C.I.R.P.É.E

Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **10-27**

Optimal Policy Restrictions on Observable Outcomes

Federico Ravenna

Septembre/September 2010

Ravenna : HEC Montréal, University of California-Santa Cruz and CIRPÉE
fravenna@ucsc.edu

Earlier drafts of this paper were circulated under the title: "How to distinguish bad policy from bad luck". I would like to thank Bart Hobijn, Oscar Jorda, Andre Kurmann, Andreas Schabert, Luca Sala, Ulf Soderstrom, Peter Tillmann, Mathias Trabandt, Carl Walsh for very helpful comments and suggestions, and Daniel Beltran for excellent research assistance.

Abstract:
We study the restrictions implied by optimal policy DSGE models for the volatility of observable endogenous variables. Our approach uses a parametric family of singular models to discriminate which volatility sample outcomes have zero probability of being generated by an optimal policy. Thus the set of volatility outcomes generated by the model is not of measure zero even if there are no random deviations from optimal policymaking. This methodology is applied to a new Keynesian business cycle model widely used in the optimal monetary policy literature, and its implications for the assessment of US monetary policy performance over the 1984-2005 period are discussed.

Keywords: Optimal monetary policy, business cycle, DSGE model, policy performance

JEL Classification: E30

1 Introduction

The business cycle theory that has become prevalent in the last two decades assumes that business cycle volatility is the result of exogenous shocks. Fiscal and monetary policy can affect the propagation of these shocks throughout the economy, and the resulting volatility in aggregate economic variables.

A central question in assessing the historical performance of monetary and fiscal policies is how to distinguish the amount of economic volatility that is an efficient outcome given the shocks driving the business cycle - that is, the volatility that would obtain conditional on the optimal policy - and the volatility resulting from suboptimal policymaking. Because exogenous shocks are typically unobservable, any assessment of the policy performance must rely on the restrictions implied by a DSGE model for the co-movement of observable endogenous variables.

This paper investigates the restrictions implied by optimal policy DSGE models for the volatility of observable endogenous variables. Optimal policy DSGE models are by construction singular - they predict the time series for one variable is a nonstochastic function of other variables' time series. Unless random deviations from optimal policy are introduced, the data will reject the restrictions of optimal policy models almost surely. We propose a way to use singular models to define a set of outcomes with nonzero probability, in terms of observable variables' volatilities. While this set of outcomes, which we label the *optimal policy space*, can be used as a diagnostic tool to distinguish from historical outcomes bad policies from bad luck, we rather use it as a tool to understand the restrictions implied by optimal policymaking in DSGE models.

A DSGE model defines a map M between the shocks vector U_t covariance matrix Σ_U and the endogenous variables vector Y_t covariance matrix Σ_Y. Typically, the map M implies that any volatility sample outcome has a nonzero probability of being generated by the model.

This is the consequence of two assumptions macroeconomists often make. First, business cycle models are solved using a linear approximation, resulting in equilibrium law of motion of the form, at its simplest, $Y_t = \mathbf{A}U_t$. Second, the linear solution is assumed nonsingular by ensuring that the number of exogenous shocks and observable endogenous variables are identical. In optimal policy models, this implies including a random shock in the policy optimality condition. Then, regardless of the restrictions imposed by optimal policymaking on the model $\mathbf{A}$, any outcome Y_t can be explained by some random vector U_t, since for any given nonsingular model and covariance outcome Σ_Y it holds $\Sigma_U = \mathbf{A}^{-1}\Sigma_Y\mathbf{A}'^{-1}$.

Rather than building the map M as a linear function of Σ_U for a nonsingular model with random deviations from the optimal policy, we include in the argument β of the map M also some of the deep parameters of the DSGE model, and build the map $M(\beta)$ for a parametric family of optimal-policy singular models. Therefore, the set of volatility outcomes generated by optimal policymaking - the image of $M(\beta)$ - is not of measure zero. At the same time, the nonlinearity of the map implies that there may exist volatility outcomes with zero probability.

We use this approach to show how truly optimal policy would restrict the volatility outcome for observable variables in a widely used monetary business cycle model. Based on this model, the 1985-2004 sample observation for US macroeconomic variables would have zero probability of being generated by optimal policymaking. Given our methodology can only identify the set of sample outcomes with zero probability, but cannot determine the likelihood that an outcome belonging to the optimal policy space was in fact generated by optimal policymaking, we interpret this result as evidence that popular models used to provide monetary policy prescriptions impose tighter restrictions on the behaviour of the economy than is readily apparent. Intuitively, alternative models belonging to a parametric family may imply a very different mapping between the volatility of exogenous shocks and endogenous variables - and very different impulse responses conditional on a one standard deviation exogenous shock. Yet the same models may be unable to generate very different sets of *unconditional* volatility outcomes. This is indeed the case for the parametric family of DSGE models we examine.

The paper is organized as follows. Section 2 defines the optimal policy space. Section 3 introduces a simple example to illustrate the restrictions on the volatility outcomes imposed by the optimal policy space, and evaluates the US policy performance. Section 4 discusses related literature and section 5 concludes.

2 The Optimal Policy Space

Let the map $M(\beta)$ associate to any DSGE model parameter space the set of the endogenous variables' volatility outcomes. The image of $M(\beta)$ conditional on the optimal policy is a set no larger than the image conditional on all possible policies. In general, $M(\beta)$ is defined as the map between the model's parameters and all the entries in the covariance matrix Σ_Y. In the following we specialize $M(\beta)$ to map into the main diagonal of Σ_Y only. This assumption is without loss of generality, and allows a useful graphical representation of the image of $M(\beta)$. It is convenient to start with some formal definitions.

Definition 1 *Let β be a vector of parameters, $\mathbf{p}$ a policy rule and $Z(\beta;\mathbf{p})$ a law of motion for n endogenous variables conditional on policy $\mathbf{p}$. Let the vector-valued function $M(\beta;\mathbf{p}) : D \subseteq \mathbf{R}^r \to \mathbf{R}^n$*

*associated with $Z(\beta;\mathbf{p})$ map every vector $\beta \in \mathbf{R}^r$ to a unique vector of variances for the n endogenous variables. Define the set $V_{\mathbf{p}}$ as the image of $M(\beta;\mathbf{p})$. The set $V_{\mathbf{p}}$ is called the **volatility space** for model Z conditional on policy $\mathbf{p}$ and parameter vector β.*

Definition 2 *Define the set $V_{\mathbf{o}}$ as the volatility space $V_{\mathbf{p}}$ associated with $M(\beta;\mathbf{o})$ conditional on the optimal policy $\mathbf{p} = \mathbf{o}$. The set $V_{\mathbf{o}}$ is called the **optimal policy space**.*

In most business cycle models, for an appropriate choice of n it holds that $V_{\mathbf{o}} \subseteq \mathbf{R}^n$ and $V_{\mathbf{o}} \subsetneq \mathbf{R}^{n-1}$ for $n > 1$, so that $V_{\mathbf{o}}$ is a non-trivial $n-$dimension subset of $\mathbf{R}^n$. In this case $V_{\mathbf{o}}$ describes a set of volatility outcomes $(\sigma^2_{Y_1}, ..., \sigma^2_{Y_n})$ which is a proper subset of the volatility space, and which is not of measure zero. In the following, we will say that an optimal policy imposes 'tight restrictions' on $V_{\mathbf{o}}$ if for any given $(\sigma^2_{Y_1}, ..., \sigma^2_{Y_{n-1}})$ belonging to the optimal policy space for the variables $(Y_1, ..., Y_{n-1})$, the range of values for $\sigma^2_{Y_n}$ belonging to $V_{\mathbf{o}}$ is bounded.

2.1 The Linear Case

Assume $M(\beta;\mathbf{o})$ is a linear map and is equal to:

$$M(\beta;\mathbf{o}) = \mathbf{C}\beta \tag{1}$$

where β is an $r \times 1$ vector and $\mathbf{C}$ is an $n \times r$ matrix. For an unrestricted vector β two outcomes are possible. When the matrix $\mathbf{C}$ is of rank n its columns span the space $\mathbf{R}^n$. Then $V_{\mathbf{o}} = \mathbf{R}^n$ and necessarily $V_{\mathbf{o}} = V_{\mathbf{p}}$ for any policy $\mathbf{p}$ such that $rank(\mathbf{C}) = n$. When $\mathbf{C}$ is of rank $s < n$ its columns span the subspace $\mathbf{R}^s$ and $V_{\mathbf{o}}$ is a s-dimension hyperplane.

