Explaining Excess Volatility in Bond Yields: Limited Information and Learning in Bond Markets*

Meredith Beechey†

October 15, 2004

Abstract

Long interest rates appear both excessively volatile and excessively sensitive to news and monetary policy innovations compared to the predictions of standard macroecomic models with constant steady state and fully informed agents. When a central bank operates a time-varying inflation target but does not communicate its value, bond markets may instead learn about the target from macroeconomic news and monetary policy movements. Employing a forward-looking model of the macroeconomy paired with the expectations hypothesis of the term struture, I show that bond-market learning about the unknown inflation target imparts additional volatility to interest rates at all maturities. Calibrations suggest this accounts for one tenth to a fifith of the observed volatility in 10 year bonds. The paper also shows that learning raises the covariance between long interest rates and transitory shocks, doubling the expected coefficients from regressions of long rates on inflation surprises and policy rate innovations.

*I am grateful to Tore Ellingsen, Nils Gottfries, Chad Jones, Torsten Persson, David Romer and John C. Williams for valuable comments.

†Department of Economics, University of California, Berkeley. Email: mbeechey@econ.berkeley.edu

1 Introduction

Bond markets exhibit puzzling behaviour in several respects. Long interest rates are more volatile than standard macroeconomic models would predict when paired with the pure expectations hypothesis of the term structure, as is the spread between long and short interest rates (Campbell and Shiller, 1991). In the wake of Shiller (1979) and Singleton's (1980) initial findings of excess volatility, an extensive literature has developed to address this behaviour. Some have turned to non-stationarity in short rates or time variation in term- and risk premiums as explanations for the failure of the expectations hypothesis.[1] Others have pointed to time-varying monetary policy as a way to reconcile the data with theory (Fuhrer (1996) and Rudebusch (1995)) but the issue remains open.

In addition to being overly volatile, long interest rates are sensitive at long horizons to current events, what may be dubbed the excess sensitivity puzzle. Interest rates on bonds as long as 30 years react on average positively and significantly to current monetary policy innovations (Cook and Hahn (1989), Kuttner (2001) and Ellingsen and Söderström (2004)). Forward rates up to 15 years ahead respond to today's news about inflation and output and exhibit as much volatility at long horizons as at short (Gürkaynak, Sack and Swanson (2003)).

Jointly, these behaviours are puzzling when judged against the benchmark predictions of a macroeconomic model with constant steady state in which shocks are transitory and agents are forward-looking and fully-informed. Even models incorporating a fair degree of backward-lookingness and persistence have difficulty reproducing the lengthy response of long rates.[2] A number of authors have pointed to the role that imperfect information may play in the sensitivity of long interest rates, in particular asymmetric information between the central bank and bond markets about economic fundamentals and preferences (see Romer and Romer (2000), Ellingsen and Söderström (2001) and Gürkaynak et al. (2003)). That theme is taken up here.

The aim of this paper is to address two questions. First, does asymmetric information and the sensitivity of long interest rates have any bearing on the puzzle of excess volatility and

[1] Tests of excess volatility are sensitive to the time series properties of short rates, with variance bounds tests misleading when short rates are non-stationary (Flavin 1981, Marsh and Merton, 1986, Cushing and Ackert 1994) and any amount of volatility can be explained with enough term premium variation.

[2] In the partially backward-looking model of Rudebusch (2002), the impulse response of forward rates to shocks dies out completely within 10 years, implying only modest responses of 10 year bonds.

if so, how? Second, does time-varying monetary policy and a departure from the standard constant steady state model also play a role in explaining the excess sensitivity puzzle?

I present a simple New Keynsian model with a time-varying inflation target and posit two different information scenarios, one in which the inflation target is communicated to bond market participants and one in which it is not. In the latter, bond markets learn about the inflation target from observable variables and update upon new information. It is the updating and revision of inflation expectations inherent in this kind of learning that will be key to several of the results in paper. Paired with the expectations hypothesis, the model yields expressions for the volatility of bonds under the two scenarios and can be used to address the overreaction of forward rates and long interest rates to new information. The framework encompasses a useful counterfactual, that of communicated but time-varying policy preferences. This device allows one to distinguish the implications for volatility and overreaction of moving from a constant to a non-stationary inflation target separately from the implications of *learning* about the inflation target.[3]

The adaptive learning mechanism used in this paper assumes that agents learn about unknown state variables via a linear updating algorithm employing their own forecasts errors. This follows recent literature on learning in macroeconomics, of which a good overview is given by Evans and Honkapohja (2001). In short, bond markets lack key information about shocks to the economy and policy preferences and derive a signal about these from unexpected macroeconomic news or monetary policy innovations. Unlike Ellingsen and Söderström (2001), who assume that only one variable at a time is unobservable, in this paper there is a true and repeated signal extraction problem which requires bond markets to decompose information about inflation into permanent and transitory elements.

Of relevance to the bond volatility results in this paper, Honkapohja and Mitra (2003) show that an economy exhibits excess volatility when memory is bounded, by which is meant that an endogenous variable has greater variance when learning about structural parameters cannot converge to the rational expectations equilibrium. Orphanides and Williams (2003) also employ finite-memory, constant-gain learning to show that learning can result in pronounced swings in inflation expectations following large transitory shocks despite parameter stability

[3] In this paper I do not attempt to model the central bank's side of the information problem or their learning behaviour.

in the macroeconomy. This paper is in a similar spirit but extends the analysis to the case when constant-gain learning is warranted by time-variation in the unknown state and draws out the implications for financial market behaviour.

The key findings are as follows. Learning results in heightened sensitivity to transitory shocks and imparts additional variance to forecast errors (Proposition 1) and additional volatility to bonds of all maturities (Propositions 2 and 3). This is the case despite the rate of learning (the gain in a Kalman filter) being optimally calibrated to the true signal-to-noise ratio in the economy. For plausible calibrations of the model and variation in the target, learning adds between a tenth and a fifth to the volatility of a 10 year bond compared to the counterfactual of full information. This is above and beyond the volatility imparted to long interest rates because of time variation in the inflation target.

Turning to the excess sensitivity behaviour described above, the introduction of a non-stationary inflation target is sufficient to qualitatively replicate the documented behaviour of forward rates. The learning story posited by previous authors affects the magnitude rather than the nature of the results. Simulated coefficient estimates from the model of the response of long interest rates to surprise innovations in the policy rate are comparable to the estimated values of Kuttner (2001) and Ellingsen and Söderström (2004) at the long end of the yield curve. The simulations in the paper show that learning raises the covariance between long and short interest rates such that the estimated coefficients are up to double those that would be observed if participants had full information about policy preferences (Proposition 4).

The paper is organized as follows. Section 2 introduces the model and solves for its behaviour with optimal policy. Section 3 introduces the information assumptions, illustrates how forecasting is carried in the two different information scenarios and derives analytical expressions for forecast errors and bond volatility. Section 4 compares calibrations of the model to empirical evidence and addresses the excess sensitivity puzzles described above. Section 5 concludes.

2 A Stylised Macroeconomy

In this section we present a stylised model of the macroeconomy and solve for its behaviour with optimal policy in terms of the shocks arriving in the model. The model consists of a forward-looking New Keynesian model based on agents' optimising behaviour akin to that of Clarida, Gali and Gertler (1999) and Woodford (1999) but with the economic environment modified to include a time-varying inflation target. Whilst more it is more forward-looking than other commonly simulated models of the macroeconomy (Rudebusch (2002), for example, is a popular choice), this characteristic keeps the model analytically tractable and key features transparent.[4] Estrella and Fuhrer (1999) criticise the ability of forward-looking New Keynesian models to match the persistence of inflation but this criticism may be less potent once persistence is added via policy goals.

2.1 Economic Environment

The model is summarised by the following equations, both of which have their roots in the microeconomic foundations of dynamic general equilibrium theory.

$$\pi_t - \pi_t^* = \beta E_t \left(\pi_{t+1} - \pi_{t+1}^*\right) + \lambda x_t + u_t \tag{1a}$$

$$x_t = -\gamma[i_t - E_t\left(\pi_{t+1}\right)] + E_t\left(x_{t+1}\right) + g_t \tag{1b}$$

where π_t is inflation, π_t^* the time-varying inflation target, x_t the output gap (defined as the log deviation of output around potential) and i_t the policy controlled nominal short interest rate. The disturbance terms obey the following laws of motion

$$u_t = \rho u_{t-1} + \hat{u}_t \text{ where } \hat{u}_t \sim iid(0, \sigma_{\hat{u}}^2) \text{ and } \rho \in [0, 1) \tag{2}$$

$$g_t = \mu g_{t-1} + \hat{g}_t \text{ where } \hat{g}_t \sim iid(0, \sigma_{\hat{g}}^2) \text{ and } \mu \in [0, 1) \tag{3}$$

where $\hat{u}_t$, $\hat{g}_t$ are independent of one another, ie: $E\left[v_t v_t'\right] = \begin{bmatrix} \sigma_{\hat{u}}^2 & 0 \\ 0 & \sigma_{\hat{g}}^2 \end{bmatrix}$ where $v_t' = [\hat{u}_t \ \hat{g}_t]$.

The aggregate demand curve (1a) is derived from the log-linearised consumption Euler

[4] Hybrid variants of such models which include both forward- and backward-looking elements can generally not be solved analytically but must rely on numerical methods (see Söderlind (1999)).

equation that solves the consumption-saving decision of the representative household. The aggregate supply equation (1b) is the log-linear approximation of the aggregate firm pricing rule that arises from individual firms' optimal pricing decisions given Calvo (1983) staggered price setting. Appendix A sets out the maximisation problems that lead to both (1a) and (1b). The economic environment in which firms optimise is augmented to include a potentially time-varying inflation target as in Smets and Wouters (2003) and Adolfson, Laseen, Linde and Villani (2004). This differs importantly from the world described by Clarida, Gali, Gertler (1999) in which the inflation target is assumed constant endowing their model with a constant nominal steady state. In brief, firms who are unable to re-optimise in a given period index their prices to a combination of past inflation and the inflation target. When indexation adjusts immediately to the new inflation target, the forward-looking nature of the Phillips curve and the tractability of the model are preserved.[5]

The highly persistent and possibly non-stationary nature of inflation in many industrialized countries suggests that a model with constant steady state and mean-reverting nominal short rates may be the wrong benchmark (for persistence see Fuhrer and Moore (1995), for evidence of unit roots see Mishkin (1992), Wallace and Warner (1993)). I choose to model the inflation target as a random walk with innovation variance σ_ε^2. Evidence for non-stationarity of the inflation target comes from several sources. Smets and Wouters (2003) find that a large share of the movement in inflation in both the US and Euro-area economies over the last 20 years can be explained by permanent shifts in a non-stationary process for the inflation target. Kozicki and Tinsley (2003) have similar success in describing the evolution of long run inflation expectations in the US by modelling the unknown inflation target as a random walk. In an exercise to back out those counterfactual policy parameters that reconcile the pure expectations hypothesis with movements in US long interest rates, Fuhrer (1996) also

[5] Assuming otherwise, that non-optimised prices are partially indexed to past inflation results in the familiar, partly backward-looking Phillips curve. The assumption that price-setters know the value of the inflation target is strong given the later scenario in which bond market participants must learn about the target. However, when price-setters' inflation expectations are also formed through adaptive learning, as for example in Preston (2002), the evolution of the macroeconomy becomes substantially more complicated and detracts from the exposition of later results.

finds that the implied series for the inflation target contains a unit root.[6]

$$\pi_t^* = \pi_{t-1}^* + \varepsilon_t \quad where \ \ \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \tag{4}$$

The specification in equation (4) has the advantage that it does not presume knowledge of the number or type of potential policy regimes, unlike when the target is modelled as a Markov switching process. One shortcoming though is that the process is evidently unbounded at long horizons. However, for sufficiently small values of σ_ε^2 the target tends to stay within plausible bounds (empirical evidence suggests a standard deviation of around ± 1 per cent per decade).[7] It also captures the idea that an apparently stable monetary policy regime may continue to exhibit small adjustments in its relative preferences and inflation target. The distribution of innovations to the target can be chosen to permit large, infrequent changes in the target with the bulk of the mass associated with very small adjustments.

Permanent shocks to the inflation target are the only type of policy change in the model, although to a first approximation, changes in the relative preference for output stability in the loss function (α) can also be characterised as adjustments in the level of the target. All other structural parameters in the economy are assumed constant and known.

2.2 Optimal Policy with Discretion

We consider the standard optimal policy problem of a central bank aiming to minimise the following loss function with discretion

$$L = -\frac{1}{2} E_t \sum \Psi^j \left(\alpha x_{t+j}^2 + (\pi_{t+j} - \pi_{t+j}^*)^2 \right) \tag{5}$$

where α is the relative preference weight on output stability and $\{\pi_t, \pi_t^*, x_t\}$ as before. The central bank chooses the pair $\{x_t, \pi_t\}$ each period to minimise its loss function and sets the appropriate value of the policy controlled interest rate i_t to achieve this. Given the purely forward-looking nature of the economy, monetary policy has only a contemporaneous effect

[6] For clarity, I do not augment the inflation target process such as accomodation of cost-push shocks as in Kozicki and Tinsley (2003) or adjustment to lagged inflation as in Gürkaynak et al. (2003).

[7] The distinction between a highly persistent, stationary target and one with a unit root is fine. Empirically these are very difficult to distinguish.

and the intertemporal policy optimisation problem reduces to a sequence of static optimisations. This leads to a first order condition representing the standard policy trade-off between the output gap and inflation gap,

$$x_t = -\frac{\lambda}{\alpha}(\pi_t - \pi_t^*). \tag{6a}$$

The model can be solved for y_t and π_t in terms of current shocks (see Appendix B);

$$\pi_t - \pi_t^* = \frac{\alpha}{\lambda^2 + \alpha(1-\beta\rho)} u_t \tag{7}$$

$$x_t = -\frac{\lambda}{\alpha}[\pi_t - \pi_t^*] = \frac{-\lambda}{\lambda^2 + \alpha(1-\beta\rho)} u_t. \tag{8}$$

Clarida et al. (1999) describe this policy as one of "leaning against the wind", with the central bank choosing how much of the inflation shock to offset in any period according to its preferences and the parameters of the economy. Larger values of α imply a larger inflation gap for a given shock u_t. Higher serial correlation in inflation shocks also raises the multiplier in (7), ie: $\frac{\partial\delta}{\partial\alpha} > 0$, $\frac{\partial\delta}{\partial\rho} > 0$).

The optimal monetary policy reaction function takes a familiar form for this class of model, resembling a Taylor rule in the sense that the policy controlled interest rate responds to the current inflation- and output gap. Here the nominal instrument is also pegged at the level of the current inflation target (the constant real interest rate is subsumed in the linearisation);

$$i_t = E_t\left(\pi_{t+1}^*\right) + \left[1 + \frac{\lambda(1-\rho)}{\alpha\gamma\rho}\right] E_t\left[\pi_{t+1} - \pi_{t+1}^*\right] + \frac{g_t}{\gamma}$$

Note that the coefficient on expected inflation exceeds unity for positive values of λ, α, γ and ρ, a necessary condition for a stabilising rule. Rewriting this in terms of shocks in the model,

$$i_t = \pi_t^* + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta u_t + \frac{g_t}{\gamma} \tag{9}$$

where g_t, can be seen as either the aggregate demand shock or a policy control error.

