

Optimal Monetary Policy with Inflation-Conditional Macroeconomic Volatility

Peter Spencer*and Vito Polito‡

Abstract

This paper develops an econometric model of the US macroeconomy that is potentially heteroskedastic and uses it to analyze optimal monetary policy. This model allows the variance structure to depend in a linear or quadratic way on a lagged endogenous variable such as the inflation or interest rate, consistent with the Okun (1971) and Friedman (1977) conjecture that macroeconomic volatility depends upon the underlying rate of inflation. The linear specification is the analogue of the Cox, Ingersoll, and Ross (1985) ‘square root’ volatility specification found in the term structure literature, but the quadratic specification is novel. Both models fit nicely into the linear-quadratic framework used in the optimal control literature. Theoretically, the quadratic component makes policy more responsive to inflation shocks in the same way that an increase in the welfare weight attached to inflation does, while the linear component reduces the steady state rate of inflation. Empirical results for the period 1961-2009 underline the statistical significance of inflation-dependent US macroeconomic volatility revealed by research on the term structure. Analysis of the welfare losses associated with inflation and macroeconomic volatility shows that the conventional homoskedastic model seriously underestimates the potential welfare gains from policy optimization.

Keywords: US monetary policy, Macroeconomic volatility, Optimal control

JEL classification: C32, C61, E52

*Department of Economics and Related Studies; ps35@york.ac.uk.

†Cardiff Business School, politov@cardiff.ac.uk.

‡We are grateful to Richard Dennis for sharing his earlier data set and Gauss code with us and to Peter Burridge, Peter Smith, Andy Tremayne, Mike Wickens and other colleagues for useful comments.

1 Introduction

This paper analyzes optimal US monetary policy using a version of the ‘central bank model’ (CBM) developed by Svensson (1997) and Svensson (1999), Rudebusch and Svensson (1998), Smets (1999), Rudebusch (2002) and others. This model represents the behavior of the macroeconomy in terms of the output gap, inflation and the short term interest rate and provides a basic dynamic description of an economy in which the central bank implicitly targets inflation using a linear relationship resembling a Taylor (1993) rule. The model is estimated by maximum likelihood and initially determines the policy rate by regressing it upon inflation and the output gap. The optimal control literature analyses the effect of replacing this regression equation with an optimal linear reaction function. An incomplete list of the large literature on optimal control applied to the analysis of monetary policy includes Taylor (1979), Fuhrer and Moore (1995), Fuhrer (1997), Levin (1998), Sack (2000), Favero and Rovelli (2003), Ozlale (2003), Dennis (2006) and Salemi (2006).

Although the optimal control literature employs homoskedastic variance structures, the term structure has long recognized the significance of ‘square root volatility’.¹ In these models, the yield curve is determined by a version of the arbitrage-free model of Cox, Ingersoll, and Ross (1985) which conveniently keeps the model linear-in-variables. For example, recent heteroskedastic macro-finance models include Spencer (2008) and Bekaert, Cho, and Moreno (2010). Spencer finds that US macroeconomic volatility is strongly influenced by the steady state rate of inflation, consistent with the Okun-Friedman hypothesis that macroeconomic volatility is influenced by the underlying rate of inflation (Okun (1971) and Friedman (1977)).²

¹In this specification the volatility (standard deviation) of the short term interest rate driving the yield curve depends upon the square root of the interest rate itself. In other words, the variance structure is linear in the interest rate or other model variables.

²Ball (1992) offers a theoretical analysis of this phenomenon and the empirical evidence is examined by Holland (1995), Caporale and McKiernan (1997) and others. There is also an extensive

This finding has an important policy implication: if as this evidence suggests macroeconomic volatility is related to the level of nominal macroeconomic variables like interest or inflation rates that can be controlled by the monetary authorities they should take account of this when setting policy. Conveniently, as Spencer (2008) observes, because square root volatility models are linear-in-variables they fit into the standard linear-quadratic control framework. This is the basis of the linear variance model (M1) used in this paper. Furthermore, we show that specifications such as those of Dothan (1978) and Courtadon (1982) that relate volatility in a linear (and thus variances in a quadratic) way to the level of nominal variables also fit into this framework. This is the basis of our quadratic variance model (M2). Model M3 allows for both linear and quadratic dependence.

Theoretically, we show that the quadratic component makes policy more responsive to inflation shocks in the same way that an increase in the welfare weight attached to inflation does, while the linear component reduces the steady state rate of inflation. We show that empirically, inflation-conditional heteroskedasticity is a very significant feature of US macroeconomic data. We then ask how the leverage over volatility implied by a heteroskedastic macroeconomic model would have influenced an optimal economic policy. We follow the standard optimal monetary policy literature (see, for example, Rudebusch and Svensson (1998), Sack (2000) and Woodford (2003)) and compute the optimal policy rules implied by a quadratic central bank objective function using a dynamic programming algorithm. The estimated interest rate equation is replaced by the rule that minimizes this objective function, given estimates of the dynamic constraints represented by the other equations of the sys-

literature on the effect of inflation and macroeconomic volatility on the equity risk premium (Brandt and Wang (2003), Lettau, Ludvigson, and Wachter (2004)). Finally, there is a large literature following Chen, Karolyi, Longstaff, and Sanders (1992) and Nowman (1999) that suggests that the first and second moments of interest rates are positively associated.

tem. This approach is known as ‘direct control’ (DC). The numerical results suggest that the conventional homoskedastic control approach seriously under-estimates the potential welfare gains from policy optimization.

The paper is set out along the following lines. The next section specifies the CBM dynamic and stochastic structures and the data sources. Section 3 reports the results of estimating the CBM equations (including the interest rate equation) by maximum likelihood to obtain the four ML models. Section 4 sets out the optimal control algorithm and the theoretical results. Section 5 then reports the results obtained following the DC methodology. Section 6 provides a summary of the main findings and a conclusion.

2 Modelling the macroeconomy

In this section we specify a simple linear structure for the US economy that is potentially heteroskedastic. This based on the CBM of Rudebusch and Svensson (1998) which represents the behavior of the macroeconomy in terms of the output gap (g_t), inflation (π_t) and the policy interest rate (r_t). The inflation and interest rate data were provided by Datastream: r_t is the Fed Funds rate and π_t is the annual change in PCEX (the consumer price deflator excluding food & energy) which is the ‘core inflation’ measure the Federal Open Market Committee (FOMC) is believed to target. The output gap series g_t is the quarterly OECD measure (vintage 2009), derived from a Hodrick-Prescott filter. Table (1) reports the basic summary statistics for the data before they were de-meaned for use in subsequent analysis. Hereafter g_t , π_t and r_t refer to deviations from mean values. The ADF and KPSS statistics suggest that the inflation rate, interest rate and yields have a unit root. Further analysis suggests that this is due to a single common trend.

[INSERT TABLE 1 HERE]

2.1 The dynamic structure

The model distinguishes the current state variables of the system $\mathbf{x}'_t = \{g_t, \pi_t\}$ from the policy interest rate r_t . In the basic model these variables are all described by linear regressions. The use of a quarterly inflation rate dictates a fourth order system and we follow Rudebusch and Svensson (1998) by employing the structural relationships:

$$g_t = a_1 g_{t-1} + a_2 g_{t-2} + a_3 (r_{t-1}^a - \pi_{t-1}^a) + u_{g,t},$$

$$\pi_t = b_1 \pi_{t-1} + b_2 \pi_{t-2} + b_3 \pi_{t-3} + b_4 \pi_{t-4} + b_5 g_{t-1} + u_{\pi,t},$$

where: $\pi_t^a = \frac{1}{4} \sum_{j=0}^3 \pi_{t-j}$ and $r_t^a = \frac{1}{4} \sum_{j=0}^3 r_{t-j}$ are the annual inflation and interest rates. We impose the restriction that $b_4 = 1 - b_1 - b_2 - b_3$. This means that inflation is only stable when the output gap is zero. The first of these equations then fixes the real interest rate in line with its sample mean (i.e. $r = \pi$ for mean-adjusted data in non-accelerating inflation equilibrium). The error terms $u_{g,t}$ and $u_{\pi,t}$ are interpreted as demand and supply shocks respectively. These equations may be written in the form of a restricted VAR:

$$\mathbf{x}_t = \boldsymbol{\phi}_x \mathbf{X}_{t-1} + \boldsymbol{\theta}_x r_{t-1} + \mathbf{u}_{\mathbf{x},t}, \quad (1)$$

where the state vector $\mathbf{X}'_t = \{\mathbf{x}'_t, r_{t-1}, \mathbf{x}'_{t-1}, r_{t-2}, \mathbf{x}'_{t-2}, r_{t-3}, \mathbf{x}'_{t-3}\}$ is a 11×1 vector of current and lagged state variables, $\mathbf{u}'_{\mathbf{x},t} = \{u_{g,t}, u_{\pi,t}\}$ is the errors vector, $\boldsymbol{\theta}'_x = \{a_3/4, 0\}$ is a vector showing the lagged response of the current state vector to a change in the policy instrument and:

$$\boldsymbol{\phi}_x = \begin{bmatrix} a_1 & -a_3/4 & a_3/4 & a_2 & -a_3/4 & a_3/4 & 0 & -a_3/4 & a_3/4 & 0 & -a_3/4 \\ b_5 & b_1 & 0 & 0 & b_2 & 0 & 0 & b_3 & 0 & 0 & b_4 \end{bmatrix}.$$

The interest rate can be described by a similar regression equation, where the parameters are determined by either an estimation (section 3) or optimization (sections 4 and 5) procedure:

$$\begin{aligned}
r_t &= c_1 g_{t-1} + c_2 g_{t-2} + c_3 \pi_{t-1} + c_4 \pi_{t-2} + c_5 \pi_{t-3} + c_6 \pi_{t-4} \\
&\quad + \theta_r r_{t-1} + c_7 r_{t-2} + c_8 r_{t-3} + c_9 r_{t-4} + u_{r,t} \\
&= \phi_r \mathbf{X}_{t-1} + \theta_r r_{t-1} + u_{r,t}
\end{aligned} \tag{2}$$

where:

$$\phi_r = \{c_1, c_3, c_7, c_2, c_4, c_8, 0, c_5, c_9, 0, c_6\}.$$

Using the bar notation to denote steady state values, we note for use in section 5.2 that steady state impact of \bar{g} and $\bar{\pi}$ on the policy rate is:

$$\bar{r}_t = \bar{\kappa} + \bar{\phi}_g \bar{g} + \bar{\phi}_\pi \bar{\pi} \tag{3}$$

where:

$$\bar{\phi}_g = (c_1 + c_2) / (\theta_r + c_7 + c_8 + c_9) \tag{4}$$

$$\bar{\phi}_\pi = (c_3 + c_4 + c_5 + c_6) / (\theta_r + c_7 + c_8 + c_9). \tag{5}$$

