized by an increasing size of the weight w_r. Consequently to achieve a low probability of hitting the ZLB one needs a larger shift of the nominal interest rate distribution to the right. Whereas under commitment $\pi^* \simeq 0.15$ per quarter corresponding to an annual inflation rate of only 0.6%, under discretion this rises to $\pi^* \simeq 2.0$ or around 8% per year. Our ZLB constraint then results in a *long-run inflationary bias* in addition to the familiar *stabilization bias* highlighted by Currie and Levine (1993), Clarida *et al.* (1999) and others.

These results of imposing the ZLB are fairly uniform across all three information sets. What then are the particular implications of imperfect information then? There are two results to highlight. First under commitment with both optimal policy and optimized rules, the welfare consequences of limiting information to lagged output and inflation is now more significant than before without ZLB considerations. However this is not true for discretion. The withdrawal of information now actually *improves* the welfare outcome by some 1.63% consumption equivalent. The reason for this is that the policymaker now has less opportunity to react opportunistically to current shocks which lowers the variance of the nominal interest rate and makes it easier to satisfy the ZLB constraint. This is a kind of 'tying one's hands' result familiar in other contexts of macroeconomic policy such as exchange-rate policy.

9 Conclusions

This is the first paper to examine optimal policy in an estimated DSGE NK model where informational consistency is applied at both the estimation and policy stages. Our main results can be summarized as follows. First, a result common to all information sets is that only with a ZLB constraint do we see very substantial gains from commitment with optimized rules taking the form of a price level rule in all cases. Second, under commitment information assumptions have significant implications for welfare only when imperfect information is limited to lagged output and inflation. Third, under discretion we observe a 'tying one's hands' effect of such a limitation of information.

There are a number of areas for future research. Our model is very basic with low costs of business cycle fluctuations in the absence of ZLB considerations. If anything we underestimate the costs of imperfect information. It seems therefore worthwhile to revisit the issues raised in the context of a richer DSGE model that includes capital, sticky wages, search-match labour market frictions and financial friction is now the subject of current research. A second avenue for research extend the work to allow the policymaker to have *more* information than the private sector. This satisfies the ICP and would allow the proper examination of the benefits or otherwise of transparency. Finally we assume rational (model consistent) expectations. It would be of interest to combine some aspects of learning (for example about the policy rule) alongside model consistent expectations with imperfect as in Ellison and Pearlman (2008).

References

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

Adam, K. and Billi, R. M. (2007). Discretionary Monetary Policy and the Zero Lower Bound on Nominal Interest Rates. *Journal of Monetary Economics*. Forthcoming.

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.

Ascari, G. and Ropele, T. (2007a). Optimal Monetary Policy under Low Trend Inflation. *Journal of Monetary Economics*, **54**(8), 2568–2583.

Ascari, G. and Ropele, T. (2007b). Trend Inflation, Taylor Principle and Indeterminacy. Kiel Working Paper 1332.

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, 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.

Clarida, R., Galí, J., and Gertler, M. (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*, **37**(4), 1661–1707.

Coenen, G. and Wieland, V. (2003). The Zero-Interest Rate Bound and the Role of the Exchange Rate for Monetary Policy in Japan. *Journal of Monetary Economics*, **50**, 1071–1101.

Coibion, O., Gorodnichenko, Y., and Wieland, J. (2011). The Optimal Inflation Rate in New Keysian Models: Should Central Banks Raise their Inflation Targets in the Light of the ZLB? Mimeo. Presented to the CEF 2011 Conference in San Francisco .

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.

Eggertsson, G. (2006). The Deflation Bias and Committing to Being Irresponsible. *Journal of Money, Credit and Banking*, **36**(2), 283–322.

Eggertsson, G. and Woodford, M. (2003). The Zero Interest-Rate Bound and Optimal Monetary Policy. *Brooking Papers on Economic Activity*, **1**, 139–211.

Ellison, M. and Pearlman, J. G. (2008). Saddlepath Learning. Mimeo and forthcoming, *Journal of Economic Theory*.

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.

Fernandez-Villaverde, J., Rubio-Ramirez, J., Sargent, T., and Watson, M. W. (2007). ABC (and Ds) of Understanding VARs. *American Economic Review*, **97**(3), 1021–1026.

Gaspar, V., Smets, F., and Vestin, D. (2010). Is Time Ripe for Price Level Path Stability? In P. L. Siklos, M. T. Bohl, and M. E. Wohar, editors, *Challenges in central banking: the current institutional environment and forces affecting monetary policy*. Cambridge University Press.

Gavin, W. T. and Keen, B. D. (2011). The Zero Lower Bound and the Dual Mandate. Mimeo. Presented to the CEF 2011 Conference in San Francisco .

Gesell, S. (1934). *The Natural Economic Order*. Free-Economy Publishing Co., Phlip Pye, San Antonio.

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.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. Macmillan, New York.

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 Pierse, 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 .

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

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 Macroeeconomic Models Embodying Rational Expectations. *Manchester School*, **51**, 235–249.