For a linear model and β including only the entries for the exogenous shocks' covariance matrix the map $M(\beta;\mathbf{o})$ can be written as in eq. (1). Let the model associated with $M(\beta;\mathbf{o})$ be described by the stationary law of motion $Y_t = \mathbf{A}U_t$ where Y_t is an $n \times 1$ vector of endogenous variables with covariance matrix Σ_Y and U_t is an $m \times 1$ vector of exogenous shocks with covariance matrix Σ_U. For $\beta \equiv vec(\Sigma_U)$ we can write

$$M(\beta;\mathbf{o}) = \mathbf{T}\left(\mathbf{A} \otimes \mathbf{A}\right) vec(\Sigma_U) \tag{2}$$

where $\mathbf{T}$ is an $n \times nn$ matrix with unitary value at entry $[i, (i-1)n + i]_{i=1}^n$ and zero otherwise, so that $M(\beta;\mathbf{o})$ is equal to the diagonal of Σ_Y. If $\mathbf{A}$ is of rank n the linear map $vec(\Sigma_Y) = (\mathbf{A} \otimes \mathbf{A})\, vec(\Sigma_U)$ spans the space defined by the vectorization of $n \times n$ positive semi-definite symmetric matrices, and the matrix $\mathbf{T}\left(\mathbf{A} \otimes \mathbf{A}\right)$ is of rank n. Because Σ_U is a positive semi-definite symmetric matrix, $M(\beta;\mathbf{o})$ does not span $\mathbf{R}^n$. It will though span $\mathbf{R}^{n^+}$, since $M(\beta;\mathbf{o})$ is just the main diagonal of Σ_Y, and any vector $g \in R^{n^+}$ is the main diagonal of at least one positive semi-definite matrix. If $\mathbf{A}$ is of rank $s < n$, also $\mathbf{T}(\mathbf{A} \otimes \mathbf{A})$ is of rank $s < n$. This is the case of a singular model, where $V_{\mathbf{o}}$ is a s-dimension hyperplane in $\mathbf{R}^n$. Therefore, conditional on the model $\mathbf{A}$ either *all* vectors $[\sigma^2_{Y_1}, \sigma^2_{Y_2}, ..., \sigma^2_{Y_n}]'$ belong to the optimal policy space (and $V_{\mathbf{o}}$ is an improper subset of $\mathbf{R}^{n^+}$) if $s = n$, or any vector $[\sigma^2_{Y_1}, \sigma^2_{Y_2}, ..., \sigma^2_{Y_n}]'$ almost surely does not belong to the optimal policy space if $s < n$.

2.2 The General Case

Assume any model parameter k is allowed to belong to the domain of M so that $\beta = [vec(\Sigma_U), k_1, ..., k_h]'$ and $M(\beta;\mathbf{o})$ is a nonlinear vector-valued function $M : D \subseteq \mathbf{R}^r \rightarrow \mathbf{R}^n$. Recall that for $M(\beta;\mathbf{o}) = \mathbf{C}\beta$ and β unrestricted only two outcomes are possible in the linear case: either $V_{\mathbf{o}} = \mathbf{R}^n$, or $V_{\mathbf{o}}$ is a lower-dimension hyperplane. When $M(\beta;\mathbf{o})$ is a nonlinear function, it is possible for $V_{\mathbf{o}}$ to be a proper subset of $\mathbf{R}^n$ and *at the same time* not to be contained in any lower-dimension subspace, even if the associated $Z(\beta;\mathbf{o})$ model's law of motion is described by the linear map $Y_t = \mathbf{A}U_t$ and $\mathbf{A}$ is of rank $s < n$. This property ensures that in general $V_{\mathbf{o}}$ is a non-trivial subset of $\mathbf{R}^n$. Effectively, verifying whether an outcome $(\sigma^2_{Y_1}, ..., \sigma^2_{Y_n})$ is optimal amounts to checking whether a vector $[\sigma^2_{Y_1}, \sigma^2_{Y_2}, ..., \sigma^2_{Y_n}]'$ belongs to the image of the function $M(\beta;\mathbf{o})$. If $M(\beta;\mathbf{o})$ were bijective this could be established by checking whether the value of the inverse function $M^{-1}(\sigma^2_{Y_1}, \sigma^2_{Y_2}, ..., \sigma^2_{Y_n})$ for a given outcome belongs to the domain D of $M(\beta;\mathbf{o})$. Since $M(\beta;\mathbf{o})$ is generally surjective but not injective, its inverse must be computed employing numerical methods.

Notice that if $\beta = vec(\Sigma_U)$ and the model is singular, any outcome $(\sigma^2_{Y_1}, ..., \sigma^2_{Y_n})$ does not belong to $V_{\mathbf{o}}$ almost surely, whereas if the model is nonsingular any outcome $(\sigma^2_{Y_1}, ..., \sigma^2_{Y_n})$ belongs to $V_{\mathbf{o}}$ with probability one. By including in β behavioral parameters in addition to the entries in Σ_U, the set $V_{\mathbf{o}}$ of a singular model can be of nonzero measure in $\mathbf{R}^n$ - intuitively, the nonlinearity of the mapping $M(\beta;\mathbf{o})$ allows $V_{\mathbf{o}}$ to be "large" or "small" with respect to $\mathbf{R}^n$.

In the linear case we saw that when $rank(\mathbf{C}) = s < n$ (as will happen whenever $rank(\mathbf{A}) = s < n$) $V_{\mathbf{o}}$ is a s-dimension hyperplane, implying $M(\beta;\mathbf{o})$ can be rewritten as a map between vectors in $\mathbf{R}^s$ and vectors in $\mathbf{R}^n$ even if the domain of $M(\beta;\mathbf{o})$ is $\mathbf{R}^r$, $r > s$. A similar notion can be extended to the case when $M(\beta;\mathbf{o})$ is nonlinear using the following definitions (Baxandall and Liebeck, 1986):

Definition 3 *A function $M : S \subseteq \mathbf{R}^s \rightarrow \mathbf{R}^n$ is smooth if it is a C^1 function and if for all $\mathbf{g} \in S$ the Jacobian $J_{M,\mathbf{g}}$ is of maximum possible rank* $\min(s, n)$.

Definition 4 *A subset $K \subseteq \mathbf{R}^n$ is called a smooth $s - surface$ if there is a region of S in $\mathbf{R}^s$ and a smooth function $\rho : S \subseteq \mathbf{R}^s \to \mathbf{R}^n$ such that $\rho(S) = K$.*

The latter definition implies that if a smooth $\rho(S)$ exists the image K of $M(\beta; \mathbf{o}) : D \subseteq \mathbf{R}^r \to \mathbf{R}^n$ can be parametrically described by a vector-valued function ρ of s variables. The smoothness condition on ρ means that the Jacobian matrix of ρ at any point in the domain has at least s independent column vectors. When for all $\mathbf{g} \in D$ it holds that $rank(J_{M,\mathbf{g}}) = n$, then for $S = D$ the function $\rho(S) \equiv M(\beta; \mathbf{o})$ maps into a smooth $n - surface$ and the probability that $[\sigma^2_{Y_1}, \sigma^2_{Y_2}, ..., \sigma^2_{Y_n}]' \in V_{\mathbf{o}} = K$ is non-zero. On the contrary, when $rank(J_{M,\mathbf{g}}) = s < n$ the function $M(\beta; \mathbf{o})$ cannot describe a smooth $n - surface$ in $\mathbf{R}^n$ and the image K will be a smooth $s - surface$ described by $\rho : S \subseteq \mathbf{R}^s \to \mathbf{R}^n$. The constant rank theorem (Conlon, 2001) ensures existence of $\rho(S)$. In this case any given vector $[\sigma^2_{Y_1}, \sigma^2_{Y_2}, ..., \sigma^2_{Y_n}]'$ almost surely does not belong to the optimal policy space.

3 A Monetary Policy Example

Consider a log-linear new Keynesian model, as in Walsh (2005) and Benigno and Woodford (2005), describing the dynamics of inflation π_t, the interest rate i_t, the welfare-relevant output gap $\widetilde{x}_t = y_t - y^*_t$, where y_t is output and y^*_t is its efficient level:

$$\widetilde{x}_t = -\frac{1}{\varphi}(i_t - E_t\pi_{t+1} - \widetilde{r}^n_t) + E_t(\widetilde{x}_{t+1}) \tag{3}$$

$$\pi_t - \gamma\pi_{t-1} = \lambda\widetilde{x}_t + \widetilde{\beta}E_t(\pi_{t+1} - \gamma\pi_t) + \lambda u_t \tag{4}$$

where φ is the coefficient of relative risk aversion for the representative household divided by the consumption share of output, $\widetilde{\beta}$ is the household's discount rate, λ is a function of behavioral parameters. It is assumed that a constant share of firms can adjust the price in each period, while the remaining share indexes the price to a fraction γ of last period's aggregate inflation rate. The variables u_t and $\widetilde{r}^n_t$ are linear combinations of all the exogenous shocks (a technology shock a_t, a tax shock τ_t, a government spending shock G_t), and are correlated. The appendix provide details on the model's derivation, and the mapping between the reduced form and structural parameters.

Let the policymaker's objective function be:

$$W_t = -\frac{1}{2}\Omega E_t \sum_{i=0}^{\infty} \widetilde{\beta}^i \left\{ \alpha\widetilde{x}^2_{t+i} + (\pi_{t+i} - \gamma\pi_{t+i-1})^2 \right\} \tag{5}$$

The parameter α specifies how the policymaker trades off fluctuations in output gap and inflation. While we assume that α depends on exogenous policymaker preferences, W_t is a second order approximation to the representative household's utility for $\alpha = \alpha^*$, where α^* is a well-defined function of the model's deep parameters.