$\bar{\kappa} = 0$ for the mean-adjusted ML models, but can be non-zero when the equilibrium is shifted by a change in the rate setting equation, as in some of the optimized models for example. With $\bar{r} = \bar{\pi}, \bar{g} = 0$ in a non-accelerating inflation equilibrium we have:

$$\bar{r} = \bar{\pi} = \bar{\kappa} / (1 - \bar{\phi}_\pi)$$

where the denominator is negative under the restriction $\bar{\phi}_\pi > 0$. Stacking these equations gives the system:

$$\begin{bmatrix} \mathbf{x}_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_x \theta_x \\ \phi_r \theta_r \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{\mathbf{x},t} \\ u_{r,t} \end{bmatrix}. \quad (6)$$

The evolution of the state vector is described by the companion form:

$$\mathbf{X}_t = \Theta r_{t-1} + \Phi \mathbf{X}_{t-1} + \mathbf{U}_t, \quad (7)$$

where $\Theta' = \{\theta'_r, \mathbf{0}'_9\}$, and $\mathbf{U}'_t = \{\mathbf{u}'_{\mathbf{x},t}, \mathbf{0}'_9\}$ are 11×1 deficient coefficient and error vectors and:

$$\Phi = \begin{bmatrix} \{a_1, -a_3/4\} & a_3/4 & \{a_2, -a_3/4\} & a_3/4 & \{0, -a_3/4\} & a_3/4 & \{0, -a_3/4\} \\ \{b_5, b_1\} & 0 & \{0, b_2\} & 0 & \{0, b_3\} & 0 & \{0, b_4\} \\ \mathbf{0}'_2 & 0 & \mathbf{0}'_2 & 0 & \mathbf{0}'_2 & 0 & \mathbf{0}'_2 \\ \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} \\ \mathbf{0}'_2 & 1 & \mathbf{0}'_2 & 0 & \mathbf{0}'_2 & 0 & \mathbf{0}'_2 \\ \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} \\ \mathbf{0}'_2 & 0 & \mathbf{0}'_2 & 1 & \mathbf{0}'_2 & 0 & \mathbf{0}'_2 \\ \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_{2,2} \end{bmatrix}.^3$$

The use of de-meanded data means that these regression systems do not include intercept constants.

³In this paper, $\mathbf{0}_a$ is the $(a \times 1) \times 1$ zero vector; $\mathbf{1}_a$ is the $(a \times 1) \times 1$ summation vector; $\mathbf{0}_{a,b}$ the $(a \times b)$ zero matrix; and \mathbf{I}_a the a^2 identity matrix. $Diag\{\delta\}$ represents a matrix with the elements of the row vector δ in the main diagonal and zeros elsewhere. $tr\{A\}$ represents the trace of the square matrix A : the sum of its diagonal elements.

2.2 The stochastic structure

The residuals in (6) and (7) are potentially heteroskedastic. Indeed, preliminary Breusch (1979) tests immediately confirmed the conditional heteroskedasticity suggested by the finance literature. These tests regressed the squared errors from regressions of the output gap, inflation and interest rates in the homoskedastic ML0 model on various lagged indicator variables (and their squares). We began with the 10 year Treasury yield, which is normally a good proxy for the underlying inflation rate in a macro-finance model. However we found that annual inflation rates fitted much better, with PCEX inflation performing best of all.⁴ Thus, following the empirical yield curve literature we adopted a single factor (A_1) volatility model,⁵ which is driven by the annual PCEX inflation rate (lagged once):

$$\begin{bmatrix} \mathbf{u}_{\mathbf{x},t} \\ u_{r,t} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0}_2 \\ \mathbf{g}' & 1 \end{bmatrix} \begin{bmatrix} \mathbf{D}_{x,t}^{1/2} & \mathbf{0}_2 \\ \mathbf{0}'_2 & \delta_{r,t}^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{\mathbf{x},t} \\ v_{r,t} \end{bmatrix} \quad (8)$$

where $\mathbf{G} = \begin{bmatrix} 1 & 0 \\ g_1 & 1 \end{bmatrix}$ and $\mathbf{g}' = \{g_2, g_3\}'$ include constant parameters, whereas $\mathbf{v}_{\mathbf{x},t} = \{v_{g,t}, v_{\pi,t}\}' \sim i.i.d. (\mathbf{0}_2, \mathbf{I}_2)$ and $v_{r,t} \sim i.i.d. (0, 1)$ are unit i.i.d. Gaussian homoskedastic shocks to output, inflation and interest rate respectively; and

$$\begin{aligned} \mathbf{D}_{x,t} &= \mathbf{D}_{x,0} + \mathbf{D}_{x,1}\pi_{t-1}^a + \mathbf{D}_{x,2}(\pi_{t-1}^a)^2 \\ &= \mathbf{D}_{x,0} + \mathbf{D}_{x,1}\mathbf{s}'_{\pi}\mathbf{X}_{t-1} + \mathbf{D}_{x,2}\mathbf{X}'_{t-1}\mathbf{S}_{\pi}\mathbf{X}_{t-1} \\ \delta_{r,t} &= \delta_{r,0}\mathbf{D}_{x,0} + \delta_{r,1}\mathbf{s}'_{\pi}\mathbf{X}_{t-1} + \delta_{r,2}\mathbf{X}'_{t-1}\mathbf{S}_{\pi}\mathbf{X}_{t-1} \end{aligned} \quad (9)$$

⁴This result was confirmed in the final model specification. Replacing the annual PCEX π_{t-1}^a inflation in (9) by the 10 year bond yield increases the loglikelihood for ML0 shown in table 2 from (-)156 to (-)104.2, but this is greater than that for the PCEX-based ML3 ((-)62.6). Replacing inflation by the combined volatility-inflation factor estimated in the A_1 macro-finance model (M1) in Spencer also gives a lower loglikelihood ((-) 96.5) than ML3-PCEX.

⁵Appendix 1 reports the restrictions that we use to ensure that this is 'admissible', i.e. that the variance structure remains non-negative definite.

where: $\mathbf{D}_{x,i} = \text{diag}\{\delta_{g,i}, \delta_{\pi,i}\}$, $i = 0, 1, 2$, with $\delta_{g,i}$, and $\delta_{\pi,i}$ denoting constant parameters; $\mathbf{s}_\pi = 0.25[0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0]'$ and $\mathbf{S}_\pi = \mathbf{s}_\pi \mathbf{s}'_\pi$ select annual inflation and squared annual inflation respectively from the state vector.⁶ Similarly, for the state vector (7) we define

$$\mathbf{U}_t \sim N(\mathbf{0}_{11}, \boldsymbol{\Sigma}_t)$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_t &= \mathbf{E}[\mathbf{U}_t \mathbf{U}'_t | \mathbf{X}_{t-1}] = \boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1 \pi_{t-1}^a + \boldsymbol{\Sigma}_2 (\pi_{t-1}^a)^2 \\ &= \boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1 \mathbf{s}'_\pi \mathbf{X}_{t-1} + \boldsymbol{\Sigma}_2 \mathbf{X}'_{t-1} \mathbf{S}_\pi \mathbf{X}_{t-1} \end{aligned} \quad (10)$$

with

$$\begin{aligned} \boldsymbol{\Sigma}_i &= \Gamma \boldsymbol{\Delta}_i \Gamma', \\ \Gamma &= \begin{bmatrix} \mathbf{G} & \mathbf{0}_{2,9} \\ \mathbf{0}_{9,2} & \mathbf{0}_{11,11} \end{bmatrix}, \end{aligned}$$

and: $\boldsymbol{\Delta}_i = \text{diag}\{\delta_{g,i}, \delta_{\pi,i}, \mathbf{0}'_9\}$, $i = 0, 1, 2$.

We label this general model ML3 and test for linear and quadratic dependence by comparing its performance with three special cases using maximum likelihood. Model ML0 assumes that the volatility structure is homoskedastic using the restrictions $\mathbf{D}_{x,i} = \mathbf{0}_{2,2}$ and $\delta_{r,i} = 0$, for $i = 1, 2$, so that the covariance matrix in equation (10) simplifies to $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_0$. ML1 assumes that the variance-covariance matrix is linear in the lagged annual inflation rate (using the restrictions $\mathbf{D}_{x,2} = \mathbf{0}_{2,2}, \delta_{r,2} = 0$) while ML2 assumes that it is quadratic ($\mathbf{D}_{x,1} = \mathbf{0}_{2,2}, \delta_{r,1} = 0$).

⁶This arrangement is more general than the specific model employed in this paper since (by redefining \mathbf{s} and \mathbf{S}) it could allow volatility to be a quadratic function of any linear combination of state variables.

3 The empirical (ML) models

Estimating (1) and (2) simultaneously by maximum likelihood (as outlined in appendix 2) gives the basic ML model structures:

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_t \\ r_t \end{bmatrix} &= \begin{bmatrix} \hat{\mathbf{x}}_t \\ \hat{r}_t \end{bmatrix} + \begin{bmatrix} \hat{\mathbf{u}}_{\mathbf{x},t} \\ \hat{u}_{r,t} \end{bmatrix} \\ &= \begin{bmatrix} \hat{\phi}_{\mathbf{x}} & \hat{\theta}_{\mathbf{x}} \\ \hat{\phi}_r & \hat{\theta}_r \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{G}\mathbf{D}_{x,t}^{1/2} & \mathbf{0}_3 \\ \mathbf{g}'\mathbf{D}_{x,t}^{1/2} & \delta_{r,t}^{1/2} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}}_{\mathbf{x},t} \\ \hat{v}_{r,t} \end{bmatrix}. \end{aligned} \quad (11)$$

In this paper ‘hats’ are used like this to denote ML estimators. Table (2) shows the loglikelihood values for the four ML specifications.

The linear-quadratic variance model ML3 has 29 parameters⁷ and a loglikelihood value of (-) 62.6. This nests the other three models, allowing their restrictions to be tested using loglikelihood ratio tests. The quadratic variance model ML2 has a loglikelihood value of (-) 106.4 and is rejected against the encompassing model as is the linear variance model ML1. The homoskedastic model ML0 is decisively rejected against all of the other three models. These tests strongly support the Okun-Friedman hypothesis: conditioning the variance structure on lagged inflation dramatically increases the likelihood.