3 Information Assumptions and Learning

The key relationships of the economy (equations (7), (8) and (9)) can be summarised as

$$
\begin{aligned}
\pi_t &= \pi_t^* + \delta u_t \quad \text{where } \delta = \frac{\alpha}{\lambda^2 + \alpha(1-\beta\rho)} \\
\pi_t^* &= \pi_{t-1}^* + \varepsilon_t \\
\text{and} \quad i_t &= \pi_t^* + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta u_t + \frac{g_t}{\gamma}
\end{aligned}
$$

where u_t, g_t and ε_t have the properties described in Section 2.1. In this environment, forecasting the nominal short rate is a matter of forecasting the time varying inflation target and macroeconomic shocks in the economy.

We now present our information assumptions. The central bank is assumed to know the structure of the economy at time t and can observe all current variables and shocks but has no advantage over bond market analysts in forecasting them. That is, the central bank possesses potentially superior information about the current state of the economy which may result in more accurate forecasts.[8] Bond market participants are assumed not to be able to observe shocks but we consider two different scenarios regarding their information about the inflation target;

i) Full information - the inflation target is communicated by the central bank every period and thus is fully observable to bond market participants. Shocks, whilst not observable, can be perfectly inferred from a combination of inflation, the inflation target and the nominal policy short rate.

ii) Limited Information - the inflation target is not communicated. Thus bond market participants are unable to accurately decompose observed inflation into its permanent policy and transitory shock components. In all other respects bond market actors are homogeneously well informed, knowing the structure and parameterization of the economy as well as the central bank's preference for output stability. Furthermore, they believe correctly that the inflation target follows a random walk and know the relative variance of innovations to the target and aggregate supply shocks, $\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}$.

[8] This is consistent with the evidence provided by Romer and Romer (2000) of the superiority of Federal Reserve information due to better data processing.

The first scenario is clearly information rich, more so than would be expected if a central bank were conducting policy with a time-varying inflation target, but it will later serve as a useful counterfactual benchmark. It also incorporates the case of a central bank conducting explicit, constant inflation targeting.

The second scenario is a more realistic depiction of policy communication for most central banks and requires that bond markets learn about shifts in policy preferences over time in order to form long-run inflation expectations. When the inflation target is time-varying and non-stationary it becomes optimal to place greater weight on recent observations and discount older observations.[9]

The timing of the model is that all shocks ($\hat{u}_t$, $\hat{g}_t$ and ε_t) are realised at the beginning of period t and taking these into account the central bank sets its policy interest rate so as to affect the outcomes $\{\pi_t, x_t\}$ in the same period. If it is communicated, the inflation target is announced at the beginning of the period.

3.1 Forecasting inflation and nominal short interest rates

In the following section, we illustrate how bond market participants form their inflation expectations and forecast the nominal short interest rate. In Section 3.2 we take a special case of the general problem leading to several propositions about forecast errors and the interest rate on a zero coupon m-period bond.

Full Information Forecasting is straightforward for the case when bond markets know the inflation target at time t. Denote the current information set as Ω_t^{FI}, which includes all information up to and including period t. Given the random walk property of the inflation target, the optimal forecast j-periods ahead conditional on Ω_t^{FI} (denoted $\pi_{t+j/t}^{*FI}$) is

$$\begin{aligned} \pi_{t+j/t}^{*FI} &= \pi_t^* + E_t\left(\sum\nolimits_{i=1}^{j} \varepsilon_{t+i} \mid \Omega_t^{FI}\right) \\ &= \pi_t^* \quad \text{for all } j \geqslant 1. \end{aligned} \tag{10}$$

[9] Such a strategy could also be motivated by arguing that agents suffer from finite memory, as in Orphanides and Williams (2003), or fear structural change and are ever alert to the possibility. Generally, however, the optimal learning strategy when there is no strutural change is to allow the sample to grow indefinitely with equal weights.

The optimal projection of inflation, $\pi^{FI}_{t+j/t}$, j periods ahead and conditional on Ω^{FI}_t is found by leading (7) and employing the serial correlation of u_{t+j};

$$\begin{aligned} \pi^{FI}_{t+j/t} &= \pi^{*FI}_{t+j/t} + \delta u_{t+j/t} \\ &= \pi^*_t + \delta \rho^j u_t \qquad \text{for all } j \geqslant 1 \end{aligned} \tag{11}$$

Similarly leading the policy reaction function in (**??**) and substituting the optimal projections $\pi^{*FI}_{t+j/t}$, $\pi^{FI}_{t+j/t}$, $u_{t+j/t}$ and $g_{t+j/t}$ yields

$$\begin{aligned} i^{FI}_{t+j/t} &= \pi^{*FI}_{t+j/t} + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right] \delta u_{t+j/t} + \frac{1}{\gamma} g_{t+j/t} \\ &= \pi^*_t + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right] \delta \rho^j u_t + \frac{1}{\gamma}\mu^j g_t \quad \text{for all } j \geqslant 1 \end{aligned} \tag{12}$$

From (12) we see that the effect of transitory aggregate supply or aggregate demand shocks on the predicted path of short rates dies out geometrically. If shocks are serially uncorrelated, ie: $\rho = \mu = 0$, the forward rate for horizons $j \geqslant 1$ is simply today's inflation target. A positive shock to the inflation target is interpreted as permanent and raises the forecast of the nominal short rate at all horizons.

Limited Information In order to forecast inflation and the nominal short rate in the limited information scenario, bond market participants need to learn the value of the inflation target. We assume they do so recursively, employing a linear algorithm to update their estimate of the unobserved state variables via their forecast errors. This type of learning has been treated in great detail by Evans and Honkapohja (2001) and is a straightforward application of a Kalman filter to the specific state space of this model (see Hamilton 1994 chapter 13 for a thorough discussion).

The stylised economy described above can be given a state-space representation in which inflation and the nominal short rate are observable variables whilst the inflation target and aggregate supply and demand shocks are unobservable state variables.[10]

[10] The output gap (x_t) is unobservable in this setup and we assume that agents do not employ it as part of their filtering program to extract π^*_t. However, it could be broken into announced output and unobservable potential output where agents would use an appropriate updating filter to extract potential output - much as is done today by professional economists.

Observation equations:

$$\begin{bmatrix} \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 1 & \delta & 0 \\ 1 & \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta & \frac{1}{\gamma} \end{bmatrix} \begin{bmatrix} \pi_t^* \\ u_t \\ g_t \end{bmatrix} \tag{13}$$

State equations:

$$\begin{bmatrix} \pi_{t+1}^* \\ u_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \mu \end{bmatrix} \begin{bmatrix} \pi_t^* \\ u_t \\ g_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ \hat{u}_{t+1} \\ \hat{g}_{t+1} \end{bmatrix} \tag{14}$$

where ε_{t+1}, $\hat{u}_{t+1}$ and $\hat{g}_{t+1}$ are distributed *iid* as before.[11]

The optimal Kalman updating algorithm is

$$\begin{bmatrix} \pi_{t+1/t+1}^* \\ u_{t+1/t+1} \\ g_{t+1/t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \mu \end{bmatrix} \begin{bmatrix} \pi_{t/t}^* \\ u_{t/t} \\ g_{t/t} \end{bmatrix} + \begin{bmatrix} \kappa_{\pi,\pi} & \kappa_{\pi,i} \\ \kappa_{u,\pi} & \kappa_{u,i} \\ \kappa_{g,\pi} & \kappa_{g,i} \end{bmatrix} \begin{bmatrix} \pi_t - \pi_{t/t-1}^{LI} \\ i_t - i_{t/t-1}^{LI} \end{bmatrix} \tag{15}$$

where $\begin{bmatrix} \kappa_{f,\pi} & \kappa_{f,i} \end{bmatrix}$ $f = \pi, u, g$ are the steady state Kalman gains.[12] Individuals update their estimate of the inflation target by attributing a constant fraction of the forecast error in inflation and the nominal short rates to the state variables. In the following section we will illustrate how such gain is related to fundamental variances in the model.

This system nests three different learning strategies. When the first equation in (13) is treated as the only observation equation, agents learn about the inflation target through the signal contained in inflation and its forecast errors. When the second equation is treated in isolation, bond markets update their estimates of the inflation target and shocks in the economy through unexpected innovations to the nominal short rate. This corresponds to the ideas of Romer and Romer (2000) and Ellingsen and Söderström (2003) in which actions of the central bank are a valuable signal because of the asymmetric information they possess. The latter assume a simplified environment in which *either* the policy position *or* a macroeconomic

[11] For the Kalman filter to be the minimum variance estimator of the unknown states, the errors need to be distributed normally. For all other distributions it is the best *linear* estimator.

[12] To possess steady state values, the eigenvalues of the coefficient matrix in (14) must be in or on the unit circle. This is closely related to Evans and Honkapojha's (2001) concept of expectational stability.

shock is unobservable but not both at the same time. Thus, the signal extraction problem is trivial as the missing information is revealed immediately and completely upon observation of the central bank's policy movement.[13]

In this model, the policy short rate is in fact a noisier signal of the inflation target due to the additional variation contributed by g_t and, depending on the parameters of the economy, additional amplitude in $\left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta$. Gurkaynak et al and Kozicki and Tinsley (2003) also assume learning via the short rate because in their more backward-looking models it may be a more timely signal of policy change. The third learning strategy is to employ both observation equations in tandem. Bond markets almost surely gain information from both sources - confirmed by their apparent overreaction to both inflation news and monetary policy innovations - but the timing of policy announcements rarely coincides with inflation releases as they do in this model.

Forecasting is still relatively straightforward and the optimal projections take a similar form to the full information case. From the Law of Iterated Expectations we have that

$$E_t \begin{bmatrix} \pi^*_{t+j} \\ u_{t+j} \quad \mid \Omega_t^{LI} \\ g_{t+j} \end{bmatrix} = \begin{bmatrix} \pi^*_{t+j/t} \\ u_{t+j/t} \\ g_{t+j/t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \mu \end{bmatrix}^j \begin{bmatrix} \pi^*_{t/t} \\ u_{t/t} \\ g_{t/t} \end{bmatrix}. \tag{16}$$

Thus the optimal forecast of the target j periods ahead conditional on Ω_t^{LI} is

$$\pi^{*LI}_{t+j/t} = \pi^*_{t/t} + E_t\left(\sum\nolimits_{i=1}^{j} \varepsilon_{t+i} \mid \Omega_t^{LI}\right) = \pi^*_{t/t} \quad \text{for all } j \geqslant 1 \tag{17}$$

and the optimal projection of inflation becomes

$$\pi^{LI}_{t+j/t} = \pi^{*LI}_{t+j/t} + \delta u^{LI}_{t+j/t} = \pi^*_{t/t} + \delta\rho^j u_{t/t} \qquad \text{for all } j \geqslant 1. \tag{18}$$

The predicted path of short rates resembles that for full information but state values are

[13] Technically, the signal-to-noise ratio is either zero or infinity, implying respectively a gain of zero or one.

replaced by their unobserved estimates,

$$
\begin{aligned}
i^{LI}_{t+j/t} &= \pi^{*LI}_{t+j/t} + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta u^{LI}_{t+j/t} + \frac{1}{\gamma} g^{LI}_{t+j/t} \\
&= \pi^{*}_{t/t} + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta \rho^{j} u_{t/t} + \frac{1}{\gamma}\mu^{j} g_{t/t} \quad \text{for all } j \geqslant 1 \qquad (19)
\end{aligned}
$$

In the following section we will focus on the simplest possible case in order to illustrate analytically the connection between learning, excess sensitivity and volatility. In the empirical section we present evidence showing that the three learning strategies imply very similar results.

3.2 Basic case

In this section we simplify the system in two ways. First by taking the case when bond markets update through inflation forecast errors and second, assuming that disturbances are serially uncorrelated (ie: $\rho = \mu = 0$).

With these assumptions, the state space simplifies to

$$
\begin{aligned}
&\text{Observation equation} \qquad \pi_t = \pi^*_t + \delta \hat{u}_t \qquad (20) \\
&\text{State equation} \qquad \pi^*_t = \pi^*_{t-1} + \varepsilon_t \qquad (21)
\end{aligned}
$$

where $\hat{u}_t$ can now be treated as the serially uncorrelated disturbance of the observation equation.

The optimal linear projections of inflation, the inflation target and nominal forward rates simplify in the full information scenario to

$$
\begin{aligned}
\pi^{FI}_{t+j/t} &= \pi^*_{t+j/t} = \pi^*_t \quad \text{for all } j \geqslant 1 \qquad (22) \\
i^{FI}_{t+j/t} &= \pi^*_t \qquad \text{for all } j \geqslant 1
\end{aligned}
$$

and when the target is not observed, to

$$\begin{aligned} \pi^{LI}_{t+j/t} &= \pi^{*LI}_{t+j/t} = \pi^*_{t/t} \quad \text{for all } j \geqslant 1 \\ i^{LI}_{t+j/t} &= \pi^*_{t/t} \qquad\qquad\quad \text{for all } j \geqslant 1. \end{aligned} \tag{23}$$

In the limited information environment, the task of bond market participants is to decompose equation (20) into transitory shocks to inflation and permanent shifts in the inflation target. In this univariate state-space, the Kalman filter is still the optimal tool for updating the estimate of the unknown inflation target and takes a simpler form,

$$\pi^*_{t/t} = \pi^*_{t-1/t-1} + (1-\phi)(\pi_t - \pi^{LI}_{t/t-1}) \tag{24}$$

where $(1-\phi)$ is the steady state Kalman gain that regulates the proportion of the inflation forecast error attributed to the inflation target. The optimally calibrated gain is a non-linear function of the signal-to-noise ratio, $\phi = \phi\left(\frac{\sigma^2_\varepsilon}{\delta^2\sigma^2_u}\right)$, and bounded between 0 and 1 (see Appendix C for specific functional form). The noisier innovations to the inflation target become relative to aggregate supply shocks, the more any forecast error is attributed to adjustment of the inflation target, that is, as $\frac{\sigma^2_\varepsilon}{\delta^2\sigma^2_u} \to \infty$, $(1-\phi) \to 1$ and $\frac{\sigma^2_\varepsilon}{\delta^2\sigma^2_u} \to 0$, $(1-\phi) \to 0$.

Figure 1 provides a graphical illustration from a simulation of the model of the inferred and actual inflation target, with a signal-to-noise ratio of 0.14 and optimal gain 0.28. The perceived inflation target, $\pi^*_{t/t}$, broadly tracks the actual inflation target but with persistent deviations that reflect the dynamics of adaptive, constant-gain learning. With constant-gain learning, even when the gain is optimally calibrated to the signal-to-noise ratio, the updated estimate of the inflation target will overreact to transitory shocks (u_t) and under-react to true changes in the target (ε_t).

