

Ordering Policy Rules with an Unconditional Welfare Measure*

Tatiana Damjanovic,^a Vladislav Damjanovic,^a and
Charles Nolan^b

^aUniversity of Exeter

^bUniversity of Glasgow

The unconditional expectation of social welfare is often used to assess alternative macroeconomic policy rules in applied quantitative research. This paper provides a detailed analysis of such policies. It sets out the unconditionally optimal (UO) policy problem and derives a linear-quadratic (LQ) version of that problem that approximates the exact non-linear problem. The properties of UO policies are analyzed through a series of examples and contrasted with the timeless perspective (TP), expounded in Benigno and Woodford (2012). Some substantive implications for optimal monetary policy are explored.

JEL Codes: E20, E32, F32, F41.

1. Introduction

DSGE models can be difficult to analyze in their non-linear form, so a linear-quadratic (LQ) approximation is often adopted. That approach has been extended by Sutherland (2002), Benigno and Woodford (2004, 2005, 2006b), and many others in the context of specific models. More recently, Benigno and Woodford (2012)

*This paper was substantially improved following comments by the editor, Pierpaolo Benigno. We should also like to thank Jinill Kim and two anonymous referees for helpful comments. We also thank the participants at the 2008 North American meeting of the Econometric Society at Carnegie Mellon University, the 2008 Norges Bank Workshop on Optimal Monetary Policy in Oslo, and the 2007 Royal Economic Society Conference and research seminars at the University of Exeter, the Bank of England, and the Dutch Central Bank for helpful comments and discussions. The usual disclaimer, of course, applies. Author e-mails: t.damjanovic@exeter.ac.uk; v.damjanovic@exeter.ac.uk; Charles.Nolan@glasgow.ac.uk.

demonstrate that the LQ approach can be applied quite generally to optimal policy problems. They show that if that approximation is performed around the optimal steady state, the solution to the LQ problem represents a first-order approximation to the (non-linear) optimal policy. The steady state is optimal in the sense that it maximizes the original non-linear objective function with respect to the original non-linear structural constraints of the model. In particular, Benigno and Woodford (2012) show that it is possible to apply the LQ method to a model where the government searches for an optimal policy from a timeless perspective (TP). Moreover, their method delivers a pure second-order approximation to TP social welfare which permits one to rank alternative policies from a timeless perspective.¹

One objective of this paper is to show that it is also possible to apply an LQ approach to the case where a policymaker maximizes the unconditional expectation of social welfare. Precisely, we will show that an LQ approximation to the unconditional expectation of social welfare around the unconditionally optimal (UO) steady state will also generate a welfare function with purely second-order terms. The solution of that LQ-UO problem yields a first-order approximation to the policy which maximizes the unconditional expectation of social welfare. Moreover, a pure second-order welfare function can be obtained and used to measure the performance of simple policy rules from an UO perspective.

The more important objective of the paper is to offer a detailed analysis of UO policies since that is lacking in the literature, despite the prevalence of the UO perspective in quantitative macroeconomic research. For example, Taylor (1979) suggested that, in quantitative theoretical investigations under rational expectations, macroeconomic stabilization policies ought to optimize the unconditional expectation of the policymaker's objective function. That perspective on policy assessment has been popular; some prominent examples include Whiteman (1986), Rotemberg and Woodford (1998), Clarida, Gali, and Gertler (1999), Woodford (1999b), Erceg, Henderson, and Levin (2000), Kollman (2002), Kim and Henderson (2005), and Schmitt-Grohe and Uribe (2007). More recently, many

¹See also Kim and Kim (2007).

researchers have adopted the TP in quantitative and theoretical research. Consequently, it will be useful to compare UO and TP approaches in what follows in order to highlight the key features of these differing perspectives on optimal policy.

The potential attractions of unconditionally optimal policy, as opposed to the timelessly optimal policy, are pursued in Jensen and McCallum (2010). The timeless perspective minimizes the variance but does not account for the initial conditions (see Woodford 2002, p. 508). However, the distribution of initial conditions depends on the policy adopted in preceding periods. That relation is explicitly taken into account by the unconditionally optimal approach. We demonstrate through a number of examples that adopting the unconditionally optimal methodology internalizes the initial distribution and delivers a policy which is best on average. That is, UO policy performs better on average precisely because it takes into consideration the initial conditions. UO policy is optimal given that a similar policy choice was made by past policymakers. Hence, the principal difference between TP policy and UO policy is that the former takes the distribution of initial conditions as given and ignores impact of policy on them in its design. In section 3 we explore in depth the fundamental reasons why UO and TP policies differ. We also extend Jensen and McCallum (2010) by applying the UO methodology to a DSGE model with no forward-looking constraints on the policy problem. That provides a very clear example of what internalizing initial conditions actually means.

This paper also highlights important differences between UO and TP policies when the steady state is distorted. Jensen and McCallum (2010) compare UO (what they call “optimal continuation” policies) with TP policies, when the steady state is efficient. In that case, it is shown that the form of the welfare function to be optimized is the same across programs. Here we show that the corresponding LQ problems can be significantly different when the steady state is distorted. We find that UO and TP approaches imply different steady states, different arguments in the quadratic social welfare function, and different linear dynamic constraints. Even the number of non-degenerate dynamic constraints may differ across the TP and UO policy problems.

To design our algorithm, we extend the methodology of Damjanovic, Damjanovic, and Nolan (henceforth DDN) (2008), which

derives the first-order necessary conditions for the policy optimizing the unconditional expectation of welfare.² Then, similar to Judd (1999) and Benigno and Woodford (2012), the linear-quadratic approximation is done around the optimal deterministic steady state; in our case, around the unconditionally optimal steady-state, and in Benigno and Woodford's case, the timeless-perspective (TP) steady state.³

As a result, we develop a useful approach for constructing the pure second-order approximation to unconditional welfare. Since this measure can be presented in the form of a linear combination of the second moments, one can apply the Anderson et al. (1996) algorithm, which has good convergence properties. Consequently, it is also straightforward numerically to analyze policies from a UO perspective.

A specific application of the approach is provided employing the canonical New Keynesian model. A number of insights emerge. First, unconditionally optimal monetary policy is characterized by trend inflation. That trend in inflation complicates the linear quadratification.⁴ That explains a second insight: The second-order approximate loss function is no longer defined solely over terms in output and inflation as found in DDN (2008) for the non-distorted steady-state case. However, the loss function that one obtains is easily interpreted in light of the underlying distortions in the economy. The approximate loss function is used to evaluate and rank different simple rules for monetary policy (i.e., the nominal interest rate). The welfare implications of nominal income targeting versus inflation targeting are explored, and our results are contrasted with some of those of Kim and Henderson (2005).

The rest of the paper is organized as follows. In section 2 we recall the basic setup of the problem and the necessary first-order conditions for the optimal steady state.⁵ In section 3 we discuss further and contrast the attractions of UO and TP policies. It is

²See also Whiteman (1986), Blake (2001), and Jensen and McCallum (2002), (2010).

³See also Sutherland (2002), Debortoli and Nunes (2006), and Levine, Pearlman, and Piers (2008).

⁴As shown in Ascari and Ropele (2007) and Damjanovic and Nolan (2010).

⁵That is the only section which overlaps with DDN (2008). Moreover, here we provide a new derivation of the UO policy problem.

shown in section 4 that one can derive a purely quadratic approximation to the unconditional expectation of the objective function. In section 5 a canonical New Keynesian, Calvo price-setting model is set up. Section 6 formalizes the policy problem and demonstrates the application of the various steps in the approach of section 4. There is then a brief discussion of the implications for optimal monetary policy when the steady state is distorted and the authorities are optimizing over the unconditional loss function. We also contrast the linear-quadratic form of UO and TP approaches, including the difference in the point of approximation. In section 7 we use the unconditional welfare criterion to explore briefly the impact of different simple rules for monetary policy. There is also a short discussion of optimal monetary policy following productivity and markup shocks under TP and UO policies. Section 8 offers some conclusions. Appendices contain proofs and details of key derivations.

2. The General UO Problem

Consider a discounted loss function of the form

$$L_t = (1 - \beta) E_t \sum_{j=0}^{\infty} \beta^j l(x_{t+j}), \quad (1)$$

where E_t is the expectations operator *conditional on information up through date t* , β is the time discount factor, $l(x_{t+j})$ is the period loss function, and x_t is a vector of target variables. Specifically, $x_t = [Z_t, z_t, i_t]$, where Z_t is a vector of predetermined endogenous variables (lags of variables that are included in z_t and i_t); z_t is a vector of non-predetermined endogenous variables (including “jump” variables), the value of which will generally depend upon both policy actions and exogenous disturbances at date t ; and i_t is a vector of policy instruments, the value of which is chosen in period t . Let μ_t denote a vector of exogenous disturbances. For simplicity, assume that μ_t is a function of primary i.i.d. shocks, $(e_i)_{-\infty}^t$.

Further, let the evolution of the endogenous variables z_t and Z_t be determined by a system of simultaneous equations,

$$E_t F(x_{t+1}, x_t, \mu_t) = 0. \quad (2)$$

Let us further assume, following Taylor (1979), that the policymaker seeks to minimize the unconditional expectation of the loss function (1), subject to constraints, (2).⁶ That is, he or she searches for a policy rule

$$\varphi(E_t x_{t+1}, x_t, \mu_t) = 0, \quad (3)$$

such that

$$\varphi^* = \arg \min EL_t(\varphi), \quad (4)$$

where E is the unconditional expectations operator. We call such a policy “unconditionally optimal” and denote it “UO policy.”

Formally, the unconditional expectation of any function $l(x)$ can be represented in Lebesgue integral form as

$$El_t(x_t(\varphi)) = \int l_t(x_t(\varphi, e)) de,$$

where de is the Cartesian product probability measure of i.i.d. primary shocks with history, $(de_{t-k})_{k=0}^{\infty}$. We emphasize that de is given exogenously and does not change with policy.

2.1 Solution

The first step is to formulate the non-linear policy problem and identify the non-stochastic steady state around which approximation needs to take place. For this purpose we will use constraints (2) and necessary first-order conditions.

2.1.1 Necessary Conditions for an Optimum

Consider the following Lagrangian function which is implied by the above optimal policy problem:

$$\mathcal{L}^{UO}(\{y_t, x_t, \mu_t\}) = E(l(x_t) + \xi_t F(y_t, x_t, \mu_t) + \rho_t(x_{t+1} - y_t)), \quad (5)$$

⁶Taylor’s approach may be interpreted as a recommendation: Policymakers *ought* to seek to minimize the unconditional value of the loss function. This appears partly, perhaps largely, in response to the issue of time inconsistency. See Taylor (1979) for further discussion. McCallum (2005) is an interesting discussion of these, and related, issues.

where we define

$$y_t = x_{t+1}. \tag{6}$$

We introduce definition (6) solely for presentational purposes. The formulation of (5) is justified as follows. The Lagrangian is defined as a function of all variables in all periods of time. Constraints (2) and (6) should be satisfied for all time periods and for any realization of the history of the shocks, $(e_i)_{-\infty}^{t+1}$. In each period t for each realization of the history e , we will have a pair of constraints corresponding to (2) and (6) and a corresponding pair of Lagrange multipliers $(\xi_t(e); \rho_t(e))$. We need to sum up all these constraints across all possible histories of the shocks' realizations and across time. To that end, first define \mathcal{L}_t , which is the sum of all constraints at time t , multiplied by their Lagrange multipliers:

$$\mathcal{L}_t := \int \xi_t(e) F(y_t, x_t, \mu_t, e) + \rho_t(e) (x_{t+1}(e) - y_t(e)) de. \tag{7}$$

By definition, \mathcal{L}_t is the unconditional expectation, and its value does not depend on time. Thus, one may write

$$\mathcal{L}_t = E(\xi_t F(y_t, x_t, \mu_t) + \rho_t(x_{t+1} - y_t)). \tag{8}$$

Now formally one can sum over time periods for any discount rate; we use β for consistency. Thus, $(1 - \beta) E_t \sum_{j=0}^{\infty} \beta^j \mathcal{L}_t = \mathcal{L}_t$, as \mathcal{L}_t represents the unconditional expectation which is independent of time. Therefore expression (5) is the sum of the objective and all the constraints multiplied by corresponding Lagrange multipliers.