[INSERT TABLE 2 HERE]

3.1 ML3 parameter estimates and residuals

We now discuss the parameters of the best baseline model: ML3. Estimates of the model parameters are shown in table (3). The one-quarter-ahead forecast values and 95% confidence intervals for the three macro variables are shown in figures 1-3. This

⁷These are: a vector \mathbf{a} comprising three parameters a_1 , a_2 and a_3 , i.e. $\mathbf{a}(3)$; $\mathbf{b}(4)$; $\mathbf{c}(9)$; θ_r ; $\mathbf{G}(1)$; $\mathbf{g}(2)$; $\mathbf{D}_0(3)$; $\mathbf{D}_1(3)$ and $\mathbf{D}_2(3)$.

model conditions the variance structure using both linear and quadratic inflation terms. However, the latter are only significant in the case of the output gap and interest rate (see the parameters $\delta_{g,2}$ and $\delta_{r,2}$ in the table). This effect is quite evident in figures 1 and 3, meaning that volatility is particularly high and variable between 1974 and 1984, relative low and stable otherwise. The last of these charts suggests that the likelihood of policy mistakes increases as a quadratic function of the inflation rate. These error bands reflect the large residuals seen over the period (shown in the lower panels). In contrast, $\delta_{\pi,2}$ is small and insignificant, so the inflation rate follows a standard CIR (1985) process.⁸ These results are consistent with the general finding of stochastic volatility in term structure (Chen and Scott (1993), Spencer (2008)) and univariate models (Chen, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1996), Stanton (1997) and others). These charts all show a significant reduction in volatility since the mid 1990s - the period of the ‘Great Moderation’. This low volatility period was interrupted by the recent credit crunch, which is reflected in large negative shocks in the final quarter of 2008 and first quarter of 2009, following the collapse of Lehman Brothers. However, conditional upon these shocks, the subsequent residuals remain low, consistent with the low volatility implied by the low level of inflation.

[INSERT TABLE 3 HERE]

[INSERT FIGURES 1, 2 AND 3 HERE]

3.2 The empirical impulse responses

The dynamic properties of these models can be seen from the impulse responses, which show the effects of innovations in the macroeconomic variables on the system. Because these innovations are correlated empirically, we work with orthogonalized

⁸The discrete time CIR model has a non-central χ^2 distribution analysed in Spencer (2008). Sun (1982) offers a Gaussian approximation.

innovations using the triangular factorization defined in (8). The orthogonalized impulse responses show the effect on the macroeconomic system of increasing each of these shocks by one percentage point for just one period using the Wald representation of the system. This arrangement is affected by the ordering of the macroeconomic variables in the vector x_t . We adopt the standard ordering: $\{g_t, \pi_t, r_t\}$, interpreting v_g as a positive demand shock and (v_π) as a negative supply shock.

These responses are shown in Figure 4. The impulse response functions for ML0 are shown by dotted lines, those for ML1 by dashed lines and those for ML3 by continuous lines. The shape of the ML0 responses is similar to that reported in Figure 1 of Rudebusch and Svensson (1998). However, the most striking feature of these ML results is that these temporary shocks appear to have permanent effects on inflation and interest rates. This is also clear from Rudebusch and Svensson (1998) and reflects the comment we made on table (1): these variables are non-stationary over this period.⁹ However, replacing the empirical monetary policy relationship used in the ML models by an optimal rule should eliminate the persistence of these impulse responses (see Lungqvist and Sargent (2004), pages 114-115). The next section explains how this rule is derived.

[INSERT FIGURE 4 HERE]

⁹Macromodellers have reacted to this apparent non-stationarity in several different ways. On the one hand, monetary modellers (for example, Rudebusch and Svensson (1998), Clarida, Gali, and Gertler (2000), Stock and Watson (2001)) have typically argued that ADF tests have low power against a near unit root alternative and proceed under the assumption that inflation and nominal interest rate series mean-revert slowly. On the other hand, term structure modellers are quite happy to acknowledge that bond yields are characterized by a non-stationary common trend (Ang and Piazzesi (2003), Rudebusch and Wu (2008) and Dewachter, Lyrio, and Maes (2006)). This is typically interpreted as the underlying rate of inflation. However, our paper suggests that this phenomenon reflects an accommodative monetary policy stance: one that only partially offsets the effect of shocks, allowing them to have permanent effects on nominal rates.

4 Optimal monetary policy

4.1 The quadratic loss function

We now adopt the standard specification introduced by Rudebusch and Svensson (1998), which assumes that the FOMC aims to minimize the variation in the output gap; the variation in the annual inflation rate around its mean, and the change in the funds rate. Specifically, the loss function is quadratic in these three arguments:

$$L_t = E_t \sum_{j=0}^{\infty} \beta^{t+j} \left[\lambda (\pi_{t+j}^a)^2 + \mu g_{t+j}^2 + v (\Delta r_{t+j})^2 \right], \quad (12)$$

where E_t is the conditional expectations operator; β is the discount factor; $\Delta r_{t+j} = r_{t+j} - r_{t+j-1}$; and the parameters $\lambda \geq 0$; $\mu \geq 0$ and $v \geq 0$ show the weights given to inflation, output gap and interest rate stabilization. In the empirical model we use the standard normalization $\lambda = 1$, but at this point we keep the welfare weight attached to inflation explicit in order to compare its effect with that of quadratic volatility in section 4.3. Quadratic loss functions such as (12) are widely employed in the inflation targeting literature and typically justified on the ground that they provide a good approximation to the expected lifetime utility of a representative household derived from a fully micro-founded macroeconomic model of the economy (See, for example, Rotemberg and Woodford (1997) and Woodford (2003)). It is convenient to express the loss function as:

$$L_t = \sum_{j=0}^{\infty} \beta^{t+j} E_t [\mathbf{X}'_{t+j} \mathbf{\Lambda} \mathbf{X}_{t+j} + v r_{t+j}^2 + 2 \mathbf{X}'_{t+j} \mathbf{H} r_{t+j}], \quad (13)$$

where:

$$\mathbf{\Lambda} = \lambda \mathbf{S}_{\pi} + \mu \mathbf{S}_{\mathbf{g}} + v \mathbf{S}_{\mathbf{r}}; \mathbf{H} = -v \mathbf{s}_r, \quad (14)$$

$\mathbf{S}_g = \mathbf{s}_g \mathbf{s}_g'$ and $\mathbf{S}_r = \mathbf{s}_r \mathbf{s}_r'$, and where $\mathbf{s}_g = \{1, \mathbf{0}'_{10}\}'$ and $\mathbf{s}_r = \{\mathbf{0}'_2, 1, \mathbf{0}'_8\}'$ are selection vectors that pick g_t and r_{t-1} respectively from \mathbf{X}_t .

4.2 The optimal policy rule

The FOMC is assumed to choose the intertemporal sequence of policy rates $\{r_{t+j}\}_{j=0}^{\infty}$ that minimizes the loss function in (13) given the model of the economy in (7) and the initial state vector \mathbf{X}_{t-1} . Since the loss is quadratic in every period and the constraints are linear, the value function is a quadratic form in \mathbf{X}_t given by

$$J(\mathbf{X}_t) = \min_{\{r_{t+j}\}} L_t \quad (15)$$

$$= \mathbf{X}'_t \mathbf{P} \mathbf{X}_t - 2\mathbf{X}'_t \mathbf{p} + c. \quad (16)$$

Substituting (13) into (15) and using (16) to replace $J(\mathbf{X}_{t+1})$ gives the Bellman equation:

$$\begin{aligned} J(\mathbf{X}_t) &= \min_{r_t} [\mathbf{X}'_t \mathbf{\Lambda} \mathbf{X}_t + v r_t^2 + 2\mathbf{X}'_t \mathbf{H} r_t + \beta E_t J(\mathbf{X}_{t+1})] \\ &= \min_{r_t} [\mathbf{X}'_t \mathbf{\Lambda} \mathbf{X}_t + v r_t^2 + 2\mathbf{X}'_t \mathbf{H} r_t + \beta E_t (\mathbf{X}'_{t+1} \mathbf{P} \mathbf{X}_{t+1} - 2\mathbf{X}'_{t+1} \mathbf{p} + c)] \end{aligned}$$

which is optimized subject to (7). Evaluating expectations using (7) and (10) gives:

$$J(\mathbf{X}_t) = \min_{r_t} \left\{ \begin{aligned} &\mathbf{X}'_t \mathbf{\Lambda} \mathbf{X}_t + v r_t^2 + 2\mathbf{X}'_t \mathbf{H} r_t + \beta (\mathbf{\Theta} r_t + \mathbf{\Phi} \mathbf{X}_t)' \mathbf{P} (\mathbf{\Theta} r_t + \mathbf{\Phi} \mathbf{X}_t) - 2\beta (\mathbf{\Theta} r_t + \mathbf{\Phi} \mathbf{X}_t)' \mathbf{p} \\ &+ \beta c + \beta tr(\mathbf{P} \Sigma_0) + \beta tr(\mathbf{P} \Sigma_1) (\mathbf{s}'_{\pi} \mathbf{X}_t) + \beta tr(\mathbf{P} \Sigma_2) (\mathbf{X}'_t \mathbf{S}_{\pi} \mathbf{X}_t) \end{aligned} \right\}, \quad (17)$$

where the last three terms capture the impact of the variance structure on the optimal policy. Differentiating this w.r.t. r_t yields the closed loop solution:

$$r_t^* = \zeta + \boldsymbol{\xi} \mathbf{X}_t \quad (18)$$

$$\zeta = (v + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Theta})^{-1} \beta \boldsymbol{\Theta}' \mathbf{p} \quad (19)$$

$$\boldsymbol{\xi} = - (v + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Theta})^{-1} (\mathbf{H}' + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Phi}). \quad (20)$$

Equation (18) shows that the optimal policy rule is linear in the current state vector, with (19) and (20) determining its intercept and slope coefficients. Substituting (18) - (20) back into (17) and equating with (16) gives:

$$\mathbf{p} = -\frac{1}{2} \beta \text{tr}(\mathbf{P} \boldsymbol{\Sigma}_1) [\mathbf{I} - \beta (\boldsymbol{\Theta} \boldsymbol{\xi} + \boldsymbol{\Phi})']^{-1} \mathbf{s}_\pi, \quad (21)$$

and the Riccati equation:

$$\mathbf{P} = \{\boldsymbol{\Lambda} + \beta \text{tr}(\mathbf{P} \boldsymbol{\Sigma}_2) \mathbf{S}_\pi\} - (\mathbf{H}' + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Phi})' (v + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Theta})^{-1} (\mathbf{H}' + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Phi}) + \beta \boldsymbol{\Phi}' \mathbf{P} \boldsymbol{\Phi}. \quad (22)$$

The optimal response coefficients are obtained by solving (22) numerically and substituting \mathbf{P} back into (21), (19) and (20). The standard homoskedastic solution occurs as a special case by setting $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{0}$ in (21) and (22) respectively. In this case, inspection of (21) shows that $\mathbf{p} = \mathbf{0}$ and then (19) shows the intercept is also zero. This means that we can infer the target rate of inflation from the sample mean in the standard model: if the (demeaned) state vector $\mathbf{X}_t = \mathbf{0}$ then $r = \bar{r}$ and the system is at rest. The certainty equivalence principle also holds in this case: since the covariance matrix (10) does not affect the coefficients of the optimal policy rule, this rule is the same as it would be in the equivalent deterministic problem.