The recursive nature of the updating in equation (24) means that the current estimate of the target can be expressed as a geometric lagged polynomial of the history of observed inflation outcomes,

$$\pi^*_{t/t} = \frac{(1-\phi)}{1-\phi L}\pi_t. \tag{25}$$

3.2.1 Forecast Errors

The one period ahead inflation forecast error is central to updating the perceived target. The nominal short rate forecast error Is a linear combination of the same ingredients and comparing its variance between the full and limited information scenarios gives us a sense of how learning adds noise via the revision of expectations.

Full Information Subtracting the one-period-ahead forecast of inflation at $t+1$ from the actual inflation outcome yields a forecast error in terms of the independent shocks arriving to the inflation target and aggregate supply at $t+1$,

$$\begin{aligned} \pi_{t+1} - \pi^{FI}_{t+1/t} &= \left(\pi^*_{t+1} + \delta \hat{u}_{t+1}\right) - \pi^*_t = \varepsilon_{t+1} + \delta \hat{u}_{t+1} \\ var(\pi_{t+1} - \pi^{FI}_{t+1/t}) &= \sigma^2_\varepsilon + \delta^2 \sigma^2_{\hat{u}} \end{aligned}$$

Similarly, subtracting the optimal forecast of the nominal short rate in (22) from the nominal policy rate dictated by the reaction function (12) (note, with $\rho = \mu = 0$) yields a forecast error which also reflects shocks to aggregate demand,

$$i_{t+1} - i^{FI}_{t+1/t} = \varepsilon_{t+1} + \frac{\lambda}{\alpha\gamma}\delta \hat{u}_{t+1} + \frac{1}{\gamma}\hat{g}_{t+1} \tag{26}$$

This forecast error has variance

$$var(i_{t+1} - i_{t+1/t}) = \sigma^2_\varepsilon + \left(\frac{\lambda}{\alpha\gamma}\right)^2 \delta^2 \sigma^2_{\hat{u}} + \frac{\sigma^2_{\hat{g}}}{\gamma^2}$$

that intuitively is increasing in all innovation variances. The latter two are scaled by the combination of structural and preference parameters that enter the reaction function. A stronger preference for output stability (higher α) reduces the variance of the forecast error; higher α calls for proportionately less movement in the nominal short rate via $\frac{\lambda}{\alpha\gamma}$ for a given shock and is less than fully offset by the widening of the tolerated inflation gap $\left(\delta = \frac{\alpha}{\lambda^2+\alpha}\right)$.

Limited Information Forecast errors can be given similar expression in the limited information scenario by subtracting the one-period-ahead projection of inflation and the nominal short rate in (23) from their respective outcomes at $t+1$. As before, forecast errors are due to innovations at $t+1$ but also incorporate a new term that reflects uncertainty about the state of policy preferences, that is, the gap between the true and inferred values of the inflation target $\left(\pi_t^* - \pi_{t/t}^*\right)$.

$$\pi_{t+1} - \pi_{t+1/t}^{LI} = \varepsilon_{t+1} + \left(\pi_t^* - \pi_{t/t}^*\right) + \delta \hat{u}_{t+1}$$

$$i_{t+1} - i_{t+1/t}^{LI} = \varepsilon_{t+1} + \left(\pi_t^* - \pi_{t/t}^*\right) + \frac{\lambda}{\alpha\gamma}\delta \hat{u}_{t+1} + \frac{1}{\gamma}\hat{g}_{t+1} \tag{27}$$

The lagged polynomial form of $\pi_{t/t}^*$ implies that $\left(\pi_t^* - \pi_{t/t}^*\right)$ is a function of the history of shocks to the economy. This complicates the derivation of the variance somewhat and the reader is referred to Appendix D. The result is stated here,

$$var(i_{t+1} - i_{t+1/t}^{LI}) = \sigma_\varepsilon^2(\frac{1}{1-\phi^2}) + \delta^2\sigma_{\hat{u}}^2\left[\frac{\lambda}{\alpha\gamma}^2 + \frac{(1-\phi)^2}{1-\phi^2}\right] + \frac{\sigma_{\hat{g}}^2}{\gamma^2} \tag{28}$$

Proposition 1 *The variance of the nominal short rate forecast error is unambiguously larger when bond market participants are learning about the inflation target than when the target is perfectly observed. Because*

$$\begin{aligned} \sigma_\varepsilon^2(\frac{1}{1-\phi^2}) + \delta^2\sigma_{\hat{u}}^2\left[\frac{\lambda}{\alpha\gamma}^2 + \frac{(1-\phi)^2}{1-\phi^2}\right] + \frac{\sigma_{\hat{g}}^2}{\gamma^2} &> \sigma_\varepsilon^2 + \frac{\lambda}{\alpha\gamma}^2\delta^2\sigma_{\hat{u}}^2 + \frac{\sigma_{\hat{g}}^2}{\gamma^2} \\ var(i_{t+1} - i_{t+1/t}^{LI}) &> var(i_{t+1} - i_{t+1/t}^{FI}) \end{aligned}$$

for any constant gain $\phi \in (0,1)$. Proof in Appendix D.

Any degree of learning generates additional forecast error variance by enlarging the coefficients attached to σ_ε^2 and $\sigma_{\hat{u}}^2$. The source of this additional variance is revision to $\left(\pi_t^* - \pi_{t/t}^*\right)$ and intuitively the presence of ϕ reflects the dependence of $\pi_{t/t}^*$ on the historical sequence of transitory shocks u_{t-j}, $j = 0, ..., \infty$. The gain determines the extent to which these shocks are attributed to $\pi_{t/t}^*$.

For given σ_ε^2, $\sigma_{\hat{u}}^2$ and $\sigma_{\hat{g}}^2$, $\left(\frac{1}{1-\phi^2}\right)$ is in rising in ϕ whilst $\frac{(1-\phi)^2}{1-\phi^2}$ is falling. However, these variances interact to determine the value of $\phi\left(\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}\right)$ when the gain is calibrated optimally to the signal-to-noise ratio. An increase in the signal-to-noise ratio (that is, a rise in σ_ε^2 for a given value of $\delta^2\sigma^2_{\hat{u}}$) lowers $\left(\frac{1}{1-\phi^2}\right)$ but not enough so to offset the rise in σ^2_ε.

As one approaches the case of a constant inflation target, ie: $\sigma^2_\varepsilon \to 0$ and $\phi\left(\frac{\sigma^2_\varepsilon}{\delta^2\sigma^2_u}\right) \to 1$, the forecast error variances of the the two information scenarios coincide. Forecast errors reflect only the arrival of transitory shocks when long-run inflation expectations are well-anchored by a constant target.

3.2.2 Building Bonds

Assuming that the expectations hypothesis of the term structure holds, it is straightforward to characterise the interest rates on long bonds and derive bond volatility for the different information scenarios. Denoting the interest rate on a zero-coupon bond with maturity m at time t as i^m_t, this interest rate is set as the average expected future short interest rate during the time to maturity plus a term premium,

$$i^m_t = \frac{1}{m}\sum\nolimits_{j=0}^{m-1} i_{t+j/t} + \zeta^m_t$$

where $i_{t+j/t}$ is as before the expected short interest rate j periods ahead and ζ^m_t is the term premium at time t for maturity m. As in Ellingsen and Söderström (2001) and Fuhrer (1996), I do not attempt to model time-variation in the term premium but assume that it is independent of all relevant variables in the model. In the calibrations below, time variation in the inflation target rather than the term premium is used to match historical volatility data, although time variation in term premiums would not change the spirit of the exercise. To build long rates, bond market participants form expectations about the future path of short interest rates based on their current information set, Ω^{FI} or Ω^{LI}. With the real interest rate assumed to be constant, forecasts of shocks and inflation expectations drive movements in the term structure.[14]

Full Information Denote an m-period bond in the full information scenario as $i^{m^{FI}}_t$. Combining the short rate at period t from the central bank reaction function in (**??**) with the

[14] This is broadly consistent with the findings of Ang and Bekaert (2004) who detect very little movement in the real component of the term struture. They also conclude that the majority of movements in long term nominal interest rates are due to changes in expected inflation rather than inflation risk premia.

optimal projection of nominal short rates in (22) yields the following;

$$\begin{aligned} i_t^{m^{FI}} &= \frac{1}{m}\left(i_t + i_{t+1/t}^{FI} + \ldots + i_{t+m-1/t}^{FI}\right) + \zeta_t^m \\ &= \pi_t^* + \frac{1}{m}\left(\frac{\lambda}{\alpha\gamma}\delta\hat{u}_t + \frac{\hat{g}_t}{\gamma}\right) + \zeta_t^m. \end{aligned} \tag{29}$$

The nominal yield is pegged at the level of the current inflation target and reflects transitory shocks, $\hat{u}_t$ and $\hat{g}_t$, only to the extent that they drive the current nominal short rate away from its equilibrium level. As m increases, averaging ensures that the effect of these shocks on longer yields diminishes.

The non-stationary property of the nominal short rate means that variances in levels are unbounded as t goes to infinity. Instead we focus on a standard measure of volatility, the variance (or standard deviation) of the period-to-period change in an m-period bond. Differencing (29) yields an expression for the change in an m-period bond between two periods;

$$i_t^{m^{FI}} - i_{t-1}^{m^{FI}} = \varepsilon_t + \frac{1}{m}\left(\frac{\lambda}{\alpha\gamma}\delta\left(\hat{u}_t - \hat{u}_{t-1}\right) + \frac{\hat{g}_t - \hat{g}_{t-1}}{\gamma}\right) + \zeta_t^m - \zeta_{t-1}^m$$

Two simple features are worth pointing out. First, the change in the bond yield at any maturity is one-for-one with any change in the inflation target. Secondly, we have the intuitive result that the effect of transitory shocks diminishes with maturity. That is,

$$\begin{aligned} \frac{\partial\left(i_t^{m^{FI}} - i_{t-1}^{m^{FI}}\right)}{\partial\varepsilon_t} &= 1 \\ \frac{\partial\left(i_t^{m^{FI}} - i_{t-1}^{m^{FI}}\right)}{\partial\hat{u}_t} &= \frac{1}{m}\frac{\lambda}{\alpha\gamma}\delta \end{aligned}$$

The variance of the change in bond yield obeys the following:

Proposition 2 *The volatility of an m-period bond with full information and serially uncorrelated errors is:*

$$var(i_t^{m^{FI}} - i_{t-1}^{m^{FI}}) = \sigma_\varepsilon^2 + \frac{1}{m^2}\left(2\left(\frac{\lambda}{\alpha\gamma}\right)^2\delta^2\sigma_{\hat{u}}^2 + \frac{1}{\gamma^2}2\sigma_{\hat{g}}^2\right) + \sigma_\varsigma^2$$

Several observations can be made. First, this variance is rising one-for-one with σ_ε^2 and can be interpreted as substitutable for σ_{ς}^2. As σ_{ε}^2 approaches zero, the case of constant inflation targeting and a constant steady state, bond volatility declines until the only source of volatility is the realisation of unforecastable transitory shocks. This has the natural implication that for given variances of transitory shocks, bond volatility with inflation targeting should be less than under other monetary policy regimes.

Secondly, the bond's variance is rising in $\sigma_{\hat{u}}^2$ and $\sigma_{\hat{g}}^2$, both of which are scaled by their respective effect on the current short rate. As expected, the effect of transitory shocks on bond volatility diminishes with maturity. From Appendix D, which derives the same variance in the presence of serially correlated errors, it can be seen that more persistent shocks raises bond volatility. A higher value of ρ raises the tolerated inflation gap for a given aggregate supply shock but has the offsetting effect of reducing the unconditional variance of $(u_t - u_{t-1})$. Higher values of α are associated with less volatile bond rates, as was the case for forecast errors. A stronger preference for output stability results in less pronounced movements of the nominal short rate to counteract inflationary shocks, with the combined coefficient $\left(\frac{\lambda}{\alpha\gamma}\right)^2 \delta^2$ declining in α.

Limited Information An expression for the m-period bond for the limited information case, $i_t^{m^{LI}}$, can be derived in a similar manner. Combining the short rate at period t dictated by the central bank's reaction function in (9) with predicted values of the short rate as in (23) yields

$$\begin{aligned} i_t^{m^{LI}} &= \frac{1}{m}\left(i_t + i_{t+1/t}^{LI} + \ldots + i_{t+m-1/t}^{LI}\right) + \zeta_t^m \\ &= \frac{1}{m}\left(\pi_t^* + \frac{\lambda}{\alpha\gamma}\delta u_t + \frac{1}{\gamma}g_t + (m-1)\pi_{t/t}^*\right) + \zeta_t^m. \end{aligned}$$

The nominal component is pegged to a combination of the true inflation target, π_t^*, and the inferred value, $\hat{\pi}_{t/t}^*$, which is projected for $(m-1)$ periods of the bond. Taking first differences, change in a bond's yield reflect true changes to the target and revisions to the perceived target,

$$i_t^{m^{LI}} - i_{t-1}^{m^{LI}} = \frac{1}{m}\left(\varepsilon_t + (m-1)(\pi_{t/t}^* - \pi_{t-1/t-1}^*) + \frac{\lambda}{\alpha\gamma}\delta(\hat{u}_t - \hat{u}_{t-1}) + \frac{1}{\gamma}(\hat{g}_t - \hat{g}_{t-1})\right)$$

The dynamics of learning complicate derivation of bond volatility so we refer the reader to Appendix E and simply present the result. The Kalman updating algorithm in (24) expresses $(\pi^*_{t/t} - \pi^*_{t-1/t-1})$ in terms of current shocks (ε_t and $\hat{u}_t$) so that we can write

$$\begin{aligned}\frac{\partial\left(i_t^{m^{FI}} - i_{t-1}^{m^{FI}}\right)}{\partial\varepsilon_t} &= 1 - \frac{m-1}{m}\phi \\ \frac{\partial\left(i_t^{m^{FI}} - i_{t-1}^{m^{FI}}\right)}{\partial\hat{u}_t} &= \frac{1}{m}\frac{\lambda}{\alpha\gamma}\delta + \frac{m-1}{m}(1-\phi).\end{aligned}$$

Comparing these to the partial derivatives for full information encapsulates the story of under- and over-reaction of bond yields to inflation news. The term structure underreacts to permanent changes in the target, ε_t, relative to full information but overreacts to transitory disturbances. This is illustrated in Figure 2 using the impulse reponses of predicted short rates to both kinds of shocks These effects interact with the maturity of the bond (the number of periods for which the mistake is projected) and the gain (how severe the learning problem is).