In order to illustrate the main result, it is useful to start from a simplified model where $\gamma = 0$ and appropriate transfers ensure that the steady state is efficient. Then the model in eqs. (3), (4), (5) simplifies to the basic new Keynesian model, as found for example in Clarida, Gali and Gertler (1999), where movements in $\widetilde{r}_t^n$ can be interpreted as 'demand shocks', since they are not correlated with u_t, and can be perfectly offset by the policymaker. The time-consistent solution to the optimal policy problem requires:

$$\pi_t = -\frac{\alpha}{\lambda}\widetilde{x}_t \tag{6}$$

The law of motion for $\pi_t, \widetilde{x}_t$ under the optimal policy is:

$$\pi_t = \alpha q u_t \;\; ; \;\; \widetilde{x}_t = -\lambda q u_t$$

When u_t is described by an AR(1) stochastic process with autocorrelation parameter ρ_u, we obtain $q = \frac{1}{\lambda^2 + \alpha(1-\widetilde{\beta}\rho_u)}$. In this model any outcome $(\sigma^2_{\pi_t}, \sigma^2_{\widetilde{x}_t})$ could be generated by an optimal policy for $\alpha, \sigma^2_{u_t} \in [0, \infty]$. Using definition 1 and 2, the optimal policy space of the variables $(\pi_t, \widetilde{x}_t)$ associated with $Z(\beta; \mathbf{o})$ for $\beta = [\sigma^2_{u_t}, \alpha]'$ is $V_{\mathbf{o}} = \mathbf{R}^{2^+}$. Since any vector $[\sigma^2_{\pi_t}, \sigma^2_{\widetilde{x}_t}]'$ belongs to the image of $M(\beta; \mathbf{p})$ for $\mathbf{p} = \mathbf{o}$ any outcome can be generated by an optimal policy.

Consider the optimal policy space of the variables $(\pi_t, \widetilde{x}_t, i_t)$ for $\beta = [\sigma^2_{u_t}, \sigma^2_{\widetilde{r}_t^n}, \sigma_{u_t \widetilde{r}_t^n}, \alpha]'$. The law of motion for (π_t, i_t) implies:

$$\sigma^2_{\pi_t} = \left(\frac{\alpha}{\lambda}\right)^2 \sigma^2_{x_t} \tag{7}$$

$$\sigma^2_{i_t} = \left(\frac{\alpha}{\lambda}\gamma_\pi\right)^2 \sigma^2_{x_t} + \sigma^2_{\widetilde{r}_t^n} - 2\alpha q \gamma_\pi \sigma_{u_t \widetilde{r}_t^n} \tag{8}$$

where $\gamma_\pi = \left[\rho_u + \varphi\frac{\lambda}{\alpha}(1-\rho_u)\right]$. The optimal policy space is a $3 - surface$, and is a proper subset of $\mathbf{R}^{3^+}$ even if we allow the covariance $\sigma_{u_t \widetilde{r}_t^n}$ to be nonzero, since for given $(\sigma^2_{\pi_t}, \sigma^2_{x_t})$ the value of $\sigma^2_{i_t}$ is bounded by below, as shown in eq. (8). But $\sigma^2_{i_t}$ does not have an upper limit for any given $(\sigma^2_{\pi_t}, \sigma^2_{x_t})$, so the range of observable outcomes for $\sigma^2_{i_t}$ is infinite. Figure 1 shows a subset of the hyperplanes in $V_{\mathbf{o}}$. The set $V_{\mathbf{o}}$ is composed by an infinite number of hyperplanes, each indexed by a value for $\sigma^2_{\widetilde{r}_t^n}$.

Optimal policymaking puts tight restrictions on $V_{\mathbf{o}}$ for the set of endogenous variables (π_t, y_t, i_t), as shown in figure 2. The parameterization for $\varphi, \lambda, \widetilde{\beta}, \rho_u$ follows Walsh (2005). Since $y_t = y_t^* + \widetilde{x}_t$ it holds that:

$$y_t = -\left[\frac{1}{\varphi(1-\rho_a)}\widetilde{r}_t^n + \widetilde{x}_t\right] \tag{9}$$

where the technology shock a_t is an AR(1) stochastic processes with autocorrelation parameter ρ_a.[1] The set $V_{\mathbf{o}} \subseteq \mathbf{R}^{3^+}$ for (π_t, y_t, i_t) includes a bounded set of outcomes for $\sigma^2_{i_t}$ conditional on any$(\sigma^2_{\pi_t}, \sigma^2_{x_t})$. The intuition for the result is straightforward. Even if conditional on the optimal policy demand shocks do not affect π_t and $\widetilde{x}_t$, they affect y_t and i_t. As a consequence, for given $\sigma^2_{\pi_t}$ optimal outcomes where $\sigma^2_{y_t}$ is larger imply that $\sigma^2_{i_t}$ is larger too. Cost-push shocks increase the volatility of all three variables.

Optimal outcomes do not align on a two-dimension hyperplane because for different combinations $(\sigma^2_{u_t}, \sigma^2_{\widetilde{r}_t^n}, \alpha)$ there may exist more than one outcome for $\sigma^2_{i_t}$ corresponding to the same outcome for $(\sigma^2_{\pi_t}, \sigma^2_{y_t})$. Nevertheless, parameterizations where $V_{\mathbf{o}} \subseteq \mathbf{R}^{2^+}$ do exist. Conditional on the optimal policy (6), define:

$$M(\beta;\mathbf{o}) \equiv \begin{bmatrix} \sigma^2_{\pi_t} \\ \sigma^2_{y_t} \\ \sigma^2_{i_t} \end{bmatrix} = \begin{bmatrix} \alpha^2 q^2 \sigma^2_{u_t} \\ \lambda^2 q^2 \sigma^2_{u_t} + \frac{1}{\varphi^2(1-\rho_a)^2}\sigma^2_{\widetilde{r}_t^n} - \frac{2}{\varphi(1-\rho_a)}\lambda q \sigma_{u_t \widetilde{r}_t^n} \\ (\alpha q \gamma_\pi)^2 \sigma^2_{u_t} + \sigma^2_{\widetilde{r}_t^n} - 2\alpha q \gamma_\pi \sigma_{u_t \widetilde{r}_t^n} \end{bmatrix}$$

where $\beta = [\sigma^2_{u_t}, \sigma^2_{\widetilde{r}_t^n}, \sigma_{u_t \widetilde{r}_t^n}, \alpha]'$. The set $V_{\mathbf{o}}$ for this model is a $3-surface$, as can be checked by computing $\det[J_{M,\mathbf{g}}]$, and as shown in figure 2. If $\rho_u = \rho_a = 0$, using the definitions of q and γ_π we obtain:

$$M(\beta;\mathbf{o}) \equiv \begin{bmatrix} \sigma^2_{\pi_t} \\ \sigma^2_{y_t} \\ \sigma^2_{i_t} \end{bmatrix} = \begin{bmatrix} \alpha^2(\lambda^2+\alpha)^{-2}\sigma^2_{u_t} \\ \lambda^2(\lambda^2+\alpha)^{-2}\sigma^2_{u_t} + \frac{1}{\varphi^2}\sigma^2_{\widetilde{r}_t^n} - \frac{2}{\varphi}\frac{\lambda}{\lambda^2+\alpha}\sigma_{u_t \widetilde{r}_t^n} \\ (\varphi\lambda)^2(\lambda^2+\alpha)^{-2}\sigma^2_{u_t} + \sigma^2_{\widetilde{r}_t^n} - 2\varphi\frac{\lambda}{\lambda^2+\alpha}\sigma_{u_t \widetilde{r}_t^n} \end{bmatrix} \tag{10}$$

Eq. (10) shows that $\sigma^2_{y_t} = \frac{1}{\varphi^2}\sigma^2_{i_t}$ for any value of λ and φ. Therefore the Jacobian of $M(\beta;\mathbf{o})$ has two proportional columns for any β. Since $rank(J_{M,\mathbf{g}}) = 2$ over the domain D, the optimal policy space cannot be a $3-surface$. The image K can be parameterized by the function $\rho : S \subseteq \mathbf{R}^2 \to \mathbf{R}^3$:

$$\rho(S) \equiv \begin{bmatrix} g_1 \\ g_2 \\ \varphi^2 g_2 \end{bmatrix}$$

[1] Eq. (9) also assumes the government spending shock G_t is an AR(1) stochastic process with autocorrelation parameter $\rho_G = \rho_a$ and steady state government spending is zero.

where $g_1 = \alpha^2(\lambda^2+\alpha)^{-2}\sigma^2_{u_t}$ and $g_2 = \lambda^2(\lambda^2+\alpha)^{-2}\sigma^2_{u_t} + \frac{1}{\varphi^2}\sigma^2_{\widetilde{r}^n_t} - \frac{2}{\varphi}\frac{\lambda}{\lambda^2+\alpha}\sigma_{u_t\widetilde{r}^n_t}$. In this case, $V_{\mathbf{o}}$ is a $2-surface$ in $\mathbf{R}^3$, implying any outcome is suboptimal almost surely.

In general, by finding the appropriate combination of n endogenous variables, it may be possible to obtain an optimal policy space conditional on a model $Z(\beta;\mathbf{o})$ that includes only a bounded set of outcomes for at least one variable. The complement set $V^{\mathbf{C}}_{\mathbf{o}} = \mathbf{R}^{n^+}\backslash V_{\mathbf{o}}$ includes only suboptimal outcomes. Since the optimal policy space is defined in terms of variance of observable variables, it can be used to assess the restrictions of the optimal policy model for observable economic volatility.