4.3 The effects of conditional volatility on the optimal policy

When $\Sigma_1 \neq \mathbf{0}$ and/or $\Sigma_2 \neq \mathbf{0}$, the optimal policy rule is affected by these matrices and certainty equivalence does not hold. Quadratic dependence (the matrix Σ_2) affects \mathbf{P} in (22), which affects the slope coefficients (20) as well as the intercept (19). It works through the scalar $\beta \text{tr}(\mathbf{P}\Sigma_2) \geq 0$ which is non-negative because, like Σ_2 , \mathbf{P} is a positive semi-definite matrix under standard regularity conditions (Lungqvist and Sargent (2004), page 113).¹⁰ This term makes policy more aggressive in the sense that it has exactly the same effect on the response coefficients in (20) as an increase in the homoskedastic welfare weight λ does. That is because as inspection of (18) to (22) shows, the matrices $\beta \text{tr}(\mathbf{P}\Sigma_2) \mathbf{S}_\pi$ and Λ only affect the slope coefficients through their effect on \mathbf{P} , working through the combined term shown in curly brackets in (22). We can replace this term by the adjusted Λ coefficient matrix $\tilde{\Lambda} = \Lambda + \beta \text{tr}(\mathbf{P}\Sigma_2) \mathbf{S}_\pi = \tilde{\lambda} \mathbf{S}_\pi + \mu \mathbf{S}_g + \nu \mathbf{S}_r$ where, using (14), the scalar $\tilde{\lambda} = (\lambda + \beta \text{tr}(\mathbf{P}\Sigma_2)) \geq \lambda$. In other words, introducing quadratic inflation-dependent volatility has the effect of increasing the effective weight $\tilde{\lambda}$ given to inflation: we could get the same level of aggression by increasing λ in the standard model.

Linear dependence only affects the intercept coefficient: the matrix Σ_1 only appears the right hand side of equation (21), which affects (19), but not (20). We now show that the term $\beta \text{tr}(\mathbf{P}\Sigma_1)/2 \geq 0$ has the same effect as a reduction in the annual inflation target. To analyze the effect of a change in the target of π^* (compared to the initial target or sample mean, which is removed by de-meaning) we define the state vector relative to the new target as $(\mathbf{X}_t - \mathbf{X}^*)$ and re-work the algebra with a new objective function that replaces $\mathbf{X}_t' \Lambda \mathbf{X}_t$ in the first period welfare loss term in

¹⁰The inequality follows from the well-known properties of positive semidefinite Hermitian matrices. These are square symmetric matrices with non-negative eigenvalues. They have a non-negative trace and can be factorised as $\Sigma_2 = \mathbf{F}'\mathbf{F}$. Using this factorisation with the basic rule for the trace of a product: $\text{tr}(\mathbf{P}\Sigma_2) = \text{tr}(\mathbf{P}\mathbf{F}'\mathbf{F}) = \text{tr}(\mathbf{F}\mathbf{P}\mathbf{F}') \geq 0$, where the inequality follows from the fact that if \mathbf{P} is positive semidefinite then so is $\mathbf{F}\mathbf{P}\mathbf{F}'$.

(17) by $(\mathbf{X}_t - \mathbf{X}^*)' \Lambda (\mathbf{X}_t - \mathbf{X}^*)$, where $\mathbf{X}^* = \mathbf{s}_\pi \pi^*$. This does not affect \mathbf{P} , but has the effect of introducing a new term into (21), which shows this equivalence:

$$\begin{aligned} \mathbf{p} &= \{\mathbf{I} - \beta (\Theta \boldsymbol{\xi} + \Phi)'\}^{-1} \left[\Lambda \mathbf{X}^* - \frac{1}{2} \beta \text{tr}(\mathbf{P} \Sigma_1) \mathbf{s}_\pi \right] \\ &= \{\mathbf{I} - \beta (\Theta \boldsymbol{\xi} + \Phi)'\}^{-1} \mathbf{s}_\pi \left[\lambda \pi^* - \frac{1}{2} \beta \text{tr}(\mathbf{P} \Sigma_1) \right]. \end{aligned} \quad (23)$$

(using $\Lambda \mathbf{X}^* \mathbf{s}_\pi = \lambda \mathbf{S}_\pi \mathbf{s}_\pi \pi^* = \mathbf{s}_\pi \lambda \pi^*$).

5 Numerical evaluation

The previous section shows how the coefficients of the optimal policy rule (18) can be determined numerically, given values of the preference parameters β , λ , μ and ν and parameter values for the state equations (1). In this section we follow Rudebusch and Svensson (1998), overwriting the empirical parameters of the policy rate equations in ML0 and ML3 by their optimized values in order to compare the behavior of the system under control with the historical data.¹¹ We also look at the cost of informing the policy rule using a homoskedastic macromodel (M0) when the true model (M3) exhibits linear-quadratic inflation-conditional volatility.

5.1 The coefficients of the optimal policy rule

The model under control consists of the state equations (1) and the policy rule (18). Following Polito and Wickens (2010), we assume that there are no policy errors; partition the vector $\boldsymbol{\xi} = \{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3\}$ conformably with $\{\mathbf{x}_t, \mathbf{X}_{t-1}, r_{t-1}\}$ and

¹¹In view of the superior performance of the M3 model we do not report the results of running this exercise for the linear (M1) and quadratic (M2) dependence models here. However, these results are available upon request.

stack these equations to obtain

$$\begin{bmatrix} \mathbf{I}_2 & 0 \\ -\boldsymbol{\xi}_1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ r_t^* \end{bmatrix} = \begin{bmatrix} \mathbf{0}_2 \\ \zeta \end{bmatrix} + \begin{bmatrix} \boldsymbol{\phi}_x & \boldsymbol{\theta}_x \\ \boldsymbol{\xi}_2 & \xi_3 \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{x,t} \\ 0 \end{bmatrix}. \quad (24)$$

This can be solved to obtain the reduced form system which is congruent with (11):

$$\begin{bmatrix} \mathbf{x}_t \\ r_t^* \end{bmatrix} = \begin{bmatrix} \mathbf{0}_2 \\ \zeta \end{bmatrix} + \begin{bmatrix} \boldsymbol{\phi}_x & \boldsymbol{\theta}_x \\ \boldsymbol{\phi}_{r^*} & \theta_{r^*} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{x,t} \\ \xi_1 \mathbf{u}_{x,t} \end{bmatrix}; \quad (25)$$

where:

$$\boldsymbol{\phi}_{r^*} = \boldsymbol{\phi}_{r^*}(\mathcal{P}) = \boldsymbol{\xi}_1(\mathcal{P})\boldsymbol{\phi}_x + \boldsymbol{\xi}_2(\mathcal{P}), \quad (26)$$

$$\theta_{r^*} = \theta_{r^*}(\mathcal{P}) = \boldsymbol{\xi}_1(\mathcal{P})\boldsymbol{\theta}_x + \xi_3(\mathcal{P}),$$

$$\boldsymbol{\xi}_1 = \boldsymbol{\xi}_1(\mathcal{P})$$

$$\zeta = \zeta(\mathcal{P})$$

and where the relationships $\zeta(\mathcal{P})$, $\boldsymbol{\xi}_1(\mathcal{P})$, $\boldsymbol{\xi}_2(\mathcal{P})$ and $\xi_3(\mathcal{P})$ follow from the restrictions (19)-(22) and (26), with $\mathcal{P} = \{\boldsymbol{\phi}_x, \boldsymbol{\theta}_x, \mu, v, \beta\}$.¹² Substituting stylized values for β , μ and v as well as the ML parameters $\hat{\boldsymbol{\phi}}_x$ and $\hat{\boldsymbol{\theta}}_x$ from (11) into (25) we can formulate the Direct Control (DC) model that optimizes the welfare function.

5.2 Steady state interest rate responses

Table (4) presents the long run coefficients implied by the DC0, DC1 and DC3 optimal policy rules under the five different specifications of the welfare weights used by Rudebusch and Svensson (1998), along with the long run coefficients of policy

¹²We normalise the welfare weight to inflation to unity, so that μ and v measure the welfare weight attached to output gap stabilisation and interest rate smoothing relative to inflation stabilisation.

rules estimated from ML0, ML1 and ML3. We used a stylized assumption about the discount rate, which is set at four percent in this section.

The first parameter set, which gives inflation, output and interest rate smoothing equal weight in the loss function ($\mu = \nu = 1$), makes the FOMC respond more aggressively to changes in inflation and output if heteroskedasticity is ignored (compare ML0 and DC0). As in previous optimal control studies, the optimal long run responses to inflation and output are larger in absolute value than those from the empirical estimates. This is true for both models, although the long run responses of the optimal policy rule obtained from ML3 are much larger than those obtained from ML0. Also, the intercept constants are large and positive, subtracting about one percentage point from the steady state inflation and interest rates compared to the homoskedastic model. These responses are all much larger than those proposed by Taylor (1993).

[INSERT TABLE 4 HERE]

5.3 The optimal impulse responses

Section 3.2 discussed the impulse responses for the three ML models. In this section we compare the results for DC0, DC1 and DC3, which are shown in figure 5 as the dotted, dashed and continuous schedules respectively. We present the results for the benchmark welfare parameters $\mu = \nu = 1$ since the difference between the impulse response functions from DC1 and DC3 models relative to those from DC0 is not markedly affected by alternative choices of the welfare parameters. The policy rate is far more responsive to output and inflation shocks than under the empirical rules, no matter which model is employed. This is reflected in the much sharper response in output shown in the top panel. Moreover, these shocks have only temporary effects on inflation and interest rates. In other words, the more active policy response

has the effect of stabilizing the system, making it stationary. This is true for the homoskedastic model, but the point emerges more clearly from the DC3 responses.¹³ The impulse responses for DC0 and DC1 are similar, but consistent with the theory of the previous section, those for DC3 are far more aggressive.

