CARDIFF BUSINESS SCHOOL
WORKING PAPER SERIES

Cardiff Economics Working Papers

Vito Polito and Peter Spencer

UK Macroeconomic Volatility and the Welfare Costs of Inflation

E2011/23

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

ISSN 1749-6101
October 2011

UK Macroeconomic Volatility and the Welfare Costs of Inflation

Vito Polito*and Peter Spencer‡

Abstract

This paper explores the implications of time varying volatility for optimal monetary policy and the measurement of welfare costs. We show how macroeconomic models with linear and quadratic state dependence in their variance structure can be used for the analysis of optimal policy within the framework of an optimal linear regulator problem. We use this framework to study optimal monetary policy under inflation conditional volatility and find that the quadratic component of the variance 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 1979-2010 underline the statistical significance of inflation-dependent UK macroeconomic volatility. Analysis of the welfare losses associated with inflation and macroeconomic volatility shows that the conventional homoskedastic model seriously underestimates both the welfare costs of inflation and the potential gains from policy optimization.

Keywords: Monetary policy, Macroeconomic volatility, Optimal control, Welfare costs of inflation.

JEL classification: C32, C61, E52

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

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

‡We are grateful to Richard Dennis for sharing his Gauss code with us and to Peter Burridge, Fabio Canova, Christos Ioannidis, Gary Koop, Jeff Sheen, James Morley, Peter Smith, Ian Tonks, Mike Wickens and other participants at seminar presentations at the Reserve Bank of Australia, Macquarie University, the University of Technology Sidney, University of New South Wales and the Universities of Bath, Strathclyde and York for insightful comments.

1 Introduction

Since the Second World War, the world's developed economies have experienced marked fluctuations in macroeconomic volatility. The early post-war years were relatively stable but volatility rose from decade to decade, reaching a peak in the late 1970s and early 1980s. It then began to subside in the mid 1980s, remaining remarkably low until the 'credit crunch' that began in August 2007. Stock and Watson (2002) refer to the decline in US macroeconomic volatility started from the mid 1980s as the 'Great Moderation'. Similar declines in volatility occurred over the same period of time in almost all G7 countries (Bernanke (2004)), and have been particularly marked in the UK, where Mervyn King dubbed the period 1994-2003 as the 'NICE' or 'Non-Inflationary Consistent Expansion' decade (King (2003)). However, economic volatility has risen markedly since the credit crunch of 2007, provoking the sharpest recession seen in the G7 countries since the Second World War.

The recognition of time-varying volatility in macroeconomic data has lead economists to abandon the traditional constant variance (homoskedasticity) assumption and to develop macroeconomic models in which volatility can fluctuate. This literature has demonstrated that modelling movements in volatility increases the accuracy of parameter estimates and macroeconomic forecasts. It also provides a framework for analyzing changes in volatility, asking whether they are for instance driven by changes in the systematic component of macroeconomic policy or the magnitude of shocks. Recent examples of this burgeoning literature include Primiceri (2005), Sims and Zha (2006), Canova, Gambetti, and Pappa (2008), Justiniano and Primiceri (2008) and Benati and Surico (2009) for the US. Benati (2004) and Bianchi, Mumtaz, and Surico (2009) look at the evidence for the UK and Mumtaz and Surico

(2011) for a sample of industrialized countries.

This paper attempts to explore theoretically and measure empirically the implications of time-varying macroeconomic volatility for macroeconomic policy and welfare. To the best of our knowledge, it is the first time that this type of exercise has been conducted. We focus on state dependent volatility, a class of stochastic volatility models that relates this to the (lagged) value of one or more state variables. This type of model is extensively used in the literature on the term structure of interest rates, where the effect of volatility upon risk premia is a vital consideration. We start with a general framework that allows the variance of the shocks to depend in both a linear and a quadratic way on state variables reflecting inflation and the state of the business cycle. The linear specification is the analogue of the Cox, Ingersoll, and Ross (1985) 'square root' volatility specification of the term structure literature, while the quadratic specification is the analogue of Dothan (1978) and Courtadon (1982). We combine these two effects and analyze the implications for welfare and optimal policy assuming that policy makers care about squared deviations in goal variables such as inflation and the output gap around their target values, as in the canonical homoskedastic control problem (Lungqvist and Sargent (2004)).

Two general results emerge from this analysis. First, unlike the standard homoskedastic control problem which is certainty equivalent (meaning that the optimal rule is the same as it would be in the absence of uncertainty) we show that the coefficients of the optimal policy rule are affected by the stochastic structure when this is state dependent. Second, we show how the optimal linear regulator problem for a state dependent model can be re-parameterized so that standard control techniques can be employed to quantify optimal policy rules and welfare losses. The first result highlights the importance of allowing for time-varying volatility in the design of optimal macroeconomic policy, while the second shows how this can be performed using

the algorithms and insights provided by existing methodologies.

We then use this framework to examine how the design of optimal monetary policy and welfare analysis change when macroeconomic volatility depends upon the rate of inflation.[1] We show that linear-quadratic inflation conditional volatility adds to the welfare cost of inflation, thus altering the conduct of optimal monetary policy. Specifically, we show that this effect makes the optimal monetary policy response to a rise in inflation more aggressive than that implied by a homoskedastic model and leads to a lower average inflation rate. A number of papers, Holland (1995), Fountas, Karanasos, and Kim (2002) and Caporale and Kontonikas (2009) among others, have argued that, if policy makers care about economic stability then an increase in inflation should lead to a monetary tightening response to constrain the increase in macroeconomic volatility. Our theoretical results formalize this conjecture.

We have looked at the significance of conditional heteroskedasticity using several econometric methodologies for several different countries. In this paper we present the results for the UK using a version Rudebusch and Svensson (1999) model which we modify to allow the variance structure of the shocks to exhibit linear-quadratic inflation dependence. We focus on the UK country because we find that inflation conditional volatility is empirically much stronger than for other countries that we have looked at. The Rudebusch and Svensson (1999) model is chosen because it provides a simple and effective description of the dynamic evolution of the key variables describing the macroeconomy and as such is a workhorse for the analysis of monetary

[1] The relationship between the inflation rate and macroeconomic volatility was first documented by Okun (1971) and the implications for welfare analysis were first noted by Friedman in his Nobel lecture, Friedman (1977). Early theoretical work suggesting that the inflation rate affects macroeconomic volatility includes Ball (1992) and Ungar and Zilberfarb (1993). Empirical evidence in support of this relationship was initially provided by Ball and Cecchetti (1990), Brunner and Hess (1993), Holland (1995) and Fountas, Karanasos, and Kim (2002) for the US. Evidence for other industrialized countries is found in Fountas and Karanasos (2007) and for the UK in Conrad, Karanasos, and Zeng (2010). The empirical term structure literature also suggests that the volatility of short term interest rates is state-dependent; see Chen, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1996), Stanton (1997) and, for the UK, Nowman (1999). Reflecting this literature, we find that the inflation rate has a very significant effect on UK macroeconomic volatility.

policy. See for example, Ozlale (2003), Favero and Rovelli (2003), Dennis (2006) and Cogley et al (2011).

The maximum likelihood estimates of the model based on UK quarterly time series for the output gap; inflation and nominal interest rate, show that linear-quadratic dependence provides a much better explanation of the UK data over the period 1979-2010 than the assumption of constant variance or linear dependence. This result is consistent with the finding of Sims and Zha (2006) that time-varying volatility models typically outperform homoscedastic models when estimated over long periods of time. In particular, the empirical model captures the high level of volatility seen in the UK until the mid 1980s and the subsequent decline, consistent with the time-varying macroeconomic volatility literature.

We then ask how recognizing the dependence of volatility upon inflation would have influenced the design of an optimal monetary policy rule derived from the minimization of a quadratic inflation targeting loss function consistent with Rotemberg and Woodford (1997) and Woodford (2003). As in Rudebusch and Svensson (1999), Sack (2000) and Woodford (2003), we replace the estimated interest rate equation (or Taylor rule) by the optimal policy rule and study the implied dynamics of the model under optimal policy. The optimal policy calculations suggest that mis-specification of the variance structure can lead researchers to seriously understate both the welfare cost of inflation and the potential gains from optimization.

The rest of the paper proceeds as follows. Section 2, supported by Appendices 1 and 2, presents a general analysis of state dependent volatility and its implications for the optimal conduct of macroeconomic policy. Section 3, supported by Appendix 3, applies this to the study of optimal monetary policy under inflation-conditional volatility and outlines the version of the macroeconomic model of the economy used here to allow for linear-quadratic inflation dependence in the variance structure. Sec-

tion 4 presents the maximum likelihood estimation results, while section 5 quantifies the coefficients of the optimal policy rule and welfare losses. Section 6 concludes with a brief summary of the empirical findings and their relevance for UK monetary policy, together with an agenda for future research.

2 Optimal macroeconomic control with state dependent volatility

In this section we explore the general implications of state dependent volatility for the optimal design of macroeconomic policy. We employ a canonical specification, which we then specialize in subsequent sections.

2.1 The dynamic structure

We start with a general linear dynamic model. This is expressed in state space form as:

$$\mathbf{X}_{t+1} = \mathbf{\Phi}\mathbf{X}_t + \mathbf{\Theta} i_t + \mathbf{U}_{t+1} \tag{1}$$

$$\mathbf{U}_{t+1} \sim N(0, \mathbf{\Sigma}_{t+1}), \tag{2}$$

where $\mathbf{X}_t$ is an $n \times 1$ vector of state variables observed by the decision maker describing the position of the macroeconomy at any time t; i_t is a policy instrument available to the decision maker in period t; $\mathbf{\Phi}$ an n^2 matrix of coefficients; $\mathbf{\Theta}$ is an $n \times 1$ vector; and $\mathbf{U}_t$ is a an $n \times 1$ vector of Gaussian error terms, with $\mathbf{\Sigma}_{t+1}$ denoting a n^2 variance-covariance matrix, discussed further below.

Variables are expressed as deviations from sample mean, so there is no intercept constant vector in this system. Obviously, the state vector must include the variables targeted by the policy maker, notably the output gap ($g_t = \mathbf{s}_g \mathbf{X}_t$) and the annual

rate of inflation ($\pi_t^a = \mathbf{s}_{\pi^a}\mathbf{X}_t$), where $\mathbf{s}_z$ is a selection vector that picks any variable or linear combination of variables z_t (including g_t and π_t^a) from $\mathbf{X}_t$. If the decision maker is a central bank, the policy instrument i_t can be either the monetary base, the exchange rate or the policy interest rate.

A wide range of macro models can be written in the state space form (1). For example, this describes VAR models such as those used for the measurement of macroeconomic shocks by Bernanke and Blinder (1992) and Bernanke and Mihov (1998); for optimal control exercises by Sack (2000) and Polito and Wickens (2011). It also encompasses the Rudebusch-Svensson central bank model, which has been extensively employed for the analysis of US monetary policy (Ozlale (2003), Favero and Rovelli (2003) and Dennis (2006)) used in the next section. Since the state vector $\mathbf{X}_{t+1}$ can also include variables representing private sector expectations, (1) is consistent with the state space representation of linear rational expectations models, as in Blanchard and Kahn (1980), Soderlind (1999), Woodford (2003), Lungqvist and Sargent (2004) and Salemi (2006). Walsh (2010) explores the implications for optimal monetary policy when the parameters of (1) are derived from Gali and Gertler (2007) specification of the New Keynesian model.

2.2 The stochastic structure

The optimal control literature has hitherto assumed that volatility is constant over time. However, the term structure literature departs from this homoskedastic framework by assuming that the error structure exhibits linear-quadratic dependence. The workhorse is provided by the A_1 specification,[2] which assumes that there is a single

[2] The subscript indicates the number of variables or combinations driving volatity, so the homoskedastic model is denoted as A_0.

variable or combination of variables driving volatility:

$$\mathbf{U}_{t+1} = \mathbf{\Sigma}_{t+1}^{\frac{1}{2}} \mathbf{V}_{t+1}, \tag{3}$$

where $\mathbf{V}_{t+1}$ is a an $n \times 1$ vector of standard normally independently distributed error terms and

$$\begin{aligned} \mathbf{\Sigma}_{t+1} &= \mathbf{E}[\mathbf{U}_{t+1}\mathbf{U}'_{t+1}|\mathbf{X}_t] \\ &= \mathbf{\Sigma}_0 + \mathbf{\Sigma}_1 \mathbf{s}'_z \mathbf{X}_t + \mathbf{\Sigma}_2 \mathbf{X}'_t \mathbf{S}_z \mathbf{X}_t \\ &= \mathbf{\Sigma}_0 + \mathbf{\Sigma}_1 z_t + \mathbf{\Sigma}_2 z_t^2, \end{aligned} \tag{4}$$

where $\mathbf{\Sigma}_i$, $i = 0, 1, 2$ are n^2 matrices of coefficients.[3] This allows the variance of the stochastic shocks entering (1) and (2) to depend on a single linear combination of lagged dependent variables (z_t) and its square. We can write the squared value of a variable affecting the variance as $z_t^2 = \mathbf{X}'_t \mathbf{S}_z \mathbf{X}_t$ where $\mathbf{S}_z = \mathbf{s}_z \mathbf{s}'_z$. In this paper we assume that the policy instrument does not affect volatility directly since $\mathbf{X}_t$ does not include i_t. We use the selection vector $\mathbf{s}'_i$ to pick out i_{t-1} from $\mathbf{X}_t : i_{t-1} = \mathbf{s}'_i \mathbf{X}_t$ and $i_{t-1}^2 = \mathbf{X}'_t \mathbf{S}_i \mathbf{X}_t$, where $\mathbf{S}_i = \mathbf{s}_i \mathbf{s}'_i$.

Equation (4) encompasses a wide range of volatility models. The standard homoskedastic specification (A_0) is obtained when $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{0}$, so that $\mathbf{\Sigma}_t = \mathbf{\Sigma}_0$. The linear dependence specification, consistent with Cox, Ingersoll, and Ross (1985), is obtained when $\mathbf{\Sigma}_2 = \mathbf{0}$. A quadratic dependence specification that relates variances to the square of a state variable, as proposed by Dothan (1978) and Courtadon (1982), is obtained when $\mathbf{\Sigma}_1 = \mathbf{0}$.

Equation (4) shows how the responses of the state variables to shocks depend on

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

the specification of the variance structure. In a homoskedastic model, these responses do not depend upon the state variables and are entirely determined by the dynamic model (1). They are linear in the shocks, symmetric for positive and negative shocks for example. However, in a conditional volatility model the impulse responses also depend on the initial values of the variables driving volatility. Moreover, because changes in the values of the driving variables in one period affect the impact of subsequent shocks (modelled by the Gaussian vector $\mathbf{V}_{t+1}$ in (3)), the responses depend upon the size, sign and duration of the shocks being simulated. These non-linear effects are not apparent in conventional simulations of one period shocks but are evident in simulations of longer-lasting shocks (as section 4.3 will demonstrate).

For example, in the linear dependence model (with $\mathbf{\Sigma}_2 = \mathbf{0}$), the model responses are amplified by any series of shocks that increases the variable (or combination of variables) z_t driving volatility. A series of negative shocks has the opposite effect, depressing this variable, attenuating the effect of further negative shocks and making the responses asymmetric. As explained in appendix 1, this effect implies a lower bound for the variable driving volatility. The amplification effect is also a feature of the quadratic dependence specification (with $\mathbf{\Sigma}_1 = \mathbf{0}$). However, starting with an initial value of $z_t = 0$ keeps the responses to positive and negative shocks symmetric. The general model incorporates both amplification and attenuation effects but the asymmetries in numerical simulations are less marked than in the linear model and the driving variable is unbounded.

2.3 The conditional volatility control framework

Appendix 2 shows how state dependent volatility affects the determination of an optimal decision rule within the general framework of the stochastic linear regulator problem. The remainder of this section (and the appendix) can be skipped by readers

who are primarily interested in the inflation-conditional volatility specification, which is set out in the next section. However the general intuition is quite straightforward. When a decision maker has a concave utility or loss function (like a quadratic) defined over deviations of variables like inflation and output from bliss values, volatility in the goal variables reduces expected utility in the same way that volatility of asset prices reduces the utility of an investor with a concave utility function. In the standard stochastic optimal regulator problem volatility is constant and the decision rule exhibits 'certainty equivalence': it does not depend in any way upon the degree of volatility in the system. In that case, volatility just lowers welfare and there is nothing the decision maker can do about it. If however, macroeconomic volatility is state dependent and the decision maker can influence the state of the system, this should influence his behavior. The certainty equivalence principle does not hold, since the coefficients of the optimal policy rule as well as the minimum value of the loss depend upon the variance structure.