Proposition 3 *The volatility of an m-period bond with limited information is:*

$$var(i_t^{m^{LI}} - i_{t-1}^{m^{LI}}) = A\sigma_\varepsilon^2 + \frac{1}{m^2}\left(B\delta^2\sigma_{\hat{u}}^2 + \frac{1}{\gamma^2}2\sigma_{\hat{g}}^2\right) + \sigma_\varsigma^2$$

where

$$A = \frac{1}{m^2}\left[[1+(m-1)(1-\phi)]^2 + \frac{((m-1)(1-\phi)\phi)^2}{1-\phi^2}\right] < 1$$

$$B = \left[\frac{\lambda}{\alpha\gamma} + (m-1)(1-\phi)\right]^2 + \left[\frac{\lambda}{\alpha\gamma} + (m-1)(1-\phi)^2\right]^2 + \left[\frac{((m-1)(1-\phi)^2\phi)^2}{1-\phi^2}\right] > 2\left(\frac{\lambda}{\alpha\gamma}\right)^2$$

That is, $var(i_t^{m^{LI}} - i_{t-1}^{m^{LI}}) > var(i_t^{m^{FI}} - i_{t-1}^{m^{FI}})$ *for any* $m > 1$ *and any* $\phi\left(\frac{\sigma_\varepsilon^2}{\delta^2\sigma_{\hat{u}}^2}\right) \in (0,1)$.

As in Proposition 2 for the variance of a bond with full information about the target, the variance in Proposition 3 is rising σ_ε^2, $\sigma_{\hat{u}}^2$ and $\sigma_{\hat{g}}^2$. However, there is now an additional interaction with the gain, ϕ, and maturity, m. Intuitively, because bond market participants update their estimate of the target by less than the change in the true target $A < 1$. But systematic overreaction to transitory disturbances results in $B > \frac{1}{m^2}2\left(\frac{\lambda}{\alpha\gamma}\right)^2$. The severity of the under-reaction is captured by ϕ; as the signal grows stronger and $(1-\phi)$ rises, A approaches 1. Likewise, the variance contributed by overreaction is regulated by ϕ; B is also rising in $\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}$ but less rapidly than $A\sigma_\varepsilon^2$.

The variance is declining in m^2 but with a partially offsetting positive effect of maturity through both A and B. As discussed above, the interaction with m reflects the number of periods for which the mistaken estimate is projected.

As the signal-to-noise ratio increases $\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2} \to \infty$, $\phi \to 0$ the bond variances under full and limited information converge.[15] The flip-side of this is that the ratio of bond volatility between the limited versus full information scenarios is largest when the signal-to-noise ratio is weak (see Figure 5). The *relative* reduction in bond volatility achieved by communicating the target is greatest when that target was already close to being stable, an important message for implicit inflation targeters. When $\sigma_{\hat{u}}^2$ is large, the gains to communication also grow.

Finally, when $\sigma_\varepsilon^2 = 0$, the case of explicit inflation targeting, there is no learning behaviour and the full and limited information variances coincide. In this situation, bond volatility is powerfully declining in maturity.

4 Empirical Evidence

4.1 Calibration

We have shown analytically for a simple case how non-stationarity of the inflation target contributes to the variance of bond yields and that learning about an uncommunicated target unambiguously raises that variance. The next question to ask is what are the relative magnitudes of these effects? In order to answer this we present calibrated values of bond volatility and then address the implications of the model for the sensitivity of long interest rates.

Whilst the model is relatively sparse, it has sufficient degrees of freedom in its parameters to make it possible to match, say, the observed volatility of 10 year bonds. I do not claim that the factors in the model are the only sources of bond volatility, as several others are not addressed here (such as term and risk premia, time variation in the real interest rate). Rather, the aim is to assess the *relative* contributions to volatility of the mechanisms described in the

[15] Note that only $A \to 1$. $B \to 2(\frac{\lambda}{\alpha\gamma})^2 + 4\frac{\lambda}{\alpha\gamma}(m-1) + 2(m-1)^2 > 2(\frac{\lambda}{\alpha\gamma})^2$ but the relative size of the variances scale the coefficients appropriately such that the LI variance approaches the FI variance.

model for reasonable parameters that generate bond volatility of a plausible magnitude.

Structural parameters for a forward-looking New Keynsian model are suggested by Clarida, Gali and Gertler (2000) - in the Phillips curve, an elasticity of inflation with respect to the output gap of 0.3 and in the aggregate demand equation a one-to-one relationship between the output gap and the real interest rate. To match the persistence in inflation and output they assume highly serially correlated disturbances in their simulations ($\rho = \mu = 0.9$) which imply that a shock has a half-life of over 6 quarters. This degree of persistence is unlikely to be appropriate here as the non-stationary inflation target accounts for much of the persistence in inflation.[16] Clarida et al do not suggest variances for the transitory shocks in their model and the inflation target is assumed constant.

Rudebusch (2002) estimates a partially backward-looking variant of a New Keynsian model, also assuming a constant steady state, and reports estimated variances of serially uncorrelated aggregate supply and demand shocks. Expressed as shocks to annualised quarterly inflation and the level of the output gap, these variances are approximately 1 and 0.7 respectively. Whilst it is not clear that these are the correct variances for the model at hand, we employ them in the calibrations.

There are few estimates of the variance of innovations to the inflation target. Smets and Wouters (2003) estimate the quarterly innovation variance for a random walk inflation target to be 0.055 (median estimate) for the US between 1973 and 2003, 0.099 for the Euro-area. Kozicki and Tinsley (2003) get a similar estimate of 0.044 using US data from 1960, although they use a dummy variable to account for the changes in the early Volcker years.

To put these estimates in perspective, a quarterly innovation variance of 0.05 implies a standard deviation of approximately 1.5 percentage points in the inflation target over one decade. Likewise, an innovation variance of 0.02 implies a standard deviation of 0.9 percentage points. The basic parameter choices are summarised in Table 1. Note that by assuming $\rho = \mu = 0.5$, the unconditional variances σ_u^2 and σ_g^2 are one third larger than $\sigma_{\hat{u}}^2$ and $\sigma_{\hat{g}}^2$ respectively.[17]

[16] Beechey, Carlsson and Österholm (2004) use the decomposition suggested by this model to re-examine the time series properties of transitory shocks to the economy once a random walk in the inflation target has been filtered out. The serial correlation in the residuals is substantially lower (around 0.4 to 0.7) than typcially needed to have a New Keynsian model fit the data.

[17] Combinations of ($\sigma_{\hat{u}}^2$,ρ) and ($\sigma_{\hat{g}}^2$,μ) can be chosen to fall on an isoquant for given values of σ_u^2 and σ_g^2.

Table 1: Parameter Calibrations

Structural Parameters		Variances	
β (Consumption discount rate)	0.99	$\sigma^2_{\hat{u}}$	1.0
λ (Elasticity of $\pi_t - \pi^*_t$ wrt y_t)	0.3	$\sigma^2_{\hat{g}}$	0.7
γ (Elasticity of output wrt real interest rate)	1	σ^2_{ε}	$(0, 0.7)$
α (Preference in central bank's loss function)	0.5		
ρ, μ (Persistence of transitory shocks)	0.5		

Notes: Variances pertain to annualised quarterly observations (that is, $\pi_t = (\ln p_t - \ln p_{t-1}) * 400$). Thus the variance of innovations to the target is 16 times that discussed in the text.

4.2 Bond Volatility

Table 2 shows the variance of quarterly changes in constant maturity bonds in the United States for the period 1981 to 2004. Bond volatility at all maturities compressed substantially in the 1990s relative to the preceding decade but in all samples behaves broadly as expected with volatility declining with maturity.

Table 2: Volatility of Selected Constant Maturity Treasuries, US, January 1981 to September 2004

Maturity(years)	*Mar* 81 − *Sep* 04	*Mar* 81 − *Dec* 89	*Mar* 90 − *Sep* 04
1	0.74	1.51	0.27
2	0.72	1.35	0.34
5	0.61	1.06	0.33
10	0.46	0.82	0.24
20	*	*	*
30	0.37	0.67	0.15**

Notes: Volatility is calculated as the variance of the quarter-to-quarter change in the reported bond yield. Data are end quarter observations March, June, September, December. *Missing data 1987 to 1994. ** Missing data March 2002 to end of sample Source: Board of Governors H.15 Database, selected constant maturity treasury bonds

Simulated bond volatility from the version of the model in which bond markets learn through inflation forecast errors are shown in Table 3 for various inflation target innovation variances. These have been chosen to roughly match the volatility at the long end of the curve for the periods shown in Table 2 and are correspondingly labelled moderate, high and low. The second column shows the predicted volatility of interest rates assuming a constant inflation target. Bond volatility declines rapidly with maturity leaving very little variance in even a 5 year bond and is plotted against actual bond volatility in Figure 3. Observed bond yields clearly exhibit substantially greater volatility and the benchmark of a constant steady state appears to be a poor approximation.

Table 3: Volatility of Calibrated m-year bonds from the model (Learning via inflation forecast error).

	Constant	Moderate		High		Low	
maturity	$\sigma_\epsilon^2 = 0$	$\sigma_\varepsilon^2 = 0.35$		$\sigma_\varepsilon^2 = 0.7$		$\sigma_\varepsilon^2 = 0.2$	
(years)		*FI*	*LI*	*FI*	*LI*	*FI*	*LI*
1	0.61	0.95	1.18	1.30	1.59	0.80	1.00
2	0.17	0.52	0.69	0.86	1.08	0.37	0.51
5	0.03	0.38	0.46	0.72	0.82	0.23	0.30
10	0.01	0.36	0.40	0.70	0.76	0.21	0.24
20	0.00	0.35	0.37	0.70	0.72	0.20	0.22
30	0.00	0.35	0.36	0.70	0.71	0.20	0.21

Notes: The data in the table are generated with Monte Carlo simulations using 1000 draws of the economy observed for 50 years. The reported numbers are the mean of the sample bond variance over all simulations.

In the subsequent columns, bond volatility under the full information (FI) and limited information (LI) scenarios are reported for the three different calibrations of σ_ϵ^2 (moderate, high and low). There are several points worth making here. First, the contribution of a time-varying inflation target to bond volatility is substantial, even for the counterfactual scenario in which it is fully communicated. Even for the low calibration (an annualised value of σ_ϵ^2 of 0.2 implies a standard deviation of the target of only 0.7 over a decade) it is movement in the target that generates bond volatility once transitory shocks die out. Second, learning contributes additional volatility compared to the full information counterfactual; for a ten year bond around 4 to 6 basis points. For the high scenario, this represents a little less than a tenth of total volatility in the ten year bond, for the low scenario, one fifth. Third, the volatility wedge created by learning diminishes with maturity.

Figure 4 plots the three columns of data in Table 2 against the simulated values from Table 3. The expectations hypothesis paired with a time varying inflation target does a surprisingly good job of matching the volatility of the term structure in the 1980s. For the latter half of the sample the model does a relatively poor job of mimicking the short end of the yield curve, as short term interest rates predicted by the model are more variable than in the data. In part this reflects that the optimal monetary policy reaction function in the model does not incorporate an interest rate smoothing term at a time when policy movements have become smoother, either due to conscious interest rate smoothing or a change in the nature of shocks arriving in the economy.[18] The model does a better job for longer maturities, and comparing Figures 3 and 4 it is clear that some persistent variance in the economy is needed to match the volatility that remains in very long interest rates.

These calibrations have assumed that variation in the inflation target accounts for the lions share of bond volatility. If time variation in the term premium accounts for some of the volatility in long interest rates, this would not lower the estimated contribution made by learning. As the signal-to-noise ratio for an uncommunicated target grows weaker, the *relative* contribution of learning to volatility rises. This can be seen more clearly in Figure 5 which plots the ratio of the variance of a 10 year bond between the limited and full information scenarios for a richer range of σ_ϵ^2 calibrations. The ratio of variances is greatest when the signal-to-noise is weakest.

The results at the long end of the term structure are not particularly sensitive to alternative parameter calibrations. Lowering the central bank's preference for output stability (α) to 0.2 raises the volatility of short interest rates (1 and 2 years) because of the central bank's greater willingness to create output deviations to restore inflation to the target. Lowering the persistence of the transitory shocks to $\rho = \mu = 0.3$ lowers the volatility of interest rates at all maturities, although again, the effect is most pronounced at the short end when transitory shocks feature more heavily in interest rate forecasts. For the 10 year bond, the difference is only a matter of 2 basis points. When the common serial correlation parameter is raised to 0.8, persistence raises the variance of a 10 year bond to 0.11 even with a constant inflation target. However, this degree of autocorrelation seems implausibly high, causing the variance of 1 and 2 year bonds to reach 3.8 and 1.9 respectively. Lastly, lowering the elasticity of output with respect to the real interest rate (γ) from 1 to 0.5 so that the central bank needs to move interest rates by more to achieve the same effect on inflation has the effect of significantly raising volatility at the short end (1 and 2 year bonds) without imparting much additional variance at or beyond the 10 year bond.

In Section 3, it was noted that bond markets may take advantage of the signal in contained in the policy rate to infer the inflation target. Table 4 presents bond variances for this case

[18] Introducing an ad hoc smoothing term in the monetary policy reaction function lowers the immediate response of the policy controlled short rate to current shocks and thus lowers the variance in 1 and 2 year bonds. Intuitively, it has little effect on longer bonds.

Table 4: Volatility of Calibrated m-year bonds from the model (Learning via nominal short rate forecast error).

m	$\sigma_\epsilon^2 = 0$	$\sigma_\epsilon^2 = 0.35$		$\sigma_\epsilon^2 = 0.7$		$\sigma_\epsilon^2 = 0.2$	
$(years)$		FI	LI	FI	LI	FI	LI
1	0.61	0.96	1.28	1.31	1.70	0.80	1.07
2	0.17	0.52	0.76	0.88	1.16	0.37	0.57
5	0.03	0.37	0.49	0.74	0.87	0.23	0.32
10	0.01	0.35	0.42	0.72	0.78	0.21	0.26
20	0.00	0.35	0.38	0.71	0.74	0.20	0.23
30	0.00	0.35	0.37	0.71	0.72	0.20	0.22

Notes: The data in the table are generated with Monte Carlo simulations using 1000 draws of the economy observed for 50 years. The reported numbers are the mean of the sample bond variance over all simulations.

using the same parameters as above. Because the policy short rate is a noisier signal of the inflation target in this model, learning is less accurate than when agents focus on just inflation releases. Correspondingly, bond volatility is slightly higher for all maturities for the limited information scenario.

4.3 Two Sensitivity puzzles

1. **The Volatility and Sensitivity of Forward Rates**

Gürkaynak et al (2003) document two empirical facts about the behaviour of forward rates. First, the volatility of forward rates is not downward sloping with horizon; 10 and 15 year forward rates are as volatile as 2 year forward rates. Second, forward rates respond at very long horizons to current news, in particular news about inflation, with the magnitude of the long response often similar to that of one year rate. Prompted by the observation that forward rates derived from inflation-indexed debt in the United States are less sensitive to news surprises,the authors suggest that the sensitivity of forward rates may reflect learning about an unknown, time-varying inflation target.