This finding has implications for the yield curve literature that originally motivated this paper. It suggests that the non-stationary common trend observed in bond yields, inflation and interest rates is the result of a passive monetary policy rather than a non-stationary drift in the target inflation rate. These results suggest that it would not be necessary to allow for unit roots in term structure models had the authorities pursued a fixed nominal objective more aggressively. Similarly, we would not be wondering how to deal with the unit roots revealed by the macroeconomic literature.

Having said that, we would accept that we are working with the benefit of hindsight, or at least a much better data set than was available to the Fed in real time. In particular, we have much better estimates of potential output and the output gap than were available historically. Indeed, Orphanides (2003) suggests that ‘the dismal economic outcomes of the Great Inflation may have resulted from the pursuit of activist policies in the face of overoptimistic assessments of the output gap associated with the productivity slowdown of the late 1960s and early 1970s’. This may explain the superior performance of the DC0 rule over the empirical relationship embedded in the historical data. However even if the Fed had the same data as we now do and had followed the optimal DC0 rule, there remains the question of how much a failure to recognize the significance of inflation-conditional heteroskedasticity would have cost in terms of welfare. We address this question in the next section.

¹³Again, the shape of the impulse response functions for DC0 is similar to that reported in Figure 1 of Rudebusch and Svensson (1998).

[INSERT FIGURE 5 HERE]

5.4 Counterfactual analysis

Recall that although the output gap is mean-reverting in the empirical ML models, the observed inflation and interest rates are non-stationary over this period, consistent with the passive policy stance. What would have happened if the FOMC had been following an optimal policy and stabilized the system?

Figures 6, 7 and 8 plot the difference between the paths of g , π and r generated by DC0, DC1 and DC3 and their historical paths. The associated welfare losses are reported in Table (5) and discussed below. This exercise uses the benchmark preference parameters. The simulated paths begin with the initial state vector in the first quarter of 1962. This is first perturbed using the homoskedastic residuals (u) generated by ML0 and iterated forward over time using the ML0 transition equations (1) for the non-policy variables g and π alongside the ML0-based optimal policy rule. This gives the path shown by the dotted line and and labelled DC0. This is then perturbed using the homoskedastic residuals (v) generated by ML1, mapping them into the heteroskedastic residuals (u) using (8) and iterating forward over time using the ML1 transition equations (1) for g and π alongside the implied optimal policy rule. This gives the path shown by the dashed line and and labelled DC1. Repeating this using the output of the ML3 model gives the results for DC3, shown as a continuous line.

These simulations illustrate the theoretical results quite nicely. The steady state inflation effect is apparent in figure 7 which shows that the paths of inflation in DC1 and DC3 are both generally lower than for DC0. The relative aggressiveness of the DC responses is apparent in figure 6 which shows that although the authorities would have reacted to the positive output gap and the upward drift in inflation in the 1960s

by hiking interest rates under both DC0 and DC1, the hike under DC3 would have been more pronounced. A similar pattern is evident in response to the oil shock of 1973-74. Then, with inflation under much tighter control, the policy rate would have been lower following the oil price shock of 1979-80. Indeed all three models suggest that the authorities would have been in a position to have reduced interest rates just as they did following the ‘credit crunch’ of 2007-08.

[INSERT FIGURES 6, 7 AND 8 HERE]

5.5 Welfare analysis

We have seen that policy optimization would have corrected the drift in the steady state nominal rates implied by the macro ML models. How would this have affected the volatility of the system and the imputed welfare losses? To answer this we follow Rudebusch and Svensson (1998) and assume for simplicity in this section that $\beta = 1$. They show that if the policy coefficients and counterfactual simulations are re-run under this assumption then the welfare loss is the welfare-parameter-weighted average of the historical or simulation variances. For example, if we use the benchmark weights $v = \mu = 1$, the loss is simply the sum of the squared standard deviations. This welfare specification is analyzed in table (5). This table shows, in line with the results of Rudebusch and Svensson (1998) and others, that policy optimization using the homoskedastic model reduces the welfare loss, from 10.4 in the historical data (and equivalently all the ML models) to 9.83 for DC0, a fall of 0.57. However, taking into account the heteroskedasticity of the system reduces the loss to 8.59 in DC1 and then to 8.08 in DC3, a fall of 2.32 that is nearly four times larger.

In the DC0 specification, the gain is achieved by reducing inflation variability at the expense of increasing the variability of both output and interest rates. The variances are fixed in this specification ($\Sigma_t = \Sigma_0$), so one type of uncertainty has

to be traded off against another. However, lowering the steady state inflation rate reduces the overall variability of the heteroskedastic systems. DC1 is able to reduce the variability of both inflation and interest rates compared to DC0, with a negligible increase in the variability of output, achieving a very significant improvement in welfare. The standard deviations fall further in DC3. The standard deviation of the annual inflation rate is 1.06, about half the 2.11 standard deviation measured in the sample, while the standard deviation of the output gap falls from 2.24 to 2.08. This is obtained at the cost of an increase in the volatility of policy rate changes of about 2/3: from 0.97 to 1.62. The gains would appear to be very large once heteroskedasticity is taken into account, much larger than suggested by the conventional model. Thus, it seems that model mis-specification leads the researcher to seriously understate the potential gains from optimization.

However, there remains the question of what would have happened if the Fed had access to the same data as we now do but had failed to recognize the significance of inflation-conditional heteroskedasticity and to take this into account in the formulation of monetary policy. To address this, we run a simulation in which the FOMC believes that the variance of the errors is homoskedastic and uses the optimal policy rule estimated for ML0. However, the true model of the economy is heteroskedastic, given by the transition equations (1) for g and π ; residuals (v) and stochastic structure (8) from ML3. This is simulated with the policy rule from DC0 to give the welfare loss shown as DC03 in table 5. The response to shocks to the output gap and the inflation rate is less aggressive than it would optimally be and the effect of temporary shocks is thus more persistent than in DC3. Table 5 shows that this gives a welfare loss of 8.52, 0.4 higher than for DC3. This shows that in the benchmark case, the cost of ‘policy mis-specification’ (DC03-DC3) is numerically greater than the benefits suggested by the standard model of moving from the historical path to

an optimal policy (Sample-DC0).

[INSERT TABLE 5 HERE]

How sensitive are these findings to the choice of policy preference parameters? We check this by repeating the simulations using the Rudebusch and Svensson parameter sets used earlier in table 4. The resulting welfare losses are shown in table 6. For the homoskedastic model DC0, the losses over the different welfare assumptions are broadly similar to those of Rudebusch and Svensson. However, the table shows that under all five specifications of the welfare weights in the objective function the gain from ‘heteroskedastic optimization’ (Sample-DC3) is significantly higher than that from ‘homoskedastic optimization’ (Sample-DC0). The final two rows of the table indicate that the cost of ‘policy mis-specification’ is remarkably constant over these parameter sets, except for the fourth specification which attaches a high weight to output stabilization.

[INSERT TABLE 6 HERE]

6 Conclusion

Our empirical results show that the Okun-Friedman effect is remarkably significant in US macroeconomic data. The variance of the inflation rate exhibits linear dependence as in the Cox, Ingersoll, and Ross (1985) specification while output and interest variances are dominated by quadratic dependence. Either way, these macro and policy errors both seem to fluctuate with the rate of inflation. This phenomenon helps explain the Great Moderation or low volatility era that characterized the years between the recessions of the early 1990s and the late 2000s. The empirical model regards the large output and interest shocks in the final quarter of 2008 following the collapse of Lehman Brothers as outliers, since subsequent surprises have returned to

the low volatility implied by the low rate of inflation.

We show that theoretically, linear dependence reduces the optimal steady state inflation and interest rates relative to both the data and the homoscedastic optimization model. The optimal policy calculations reveal that the average rate inflation would have been about one percentage point lower. Quadratic dependence makes policy react more aggressively to deviations of inflation from target than in the homoscedastic and linear dependence models. Reflecting this, the model simulations show that the authorities would have reacted more aggressively to the upward drift in inflation in the 1960s and the oil price shock of 1973-74, particularly in the quadratic variance model, leaving them in a much better position to deal with subsequent shocks.

We use these models to analyze the welfare losses implied by historical and optimal interest rate behavior. Following Rudebusch and Svensson (1998) we use various welfare parameter sets, starting with the benchmark set that give inflation, output and interest rate smoothing terms equal weight. The gains from the homoskedastic optimization approach are relatively small in this case. However the welfare gains increase nearly fourfold once the heteroscedastic nature of the system is recognized, allowing the volatility of the system to be reduced by lowering the steady state inflation rate. Parameter sets that reduce the emphasis given to output and interest rate smoothing naturally make the authorities react much more aggressively towards inflation shocks and imply much larger welfare gains in the conventional homoscedastic optimization model. Even so, the conventional model still seriously underestimates the potential gains from policy optimization.

A Appendix 1: Admissibility

A stochastic volatility specification is said to be ‘admissible’ if it ensures that the variance structure remains non-negative definite. This is guaranteed in a mean-reverting

continuous-time model when there is a single square root volatility factor, essentially because the volatility goes to zero gradually as the interest rate or other variable driving the volatility goes to zero, allowing the system to mean revert.¹⁴ Admissibility is more problematic in discrete time square root volatility models because these use a Gaussian approximation (due originally to Sun (1992)) allowing the driving variable to turn negative during a discrete time interval. However this is not likely to be a problem in our linear-quadratic specification. In this case, we simply need to ensure that the eigenvalues of the variance structure remain non negative for all values of the driving variable (in this paper, π^a). These are given by (9) requiring $\delta_{k,t} \geq 0, k = g, \pi, r$:

$$\delta_{k,t} = \delta_{k,0} + \delta_{k,1}\pi^a + \delta_{k,2}(\pi^a)^2 \geq 0$$

This is ensured provided that $4\delta_{k,0}\delta_{k,2} \geq \delta_{k,1}^2$ so that the roots of the associated quadratic equation are complex. Empirically, the linear term $\delta_{k,1}$ is typically small compared to the constant and quadratic terms so that this is not an issue. Our *Matlab* code automatically checks that this restriction is satisfied.