Appendix 2 formalizes this observation and analyses the general implications, assuming that the preferences of a decision maker are characterized by a canonical quadratic loss function and that the law of motion of the state variables is described by the linear-quadratic state space framework set out in equations (1) and (4). The appendix shows that because the conditional volatility terms in (4) are linear and quadratic they have an effect which is mathematically equivalent to the linear and quadratic terms describing the welfare loss. This isomorphism means that we can re-write any linear-quadratic state dependent volatility control problem as an equivalent homoskedastic problem by a suitable re-parameterization of the targets and welfare weights. Given a volatility process of the general form (4), we set the linear ($\mathbf{\Sigma}_1$) and quadratic ($\mathbf{\Sigma}_2$) components to zero and appropriately adjust the target values and welfare weights in the loss function. Specifically, appendix 2 shows that

the linear dependence component has an effect equivalent to changing the target values in the welfare function while the quadratic component is equivalent to a change in the welfare weights given to squared deviations from target. This means that researchers can draw upon standard optimal control algorithms and insights in solving heteroskedastic control problems and discussing the results. An illustration is provided by the model set out in the next section.

3 An inflation-conditional volatility model

The adjustments to the welfare parameters in the canonical model of the previous section depend upon the choice of the state variables conditioning volatility. These adjustments could make policy either more or less responsive to economic disturbances and little can be said without specifying the nature of this dependence. In this section we provide an example based upon the inflation-conditional volatility model discussed in the introduction and examine its broad qualitative implications for economic policy.

First we briefly describe the specification search that led us to employ this specification. We started by using the Breusch and Pagan (1979) test to confirm the significance of conditional heteroskedasticity in various UK macroeconomic data. The first stage of this test is to regress a variable such as the output gap, inflation or interest rate on its lagged values. The second stage takes the squared residuals from these regressions as a measure of volatility and regresses them on various lagged indicator variables (and their squares). We began by using the base rate and the 10 year Treasury yield as explanatory variables in these second stage regressions, as suggested by the term structure literature. These variables performed reasonably well, but we found that the annual Consumer Expenditure Deflator (CED) inflation rate gave a much better explanation, consistent with the hypothesis that macroeco-

nomic volatility is driven by the underlying inflation rate. The output gap was not significant either on its own or in combination with other variables. We thus adopted a single factor volatility structure, driven by the lagged annual (CED) inflation rate. Thus we specify: $z = \pi$ in equation (4).

In this section, we assume that the central bank uses the policy interest rate (i_t) to minimize an intertemporal loss function including as arguments the variation in the output gap; the variation in the annual inflation rate around its target and the change in the interest rate:

$$L_t = E_t \sum_{j=0}^{\infty} \beta^{t+j} \left[\lambda \left(\pi_{t+j}^{a} - \pi^{*} \right)^2 + \mu g_{t+j}^2 + \upsilon \left(\Delta i_{t+j} \right)^2 \right], \tag{5}$$

where E_t is the time $t-$conditional expectations operator; β is the discount factor, π_{t+j}^{a} is the annual inflation rate, $\Delta i_{t+j} = i_{t+j} - i_{t+j-1}$ is the change in the base rate; π^* is the inflation target; and the parameters $\lambda \geq 0$; $\mu \geq 0$ and $\upsilon \geq 0$ are weights given to inflation, output gap and instrument stabilization. Rotemberg and Woodford (1997) and Woodford (2003) show that this quadratic loss function provides a good approximation to the expected lifetime utility of a representative household derived from a fully micro-founded macroeconomic model of the economy, in which inflation brings efficiency costs by distorting relative prices.

The policy maker is assumed to choose the intertemporal sequence of policy instruments $\{i_{t+j}\}_{j=0}^{\infty}$ that minimizes this loss function given the model of the economy in (1)-(4) and the initial state vector $\mathbf{X}_t$. This minimization problem can be solved using standard dynamic programming techniques. Because the control problem is entirely linear-quadratic the solution or value function (which shows the minimum expected loss in any period t aggregating over current and future periods), can be written as a linear-quadratic function of the state variables $\mathbf{X}_t$ observed by the policy

maker at the beginning of that period:

$$J(\mathbf{X}_t) = \min_{\{i_{t+j}\}} L_t \tag{6}$$

$$= c - 2\mathbf{X}_t'\mathbf{p} + \mathbf{X}_t'\mathbf{P}\mathbf{X}_t, \tag{7}$$

where $\mathbf{p}$ is an $n \times 1$ vector of constant coefficients and $\mathbf{P}$ an n^2 positive semidefinite matrix of coefficients that depend upon the nature of the problem.

The optimal policy is then found from the solution of a recursive Bellman equation that is obtained by substituting (5) into (6); using $J(\mathbf{X}_{t+1})$ to represent the minimum expected value of future losses and then substituting (7):

$$J(\mathbf{X}_t) = \min_{i_t}[\lambda(\pi_t^a - \pi^*)^2 + \mu g_t^2 + \upsilon(\Delta i_t)^2 + \beta E_t\left(\mathbf{X}_{t+1}'\mathbf{P}\mathbf{X}_{t+1} - 2\mathbf{X}_{t+1}'\mathbf{p} + c\right)].$$

Evaluating expectations using the constraints (1) and (4) gives the conditional volatility problem:

$$J(\mathbf{X}_t) = \min_{i_t}\left\{\begin{array}{c}\lambda(\pi_t^a - \pi^*)^2 + \mu g_t^2 + \upsilon(\Delta i_t)^2 + \beta c + I_t \\ +\beta tr(\mathbf{P}\mathbf{\Sigma}_1)\pi_t^a + \beta tr(\mathbf{P}\mathbf{\Sigma}_2)(\pi_t^a)^2\end{array}\right\}, \tag{8}$$

where:

$$I_t = I(\mathbf{X}_t, i_t) = \beta[(\mathbf{\Theta} i_t + \mathbf{\Phi}\mathbf{X}_t)'\mathbf{P}(\mathbf{\Theta} i_t + \mathbf{\Phi}\mathbf{X}_t) - 2(\mathbf{\Theta} i_t + \mathbf{\Phi}\mathbf{X}_t)'\mathbf{p} + tr(\mathbf{P}\mathbf{\Sigma}_0)]. \tag{9}$$

The first line of (8) shows the Bellman form of the standard homoskedastic control problem, in which: $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{0}$. In this case, the target rate for the variables in $\mathbf{X}_t$ is normally assumed to be in line with the sample mean and $\pi^* = 0$ if the model is specified in terms of mean-differences. The certainty equivalence principle also

holds in this case. The last two terms of (8) are non-standard and capture the effect of the state dependent variance structure on the control problem. Since these are respectively linear and quadratic in π_t^a, they affect the coefficients of the optimal policy rule. However, we can consolidate these with the inflation term in the first period loss writing the Bellman equation in the canonical form by setting $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{0}$ and replacing the welfare parameters λ, π^* and c by $\tilde{\lambda}$, $\tilde{\pi}^*$ and $\tilde{c}$ in the first term of the loss function (8), where:

$$\tilde{\lambda} = [\lambda + \beta tr \mathbf{P}\mathbf{\Sigma}_2] \geq \lambda \tag{10}$$

$$\tilde{\pi}^* = [\lambda\pi^* - \beta tr(\mathbf{P}\mathbf{\Sigma_1})/2]/\tilde{\lambda} \tag{11}$$

$$\tilde{c} = c + \lambda[(\pi^*)^2 - (\tilde{\pi}^*)^2]/\beta. \tag{12}$$

This shows that the quadratic-dependent volatility term $\beta tr \mathbf{P}\mathbf{\Sigma}_2$ stemming from $\mathbf{\Sigma}_2$ in (4) has the effect of making policy more aggressive in the sense that it has exactly the same effect on the optimal policy responses as would an increase in the welfare weight λ due to the macroeconomic costs of inflationary price distortions in the standard problem (Rotemberg and Woodford (1997)). The linear dependence term $\mathbf{\Sigma}_1$ has the effect of reducing the effective target rate of inflation from π^* to: $\tilde{\pi}^* = [\lambda\pi^* - \beta tr(\mathbf{P}\mathbf{\Sigma_1})]/\tilde{\lambda}$, where $\beta tr(\mathbf{P}\mathbf{\Sigma_1}) \geq \mathbf{0}$. Provided that $\pi^* \geq 0$ then $\tilde{\pi}^* \leq \pi^*$ (since $\lambda \leq \tilde{\lambda}$). Thus we see that the effect of linear dependence is to reduce the effective target and hence the steady state rate of inflation in exactly the same way as a negative structural shift in π^* would. Finally, the intercept in the value function shifts from c to $\tilde{c}$, but this does not affect the decision rule.

Transforming (8) using this re-parameterization allows the optimization problem to be written in the form of a standard Bellman problem:

$$J(\mathbf{X}_t) = \min_{i_t}[\tilde{\lambda}(\pi_t^a - \tilde{\pi}^*)^2 + \mu g_t^2 + \upsilon \Delta i_t^2 + \beta \tilde{c} + I(\mathbf{X}_t, i_t)]. \tag{13}$$

The solution to (13) can then be obtained using the techniques developed for the standard homoskedastic problem (for a review, see Lungqvist and Sargent (2004)). If the parameters of (1) are independent of the policy rule, as for example in the model developed in the rest of this section, the optimal policy rule is obtained simply by (13) differentiating w.r.t. i_t and solving for the optimal policy rate. Using the notation described in section 2.1 to write Δi_t^2 as $(i_t - \mathbf{s}_i \mathbf{X}_t)^2$ in (13), this procedure gives:

$$i_t = \zeta + \boldsymbol{\xi} \mathbf{X}_t \tag{14}$$

$$\text{where:} \quad \zeta = \frac{\beta \boldsymbol{\Theta}' \mathbf{p}}{\upsilon + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Theta}}; \ \boldsymbol{\xi} = \frac{\upsilon \mathbf{s}_i' - \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Phi}}{\upsilon + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Theta}}. \tag{15}$$

Substituting these expressions back into (13) and equating with (7) then allows us to solve for the parameters using this notation as:

$$\mathbf{p} = -\frac{1}{2}\tilde{\lambda}\tilde{\pi}^* \left[\mathbf{I} - \beta(\boldsymbol{\Theta}\boldsymbol{\xi} + \boldsymbol{\Phi})'\right]^{-1} \mathbf{s}_{\pi^a} \tag{16}$$

$$\begin{aligned} \mathbf{P} &= \tilde{\lambda}\mathbf{S}_{\pi^a} + \mu \mathbf{S}_{g^a} + \upsilon \mathbf{S}_{i^a} \\ &\quad - (\upsilon \mathbf{s}_i' - \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Phi})' (\upsilon + \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Theta})^{-1} (\upsilon \mathbf{s}_i' - \beta \boldsymbol{\Theta}' \mathbf{P} \boldsymbol{\Phi}) + \beta \boldsymbol{\Phi}' \mathbf{P} \boldsymbol{\Phi}. \end{aligned} \tag{17}$$

Equation (14) shows that the optimal policy rule is linear in the current state vector, with (15) determining its intercept and slope coefficients as in the standard model. The linear and quadratic volatility terms $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ do affect the policy rule. They work indirectly through the parameters $\mathbf{p}$ and $\mathbf{P}$ of the value function (7), defined in (10), (11), (16) and (17). Inspection of (14) and (15) shows that $\mathbf{p}$ and

hence linear dependence only affects the intercept coefficient in the policy response function, while quadratic dependence also affects the slope parameters, via its effect on $\mathbf{P}$.

The observation that the monetary authorities should be more aggressive in responding to increases in inflation, if this increases macroeconomic volatility is formalized in equations (10) and (11), but is not new. In his seminal paper, Friedman (1977) suggests that a burst in inflation increases variability of both actual and anticipated inflation. A number of authors, for example, Holland (1993), Holland (1995), Fountas, Karanasos, and Kim (2002), Fountas, Ioannidis, and Karanasos (2004) and Caporale and Kontonikas (2009), follow Friedman (1977) in suggesting that the associated rise in macroeconomic volatility is part of the welfare cost of inflation and that taking account of this effect should make the authorities react more aggressively towards an increase in inflation, thus leading to a lower average inflation rate.

The next section adapts a standard macro model to quantify empirically the impact of inflation conditional volatility on optimal monetary policy in the UK.

3.1 Modelling the macroeconomy

In this section we specify a simple linear structure for the UK economy that is potentially heteroskedastic. This based on the semi-structural dynamic model of Rudebusch and Svensson (1999), which represents the behavior of the macroeconomy in terms of the output gap, inflation and the policy interest rate. This type of model has been extensively used in macro-finance literature on the term structure of interest rates (Dewachter and Lyrio (2006), Rudebusch and Wu (2008)) and in the macroeconomic literature on optimal monetary policy (Favero and Rovelli (2003), Dennis (2006), Cogley, De Paoli, Matthes, Nikolow, and Yates (2011)). We represent inflation (π_t) by the quarterly percentage change in the CED, averaging this over

four quarters to get the annual rate (π_t^a).[4] The Bank of England's base rate is used to represent the policy instrument (i_t). Both of these series were supplied by Datastream. The GDP output gap (g_t) is the OECD measure, based on a trend filtering approach. The estimation sample begins in 1979Q3 following the election of the Thatcher government in May 1979, which saw a move to a more aggressive monetary stance and it ends in 2010Q4. 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 i_t refer to deviations from mean values. The ADF and KPSS statistics suggest that the inflation rate and the base rate have a unit root.

3.2 The dynamic structure

The model describes the dynamic evolution of the output gap and the inflation rate according to the structural relationships:

$$g_t = a_1 g_{t-1} + a_2 g_{t-2} + a_3(i_{t-1}^a - \pi_{t-1}^a) + u_{g,t} \tag{18}$$

$$\pi_t = b_1 \pi_{t-1} + b_2 \pi_{t-2} + b_3 \pi_{t-3} + b_4 \pi_{t-4} + b_5 g_{t-1} + u_{\pi,t}, \tag{19}$$

where $\pi_t^a = \frac{1}{4}\sum_{j=0}^{3} \pi_{t-j}$ and $i_t^a = \frac{1}{4}\sum_{j=0}^{3} i_{t-j}$ are the annual inflation and interest rates. The first equation represents the IS curve, while the second is the Phillips curve. As the restriction $b_4 = 1 - b_1 - b_2 - b_3$ is imposed, the Phillips curve is vertical in the long run. The IS equation implies that in the long run the nominal rate equals the inflation rate, i.e. $i = \pi$ for mean-adjusted data in non-accelerating inflation equilibrium. The error terms $u_{g,t}$ and $u_{\pi,t}$ are conventionally interpreted as demand and supply shocks respectively. The model is augmented to include an

[4] This series is used because it is (implicitly) seasonally adjusted. The annual percentage changes in the CED closely track those in the RPIX (Retail Price Index excluding mortgage interest payments) which was the policy objective (with a target rate of 2.5 %) between November 1992 and April 2004 when it was replaced by the Consumer Price Index (with a target rate of 2%). The CPI data is only available since 1996, but also follows RPIX closely once a time trend is allowed for.

interest rate equation, which describes the actual behavior of the central bank in terms of a systematic component, including on lagged values of the output gap, the inflation rate and the interest rate; and an idiosyncratic component $u_{i,t}$, which is interpreted as the monetary policy shock. The full model is described as:

$$\mathbf{x}_t = \boldsymbol{\phi}_{\mathbf{x}} \mathbf{X}_{t-1} + \boldsymbol{\theta}_{\mathbf{x}} i_{t-1} + \mathbf{u}_{\mathbf{x},t}, \tag{20}$$

$$i_t = \boldsymbol{\phi}_i \mathbf{X}_{t-1} + \theta_i i_{t-1} + u_{i,t}, \tag{21}$$

where: $\mathbf{x}_t = \{g_t, \pi_t\}'$, $\mathbf{u}_{\mathbf{x},t} = \{u_{g,t}, u_{\pi,t}\}$ and $\mathbf{X}_t$ includes current and lagged values of $\mathbf{x}_t$ and i_{t-1}. The complete description of all the equations of the model, their state space representation and how they are mapped into the system in equation (1) is in Appendix 3. The parameters of the policy rule (21) are initially determined by maximum likelihood (ML) estimation along with those of the state equations (20). We refer to these as the ML models. The likelihood-based estimation of the parameters entering (20) and (21) is straightforward: given an initial set of parameters values (typically this employs the OLS estimates of the homoskedastic model), a vector of shocks and a sample of the observations, the likelihood function is computed through the Kalman filter, and the original parameter values updated using simplex methods. Appendix 3 also describes the likelihood function, while appendix 1 specifies the admissibility conditions imposed on the covariance structure to estimate the model under state dependence.