In fact, both features are predictions of the model with a random walk inflation target, with or without the informational problem. With a non-stationary process for the inflation target, long run inflation expectations are not anchored to a fixed point. Using the simple analytical case built above, we can calculate the volatility of forward rates as follows;

$$\begin{aligned} i^{FI}_{t+j/t} - i^{FI}_{t+j/t-1} &= \pi^*_t - \pi^*_{t-1} = \varepsilon_t \\ var(i^{FI}_{t+j/t} - i^{FI}_{t+j/t-1}) &= \sigma^2_\varepsilon \quad \forall \ j > 1 \end{aligned}$$

$$\begin{aligned} i^{LI}_{t+j/t} - i^{LI}_{t+j/t-1} &= (1-\phi)(\pi_t - \pi^{LI}_{t/t-1}) = (1-\phi)k_t \\ var(i^{FI}_{t+j/t} - i^{FI}_{t+j/t-1}) &= (1-\phi)^2\sigma^2_k = \sigma^2_\varepsilon \quad \forall \ j > 1 \end{aligned}$$

These variances are not only constant over horizon j but with the gain calibrated optimally to the signal-to-noise ratio, they take the same value. The flat profile of forward rate volatility contrasts strongly with the predictions of a partly backward-looking model such as in Rudebusch (2002) with constant steady state. Whilst such a model is able to generate substantial persistence in inflation, the purely transitory nature of the shocks means that the volatility profile is strongly downward sloping. The response of forward rates to an inflation surprise $(\pi_t - \pi_{t/t-1})$ as would be estimated by the following regression

$$i_{t+j/t} - i_{t+j/t-1} = \alpha + b_{1,j}\frac{\pi_t - \pi_{t/t-1}}{stdev(\pi_t - \pi_{t/t-1})} + \epsilon_{t,j}$$

yields the following coefficients

$$\begin{aligned} b^{FI}_{1,j} &= \frac{\sigma^2_\varepsilon}{\sqrt{\sigma^2_\varepsilon + \delta^2\sigma^2_{\hat{u}}}} \quad \forall \ j > 1 \\ b^{LI}_{1,j} &= (1-\phi)\sqrt{\frac{\sigma^2_\varepsilon}{(1-\phi)^2}} = \sigma_\varepsilon \quad \forall \ j > 1 \\ where \ \ \sigma_\varepsilon &> \frac{\sigma^2_\varepsilon}{\sqrt{\sigma^2_\varepsilon + \delta^2\sigma^2_{\hat{u}}}} \quad given \ \ \sigma^2_{\hat{u}} > 0 \end{aligned}$$

Again, both information scenarios deliver the same qualitative properties with the response of forward rates to macroeconomic news constant over all horizons.[19] Figure 6 plots the

[19] Introducing serially corellated errors imparts slightly greater volatility to near horizon forward rates, but relatively little as in Gurkaynak et al's empirical findings.

$$var(i^{FI}_{t+j/t} - i^{FI}_{t+j/t-1}) = \sigma^2_\varepsilon + \left[\rho + \frac{\lambda}{\alpha\varphi}(1-\rho)\right]^2 \delta^2\rho^{2j}\sigma^2_{\hat{u}} + \mu^{2j}\frac{\sigma^2_{\hat{g}}}{\varphi^2} \quad \text{for all } j \geqslant 1 \tag{31}$$

coefficients $b^{FI}_{1,j}$ and $b^{LI}_{1,j}$ for the model calibrated according to Table 1, as well as the predicted coefficient when the inflation target is constant. Learning approximately doubles the expected coefficient from such a regression. Qualitatively it is the addition of a moving inflation target, rather than excess sensitivity due to learning, that explains the behaviour of short rates but learning has implications for the magnitude of observed forward rate volatility and heightens the response of forward rates to inflation news

For 1990 to 2002, Gürkaynak et al (2003) report that the volatility of forward rates (standard deviation of quarterly changes) declines from about 1.3 percentage points at the 1 year to horizon to 1 percentage point at 15 years. In the calibrations described above, the moderate innovation variance corresponds to a standard deviation $\sigma_\epsilon = 0.59$ and the low to $\sigma_\epsilon = 0.44$, around half the reported volatility.

2. **The Sensitivity of Bond Yields to monetary policy innovations**

Having derived expressions for long interest rates and the nominal short rate, we can also shed light on a question addressed empirically by previous authors. By how much should we expect an m-period bond respond to a movement in the policy controlled short rate? Previous authors have tried to answer this question by estimating the following regression

$$\Delta i^m_t = a + b_2 \Delta i_t + e_t. \tag{32}$$

Cook and Hahn (1989) performed this regression for data from the 1970s with some success, although the same estimation for later samples yields only a weak relationship perhaps due to a greater anticipated component of target rate movements in recent years (Kuttner 2001). To better assess the impact of monetary policy actions on bonds, Kuttner concentrates on the response of long interest rates to *surprise* monetary policy actions where market expectations are derived using Fed futures contracts. That is, he estimates the following regression

$$\Delta i^m_t = a + b_3 (i_t - i_{t/t-1}) + e_t \tag{33}$$

which yields large and significant estimates of b_3 (reproduced in Table 5). Ellingsen and

Table 5: Estimated and Calibrated response of long interest rates to surprise Fed Funds rate innovations

m	Kuttner	Ellingsen & Söderström	Calibrated Values						
			$\sigma_\epsilon^2 = 0$	$\sigma_\epsilon^2 = 0.35$			$\sigma_\epsilon^2 = 0.2$		
(years)			$\hat{b}_{3,FI} = \hat{b}_{3,LI}$	$\hat{b}_{3,FI}$	$\hat{b}_{3,LI}$	*ratio*	$\hat{b}_{3,FI}$	$\hat{b}_{3,LI}$	*ratio*
1	0.72 (0.08)	0.83 (0.09)	0.47	0.56	0.56	*0.99*	0.52	0.53	*1.03*
2	0.61 (0.06)	0.68 (0.10)	0.25	0.38	0.42	*1.10*	0.32	0.38	*1.19*
5	0.48 (0.04)	0.50 (0.11)	0.10	0.25	0.32	*1.30*	0.19	0.28	*1.50*
10	0.32 (0.03)	0.29 (0.11)	0.05	0.20	0.29	*1.41*	0.14	0.24	*1.72*
20	–	–	0.03	0.18	0.27	*1.50*	0.12	0.23	*1.90*
30	0.19 (0.02)	0.17 (0.09)	0.02	0.17	0.27	*1.52*	0.11	0.22	*1.97*

Notes: Coefficient estimates with standard errors in parentheses reported from Kuttner (2001) and Ellingsen and Soderstrom (2004). Calibrated values generated with monte carlo simulations as in Tables 3 and 4.

Söderström (2004) estimate the same regression, measuring the unanticipated component of monetary policy as the change in the 3 month rate on days when the Fed funds rate was moved. As is evident in the table, both methods yield very similar estimates.

The coefficient b_3 can be given analytical form in terms of the model in Section 3 and leads to the following proposition.

Proposition 4 $b_3^{LI} = \frac{cov(\Delta i_t^m, i_t - i_{t/t-1})^{LI}}{var(i_t - i_{t/t-1})^{LI}} > \frac{cov(\Delta i_t^m, i_t - i_{t/t-1})^{FI}}{var(i_t - i_{t/t-1})^{FI}} = b_3^{FI}$ *for all* $m > 1$ *and* $\frac{\sigma_\varepsilon^2}{\delta^2 \sigma_u^2} > 0$

See Appendix H for proof.

When bond markets employ constant gain learning via forecast errors, the overreaction of the perceived inflation target to transitory shocks that also affect the short rate raises the covariance between in long rates and surprise movements of the short rate.

Table 5 shows calibrated values of the coefficient b_3 for the moderate and low values of target innovation variance (σ_ϵ^2) likely to describe the period after 1989 as well as for a constant inflation target. The underlying parameters of the model are the same as those outlined in Table 1. For longer maturity bonds, the averaging inherent in the expectations hypothesis reduces the covariance between long and short rates. In the case of a constant inflation target, this is powerfully so with long maturity bonds well anchored despite current transitory shocks. Relative to this benchmark, a random walk in the inflation target raises the estimates substantially and learning about an uncommunicated target even more so.

Comparing the columns indicating ratios, we see that constant gain learning significantly raises the coefficient in a regression such as (33), especially for longer rates, with predicted coefficients range between one third to twice as large. Figure 7 plots $\hat{b}_{3,FI}$ and $\hat{b}_{3,LI}$ for the case when $\sigma_{\epsilon}^2 = 0.35$ as well as the predicted coefficients from the constant steady state model. The calibrations shown in the table coincide roughly with those of Kuttner, Ellingsen and Söderström for longer maturities, although at times both $\hat{b}_{3,FI}$ and $\hat{b}_{3,LI}$ fall within one standard deviation of the estimates.

Results change very little when coefficients are calculated using Ellingsen and Söderström's measure of surprise policy innovations (the change in the 3 month rate) or when learning is via the nominal short rate instead of inflation as hypothesised by Romer and Romer. (Present tables.)

4.4 Extensions

1. **Learning with the Wrong Gain**

In the analysis above, the rate at which bond market participants learn about the inflation target is assumed to be calibrated to the true signal-to-noise ratio in the economy. However, agents may be learning at the wrong rate for a number of reasons; insufficient information about the true gain, an attempt to learn more quickly the character of a new policy regime following a change, or a lack of credibility in an announced inflation target. Small deviations from the optimal gain can have sizeable implications for volatility. With faster learning, agents adjust to permanent structural shocks more rapidly but also incorporate more transitory shocks into their estimate of the state variables. For example, building upon the structurally stable model posited by Orphanides and Williams (2003), a small and positive gain imparts sufficient volatility to long interest rates to reject the null in a Shiller style test of excess volatility (Beechey (2004)).

2. **Inflation Targeting**

Proponents of inflation targeting claim greater financial market stability as one of the potential benefits of such a policy (Edey and Stone, 2004). The mechanism described in this paper suggests that this benefit should arise through two channels - by stabilising the nominal target and by communicating its value. One implication is that, conditional on the mean and variance of macroeconomic shocks, long term interest rates should exhibit less volatility under inflation targeting regimes. Bond volatility ought to be able to offer a test of the success of such targets in anchoring long run expectations. This is easier said than done, however, as controlling for differences in the magnitude of macroeconomic shocks and idiosyncratic differences in term premium variation will confound the exercise. In addition, some inflation targets operate with only loosely defined targets that functionally may not differ from the implicit inflation targeting practised in the US (Australia, for example, has a target band with loosely defined assessment horizon) whilst others may not yet have earnt sufficient credibility to affect bond market outcomes.

5 Conclusions

The puzzles of volatility and sensitivity of long interest rates are closely related. As previous authors have noted, non-stationary nominal short rates are more likely to be able to reconcile observed bond volatility with an expectations hypothesis explanation of the term structure than mean-reverting short rates. The model in this paper has built non-stationarity into the economy via the inflation target and introduced an asymmetric information problem which requires key agents to learn about the time-variation of policy preferences.

By adding an asymmetric information dimension to the problem, it has been possible to show that learning contributes not only to the sensitivity of long rates but also imparts volatility to long interest rates of all maturities. The fundamental source of the additional variance is the tendency of the perceived value of the target to overreact to transitory shocks in the economy and confound them for permanent shocks to the inflation target. While this is by no means a claim that asymmetric information about policy goals is the only source of bond volatility - rather, around one fifth of the volatility in a 10 year bond is a more reasonable estimate - it suggests that the revision of long run inflation expectations due to continual learning does contribute to movement in the yield curve. This channel is potent even when there is relatively little time-variation in the inflation target, or to an approximation, the relative preference for output stability.

The addition of a time-varying inflation target also appears important to explain the qualitative reaction of long interest rates to inflation surprises and monetary policy innovation. Learning about the target increases the magnitude of these responses and offers an explanation for the apparent power of monetary policy innovations to affect bonds as long as 30 years.

The framework proposed in this paper is also helpful for addressing such questions as the source of the compression in bond volatility observed in much of the OECD during the mid-1990s and the decline in the amplitude of forecast errors of the nominal short rate.[20] Whilst some have answered that the amplitude of shocks arriving in the economy has compressed, others have pointed to a shift toward explicit or implicit inflation targeting in certain countries and to improvements in central bank transparency and communication. All three channels are at work in the model presented in this paper and the view that reducing the degree of time-variation in policy preferences is a factor finds support here. However, a central bank operating a relatively stable yet uncommunicated target stands to gain the most in reducing financial market volatility by regularly announcing its policy goals.

[20] Swanson (2004) shows using federal funds futures data that the standard deviation of the three month ahead forecast error dropped 15 basis points in the decade after 1994 compared to the 5 years prior, from 39.5 to 24 basis points.

References

Adolfson, M., Laseen, S., Linde, J. and Villani, M.: 2004, Bayesian estimation of an open economy dsge model with incomplete pass-through. mimeo.

Ang, A. and Bekaert, G.: 2004, The term structure of real rates and expected inflation. Unpublished manuscript, Columbia Business School.

Beechey, M.: 2004, Limited information, learning and bond volatility. Unpublished manuscript, U.C. Berkeley.

Calvo, G.: 1983, Staggered prices in a utility-maximizing framework, *Journal of Monetary Economics* .

Clarida, R., Gali, J. and Gertler, P.: 1999, The science of monetary policy: A new keynsian perspective, *Journal of Economic Literature* **37**(4), 1661–1707.

Clarida, R., Gali, J. and Gertler, P.: 2000, Monetary policy rules and macroeconomic stability: Evidence and some theory, *Quarterly Journal of Economics* (1), 147–180.

Cook, T. and Hahn, T.: 1989, The effect of changes in the federal funds rate target on market rates in the 1970s, *Journal of Monetary Economics* **24**(3), 331–351.

Ellingsen, T. and Söderström, U.: 2001, Monetary policy and market interest rates, *American Economic Review* **91**(5), 1594–1607.

Ellingsen, T. and Söderström, U.: 2004, Why are long rates sensitive to monetary policy? Sveriges Riksbank Working Paper Series No. 160.

Evans, G. W. and Honkapohja, S.: 2001, *Learning and Expectations in Macroeconomics*, Princeton University Press.

Fuhrer, J. C.: 1996, Monetary policy shifts and long-term interest rates, *Quarterly Journal of Economics* **111**(4), 1183–1209.

Fuhrer, J. C. and Moore, G.: 1995, Inflation persistence, *Quarterly Journal of Economics* **110**(1), 127–159.

Gali, J.: 2003, *Advances in Economic Theory*, Cambridge University Press, chapter New perspectives on monetary policy, inflation and the business cycle, pp. 151–197.

Gürkaynak, R., Sack, B. and Swanson, E.: 2003, The excess sensitivity of long term interest rates: Evidence and implications for macroeconomic models. Mimeo, Federal Reserve Board of Governors.