B Appendix 2: The likelihood function

This appendix derives the likelihood function and describes the numerical optimization procedure. Write (6) and (8) as:

$$\mathbf{z}_t = \boldsymbol{\phi}\mathbf{X}_{t-1} + \theta r_{t-1} + \mathbf{u}_t$$

¹⁴It is however a problem in multi-factor correlated square root (CSR) volatility models. Dai and Singleton (2000) show that these are admissible only if the factors are negatively correlated, while empirical evidence is that they are positively correlated.

$$\mathbf{u}_t = \mathbf{C}\mathbf{D}_t^{1/2}\mathbf{v}_t \quad (27)$$

$$\mathbf{v}_t \sim N(\mathbf{0}_4, \mathbf{I}_4)$$

where: $\mathbf{z}'_t = \{\mathbf{x}'_t, r_t\}$, $\mathbf{u}'_t = \{\mathbf{u}'_{\mathbf{x},t}, u_{r,t}\}$; $\mathbf{v}'_t = \{\mathbf{v}'_{\mathbf{x},t}, v_{r,t}\}$ and:

$$\mathbf{C} = \begin{bmatrix} \mathbf{G} & \mathbf{0}_3 \\ \mathbf{g}' & 1 \end{bmatrix}; \quad \mathbf{D}_t = \begin{bmatrix} \mathbf{D}_{\mathbf{x},t} & \mathbf{0}_3 \\ \mathbf{0}'_3 & \delta_{r,t} \end{bmatrix}; \quad \mathbf{D}_i = \begin{bmatrix} \mathbf{D}_{\mathbf{x},i} & \mathbf{0}_3 \\ \mathbf{0}'_3 & \delta_{r,i} \end{bmatrix}, i = 0, 1, 2. \quad (28)$$

Then using (27):

$$\begin{aligned} \mathbf{v}_t &= \mathbf{C}^{-1}\mathbf{D}_t^{-1/2}\mathbf{u}_t \\ &= \mathbf{C}^{-1}\mathbf{D}_t^{-1/2}[\mathbf{z}_t - \phi\mathbf{X}_{t-1} - \theta r_{t-1}] \\ \mathbf{v}_t &\sim N(\mathbf{0}_4, \mathbf{I}_4) \end{aligned}$$

Thus the loglikelihood for period t can be written as:

$$L_t = -\frac{4}{2}\ln(2\pi) - \frac{1}{2}\ln(|\mathbf{D}_t|) - \frac{1}{2}\mathbf{v}'_t\mathbf{D}_t^{-1}\mathbf{v}_t \quad (29)$$

Summing this over T periods gives the loglikelihood for the estimation period:

$$L = -2T\ln(2\pi) - \frac{1}{2}\sum_{t=1}^T \ln(|\mathbf{D}_t|) - \frac{1}{2}\sum_{\tau=1}^k \mathbf{v}'_\tau\mathbf{D}_\tau^{-1}\mathbf{v}_\tau.$$

This likelihood function was maximized using the *FindMinimum* numerical optimization package on *Matlab*.

References

AIT-SAHALIA, Y. (1996): "Testing Continuous Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–42.

- ANG, A., AND M. PIAZZESI (2003): “A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables,” *Journal of Monetary Economics*, 43, 75–87.
- BALL, L. (1992): “Why Does Inflation Raise Inflation Uncertainty,” *Journal of Monetary Economics*, 29, 371–388.
- BEKAERT, G., S. CHO, AND A. MORENO (2010): “New Keynesian Macroeconomics and the Term Structure,” *Journal of Money, Credit and Banking*, 42(1), 33–62.
- BRANDT, M., AND K. WANG (2003): “Time Varying Risk Aversion and Unexpected Inflation,” *Journal of Monetary Economics*, 50, 1457–1498.
- BREUSCH, T.S.; PAGAN, A. (1979): “Simple test for heteroscedasticity and random coefficient variation,” *Econometrica*, 47, 1287–1294.
- CAPORALE, T., AND B. MCKIERNAN (1997): “High and Variable Inflation: Further Evidence on the Friedman Hypothesis,” *Economics Letters*, 54, 65–68.
- CHEN, K., F. KAROLYI, F. LONGSTAFF, AND A. SANDERS (1992): “An Empirical Comparison of Alternative Models of the Short Term Interest Rate,” *The Journal of Finance*, 47, 1209–1227.
- CHEN, R., AND L. SCOTT (1993): “Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates,” *Journal of Fixed Income*, 3, 14–31.
- CLARIDA, R., J. GALL, AND M. GERTLER (2000): “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory,” *Quarterly Journal of Economics*, 115, 147–180.
- COURTADON, G. (1982): “The Pricing of Options on Default-Free Bonds,” *Journal of Banking and Finance*, 17, 75–100.
- COX, J., J. INGERSOLL, AND S. ROSS (1985): “A Theory of the Term Structure of Interest Rates,” *Econometrica*, 52, 385–407.

- DAI, Q., AND K. SINGLETON (2000): “Specification Analysis of Affine Term Structure Models,” *Journal of Finance*, 55, 415–441.
- DENNIS, R. (2006): “The Policy Preferences of the US Federal Reserve,” *Journal of Applied Econometrics*, 21(1), 55–77.
- DEWACHTER, H., M. LYRIO, AND K. MAES (2006): “A Joint Model for the Term Structure of Interest Rates and the Macroeconomy,” *Journal of Applied Econometrics*, 21, 439–462.
- DOTHAN, U. (1978): “On the Term Structure of Interest Rates,” *Journal of Financial Economics*, 6, 59–69.
- FAVERO, C., AND R. ROVELLI (2003): “Macroeconomic Stability and the Preferences of the Fed. A formal Analysis, 1961-98,” *Journal of Money Credit and Banking*, 35, 545–556.
- FRIEDMAN, M. (1977): “Nobel Lecture: Inflation and Unemployment,” *Journal of Political Economy*, 85, 451–472.
- FUHRER, J. (1997): “Inflation/Output Variance Trade-offs and Optimal Monetary Policy,” *Journal of Money, Credit and Banking*, 29, 214–234.
- FUHRER, J., AND J. MOORE (1995): “Monetary policy trade-offs and the correlation between nominal interest rates and real output,” *American Economic Review*, 85, 219–239.
- HOLLAND, A. S. (1995): “Inflation and Uncertainty: Tests for Temporal Ordering,” *Journal of Money, Credit and Banking*, 27, 827–837.
- KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT, AND Y. SHIN (2001): “Testing the Null Hypothesis of Stationary Against the Alternative of A Unit Root,” *Journal of Econometrics*, 54, 159–78.
- LETTAU, M., S. LUDVIGSON, AND J. WACHTER (2004): “The Declining Equity Premium: What Role Does Macroeconomic Risk Play,” *NBER and New York*

University Working Paper Series.

- LEVIN, A., W. V. W. J. (1998): “The robustness of simple policy rules under model uncertainty,” in *Monetary Policy Rules*, ed. by J. Taylor, p. 263.
- LUNGQVIST, L., AND J. SARGENT, T. (2004): *Recursive Macroeconomic Theory*. The MIT Press, Cambridge, Massachusetts.
- NOWMAN, B. (1999): “Gaussian Estimation of Single Factor Continuous Time Models of the Term Structure of Interest Rates,” *The Journal of Finance*, 52, 1695–1706.
- OKUN, A. (1971): “The Mirage of Steady Inflation,” *Brookings Papers on Economic Activity*, 2, 485–498.
- ORPHANIDES, A. (2003): “The quest for prosperity without inflation,” *Journal of Monetary Economics*, 50(3), 633 – 663, Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.
- OZLALE, U. (2003): “Price Stability vs. Output Stability: Tales from Three Federal Reserve Administrations,” *Journal of Economic Dynamics and Control*, 27, 1595–1610.
- POLITO, V., AND M. WICKENS (2010): “Optimal Monetary Policy Using an Unrestricted VAR,” *Journal of Applied Econometrics*, *Forthcoming*.
- ROTEMBERG, J., J., AND M. WOODFORD (1997): “An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy,” *NBER Macroeconomics Annual*, 12, 297–346.
- RUDEBUSCH, G. (2002): “Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty,” *The Economic Journal*, 112, 402–32.
- RUDEBUSCH, G., AND L. SVENSSON (1998): “Monetary Policy Rules,” *NBER Working Papers 6512*.
- RUDEBUSCH, G. D., AND T. WU (2008): “A Macro-Finance Model of the Term

- Structure, Monetary Policy and the Economy,” *The Economic Journal*, 118(530), 906–926.
- SACK, B. (2000): “Does the Fed Act Gradually? A VAR Analysis,” *Journal of Monetary Economics*, 46, 229–256.
- SALEMI, M. (2006): “Econometric Policy Evaluation and Inverse Control,” *Journal of Money, Credit and Banking*, 38(7), 1737–1764.
- SMETS, F. (1999): “Output Gap Uncertainty: Does It Matter for the Taylor Rule,” in *Monetary Policy under Uncertainty*, ed. by B. Hunt, and A. Orr, pp. 10–29, Wellington. Reserve Bank of New Zealand.
- SPENCER, P. D. (2008): “Stochastic Volatility in a Macro-Finance Model of the US Term Structure of Interest Rates 1961-2004,” *Journal of Money, Credit, and Banking*, 40, 1177–1215.
- STANTON, R. (1997): “A Non-Parametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk,” *The Journal of Finance*, 52, 1973–2002.
- STOCK, J., AND M. WATSON (2001): “Vector Autoregressions,” *Journal of Economic Perspectives*, 15(4), 101–115.
- SVENSSON, L. (1999): “Inflation Targetting as a Monetary Policy Rule,” *Journal of Monetary Economics*, 43, 607–54.
- SVENSSON, L. E. (1997): “Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets,” *European Economic Review*, 41, 1111–1146.
- TAYLOR, J. (1979): “Estimation and Control of a Macroeconomic Model with Rational Expectations,” *Econometrica*, 47, 1267–1286.
- (1993): “Discretion versus Policy Rules in Practice,” *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.
- WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

C Tables and Figures

Table 1: Data Summary Statistics: 1961Q4-2009Q4

	Mean	Standard deviation	Skewness	Excess kurtosis	First order Autocorrel.	KPSS	ADF
g	-0.03	0.55	-0.47	3.72	0.914	0.174	-4.084
π	0.91	0.55	1.03	3.34	0.981	0.491	-1.934
r	1.37	0.72	1.04	4.95	0.934	0.431	-1.889

Note: Inflation (π) and Fed funds rate (r) are from Datastream. Output gap (g) is from OECD. Mean denotes sample arithmetic mean expressed as percentage per annum. The standard deviation is in percentage (per annum). Skewness and excess kurtosis are the standard measures of the third and fourth moments. Autocorrel. is the first order quarterly autocorrelation coefficient. KPSS is the Kwiatkowski, Phillips, Schmidt, and Shin (2001) statistic for the null hypothesis of level stationarity: the 10% and 5% significance levels are 0.347 and 0.463 respectively. ADF is the Adjusted Dickey-Fuller statistic for the null hypothesis of non-stationarity: the 10% and 5% significance levels are 2.575 and 2.877 respectively.