Having then obtained the maximum likelihood estimates of the model, we use the estimated equations for the state variables to mimic the law of motion in the optimization problem and then replace the coefficients of the policy rule in the ML models with those obtained from the optimization procedure. We refer to these as the Direct Control (DC) models.

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

$$i_t = \bar{\kappa} + \bar{\phi}_g \bar{g} + \bar{\phi}_\pi \bar{\pi}. \tag{22}$$

This denotes the long run policy rule, and the actual definition of the coefficients $\bar{\kappa}$, $\bar{\phi}_g$ and $\bar{\phi}_\pi$ is also in appendix 3. Evidently the ML and DC models will incorporate different types of policy rules, which in turn will lead to different long run values. In particular, we note that $\bar{\kappa} = 0$ for the mean-adjusted ML models, but can be non-zero when the equilibrium is shifted by the linear dependence effect in the heteroskedastic DC models. With $i = \bar{\pi}, \bar{g} = 0$ in a non-accelerating inflation equilibrium we have: $i = \bar{\pi} = \bar{\kappa}/(1 - \bar{\phi}_\pi)$, where the denominator is negative under the Taylor (1993) principle $\bar{\phi}_\pi > 1$. The stochastic structure of the model is of the form (4) and specified in appendix 3.

4 The empirical (ML) models

Appendix 3 derives the likelihood function for the model formed by (20), (21) and (38) and outlines the ML estimation procedure. We start by estimating the homoskedastic model ML0, which provides a set of baseline parameters for the dynamic model (1). This model has a likelihood value of (-) 122.7, as reported in Table 2. Then the likelihood is optimized with respect to the parameters of the stochastic structure (4) keeping these baseline parameters fixed. This two-stage exercise immediately reveals the significance of state-dependent volatility. Table 2 shows that the likelihood increases to (-) 70.2 once linear dependence is allowed for (in model ML01) and increases further to (-) 63.1 when an additional allowance is made for quadratic dependence (in model ML03). Quadratic dependence on its own (ML02) does not

produce as large an improvement in fit as these two models. These ML0X, where X=1,2,3, models are used in the next section to study the effects on optimal monetary policy and welfare of changing the stochastic structure (4) while keeping the parameters of the dynamic model (1) unchanged.

We then used these results as the starting values for a fully optimized set of models (respectively MLX, for X=1,2,3) in which all relevant parameters were optimized. This produces a further modest improvement in fit. Table 3 reports these fully optimized ML results. Model ML3 has 29 parameters and a likelihood value of (-) 56.14.[5] The table compares this value ($\ln L_u$) with that of each restricted model ($\ln L_r$). It reports the loglikelihood ratio test statistic $2\times(\ln L_u - \ln L_r)$, which has the 1% critical values $\chi^2(3)$ =11.35 and $\chi^2(6)$ =15.09. All three restricted models are rejected on test. However, to guard against overfitting the table also reports the difference in the Schwarz approximation to the Posterior Odds ratio (SCA=($\ln L_u - \ln L_r$)-0.5$\times$($k_u - k_r$)$\times \ln(T)$) as proposed by Canova (2007). On this criterion the pure quadratic variance model ML2 is decisively rejected against the encompassing model, as is the homoskedastic model ML0. The performance of the linear variance model ML1 is very similar to that of ML3 on this criterion. These tests strongly support the inflation-conditional volatility hypothesis. The results in Table 3 are consistent with a large body of the empirical literature in support of the relevance of this hypothesis in UK data, Grier and Perry (1996), Fountas, Karanasos, and Kim (2002), Fountas, Ioannidis, and Karanasos (2004), Conrad, Karanasos, and Zeng (2010).

4.1 Parameter estimates and residuals

Table 4 reports estimates of the parameters of these MLX models. The one-quarter-ahead forecast values and 95% confidence intervals for the three macro variables in

[5] These 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_i; $\mathbf{G}(1)$; $\mathbf{g}(2)$; $\mathbf{D}_0(3)$; $\mathbf{D}_1(3)$ and $\mathbf{D}_2(3)$.

ML3 are shown in the upper panels of figures 1-3, while the lower panels show the unconditioned residuals (the u's in (38)) and their error bands. ML3 conditions the variance structure using both linear and quadratic inflation terms, although the quadratic component is not significant in the base rate equation. This effect is evident in these figures, meaning that volatility is particularly high when inflation is elevated between 1979 and 1982. These figures all show a very low level of volatility between 1994 and 2003 - the period of the NICE decade. This was interrupted by the recent credit crunch, which is reflected in large negative outliers in the final quarter of 2008 and first quarter of 2009, following the collapse of Lehman Brothers. However, the subsequent residuals remain low, consistent with the low volatility implied by the relatively low level of inflation.

Importantly, conditioning the error structure in this way means that the likelihood function in models ML1-3 discount the large errors (the u's in (38)) that occur during the high inflation period. This means that the errors in that period tend to be larger than in ML0. Consequently the sum of squared errors (u^2) is higher in ML3 (635.3) than in ML1 (619.7) and ML0 (579.9). In this sense, the standard homoskedastic model underestimates the degree of macroeconomic volatility. Thus, we see that neglecting the inflation conditional volatility effect leads researchers to significantly understate the volatility of the system. This conditioning also affects the deterministic parameters in these models because it acts like a weighted regression system that gives observations a weight that varies inversely with inflation.

4.2 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 the reduced form innovations ($u_t's$) are correlated empirically, we work with

orthogonalized innovations using the triangular factorization defined in (38). The orthogonalized impulse responses show the effect on the macroeconomic system of increasing each of these shocks by one percentage point for one or several periods using the Wald representation of the system. This arrangement is affected by the ordering of the macroeconomic variables. We adopt the standard ordering: $\{g_t, \pi_t, i_t\}$, interpreting v_g as a positive demand shock, v_π as a negative supply shock, and v_i as a contractionary monetary shock. As noted in section 2.2, we need to distinguish one period shocks in which the impulse response functions depend only on the parameters of the deterministic equation (1) and longer-lasting shocks in which non-linear amplified and asymmetric responses can occur.

Figure 4 shows the effect of unit one period shocks in $\{g_t, \pi_t, i_t\}$. The impulse response functions for ML0, ML1, ML2 and ML3 are shown by dotted; dashed; thin continuous and thick continuous continuous schedules respectively. A temporary shock in real output leads to a decline of inflation on impact and an increase in the nominal rate. The initial effect on the nominal rate in the heteroskedastic models is larger than in ML0. The increase in inflation pushes up output initially, but then output declines in all four models as the real rate responds. A nominal rate shock causes output and inflation to fall in all three models.

Figure 5 plots the responses of the output gap, annual inflation and annual interest rate to a 5 year sequence of positive (continuous lines) and negative (dashed lines) unit shocks to inflation.[6] To analyze the amplification and asymmetric effects emerging under conditional volatility we keep the dynamic parameters of equation (1) fixed across these models and use the baseline ML0X variants shown in table 2. In the homoskedastic model ML0, the increase in inflation looks like a linear trend

[6] The amplification and asymmetric effects under linear quadratic dependence are also visible when cosidering permanent shocks to both output and inflation.

over the 5 year period and output declines because the nominal rate increases more than inflation. The patterns for negative shocks are mirror images of these. The amplification effect is clearly evident for a sequence of positive shocks in the heteroskedastic models: inflation increases more than in the homoskedastic model and the amplification effect is largest under ML03. This leads to larger responses of the nominal rate, which depress output.

In the pure quadratic models ML2 and DC02, the linear dependence effect is absent and the simulated values from pairs of antithetical simulations are mirror images of each other as in ML0. However, the asymmetry of the responses is evident under both ML01 and ML03 when the economy is hit by a sequence of negative inflation shocks. As explained in section 2.2 and appendix 1, this attenuation effect puts a lower bound on the variable driving volatility,[7] as in the Cox, Ingersoll, and Ross (1985) model.[8] Inflation approaches the lower bound after about 10 periods in the ML01 simulation. This linear dependence effect is also a feature of the ML03 and DC03 models but is offset by the quadratic amplification effect which dominates at high or low inflation rates. This effect removes the lower bound on inflation, as appendix 1 explains.

5 Optimal control

Sections 2 and 3 showed how dynamic linear models with state-dependent variance structures could be employed for the study of optimal macroeconomic policy. The previous section used a simple macroeconomic model with a variance structure that allows for linear and quadratic dependence on the lagged inflation rate to capture

[7] If the simulated value of inflation approaches the lower bound this shuts down the volatility structure temporarily. However mean reversion means that the inflation rate then tends to move back up slowly, switching the volatility back on.

[8] In the continuous time CIR model, the interest rate drives its own volatility and is non-negative, having an asymmetric non-central χ^2 distribution.

the inflation conditional volatility effect found in empirical studies of UK data. In this section we use these results to quantify the welfare costs of UK macroeconomic volatility and its effect on the conduct of optimal monetary policy. In order to analyze the effect of different stochastic and welfare specifications, in this section we use the benchmark parameter estimates (shown as ML0 in table 4) for the two state equations (18) and (19) throughout, alongside the stochastic parameters shown in table 2 for the benchmark ML0X heteroskedastic models.[9] We then replace the empirical Taylor rules by the optimal base rate rules implied by different stochastic structures and welfare weights. We label the models under control derived from ML0, ML01, ML02 and ML03 as DC0, DC01, DC02 and DC03 respectively. Appendix 2 describes how we combine the state equations and the optimal policy rule to form the model under control used for impulse response and welfare analysis.

5.1 Steady state base rate responses

Columns 2-6 of Table 5 present the long run coefficients implied by the optimal policy rules for the four different ML0X stochastic specifications under the five different specifications of the welfare weights used by Rudebusch and Svensson (1999).[10] The first column reports the long run coefficients of the benchmark empirical policy rule implied by ML0. We used a stylized assumption about the discount rate, which is set at six percent.

Several patterns are apparent in this table. First, reading across the table we see the effect of different welfare weights. The first set of weights is a benchmark that gives inflation, output and base rate smoothing equal weight in the loss function

[9] Results for the fully estimated heteroskedastic MLX models (ML1-ML3), which include the additional effect of changes in the estimated parameters (18) and (19), are qualitatively similar and available upon request.

[10] These parameter sets are also used by Rudebusch (2002) in his analysis of money GDP rules for the UK.

($\mu = \upsilon = 1$). The second set shows the effect of reducing the base rate smoothing parameter, which makes the monetary authorities more responsive to increases in both inflation and output. The third set shows the effect of then reducing the weight given to the output gap, making the monetary authorities still more aggressive in response to rising inflation. The fourth set shows the effect of attaching a much higher weight to output volatility than in cases 2 and 3, while the last specification considers a case in which much less weight is given to changes in the policy instrument than in case 1.

Reflecting the results of sections 2 and 3, we see that because they use the same state equations, the policy responses in DC01 are identical to those of the homoskedastic model DC0. However linear dependence has the effect of introducing a positive intercept ($\bar{\kappa}$) into the base rate equation in DC01, depressing the steady state rate of inflation ($\bar{\pi}$). In DC02, quadratic dependence has the effect of making policy much more responsive to inflation, without affecting the steady state. In DC03 both effects are present: the shift in $\bar{\kappa}$ is slightly larger, but the more aggressive inflation response has the effect of damping the effect on the steady state rate of inflation compared to DC01.

In contrast to optimal control studies of US monetary policy, which generally find that the optimal long run responses to inflation in a standard homoskedastic model are larger in absolute value than those implied by the empirical estimates, the optimal control responses in our equivalent DC0 model tend to straddle the empirical benchmark shown in the first column. The optimal inflation responses in the heteroskedastic models DC02 and DC03 are naturally more aggressive. These responses are all larger than those proposed by Taylor (1993). The output responses are much higher than observed empirically since 1979. This is true for all models and welfare specifications. We also find that the optimal output gap response is still

much larger than the empirical one even when the model is simulated with $\mu = 0$. This is because the output gap acts as a leading indicator of movement in inflation, an observation that might justify the prominence of the negative output gap in the Monetary Policy Committee's deliberations of monetary policy at the moment.

5.2 The optimal impulse responses

Section 4.2 discussed the impulse responses for the three ML models. In this section we compare the responses to unit one period shocks for the benchmark models under control, DC0, DC02 and DC03.[11] These are shown in figure 6 as the dotted, dashed and continuous schedules respectively. We only present the results for the first set of welfare parameters $\mu = \upsilon = 1$ since the results are not markedly affected by alternative choices. The policy rate is far more responsive to output and inflation shocks than under the empirical rules in the homoskedastic models. This is reflected in the much sharper response in output shown in the top panel. Consistent with the theory, the policy response for DC03 is more aggressive than for the other models. The simulations reported in the next section are designed to show the responses of the model to longer-lasting sequences of random shocks.[12]

5.3 Welfare analysis

How would policy optimization have affected the volatility of the system and the imputed welfare losses? To answer this we follow Sack (2000) and simulate the various models and welfare specifications stochastically, recording the standard deviations of output, inflation and interest rate changes and the welfare losses. We start by creating a flat benchmark path setting the starting values in the state vector to zero, to keep the steady state values in line with the zero mean of the empirical sample.

[11] The responses for DC01 are identical to those for DC0, as in table 4.

[12] A figure showing the effect of 5 year sequence of unit shocks to inflation comparable with that of figure 5 for the ML0X models is available upon request.

Then, reflecting the change in the steady state rate of inflation implied by DC01 and DC03, we reduce the benchmark path for inflation for these two models by the respective value of $\bar{\pi}$ shown in table 5. We then use the Matlab function *randn*, which produces standardized random normal variates, to generate 500 random paths for the homoskedastic errors (the v's in (3) and (38)). Changing the sign on these then gives a total of 1,000 antithetical normal variates.[13] This set is then used to perturb the benchmark path using (1) and (3) in each model/weight combination. Although these sets of shocks are the same for each such combination, it is important to note that the final disturbances (the u's) generated by (3) can be very different, particularly for sequences of large shocks with the same sign. The use of antithetical variable shocks makes the simulation results for positive and negative residual tracks symmetric for the ML0, DC0, ML02 and DC02 models. However, recall that in the linear dependence models ML01 and DC01, a series of negative output and inflation shocks has the cumulative effect of lowering the simulated value of inflation, thus attenuating the effect of future shocks. This linear dependence effect is also a feature of the ML03 and DC03 models but is offset by the quadratic dependence effect which dominates at high or low inflation rates, amplifying volatility.

Table 6 shows the results obtained by simulating the ML models with their empirical policy rules. The first three columns of numbers show the standard deviations of the three goal variables and the remaining columns show the welfare losses implied by the five welfare specifications used in table 5. The losses fall as we move from case 1 to 3 reducing respectively the base rate smoothing and output weights. Case 4 shows the effect of a large output weight and case 5 that of a very low rate smoothing weight. This table shows a fall in the volatility of inflation and interest rates in ML01 compared to ML0. This reflects the asymmetric volatility attenuation effect of low

[13] We trim 10% of the simulated series to eliminate the impact of extreme draws.

inflation paths noted earlier. The quadratic amplification effect in ML02 increases the variances of the goal variables and hence the welfare losses compared to the other models. Again, ML03 combines both effects and implies a welfare loss somewhere between that of ML01 and ML02.