Kozicki, S. and Tinsley, P.: 2003, Permanent and transitory policy shocks in an empirical macro model with asymmetric information. CFS Working Paper No. 2003/41.

Kuttner, K.: 2001, Monetary policy surprises and interest rates: Evidence from the fed funds futures market, *Journal of Monetary Economics* **47**(3), 523–544.

Mishkin, F.: 1992, Is the fisher effect for real?: A reexamination of the relationship between inflation and interest rates, *Journal of Monetary Econonmics* **30**(2), 195–215.

Orphanides, A. and Williams, J. C.: 2003, Imperfect knowledge, inflation expectations and monetary policy. NBER Working Paper 9884.

Preston, B.: 2002, Learning about monetary policy rules when long-horizon expectations matter. Unpublished, Princeton University, http//www.princeton.edu/bpreston/preston.pdf.

Romer, C. and Romer, D.: 2000, Federal reserve information and the behavior of interest rates, *American Economic Review* **90**(3), 429–457.

Rudebusch, G.: 1995, Federal reserve interest rate targeting, rational expectations and the term structure, *Journal of Monetary Economics* **35**, 245–274.

Rudebusch, G.: 2002, Assessing nominal income rules for monetary policy with model and data uncertainty, *Economic Journal* **112**, 1–31.

Söderlind, P.: 1999, Solution and estimation of re macromodels with optimal policy, *European Economic Review* **43**(4).

Shiller, R.: 1979, The volatility of long-term interest rates and expectations models of the term structure, *Journal of Political Economy* **87**(6), 1190–1219.

Smets, F. and Wouters, R.: 2003, Shocks and frictions in us and euro area business cycles: A bayesian dsge approach. mimeo.

Swanson, E. T.: 2004, Federal reserve transparency and financial market forecasts of short-term interest rates. Board of Governors, Federal Reserve.

Wallace, M. and Warner, J.: 1993, The fisher effect and the term structure of interest rates: Tests of cointegration, *Review of Economics and Statistics* **75**(2), 320–324.

Woodford, M.: 2003, *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton Univeristy Press.

6 Appendix A - Phillips's curve with π_t^*

Aggregate Demand A infinitely-lived, representative household chooses C_t (consumption), N_t (household size) and B_t (assets) to maximise lifetime utility

$$\max_{C_t, N_t, B_t} \mathsf{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} \right]$$

s.t. $P_t C_t + (1+i_t)^{-1} B_t = B_{t-1} + W_t N_t$ for all t

where C_t and P_t are constant elasticity of substitution combinations of goods over measure 1. The first order conditions of the standard Lagrangian for this problem yield the consumption Euler equation

$$1 = \mathsf{E}_t \left\{ \beta(1+i_t) \left[\frac{C_{t+1}}{C_t} \right]^{-\sigma} \frac{P_t}{P_{t+1}} \right\}.$$

Log-linearising the Euler equation around the steady state yields

$$c_t = \mathsf{E}_t c_{t+1} - \frac{1}{\sigma} \left[i_t - \mathsf{E}_t \pi_{t+1} \right]$$

where lower case letters denote log deviations. Paired with the market clearing condition $c_t = y_t$ and rewritten in terms of the output gap, $x_t = y_t - y_t^P$ (the deviation of output from the flexible price "potential" level y_t^P) yields the aggregate demand equation in the text. (This can also encompass exogenously evolving government spending). The steady state implies a constant real interest rate around which the Euler equation is linearised.

Aggregate Supply The market for final goods is perfectly competitive and the production function of the final good firm transforms intermediate inputs into final output according to

$$Y_t = \left[\int_0^1 Y_{i.t}^{\frac{1}{\lambda}} \right]^{\lambda_t}, 1 \leq \lambda < \infty.$$

Profit maximisation leads to the following relationship between the average price level of final goods, P_t, and the prices of intermediate goods, $P_{i,t}$,

$$P_t = \left[\int_0^1 P_{i.t}^{\frac{1}{1-\lambda_t}} \right]^{1-\lambda_t}. \tag{34}$$

A continuum of intermediate firms, each producing a differentiated good, faces monopolistic competition. As in Calvo (1983), the probability that a firm can re-optimise its price in a given period is constant and equal to $(1-\xi_p)$. For a given firm i, its re-optimised price is $P_{i,t}^{new}$. If a firm does not re-optimise then its price at $t+1$ is indexed according to $P_{i,t+1} = \pi_t^{\gamma_p} (\pi_t^*)^{1-\gamma_p} P_{i,t}^{new}$ where $\gamma_p \in [0,1]$. Thus the price the firm can charge if it has not re-optimised in j periods is $(\pi_t \pi_{t+1} \ldots \pi_{t+j-1})^{\gamma_p} \left(\pi_{t+1}^* \pi_{t+2}^* \ldots \pi_{t+j}^* \right)^{1-\gamma_p} P_{i,t}^{new}$.

The representative firm faces the following optimisation problem when setting its price,

$$\max_{P_{i,t}^{new}} \mathsf{E}_t \sum_{j=0}^{\infty} (\beta\xi)^j \, \upsilon_{t+j} \left[\begin{array}{c} \left((\pi_t \pi_{t+1} \dots \pi_{t+j-1})^{\gamma_p} \left(\pi_{t+1}^* \pi_{t+2}^* \dots \pi_{t+j}^* \right)^{1-\gamma_p} P_{i,t}^{new} \right) Y_{i,t+j} \\ -MC_{i,t+j} Y_{i,t+j} - MC_{i,t+j} z_{t+j} \phi \end{array} \right]$$

where $\beta\upsilon_{t+j}$ is the stochastic discount factor between periods t and $t+j$ used to discount profits, $Y_{i,t+j}$ is the output of the i^{th} intermediate firm, $MC_{i,t+j}$ its real marginal cost and z_{t+j} a permanent technology shock in the intermediate goods production function.

The optimisation problem leads to a first order condition that when combined with the average price level in (34) yields two equations that can be log-linearised. Together these yield an aggregate Phillips curve showing the relationship between inflation, real marginal cost and the inflation target when the latter follows a random walk;

$$\pi_t - \pi_t^* = \frac{\beta}{1+\gamma_p \beta}(\mathsf{E}_t \pi_{t+1} - \pi_t^*) + \frac{\gamma_p}{1+\gamma_p \beta}(\pi_{t-1} - \pi_t^*) + \frac{(1-\xi_p)}{\xi_p} \frac{(1-\beta\xi_p)}{1+\gamma_p \beta} (mc_t + \lambda_t)$$

where π_t and mc_t are log deviations from their respective steady state values. In this form, the deviation of inflation from the target depends on past and expected future inflation as well as marginal cost. When $\gamma_p = 0$ (ie: non-optimised prices are fully indexed to the current inflation target), the equation simplifies to the forward-looking aggregate supply curve shown in the text with the additional simplification that $mc_t = \beta_1 x_t + u_t$ where u_t can be interpreted as a cost-push shock or as a markup shock as in Gali (2003) and Woodford (2003).

7 Appendix B - Solving for Optimal Policy

Given the model shown in equations (1a) to (**??**) and (4), the central bank minimises its loss function each period. That is, they consider max $\left\{ \frac{1}{2}\alpha x_t^2 + (\pi_t - \pi_t^*)^2 \right\}$ with respect to x_t which yields first order condition

$$\begin{aligned} \alpha x_t + \lambda(\pi_t - \pi_t^*) &= 0 \\ x_t &= -\frac{\lambda}{\alpha}(\pi_t - \pi_t^*). \end{aligned}$$

To solve for x_t and π_t in terms of shocks arriving in the model, substitute this first order condition into the Phillips curve and solve for the inflation deviation at time t

$$\begin{aligned} \pi_t - \pi_t^* &= \beta E_t[\pi_{t+1} - \pi_{t+1}^*] - \frac{\lambda^2}{\alpha}[\pi_t - \pi_t^*] + +u_t \\ \pi_t - \pi_t^* &= \frac{\alpha\beta}{\alpha+\lambda^2} E_t[\pi_{t+1} - \pi_{t+1}^*] + \frac{\alpha}{\alpha+\lambda^2} u_t \end{aligned}$$

Assuming that private sector (price setters') expectations are rationally forward looking, that is $E_t[\pi_{t+1} - \pi_{t+1}^*] = E_t \left(\frac{\alpha\beta}{\alpha+\lambda^2} E_{t+1}[\pi_{t+2} - \pi_{t+2}^*] + \frac{\alpha}{\alpha+\lambda^2} u_{t+1} \right)$, repeatedly substitute and take expectations at t

$$\begin{aligned}
\left(\alpha+\lambda^2\right)[\pi_t-\pi_t^*] &= \alpha\beta E_t\left[\frac{\alpha\beta}{\alpha+\lambda^2}E_{t+1}[\pi_{t+2}-\pi_{t+2}^*]+\frac{\alpha}{\alpha+\lambda^2}u_{t+1}\right]+\alpha u_t \\
&\vdots \\
&= \alpha E_t\left[\sum_{j=0}^{\infty}\left(\frac{\alpha\beta}{\alpha+\lambda^2}\right)^j u_{t+j}\right] \\
&= \alpha\sum_{j=0}^{\infty}\left(\frac{\alpha\beta}{\alpha+\lambda^2}\right)^j \rho^j u_t
\end{aligned}$$

Taking an infinite geometric sum, we can express the deviation of inflation from the target at time t as a function of the current inflation shock,

$$\pi_t-\pi_t^*=\frac{\alpha}{\lambda^2+\alpha(1-\beta\rho)}u_t.$$

Substituting into the policy trade-off likewise yields the expression for the output gap,

$$x_t=-\frac{\lambda}{\alpha}[\pi_t-\pi_t^*]=\left(\frac{-\lambda}{\lambda^2+\alpha(1-\beta\rho)}\right)u_t.$$

To find the optimal interest rate rule, employ the serial correlation in u_t to write $E_t[\pi_{t+1}-\pi_{t+1}^*]=\rho[\pi_t-\pi_t^*]$ and substitute into the policy trade-off,

$$x_t=-\frac{\lambda}{\alpha}[\pi_t-\pi_t^*]=-\frac{\lambda}{\alpha\rho}E_t[\pi_{t+1}-\pi_{t+1}^*].$$

Replacing x_t in the LHS of the aggregate demand equation and substituting for $E_t\left(x_{t+1}\right)$ in the RHS, we can solve for the optimal setting of the policy controlled nominal short interest rate,

$$\begin{aligned}
-\frac{\lambda}{\alpha\rho}E_t[\pi_{t+1}-\pi_{t+1}^*] &= -\gamma\left(i_t-E_t\left(\pi_{t+1}\right)\right)+E_t\left(y_{t+1}\right)+g_t \\
i_t &= \left[1+\frac{\lambda(1-\rho)}{\alpha\gamma\rho}\right]E_t\left[\pi_{t+1}-\pi_{t+1}^*\right]+E\left(\pi_{t+1}^*\right)+\frac{g_t}{\gamma}
\end{aligned}$$

8 Appendix C - Univariate Steady State Kalman Filter

The following derivation of the univariate steady state Kalman gain owes heavily to Sargent's (1983) treatment of Muth's solution to the permanent income inference problem, which conveniently takes a similar form to the problem here. From Section 3.2 we have

$$\begin{array}{lrcl} \text{Observation equation} & \pi_t & = & \pi_t^* + \delta\hat{u}_t \\ \text{State equation} & \pi_t^* & = & \pi_{t-1}^* + \varepsilon_t \end{array}$$

Add $\delta\hat{u}_t$ to both sides of the state equation (21) and rewrite the change in inflation as as an AR(1) process with a MA error term,

$$\pi_t - \pi_{t-1} = \delta(u_t - u_{t-1}) + \varepsilon_t.$$

The auto-covariance structure of this error term is

$$cov\left[\delta(u_t - u_{t-1}) + \varepsilon_t, \delta(u_{t-j} - u_{t-j-1}) + \varepsilon_{t-j}\right] = \left\{ \begin{array}{c} 2\delta^2\sigma_u^2 + \sigma_\varepsilon^2 \text{ for } j = 0 \\ -\delta^2\sigma_u^2 \text{ for } j = 1 \\ 0 \text{ for } j \geqslant 2 \end{array} \right\}.$$

We can replicate the covariance properties of this process with a MA process (Wold's theorem) as follows

$$\delta(u_t - u_{t-1}) + \varepsilon_t = k_t - \phi k_{t-1} \tag{35}$$

where k_t is a stationary, serially uncorrelated random process with mean zero and variance σ_k^2. The auto-covariance structure of k_t is

$$cov\left[k_t - \phi k_{t-1}, k_t - j - \phi k_{t-j-1}\right] = \left\{ \begin{array}{c} (1+\phi^2)\sigma_k^2 \text{ for } j = 0 \\ -\phi^2\sigma_k^2 \text{ for } j = 1 \\ 0 \text{ for } j \geqslant 2 \end{array} \right\}.$$

Matching coefficients we have two relationships

$$\begin{array}{rcl} 2\delta^2\sigma_u^2 + \sigma_\varepsilon^2 & = & (1+\phi^2)\sigma_k^2 \text{ and} \\ -\delta^2\sigma_u^2 & = & -\phi^2\sigma_k^2 \end{array} \tag{36}$$

which can be solved for $\sigma_k^2\left(\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}\right)$ and

$$\phi = 1 + \frac{1}{2}\left(\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}\right) - \sqrt{\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}\left(1 + \frac{1}{4}\left(\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2}\right)\right)}$$

As $\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2} \to \infty$, $\phi \to 0$ and $\frac{\sigma_\varepsilon^2}{\delta^2\sigma_u^2} \to 0$, $\phi \to 1$.

It will be useful to write the optimal projection as a geometrically declining lagged polynomial of previously observed inflation outcomes. From above we have

$$\begin{aligned}(1-L)\pi_t &= (1-\phi L)k_t \\ &\Rightarrow \frac{(1-L)}{(1-\phi L)}\pi_t = \pi_t - \frac{(1-\phi)}{(1-\phi L)}\pi_{t-1} = k_t\end{aligned}$$

Because k_t is orthogonal to the information set Ω_{t-1} and using the optimal projections we have

$$\pi_{t/t-1} = \pi_{t-1/t-1} = \pi^*_{t-1/t-1} = \frac{(1-\phi)}{1-\phi L}\pi_{t-1}$$

Manipulation of the above results yields

$$\pi_t - \pi^*_{t/t} = \phi k_t \qquad \text{and}$$

$$\pi^*_{t/t} - \pi^*_{t-1/t-1} = (1-\phi)\left[\pi_t - \frac{(1-\phi)}{1-\phi L}\pi_{t-1}\right] = (1-\phi)k_t$$

where $k_t = (\pi_t - \pi_{t/t-1})$ is the one-period ahead inflation forecast error.