3.1 Optimal Policy Restrictions from the New Keynesian Model and U.S. Monetary Policy

As an illustration of our methodology, consider the optimal policy space for the variables (π_t, y_t, i_t) conditional on the model in eqs. (3), (4), (5) and $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \sigma_{a_t\tau_t}, \alpha]'$. [2] We allow for endogenous inflation persistence by setting $\gamma = 0.5$ and consider an economy with a distorted steady state, so that any shock will affect all the endogenous variables, and consider the time-consistent optimal policy. While this is a stylized model, it is widely used in theoretical and empirical work. Figure 3 plots $V_{\mathbf{o}}$ (similar in shape to the plot in figure 2) together with the outcome $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ for the US over the period 1984:1 - 2005:1. There is no combination of the volatility of exogenous shocks and policymaker preferences that could have generated the observed $(\sigma^{US}_{\pi_t}, \sigma^{US}_{y_t}, \sigma^{US}_{i_t})$ as an optimal policy outcome.

Enlarging the parametric family of singular models leaves the result for the US sample unchanged. We build the function $M(\beta;\mathbf{o})$ for $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu]'$ where χ is the share of firms that cannot optimally adjust the price in each period, γ is the fraction of last period's aggregate inflation rate to which the share χ of firms indexes the price, θ is the firms' demand elasticity, ν is the inverse of labor supply wage elasticity. We assume the policymaker maximizes the representative household's utility. Table 1 reports the range of variation for the model's parameters. The mapping still results in $(\sigma^{US}_{\pi_t}, \sigma^{US}_{y_t}, \sigma^{US}_{i_t}) \notin V_{\mathbf{o}}$. [3] Including additional parameters in β may eventually result in a large enough optimal policy space such that $(\sigma^{US}_{\pi_t}, \sigma^{US}_{y_t}, \sigma^{US}_{i_t}) \in V_{\mathbf{o}}$, but does not need to because of the nonlinearity of the mapping $M(\beta;\mathbf{o})$.

[2] Using $\beta = [\sigma_{a_t}, \sigma_{G_t}, \sigma_{\tau_t}, \alpha]'$ would generate the same image for $M(\beta;\mathbf{o})$. To ease the reading of the plot in figure 2 the set $V_{\mathbf{o}}$ is defined in terms of the standard deviation of a variable rather than of its variance.

[3] We verified that $(\sigma^{US}_{\pi_t}, \sigma^{US}_{y_t}, \sigma^{US}_{i_t}) \notin V_{\mathbf{o}}$ by searching for a vector $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu,]'$ such that $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ is in the $\pm 2.5\%$ interval around the data point $(\sigma^{US}_{\pi_t}, \sigma^{US}_{y_t}, \sigma^{US}_{i_t})$. Allowing for a range of variation in $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ lets us account for the numerical error in the approximation to $M(\beta;o)$. The map $M(\beta;o)$ is computed through a discrete approximation over 3,686,000 simulated data points. We verified that admissable parameter values outside the range in table 1 result in outcomes further away from the historical observation for the US.

The result can be explained by two observations. First, all the model parameterizations imply different responses of endogenous variables to exogenous shocks. But many of the resulting models are nearly observationally equivalent in terms of unconditional volatility outcomes $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$. A measure of volatility is a coarse, low-level characterization of the behaviour of endogenous variables. What does change across model parameterizations is the mapping between the volatility of exogenous shocks and endogenous variables. That is, the same outcome $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ can be generated with alternative parameterizations by different vectors $[\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu]'$. Second, changes in a parameter do not necessarily add useful degrees of freedom to enlarge $V_{\mathbf{o}}$. For example, in the optimal policy space for $(\sigma_{\pi_t}, \sigma_{\widetilde{x}_t}, \sigma_{i_t})$ of the basic new Keynesian model a change in λ is observationally equivalent to a change in α, since the relationship between $\widetilde{x}_t$ and π_t and between $\widetilde{x}_t$ and i_t in eqs. (7) and (8) depends on the ratio α/λ.

The difficulty in finding a model within the parametric family such that the US outcome belongs to the optimal policy space has two alternative interpretations. First, US monetary policymaking was indeed suboptimal. After all, the building of the optimal policy space does allow for any possible parameterization in the vector $[\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu]'$, including parameterizations that may be inconsistent with available empirical evidence. Moreover, the optimal policy space has by construction weak power against detecting suboptimal policies: historical outcomes may belong to $V_{\mathbf{o}}$ even if they are the result of period-by-period suboptimal policies. Finally, it can be shown that the outcome $(\sigma_{\pi_t}^{US}, \sigma_{y_t}^{US}, \sigma_{i_t}^{US})$ does not belong to $V_{\mathbf{o}}$ for a number of alternative policies, including the timeless perspective optimal commitment policy, or the policymaker adopting the the wrong objective function and assuming $\gamma = 0$ in eq. (5), or even the policymaker adopting an objective function quadratic in π, $\widetilde{x}_t$ and Δi_t, for any relative weight of the three objectives.

Second, the DSGE model propagation mechanism is incomplete or inaccurate. Conditional on optimal monetary policy, it puts implausible restrictions on the endogenous variables' variances. This conclusion leads to question whether the optimal policy prescriptions derived from stylized DSGE models such as the one used are appropriate to guide real-world policymaking.

4 A Probabilistic Interpretation

Consider an optimal policy DSGE model with associated law of motion $Z(\beta; \mathbf{o})$ described by the linear map $Y_t = \mathbf{A}U_t$ where $\mathbf{A}$ is an $n \times r$ matrix. Partition the vector β into $\beta_\sigma = [\sigma_{U_{1,t}}, ...\sigma_{U_{r,t}}]$ and $\beta_k = [k_1, k_2, ...k_s]$. It is assumed the matrix $\mathbf{A}$ is a function of β_k, a vector of structural parameters of the model.

In general, when $r < n$ the support of the probability measure associated with the random vector Y_t lies on an $r-$dimension hyperplane in $\mathbf{R}^n$. The sample space $\Omega_\mathbf{r}$ is a null set with respect to Lebesgue measure in $\mathbf{R}^n$, and a density function is not defined with respect to the $n-$dimension Lebesgue measure, while it exists with respect to Lebesgue measure in $\mathbf{R}^r$ for events belonging to the $r-$dimension sample space $\Omega_\mathbf{r}$. This is the relevant sample space for most DSGE models used in business cycle analysis, and for every optimal policy model by construction, since optimal policy implies movements in the policy instrument can be written as a function of endogenous variables only, so that $r < n$.

If $r < n$ and U_t is normally distributed, the random vector Y_t is said to have a singular normal distribution. With a slight abuse of notation, we can write $Y_t \sim N_n[\mathbf{A}\mu_U,\ \mathbf{A}\Sigma_U\mathbf{A}']$. A singular normal distribution has a covariance matrix with rank strictly smaller than the dimension of the random vector. [4] To each parameter vector β_k corresponds a null set $\Omega_\mathbf{r}(\beta_k)$ in $\mathbf{R}^n$. Since the sample space $\Omega_\mathbf{r}$ is not a function of β_σ, a set $V_\mathbf{o}(\beta_k) = M(\beta_k, \beta_\sigma; \mathbf{o})$ can be associated with $\Omega_\mathbf{r}(\beta_k)$. Section 2 showed that $V_\mathbf{o}(\beta_k)$ and $\Omega_\mathbf{r}(\beta_k)$ have the same dimension, since if Y_t has singular covariance matrix with rank r, the set $V_\mathbf{o}(\beta_k)$ is a r-dimension hyperplane in $\mathbf{R}^n$.

The space $V_\mathbf{o}$ encompasses all sets $V_\mathbf{o}(\beta_k)$ for any parameterization of the vector β_k. Notice that since $V_\mathbf{o}$ simply maps entries of $\mathbf{A}$ and Σ_U into Σ_Y, the set $V_\mathbf{o}$ can be built regardless of the rank of Σ_U. On the contrary, we cannot define a joint density for the model $Z(\beta; \mathbf{o})$ since the sample space is the null set in $\mathbf{R}^n$, nor can we write a likelihood function for an observed sample.

A vector Σ_Y in $\mathbf{R}^n$ belonging to $V_\mathbf{o}$ must also belong to $V_\mathbf{o}(\beta_k)$ for some β_k, and therefore Σ_Y is the outcome of an optimal policy singular model. A model may impose restriction on Σ_U, for example requiring that the structural shocks U_t be uncorrelated, and Σ_U diagonal. The set $V_\mathbf{o}$ satisfying these restrictions is thus the *population* optimal policy space. If the vector β_σ includes all the elements of Σ_U we can build a *sample* optimal policy space, and incorporate the impact of small sample uncertainty. In the example discussed in section 3, assuming $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \sigma_{a_t\tau_t}, \alpha]'$ implies the set $V_\mathbf{o}$ includes all the realizations of the random vector $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ for any possible sample draw from the distribution of the random vector U_t, regardless of the population value for Σ_U. Therefore $V_\mathbf{o}$ is the space of all possible *sample* outcomes $(S_{\pi_t}, S_{y_t}, S_{i_t})$ for $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$. If a sample observation $(S_{\pi_t}, S_{y_t}, S_{i_t})$ does not belong to $V_\mathbf{o}$, then the sample $\{Y_t\}_{t=0}^T$ does not belong to the sample space $\Omega_\mathbf{r}(\beta_k)$ for any β_k.