Table 7 compares the results obtained by simulating the three models with the optimal policy rule using the five welfare specifications to get the DC results. These welfare specifications affect the policy rules and hence the volatilities of the goal variables. Optimization reduces the welfare loss relative to the empirical rules, with the percentage welfare gains being shown in the final column. In the DC0 specification, the gain is achieved by reducing the variability of inflation at the expense of increasing that of output and/or interest rates. The variances are fixed in this specification ($\mathbf{\Sigma}_t = \mathbf{\Sigma}_0$), so a reduction in the variance of one variable has to be traded off against an increase in that of another. However, the low downside risk in DC01 makes it optimal to lower the steady state inflation rate (as shown in table 5), thus shifting the trade-off and reducing the overall variability of the system. DC01 is able to reduce the variability of all the goal variables compared to DC0, for all of these welfare parameter values. Optimization makes a much bigger difference in this case than it does in the standard model, even though it comes through a reduction in the steady state (like a reduction in the inflation target) and not an increase in the level of aggression. As we saw in table 6, the quadratic amplification effect increases the welfare losses in ML02 compared to the other models, but the gains from optimization are nevertheless bigger than in the standard model, taking the form of a change in the level of aggression without affecting the steady state. DC03 combines both a more aggressive stance and a shift in the steady state (shown in table 5) and reduces the welfare losses to about a fifth of those implied in ML03. These results suggest that mis-specification of the variance structure can lead the researchers to

seriously understate the potential gains from optimization.

6 Conclusion

Time varying stochastic volatility is remarkably significant in UK macroeconomic data; our empirical results show that this feature can be well captured by a simple state dependent macroeconomic model in which the variance of the output gap, inflation and interest rate data exhibits quadratic as well as linear dependence. This phenomenon helps explain the NICE decade, the UK equivalent of the US 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 apparently returned to the low volatility implied by the relatively low rate of inflation. The empirical results suggest that the conventional model significantly understates the degree of volatility in the macroeconomy as well as neglecting its association with the level of inflation.

There is a burgeoning empirical literature highlighting the role of time varying macroeconomic volatility as a feature of macroeconomic data in most industrialized countries. This paper is the first to explore the implications of this for the design of optimal monetary policy and the analysis of the welfare costs of inflation. Theoretically, linear dependence reduces the risk of deflation and makes it optimal for monetary authorities to reduce the inflation target relative to both the sample mean of the data and the homoskedastic optimization model target. The optimal policy calculations reveal that, depending upon the welfare specification, this effect would have reduced the target by one or two percentage points compared to the sample mean of $4\frac{3}{4}\%$, in turn reducing the welfare losses (in DC01) relative to those implied by the optimization of the homoskedastic specification (DC0).

We find that optimizing the standard homoskedastic model implies a level of aggression comparable to that seem empirically since 1979. However, this model is dominated empirically by the heteroskedastic model ML3, which implies a higher level of macroeconomic volatility and makes it optimal for the monetary authorities to react more aggressively to inflation shocks. The welfare costs under the optimal rule are about a fifth of those implied by the empirical rule for this model. These cost reductions are much larger than those suggested by the conventional model.

We believe that this paper opens the way to the more general use of time-varying macroeconomic volatility models in optimal policy analysis. The paper analyses the policy implications of state dependent volatility in the conduct of optimal monetary policy abstracting from issues such as parameter uncertainty (Sack (2000), Soderstrom (2002)) or learning about the unknown state of the economy (Ellison and Valla (2001)). These additional considerations could make optimal policy rules more or less aggressive relative to the benchmark analysis in this paper, and are potentially interesting extensions of the analysis provided here. However, they are unlikely to change the central result of this paper, which is that mis-specification of the variance structure can lead researchers to seriously understate both the welfare cost of inflation and the potential gains from optimization.

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] If the variance of inflation is driven by (39) with $k = \pi$ and $\delta_{\pi,2} = 0$ (which gives a model of the Cox, Ingersoll, and Ross (1985) type):

$$\delta_{\pi,t} = \delta_{\pi,0} + \delta_{\pi,1}\pi^a_{t-1} \geq 0$$

then the variance exhibits linear dependence and is shut off at $\delta_{\pi,t} = 0$. This puts a lower bound on the driving variable of $\pi^a_{\min} = -\delta_{\pi,0}/\delta_{\pi,1} \leq 0$. For example using the parameters for ML01 shown in table 0 gives a lower bound of $\pi^a_{\min} = -1.01/0.3185 = -3.1711$ on the Rudebusch and Svensson (1999) definition (that follows (19)), as reflected in figure 5. This is -12.684% when expressed as an annual logarithmic change. .

Admissibility is more problematic in discrete time square root volatility models (i.e. linear-dependent variance structures) 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 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 possible values of the driving variable (in

[14] It 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.

this paper, $z = \pi^a$). These are given by (39) requiring $\delta_{k,t} \geq 0, k = g, \pi, i$:

$$\delta_{k,t} = \delta_{k,0} + \delta_{k,1}\pi^a_{t-1} + \delta_{k,2}{\pi^a_{t-1}}^2 \geq 0.$$

This is ensured provided that $4\delta_{k,0}\delta_{k,2} \geq \delta^2_{k,1}$ so that the roots of the associated quadratic equation are complex. The absence of a real valued solution means that there is no lower bound, although the linear component does reduce downside risk relative to the pure quadratic model 2. 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.

Appendix 2: Representing the general problem in canonical form

This appendix explores the general implications of the dynamic stochastic framework set out in sections 2.1 and 2.2, assuming that the policy maker has a canonical quadratic loss function defined over state variables and instruments:

$$L_t = \sum_{j=0}^{\infty} \beta^{t+j} E_t[(\mathbf{X}_{t+j} - \mathbf{X}^*)'\mathbf{\Lambda}(\mathbf{X}_{t+j} - \mathbf{X}^*) + \upsilon\ (i_{t+j} - i^*)^2 + 2(\mathbf{X}_{t+j} - \mathbf{X}^*)'\mathbf{H}\ (i_{t+j} - i^*)], \tag{23}$$

where E_t is the time $t-$conditional expectations operator; β is the discount factor; $\mathbf{X}^*$ and i^* are target or bliss vectors for $\mathbf{X}_{t+j}$ and i_{t+j} respectively; $\mathbf{\Lambda}$ and $\mathbf{H}$ are matrices of constant welfare weights; and υ is a weight attached to deviations of the policy instruments from target. The policy maker is assumed to choose the intertemporal sequence of policy instruments $\{i_{t+j}\}_{j=0}^{\infty}$ that minimizes the loss function (23) given the model of the economy in (1)-(4) and the initial state vector $\mathbf{X}_t$.

This minimization problem can be solved using standard dynamic programming

techniques. Since the per-period loss function is quadratic and the dynamics are linear, the value function in the canonical problem takes the form (7). The optimal policy is then found from the solution of a recursive Bellman equation that is obtained by substituting (23) into (6) and using (7) to replace $J(\mathbf{X}_{t+1})$:

$$J(\mathbf{X}_t) = \min_{i_t}[(\mathbf{X}_t - \mathbf{X}^*)' \mathbf{\Lambda} (\mathbf{X}_t - \mathbf{X}^*) + \upsilon(i_t - i^*)^2 + 2(\mathbf{X}_t - \mathbf{X}^*)' \mathbf{H}(i_t - i^*)$$
$$+\beta E_t \left(\mathbf{X}'_{t+1}\mathbf{P}\mathbf{X}_{t+1} - 2\mathbf{X}'_{t+1}\mathbf{p} + c\right)].$$

This is optimized subject to the constraints (1) and (4). Evaluating expectations using these constraints gives the conditional volatility problem:

$$J(\mathbf{X}_t) = \min_{i_t} \left\{ \begin{array}{c} (\mathbf{X}_t - \mathbf{X}^*)' \mathbf{\Lambda} (\mathbf{X}_t - \mathbf{X}^*) + \upsilon(i_t - i^*)^2 + 2(\mathbf{X}_t - \mathbf{X}^*)' \mathbf{H}(i_t - i^*) + \beta c + I_t \\ +\beta tr(\mathbf{P}\mathbf{\Sigma}_1)\mathbf{X}'_t\mathbf{s}_z + \beta tr(\mathbf{P}\mathbf{\Sigma}_2)\mathbf{X}'_t\mathbf{S}_z\mathbf{X}_t. \end{array} \right\}, \tag{24}$$

where I_t is defined in equation (9). The first line of (24) shows the Bellman form of the standard homoskedastic control problem, in which: $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{0}$. In this case, the target rate for the variables in $\mathbf{X}_t$ is the sample mean and if the model is specified in terms of mean-differences then: $\mathbf{X}^* = 0$. The certainty equivalence principle also holds in this case. The last two terms of (24) capture the effect of the state dependent variance structure on the control problem. Since these are, respectively, linear and quadratic in $\mathbf{X}_t$, they affect the coefficients of the optimal policy rule. However, we can consolidate these with the other linear and quadratic terms and write the loss function in the canonical form given by the first two lines of (24) by setting: $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{0}$ and replacing the welfare parameters $\mathbf{\Lambda}, \mathbf{X}^*, i^*$ and c in (23) by

$\tilde{\mathbf{\Lambda}}$, $\tilde{\mathbf{X}}^*, \tilde{\imath}^*$ and $\tilde{c}$ to get:

$$L_t = \sum_{j=0}^{\infty} \beta^{t+j} E_t[(\mathbf{X}_{t+j} - \tilde{\mathbf{X}}^*)' \tilde{\mathbf{\Lambda}} (\mathbf{X}_{t+j} - \tilde{\mathbf{X}}^*) + \upsilon \ (i_{t+j} - \tilde{\imath}^*)^2 + 2(\mathbf{X}_{t+j} - \tilde{\mathbf{X}}^*)' \mathbf{H} \ (i_{t+j} - \tilde{\imath}^*)], \tag{25}$$

where:

$$\tilde{\mathbf{\Lambda}} = \mathbf{\Lambda} + \beta tr\left(\mathbf{P}\Sigma_2\right) \mathbf{S}_z \tag{26}$$

$$\tilde{\mathbf{X}}^* = (\tilde{\mathbf{\Lambda}} - \mathbf{H}\mathbf{H}'/\upsilon)^{-1}[(\mathbf{\Lambda} - \mathbf{H}\mathbf{H}'/\upsilon)\mathbf{X}^* - \beta tr\left(\mathbf{P}\mathbf{\Sigma}_1\right) \mathbf{s}_z/2] \tag{27}$$

$$\tilde{\imath}^* = i^* + \mathbf{H}'\left(\mathbf{X}^* - \tilde{\mathbf{X}}^*\right)/\upsilon \tag{28}$$

$$\tilde{c} = c + \left[\mathbf{X}^{*\prime}\mathbf{\Lambda}\mathbf{X}^* - \tilde{\mathbf{X}}^{*\prime}\mathbf{\Lambda}\tilde{\mathbf{X}}^* - \upsilon(i^* - \tilde{\imath}^*)^2\right]/\beta - 2\mathbf{X}^{*\prime}\mathbf{H}\mathbf{H}'\left(\mathbf{X}^* - \tilde{\mathbf{X}}^*\right)/\upsilon. \tag{29}$$

This allows the optimization problem to be expressed in the form of the standard Bellman equation:

$$J\left(\mathbf{X}_t\right) = \min_{i_t}[(\mathbf{X}_t - \tilde{\mathbf{X}}^*)' \tilde{\mathbf{\Lambda}} (\mathbf{X}_t - \tilde{\mathbf{X}}^*) + \upsilon(i_t - \tilde{\imath}^*)^2 + 2\left(\mathbf{X}_t - \tilde{\mathbf{X}}^*\right)' \mathbf{H}(i_t - \tilde{\imath}^*) + \beta\tilde{c} + I_t]. \tag{30}$$

To demonstrate the equivalence of (24) and (30), we expand the first quadratic term in (24) and use (26) to rearrange this as:

$$J\left(\mathbf{X}_t\right) = \min_{i_t} \left\{ \begin{array}{c} \mathbf{X}_t' \tilde{\mathbf{\Lambda}} \mathbf{X}_t - 2\mathbf{X}_t'[\mathbf{\Lambda}\mathbf{X}^* - \beta tr\left(\mathbf{P}\Sigma_1\right) \mathbf{s}_z/2] + \mathbf{X}^{*\prime}\mathbf{\Lambda}\mathbf{X}^* + \upsilon(i_t - i^*)^2 \\ +2\left(\mathbf{X}_t - \mathbf{X}^*\right)' \mathbf{H}(i_t - i^*) + \beta c + I_t \end{array} \right\}. \tag{31}$$

We then write (30) as:

$$J\left(\mathbf{X}_t\right)=\min_{i_t}\left\{\begin{array}{c}\mathbf{X}_t'\tilde{\boldsymbol{\Lambda}}\mathbf{X}_t-2\mathbf{X}_t'\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{X}}^*+\tilde{\mathbf{X}}^{*\prime}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{X}}^*\\ +\upsilon[(i_t-i^*)^2+(i^*-\tilde{\imath}^*)^2+2(i_t-i^*)(i^*-\tilde{\imath}^*)]+I_t+\beta\tilde{c}\\ 2\left(\mathbf{X}_t-\mathbf{X}^*\right)'\mathbf{H}(i_t-i^*)+2\left(\mathbf{X}_t-\mathbf{X}^*\right)'\mathbf{H}(i^*-\tilde{\imath}^*)+2\left(\mathbf{X}^*-\tilde{\mathbf{X}}^*\right)'\mathbf{H}(i_t-\tilde{\imath}^*)\end{array}\right\}. \tag{32}$$

Equating (31) and (32) and cancelling common terms, we require:

$$\begin{aligned}&\mathbf{X}^{*\prime}\boldsymbol{\Lambda}\mathbf{X}^*+\beta c-2\mathbf{X}_t'[\boldsymbol{\Lambda}\mathbf{X}^*-\beta tr\left(\mathbf{P}\Sigma_1\right)\mathbf{s}_z/2]\\ =&\tilde{\mathbf{X}}^{*\prime}\boldsymbol{\Lambda}\tilde{\mathbf{X}}^*-2\mathbf{X}_t'\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{X}}^*+\upsilon(i^*-\tilde{\imath}^*)^2+2\left(\mathbf{X}_t-\mathbf{X}^*\right)'\mathbf{H}(i^*-\tilde{\imath}^*)+\beta\tilde{c},\\ &+2[\upsilon(i^*-\tilde{\imath}^*)+2\left(\mathbf{X}^*-\tilde{\mathbf{X}}^*\right)'\mathbf{H}](i_t-i^*)\\ =&\tilde{\mathbf{X}}^{*\prime}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{X}}^*-2\mathbf{X}_t'\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{X}}^*+\upsilon[(i^*-\tilde{\imath}^*)^2]-2\left(\mathbf{X}_t-\mathbf{X}^*\right)'\mathbf{H}\mathbf{H}'\left(\mathbf{X}^*-\tilde{\mathbf{X}}^*\right)/\upsilon+\beta\tilde{c},\end{aligned}$$

where the last line follows by substituting condition (28). Equating the coefficients of $\mathbf{X}_t$ gives $[\boldsymbol{\Lambda}\mathbf{X}^*-\beta tr\left(\mathbf{P}\Sigma_1\right)\mathbf{s}_z/2]=\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{X}}^*+\mathbf{H}\mathbf{H}'\left(\mathbf{X}^*-\tilde{\mathbf{X}}^*\right)/\upsilon$, with the solution (27). Finally, equating the respective intercept terms gives (29).