DDN (2008) show that the necessary conditions for the optimality of policy, φ , is that it implies a path for the endogenous variables, x_t and y_t , and that there exist Lagrange multipliers, (ξ_t, ρ_t) , that together satisfy the first-order conditions (9), (10) and constraints (2),⁷

$$\frac{\partial H}{\partial x_t} = \frac{\partial l(x_t)}{\partial x} + \xi_t \frac{\partial F(y_t, x_t, \mu_t)}{\partial x} + \rho_{t-1} = 0; \tag{9}$$

$$\frac{\partial H}{\partial y_t} = \xi_t \frac{\partial F(y_t, x_t, \mu_t)}{\partial y_t} - \rho_t = 0, \tag{10}$$

⁷The notation ξF is a shorthand for the tensor product, $\sum_{i=1}^n \xi_i F_i$.

where $H(y_t, x_t, \mu_t)$ is the Hamiltonian for (5), such that $\mathcal{L}(y_t, x_t, \mu_t) = E(H(y_t, x_t, \mu_t))$.

Judd (1999), Woodford (2002), and Benigno and Woodford (2005) demonstrate very clearly that the choice of the steady state is crucial (along with the solution concept for forward-looking policy problems) in being able to obtain LQ approximations to general non-linear, forward-looking policy problems. To choose the deterministic steady state around which log-linearization takes place, one needs to solve the system of first-order conditions (9), (10) and constraints (2). This leads to the following proposition.

PROPOSITION 1. *The steady state (X, ξ) is defined by the system (11)–(12):*

$$F(X, X, \mu) = 0; \quad (11)$$

$$\frac{\partial l(X)}{\partial x_t} + \xi \frac{\partial F(X, X, \mu)}{\partial x} + \xi \frac{\partial F(X, X, \mu)}{\partial y} = 0, \quad (12)$$

where X , ξ , and μ indicate the vectors of steady-state values of endogenous variables, Lagrange multipliers, and the average value of shocks, respectively.

We refer to (X, ξ) as the “unconditionally optimal steady state.”⁸ In the absence of shocks, solution (11) shows that unconditionally optimal policy delivers the steady state with the highest level of steady-state welfare “on average,” where the averaging is with respect to the unconditional measure. It is worth emphasizing that the TP approach discussed in, e.g., Woodford (2002) implies different first-order conditions and therefore a different center of approximation. That difference will be shown to lead to a different optimal monetary policy.

3. Comparing UO and TP

In this section we begin our analysis of UO and TP policies. Before proceeding, it is noted that comparing TP and UO policies should be done with care⁹: TP policies are optimal (and time consistent)

⁸It is assumed throughout that system (11) has a unique solution.

⁹The editor has emphasized this point to us.

when one employs a TP welfare measure, and the same is true vis-à-vis UO policy. However, we will on occasion ask whether agents would prefer to be situated in economies with TP or UO policies.¹⁰

3.1 TP Program and Social Time Discounting

The TP approach was defined first in Woodford (1999b), where a mathematical description of the policy was provided via the first-order conditions for TP policy. One first forms an appropriate Lagrangian for the TP program:

$$\begin{aligned} \mathcal{L}^{TP}(y_t, x_t, \mu_t) = E_t \sum_{j=0}^{\infty} \beta^j & (l(x_{t+j}) + \xi_t F(y_{t+j}, x_{t+j}, \mu_{t+j}) \\ & + \rho_t (x_{t+1+j} - y_{t+j})). \end{aligned} \quad (13)$$

The necessary first-order conditions with respect to future variables will be

$$\frac{\partial \mathcal{L}^{TP}}{\partial x_t} = \frac{\partial l(x_t)}{\partial x} + \xi_t \frac{\partial F(y_t, x_t, \mu_t)}{\partial x} + \frac{1}{\beta} \rho_{t-1} = 0; \quad (14)$$

$$\frac{\partial \mathcal{L}^{TP}}{\partial y_t} = \xi_t \frac{\partial F(y_t, x_t, \mu_t)}{\partial y_t} - \rho_t = 0. \quad (15)$$

Based on these relations, one concludes that TP policy coincides with fully optimal Ramsey policy conditional on the economy starting from the TP-optimal steady state. Thus, (14), (15), and (2) determine the TP optimal steady state. It is also the convergence state of the fully optimal Ramsey policy, conditional on the government possessing appropriate commitment technology.

Comparing the first-order conditions for the UO program (9) and (10) with those of the TP (14) and (15), one concludes that TP policy coincides with UO policy if the policymaker's discount factor is equal to unity, $\beta = 1$. Hence, the policymaker's time discounting is central to the differing perspectives of UO and TP policies. We explore these differences in a little more detail now.

¹⁰In much the same way as public finance theorists might inquire whether agents would prefer to live in a world with Benthamite or Rawlsian tax and redistribution policies.

3.2 Policies Compared

One may compare TP and UO policies along two related dimensions. First, which concept of optimality is more appropriate in normative study of policy problems? Second, if actual policymakers are able to commit, which of the two policies are preferable? Jensen and McCallum (2010) pursue the first of these, concluding that UO policies are more appropriate in many instances. We refer the reader to their insightful discussion. Assuming both policies are achievable, which policy ought a policymaker adopt? That is perhaps something of a philosophical question.

Unconditionally optimal policy maximizes average utility, given that all generations follow one rule. To design that rule, one takes into account not only current actions but also the actions one would have wished our predecessor policymakers to have followed; such a rule would bequeath preferable initial conditions to the current policymakers. By the same token, whatever current policymakers would have asked of our predecessors is what current policymakers undertake to provide to future policymakers. Therefore the benefits to the current generation of the rule, due to past commitments, are traded against costs associated in passing benefits to future generations. In other words, the UO methodology entails treating future generations the same as the current generation, or without any discounting ($\beta = 1$). Of course, that idea is not new in economic philosophy.

According to Ramsey (1928), discounting future generations' welfare is unethical. Harrod (1948) recommends that government "correct" individuals' savings decisions because they reflect positive time discounting and a resulting "palpable improvidence." In this he agrees with Pigou (1932), who comments on agents' "defective telescopic faculty" as a reason why private discount rates are excessive. In particular, Harrod argues that an individual will ex post be grateful to a government which induces him or her to invest the amount corresponding to a decision predicated on a zero discount rate. Solow (1974) also argues that the social discount factor perhaps should be higher than the private one.

More concretely, a recent example of the importance of time discounting is in the area of environmental policy, where the current generation is expected to invest in environmental protection for the benefit of future generations. The size of that investment crucially

depends on the discount factor attached to the welfare of future generations. Stern (2007, p. 31) proposed to “treat the welfare of future generations on a par with our own,” which on a strict interpretation is to set the discount factor equal to 1. The same view has been expressed in Anand and Sen (2000) and many others.

Another reason for a “low” social discount factor follows from the assumption that utility depends on past as well as present and future variables. For example, good memories may make one happier. Strotz (1956) introduced that idea and Caplin and Leahy (2004) show that, in such a setting, policymakers should be more patient than private individuals.

In the next section a simple example is provided that highlights, we think, the attractions, and difficulties, in implementing UO policy relative to TP policy. We show that it may be not altruism towards future generations but rather considerate behavior of the predecessors that makes UO policy more attractive. In particular, in the following example, it is preferable for agents to live in a UO world given that the previous generations acted under UO policy, rather than living in a TP world inherited from TP predecessors. These benefits notwithstanding, there are costs in a transition to UO policy, and some generation will have to bear them.

3.3 Treatment of the Initial Conditions

The central difference between UO and TP policies is in the treatment of initial conditions. TP policy takes the initial conditions as given and optimizes future losses accordingly. UO policy acknowledges the fact that policy affects the distribution of initial conditions as well. To show the difference, we apply these differing policies to a simple model with only backward-looking constraints on policy—in that case, the value of the state variable is the relevant “initial condition.” In this “backward-looking” model, the TP policy will coincide with optimal Ramsey policy (i.e., one with discounting). UO policy will give a different outturn, which we describe below. Consider the following example.

EXAMPLE 1. The technological process is the following. Generation t makes a costly investment N_t (like planting edible seeds). Generation $t + 1$, consumes the fruit from the crop

$$C_{t+1} = A_{t+1}N_t, \quad (16)$$

and makes an investment for the next generation. Here A_{t+1} is an index of productivity. Society is altruistic and cares about succeeding generations applying a discount rate β to their welfare. So utility of the current generation t is

$$U_t = \log(C_t) - kN_t + \beta U_{t+1}, \quad k > 0.$$

Upon integrating forward, social welfare is found to be

$$U_t = \sum \beta^{t+s} (\log(C_{t+s}) - N_{t+s}).$$

Then, on substituting in the production technology (16), the policy maximand is

$$U_t = \sum \beta^{t+s} (\log(A_{t+s}) + \log(N_{t+s-1}) - N_{t+s}). \quad (17)$$

The first-order condition with respect to N_t is

$$\frac{\partial U_t}{\partial N_t} = \beta \frac{1}{N_t} - 1 = 0.$$

Clearly, therefore, the Ramsey solution and the TP solution coincide and $N_t = \beta$.

Now consider a rule for N that has applied for all time; it has been followed by past generations and it will be followed by the current and future generations. In that case, one seeks an optimal choice, $N_t = N$ for all t . Thus, utility becomes

$$U = \sum \beta^{t+s} (\log(A_{t+1}) + \log(N) - N) \quad (18)$$

and the first-order condition gives $N = 1$. One might inquire whether or not one would wish to live in an economy with a TP policy or a UO policy. In fact, the UO policy generates much larger utilities for all generations. In this particular case, the gain in consumption equivalent is $\exp(-\log(\beta) - (1 - \beta))$, which can vary from 0.5 percent if β is relatively large, say $\beta = 0.9$, to 6 percent for $\beta = 0.7$.

So one may assume that, given the choice, an individual would prefer to live in an economy where the government runs a policy corresponding to a higher social discount factor.

The difference in TP and UO policies can be explained by the treatment of initial conditions. TP methodology considers N_{t-1} as given and treats it as a “term independent of policy.” UO methodology is designed as the best policy, conditional that it is accepted by all generations, including generation $t - 1$. That is why N_{t-1} is taken into account.

However, in the present example, there are two major drawbacks of the UO approach relating to transition and time consistency. If a country decides to switch from TP to UO policy, the transition generation will be worse off; they will inherit a lower investment left by the previous generation but will be asked to invest more for the sake of the next generation. Moreover, every generation *irrespective of their inheritance* will have an incentive to deviate from UO policy, switching to Ramsey policy (which coincides in this instance with the TP). Therefore, UO policy is time inconsistent in the sense of Kydland and Prescott (1977).

It is an open question whether one is able to identify actual policies that may have been (approximately) optimal from an unconditional perspective. However, there is circumstantial evidence that sometimes transition costs are incurred following major policy changes, that is, when the current generation is forced to suffer for a better future. It happens during wartime. It also often appears to happen during pension reform; all future generations will benefit from a less distorted economy if the current generation sacrifices part of their pension benefits. Despite the costly transition, more than eighty countries recently undertook some degree of pension reform (see Holzmann and Hinz 2005). Moreover, although the transition cost is clear in a deterministic environment, it may be less visible, and therefore more politically implementable, when an economy is subject to stochastic shocks and the favorable shocks can compensate the present generation for lower investment made by the previous generation.

So, with only backward-looking constraints, TP policy, unlike UO policy, is in a sense “stable”; it is time consistent in the sense of Kydland and Prescott (1977) and credible. However, matters are

somewhat different when forward-looking structural equations are present, as we now discuss.

3.4 *Forward-Looking Models*

As Soderlind (1999) shows, the best time-invariant policy depends on initial conditions at the time when the decision is made. Moreover, the initial conditions are changing over time so that there is always an incentive to deviate from any time-invariant policy. DDN (2008) show that UO policy maximizes objectives over all possible initial conditions, which is presented as the history of the realization of the exogenous shocks. UO policy acknowledges that the initial conditions depend on the policy run by predecessors and internalizes this. Another feature of the UO policy is that it maximizes over all initial conditions implied by the policy run by the same policymaker acting optimally in the past. In contrast, TP policy commits to time-zero expectations in the same way (that is, with the same functional form) as in the future. However, initial conditions are not necessarily the steady state, but can be any state. What is key is that TP does not internalize initial conditions. We now set these differences out explicitly.

Any policy ϕ together with constraints (2) define choice variables x_t as a function of initial values and shocks e_t :

$$\phi : \{x_{t-1}, e_t\} \rightarrow x_t.$$

Therefore, policy ϕ generates a distribution of initial conditions, $F_\phi(x_t)$. The timeless-perspective methodology takes that distribution as given and ignores the fact that policy influences the distribution of initial conditions. In particular, Woodford (2002, p. 509) decomposes the objective function into two components:

$$L = L^{\text{det}} + L^{\text{stab}},$$

where L^{det} depends on initial conditions, and L^{stab} depends only on the responses to unexpected shocks. He explains that the TP method minimizes L^{stab} and so does not internalize the influence of policy on the distribution of initial conditions. Hence,

$$\phi^{TP} = \arg \min L^{\text{stab}}(\phi).$$

UO policy takes the influence of ϕ on the distribution of initial conditions into account. The simplest way to see this is by noting that the distribution of L^{det} depends on the policy which was implemented by predecessors. In particular, L^{det} can be completely defined once we know the complete history of shocks and the policy: $L^{\text{det}} = L^{\text{det}}(\phi, e_{t-})$. Therefore, UO policy is defined as

$$\phi^{UO} = \arg \min \left(\int L^{\text{det}}(\phi, e_{t-}) de_{t-} + L^{\text{stab}}(\phi) \right).$$

In other words, UO methodology internalizes the initial distribution and delivers the policy which is best on average.