[4] While a likelihood function for Y_t does not exist, various authors have proposed methods for maximum likelihood estimation of singular systems. See Bierens (2007), Kwakernaak (1979), Lai (2008).

The optimal policy space depicted in figure 3 is in fact drawn without imposing any restriction on Σ_U. The result implied by figure 3 that $(S_{\pi_t}^{US}, S_{y_t}^{US}, S_{i_t}^{US}) \notin V_{\mathbf{o}}$ has the interpretation that the sample $\{\pi_t, y_t, i_t\}_{t=0}^{T}$ such that $(S_{\pi_t}, S_{y_t}, S_{i_t}) = (S_{\pi_t}^{US}, S_{y_t}^{US}, S_{i_t}^{US})$ does not belong to the sample space of the data-generating process in eqs. (3), (4), (5) for all possible values of the preference parameter α. An equivalent way of stating the same result is that, while unconditionally the probability of any draw for $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ is always nonzero (and a confidence ellipse could be computed using standard statistical results for random sampling), no amount of sampling uncertainty could have generated the draw $(S_{\pi_t}^{US}, S_{y_t}^{US}, S_{i_t}^{US})$ conditional on our assumptions for the data generating process. Whatever the amount of sampling uncertainty, and the true population value for $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$, the data imply the model is false: either the propagation mechanism in eqs. (3), (4) is mistaken, or the policymaker deviated from the optimal policy in a way that the optimal policy space is able to discriminate.

The set $V_{\mathbf{o}}$ computed earlier for $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu]'$ imposed the restriction that Σ_U be diagonal, thus it did not include all possible sample outcomes $(S_{\pi_t}, S_{y_t}, S_{i_t})$. Building $V_{\mathbf{o}}$ accounting for sample uncertainty is straightforward, but computationally burdensome. For the case $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu]'$ we perform a different exercise, that illustrates the impact of the covariance matrix singularity on the optimal policy space. Assume the observable interest rate i_t^{obs} is described by

$$i_t^{obs} = i_t + w_t$$

where w_t is random variable with variance $\sigma_{w_t}^2 = \frac{x}{100}\sigma_{i_t}^2$. The value x gives the variance of the variable w_t as a percent share of the variance of the unobservable variable i_t, which is assumed to behave according to the optimal policy. In the econometric literature w_t is assumed to represent a measurement error. It can be interpreted as summarizing the volatility in i_t^o which is not explained by the DSGE model.

By adding a third source of randomness, we enlarge the set $V_{\mathbf{o}}$ of optimal policy outcomes, and obtain a measure of how large deviations of the observed σ_{i_t} from the volatility implied by the optimal policy need to be to have a nonzero probability of observing a given $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ conditional on the data-generating process in eqs. (3), (4), (5) and on all possible vectors $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \sigma_{w_t}, \chi, \gamma, \theta, \nu]'$.

Given our model, we can now ask what is the probability of a population value $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ equal to the US observation and belonging to $V_{\mathbf{o}}$ for different values of x. The probability is calculated for the standard deviation of a variable z_t belonging to the 5% interval $[b_{z_t,US}^L, b_{z_t,US}^H]$ centered around the observation $S_{z_t}^{US}$. Finally, let $V_{\mathbf{o}}^i \subseteq \mathbf{R}^+$ be the optimal policy space for the variable i_t and $V_{\mathbf{o}}^{\pi,y} \subseteq \mathbf{R}^{2+}$ be the optimal policy space for the variables (π_t, y_t). To scale the result we compute the probability of an outcome $\sigma_{i_t} \in [b^L_{i_t,US}, b^H_{i_t,US}]$ belonging to $V^i_{\mathbf{o}}$ conditional on any value within the 5% interval for $(\sigma_{\pi_t}, \sigma_{y_t})$ belonging to $V^{\pi,y}_{\mathbf{o}}$. Formally, we compute

$$\Pr\left\{\begin{array}{c}\left[(\sigma_{i_t} \in V^i_{\mathbf{o}}) \cap (b^L_{i_t,US} \leq \sigma_{i_t} \leq b^H_{i_t,US})\right] \mid \\ \left[(\sigma_{\pi_t}, \sigma_{y_t}) \in V^{\pi y}_{\mathbf{o}} \cap (b^L_{\pi_t,US} \leq \sigma_{\pi_t} \leq b^H_{\pi_t,US}) \cap (b^L_{y_t,US} \leq \sigma_{y_t} \leq b^H_{y_t,US})\right]\end{array}\right\}$$

Figure 4 plots the conditional probability against the variance $\sigma^2_{w_t}$ as a percent share x of the variance $\sigma^2_{i_t}$. Allowing for a third source of randomness implies that the US observation outcome can be the result of optimal policymaking, even without allowing for sampling uncertainty. The variable x provides a simple measure of the additional randomness needed for the US observation to belong to $V_{\mathbf{o}}$.

5 Related Literature

A growing literature investigates the fit of micro-founded DSGE models to the data conditional on an optimal monetary policy. Research focused on forward and backward-looking small macroeconomic models used in monetary policy work. Soderstrom et al. (2002) use informal calibration to match a new Keynesian model dynamics to US data. Dennis (2004), Favero and Rovelli (2003) and Salemi (2006) estimate structural models subject to the restriction that the policy rule minimizes the policymaker loss function.

Given a time series for the observables $(Y_{1_t}...Y_{n_t})$ with covariance matrix Σ_Y the approach adopted by these authors produces estimates for the deep parameters, the policymaker preferences, and a time series for a vector of shocks with nonsingular covariance matrix such that the theoretical model can generate the historical data, and such that a given function, depending on the econometric technique adopted, is maximized. This also implies that there will exist an estimated parameter vector, including random deviations from the optimal policy, such that the historical volatility outcome can be generated by the model.

Salemi (2006) shows how to use the nonsingular model estimation approach to compute a statistical test for optimal policymaking. The optimal policy imposes cross-equation restrictions on the estimated parameters, and their impact on the likelihood of the model can be exploited for testing. The optimal policy space is instead built exploiting the restrictions imposed by truly optimal policymaking in a parametric family of singular models on the volatility of observable variables. Compared to the estimation assumptions, the singular-model approach makes stronger assumptions on the behaviour of the policymaker, who is assumed to always implement the optimal policy. On the other hand, it relaxes the demand on the data fit since policies that are period-by-period suboptimal may still result in volatility outcomes belonging to the optimal policy space.

Clearly a three-equations model, as the one adopted in this paper, can only provide a stylized description of the economy's behaviour. Yet small optimal policy DSGE models are estimated to gain insight into the preferences of the policymaker, and are often relied upon by economists to illustrate and generate policy prescriptions and guidelines. Computing the optimal policy space for such models provides important insights into the restrictions on the data that the models imply.

6 Conclusions

This paper studied the restrictions implied by optimal policy DSGE models for the volatility of observable endogenous variables.

Our approach relies on the restrictions imposed by optimal policymaking on the variance of the endogenous variables in singular models. To generate a non-trivial set for the volatility of observable variables - which we label the optimal policy space - we introduce variation in the behavioral parameters when building the set of outcomes consistent with the model. We show that a DSGE model can be associated with a well-defined subset of all the possible volatility outcomes, which is not of measure zero. This is the result of the nonlinearity of the mapping between a DSGE model parameter space and the implied volatility of the endogenous variables. Nonsingular models, which assume random perturbations to optimal policymaking, imply no observable outcome has zero probability.

We illustrated our method by building the optimal policy space of a widely used new Keynesian model. Conditional on this model, recent US monetary policymaking would have zero likelihood of being the result of optimal policymaking. Since this approach has by construction low power in discriminating optimal policy outcomes, we interpret the result as evidence that widely used optimal policy models can only be consistent with a very limited set of volatility outcomes, regardless of the parameterization adopted.

References

[1] Baxandall, P. and Liebeck, H., (1986), Vector Calculus, Oxford: Clarendon Press.

[2] Benigno, P. and Woodford, M., (2005), ”Inflation Stabilization and Welfare: the Case of a Distorted Steady State”, *Journal of the European Economic Association* 3(6): pp. 1185-1236.

[3] Bierens, H., (2007) "Econometric analysis of linearized singular dynamic stochastic general equilibrium models", *Journal of Econometrics* 136: pp. 595-627.

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

[5] Conlon, L., (2001), Differentiable Manifolds, Boston: Birkhauser.

[6] Dennis, R., (2004), ”Inferring Policy Objectives from Economic Outcomes”, *Oxford Bulletin of Economics and Statistics* 66: pp. 735-764.

[7] Favero, C. and Rovelli, R., (2003), ”Macroeconomic Stability and the Preferences of the Fed: A Formal Analysis, 1961-1998”, *Journal of Money, Credit and Banking* 35: pp. 546-556.

[8] Kwakernaak, H., (1979), "Maximum likelihood parameter estimation for linear systems with singular observations", *IEEE Transactions on Automatic Control* AC-24:3.

[9] Lai, Hung-pin, (2008), "Maximum likelihood estimation of singular systems of equations", *Economic Letters* 99: pp. 51-54.