The solution can then be obtained by appropriate use of the algorithms developed for the standard homoskedastic problem (Lungqvist and Sargent (2004)). If the parameters of (1) and (4) are independent of the policy rule as they are for example in the model developed in section 3, the optimal value is obtained simply by differentiating (30) w.r.t. i_t to get the closed loop solution:

$$i_t=\zeta+\boldsymbol{\xi}\mathbf{X}_t \tag{33}$$

$$\zeta=\left(\upsilon+\beta\boldsymbol{\Theta}'\mathbf{P}\boldsymbol{\Theta}\right)^{-1}\beta\boldsymbol{\Theta}'\mathbf{p} \tag{34}$$

$$\boldsymbol{\xi}=-\left(\upsilon+\beta\boldsymbol{\Theta}'\mathbf{P}\boldsymbol{\Theta}\right)^{-1}\left(\mathbf{H}'+\beta\boldsymbol{\Theta}'\mathbf{P}\boldsymbol{\Phi}\right). \tag{35}$$

The matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ showing the effect of the linear and quadratic volatility terms affect the policy rule indirectly, through the parameters $\mathbf{p}$ and $\mathbf{P}$ of the value function (7). These values are obtained by substituting (33) - (35) back into (30) and equating with (7) to obtain the standard formula:

$$\mathbf{p} = \left[\mathbf{I} - \beta\left(\mathbf{\Theta}\boldsymbol{\xi} + \mathbf{\Phi}\right)'\right]^{-1}\left[\left(\tilde{\Lambda} + \boldsymbol{\xi}'\mathbf{H}'\tilde{\mathbf{X}}^*\right) + \left(\boldsymbol{\xi}'\upsilon + \mathbf{H}\right)i^*\right] \tag{36}$$

and the standard Riccati equation:

$$\mathbf{P} = \tilde{\mathbf{\Lambda}} - \left(\mathbf{H}' + \beta\mathbf{\Theta}'\mathbf{P}\mathbf{\Phi}\right)'\left(\upsilon + \beta\mathbf{\Theta}'\mathbf{P}\mathbf{\Theta}\right)^{-1}\left(\mathbf{H}' + \beta\mathbf{\Theta}'\mathbf{P}\mathbf{\Phi}\right) + \beta\mathbf{\Phi}'\mathbf{P}\mathbf{\Phi}. \tag{37}$$

The optimal response coefficients are then obtained by solving (37) numerically and substituting $\mathbf{P}$ back into (36), and (33) - (35).

A Appendix 3: Estimating and optimizing the empirical model

This canonical model of section 2 and appendix 1 can be used to describe various models appearing in the control literature. In this appendix we show how the empirical model of section 3 fits into this general structure and describe the estimation procedure. We then show how we obtain the policy rule implied by the state equations and the loss function of section 2.4.

A.1 The stochastic structure

The residuals in (20) and (21) are potentially heteroskedastic, driven by the annual CED inflation rate (lagged one quarter):

$$\begin{bmatrix} \mathbf{u}_{\mathbf{x},t} \\ u_{i,t} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0}_2 \\ \mathbf{g}' & 1 \end{bmatrix} \begin{bmatrix} \mathbf{D}_{x,t}^{1/2} & \mathbf{0}_2 \\ \mathbf{0}_2' & \delta_{i,t}^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{\mathbf{x},t} \\ v_{i,t} \end{bmatrix} \tag{38}$$

where:

$$\mathbf{D}_{x,t} = \mathbf{D}_{x,0} + \mathbf{D}_{x,1}\pi^a_{t-1} + \mathbf{D}_{x,2}(\pi^a_{t-1})^2 \tag{39}$$

$$\delta_{i,t} = \delta_{i,0} + \delta_{i,1}\pi^a_{t-1} + \delta_{i,2}(\pi^a_{t-1})^2$$

and where $\mathbf{G} = \begin{bmatrix} 1 & 0 \\ g_1 & 1 \end{bmatrix}$ and $\mathbf{g}' = \{g_2, g_3\}'$ include constant parameters; $\mathbf{v}_{\mathbf{x},t} = \{v_{g,t}, v_{\pi,t}\}' \sim N(\mathbf{0}_2, \mathbf{I}_2)$ and $v_{i,t} \sim N(0,1)$ are Gaussian homoskedastic shocks to output, inflation and interest rate respectively; $\mathbf{D}_{x,j} = diag\{\delta_{g,j}, \delta_{\pi,j}\}$, $j = 0, 1, 2$, with $\delta_{g,j}$, and $\delta_{\pi,j}$ again denoting constant parameters.[15]

The state space form

Equations (18) and (19) describing the current state variables of the system $\mathbf{x}'_t = \{g_t, \pi_t\}$ can be written in the matrix form (20) using

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

$$\mathbf{X}'_t = \left\{\mathbf{x}'_t, i_{t-1}, \mathbf{x}'_{t-1}, i_{t-2}, \mathbf{x}'_{t-2}, i_{t-3}, \mathbf{x}'_{t-3}\right\}$$

$$\boldsymbol{\theta}'_{\mathbf{x}} = \{a_3/4, 0\}.$$

The policy interest rate can be written as

$$i_t = \theta_i i_{t-1} + c_1 g_{t-1} + c_2\pi_{t-1} + c_3 i_{t-2} + c_4 g_{t-2} + c_5\pi_{t-2} + c_6 i_{t-3} + c_7\pi_{t-3}$$
$$+ c_8 i_{t-4} + c_9\pi_{t-4} + u_{i,t}$$

[15] In this paper, $diag\{\delta\}$ represents a matrix with the elements of the row vector δ in the main diagonal and zeros elsewhere. $\mathbf{0}_a$ is the $(a \times 1) \times 1$ zero vector; 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.

which can then be arranged in a form compatible with equation (21) using:

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

The interest rate rule in the long run is written in equation (22), using:

$$\bar{\phi}_g = (c_1 + c_4)/(\theta_i + c_3 + c_6 + c_8) \tag{40}$$

$$\bar{\phi}_\pi = (c_2 + c_5 + c_7 + c_9)/(\theta_i + c_3 + c_6 + c_8). \tag{41}$$

The coefficient $\bar{\kappa} = 0$ 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. The evolution of the state vector is described by the companion form (1), where specifically: $\boldsymbol{\Theta}' = \{\theta_i', \mathbf{0}_9'\}$, and $\mathbf{U}_t' = \left\{\mathbf{u}_{\mathbf{x},t}', \mathbf{0}_9'\right\}$ are 11×1 deficient coefficient and error vectors and:

$$\boldsymbol{\Phi} = \begin{bmatrix} \{a_1, -a_3/4\} & a_3/4 & \{a_2, -a_3/4\} & a_3/4 & \{0, -a_3/4\} & a_3/4 & \{0, -a_3/4\} \\ \{b_5, b_1\} & 0 & \{0, b_2\} & 0 & \{0, b_3\} & 0 & \{0, b_4\} \\ \mathbf{0}_2' & 0 & \mathbf{0}_2' & 0 & \mathbf{0}_2' & 0 & \mathbf{0}_2' \\ \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} \\ \mathbf{0}_2' & 1 & \mathbf{0}_2' & 0 & \mathbf{0}_2' & 0 & \mathbf{0}_2' \\ \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} \\ \mathbf{0}_2' & 0 & \mathbf{0}_2' & 1 & \mathbf{0}_2' & 0 & \mathbf{0}_2' \\ \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{0}_{2,2} & \mathbf{0}_2 & \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_{2,2} \end{bmatrix}.$$ [16]

[16] In this paper, $\mathbf{0}_a$ is the $(a \times 1) \times 1$ zero vector; 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.

Finally, to put (38) into form (1) we define:

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

where $\mathbf{\Sigma}_t$ is defined in (4) specifying: $z = \pi^a$;

$$\mathbf{\Sigma}_j = \Gamma \mathbf{\Delta}_j \Gamma',$$

$$\Gamma = \begin{bmatrix} \mathbf{G} & \mathbf{0}_{2,9} \\ \mathbf{0}_{9,2} & \mathbf{0}_{11,11} \end{bmatrix},$$

and: $\mathbf{\Delta}_j = diag\{\delta_{g,j}, \delta_{\pi,j}, \mathbf{0}_9'\}$, $j = 0, 1, 2$.

The likelihood function

Next we derive the likelihood function of the model of section 3 and describe the numerical optimization procedure. Write (20) and (21) as:

$$\mathbf{z}_t = \begin{bmatrix} \boldsymbol{\phi}_{\mathbf{x}} \\ \boldsymbol{\phi}_i \end{bmatrix} + \mathbf{X}_{t-1} \begin{bmatrix} \boldsymbol{\theta}_{\mathbf{x}} \\ \theta_i \end{bmatrix} i_{t-1} + \begin{bmatrix} \mathbf{u}_{\mathbf{x},t} \\ u_{i,t} \end{bmatrix} \tag{42a}$$

$$= \boldsymbol{\phi}\mathbf{X}_{t-1} + \theta i_{t-1} + \mathbf{u}_t. \tag{42b}$$

Similarly, write (38) as:

$$\mathbf{u}_t = \mathbf{C}\mathbf{D}_t^{1/2}\mathbf{v}_t \tag{43}$$

$$\mathbf{v}_t \sim N\left(\mathbf{0}_3, \mathbf{I}_3\right)$$

where: $\mathbf{z}_t' = \{\mathbf{x}_t', i_t\}$, $\mathbf{u}_t' = \left\{\mathbf{u}_{\mathbf{x},t}', u_{i,t}\right\}$; $\mathbf{v}_t' = \left\{\mathbf{v}_{\mathbf{x},t}', v_{i,t}\right\}$ and:

$$\mathbf{C} = \begin{bmatrix} \mathbf{G} & \mathbf{0}_3 \\ \mathbf{g}' & 1 \end{bmatrix}; \quad \mathbf{D}_t = \begin{bmatrix} \mathbf{D}_{\mathbf{x},t} & \mathbf{0}_3 \\ \mathbf{0}_3' & \delta_{i,t} \end{bmatrix}. \tag{44}$$

Then using (43):

$$\begin{aligned} \mathbf{v}_t &= \mathbf{C}^{-1}\mathbf{D}_t^{-1/2}\mathbf{u}_t \\ &= \mathbf{C}^{-1}\mathbf{D}_t^{-1/2}[\mathbf{z}_t - \boldsymbol{\phi}\mathbf{X}_{t-1} - \theta i_{t-1}] \\ \mathbf{v}_t &\sim N\left(\mathbf{0}_3, \mathbf{I}_3\right) \end{aligned}$$

Thus the loglikelihood for period t can be written as:

$$L_t = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(|\mathbf{D}_t|) - \frac{1}{2}\mathbf{v}_t'\mathbf{D}_t^{-1}\mathbf{v}_t \tag{45}$$

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

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

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

Policy optimization

First, the Rotemberg-Woodford type welfare function (5) used in the model of section 3 is put into the canonical form (23) using:

$$\boldsymbol{\Lambda} = \lambda\mathbf{S}_{\pi^a} + \mu\mathbf{S}_g + \upsilon\mathbf{S}_i; \;\; \mathbf{H} = -\upsilon\mathbf{s}_i; \;\; \mathbf{X}^* = \pi^*\mathbf{s}_{\pi^a}. \tag{46}$$

where: $\mathbf{s}_{\pi^a} = 0.25[0,1,0,1,0,0,1,0,0,1,0]'$ and $\mathbf{S}_\pi = \mathbf{s}_\pi \mathbf{s}_\pi'$.

The model under control consists of the state equations (20) and the policy rule (14). Following Polito and Wickens (2011), we partition the vector $\xi = \{\xi_1, \xi_2, \xi_3\}$ conformably with $\{x_t, X_{t-1}, i_{t-1}\}$ and stack these equations to obtain

$$\begin{bmatrix} \mathbf{I}_2 & 0 \\ -\boldsymbol{\xi}_1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ i_t^* \end{bmatrix} = \begin{bmatrix} \mathbf{0}_2 \\ \zeta \end{bmatrix} + \begin{bmatrix} \boldsymbol{\phi}_\mathbf{x} & \boldsymbol{\theta}_\mathbf{x} \\ \boldsymbol{\xi}_2 & \xi_3 \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{\mathbf{x},t} \\ 0 \end{bmatrix}. \tag{47}$$

This can be solved to obtain the reduced form system that is congruent with (42a):

$$\begin{bmatrix} \mathbf{x}_t \\ i_t^* \end{bmatrix} = \begin{bmatrix} \mathbf{0}_2 \\ \zeta \end{bmatrix} + \begin{bmatrix} \boldsymbol{\phi}_\mathbf{x} \\ \boldsymbol{\phi}_{i^*} \end{bmatrix} \mathbf{X}_{t-1} + \begin{bmatrix} \boldsymbol{\theta}_\mathbf{x} \\ \theta_{i^*} \end{bmatrix} i_{t-1} + \begin{bmatrix} \mathbf{u}_{\mathbf{x},t} \\ \xi_1 \mathbf{u}_{\mathbf{x},t} \end{bmatrix}; \tag{48}$$

where:

$$\begin{aligned} \boldsymbol{\phi}_{i^*} &= \boldsymbol{\phi}_{i^*}(\mathcal{P}) = \boldsymbol{\xi}_1(\mathcal{P})\boldsymbol{\phi}_\mathbf{x} + \xi_2(\mathcal{P}), \\ \theta_{i^*} &= \theta_{i^*}(\mathcal{P}) = \boldsymbol{\xi}_1(\mathcal{P})\boldsymbol{\theta}_\mathbf{x} + \xi_3(\mathcal{P}), \\ \boldsymbol{\xi}_1 &= \boldsymbol{\xi}_1(\mathcal{P}) \\ \zeta &= \zeta(\mathcal{P}) \end{aligned} \tag{49}$$

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 (15)-(17) and (49), with $\mathcal{P} = \{\boldsymbol{\phi}_\mathbf{x}, \boldsymbol{\theta}_\mathbf{x}, \mu, \upsilon, \beta\}$.[17] Substituting stylized values for β, μ and υ as well as the ML parameters $\hat{\boldsymbol{\phi}}_\mathbf{x}$ and $\hat{\boldsymbol{\theta}}_\mathbf{x}$ from (42a) into (48) gives the Direct Control (DC) model used for impulse response and welfare analysis in section 5.

[17] We normalise the welfare weight to inflation to unity, so that μ and υ measure the welfare weight attached to output gap stabilisation and interest rate smoothing relative to inflation stabilisation.

References

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

Ball, L. (1992): "Why Does Inflation Raise Inflation Uncertainty," *Journal of Monetary Economics*, 29, 371–388.

Ball, L., and S. G. Cecchetti (1990): "Inflation and Uncertainty at Long and SHort Horizons," *Brookings Papers on Economic Activity*, 1, 215–254.

Benati, L. (2004): "Evolving post-World War II UK economic performance," *Journal of Money, Credit, and Banking*, 36, 691–718.

Benati, L., and P. Surico (2009): "VAR Analysis and the Great Moderation," *American Economic Review*, 99, 1636–1652.

Bernanke, B., and A. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82, 901–921.

Bernanke, B., and I. Mihov (1998): "Measuring Monetary Policy," *The Quarterly Journal of Economics*, 113, 869–902.

Bernanke, B. S. (2004): "The Great Moderation," *Proceedings of the Meetings of the Eastern Economic Association, Washington D.C.*

Bianchi, F., H. Mumtaz, and P. Surico (2009): "The great moderation of the term structure of UK interest rates," *Journal of Monetary Economics*, 56, 856–871.

Blanchard, O., and C. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48, 1305–1311.

Breusch, T., and A. Pagan (1979): "Simple test for heteroscedasticity and random coefficient variation," *Econometrica*, 47, 1287–1294.

Brunner, A., and G. Hess (1993): "Are Higher Levels of Inflation Less Predictable? A State Dependent Conditional Heteroscedasticity Approach," *Journal*

of Business and Economic Statistics, 11, 187–197.

Canova, F. (2007): *Methods for Applied Macroeconomic Research*. Princeton University Press, Princeton NJ.

Canova, F., L. Gambetti, and E. Pappa (2008): "The Structural Dynamics of U.S. Output and Inflation: What Explains the Changes?," *Journal of Money, Credit and Banking*, 40, 368–388.

Caporale, G. M., and A. Kontonikas (2009): "The Euro and inflation uncertainty in the European Monetary Union," *Journal of International Money and Finance*, 28, 954–971.

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.

Cogley, T., B. De Paoli, C. Matthes, K. Nikolow, and T. Yates (2011): "A Bayesian approach to optimal monetary policy with parameter and model uncertainty," *Bank of England working paper No. 414*.