3.5 Stationarity

In contrast to the TP, UO policy induces stationarity. For example, under the TP it is optimal to permit permanent increases in debt and taxes following structural shocks under nominal rigidity (see Benigno and Woodford 2004, 2006a, and Schmitt-Grohe and Uribe 2004). However, Horvath (2011) finds that in the log-linearized unconditionally optimal economy, public debt converges to its steady state following a shock. Clearly, the UO approach is not applicable to non-stationary policies. Any policy which causes some variable to evolve as a unit root would generate unlimited unconditional losses, and such a policy would not be adopted by a policymaker with a UO perspective.

An alternative measure of unconditionally optimal policy is proposed in Benigno and Woodford (2012) in which the unconditional expectations operator is applied to a stationary sub-space of all variables. The values and expected values of unit-root variables are treated as initial conditions and therefore are classified as “terms independent of policy,” or “t.i.p.” Specifically, all predetermined variables are split into “trend” and “cyclical” components, where the “trend” consists of all non-stationary variables. The unconditional measure is then applied to the “cyclical” component only. How does this alternative UO policy differ from the one developed in the current paper? Consider two policies. Assume that they generate the same volatility of inflation. However, policy 1 ($P1$) induces a unit root in output, while policy 2 ($P2$) induces stationary output with finite volatility. According to UO policy as proposed in Benigno

and Woodford (2012), $P1$ is to be preferred to $P2$ since the volatility of output will be counted as part of trend output and will not have an impact on the ranking of alternative, feasible policies. Therefore, an economy with infinitely volatile output will be preferred to an economy with low output volatility, for a given volatility of inflation.¹¹

Alternatively, under the standard UO measure (as defined in section 2), $P1$ will be rejected in favor of $P2$. That is because the rank of the “trend” sub-space depends on policy design and will be internalized by the policymaker optimizing unconditional losses.

4. The Possibility of Pure Second-Order Approximation

In this section we show that it is possible to construct a pure second-order approximation to a general unconditional optimization problem (1), subject to constraints (2). We formulate it in the following proposition.

PROPOSITION 2. *It is always possible to approximate unconditional welfare up to second order around the UO steady state (X, ξ) , defined by the system (11)–(12).*

Proof. The value of the loss function $El(x_t)$ should not change if combined with the unconditional expectation of the constraints $EF(y_t, x_t, \mu_t)$. Thus, appendix 1 demonstrates that the second-order approximation to this combination has a pure second-order form. That is,

$$\begin{aligned} El(x_t, \mu_t) &= E[l(x_t) + \xi F(y_t, x_t, \mu_t)] \\ &= EQ_l + \xi EQ_F + t.i.p + O_3. \end{aligned} \quad (19)$$

The notation O_3 denotes third- or higher-order terms. Q_l and Q_F are pure second-order terms of the log-approximation, around the

¹¹To be clear, what matters for welfare is the discounted value of second moments, which might be finite even if volatility is infinite. This is why unit-root processes might be allowed.

unconditionally optimal steady state, to the loss function $l(x_t)$ and dynamic constraints $EF(y_t, x_t, \mu_t)$:

$$Q_l = \frac{1}{2} \left(X^2 \frac{\partial^2 l}{\partial x^2} \hat{x}_t \hat{x}_t \right);$$

$$Q_F = \frac{1}{2} X^2 \left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} \right) \hat{x}_t \hat{x}_t + XX \frac{\partial^2 F}{\partial x \partial y} \hat{x}_t \hat{x}_{t+1} + X\mu \frac{\partial^2 F}{\partial x \partial \mu} \hat{x}_t \hat{\mu}_t$$

$$+ X\mu \frac{\partial^2 F}{\partial y \partial \mu} \hat{x}_{t+1} \hat{\mu}_t,$$

where we use \hat{x}_t to denote a log-deviation from steady state.

It is straightforward to show that the maximization of the unconditional objective (19) subject to the linearized analogues of equations (2) yields the same solution as log-linearization of the first-order conditions (9). This latter approach is proposed by Khan, King, and Wolman (2003) in the context of conditional optimization, and is extended in DDN (2008) to unconditional optimization. Implementing the above result turns out to be fairly straightforward both from a numerical perspective and an analytical perspective.

4.1 Numerical Methodology

Any model economy can be described by the agents' behavioral dynamics (2) and the policy rule (3). Log-linearization of those two equations can be presented in the form of

$$E_t V_{t+1} = AV_t + B\varepsilon_{t+1},$$

where V_t is the vector of endogenous variables and ε_{t+1} is the vector of exogenous shocks. In this form it is straightforward to construct the variance-covariance matrix, $R \equiv EV_t V_t'$, using standard software.¹² That is, R is recovered by solving the following matrix equation:

$$R = ARA' + B\Upsilon B', \quad (20)$$

¹²For example, Dynare gives the variance-covariance matrix as part of its standard output.

where $\Upsilon = E\varepsilon_t\varepsilon_t'$ is the unconditional variance-covariance matrix of the underlying shock processes. Equation (20) can be solved numerically using a doubling algorithm as described in Anderson et al. (1996) using an equivalent form

$$R = \sum_{j=0}^{+\infty} A^j B \Upsilon B' A'^j.$$

The social welfare function can be computed as a linear combination of the elements of matrix R .

4.2 Substitution Techniques for UO and TP Policies

Although the method of pure second-order approximation (19) is straightforward and quite efficient, it may be useful to show how one can replicate the same welfare analysis by substituting variables employing the dynamic constraints (2). In particular, it demonstrates that even though UO policy cannot ignore initial conditions, that does not prevent one from using a substitution approach for UO policy analysis. Consider a second-order approximation to the dynamic constraint equations,

$$\hat{x}_{t+1} = \alpha \hat{x}_t + \hat{y}_t + Q_t + O_3, \quad (21)$$

where Q_t is a pure quadratic form.

We first discuss the TP methodology which has been important for recovering the relationship between means and variances in second-order approximations useful for welfare analysis.¹³ The TP substitution methodology expresses the discounted sum of $\{\hat{x}_{t+s}\}_{s=0}^{+\infty}$ as a function of $\{\hat{y}_{t+s}\}_{s=0}^{+\infty}$. In that case, equation (21) is integrated forward to yield

$$\sum_{s=0}^{+\infty} \beta^s \hat{x}_{t+1+s} = a \sum_{s=0}^{+\infty} \beta^s \hat{x}_{t+s} + \sum_{s=0}^{+\infty} \beta^s \hat{y}_{t+s} + \sum_{s=0}^{+\infty} \beta^s Q_{t+s} + O_3.$$

¹³Sutherland (2002) was the first to apply this approach to the case of an economy with a distorted steady state and a particular policy rule. See also Kim and Kim (2003).

That expression can be simplified as

$$(\beta^{-1} - a) \sum_{s=0}^{+\infty} \beta^s \widehat{x}_{t+s} - \beta^{-1} \widehat{x}_t = \sum_{s=0}^{+\infty} \beta^s \widehat{y}_{t+s} + \sum_{s=0}^{+\infty} \beta^s Q_{t+s} + O_3.$$

Then an initial value, \widehat{x}_t , is ignored as a “t.i.p.” and the final expression appears as

$$\sum_{s=0}^{+\infty} \beta^s \widehat{x}_{t+s} = \frac{1}{\beta^{-1} - a} \sum_{s=0}^{+\infty} \beta^s \widehat{y}_{t+s} + \frac{1}{\beta^{-1} - a} \sum_{s=0}^{+\infty} \beta^s Q_{t+s} + O_3.$$

This expression is then used to calculate approximate utility.

To deliver the analogous expression in the case of UO policy, one applies the unconditional expectations operator to (21),

$$E\widehat{x}_{t+1} = E\alpha\widehat{x}_t + E\widehat{y}_t + EQ_t + O_3. \quad (22)$$

Then, one uses the fact that $E\widehat{x}_{t+1} = E\widehat{x}_t$, which transforms (22) into

$$E\widehat{x}_t = \frac{1}{1-a} E\widehat{y}_t + \frac{1}{1-a} EQ_t + O_3, \quad (23)$$

which is the desired expression.

5. Example: Calvo Model with Distorted Steady State

A more or less canonical dynamic New Keynesian model is now developed and two issues in particular are pursued. First, which model variables appear in the approximate loss function under UO policy? Second, some insight is sought into the nature of UO monetary policy compared with TP policy. That comparison is pursued further in section 7.

5.1 The Households

There is a large number of identical agents in this (closed) economy where the only input to production is labor. Each agent evaluates utility using the following criterion:

$$E_0 \sum_{t=0}^{\infty} \beta^t U(Y_t, N_t(i)) = E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(Y_t) - \frac{\lambda}{1+v} \left(\int_i N_t(i) di \right)^{1+v} \right). \quad (24)$$

E_t denotes the conditional expectations operator at time $t \geq 0$, β is the discount factor, Y_t is consumption, and $N_t(i)$ is the quantity of labor supplied to industry i ; labor is industry specific. $v \geq 0$ measures the labor supply elasticity, while λ is a preference parameter.

Consumption is defined over a Dixit-Stiglitz basket of goods,

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}}. \quad (25)$$

The average price level, P_t , is known to be

$$P_t = \left[\int_0^1 p_t(i)^{1-\theta} di \right]^{\frac{1}{1-\theta}}. \quad (26)$$

The demand for each good is given by

$$Y_t(i) = \left(\frac{p_t(i)}{P_t} \right)^{-\theta} Y_t^d, \quad (27)$$

where $p_t(i)$ is the nominal price of the final good produced in industry i and Y_t^d denotes aggregate demand.

Agents face the flow constraint

$$P_t Y_t + B_t = (1 + i_{t-1}) B_{t-1} + (1 - \tau) W_t N_t + \Pi_t. \quad (28)$$

As all agents are identical, the only financial assets traded in equilibrium will be those issued by the fiscal authority. Here B_t denotes the nominal value of government bond holdings, at the end of date t ; $1 + i_t$ is the nominal interest rate on this “riskless” one-period nominal asset; W_t is the nominal wage in period t (our assumptions mean that we do not need to index wages on i); and Π_t indicates any profits remitted to the individual. It is assumed that labor income is taxed at rate τ . The usual conditions are assumed to apply

to the consumer's limiting net savings behavior. Hence, necessary conditions for an optimum include

$$-\frac{U'_N(Y_t, N_t)}{U'_Y(Y_t, N_t)} = \lambda N_t^v Y_t = (1 - \tau) w_t, \quad (29)$$

$$w_t = \frac{\lambda}{1 - \tau} N_t^v Y_t, \quad (30)$$

and

$$E_t \left\{ \frac{\beta U'_Y(Y_{t+1}, N_{t+1})}{U'_Y(Y_t, N_t)} \frac{P_t}{P_{t+1}} \right\} = \frac{1}{1 + i_t}. \quad (31)$$

Here w_t denotes the real wage. The complete-markets assumption implies the existence of a unique stochastic discount factor,

$$Q_{t,t+k} = \beta \frac{Y_t P_t}{Y_{t+k} P_{t+k}}, \quad (32)$$

where

$$E_t \{Q_{t,t+k}\} = E_t \prod_{j=0}^k \frac{1}{1 + i_{t+j}}.$$

5.2 Representative Firm: Factor Demand

As noted, labor is the only factor of production. Firms are monopolistic competitors who produce their distinctive goods according to the following technology:

$$Y_t(i) = A_t [N_t(i)]^{1/\phi}, \quad (33)$$

where $N_t(i)$ denotes the amount of labor hired by firm i in period t , A_t is a stochastic productivity shock, and $1 < \phi$.

The demand for output determines the demand for labor. Hence one finds that

$$N_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\theta\phi} \left(\frac{Y_t}{A_t} \right)^\phi. \quad (34)$$

There is an economy-wide labor market so that all firms pay the same wage for the same labor. As a result, as asserted above, one may write $w_t(i) = w_t, \forall i$. All households provide the same share of labor to all firms. The total amount of labor will then be

$$N_t = \int N_t(i) di = \left(\frac{Y_t}{A_t} \right)^\phi \int \left(\frac{P_t(i)}{P_t} \right)^{-\theta\phi} di = (A_t^{-1} Y_t)^\phi \Delta_t, \quad (35)$$

where Δ_t is the measure of price dispersion:

$$\Delta_t \equiv \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\theta\phi} di. \quad (36)$$

5.3 Representative Firm: Price Setting

As in Calvo (1983), each period a fixed proportion of randomly chosen firms is allowed to adjust prices. Those firms choose the nominal price which maximizes their expected profit given that they have to charge the same price in k periods' time with probability α^k . As usual, we assume that firms are cost takers. Let $p'_t(i)$ denote the choice of nominal price by a firm that is permitted to reprice in period t .