[10] Salemi, M., (2006), ”Econometric Policy Evaluation and Inverse Control”, *Journal of Money, Credit and Banking* 38: pp. 1737-1764.

[11] Soderstrom, U., Soderlind, P. and Vredin, A., (2002), ”Can a Calibrated New -Keynesian Model of Monetary Policy Fit the Facts?”, Sveriges Riksbank Working Paper 140.

[12] Walsh, C., (2005), ”Endogenous Objectives and the Evaluation of Targeting Rules for Monetary Policy”, *Journal of Monetary Economics* 52: pp. 889-911.

New Keynesian model	parameter range for US optimal policy space		
γ	χ	v	θ
0.2-0.82	0.1-0.66	0.1-1.17	4-16

Table 1: New Keynesian model parameter space used to compute optimal policy space $V_{\mathbf{o}} = M(\beta; o)$ for $\beta = [\sigma_{a_t}, \sigma_{\tau_t}, \chi, \gamma, \theta, \nu]'$. Other parameters are set as in Walsh (2005). Model is described by the time-consistent solution to maximization of eq. (5) given eqs. (3), (4) and assuming the policymaker's objective function maximizes the utility of the representative household. Parameter χ is the share of firms that cannot optimally adjust the price in each period, γ is the fraction of last period's aggregate inflation rate to which the share χ of firms indexes the price, θ is the firms' demand elasticity, ν is the inverse of labor supply wage elasticity. Parameter values outside the range in table 1 result in outcomes $(\sigma_{\pi_t}, \sigma_{y_t}, \sigma_{i_t})$ further from the historical US observation for the sample 1984:1-2005:1.

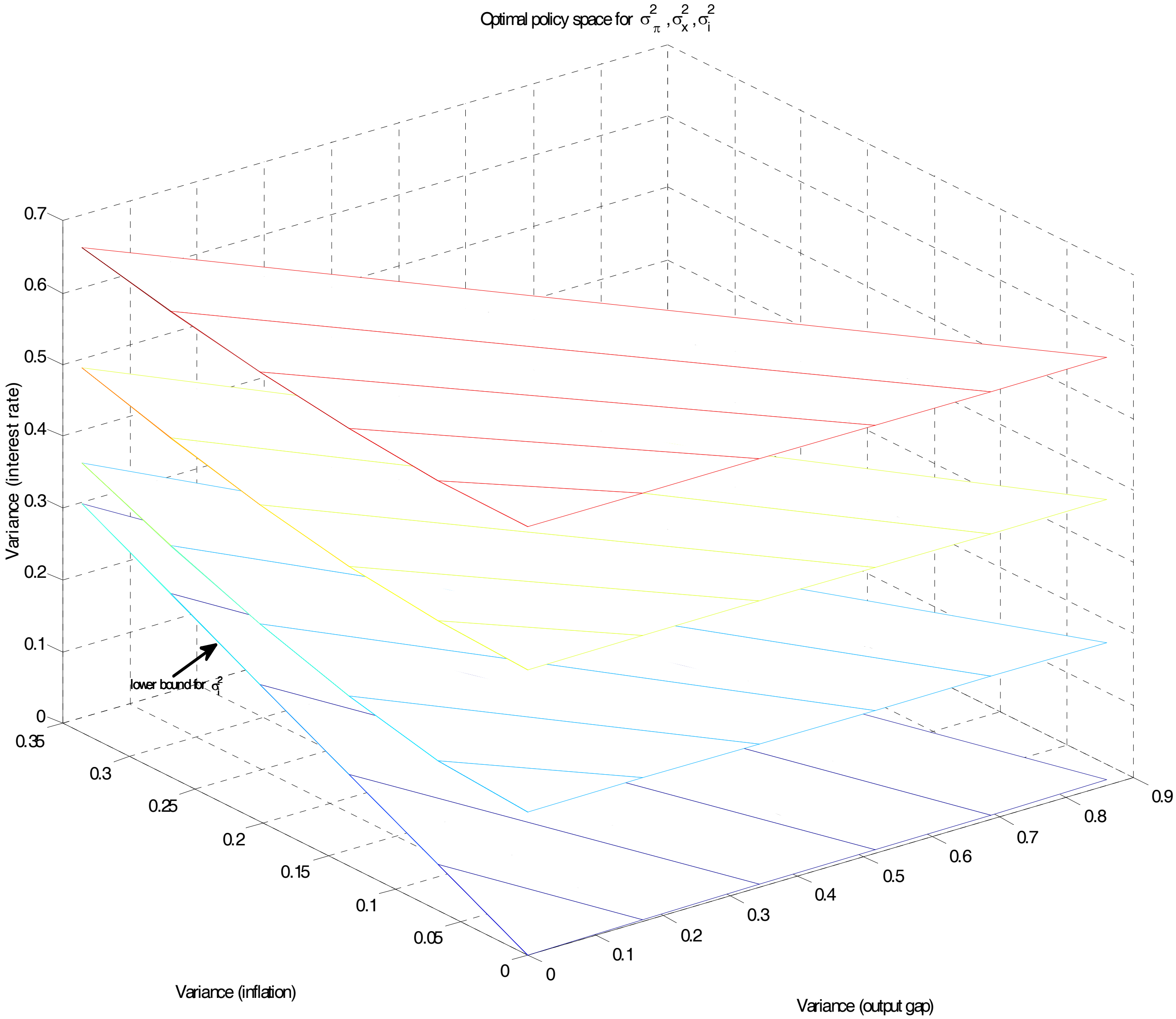

Figure 1: Sample optimal policy hyperplanes in the optimal policy space $V_{\mathbf{o}}$ for the variables $(\pi_t, \widetilde{x}_t, i_t)$ and for $\beta = [\sigma^2_{u_t}, \sigma^2_{\widetilde{r}^n_t}, \alpha]'$ using the baseline new Keynesian model. Each hyperplane is indexed by a value for $\sigma^2_{\widetilde{r}^n_t}$.

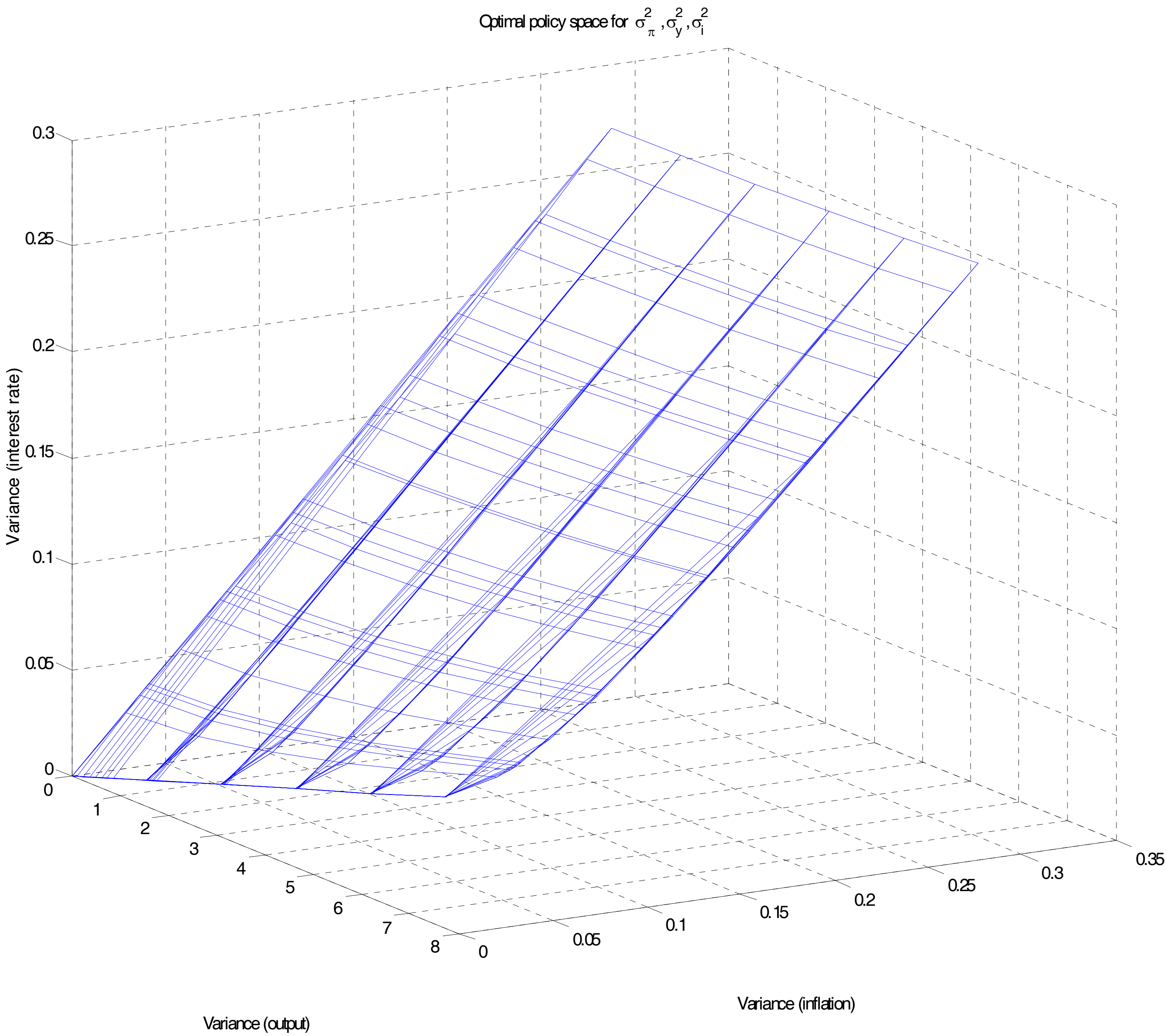

Figure 2: A subset of the optimal policy space $V_{\mathbf{o}}$ for the variables (π_t, y_t, i_t) and for $\beta = [\sigma^2_{u_t}, \sigma^2_{\tilde{r}^n_t}, \alpha]'$ using the baseline new Keynesian model.