Conrad, C., M. Karanasos, and N. Zeng (2010): "The link between macroeconomic performance and variability in the UK," *Economic Letters*, 106, 154–157.

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, 415–41.

Dennis, R. (2006): "The Policy Preferences of the US Federal Reserve," *Journal of Applied Econometrics*, 21(1), 55–77.

Dewachter, H., and M. Lyrio (2006): "Macro Factors and the Term Structure

of Interest Rates," *Journal of Money, Credit and Banking*, 38, 119–140.

Dothan, U. (1978): "On the Term Structure of Interest Rates," *Journal of Financial Economics*, 6, 59–69.

Ellison, M., and N. Valla (2001): "Learning, uncertainty and central bank activism in an economy with strategic interactions," *Journal of Monetary Economics*, 48, 153–171.

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.

Fountas, S., A. Ioannidis, and M. Karanasos (2004): "Inflation, inflation uncertainty and a common European monetary policy," *Manchester School*, 72, 221–241.

Fountas, S., and M. Karanasos (2007): "Inflation, output growth, and nominal and real uncertainty: Empirical evidence for the G7," *Journal of International Money and Finance*, 26, 229–250.

Fountas, S., M. Karanasos, and J. Kim (2002): "Inflation and output growth uncertainty and their relationship with inflation and output growth," *Economic Letters*, 75, 293–301.

Friedman, M. (1977): "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, 85, 451–472.

Gali, J., and M. Gertler (2007): "Macroeconomic Modelling for Monetary Policy Evaluation," *Journal of Economic Perspectives*, 21, 25–45.

Grier, K., and M. Perry (1996): "On inflation and inflation uncertainty in the G7 countries.," *Journal of International Money and Finance*, 17, 671–689.

Holland, A. S. (1993): "Comment on 'Inflation regimes and the Sources of Inflation Uncertainty'," *Journal of Money, Credit and Banking*, 25, 514–520.

——— (1995): "Inflation and Uncertainty: Tests for Temporal Ordering," *Journal of Money, Credit and Banking*, 27, 827–837.

Justiniano, A., and G. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98(3), 604–641.

King, M. (2003): "Speech to the East Midlands Development Agency," *Bank of England*, 14 October.

Lungqvist, L., and J. Sargent, T. (2004): *Recursive Macroeconomic Theory.* The MIT Press, Cambridge, Massachusetts.

Mumtaz, H., and P. Surico (2011): "Evolving international inflation dynamics: evidence from a time-varying dynamic factor model," *Journal of European Economic Association, forthcoming.*

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.

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 (2011): "Optimal Monetary Policy Using an Unrestricted VAR," *Journal of Applied Econometrics, Forthcoming.*

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy.," *Review of Economic Studies*, 72, 821–852.

Rotemberg, J., and M. Woodford (1997): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual*, 12, 297–346.

Rudebusch, G., and L. Svensson (1999): "Policy Rules for Inflation Targeting," in *Monetary Policy Rules*, ed. by J. Taylor, pp. 203–262, Cambridge, Massachusetts. National Bureau of Economic Research.

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.

Sims, C., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96, 54–81.

Soderlind, P. (1999): "Solution and estimation of RE macroeconomic models with optimal policy," *European Economic Review*, 43, 813–823.

Soderstrom, U. (2002): "Monetary Policy with uncertain parameters," *Scandinavian Journal of Economics*, 104, 125–145.

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 (2002): "Has the business cycle changed and why?," *NBER Macroeconomic Annuals 2002*, 17, 159–230.

Taylor, J. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

Ungar, M., and B. Zilberfarb (1993): "Inflation and its unpredictability - theory and empirical evidence," *Journal of Money, Credit, and Banking*, 25, 709–720.

Walsh, C. (2010): *Monetary Theory and Policy*. The MIT Press, Cambridge, Massachusetts.

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

B Tables and Figures

Table 1: Data summary statistics: 1979Q3-2010Q4

	Mean	Standard deviation	Skewness	Excess kurtosis	First order Autocorrel.	KPSS	ADF
g	-0.329	2.481	-0.467	0.219	0.975	0.267	-4.249
π	1.190	0.900	1.653	1.699	0.951	0.766	-1.755
i	1.941	0.962	0.335	-0.655	0.963	0.941	-1.176

Note: Output gap (g) is from OECD; CED inflation (π) and the base rate (i) are from Datastream. Mean denotes sample arithmetic mean expressed as percentage p.a.; $KPSS$ is the Kwiatowski et al (1992) statistic for the null hypothesis of level stationarity and ADF is the Adjusted Dickey-Fuller statistic for the null of non-stationarity. The 5% significance levels are 0.463 and (-)2.877 respectively.

Table 2: Baseline estimates

Model		**ML0**	**ML01**	**ML02**	**ML03**
Loglikelihood value		(-) 122.7	(-) 70.16	(-) 97.281	(-) 63.09
Number of Σ Parameters		6	9	9	12
Initial estimates					
Δ_0	$\delta_{g,0}$	0.4784	0.4953	0.301 4	0.4131
	$\delta_{\pi,0}$	1.0893	1.0100	0.4946	0.6056
	$\delta_{i,0}$	0.6551	0.6905	0..6547	0.7420
Δ_1	$\delta_{g,1}$		0.2643		0.2403
	$\delta_{\pi,1}$		0.3184		0.3283
	$\delta_{i,1}$		0.3967		0.4486
Δ_2	$\delta_{g,2}$			0.0788	0.0257
	$\delta_{\pi,2}$			0.3089	0.2943
	$\delta_{i,2}$			2.25×10^{-6}	1.26×10^{-2}
G	g_1	-0.5011	-0.4035	-0.2254	-0.1909
	g_2	0.5942	0.5052	0.6372	0.5272
	g_3	0.1559	0.1303	0.1559	0.1270

Note: Model ML0 is the baseline homoskedastic model. This provides a set of baseline parameters for the dynamic model (1). In the ML0X approach, the likelihood is then optimized with respect to the parameters of the stochastic structure (4) keeping these baseline parameters fixed. ML01 assumes that the error variances are linear in the (lagged) annual rate of inflation while ML02 assumes that the variances are quadratic in this rate. The encompassing model ML03 includes both linear and quadratic effects. The fully optimized results are reported in the next two tables..

Table 3: Loglikelihood ratio tests

Model	**ML0**	**ML1**	**ML2**	**ML3**
Loglikelihood value (logL)	-122.70	- 63.24	- 96.81	- 56.14
Number of Parameters (k)	23	26	26	29
LR test (against M3)	133.12 $\sim \chi^2(6)$	14.20 $\sim \chi^2(3)$	81.34 $\sim \chi^2(3)$	
SCA test (against M3)	- 52.06	- 0.16	- 33.40	

Note: In the ML approach, the interest rate equation (21) is estimated alongside the structural equations (20) using maximum likelihood. 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 terms. Its likelihood value (L_u) is compared with that of each restricted model (L_r) using a loglikelihood ratio test LR=($2\times(\ln L_u - \ln L_r)$), which has the 1% critical values $\chi^2(3)$=11.35 and $\chi^2(6)$=15.09. All three restricted models are rejected on this test. The Schwarz statistic (SCA=($\ln L_u - \ln L_r$)-0.5$\times$($k_u - k_r$)$\times \ln(T)$) guards against over-fitting and provides an asymptotically consistent test. On this criterion, ML0 and ML2 are decisively rejected against ML3, while the performance of ML1 and ML3 are similar.

Table 4: Parameter estimates for ML models

	ML0		**ML1**		**ML2**		**ML3**	
	par.	**t-stat**	**par.**	**t-stat**	**par.**	**t-stat.**	**par.**	**t-stat.**
g_1	-0.5011	3.91	-0.4700	3.10	-0.2647	3.00	-0.2617	2.38
g_2	0.5942	5.65	0.4353	2.90	0.6311	3.77	0.4551	4.09
g_3	0.1559	1.62	0.1502	2.80	0.1559	3.01	0.1401	3.85
a_1	1.2930	16.67	1.4636	19.06	1.3785	16.90	1.4488	20.17
a_2	-0.3495	4.41	-0.5047	6.11	-0.4204	5.07	0.4812	-6.29
a_3	-0.0236	1.03	-0.0108	0.58	-0.0236	1.12	-0.0236	1.12
b_1	0.3588	4.29	0.4213	4.82	0.3004	3.48	0.3666	4.22
b_2	0.2887	3.38	0.2232	2.40	0.2320	2.61	0.2169	2.39
b_3	0.1491	1.76	0.1044	1.17	0.1448	1.65	0.1176	1.31
b_5	0.0972	2.53	0.0180	0.61	0.0793	2.55	0.0419	1.32
θ_i	0.2602	2.34	0.2251	2.50	0.3150	2.83	0.1986	1.92
c_1	0.2584	3.72	0.1955	3.21	0.2493	3.59	0.1736	2.56
c_2	0.7941	9.11	0.9137	11.17	0.7941	9.11	0.9073	10.69
c_3	-0.1423	1.36	-0.1169	1.40	-0.1900	1.82	-0.0902	0.85
c_4	0.1204	1.69	0.0010	0.02	0.1116	1.56	0.0010	0.01
c_5	0.0863	0.78	-0.0192	-0.17	0.0863	0.78	-0.0163	0.15
c_6	0.0268	0.37	0.0010	0.01	0.0261	0.36	0.0010	0.01
c_7	-0.0476	0.45	0.0521	0.36	-0.0476	0.45	-0.0424	0.41
c_8	0.0119	0.15	0.0744	1.36	0.0306	0.40	0.0777	1.20
c_9	0.0179	0.23	0.0687	0.95	0.0179	0.23	0.0666	0.96
$\delta_{g,0}$	0.4784	15.95	0.5007	13.28	0.2894	12.47	0.4088	13.28
$\delta_{\pi,0}$	1.0893	11.50	0.9986	9.99	0.4674	10.91	0.6237	9.70
$\delta_{i,0}$	0.6553	14.95	0.6462	14.10	0.6547	15.95	0.7112	15.87
$\delta_{g,1}$			0.2792	10.01			0.2715	8.67
$\delta_{\pi,1}$			0.3718	5.05			0.3754	5.05
$\delta_{i,1}$			0.3704	9.09			0.4307	9.99
$\delta_{g,2}$					9.25×10^{-2}	3.66	3.94×10^{-2}	2.35
$\delta_{\pi,2}$					0.3174	4.00	0.2729	4.94
$\delta_{i,2}$					4.91×10^{-7}	0.01	1.23×10^{-2}	0.66

Note: The model parameters are defined in (20), (21) and (38) and are estimated using maximum likelihood (appendix 3).

Table 5: Long run responses of estimated and optimal policy rules

	ML0			**DC0**		
		μ=υ=1	μ=1,υ=0.5	μ=0.2,υ=0.5	μ=5,υ=0.5	μ=1,υ=0.1
$\bar{\kappa}$	(0)	(0)	(0)	(0)	(0)	(0)
$\bar{\phi}_g$	0.79	2.50	2.91	2.54	3.87	4.16
$\bar{\phi}_\pi$	2.80	2.89	3.13	5.06	1.73	3.79
$\bar{\pi}$	(0)	(0)	(0)	(0)	(0)	(0)
	ML01			**DC01**		
$\bar{\kappa}$	0.00	3.39	3.58	3.39	3.77	4.01
$\bar{\phi}_g$	0.79	2.50	2.91	2.54	3.87	4.16
$\bar{\phi}_\pi$	2.80	2.89	3.13	5.06	1.73	3.79
$\bar{\pi}$	(0)	-1.79	-1.68	-0.83	-5.16	-1.44
	ML02			**DC02**		
$\bar{\kappa}$	0.00	0.00	0.00	0.00	0.00	0.00
$\bar{\phi}_g$	0.79	3.23	3.63	3.23	4.57	4.85
$\bar{\phi}_\pi$	2.80	6.77	7.24	8.16	6.63	8.55
$\bar{\pi}$	(0)	(0)	(0)	(0)	(0)	(0)
	ML03			**DC03**		
$\bar{\kappa}$	0.00	3.53	3.73	3.62	3.91	4.23
$\bar{\phi}_g$	0.79	3.00	3.41	3.07	4.29	4.65
$\bar{\phi}_\pi$	2.80	5.59	6.02	7.44	4.68	7.22
$\bar{\pi}$	(0)	-0.77	-0.74	-0.56	-1.06	-0.68

Note: The ML0X models are described in the footnote to table 2 and the DC0X models replace the empirical interest rule with the appropriate optimal rule. The long run interest rate rule takes the form $i = \bar{\kappa} + \bar{\phi}_g g + \bar{\phi}_\pi \pi$, where the long run coefficients shown in (40) and (41) are computed using the coefficients of either the estimated or the optimal policy rule. The intercept is zero in the MLX models and in DC0 since the data is de-meaned. Heteroskedasticity has the effect of inducing a positive $\bar{\kappa}$ intercept, which reduces the steady state inflation rate by $\bar{\pi} = \bar{\kappa}/(1 - \bar{\phi}_\pi)$ in DC1 and DC3.

Table 6: Volatilities and welfare losses implied by the empirical rules under different welfare specifications

	Standard deviations			**Losses**				
	Goal variables			Welfare weights				
	g_t	π_t	Δi_t	Case: 1	2	3	4	5
				$\mu = 1$	$\mu = 1$	$\mu = 0.2$	$\mu = 5$	$\mu = 1$
				$\mu = 1$	$\nu = 0.5$	$\nu = 0.5$	$\nu = 0.5$	$\nu = 0.2$
Model								
ML0	2.56	2.29	1.00	12.81	12.31	7.02	38.26	11.91
ML01	2.43	2.01	0.98	10.88	10.40	5.67	33.62	10.02
ML02	3.54	2.90	1.22	22.40	21.66	11.57	71.16	21.06
ML03	2.69	2.60	1.22	15.49	14.75	8.94	43.28	14.15

Note: This table shows the results obtained by simulating the ML0X models with the benchmark ML0 empirical policy rule (and state equations) and the stochastic parameters shown in table 2. The first three columns of numbers show the standard deviations of the three goal variables and the remaining columns show the welfare losses implied by the five welfare specifications used in table 5. The losses drop as we move from case 1 to 3 reducing respectively the interest rate smoothing and output weights. Macro volatility is lower in ML1 than ML0 because sequences of shocks that lower inflation have the effect of attenuating the effect of later shocks. On the other hand, the quadratic amplification effect in ML02 increases volatility and welfare losses, while ML03 combines both linear attenuation and quadratic amplification effects.

Table 7: Welfare gains from policy optimization

Empirical rules		**Optimal rules**					
Model	Welfare loss	Model	Standard deviations g_t	π_t	Δi_t	Welfare loss	% gain
		Case 1, $\mu = \nu = 1$					
ML0	12.81	DC0	2.22	1.95	0.89	9.52	0.26
ML01	10.88	DC01	1.91	1.22	0.66	5.59	0.49
ML02	22.40	DC02	2.72	1.72	1.99	14.29	0.36
ML03	15.49	DC03	1.56	0.79	0.83	3.75	0.76
		Case 2, $\mu = 1, \nu = 0.5$					
ML0	12.31	DC0	2.15	1.87	1.21	8.85	0.28
ML01	10.40	DC01	1.80	1.13	0.89	4.91	0.53
ML02	21.66	DC02	2.58	1.61	2.39	12.08	0.44
ML03	14.75	DC03	1.48	0.72	1.04	3.24	0.78
		Case 3, $\mu = 0.2, \nu = 0.5$					
ML0	7.02	DC0	2.63	1.88	1.08	5.49	0.22
ML01	5.67	DC01	1.96	1.12	0.64	2.22	0.61
ML02	11.57	DC02	2.97	1.68	2.13	6.87	0.41
ML03	8.94	DC03	1.74	0.82	0.94	1.72	0.81
		Case 4: $\mu = 5, \nu = 0.5$					
ML0	38.26	DC0	1.82	1.98	2.12	22.64	0.41
ML01	33.62	DC01	1.87	1.37	2.09	21.49	0.36
ML02	71.16	DC02	2.21	1.54	3.45	32.80	0.54
ML03	43.28	DC03	1.18	0.61	1.55	8.54	0.80
		Case 5: $\mu = 1, \nu = 0.2$					
ML0	11.91	DC0	1.98	1.72	2.51	7.50	0.37
ML01	10.02	DC01	1.55	1.00	1.77	3.71	0.63
ML02	21.06	DC02	2.29	1.41	3.91	8.76	0.58
ML03	14.15	DC03	1.32	0.63	1.87	2.48	0.82

Note: This table compares the results obtained by simulating the three models under the optimal policy (the DC0X models) with the empirial (ML0X) models shown in the previous table. The welfare specifications affect the policy rules and hence the volatilities of the goal variables. Optimization reduces the welfare loss relative to the empirical rules, with the percentage welfare gains being shown in the final column.