Real profits can be written as

$$\frac{P_t(i)}{P_t} Y_t(i) - \mu_t w_t N_t(i), \quad (37)$$

where μ_t is a cost-push shock so that the total cost facing firm i will be $TC(i) = \mu_t w_t N_t(i)$. In combination with the production function, $TC(i) = \mu_t w_t \left(\frac{Y_t(i)}{A_t} \right)^\phi$, and therefore marginal cost for a particular firm i is $c(i) = \phi \mu_t w_t \left(\frac{Y_t(i)}{A_t} \right)^\phi / Y_t(i)$.

As all firms that are permitted to reprice will choose the same price, optimal repricing implies

$$\left(\frac{p'_t}{P_t} \right)^{1+\theta(\phi-1)} = \frac{\left(\frac{\theta}{\theta-1} \right) \sum_{k=0}^{\infty} (\alpha\beta)^k Y_{t+k}^{-1} \times \left[\phi \mu_{t+k} w_{t+k} A_{t+k}^{-\phi} Y_{t+k}^\phi (P_t/P_{t+k})^{-\theta\phi} \right]}{\sum_{k=0}^{\infty} (\alpha\beta)^k (P_t/P_{t+k})^{1-\theta}}. \quad (38)$$

The price index then evolves according to the law of motion,

$$P_t = [(1 - \alpha)p_t^{1-\theta} + \alpha P_{t-1}^{1-\theta}]^{1/(1-\theta)}. \tag{39}$$

Because the relative prices of the firms that do not change their prices in period t fall by the rate of inflation, the law of motion for the measure of price dispersion is

$$\Delta_t = \alpha \Delta_{t-1} \pi_t^{\theta\phi} + (1 - \alpha) (p'_t/P_t)^{-\theta\phi}. \tag{40}$$

6. UO Monetary Policy

Proposition 3 sets out the relevant UO Ramsey problem.

PROPOSITION 3. *The UO Ramsey plan is a choice of state-contingent paths for the endogenous variables $\{\pi_{t+k}, \Delta_{t+k}, p_{t+k}, c_{t+k}, X_{t+k}, Z_{t+k}\}_{k=0}^\infty$ from date t onwards given $\{E_t A_{t+k}, E_t \mu_{t+k}\}_{k=0}^\infty$, so as to maximize social welfare function (41) subject to constraints (42)–(45):*

$$\max E E_t \sum_{k=0}^\infty \beta^k \left(\frac{\log(c_{t+k})}{\phi} - \frac{v}{\phi} \log \Delta_{t+k} - (1 - \tau) \frac{c_{t+k} \Delta_{t+k}}{\mu_{t+k}} \right), \tag{41}$$

subject to the following:

- the Phillips block

$$\left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta\phi - \theta + 1}{1 - \theta}} X_t = \frac{\theta}{\theta - 1} Z_t; \tag{42}$$

$$X_t = 1 + \alpha \beta E_t X_{t+1} \pi_{t+1}^{\theta-1}; \tag{43}$$

$$Z_t = c_t + \alpha \beta E_t Z_{t+1} \pi_{t+1}^{\theta\phi}. \tag{44}$$

- the law of motion of prices

$$\Delta_t = \alpha \Delta_{t-1} \pi_t^{\theta\phi} + (1 - \alpha) \left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta\phi}{\theta-1}}. \tag{45}$$

Here c_t is real marginal cost for the firm which produces output Y_t ,

$$c_t = \phi \frac{\lambda}{1-\tau} \mu_t \Delta_t^v (A_t^{-1} Y_t)^{(v+1)\phi}; \quad (46)$$

discounted marginal revenue is $X_t := E_t \sum_{k=0}^{\infty} (\beta\alpha)^k \left(\frac{P_t}{P_{t+k}}\right)^{1-\theta}$;

and discounted marginal cost is $Z_t := E_t \sum_{k=0}^{\infty} (\beta\alpha)^k c_{t+k} \left(\frac{P_t}{P_{t+k}}\right)^{-\theta\phi}$.

6.1 The Steady State

We now turn in more detail to steady-state analysis. In appendix 1 we solve the policymaker's problem defined in proposition 3 and establish the following result.

PROPOSITION 4. *The steady-state inflation is positive, $\pi \geq 1$. Price stability is only optimal if either $\beta = 1$ or if $\frac{\theta-1}{\theta} \frac{1-\tau}{\mu} = 1$ (which corresponds to the non-distorted steady state). Moreover, π is unique and bounded: $\pi \leq \min(\beta^{1/(\theta-1-\phi\theta)}, \alpha^{-1/(\phi\theta)})$.*

Proof. See appendix 1.

Our result shows that UO policy delivers a different equilibrium inflation than TP optimal policy. It is well known (see Benigno and Woodford 2005) that TP optimal policy requires price stability in the steady state. The UO policy delivers a trend inflation. The intuition follows from the fact that whilst higher inflation induces price-setting firms to choose a higher markup, firms holding prices constant will see their markup erode more quickly; one effect acts to boost demand and the other to reduce it. King and Wolman (1999) show that a slightly positive inflation rate maximizes steady-state welfare by reducing the markup distortion (inverse of marginal cost). In our model the steady-state value of real marginal cost as a function of inflation is obtained from the Phillips relation (42)–(44):

$$c = \frac{1 - \alpha\beta\pi^{\theta\phi}}{1 - \alpha\beta\pi^{\theta-1}} \frac{\theta - 1}{\theta} \left(\frac{1 - \alpha\pi^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta\phi - \theta + 1}{1 - \theta}}.$$

It is easy to see that it increases with inflation at $\pi = 1$, and it follows that

$$\begin{aligned} \frac{d \log c}{d\pi} \pi &= (\theta - 1) \phi \left(\frac{\alpha \pi^{\theta-1} - \alpha \beta \pi^{\theta\phi}}{(1 - \alpha \pi^{\theta-1})(1 - \alpha \beta \pi^{\theta\phi})} \right) \\ &= \frac{(\theta - 1) \phi \alpha (1 - \beta)}{(1 - \alpha)(1 - \alpha \beta)} > 0. \end{aligned}$$

This shows that a small increase in inflation will reduce the price-to-marginal-cost ratio and reduce the monopolistic distortion in the economy. In fact, the following proposition is true:

PROPOSITION 5. *Steady-state inflation increases with the distortion, measured as $1 - \frac{\theta-1}{\theta} \frac{1-\tau}{\mu}$, and declines in the discount factor β and the labor elasticity, v .*

Proof. See appendix 1.

Using parameter values typically found in the literature, we find that the optimal steady-state inflation is of the order of 0.2 percent a year.

So, on the one hand, a small amount of inflation can boost demand, as it partially offsets the markup distortion. On the other hand, price dispersion, which is rising in inflation, acts rather like a cost shock on firms, for reasons analyzed in Damjanovic and Nolan (2010). Hence, one finds that optimal trend inflation has a U-shaped relation to price stickiness, α ; it is increasing in α when initial price dispersion is relatively small, and declines once initial price dispersion is sufficiently large. Optimal inflation declines in the discount factor, β . As discussed DDN (2008) and emphasized earlier, UO policy—in contrast to timeless-perspective policy—gives some weight to the distribution of initial conditions. In particular, it considers the distribution of the initial output gap, which is a time-invariant ergodic distribution imposed by chosen policy. That is partly why some stimulation of output via inflation is desirable. So the smaller the discount factor, the higher is the relative weight on initial conditions and the higher the optimal inflation rate.

6.1.1 Contrast to TP Policy

In the above example, UO policy exploits the Phillips curve to reduce the markup. To understand the difference with the TP approach, consider the following example.

EXAMPLE 2. Assume that the government problem can be formulated in terms of the following social objective:

$$\max E_t \sum_{k=0}^{\infty} \beta^k u(c_t, \pi_t),$$

and a price-setting constraint

$$E_t f(c_t, \pi_t, \pi_{t+1}) = 0, \quad (47)$$

where c_t is real marginal costs, π_t is inflation, and $u(c_t, \pi_t)$ is the period social objective. Social welfare increases with c , $u_c > 0$ and achieves its maximum when prices are stable:

$$u_{\pi}(1) = 0; \quad (48)$$

$u_{\pi\pi}(1) < 0$. The second equation, (47), is a Phillips curve and has the following properties:

$$-\frac{\partial f}{\partial \pi_t} = \beta \frac{\partial f}{\partial \pi_{t+1}}; \quad (49)$$

$$\frac{\partial f}{\partial \pi_t} < 0; \quad \frac{\partial f}{\partial c_t} > 0. \quad (50)$$

Equation (49) reflects the intertemporal trade-off between current and future inflation, while (50) suggests a positive correlation between inflation and marginal cost. Those properties are general and satisfied in Calvo, Rotemberg, and discounted Taylor frameworks.

The UO program is set out and solved:

$$\begin{aligned} H^{UO} &= u(c, \pi) - \lambda f(c, \pi, \pi); \\ \frac{\partial H^{UO}}{\partial c} &= \frac{\partial u}{\partial c} - \lambda \frac{\partial f}{\partial c} = 0; \end{aligned} \quad (51)$$

$$\frac{\partial H^{UO}}{\partial \pi} = \frac{\partial u}{\partial \pi} - \lambda \left(\frac{\partial f}{\partial \pi_t} + \frac{\partial f}{\partial \pi_{t+1}} \right) = 0. \quad (52)$$

Since $\frac{\partial u}{\partial c} > 0$, and $\frac{\partial f}{\partial c} > 0$, the Lagrange multiplier, λ , is positive and the last equation can be rewritten as

$$\frac{\partial H^{UO}}{\partial \pi} = \frac{\partial u}{\partial \pi} + \lambda(1 - \beta) \frac{\partial f}{\partial \pi_{t+1}} = 0.$$

Therefore, in equilibrium there is a positive inflation, $\pi > 1$, since $\frac{\partial u}{\partial \pi} < 0$.

Unconditional optimization uses positive inflation to reduce the markup, which is equivalent to increasing steady-state marginal cost. The implicit function theorem is applied to the price-setting equation to show that

$$\frac{dc}{d\pi} = - \left(\frac{\partial f}{\partial \pi_t} + \frac{\partial f}{\partial \pi_{t+1}} \right) / \frac{\partial f}{\partial c_t} = (1 - \beta) \frac{\partial f}{\partial \pi_{t+1}} / \frac{\partial f}{\partial c_t} > 0,$$

and this is exploited in order to improve welfare.

The TP approach, on the other hand, ignores the relation between inflation and marginal costs in the Phillips curve and does not use inflation to stimulate the economy by reducing the inefficient markup. The TP program and solution is

$$\begin{aligned} L^{TP} &= E_t \sum_{k=0}^{\infty} \beta^k (u(c_{t+k}, \pi_{t+k}) - \lambda_{t+k} E_t f(c_{t+k}, \pi_{t+k}, \pi_{t+1+k})); \\ \frac{\partial L^{TP}}{\partial c_t} &= \frac{\partial u}{\partial c} - \lambda_t \frac{\partial f}{\partial c} = 0; \\ \frac{\partial L^{TP}}{\partial \pi_{t+1}} &= \frac{\partial u}{\partial \pi} - \lambda_{t+1} \beta \frac{\partial f}{\partial \pi_t} - \lambda_t \frac{\partial f}{\partial \pi_{t+1}} = 0. \end{aligned} \quad (53)$$

Combining (49) with (53) in steady state shows that TP policy chooses inflation to optimize intertemporal utility in the steady state:

$$\frac{\partial u}{\partial \pi} = 0.$$

It is apparent that the effect of inflation on the reduction of the price markup is not utilized when TP optimization is applied.

Finally, we note that price stability only occurs under the TP if conditions (48) and (49) are satisfied when the price setting is one period forward looking. In a more general case, if the Phillips relation should satisfy $\sum_{k=-\infty}^{\infty} \beta^{-k} \frac{\partial f}{\partial \pi_{t+1}} = 0$, that may not be so for a wide range of models. For example, it may not be the case for an open economy (see Benigno and Lopez-Salido 2006).