Table 2: Loglikelihood ratio tests

Model	ML0	ML1	ML2	ML3
Loglikelihood value	(-) 155.96	(-) 133.10	(-) 106.41	(-) 62.60
Number of Parameters	23	26	26	29
Value of likelihood ratio test against M3	186.7 $\chi^2(6)$	141.0 $\chi^2(3)$	87.6 $\chi^2(3)$	
Critical values χ^2 : 5%	11.07	7.81	7.81	
Critical values χ^2 : 1%	15.09	11.35	11.35	

Note: Model ML0 assumes a homoskedastic error structure. ML1 assumes that the error variances are linear in the (lagged) annual rate of inflation while ML2 assumes that the variances are quadratic in this rate. The encompassing model ML3 includes both linear and quadratic effects. In the MLVAR approach, the interest rate equation (2) is estimated alongside the structural equations (1) using maximum likelihood.

Table 3: Parameter estimates for model ML3

Par.	Estimate	t-value	Par.	Estimate	t-value	Par.	Estimate	t-value
g_1	0.008	0.22	θ_r	0.267	4.54	$\delta_{g,0}$	0.625	11.84
g_2	0.043	1.36	c_1	0.139	2.27	$\delta_{\pi,0}$	0.779	14.06
g_3	0.017	0.51	c_2	1.154	16.30	$\delta_{r,0}$	-0.605	-13.42
a_1	1.250	18.32	c_3	-0.232	-3.91	$\delta_{g,1}$	0.325	5.71
a_2	-0.328	-4.74	c_4	-0.050	-0.74	$\delta_{\pi,1}$	-0.467	-10.46
a_3	-0.015	-0.49	c_5	-0.288	-2.67	$\delta_{r,1}$	0.490	10.72
b_1	0.565	7.72	c_6	0.033	0.49	$\delta_{g,2}$	-0.211	-4.60
b_2	0.189	2.28	c_7	0.176	1.65	$\delta_{\pi,2}$	-0.114	-1.38
b_3	0.143	1.74	c_8	0.000	0.00	$\delta_{r,2}$	0.255	6.67
b_4	0.105	3.91	c_9	-0.015	-2.19			

Note: Model ML3 assumes that the variance structure depends upon the lagged annual rate of inflation and its square. In the ML model, the interest rate equation (2) is estimated alongside the structural equations (1) using maximum likelihood.

Table 4: Long run responses of estimated and optimal policy rules

	$\mu=1,v=1$	$\mu=0.2,v=0.5$	$\mu=1,v=0.5$	$\mu=5,v=0.5$	$\mu=1,v=0.1$
	ML0		DC0		
$\bar{\kappa}$	0.00	0.00	0.00	0.00	0.00
$\bar{\phi}_g$	1.16	1.88	2.02	2.17	3.00
$\bar{\phi}_\pi$	1.60	2.86	4.12	3.11	3.77
$\bar{\pi}$	0	0	0	0	0
	ML1		DC1		
$\bar{\kappa}$	0.00	1.02	1.44	1.09	0.59
$\bar{\phi}_g$	1.17	1.78	1.87	2.06	2.61
$\bar{\phi}_\pi$	1.61	2.87	4.38	3.10	1.89
$\bar{\pi}$	0	-0.92	-0.62	-0.87	-1.53
	ML3		DC3		
$\bar{\kappa}$	0.00	4.64	4.75	4.92	5.26
$\bar{\phi}_g$	0.32	3.49	3.69	4.04	5.12
$\bar{\phi}_\pi$	1.15	5.85	7.80	6.43	5.10
$\bar{\pi}$	0	-1.37	-1.03	-1.28	-1.70

Note: The long run interest rate rule takes the form $r = \bar{\kappa} + \bar{\phi}_{r,g}g + \bar{\phi}_{r,\pi}\pi$, where the long run coefficients are computed from the long run solution of either the interest rate equation estimated in (2) or the optimal policy rule (18). The intercept is always zero in the homoskedastic models ML0 and DC0 since the data is de-measured. Heteroskedasticity has the effect of inducing a positive intercept, which reduces the steady state by $\bar{\pi} = \bar{\kappa}/(1 - \bar{\phi}_\pi)$.

Table 5: Welfare losses implied by historic data and optimal rules, benchmark case

Model	std.(g)	std.(π^a)	std.(Δr)	Loss
Sample	2.24	2.11	0.97	10.4
Optimal monetary policy				
DC0	2.46	1.29	1.44	9.83
DC1	2.44	1.05	1.25	8.59
DC03	2.15	1.51	1.27	8.52
DC3	2.08	1.06	1.62	8.08

Note: The policy rules in this table use the benchmark parameters, which give equal weight to inflation, output gap and funds rate stabilization: $\mu = v = 1$. The discount rate is set to $\beta = 1$. This means that the welfare loss (Loss) is the sum of the squared standard deviations.

Table 6: Welfare losses implied by historic data and optimal rules with different welfare parameters

Model	Welfare loss				
	$\mu=v=1$	$\mu=0.2; v=0.5$	$\mu=1; v=0.5$	$\mu=5; v=0.5$	$\mu=1; v=0.1$
Sample	10.40	5.93	9.93	29.92	9.56
DC0	9.83	4.75	9.15	25.55	7.52
DC03	8.52	4.18	8.03	26.34	6.98
DC3	8.08	3.82	7.56	23.91	6.61

Note: The policy rules use the preference parameters specified in Rudebusch and Svensson (1998) with $\beta = 1$ for the discount rate.

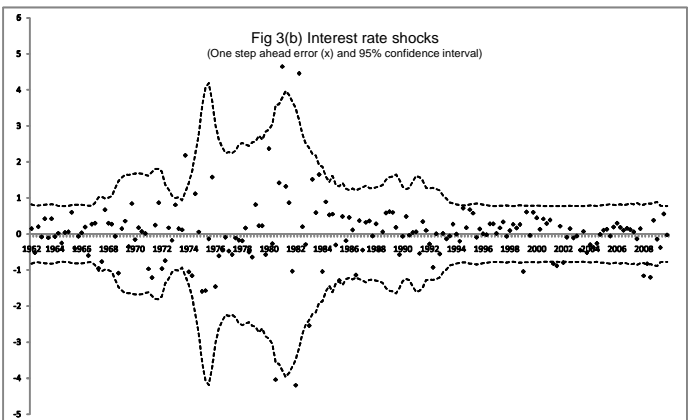
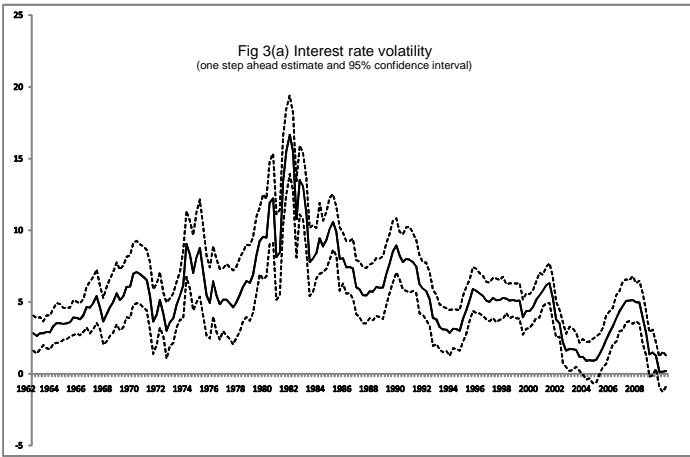
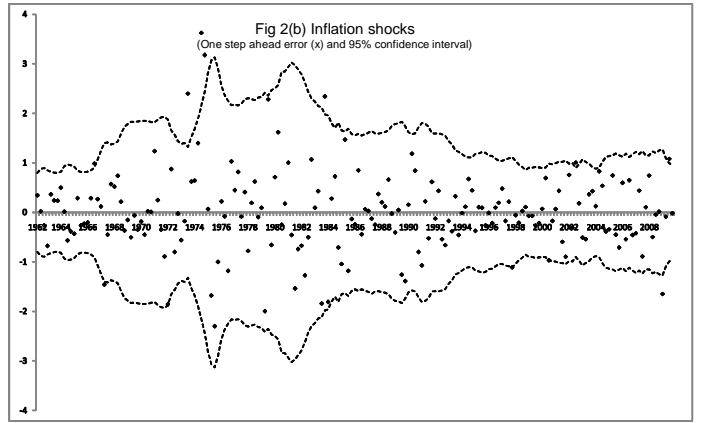
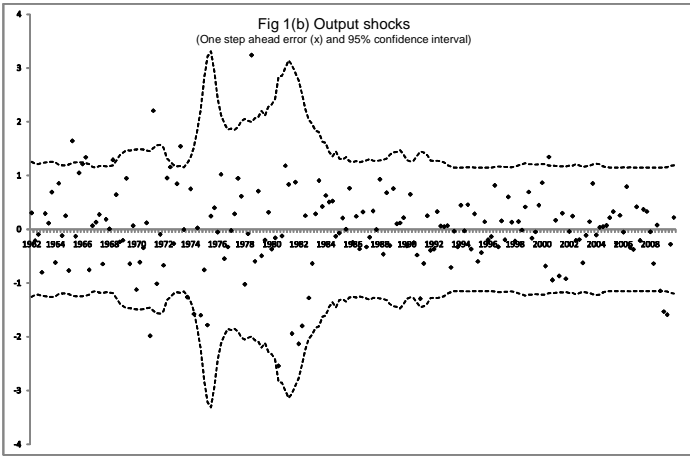
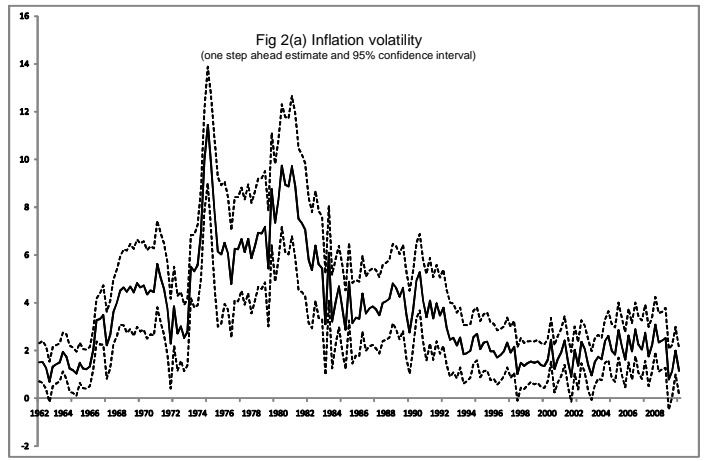
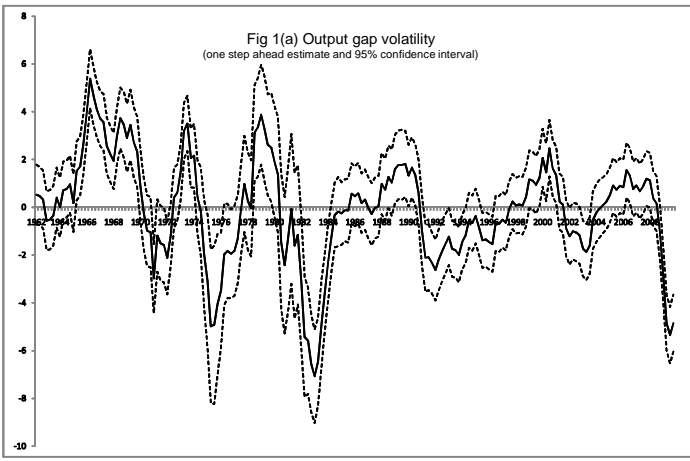