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

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 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[18] of the model about the non-stochastic steady state zero-growth[19], zero-inflation is given by

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

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

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

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

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

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

$$w_t - p_t = mu_t^L - mu_t^C \tag{A.6}$$

$$mc_t = w_t - p_t - a_t \tag{A.7}$$

$$y_t = a_t + l_t \tag{A.8}$$

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, Δ_t, disappears up to a first order approximation and therefore does not enter the linear dynamics. Finally, shock processes and the Taylor rule are given by

$$\begin{aligned}
g_{t+1} &= \rho_g g_t + \epsilon_{g,t+1} \\
a_{t+1} &= \rho_a a_t + \epsilon_{a,t+1} \\
msper_{t+1} &= \rho_{ms} msper_t + \epsilon_{msper,t+1} \\
ms_t &= msper_t + \epsilon_{mstra,t} \\
\pi_{tar,t+1} &= \rho_a \pi_{tar,t} + \epsilon_{tar,t+1} \\
r_t &= \rho_r r_{t-1} + (1-\rho_r)\theta(E_t\pi_{t+1} - \rho_{tar}\pi_{tar,t}) + \epsilon_{e,t}
\end{aligned}$$

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

[18]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.

[19]With growth we simply replace β 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	σ	Normal	1.50	0.375
Price indexation	γ	Beta	0.50	0.15
Calvo prices	ξ	Beta	0.50	0.10
Consumption habit formation	h_C	Beta	0.70	0.10
Preference parameter	ϱ	Beta	0.50	0.20
Adaptive expectations				
Error adjustment - $E_{f,t}\bar{q}^a_{t+1}$	μ_1	Beta	0.50	0.20
Error adjustment - $E_{h,t}u^a_{c,t+1}$	μ_2	Beta	0.50	0.20
Error adjustment - $E^a_{h,t}[\pi_{t+1}]$	μ_3	Beta	0.50	0.20
Proportion of rational households	λ_h	Beta	0.50	0.20
Proportion of rational firms	λ_f	Beta	0.50	0.20
Interest rate rule				
Inflation	θ_π	Normal	1.50	0.25
Output	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
Inflation target	ρ_{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 8: Prior Distributions

	AI				II			
Parameter	Model GH	Model H	Model G	Model Z	Model GH	Model H	Model G	Model Z
σ	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]
γ	0.26 [0.08:0.43]	-	0.27 [0.08:0.46]	-	0.22 [0.07:0.35]	-	0.23 [0.08:0.38]	-
ξ	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]
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]	-	-
ϱ	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]
Interest rate rule								
θ_π	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]
θ_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]
ρ_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]
AR(1) coefficient								
ρ_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]
ρ_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]
ρ_{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]
ρ_{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]
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]
$sd(\epsilon_g)$	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]]
$sd(\epsilon_{ms})$	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]
$sd(\epsilon_{targ})$	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]
Standard deviation of I.I.D. shocks/mearsument 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]
$sd(\epsilon_e)$	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]
$sd(\epsilon_\pi)$	-	-	-	-	-	-	-	-
$sd(\epsilon_y)$	-	-	-	-	-	-	-	-
Price contract length								
$\frac{1}{1-\xi}$	3.57	4.17	2.78	3.13	3.70	4.00	3.03	3.23
LL and posterior model odd								
LL	-96.95	-92.85	-103.77	-99.48	-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

TABLE 9: 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.

C Optimal Policy Under Imperfect Information: Further Details

Pearlman (1992)[20] shows that optimal policy is certainty equivalent in the sense that all the rules under imperfect information correspond to those under perfect information, but with $z_{t,t}$ and $x_{t,t}$ replacing $z_t,\ x_t$. In particular, for the fully optimal rule p_{2t} then depends only on past values $\{z_{s,s},\ x_{s,s} :\ s < t\}$, so that $p_{2t} = p_{2t,t} = p_{2t,t-1}$. (76) is then derived as follows:

$$x_{t,t} + Nz_{t,t} + N_p p_{2t} = 0 \qquad x_t - x_{t,t} = A_{22}^{-1}A_{21}(z_t - z_{t,t}) \tag{C.1}$$

where $N_p = 0$ for TCT and SIM, and N is dependent on which rule is in place, and the second equation was derived earlier. After taking expectations of each of these at $t-1$, it then follows that we can write

$$m_t - m_{t,t-1} = D(z_t - z_{t,t-1}) + v_t + (E - D)(z_{t,t} - z_{t,t-1}) \tag{C.2}$$

using the previous definitions of D and E. Now assume that

$$z_{t,t} - z_{t,t-1} = J_1(D(z_t - z_{t,t-1}) + v_t) \tag{C.3}$$

which will be verified shortly. It then follows that

$$m_t - m_{t,t-1} = (I + (E - D)J_1)(D(z_t - z_{t,t-1}) + v_t) \tag{C.4}$$

and hence the updated value $z_{t,t}$ using the measurement m_t is given by

$$\begin{aligned} z_{t,t} - z_{t,t-1} &= PD^T(DPD^T + V)^{-1}(I + (E - D)J_1)^{-1}(m_t - m_{t,t-1}) \\ &= PD^T(DPD^T + V)^{-1}(D(z_t - z_{t,t-1}) + v_t) \end{aligned} \tag{C.5}$$

Finally Pearlman (1992) shows that $E[(z_t - z_{t,t})z_{s,s}] = 0,\ s \leq t$. This enables us to rewrite the welfare loss in the form of (78), and to obtain its value in (81) using (79).

[20]This paper conducts the analysis in continuous time, but it is straightforward to translate the results and proofs to discrete time.