6.2 The Quadratic Form

Having recovered the optimal steady state, one can obtain a quadratic loss function of the form (19). The quadratic welfare derived from (41) and constraints (42)–(45) is

$$EU = -\frac{1}{2}E \left(h\Phi \widehat{u}_t^2 + (1 - h\Phi) \widehat{c}_t^2 + \Lambda_x \widehat{X}_t^2 + \Lambda_\pi \widehat{\pi}_t^2 + \Lambda_\Delta \widehat{\Delta}_t^2 \right), \quad (54)$$

where $\widehat{u}_t = \widehat{c}_t + \widehat{\Delta}_t - \widehat{\mu}_t$ is the log-linearized disutility of labor, $u_t = -\frac{\lambda N_t^{v+1}}{v+1}$.

For the model at hand, one can show that it can be written as follows:

$$EU = -\frac{1}{2}E \left[\phi(1+v) \left(\widehat{Y}_t - \widehat{Y}_t^* \right)^2 + G \widehat{g}_t^2 + \Lambda_x \widehat{X}_t^2 + \Lambda_\Delta \widehat{\Delta}_t^2 + \Lambda_\pi \widehat{\pi}_t^2 \right], \quad (55)$$

where $\widehat{g}_t = \widehat{\mu}_t - \widehat{\Delta}_t$. \widehat{g}_t has an intuitive interpretation as the log-deviation of the ratio of natural output to labor cost. To see that, note that if prices were flexible ($\alpha = 0$), the price-setting condition (38) will result in the following level of output:

$$Y_t^n = \frac{\theta}{\theta - 1} \frac{\mu_t}{\Delta_t} \phi w_t N_t. \quad (56)$$

Therefore, if we define $g_t := \frac{Y_t^n}{w_t N_t} = \phi \frac{\theta}{\theta - 1} \frac{\mu_t}{\Delta_t}$, that becomes a relationship between the log-deviation of the cost-push shock and relative price dispersion.

The term \widehat{Y}_t^* represents the “target” level of output $Y_t^* := \widehat{A}_t - v_\mu \widehat{\mu}_t - v_\Delta \widehat{\Delta}_t$. Details concerning coefficients are given in appendix 2. The “target” rate is increasing in productivity and declining in

the cost-push shock; it is also declining in price dispersion. The variable \widehat{X}_t represents the losses to firms forced to charge sub-optimal prices due to price stickiness and expected inflation, to which they may not be able to react.

This form of the loss function can easily be nested to familiar cases, either the non-distorted steady state where $\Phi = 1$, or where the steady state of the model economy remains distorted but where the social discount rate is equal to the private rate of discount, $\beta = 1$ (in which case the UO policy and the timeless-perspective policies coincide). In both special cases, optimal monetary policy corresponds to price stability, and the loss function (55) reduces to a familiar form defined simply over inflation and output. Specifically, if the optimal steady state is characterized by price stability, then $\Lambda_x = 0$. Moreover, one can easily show that price dispersion, $\widehat{\Delta}_t$, is a second-order term in that case. Lastly, the labor wedge \widehat{g}_t is then simply a cost-push shock, $\widehat{\mu}_t$, and can be considered as a term independent of policy.

7. Application: Unconditional Ordering of Simple Rules

The foregoing approach is easily used to evaluate simple rules for monetary policy and to highlight the potential significance for policy design of a distorted steady state. First, write the model in vector autoregressive form as follows:

$$E_t \widehat{\pi}_{t+1} + E_t \widehat{Y}_{t+1} = \widehat{Y}_t + (1 - \beta) \widehat{i}_t; \quad (57)$$

$$\beta \alpha \pi^{\theta \phi} E_t \left(\widehat{Z}_{t+1} + \theta \phi \widehat{\pi}_{t+1} \right) = \widehat{Z}_t - (1 - \alpha \beta \pi^{\theta \phi}) \widehat{c}_t; \quad (58)$$

$$\beta \alpha \pi^{\theta - 1} E_t \left(\widehat{X}_{t+1} + (\theta - 1) \widehat{\pi}_{t+1} \right) = \widehat{X}_t; \quad (59)$$

$$\widehat{Z}_t - (\theta \phi - \theta + 1) \frac{\alpha \pi^{\theta - 1}}{1 - \alpha \pi^{\theta - 1}} \widehat{\pi}_t - \widehat{X}_t = 0; \quad (60)$$

$$-\widehat{g}_t + \widehat{\Delta}_t - \widehat{\mu}_t = 0; \quad (61)$$

$$-\widehat{c}_t + v \widehat{\Delta}_t + (v + 1) \phi (\widehat{Y}_t - \widehat{A}_t) + \widehat{\mu}_t = 0; \quad (62)$$

$$\widehat{\Delta}_{t+1} + \theta \phi \frac{\alpha \pi^{\theta - 1} - \alpha \pi^{\theta \phi}}{1 - \alpha \pi^{\theta - 1}} \widehat{\pi}_{t+1} = \alpha \pi^{\theta \phi} \widehat{\Delta}_t; \quad (63)$$

$$-Y_t^* + \widehat{A}_t - v_\mu \widehat{\mu}_t - v_\Delta \widehat{\Delta}_t = 0; \quad (64)$$

$$-\widehat{i}_t = \phi_\pi \widehat{\pi}_t + \phi_f \widehat{f}_t + m_t. \quad (65)$$

In the above linearized system of equations, the final equation (65) is the policy rule, where \widehat{i}_t is the gross nominal interest rate, $\widehat{i}_t = \log\left(\frac{\beta}{\pi}(1+i_t)\right)$, and \widehat{f}_t represents a linear combination of policy feedback variables, while m_t is a policy shock.¹⁴

It is clear that steady-state distortions complicate the policy problem so far as the policymaker's objective function is concerned.¹⁵ However, does it make any difference so far as the design of simple rules are concerned?¹⁶

First, a simple interest rate feedback rule is considered, where the interest rate responds to current and lagged inflation only. The feedback on current inflation is fixed at $\phi_\pi = 1.5$. Given this, the optimized weight on lagged inflation, $f = \widehat{\pi}_{t-1}$, is computed. It is about 15 in the distorted steady-state case and 14 for the non-distorted steady state. However, the difference in welfare between responding and not responding to lagged inflation is up to 0.16 percentage points in terms of consumption-equivalent units (see the top-right graph in figure 1; ϕ_f is at its optimal value). As in the TP approach, relative price distortion is very costly, and the optimal simple rule may be very close to price stability ($\phi_\pi = +\infty$). However, if for any reason the policy reaction on current inflation is restricted, the economy may significantly benefit from a response to lagged inflation.

One can also show that the optimal feedback on output should be slightly *negative*, $\phi_f = -0.015$. Furthermore, inclusion of real output targeting leads to very modest welfare improvements, in the

¹⁴The following parameterization is used in the quantitative investigation: $\beta = 0.9$, $v = 1.1$, $\theta = 7$, $\alpha = 0.5$, and $\phi = 1.3$. It is assumed that shocks A_t , μ_t , and m_t follow *AR*(1) processes with $\rho_A = 0.98$, $\sigma_A = 0.008$, $\rho_m = 0.9$, $\sigma_m = 0.005$, and $\rho_\mu = 0.9$, $\sigma_\mu = 0.02$.

¹⁵That is, complicates it relative to the objective function in the non-distorted case.

¹⁶In the particular model developed above, the UO trend inflation is rather small and the policy ordering across distorted and non-distorted steady states is often the same for given simple rules. However, in simulations not reported, it was possible to find simple, plausible rules that result in welfare "reversals"; that is, where rule 1 welfare dominates rule 2 in the distorted economy, but where the ranking switched in the non-distorted economy.

order of 10^{-3} compared with targeting inflation alone. This result is consistent with Schmitt-Grohe and Uribe (2007).

The results are summed up in figure 1 (where the broken line is the non-distorted economy).

7.0.1 Targeting Nominal Income Growth

Finally, inflation targeting and nominal income targeting are compared under a UO policy criterion as in Kim and Henderson (2005). Kim and Henderson suggest, in a model with one-period price stickiness, that nominal income targeting may have superior welfare properties to inflation targeting. Two rules are compared:

$$\text{Nominal income growth targeting: } i_t = 0.05(y_t - y_{t-1} + \pi_t) + (\phi_\pi - 0.05)\pi_t + m_t; \quad (66)$$

$$\text{Inflation targeting: } i_t = \phi_\pi \pi_t + m_t. \quad (67)$$

In the case of a non-distorted steady state and a “low” feedback on inflation, the findings are similar to some of Kim and Henderson’s findings. Specifically, in the case of a distorted steady-state model, the net welfare gain from targeting nominal income growth over inflation targeting is positive. In the non-distorted case, inflation targeting is rarely dominated by nominal income targeting. In figure 2, the relative welfare gain (over inflation targeting) in targeting nominal income growth is plotted against ϕ_π .

The precise position of these net welfare schedules is quite sensitive to parameterization of the model (in particular, the persistence of shocks), but in general one finds that as the feedback on inflation rises, inflation targeting is likely to dominate nominal income targeting.

7.1 Comparing Optimal Policies

In this section we compare impulse responses when UO and TP policies are applied. The linear approximation to UO policy looks rather complex. It has three predetermined variables; besides inflation, there is price dispersion Δ_t and the Lagrange multiplier associated with one of the dynamic constraints (43). The complete system is presented in appendix 2. The TP policy system breaks naturally

Figure 1. Sub-Optimal Simple Policies

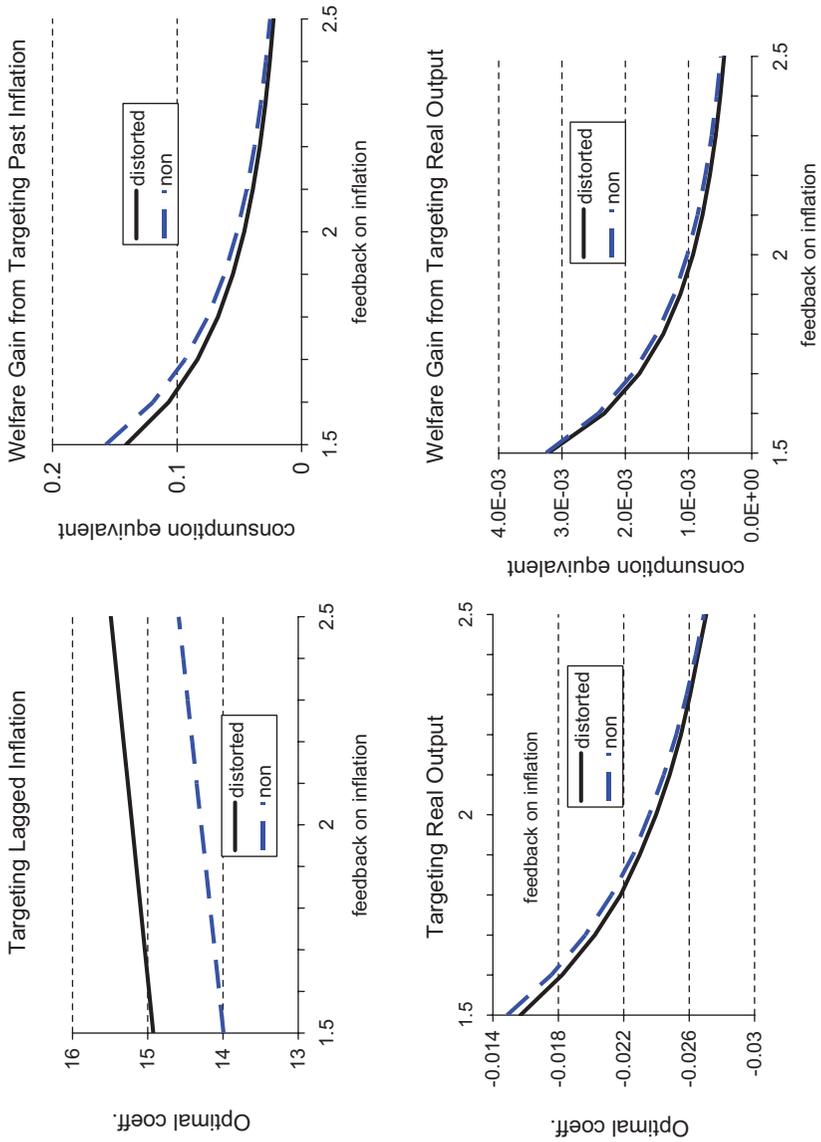
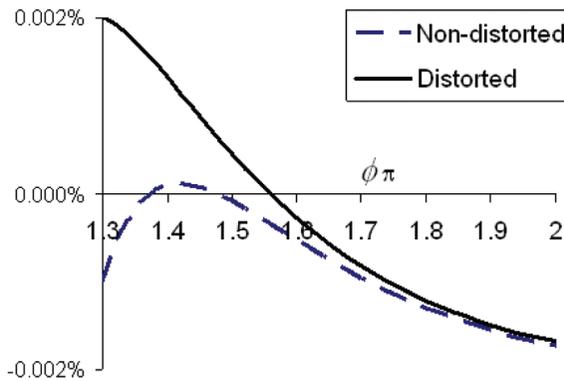


Figure 2. Relative Welfare Gain in Targeting Nominal Income Growth



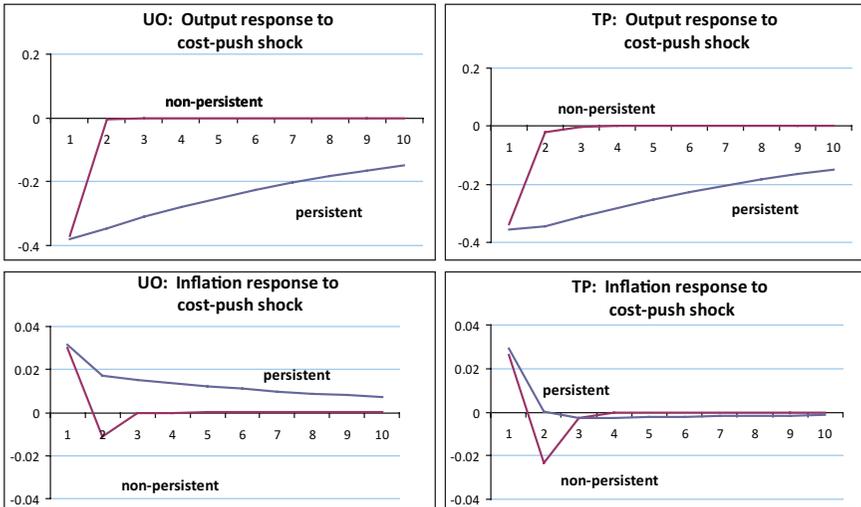
into blocks, and the dynamics for inflation and marginal cost can be solved from these two-equation blocks as shown in appendix 2. However, due to the fact that β is very close to 1, we do not expect to see quantitatively a large difference across TP and UO policy.