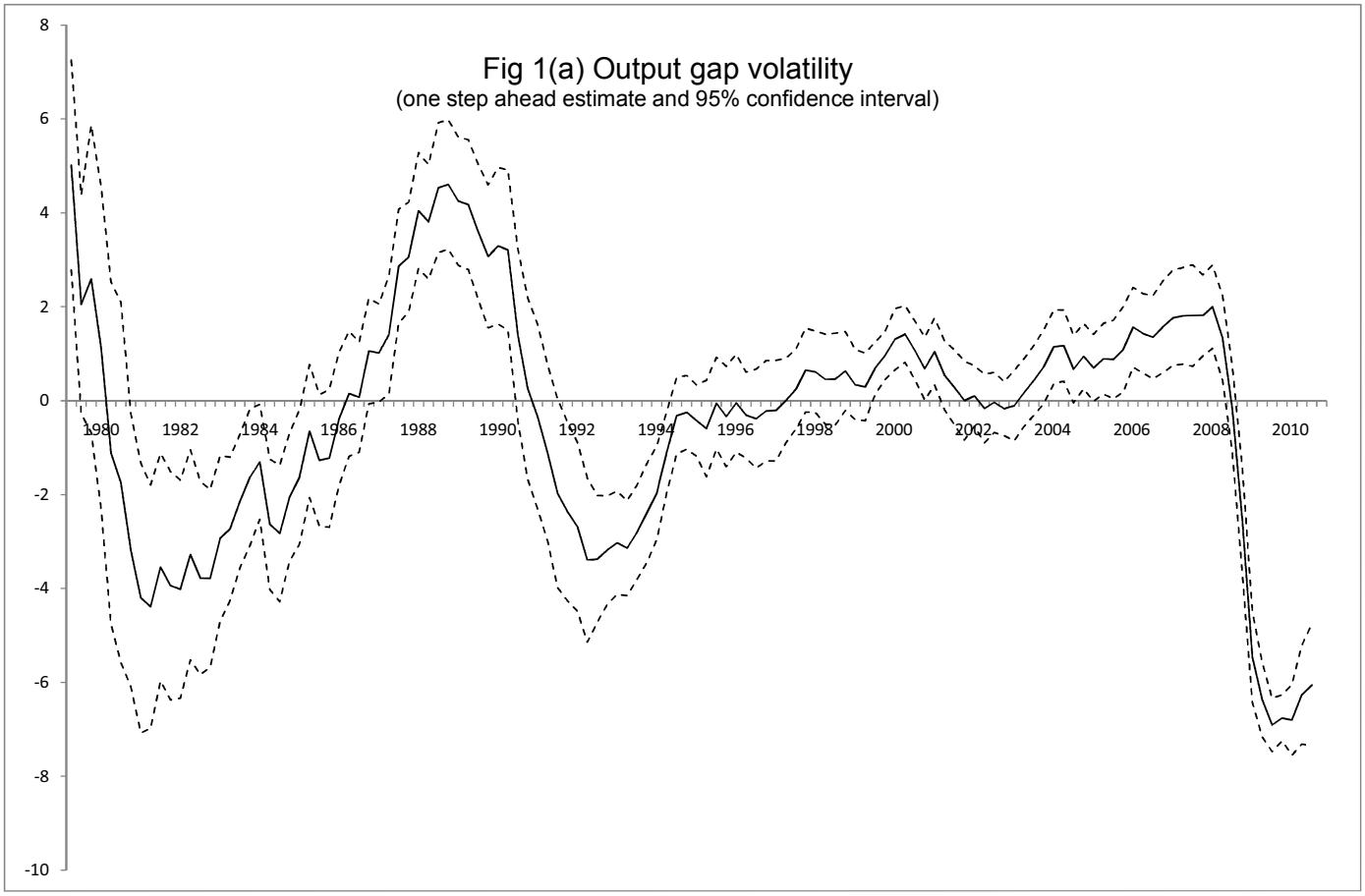
Fig 1(a) Output gap volatility
(one step ahead estimate and 95% confidence interval)
8
6
4
2
0
-2
-4
-6
-8
-10
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010

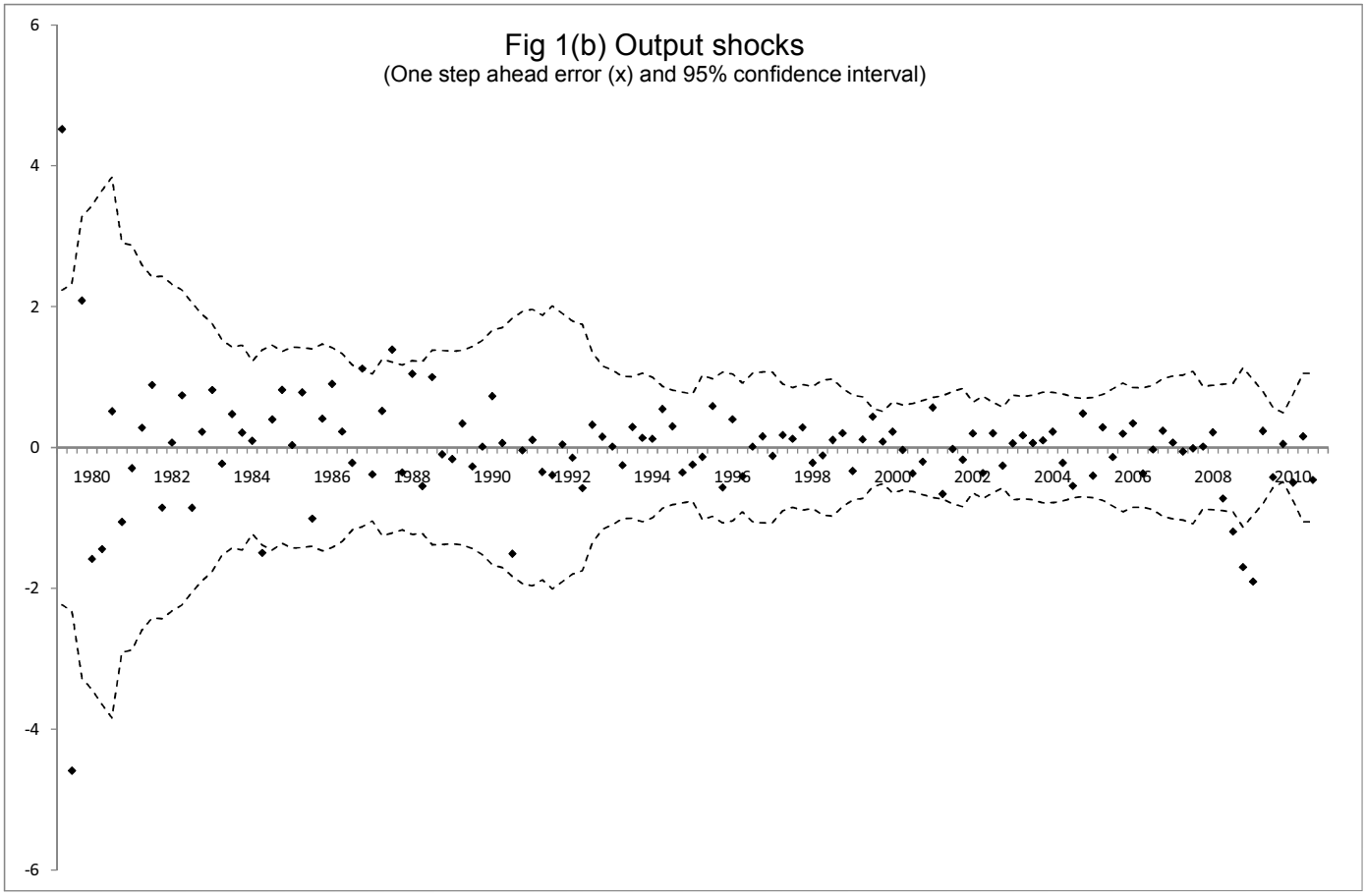
Fig 1(b) Output shocks
(One step ahead error (x) and 95% confidence interval)
6
4
2
0
-2
-4
-6
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010

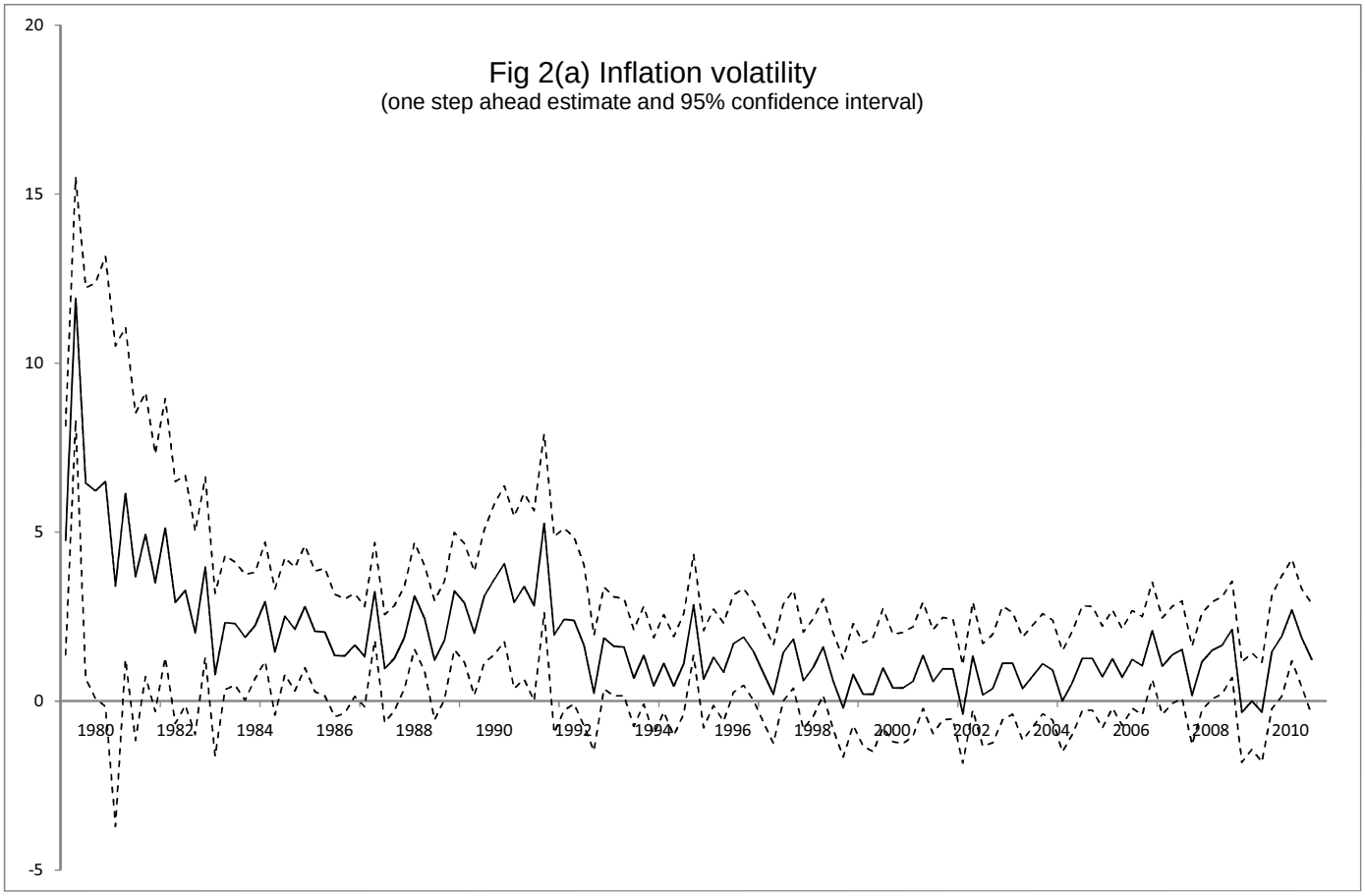
Fig 2(a) Inflation volatility
(one step ahead estimate and 95% confidence interval)
20
15
10
5
0
-5
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010

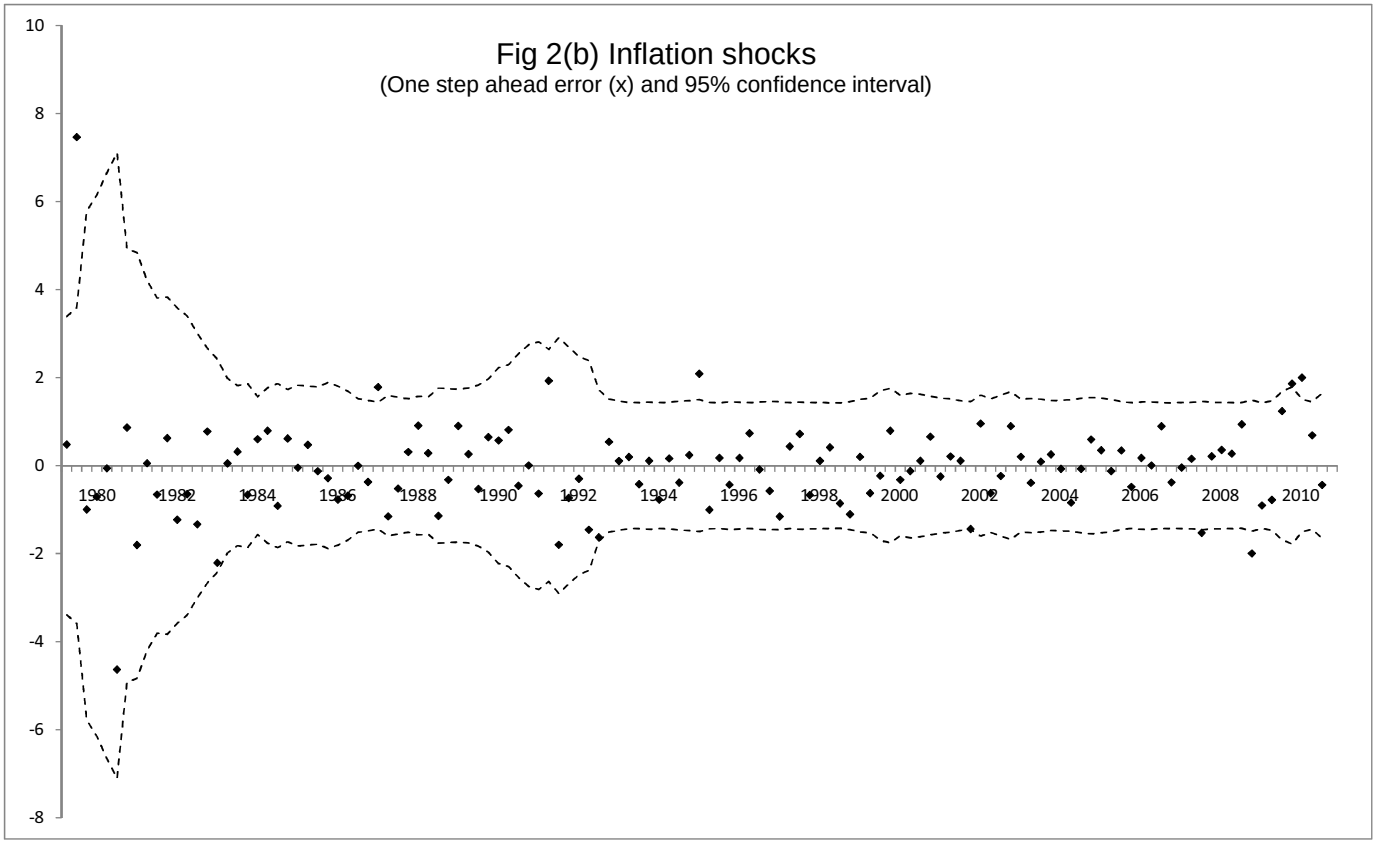
Fig 2(b) Inflation shocks
(One step ahead error (x) and 95% confidence interval)
10
8
6
4
2
0
-2
-4
-6
-8
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010