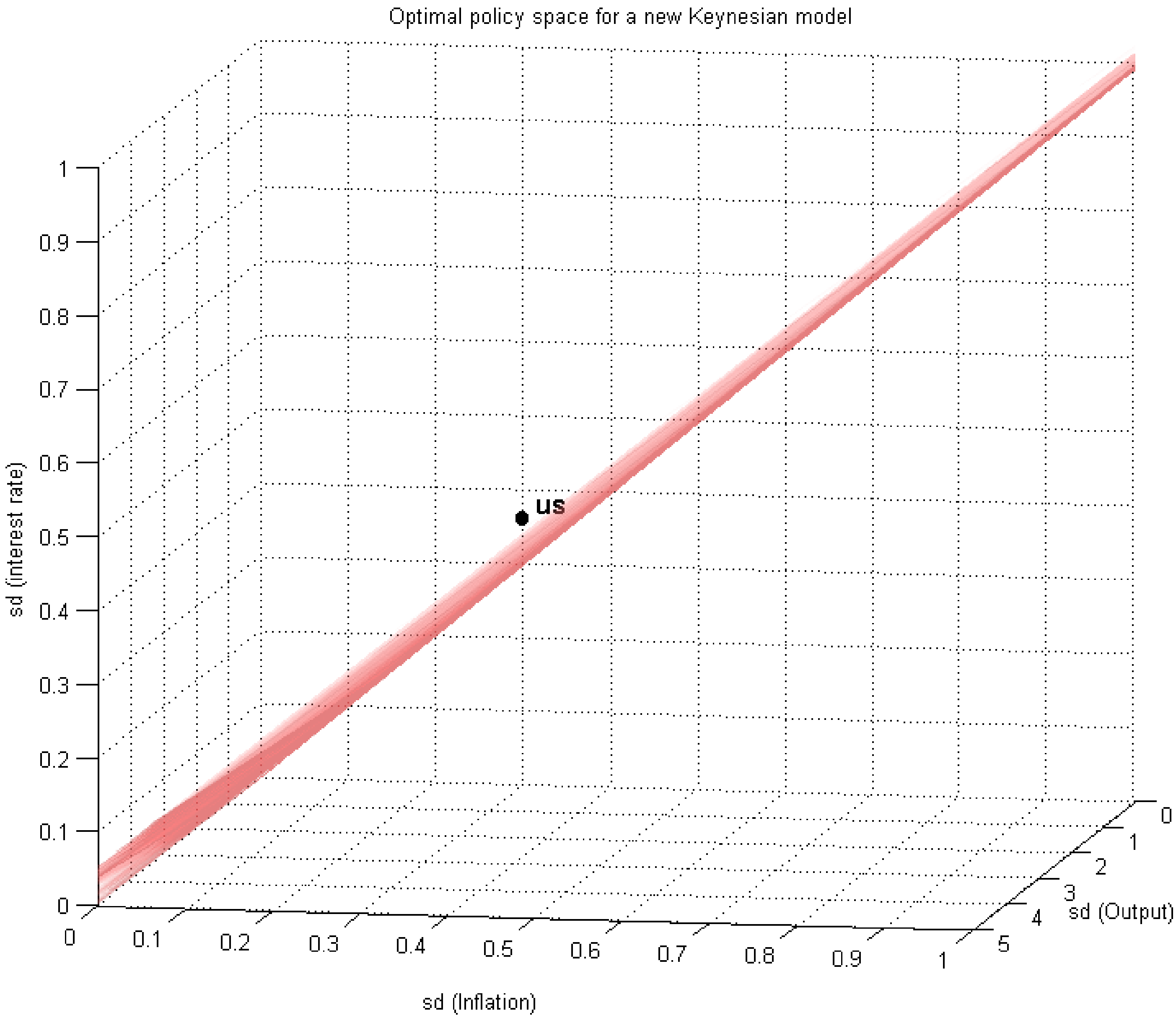

Figure 3: A subset of the optimal policy space $V_{\mathbf{o}}$ for the variables (π_t, y_t, i_t) and for $\beta = [\sigma^2_{u_t}, \sigma^2_{\tilde{r}^n_t}, \alpha]'$ using a new Keynesian model with endogenous inflation persistence and a distorted steady state. The plot shows the historical volatility outcome for the US over the period 1984:1 - 2005:1. Output y_t is detrended seasonally adjusted non-farm business sector real GDP. Inflation π_t is seasonally adjusted CPI inflation. Interest rate i_t is 3-month government bond. All data is sampled at quarterly intervals. Rates are not annualized.

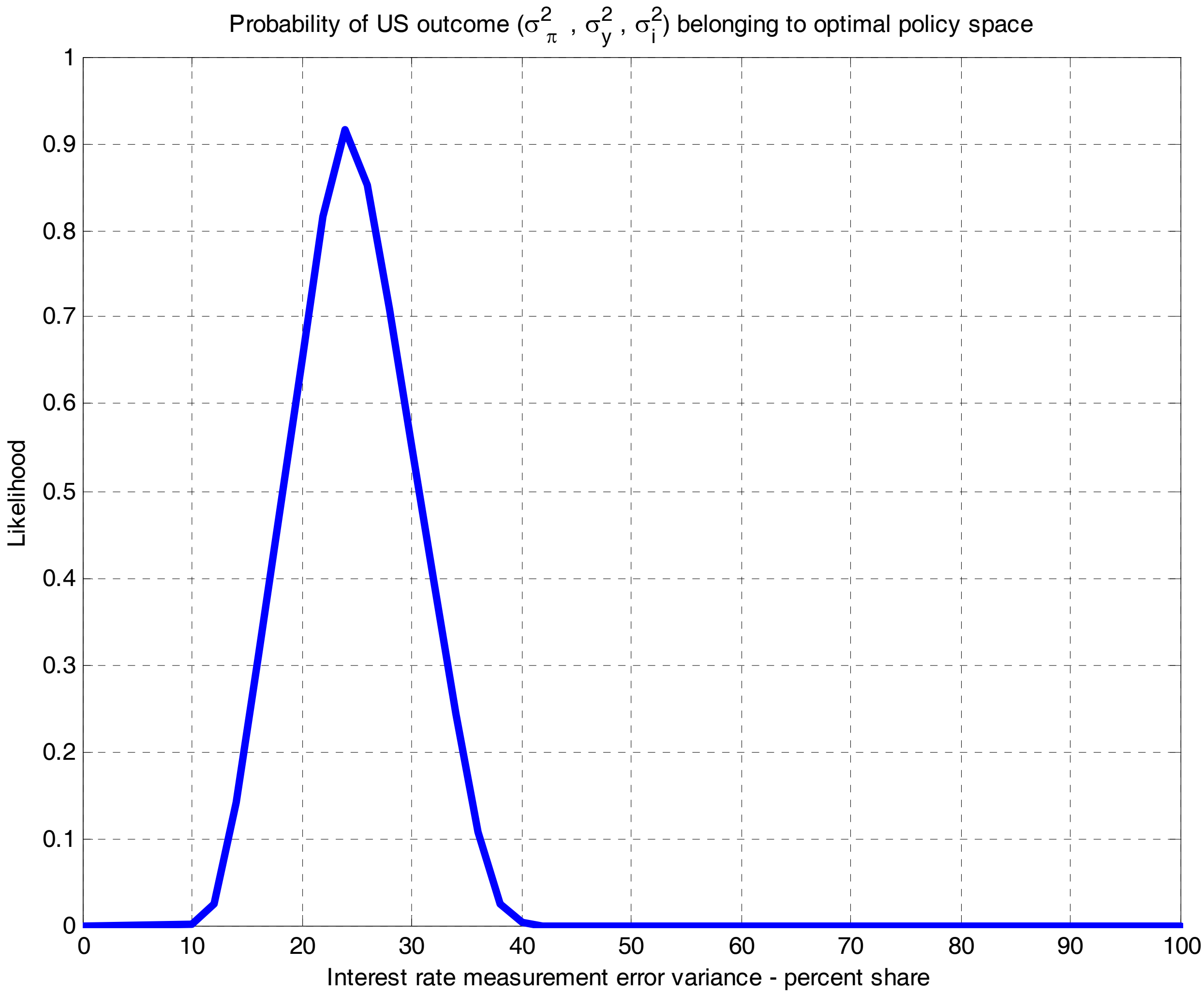

Figure 4: Probability of the outcome $\{(b^L_{i_t,US} \leq \sigma_{i_t} \leq b^H_{i_t,US}) \cap (b^L_{\pi_t,US} \leq \sigma_{\pi_t} \leq b^H_{\pi_t,US}) \cap (b^L_{y_t,US} \leq \sigma_{y_t} \leq b^H_{y_t,US})\}$ belonging to the optimal policy space $V_{\mathbf{o}}$, conditional on the outcome $\{(b^L_{\pi_t,US} \leq \sigma_{\pi_t} \leq b^H_{\pi_t,US}) \cap (b^L_{y_t,US} \leq \sigma_{y_t} \leq b^H_{y_t,US})\}$ belonging to the optimal policy space $V_{\mathbf{o}}^{\pi,y}$. Horizontal axis measures variance of the measurement error for observed interest rate i_t^{obs} as a percent share of the variance for the optimal interest rate i_t, given by $\sigma^2_{w_t} = \frac{x}{100}\sigma^2_{i_t}$.