Note that $\{(41), (42)-(45)\}$ does not include the productivity shock. Hence, one may conclude that neither UO nor TP policy should react to such shocks and that price/inflation stability will be optimal under both policies. Below we report the impulse responses from TP and UO policies following a cost-push shock. We ran two experiments, one with persistent ($\rho_\mu = 0.9$) and one with non-persistent ($\rho_\mu = 0$) shocks.¹⁷

Strict inflation targeting is sub-optimal under both approaches when a markup shock hits the economy. When the shock has low persistence, one observes overshooting in inflation, although this is absent when the shock is persistent. Perhaps the key point to observe is that when the shock is persistent, UO policy implies a slow rate of convergence of inflation.

Not surprisingly, perhaps, for the baseline New Keynesian model we find that the differences between TP and UO policies are small. In figure 3, they generate very similar responses in output to a cost-push shock, but slightly different responses in inflation. When the

¹⁷Let ρ_μ denote the autocorrelation coefficient; then the shock process that is posited is described by $\mu_t = \rho_\mu\mu_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim N(0, 1)$ is an i.i.d. shock.

Figure 3. UO and TP-Optimal Policies

shock is persistent, it affects the economy for a longer period of time, and the difference in social time discounting becomes more important in that case.

Finally, one observes that the price level returns to its initial value after the shocks under TP policy but not following UO policy. Thus, UO policy induces infinite variability in the price level (although inflation is stationary), whilst TP gives a stationary price.¹⁸

8. Conclusion

In this paper we analyze UO policies and compare them with TP policies. UO policy in spirit is actually very similar to the definition given when TP policy was introduced in Woodford (1999a, p. 19): “under the timeless perspective, one chooses to act as one believes one would have wished to commit oneself to act at a date far in the past.” The TP acknowledges that commitment results in more favorable expectations (which are considered, in effect, as state variables).

¹⁸We thank the editor for highlighting this finding in our results.

But whilst the TP recognizes that expectations can be changed by a particular policy implementation and that that effect should be internalized, it rejects the notion that policy affects the distribution of the other state variables (like capital) and that that influence ought to be internalized also. So we think of the UO perspective pushing the original intuition for the TP one step further: One chooses to act as one believes one would have wished to commit oneself to act at a date far in the past not only with respect to private expectations but with respect to other state variables like capital.

The paper also demonstrates that one can formulate a purely quadratic approximate unconditional loss function to a model economy with a distorted steady state. It develops a straightforward, efficient approach to implementing the UO approach. It contrasts this approach to policy formulation with the TP approach, giving a number of examples where policies and objectives differ. It explores reasons why one may be interested in pursuing UO policies and the difficulties so encountered. In an application, it is shown that the loss function may be somewhat more complex than in a model with no steady-state distortions; inflation and output are no longer the sole arguments in the loss function. However, the loss function so obtained is easily interpreted in terms of the underlying distortions in the economy.

Appendix 1. A Second-Order Approximation of the Welfare Function

The first part of this appendix demonstrates the key result in section 4, namely the existence of the quadratic form, (19). The first line of the following block of equations corresponds to the top line of (19), the subsequent lines being its quadratic approximation:

$$\begin{aligned}
 El(x_t) &= E[l(x_t) + \xi F(y_t, x_t, \mu_t)] \\
 &= E\left(l + X \frac{\partial l}{\partial x} \hat{x}_t + \frac{1}{2} \left(X^2 \frac{\partial^2 l}{\partial x^2} + X \frac{\partial l}{\partial x} \right) \hat{x}_t \hat{x}_t \right) \\
 &\quad + E\xi \left(F + X \frac{\partial F}{\partial x} \hat{x}_t + X \frac{\partial F}{\partial y} \hat{y}_t + \mu \frac{\partial F}{\partial \mu} \hat{\mu}_t \right) \\
 &\quad + \frac{1}{2} \xi \left(X \frac{\partial F}{\partial x} + X^2 \frac{\partial^2 F}{\partial x^2} \right) E \hat{x}_t \hat{x}_t + \frac{1}{2} \xi \left(X \frac{\partial F}{\partial y} + X X \frac{\partial F}{\partial y^2} \right)
 \end{aligned}$$

$$\begin{aligned}
& \times E\hat{y}_t\hat{y}_t + \frac{1}{2}\xi \left(\mu \frac{\partial F}{\partial x} + \mu^2 \frac{\partial^2 F}{\partial x^2} \right) E\hat{\mu}_t\hat{\mu}_t \\
& + \xi E \left(XX \frac{\partial F}{\partial x \partial y} \hat{x}_t\hat{y}_t + X\mu \frac{\partial^2 F}{\partial x \partial \mu} \hat{x}_t\hat{\mu}_t + X\mu \frac{\partial F}{\partial y \partial \mu} \hat{y}_t\hat{\mu}_t \right) \\
& + O3.
\end{aligned}$$

Using the constraints $E_t x_{t+1} = y_t$, and the property of unconditional expectations that $Ez_{t+1} = Ez_t$, this can be rewritten as

$$\begin{aligned}
El(x_t) &= XE\hat{x}_t \left(\frac{\partial l}{\partial x} + \xi \frac{\partial F}{\partial x} + \xi \frac{\partial F}{\partial y} \right) \\
&+ \frac{1}{2}XE\hat{x}_t\hat{x}_t \left(\frac{\partial l}{\partial x} + \xi \frac{\partial F}{\partial x} + \xi \frac{\partial F}{\partial y} \right) + EQ_l + \xi EQ_F \quad (68) \\
&+ l + \xi F + \xi \mu \frac{\partial F}{\partial \mu} E\hat{\mu}_t + \frac{1}{2}\xi \left(\mu \frac{\partial F}{\partial x} + \mu^2 \frac{\partial^2 F}{\partial x^2} \right) E\hat{\mu}_t\hat{\mu}_t \\
&+ O3. \quad (69)
\end{aligned}$$

Here Q_l and Q_F are pure second-order terms:

$$\begin{aligned}
Q_l &= \frac{1}{2}X^2 \frac{\partial^2 l}{\partial x^2} \hat{x}_t\hat{x}_t; \\
Q_F &= \frac{1}{2}X^2 \left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} \right) \hat{x}_t\hat{x}_t + XX \frac{\partial^2 F}{\partial x \partial y} \hat{x}_t\hat{x}_{t+1} + X\mu \frac{\partial^2 F}{\partial x \partial \mu} \hat{x}_t\hat{\mu}_t \\
&+ X\mu \frac{\partial^2 F}{\partial y \partial \mu} \hat{x}_{t+1}\hat{\mu}_t.
\end{aligned}$$

Furthermore, using the steady-state conditions (11), one can show that the first line of expression (68) equals zero. Moreover, expression (69) consists of $l + \xi F = l$, the steady-state value of the loss function and shocks. These are terms independent of policy (t.i.p.). Thus, it is proved that the loss function can be represented in a pure quadratic form.

$$El(x_t) = EQ_l + \xi EQ_F + t.i.p + O3.$$

UO Policy in a Distorted Calvo Model

In this section we apply the algorithm designed in the main text to the Calvo model summarized in section 6. One can set up

the Hamiltonian for this problem, as proposed in section 2.1.1, as follows:

$$\begin{aligned}
 H = & \left(\log(c_t) - v \log \Delta_t - \Delta_t \frac{\theta}{\theta - 1} \Phi c_t \right) \\
 & + \rho_t (X_t - 1 - \beta \alpha \pi_{t+1}^{\theta-1} X_{t+1}) \\
 & + \varphi_t \left(-Z_t + c_t + \alpha \beta E_t Z_{t+1} \pi_{t+1}^{\theta \phi} \right) \\
 & + \xi_t \left(\left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta \phi - \theta + 1}{1 - \theta}} X_t - \frac{\theta}{\theta - 1} Z_t \right) \\
 & + \eta_t \left(\Delta_t - \alpha \Delta_{t-1} \pi_t^{\theta \phi} - (1 - \alpha) \left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta \phi}{\theta - 1}} \right).
 \end{aligned}$$

The necessary conditions for an optimum include

$$\begin{aligned}
 \frac{\partial H}{\partial c_t} &= \frac{1}{c_t} - \frac{\theta}{\theta - 1} \Phi \Delta_t + \varphi_t; \\
 \frac{\partial H}{\partial \Delta_t} &= -\frac{v}{\Delta_t} - \frac{\theta}{\theta - 1} \Phi c_t + \eta_t - E_t \alpha \eta_{t+1} \pi_{t+1}^{\theta \phi}; \\
 \frac{\partial H}{\partial X_t} &= \rho_t - \rho_{t-1} \beta \alpha \pi_t^{\theta-1} + \xi_t \left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta \phi - \theta + 1}{1 - \theta}}; \\
 \frac{\partial H}{\partial Z_t} &= -\varphi_t + \varphi_{t-1} \beta \alpha \pi_t^{\theta \phi} - \xi_t \frac{\theta}{\theta - 1}; \\
 \pi_t \frac{\partial H}{\partial \pi_t} &= -(\theta - 1) \rho_{t-1} \beta \alpha \pi_t^{\theta-1} X_t + \varphi_{t-1} \beta \alpha \theta \phi \pi_t^{\theta \phi} Z_t \\
 & + \xi_t X_t (\theta \phi - \theta + 1) \left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta \phi}{1 - \theta}} \frac{\alpha}{1 - \alpha} \pi_t^{\theta-1} \\
 & - \eta_t \alpha \theta \phi \Delta_{t-1} \pi_t^{\theta \phi} + \eta_t \theta \phi \alpha \pi_t^{\theta-1} \left(\frac{1 - \alpha \pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta \phi - \theta + 1}{\theta - 1}}.
 \end{aligned} \tag{70}$$

Optimal Steady State

The value of the endogenous variables in steady state should solve the system of constraints (42)–(45) and the first-order conditions, (70). As a result, one obtains the following steady-state equations:

$$\begin{aligned}
 X &= \frac{1}{1-\beta\alpha\pi^{\theta-1}}; & [\Phi\Delta - \varphi^{\frac{\theta-1}{\theta}}] \frac{\theta}{\theta-1}c &= 1; \\
 Z &= \frac{\theta-1}{\theta}X \left(\frac{1-\alpha\pi_t^{\theta-1}}{1-\alpha} \right)^{\frac{\theta\phi-\theta+1}{1-\theta}}; & \eta\Delta(1-\alpha\pi^{\theta\phi}) &= \left(v + \frac{\theta}{\theta-1}\Phi\Delta c \right); \\
 c &= (1-\alpha\beta\pi^{\theta\phi})Z; & \xi &= -\varphi^{\frac{\theta-1}{\theta}}(1-\alpha\beta\pi^{\theta\phi}); \\
 \Delta &= \left(\frac{1-\alpha}{1-\alpha\pi^{\theta\phi}} \right) \left(\frac{1-\alpha\pi_t^{\theta-1}}{1-\alpha} \right)^{\frac{\theta\phi}{\theta-1}}; & \rho &= -\xi \left(\frac{1-\alpha\pi_t^{\theta-1}}{1-\alpha} \right)^{\frac{\theta\phi-\theta+1}{1-\theta}} X.
 \end{aligned}
 \tag{71}$$

For further convenience, we will compute the steady-state value of the Lagrange multipliers

$$\begin{aligned}
 \rho &= \Phi h - 1 < 0; \\
 \eta\Delta(1-\alpha\pi^{\theta\phi}) &= v + \Phi h,
 \end{aligned}$$

where we defined $h := \frac{1-\alpha\pi^{\theta-1}}{1-\alpha\pi^{\theta\phi}} \frac{1-\alpha\beta\pi^{\theta\phi}}{1-\beta\alpha\pi^{\theta-1}}$. The last first-order condition (70) provides an equation to find optimal inflation,

$$\begin{aligned}
 0 &= [1 - \Phi h] \phi \theta \left(\frac{\beta\alpha\pi^{\theta\phi}}{(1-\alpha\beta\pi^{\theta\phi})} - \frac{\alpha\pi^{\theta-1}}{1-\alpha\pi^{\theta-1}} \right) \\
 &+ [1 - \Phi h] (\theta - 1) \left(\frac{\alpha\pi_t^{\theta-1}}{1-\alpha\pi_t^{\theta-1}} - \frac{\beta\alpha\pi^{\theta-1}}{1-\beta\alpha\pi^{\theta-1}} \right) \\
 &+ [v + \Phi h] \theta \phi \left(\frac{\alpha\pi^{\theta-1}}{1-\alpha\pi^{\theta-1}} - \frac{\alpha\pi^{\theta\phi}}{(1-\alpha\pi^{\theta\phi})} \right).
 \end{aligned}
 \tag{72}$$

Using these equations, one can infer certain properties of the optimal steady-state inflation rate.