Fig 4 Impulse response functions under homoskedastic (ML0), linear (ML1) and linear-quadratic (ML3) dependence of errors variance

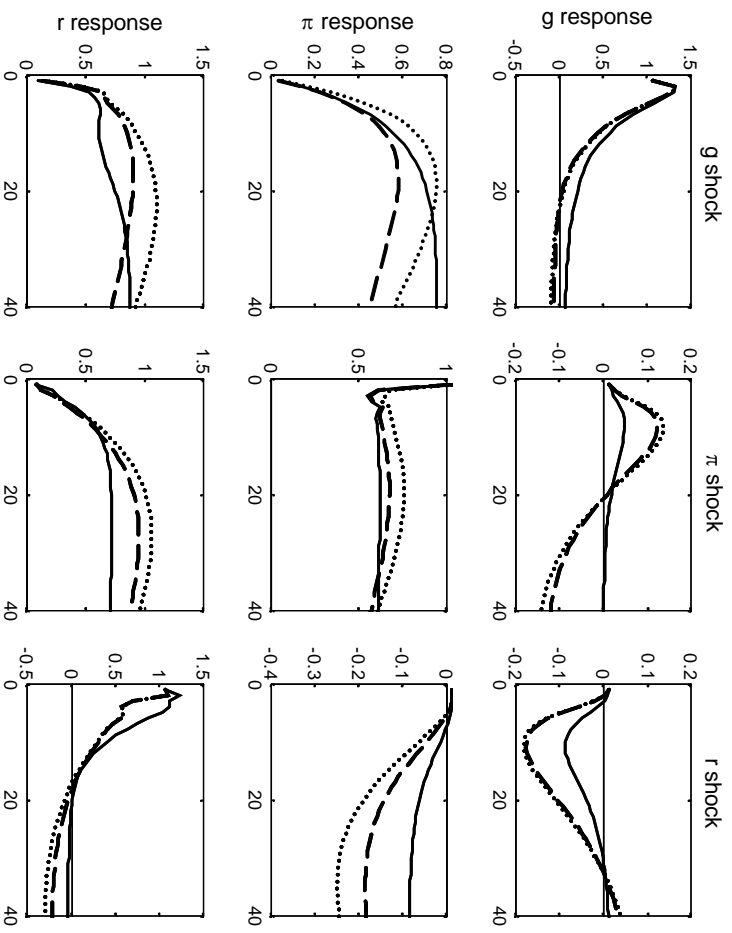


Fig 5 Impulse response functions from DC models under homoskedastic (DC0), linear (DC1) and linear-quadratic (DC3) dependence of errors variance

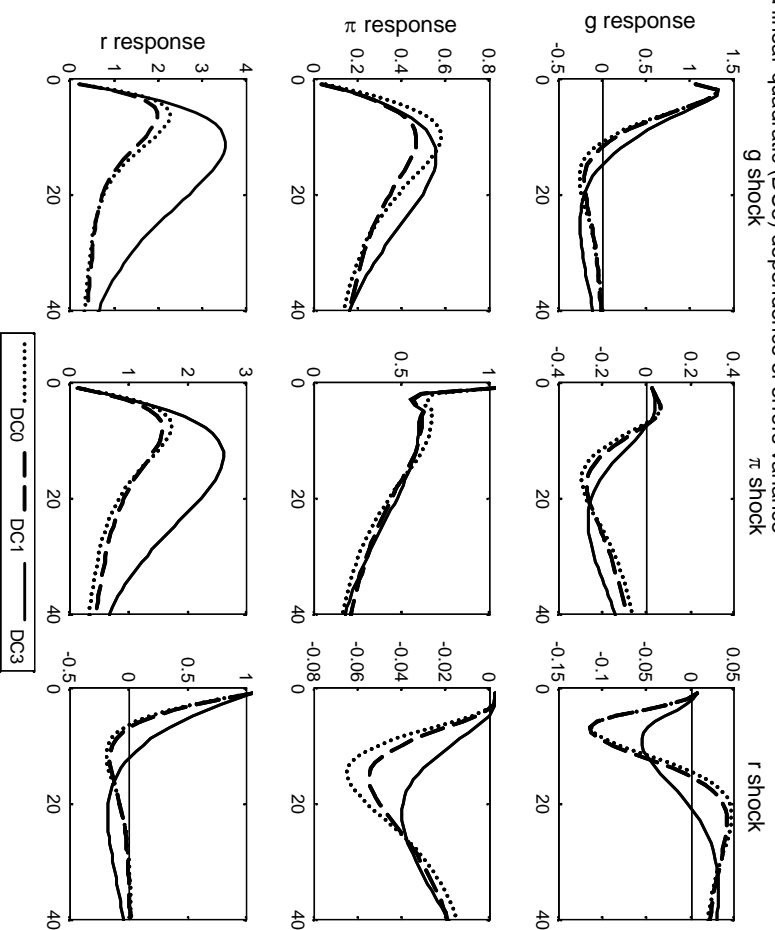


Fig 8 Difference between the output gap simulated from the DC models and the historical output gap

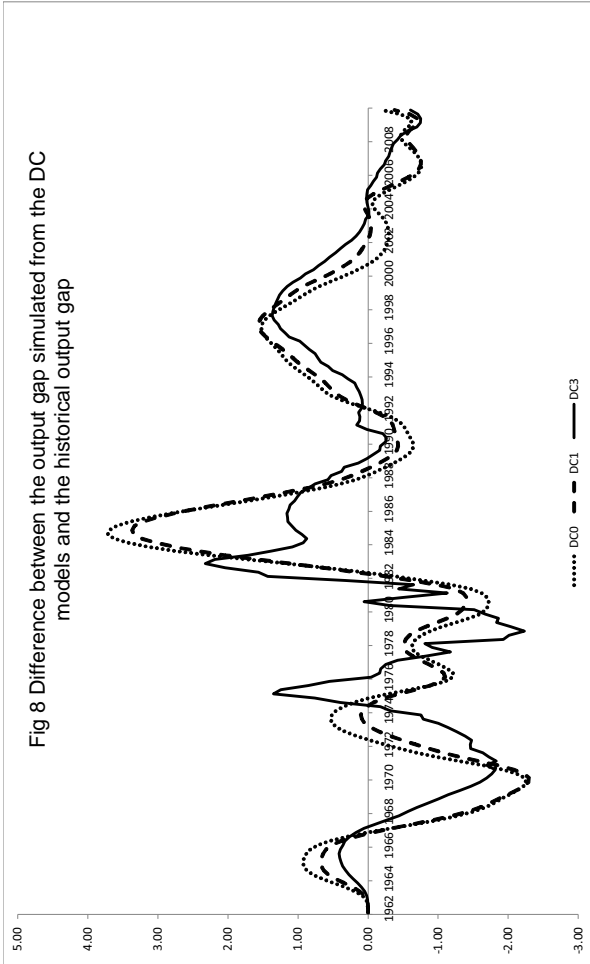


Fig 6 Difference between the Federal funds rate simulated from the DC models and the historical funds rate

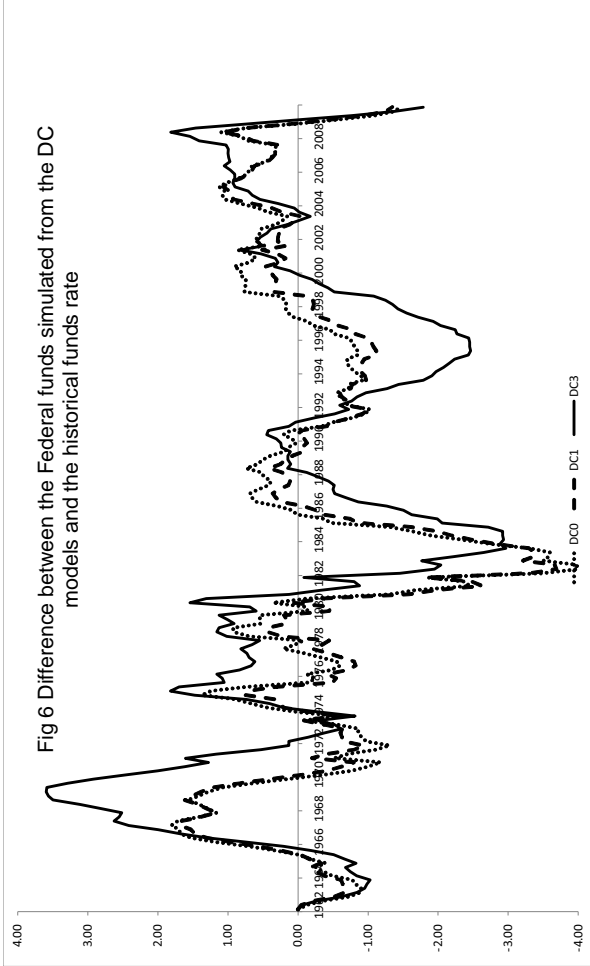


Fig 7 Difference between the inflation rate simulated from the DC models and the historical inflation rate