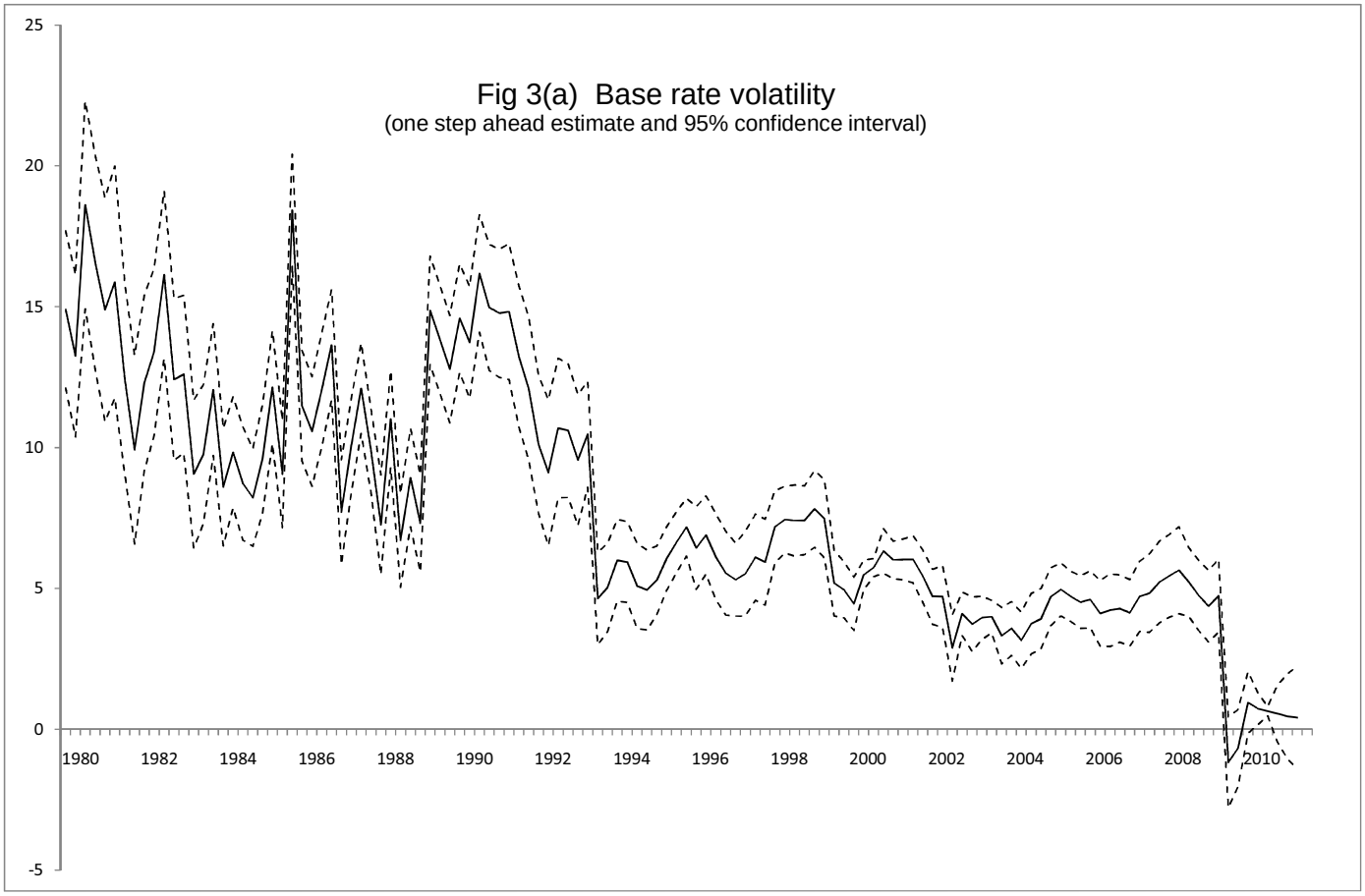
Fig 3(a) Base rate volatility
(one step ahead estimate and 95% confidence interval)
25
20
15
10
5
0
-5
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010

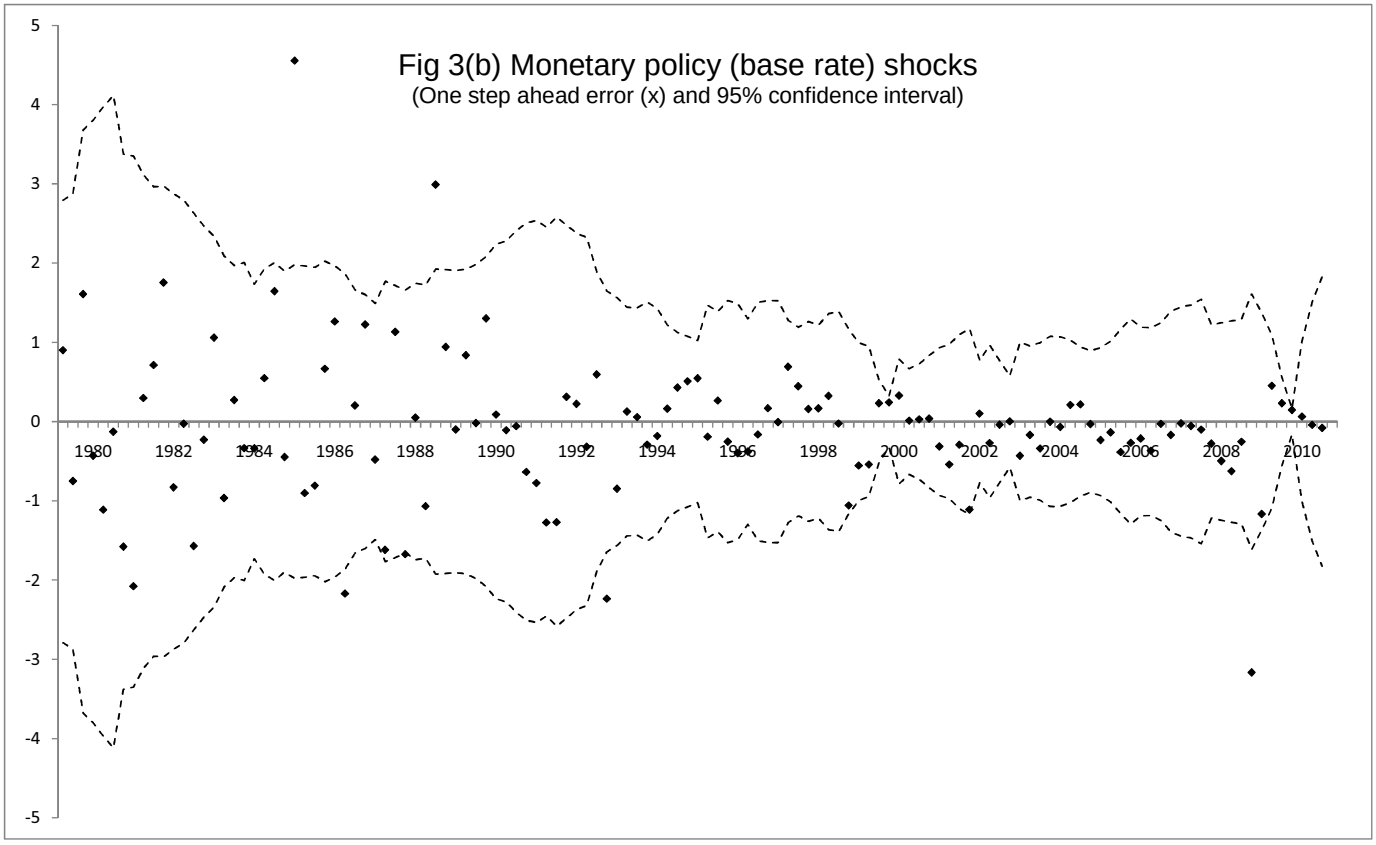
Fig 3(b) Monetary policy (base rate) shocks
(One step ahead error (x) and 95% confidence interval)
5
4
3
2
1
0
-1
-2
-3
-4
-5
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010

Fig 4. Impulse response functions for the ML models (response to one period shocks)

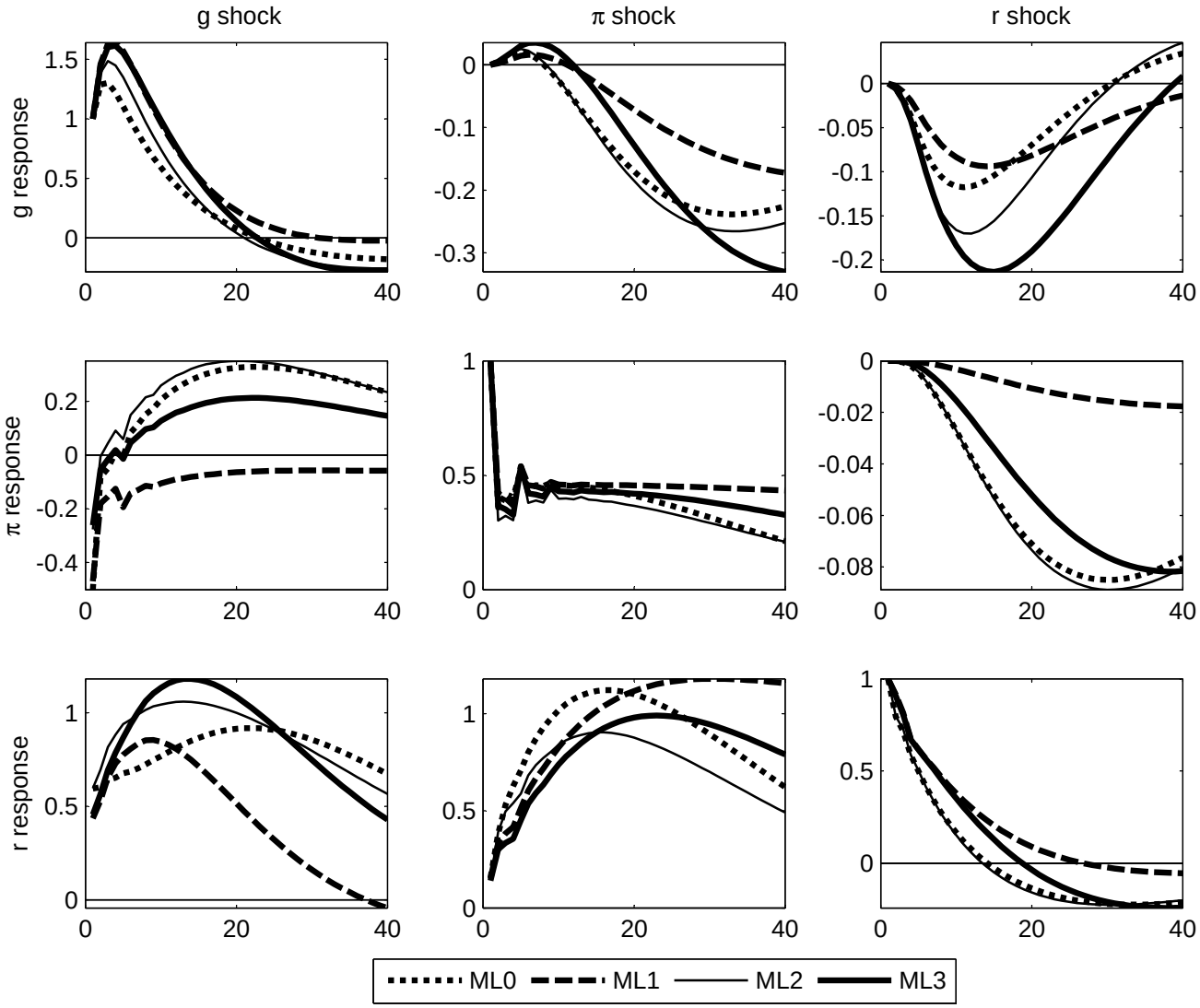

Fig 5. Responses to a sequence (20-quarters) of positve and negative inflation shocks

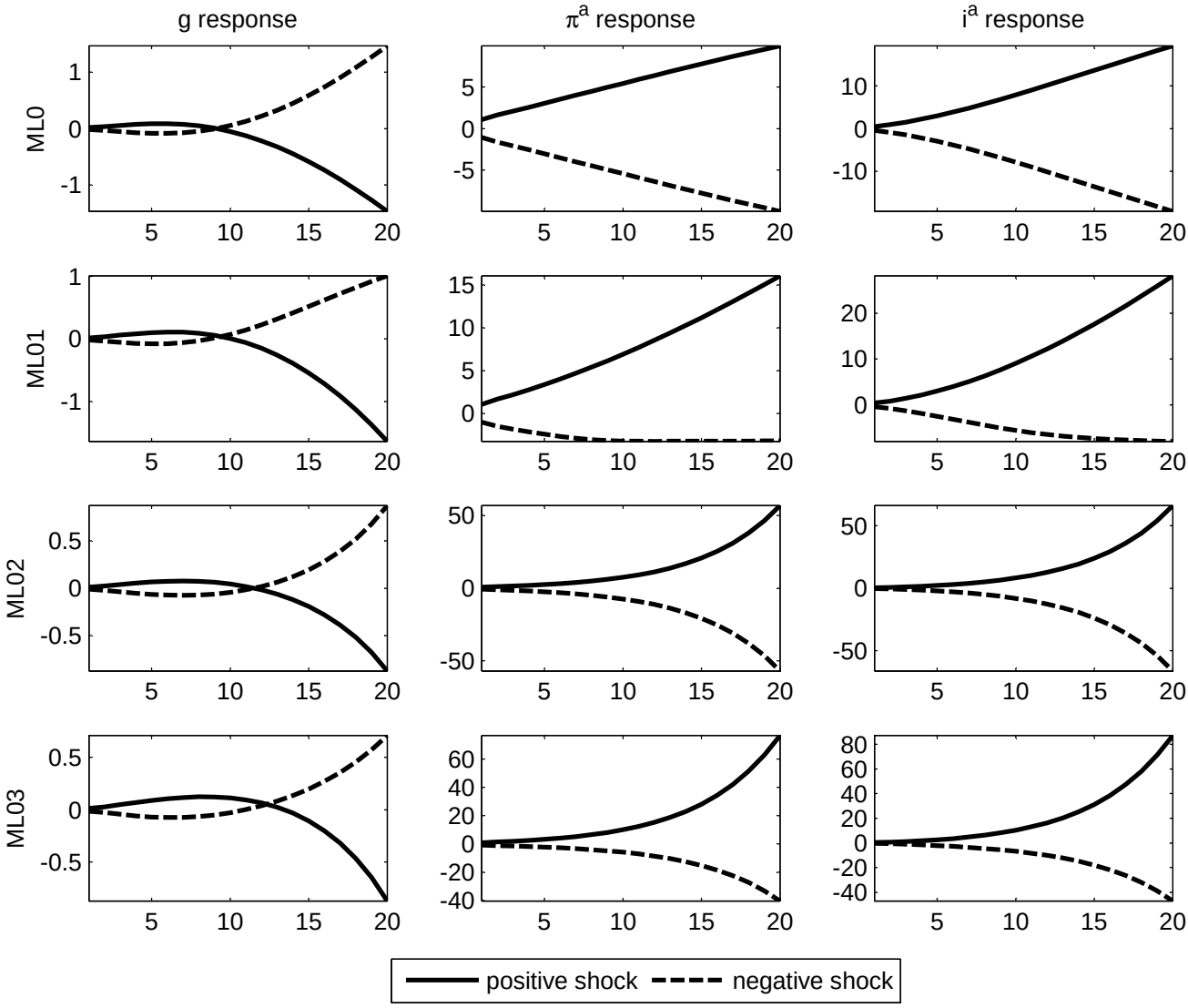

Fig 6. Impulse response functions for the DC0X models (response to one period shocks)