Proof of Proposition 4: Existence

One may rewrite (72) as (73):

$$F(\pi) = v g(\pi) + [g(\pi) - f(\pi)] + \Phi h(\pi) f(\pi) = 0, \tag{73}$$

where $h(\pi) = \frac{1-\alpha\beta\pi^{\theta\phi}}{1-\alpha\beta\pi^{\theta-1}} \frac{1-\alpha\pi^{\theta-1}}{1-\alpha\pi^{\theta\phi}} > 0$, and $g(\pi) = \frac{\theta\phi}{1-\alpha\pi^{\theta\phi}} - \frac{\theta\phi}{1-\alpha\pi^{\theta-1}}$;
 $f(\pi) = \left[\frac{\theta\phi}{1-\alpha\pi^{\theta\phi}} - \frac{\theta\phi}{1-\alpha\beta\pi^{\theta\phi}} \right] - \left[\frac{\theta-1}{1-\alpha\pi^{\theta-1}} - \frac{\theta-1}{1-\beta\alpha\pi^{\theta-1}} \right]$; $\Phi = \frac{\theta-1}{\theta} \frac{1-\tau}{\mu}$.

It is easy to see that $g(1) = 0$; $h(1) = 1$ and $f(1) = \frac{(\theta\phi-\theta+1)\alpha(1-\beta)}{(1-\alpha)(1-\alpha\beta)} > 0$, which implies that $F(1) = -(1-\Phi) \frac{(\theta\phi-\theta+1)\alpha(1-\beta)}{(1-\alpha)(1-\alpha\beta)} \leq 0$. The strict equality obtains in three cases only: first, when prices are flexible, $\alpha = 0$; second, when the future is not discounted by firms, $\beta = 1$; and finally, when there are no distortions in steady state, $\Phi = 1$.

Define $\pi_h = \alpha^{-1/(\phi\theta)}$ and note that the functions g , f , and h are defined on an interval $[1, \pi_h)$. The difference $[g(\pi_h) - f(\pi_h)]$ is bounded while $g(\pi)$, $h(\pi)$, and $f(\pi)$ tend to positive infinity as π approaches π_h . Hence, $\lim_{\pi \rightarrow \pi_h} F(\pi) = +\infty$. Since $F(\pi)$ is a continuous function, one can conclude that there is a solution to (73) on the interval $[1, \alpha^{-1/(\phi\theta)})$. One may easily show then that if $\pi_m = \beta^{1/(\theta-1-\phi\theta)}$, then it follows that $F(\pi_m) > 0$, since $g(\pi_m) - f(\pi_m) > 0$. Therefore, optimal inflation is smaller than π_m .

Proof of Proposition 4: Uniqueness

The proof is by contradiction. First it is proved that if $\beta < 1$, for any $\pi_1 < \pi_m$ such that $F(\pi_1) = 0$, it is necessary that $F'(\pi_1) > 0$. By direct differentiation, it follows that

$$F'(\pi_1) = (v + 1)g'(\pi_1) + (\Phi h(\pi_1) - 1)f'(\pi_1) + \Phi h'(\pi_1)f(\pi_1).$$

Moreover, since $F(\pi_1) = 0$, it follows that $\Phi h(\pi_1) - 1 = -(v + 1)g(\pi_1)/f(\pi_1)$. Therefore,

$$F'(\pi_1) = \frac{(v + 1)}{f(\pi_1)} [g'(\pi_1)f(\pi_1) - f'(\pi_1)g(\pi_1)] + \Phi h'(\pi_1)f(\pi_1),$$

and it is easy to show that for any $\pi_1 < \pi_m$, $g'(\pi_1)f(\pi_1) - f'(\pi_1)g(\pi_1) > 0$, and therefore $F'(\pi_1)$ is positive.

Since F is continuously differentiable, if a solution of (73) is not unique, there will be at least one solution such that $F'(\pi_1) \leq 0$. It has been demonstrated that such a solution is impossible and the necessary contradiction is obtained.

Proof of Proposition 5

By the implicit function theorem, one concludes that $\frac{d\pi}{d\Phi} = -\frac{\partial F}{\partial \pi} / \frac{\partial F}{\partial \Phi}$. From the proof of proposition 4 (uniqueness), we know that $\frac{\partial F}{\partial \pi} > 0$, while $\frac{\partial F}{\partial \Phi} = h(\pi)f(\pi) > 0$. Therefore $\frac{d\pi}{d\Phi} < 0$, and equilibrium inflation increases with steady-state distortions, measured as $1 - \Phi$.

Similarly, $\frac{d\pi}{dv} = -\frac{\partial F}{\partial \pi} / \frac{\partial F}{\partial v}$, where $\frac{\partial F}{\partial v} = g(\pi) > 0$ for $\pi > 1$, therefore $\frac{d\pi}{dv} < 0$, and optimal inflation declines with the elasticity of labor.

Moreover, $\frac{d\pi}{d\beta} = -\frac{\partial F}{\partial \pi} / \frac{\partial F}{\partial \beta}$, where $\frac{\partial F}{\partial \beta} = -(1 - \Phi h(\pi)) \frac{\partial f}{\partial \beta} + \Phi f(\pi) \frac{\partial h}{\partial \beta}$, and one may prove by direct differentiation that $\frac{\partial f}{\partial \beta} < 0$, $\frac{\partial \ln h}{\partial \beta} > 0$, and $(1 - \Phi h(\pi)) = (v + 1)g(\pi)/f(\pi) > 0$. Therefore, $\frac{\partial F}{\partial \beta} > 0$, and $\frac{d\pi}{d\beta} < 0$.

Finally, it is worth noting that steady-state inflation can both increase or decrease in price stickiness, since the sign of $\frac{\partial F}{\partial \alpha}$ may be positive or negative.

Appendix 2. The Second-Order Approximation to Unconditional Welfare

In section 6.2 of the main text, we asserted the existence of the following quadratic equation:

$$EU = E(Q_l + \rho Q_X + \varphi Q_Z + \xi Q_{ZX} + \eta Q_\Delta),$$

where Q_l is the second-order term of the loss function and Q_X , Q_Z , Q_{ZX} , Q_Δ , and Q_p are the second-order terms of the log-linear approximation to constraints in the above Hamiltonian. We outline here the key manipulations required to derive these expressions. A more detailed appendix is available upon request. For optimal π , they can be written as

$$Q_l = -\frac{1}{2}\widehat{c}_t^2 + \frac{1}{2}v\widehat{\Delta}_t^2 - \frac{1}{2}\frac{\theta}{\theta - 1}\Phi c\Delta \left(2\widehat{c}_t\widehat{\Delta}_t - 2\widehat{c}_t\widehat{\mu}_t - 2\widehat{\mu}_t\widehat{\Delta}_t \right) + tip; \tag{74}$$

$$Q_x = -(\theta - 1)\beta\alpha\pi^{\theta - 1}X \left[\widehat{X}_t\widehat{\pi}_t + \frac{1}{2}(\theta - 2)\widehat{\pi}_t^2 \right]; \tag{75}$$

$$Q_z = \theta\phi\beta\alpha\pi^{\theta\phi} Z \left(\widehat{Z}_t \widehat{\pi}_t + \frac{1}{2} (\theta\phi - 1) \widehat{\pi}_t^2 \right); \tag{76}$$

$$\begin{aligned} Q_{zx} &= (\theta\phi - \theta + 1) X \left(\frac{1 - \alpha\pi^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta\phi}{1-\theta}} \frac{\alpha\pi^{\theta-1}}{1 - \alpha} \widehat{X}_t \widehat{\pi}_t \\ &\quad + \frac{1}{2} X \frac{(\theta\phi - \theta + 1) \alpha\pi^{\theta-1}}{1 - \alpha} \left(\frac{1 - \alpha\pi^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta\phi}{1-\theta}} \\ &\quad \times \left[\frac{\theta\phi\alpha\pi^{\theta-1}}{1 - \alpha\pi^{\theta-1}} + \theta - 2 \right] \widehat{\pi}_t^2; \end{aligned} \tag{77}$$

$$\begin{aligned} Q_{\Delta} &= -\theta\phi\alpha\Delta\pi^{\theta\phi} \left(\widehat{\Delta}_{t-1} \widehat{\pi}_t + (\theta\phi - 1) \frac{1}{2} \widehat{\pi}_t^2 \right) \theta\phi\alpha\pi^{\theta-1} \left(\frac{1 - \alpha\pi_t^{\theta-1}}{1 - \alpha} \right)^{\frac{\theta\phi}{\theta-1}-1} \\ &\quad \times \left(-\frac{(\theta\phi - \theta + 1) \alpha\pi^{\theta-1}}{1 - \alpha\pi_t^{\theta-1}} + \theta - 2 \right) \frac{1}{2} \widehat{\pi}_t^2. \end{aligned} \tag{78}$$

The linear relations are

$$\widehat{X}_t - \beta\alpha\pi^{\theta-1} \left(\widehat{X}_{t+1} + (\theta - 1) \widehat{\pi}_{t+1} \right) = O2; \tag{79}$$

$$\widehat{Z}_t - (1 - \alpha\beta\pi^{\theta\phi}) \widehat{c}_t - \beta\alpha\pi^{\theta\phi} \left(\widehat{Z}_{t+1} + \theta\phi\widehat{\pi}_{t+1} \right) = O2; \tag{80}$$

$$\widehat{Z}_t - \left(\widehat{X}_t + (\theta\phi - \theta + 1) \frac{\alpha\pi^{\theta-1}}{1 - \alpha\pi_t^{\theta-1}} \widehat{\pi}_t \right) = O2; \tag{81}$$

$$-\widehat{\Delta}_t + \alpha\pi^{\theta\phi} \widehat{\Delta}_{t-1} + \theta\phi \left(\frac{\alpha\pi^{\theta\phi} - \alpha\pi^{\theta-1}}{1 - \alpha\pi^{\theta-1}} \right) \widehat{\pi}_t = O2. \tag{82}$$

Simplification of Q_{Δ}

Use (82) to derive an expression for $\widehat{\Delta}_t^2$. From that we find an expression for the cross term $\widehat{\pi}_t \widehat{\Delta}_{t-1}$:

$$\begin{aligned} E\alpha\pi^{\theta\phi}\theta\phi\widehat{\pi}_t\widehat{\Delta}_{t-1} &= E\frac{1}{2} \left(\frac{\alpha\pi^{\theta\phi} - \alpha\pi^{\theta-1}}{1 - \alpha\pi^{\theta-1}} \right)^{-1} \left(1 - (\alpha\pi^{\theta\phi})^2 \right) \widehat{\Delta}_t^2 \\ &\quad - \frac{1}{2} (\theta\phi)^2 \left(\frac{\alpha\pi^{\theta\phi} - \alpha\pi^{\theta-1}}{1 - \alpha\pi^{\theta-1}} \right) \widehat{\pi}_t^2. \end{aligned}$$