9 Appendix D - Forecast Errors

Full information: With ρ and μ reintroduced, the expression for the one-period ahead forecast error of the nominal short rate becomes

$$i_{t+1} - i^{FI}_{t+1/t} = \varepsilon_{t+1} + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]\delta\hat{u}_{t+1} + \frac{1}{\varphi}\hat{g}_{t+1} \tag{37}$$

with variance

$$var(i_{t+1} - i^{FI}_{t+1/t}) = \sigma^2_\varepsilon + \left[\rho + \frac{\lambda}{\alpha\gamma}(1-\rho)\right]^2 \delta^2\sigma^2_{\hat{u}} + \frac{\sigma^2_{\hat{g}}}{\gamma^2}$$

This variance is rising in ρ.

Limited Information ($\rho = \mu = 0$): First note that we can rewrite the expression for the forecast error as

$$\begin{aligned} i_{t+1} - i^{LI}_{t+1/t} &= \varepsilon_{t+1} + \left(\pi^*_t - \hat{\pi}^*_{t/t}\right) + \frac{\lambda}{\alpha\gamma}\delta\hat{u}_{t+1} + \frac{1}{\gamma}\hat{g}_{t+1} && (38) \\ &= \varepsilon_{t+1} - \delta u_t + \left(\pi_t - \hat{\pi}^*_{t/t}\right) + \frac{\lambda}{\alpha\gamma}\delta\hat{u}_{t+1} + \frac{1}{\gamma}\hat{g}_{t+1} && (39)\end{aligned}$$

From Appendix C that have $\pi_t - \pi^*_{t/t} = \phi k_t$ so replacing this and recalling $k_t - \phi k_{t-1} = \varepsilon_t + \delta\left(u_t - u_{t-1}\right)$ we can recursively substitute for lags of k_{t-j} and group terms until the forecast error is expressed in terms of historical errors

$$
\begin{aligned}
i_{t+1} - i^{LI}_{t+1/t} &= \varepsilon_{t+1} - \delta \hat{u}_t + \phi k_t + \frac{\lambda}{\alpha\gamma}\delta\hat{u}_{t+1} + \frac{1}{\gamma}\hat{g}_{t+1} \\
&= \varepsilon_{t+1} - \delta \hat{u}_t + \phi(\varepsilon_t + \delta\left(\hat{u}_t - \hat{u}_{t-1}\right) + \phi k_{t-1}) + \frac{\lambda}{\alpha\gamma}\delta\hat{u}_{t+1} + \frac{1}{\gamma}\hat{g}_{t+1} \\
&\vdots \\
&= \sum_{m=0}^{\infty} \phi^m \varepsilon_{t+1-m} + \frac{\lambda}{\alpha\gamma}\delta\hat{u}_{t+1} - (1-\phi)\sum_{m=0}^{\infty}\phi^m \hat{u}_{t-m} + \frac{1}{\gamma}\hat{g}_{t+1}
\end{aligned}
$$

For $\phi < 1$ these summations are finite allowing us to compute the variance (recalling the cross-independence of the errors)

$$
\begin{aligned}
var(i_{t+1} - i^{LI}_{t+1/t}) &= \sigma^2_\varepsilon(1 + \phi^2 + \phi^4 + \ldots) + \delta^2\sigma^2_{\hat{u}}\frac{\lambda}{\alpha\varphi}^2 + \delta^2\sigma^2_{\hat{u}}(1-\phi)^2(1 + \phi^2 + \phi^4 + \ldots) + \frac{\sigma^2_{\hat{g}}}{\varphi} \\
&= \sigma^2_\varepsilon(\frac{1}{1-\phi^2}) + \delta^2\sigma^2_{\hat{u}}\left[\frac{\lambda}{\alpha\gamma}^2 + \frac{(1-\phi)^2}{1-\phi^2}\right] + \frac{\sigma^2_{\hat{g}}}{\gamma^2}
\end{aligned}
$$

Proof of Proposition 1

For any $\phi \in (0,1)$

$$
\begin{aligned}
\frac{1}{1-\phi^2} &> 1 \\
\frac{(1-\phi)^2}{1-\phi^2} &> 0
\end{aligned}
$$

Thus for any given σ^2_ε, $\sigma^2_{\hat{u}}$ and $\sigma^2_{\hat{g}}$

$$
\sigma^2_\varepsilon(\frac{1}{1-\phi^2}) + \delta^2\sigma^2_{\hat{u}}\left[\frac{\lambda}{\alpha\varphi}^2 + \frac{(1-\phi)^2}{1-\phi^2}\right] + \frac{\sigma^2_{\hat{g}}}{\varphi^2} > \sigma^2_\varepsilon + \frac{\lambda}{\alpha\gamma}^2\delta^2\sigma^2_{\hat{u}} + \frac{\sigma^2_{\hat{g}}}{\gamma^2} \tag{40}
$$

□

10 Appendix E - m-period bond with full information and serially correlated errors

With serially correlated errors, the sum of the current and forecasted short rates is

$$
\begin{aligned}
i^{m^{FI}}_t &= \frac{1}{m}\left(\pi^*_t + \sum\nolimits_{j=1}^{m-1}\pi^*_{t+j/t} + \left[\rho + \frac{\lambda}{\alpha\varphi}(1-\rho)\right]\delta\left(u_t + \sum\nolimits_{j=1}^{m-1} u_{t+j/t}\right) + \frac{1}{\varphi}\left(g_t + \sum\nolimits_{j=1}^{m-1} g_{t+j/t}\right)\right) + \zeta^m_t \\
&= \pi^*_t + \frac{1}{m}\left(\left[\rho + \frac{\lambda}{\alpha\varphi}(1-\rho)\right]\delta u_t \sum_{j=0}^{m-1}\rho^j + \frac{g_t}{\varphi}\sum_{j=0}^{m-1}\mu^j\right) + \zeta^m \qquad (41)
\end{aligned}
$$

so that the one period change becomes

$$i_t^{m^{FI}} - i_{t-1}^{m^{FI}} = \varepsilon_t + \frac{1}{m}\left(\left[\rho + \frac{\lambda}{\alpha\varphi}(1-\rho)\right]\delta\sum_{j=0}^{m-1}\rho^j(u_t - u_{t-1}) + \frac{1}{\varphi}\sum_{j=0}^{m-1}\mu^j\left(g_t - g_{t-1}\right)\right) + \zeta_t^m - \zeta_{t-1}^m. \tag{42}$$

The variance of the change in this m-period bond takes account of the fact that $var(u_t - u_{t-1}) = 2(1-\rho)\sigma_u^2$

$$var(i_t^{m^{FI}} - i_{t-1}^{m^{FI}}) = \sigma_\varepsilon^2 + \frac{1}{m^2}\left(\left(\left[\rho + \frac{\lambda}{\alpha\varphi}(1-\rho)\right]\delta\sum_{j=0}^{m-1}\rho^j\right)^2 2(1-\rho)\sigma_u^2 + \left(\frac{1}{\varphi}\sum_{j=0}^{m-1}\mu^j\right)^2 2\sigma_g^2(1-\mu)\right) + \sigma_\varsigma^2$$

where $\sigma_u^2 = \frac{\sigma_{\hat{u}}^2}{1-\rho^2}$ and $\sigma_g^2 = \frac{\sigma_{\hat{g}}^2}{1-\mu^2}$ are the unconditional variances of the serially correlated disturbances, u_t and g_t. This variance is increasing in p, both through the first and second terms in parentheses, as well as due to the fact that $\frac{\partial\delta}{\partial\rho} > 0$.

11 Appendix F - m-period bond with learning when $\rho = \mu = 0$

From the text we have

$$i_t^{m^{LI}} - i_{t-1}^{m^{LI}} = \frac{1}{m}\left(\varepsilon_t + (m-1)(\pi_{t/t}^* - \pi_{t-1/t-1}^*) + \frac{\lambda}{\alpha\varphi}\delta(\hat{u}_t - \hat{u}_{t-1}) + \frac{1}{\varphi}(\hat{g}_t - \hat{g}_{t-1})\right)$$

Borrowing the result from Appendix B that in the simplest case $\pi_{t/t}^* - \pi_{t-1/t-1}^* = (1-\phi)k_t$ then substituting recursively for k_{t-j}, $j = 0, ..\infty$ using $k_t = \varepsilon_t + \delta(u_t - u_{t-1}) + \phi k_{t-1}$

$$\begin{aligned} i_t^{m^{LI}} - i_{t-1}^{m^{LI}} &= \frac{1}{m}\left(\varepsilon_t + (m-1)(1-\phi)k_t + \frac{\lambda}{\alpha\varphi}\delta(\hat{u}_t - \hat{u}_{t-1}) + \frac{1}{\varphi}(\hat{g}_t - \hat{g}_{t-1})\right) \\ &\vdots \end{aligned} \tag{43}$$

$$= \frac{1}{m}\left(\begin{array}{c} \varepsilon_t\left[1 + (m-1)(1-\phi)\right] + \frac{(m-1)(1-\phi)}{1-\phi L}\phi\varepsilon_{t-1} + \frac{1}{\varphi}(g_t - g_{t-1}) \\ + \left[\frac{\lambda}{\alpha\varphi} + (m-1)(1-\phi)\right]\delta u_t - \left[\frac{\lambda}{\alpha\varphi} + (m-1)(1-\phi)^2\right]\delta u_{t-1} \\ - \left[\frac{(m-1)(1-\phi)^2\phi}{1-\phi L}\right]\delta u_{t-2} \end{array}\right) \tag{44}$$

$$var(i_t^{m^{LI}} - i_{t-1}^{m^{LI}}) = \frac{1}{m^2}\left(\begin{array}{c} \left[1 + (m-1)(1-\phi)\right]^2\sigma_\varepsilon^2 + \frac{((m-1)(1-\phi)\phi)^2}{1-\phi^2}\sigma_\varepsilon^2 + \frac{1}{\varphi^2}2\sigma_g^2 \\ + \left[\frac{\lambda}{\alpha\varphi} + (m-1)(1-\phi)\right]^2\delta^2\sigma_u^2 + \left[\frac{\lambda}{\alpha\varphi} + (m-1)(1-\phi)^2\right]^2\delta^2\sigma_u^2 \\ + \left[\frac{\left((m-1)(1-\phi)^2\phi\right)^2}{1-\phi^2}\right]\delta^2\sigma_u^2 \end{array}\right) + \sigma_\varsigma^2 \tag{45}$$

12 Appendix G - Forward rate regression coefficients

The regression we are concerned with is

$$i_{t+j/t} - i_{t+j/t-1} = \alpha + \beta \frac{\pi_t - \pi_{t/t-1}}{stdev(\pi_t - \pi_{t/t-1})} + \epsilon$$

The coefficients can be calculated as follows:

$$\begin{aligned}
\frac{cov(i^{FI}_{t+j/t} - i^{FI}_{t+j/t-1}, \pi_t - \pi^{FI}_{t/t-1})}{var(\pi_t - \pi^{FI}_{t/t-1})} \sqrt{var(\pi_t - \pi^{FI}_{t/t-1})} &= \frac{\sigma^2_\varepsilon}{\sqrt{\sigma^2_\varepsilon + \delta^2 \sigma^2_{\hat{u}}}} \quad \forall \ \ j > 1 \\
\frac{cov(i^{FI}_{t+j/t} - i^{FI}_{t+j/t-1}, \pi_t - \pi^{FI}_{t/t-1})}{var(\pi_t - \pi^{FI}_{t/t-1})} \sqrt{var(\pi_t - \pi^{LI}_{t/t-1})} &= (1-\phi)\sqrt{\frac{\sigma^2_\varepsilon}{(1-\phi)^2}} = \sigma_\varepsilon \qquad \forall \ \ j > 1
\end{aligned}$$

Because $\sigma_\varepsilon > \frac{\sigma^2_\varepsilon}{\sqrt{\sigma^2_\varepsilon + \delta^2 \sigma^2_{\hat{u}}}}$ for any $\sigma^2_{\hat{u}} > 0$, $\beta^{LI} > \beta^{FI}$.

13 Appendix H - Proof of Proposition 4

For Proposition 5 to be true requires that

$$\frac{cov(\Delta i^m_t, i_t - i_{t/t-1})^{LI}}{cov(\Delta i^m_t, i_t - i_{t/t-1})^{FI}} > \frac{var(i_t - i_{t/t-1})^{LI}}{var(i_t - i_{t/t-1})^{FI}}$$

From Appendix D, we have the variances of the forecast error under both scenarios when $\rho = \mu = 0$

$$\begin{aligned}
var(i_{t+1} - i^{FI}_{t+1/t}) &= \sigma^2_\varepsilon [1] + \delta^2 \sigma^2_{\hat{u}} \left[\frac{\lambda}{\alpha\gamma}^2 \right] + \frac{\sigma^2_{\hat{g}}}{\gamma^2} \\
var(i_{t+1} - i^{LI}_{t+1/t}) &= \sigma^2_\varepsilon \left[1 + \frac{\phi^2}{1-\phi^2} \right] + \delta^2 \sigma^2_{\hat{u}} \left[\frac{\lambda}{\alpha\gamma}^2 + \frac{(1-\phi)^2}{1-\phi^2} \right] + \frac{\sigma^2_{\hat{g}}}{\gamma^2}.
\end{aligned}$$

The covariances are as follows

$$\begin{aligned}
cov(\Delta i^m_t, i_t - i_{t/t-1})^{FI} &= \sigma^2_\varepsilon [1] + \delta^2 \sigma^2_{\hat{u}} \left[\frac{1}{m} \frac{\lambda}{\alpha\gamma}^2 \right] + \frac{1}{m} \frac{\sigma^2_{\hat{g}}}{\gamma^2} \\
cov(\Delta i^m_t, i_t - i_{t/t-1})^{LI} &= \sigma^2_\varepsilon \left[1 + \frac{m-1}{m} \frac{\phi(1-\phi)}{1-\phi^2} \right] + \delta^2 \sigma^2_{\hat{u}} \left[\frac{1}{m} \frac{\lambda}{\alpha\gamma}^2 + \frac{\lambda}{\alpha\gamma}(1-\phi) + \frac{m-1}{m} \frac{(1-\phi)^3}{1-\phi^2} \right] + \frac{\sigma^2_{\hat{g}}}{\gamma^2}
\end{aligned}$$

The ratio of covariances must exceed the ratio of variances. When $m = 1$ the covariances are still not the same because of the presence of $\frac{\lambda}{\alpha\gamma}(1-\phi)$. That is, the change in the short rate (Δi^m_t) has a larger covariance with the forecast error in that period under LI. When $\sigma^2_\varepsilon = 0$, however, and $\phi = 1$ the two coincide because there is no longer an information problem.

$$\frac{cov(\Delta i_t^m, i_t - i_{t/t-1})^{LI}}{var(i_t - i_{t/t-1})^{LI}} > \frac{cov(\Delta i_t^m, i_t - i_{t/t-1})^{FI}}{var(i_t - i_{t/t-1})^{FI}} \quad for \quad all \quad m \geqslant 1 \quad and \quad \frac{\sigma_\varepsilon^2}{\delta^2 \sigma_u^2} > 0$$