7 Appendix: Solution of the Benigno and Woodford (2005) Model

Consider the New Keynesian model for inflation π_t, output gap x_t, interest rate i_t as described in Walsh (2005) and Benigno and Woodford (2005):

$$\begin{aligned} x_t &= -\frac{1}{\varphi}(i_t - E_t\pi_{t+1} - r_t^n) + E_t(x_{t+1}) \qquad (11) \\ \pi_t - \gamma\pi_{t-1} &= \lambda x_t + \widetilde{\beta} E_t(\pi_{t+1} - \gamma\pi_t) \qquad (12) \\ x_t &= y_t - y_t^n \end{aligned}$$

where r_t^n is the Wicksellian real rate of interest, y_t is output, y_t^n is the level of output that would obtain in the flexible-price equilibrium, φ is the coefficient of relative risk aversion for the representative household divided by the consumption share of output, $\widetilde{\beta}$ is the household's discount rate. It is assumed that a constant share of firms can adjust the price in each period, while the remaining share indexes the price to a fraction γ of last period's aggregate inflation rate. When prices can optimally adjust in every period the rational expectation equilibrium solution for y_t^n and r_t^n does not depend on i_t :

$$\begin{aligned} y_t^n &= \phi_1 G_t + \phi_2 a_t + \phi_3 \tau_t \\ r_t^n &= \phi_4 E_t(y_{t+1}^n - y_t^n) + \phi_5 E_t(G_{t+1} - G_t) \end{aligned}$$

$$\begin{aligned} \phi_1 &= \frac{\varphi}{\omega + \varphi} \\ \phi_2 &= \frac{\zeta(1+v)}{\omega + \varphi} \\ \phi_3 &= \frac{[\overline{\tau}/(1-\overline{\tau})]}{\omega + \varphi} \\ \phi_4 &= \varphi \\ \phi_5 &= (1 - s_C) \\ \omega &= \zeta(1+v) - 1 \end{aligned}$$

The variable G_t is defined as exogenous government consumption (in log-deviations from the steady state), a_t is an exogenous productivity shock, τ_t is an exogenous income tax shock. The parameter ζ is the elasticity of firm output with respect to labor input, v is the inverse of the wage elasticity of labor supply, ω is the inverse of the elasticity of firm marginal cost with respect to output, $\overline{\tau}$ is the steady state tax rate, s_C is the consumption steady state share of output, φ is the coefficient of relative risk aversion for the representative household divided by s_C. The elasticity of inflation with respect to x_t is given by:

$$\lambda = \frac{(1-\chi)(1-\chi\widetilde{\beta})}{\chi(1+\theta\omega)}(\omega+\varphi)$$

In the absence of transfers to correct the steady state distortions arising from taxes and imperfect competition, or in the case $\tau_t \neq 0$, the efficient level of output y^* is different from y^n and is given by:

$$y_t^* = w_1 y_t^n + w_2 G_t + w_3 \tau_t$$

$$\begin{aligned}
w_1 &= \frac{\omega+\varphi+\Phi(1-\varphi)}{\xi} \\
w_2 &= \frac{\Phi\sigma}{(\omega+\varphi)\xi s_C} \\
w_3 &= \overline{\tau}/(1-\overline{\tau})\xi \\
\xi &= (\omega+\varphi)+\Phi(1-\varphi)-\frac{\Phi\sigma(s_C^{-1}-1)}{(\omega+\varphi)} \\
\Phi &= 1-\frac{\theta-1}{\theta}(1-\overline{\tau})
\end{aligned}$$

where θ is the firms' demand elasticity. The second order approximation to the utility of the household can be written as:

$$\begin{aligned}
W_t &= -\frac{1}{2}\Omega E_t \sum_{i=0}^{\infty} \widetilde{\beta}^i \left\{\alpha \widetilde{x}_{t+i}^2 + (\pi_{t+i}-\gamma\pi_{t+i-1})^2\right\} \qquad (13) \\
\widetilde{x}_t &= (y_t - y_t^*)
\end{aligned}$$

where $\widetilde{x}_t$ is the welfare-relevant output gap. W_t is equal to the household's welfare for $\alpha = \alpha^*$ where

$$\alpha^* = \frac{\lambda}{w_1\theta}$$

The model in (11), (12) can be expressed in terms of the endogenous variables appearing in the objective function (13):

$$\begin{aligned}
\widetilde{x}_t &= -\frac{1}{\varphi}(i_t - E_t\pi_{t+1} - \widetilde{r}_t^n) + E_t(\widetilde{x}_{t+1}) && (14)\\
\pi_t - \gamma\pi_{t-1} &= \lambda\widetilde{x}_t + \widetilde{\beta}E_t(\pi_{t+1} - \gamma\pi_t) + \lambda u_t && (15)\\
\widetilde{r}_t^n &= \phi_4 E_t(y_{t+1}^* - y_t^*) + \phi_5 E_t(G_{t+1} - G_t) &&\\
u_t &= y_t^* - y_t^n &&
\end{aligned}$$

The variable u_t is a linear combination of all the exogenous shocks. The variable Φ is a measure of the steady state distortions in the economy. If appropriate transfers ensure, as is often assumed, that the steady state is efficient, then $\Phi = 0$. Benigno and Woodford (2005) show that in this case $w_1 = 1, w_2 = 0$, and

$$u_t = w_3\tau_t$$

Assume $\gamma = 0$. Then the problem faced by the optimal policymaker can be written as:

$$Max \ -\frac{1}{2}\Omega E_t \sum_{i=0}^{\infty} \widetilde{\beta}^i \left\{\alpha\widetilde{x}_{t+i}^2 + \pi_{t+i}^2\right\} \tag{16}$$

$$\begin{aligned}
st \ \ \widetilde{x}_t &= -\frac{1}{\varphi}(i_t - E_t\pi_{t+1} - \widetilde{r}_t^n) + E_t(\widetilde{x}_{t+1}) && (17)\\
\pi_t &= \lambda\widetilde{x}_t + \widetilde{\beta}E_t\pi_{t+1} + \lambda u_t && (18)\\
u_t &= w_3\tau_t && (19)\\
\widetilde{r}_t^n &= \phi_4 E_t\left[\frac{\varphi(G_{t+1} - G_t) + \zeta(1+v)(a_{t+1} - a_t)}{\omega + \varphi}\right] + \phi_5 E_t(G_{t+1} - G_t) && (20)
\end{aligned}$$

In this model movements in a_t or G_t can be interpreted as 'demand shocks' since they affect $\widetilde{r}_t^n$ but not u_t, therefore do not affect the trade-off between the stabilization objectives and can be perfectly offset by the policymaker. The variable u_t takes the interpretation of a 'cost push' shock, and depends only on movements in τ_t. Assuming, as in eq. (9), that $s_C = 1$, $G_t = \rho_G G_{t-1} + \varepsilon_{G_t}$,

$a_t = \rho_a a_{t-1} + \varepsilon_{a_t}$, $\varepsilon_t \sim iid$, $\rho_G = \rho_a$ it holds:

$$\begin{aligned} \widetilde{r}_t^n &= \phi_4 E_t(y_{t+1}^* - y_t^*) \\ y_t^* &= -\frac{1}{\varphi(1-\rho_a)}\widetilde{r}_t^n \end{aligned} \tag{21}$$

Eq. (21) holds also for $s_C < 1$ and $G_t = 0 \; \forall \; t$ or for $s_C < 1$ and $\rho_g = 1$.

The optimal time-consistent policy is given by the FOC:

$$\pi_t - \gamma\pi_{t-1} = -\frac{\alpha}{\lambda}(1+\widetilde{\beta}\gamma)x_t$$

The timeless perspective optimal commitment policy is given by the FOC:

$$\pi_t - \gamma\pi_{t-1} = \left(-\frac{\alpha}{\lambda}\right)(x_t - x_{t-1})$$

Baseline parameterization The parameterization follows Walsh (2005) unless otherwise stated in the main text.

$$\begin{aligned} \chi &= 0.66 \\ \gamma &= 0.5 \\ \widetilde{\beta} &= 0.99 \\ \varphi &= 0.16 \\ \phi &= 1.5 \\ \theta &= 7.88 \\ s_C &= 0.8 \\ v &= 0.49 \\ \overline{\tau} &= 0.2 \\ \rho_a &= 0.95 \\ \rho_G &= 0.95 \\ \rho_\tau &= 0.95 \end{aligned}$$