Then, substitute that into (78) to conclude that

$$\begin{aligned}
 EQ_{\Delta}\eta = & -\frac{1}{2} \frac{(1 - \alpha\pi^{\theta-1})(1 + \alpha\pi^{\theta\phi})}{\alpha\pi^{\theta\phi} - \alpha\pi^{\theta-1}} (v + \Phi h) E\widehat{\Delta}_t^2 + \theta\phi \frac{(v + \Phi h)}{1 - \alpha\pi^{\theta-1}} \\
 & \times \left[(\theta - 1 - \theta\phi) \frac{\alpha\pi^{\theta-1}}{1 - \alpha\pi^{\theta-1}} + \frac{\alpha\pi^{\theta\phi} - \alpha\pi^{\theta-1}}{1 - \alpha\pi^{\theta\phi}} \right] \frac{1}{2} E\widehat{\pi}_t^2.
 \end{aligned} \tag{83}$$

Simplification of Q_z

Using (81) in (76) results in

$$\begin{aligned}
 \varphi Q_z = \rho \frac{\theta\phi\beta\alpha\pi^{\theta\phi}}{1 - \alpha\beta\pi^{\theta\phi}} & \left(\widehat{X}_t\widehat{\pi}_t + (\theta\phi - \theta + 1) \frac{\alpha\pi^{\theta-1}}{1 - \alpha\pi_t^{\theta-1}} \widehat{\pi}_t^2 \right. \\
 & \left. + \frac{1}{2} (\theta\phi - 1) \widehat{\pi}_t^2 \right).
 \end{aligned} \tag{84}$$

Computing the Final Quadratic Form

We use the steady-state relation among Lagrange multipliers (71) and quadratic forms (74), (75), (77), (83), and (84) to write

$$\begin{aligned}
 EU = & -E\frac{1}{2}\Lambda_{\Delta}\widehat{\Delta}_t^2 - \frac{1}{2}(1 - h\Phi)\widehat{c}_t^2 - \frac{1}{2}h\Phi \left(\widehat{c}_t + \widehat{\Delta}_t - \widehat{\mu}_t \right)^2; \\
 & - E \left(\Lambda_{x\pi}\widehat{X}_t\widehat{\pi}_t + \frac{1}{2}\Lambda_{\pi\pi}\widehat{\pi}_t^2 \right).
 \end{aligned} \tag{85}$$

In that expression, parameters are gathered in the Λ -terms. For example, $\Lambda_{\Delta} = \left(\frac{1 - \alpha\pi^{\theta-1}\alpha\pi^{\theta\phi}}{\alpha\pi^{\theta\phi} - \alpha\pi^{\theta-1}} \right) (v + \Phi h)$. Finally, we exploit the linear relation (79) to find an expression for the cross term $E\widehat{X}_t\widehat{\pi}_t$: $E\widehat{X}_t^2 = (\beta\alpha\pi^{\theta-1})^2 E \left(\widehat{X}_{t+1} + (\theta - 1) \widehat{\pi}_{t+1} \right)^2$ and $E\widehat{X}_t\widehat{\pi}_t = -\frac{1}{2}(\theta - 1) E\widehat{\pi}_{t+1}^2 + \frac{1}{2} \frac{1 - (\beta\alpha\pi^{\theta-1})^2}{(\theta - 1)(\beta\alpha\pi^{\theta-1})^2} E\widehat{X}_t^2$. Hence, one recovers

$$\begin{aligned}
 EU = & -\frac{1}{2} E \left(h\Phi \left(\widehat{c}_t + \widehat{\Delta}_t - \widehat{\mu}_t \right)^2 + (1 - h\Phi)\widehat{c}_t^2 \right. \\
 & \left. + \Lambda_x\widehat{X}_t^2 + \Lambda_{\pi}\widehat{\pi}_t^2 + \Lambda_{\Delta}\widehat{\Delta}_t^2 \right).
 \end{aligned} \tag{86}$$

Alternative Presentation

There are alternative ways to compose the quadratic criterion. Recall that c_t is marginal cost defined in (46):

$$\widehat{c}_t = \widehat{\mu}_t + v\widehat{\Delta}_t + (v + 1)\phi\left(\widehat{Y}_t - \widehat{A}_t\right). \tag{87}$$

The first two terms in the quadratic loss function (86) can be simplified, defining Y_t^* , \widehat{g}_t , and G as

$$\begin{aligned} Y_t^* &:= \widehat{A}_t - v_\mu\widehat{\mu}_t - v_\Delta\widehat{\Delta}_t; \\ G &:= h\Phi(1 - h\Phi); \quad v_\mu := \frac{1 - h\Phi}{(v + 1)\phi}; \quad v_\Delta := \frac{v + h\Phi}{(v + 1)\phi}; \\ \widehat{g}_t &:= \widehat{\Delta}_t - \widehat{\mu}_t. \end{aligned}$$

Thus (86) becomes

$$EU = -\frac{1}{2}E\left[\phi(1 + v)\left(\widehat{Y}_t - \widehat{Y}_t^*\right)^2 + G\widehat{g}_t^2 + \Lambda_x\widehat{X}_t^2 + \Lambda_\Delta\widehat{\Delta}_t^2 + \Lambda_\pi\widehat{\pi}_t^2\right]. \tag{88}$$

Linear Approximation to UO Policy

This is constructed by approximating the relevant first-order conditions around the steady state defined in section 6.1. Details are available upon request.

Linear Approximation to TP-Optimal Policy

For the most part, these derivations, although a little involved in places, are straightforward. Details are available upon request.

References

Anand, S., and A. K. Sen. 2000. "Human Development and Economic Sustainability." *World Development* 28 (12): 2029–49.
 Anderson, E. W., E. R. McGrattan, L. P. Hansen, and T. J. Sargent. 1996. "Mechanics of Forming and Estimating Dynamic Linear Economies." In *Handbook of Computational Economics*,

- Vol. 1, ed. H. M. Amman, D. A. Kendrick, and J. Rust, 171–252 (chapter 4). Elsevier.
- Ascari, G., and T. Ropele. 2007. “Optimal Monetary Policy under Low Trend Inflation.” *Journal of Monetary Economics* 54 (8): 2568–83.
- Benigno, P., and D. J. Lopez-Salido. 2006. “Inflation Persistence and Optimal Monetary Policy in the Euro Area.” *Journal of Money, Credit and Banking* 38 (3): 587–614.
- Benigno, P., and M. Woodford. 2004. “Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach.” In *NBER Macroeconomics Annual 2003*, ed. M. Gertler and K. Rogoff, 271–333. Cambridge, MA: MIT Press.
- . 2005. “Inflation Stabilization and Welfare: The Case of a Distorted Steady State.” *Journal of the European Economic Association* 3 (6): 1185–1236.
- . 2006a. “Optimal Inflation Targeting under Alternative Fiscal Regimes.” NBER Working Paper No. 12158.
- . 2006b. “Optimal Taxation in an RBC Model: A Linear-Quadratic Approach.” *Journal of Economic Dynamics and Control* 30 (9–10): 1445–89.
- . 2012. “Linear-Quadratic Approximation of Optimal Policy Problems.” *Journal of Economic Theory* 147 (1): 1–42.
- Blake, A. P. 2001. “A Timeless Perspective on Optimality in Forward-Looking Rational Expectations Models.” Discussion Paper No. 154, National Institute of Economic and Social Research.
- Calvo, G. 1983. “Staggered Prices in a Utility-Maximizing Framework.” *Journal of Monetary Economics* 12 (3): 383–98.
- Caplin, A., and J. Leahy. 2004. “The Social Discount Rate.” *Journal of Political Economy* 112 (6): 1257–68.
- Clarida, R., J. Gali, and M. Gertler. 1999. “The Science of Monetary Policy: A New Keynesian Perspective.” *Journal of Economic Literature* 37 (4): 1661–1707.
- Damjanovic, T., V. Damjanovic, and C. Nolan. 2008. “Unconditionally Optimal Monetary Policy.” *Journal of Monetary Economics* 55 (3): 491–500.
- Damjanovic, T., and C. Nolan. 2010. “Relative Price Distortions and Inflation Persistence.” *Economic Journal* 120 (547): 1080–99.

- Debortoli, D., and R. Nunes. 2006. "On Linear Quadratic Approximations." MPRA Paper No. 544, University Library of Munich (revised July 2006).
- Erceg, C. J., D. W. Henderson, and A. T. Levin. 2000. "Optimal Monetary Policy with Staggered Wage and Price Contracts." *Journal of Monetary Economics* 46 (2): 281–313.
- Harrod, R. 1948. *Towards a Dynamic Economics*. London: Macmillan and Company.
- Holzmann, R., and R. Hinz. 2005. *Old-Age Income Support in the 21st Century: An International Perspective on Pension Systems and Reform*. Washington, D.C.: The World Bank.
- Horvath, M. 2011. "Alternative Perspectives on Optimal Public Debt Adjustment." *B.E. Journal of Macroeconomics* 11 (1): 34.
- Jensen, C., and B. T. McCallum. 2002. "The Non-Optimality of Proposed Monetary Policy Rules under Timeless Perspective Commitment." *Economics Letters* 77 (2): 163–8.
- . 2010. "Optimal Continuation versus the Timeless Perspective in Monetary Policy." *Journal of Money, Credit and Banking* 42 (6): 1093–1107.
- Judd, K. L. 1999. *Numerical Methods in Economics*. MIT Press.
- Khan, A., R. G. King, and A. L. Wolman. 2003. "Optimal Monetary Policy." *Review of Economic Studies* 70 (4): 825–60.
- Kim, J., and D. Henderson. 2005. "Inflation Targeting and Nominal-Income-Growth Targeting: When and Why Are They Suboptimal?" *Journal of Monetary Economics* 52 (8): 1463–95.
- Kim, J., and S. H. Kim. 2003. "Spurious Welfare Reversals in International Business Cycle Models." *Journal of International Economics* 60 (2): 471–500.
- . 2007. "Two Pitfalls of Linearization Methods." *Journal of Money, Credit and Banking* 39 (4): 995–1001.
- King, R., and A. L. Wolman. 1999. "What Should the Monetary Authority Do When Prices Are Sticky?" In *Monetary Policy Rules*, 349–404. National Bureau of Economic Research, Inc.
- Kollmann, R. 2002. "Monetary Policy Rules in the Open Economy: Effect on Welfare and Business Cycles." *Journal of Monetary Economics* 49 (5): 989–1015.
- Kydland, F., and E. C. Prescott. 1977. "Rules rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85 (3): 473–90.

- Levine, P., J. Pearlman, and R. Piersse. 2008. "Linear-Quadratic Approximation, External Habit and Targeting Rules." *Journal of Economic Dynamics and Control* 32 (10): 3315–49.
- McCallum, B. T. 2005. "What Is the Proper Perspective for Monetary Policy Optimality?" *Monetary and Economic Studies* 23 (S1): 13–24.
- Pigou, A. C. 1932. *The Economics of Welfare*. 4th ed. London: Macmillan.
- Ramsey, F. P. 1928. "A Mathematical Theory of Saving." *Economic Journal* 38 (152): 543–59.
- Rotemberg, J. J., and M. Woodford. 1998. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy: Expanded Version." NBER Technical Working Paper No. 233.
- Schmitt-Grohe, S., and M. Uribe. 2004. "Optimal Fiscal and Monetary Policy under Sticky Prices." *Journal of Economic Theory* 114 (2): 198–230.
- . 2007. "Optimal Simple and Implementable Monetary and Fiscal Rules." *Journal of Monetary Economics* 54 (6): 1702–25.
- Soderlind, P. 1999. "Solution and Estimation of RE Macromodels with Optimal Policy." *European Economic Review* 43 (4–6): 813–23.
- Solow, R. M. 1974. "The Economics of Resources or the Resources of Economics." *American Economic Review* 64 (2): 1–14.
- Stern, N. H. 2007. *The Economics of Climate Change: The Stern Review*. Cambridge University Press.
- Strotz, R. H. 1956. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies* 23 (3): 165–80.
- Sutherland, A. 2002. "A Simple Second-Order Solution Method for Dynamic General Equilibrium Models." CEPR Discussion Paper No. 3554.
- Taylor, J. B. 1979. "Estimation and Control of a Macroeconomic Model with Rational Expectations." *Econometrica* 47 (5): 1267–86.
- Whiteman, C. H. 1986. "An Analytical Policy Design under Rational Expectations." *Econometrica* 54 (6): 1387–1405.
- Woodford, M. 1999a. "Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?" In *New Challenges*

for Monetary Policy, 277–316. Proceedings of the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming, August 26–28.

———. 1999b. “Optimal Monetary Policy Inertia.” *The Manchester School* 67 (s1): 1–35.

———. 2002. “Inflation Stabilization and Welfare.” *Contributions to Macroeconomics* 2 (1): Article 1.