Figure 1: Inflation Target - Actual and Perceived

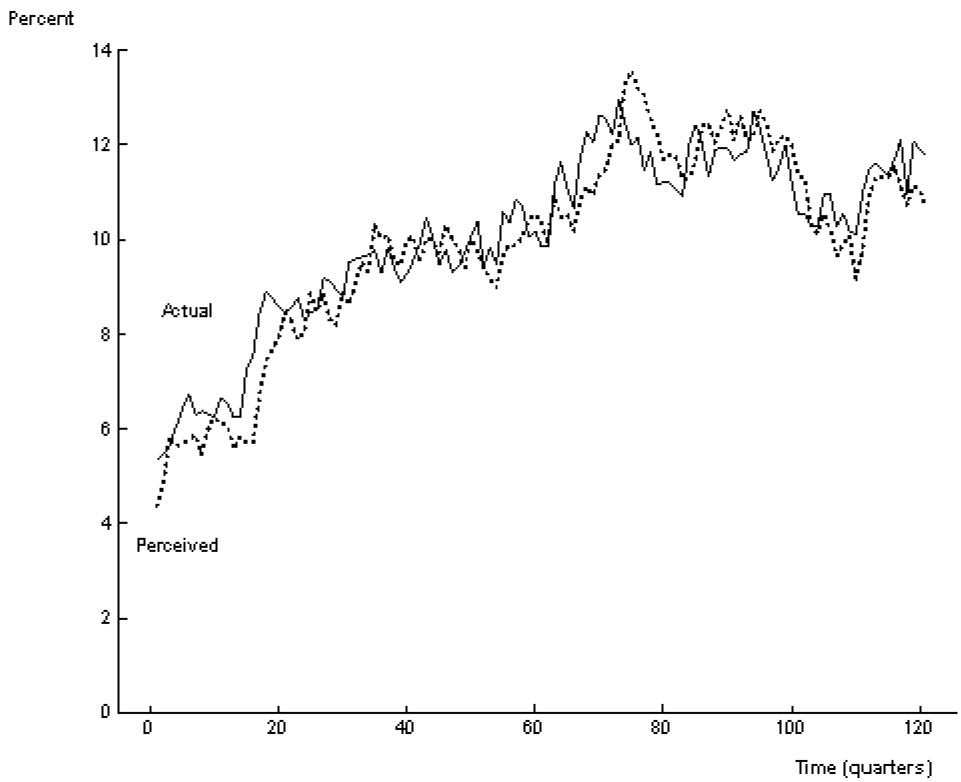

Figure 2: Impulse Response Functions of Forward Rates

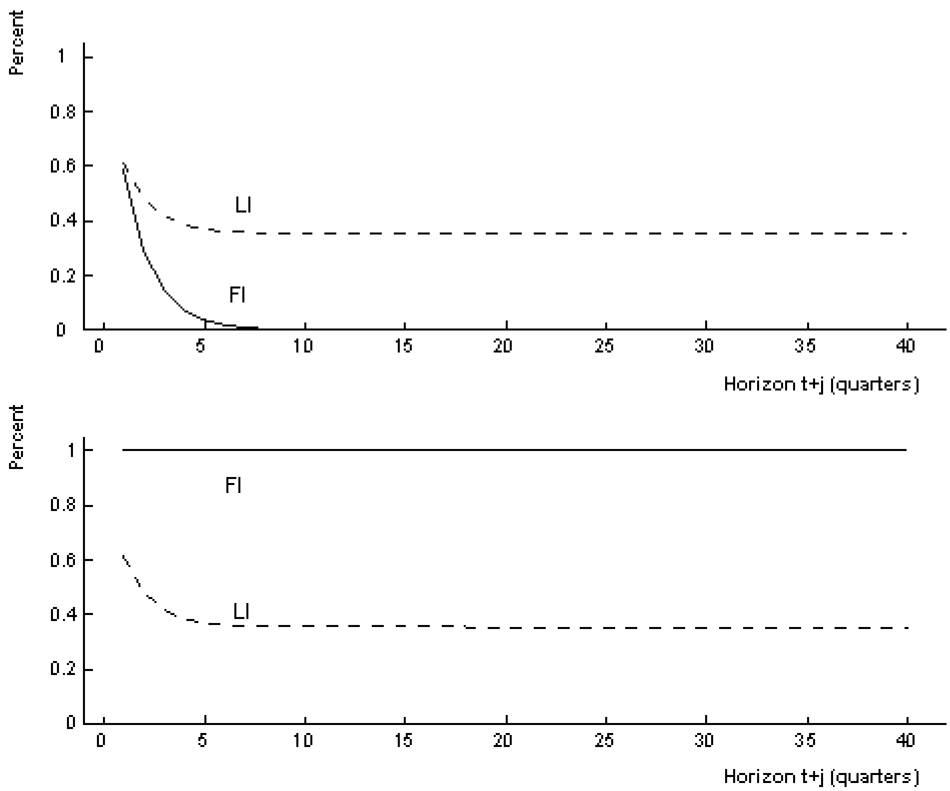

Figure 3: Volatility of the Term Structure - Actual and Calibrated

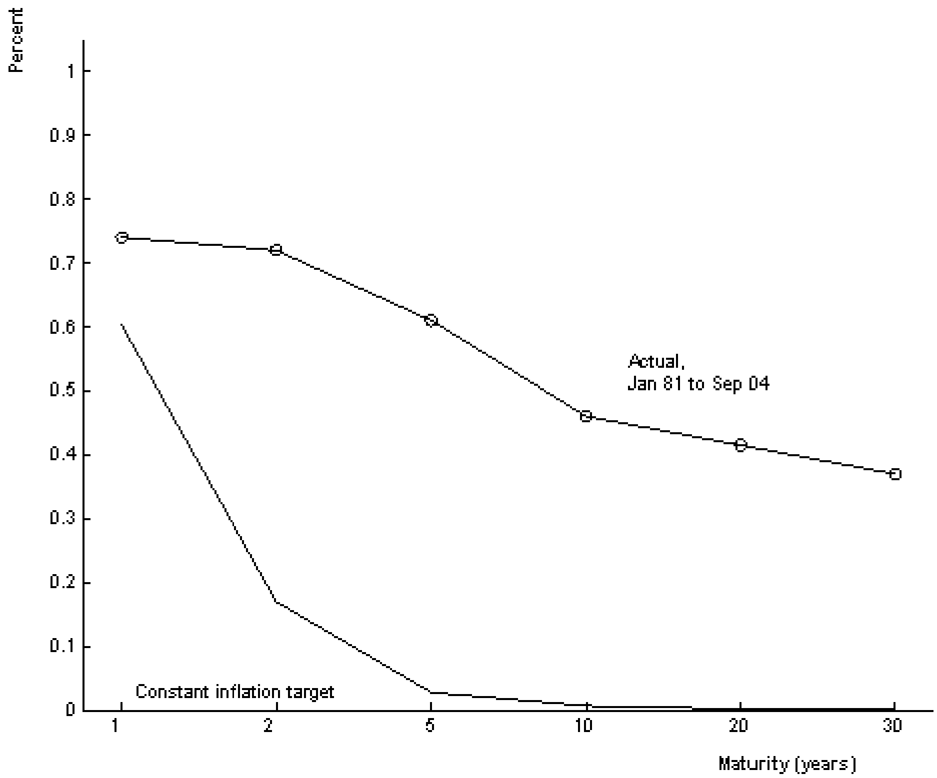

Figure 4: Volatility of the Term Structure - Actual and Calibrated

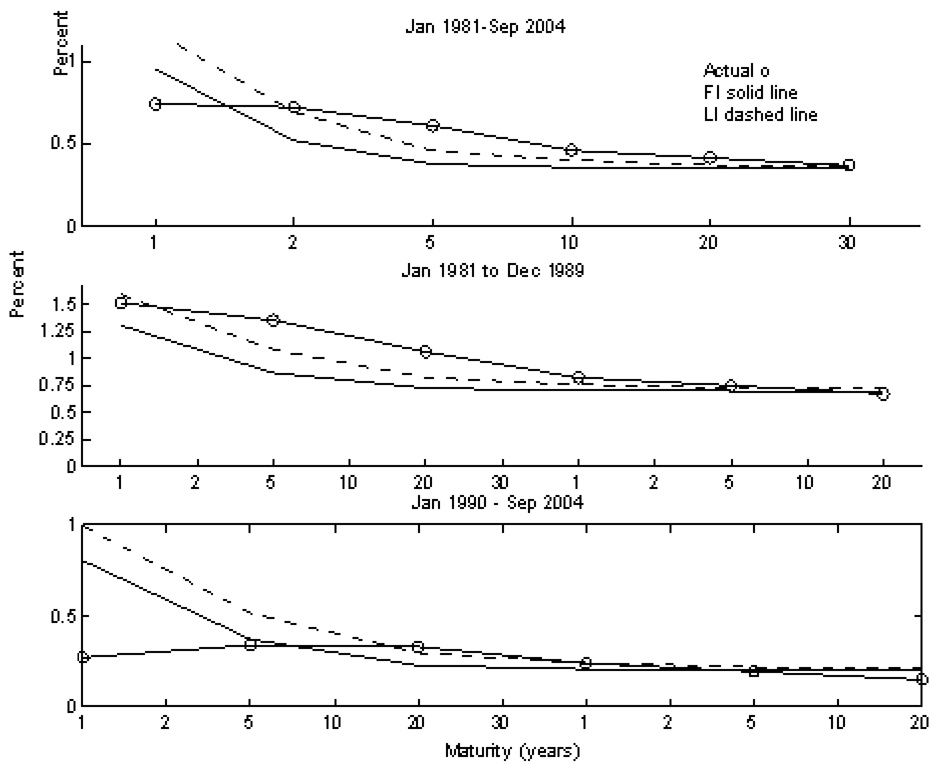

Figure 5: Simulated variance ratio of 10 year bond: LI/FI

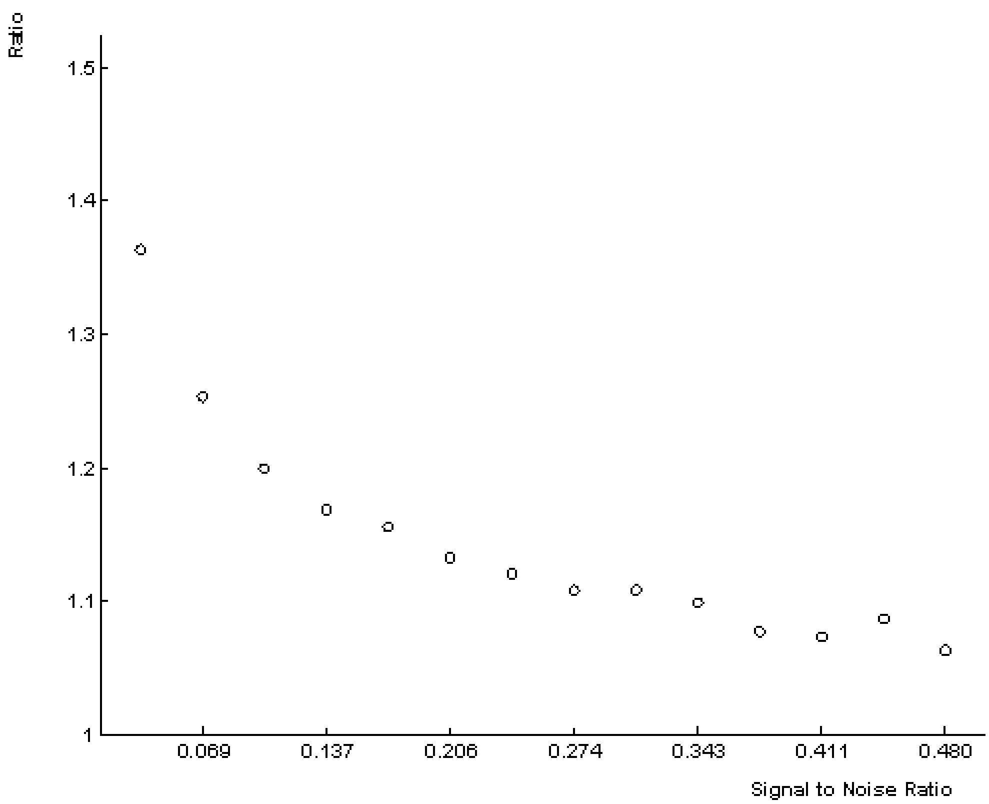

Figure 6: Regression coefficient of the response of forward rates to an inflation surprise

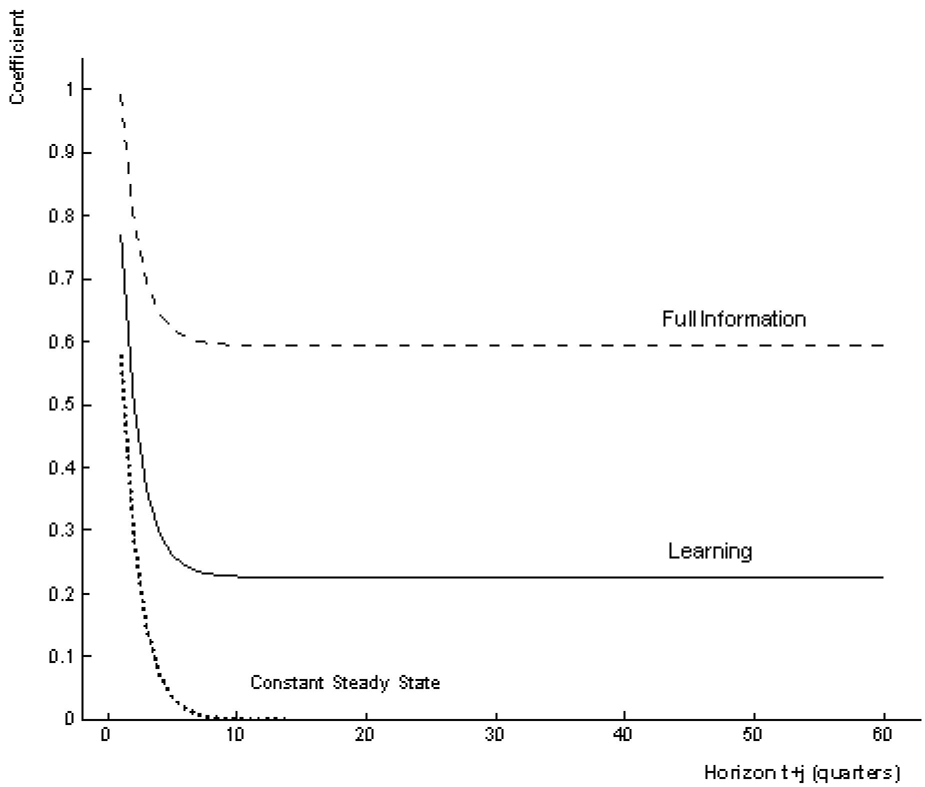

Figure 7: Regression coefficient of m-period bond response to a short rate surprise