



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Journal of Monetary Economics 50 (2003) 1425–1456

Journal of  
MONETARY  
ECONOMICS

[www.elsevier.com/locate/econbase](http://www.elsevier.com/locate/econbase)

# Optimal monetary policy in an economy with inflation persistence<sup>☆</sup>

Jón Steinsson

*Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138, USA*

Received 1 April 2001; received in revised form 6 May 2003

---

## Abstract

This paper studies optimal monetary policy in a model where inflation is persistent. Two types of price setters are assumed to exist. One acts rationally given Calvo-type constraints on price setting. The other type sets prices according to a rule-of-thumb. This results in a Phillips curve with both a forward-looking term and a backward-looking term. The Phillips curve nests a standard purely forward-looking Phillips curve as well as a standard purely backward-looking Phillips curve as special cases. A cost push supply shock is derived from microfoundations by adding a time varying income tax and by making the elasticity of substitution between goods stochastic. A central bank loss function for this model is derived from a second-order Taylor approximation of the household's welfare function. Optimal monetary policy for different relative values of the forward- and backward-looking terms is then analyzed for both the commitment case and the case of discretion.

© 2003 Elsevier B.V. All rights reserved.

*JEL classification:* E31; E32; E52; E58

*Keywords:* Optimal monetary policy; Phillips curve; Inflation persistence; Cost push shock

---

---

<sup>☆</sup>I thank Michael Woodford for priceless advice and encouragement. I thank Gauti B. Eggertsson, Marc Giannoni, Emi Nakamura, Thórarinn G. Pétursson and anonymous referees for valuable discussions and comments. Special thanks to Lonn Waters for moral support. Finally, I thank the Central Bank of Iceland for financial support.

*E-mail address:* [steins@fas.harvard.edu](mailto:steins@fas.harvard.edu) (J. Steinsson).

## 1. Introduction

Ever since the publication of Phillips' (1958) famous paper documenting the apparent tradeoff between inflation and unemployment, the Phillips curve has been a central piece of macroeconomics. Few ideas in economics have been as controversial, as influential, and undergone as many fundamental revisions. Since Friedman (1968) and Phelps (1967) it has been widely appreciated that inflationary expectations are an important element of the Phillips curve. Two very different approaches to modeling how inflationary expectations enter the Phillips curve have been most popular in the literature. One approach uses lagged values of inflation as a proxy for current inflationary expectations. According to this approach, the Phillips curve takes the following form:

$$\pi_t = A(L)\pi_{t-1} + B(L)x_t,$$

where  $\pi_t$  is inflation in period  $t$ ,  $x_t$  is the output gap in period  $t$ , while  $A(L)$  and  $B(L)$  are polynomials in the lag operator. We will refer to this as the “acceleration” Phillips curve.<sup>1</sup> Alternatively, it is often assumed that inflationary expectations are formed rationally in an environment of staggered price and wage adjustments. These assumptions result in a Phillips curve of the following form:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t,$$

where  $E_t \pi_{t+1}$  is the conditional expectation of  $\pi_{t+1}$  at date  $t$ . We will refer to this as the “new Keynesian” Phillips curve.<sup>2</sup>

Neither of these two specifications, however, seems adequate to capture the behavior of inflation in actual economies. The acceleration Phillips curve fails to capture the fact that individuals and firms do not form their expectations about inflation in a rigid and mechanical manner. For instance, it is well documented that inflationary expectations can be drastically altered by a sharp change in macroeconomic policy.<sup>3</sup> On the other hand, the new Keynesian Phillips curve fails to capture the fact that inflation is highly persistent. According to it firms completely front load changes in prices in response to “news” about future profits. Empirical studies do not validate this prediction. Several recent studies which seek to estimate Phillips curves of this type find that they fit the data poorly (see e.g., Fuhrer and Moore, 1995; Fuhrer, 1997; Gali and Gertler, 1999; Roberts, 2000). Evidence from VAR studies also show that the response of inflation to shocks is “hump-shaped” rather than front loaded.

<sup>1</sup>For recent examples of papers which specify the Phillips curve in this manner, see Ball (1997) and Svensson (1997a).

<sup>2</sup>For examples of papers which use this specification of the Phillips curve, see Roberts (1995) and Woodford (2003, Chapter 3).

<sup>3</sup>For a particularly dramatic account of this, see Thomas Sargent's essay, “The Ends of Four Big Inflations,” in Sargent (1993).

In recent years increasing attention has been given to the following hybrid specification of the Phillips curve:

$$\pi_t = \chi_1 E_t \pi_{t+1} + \chi_2 \pi_{t-1} + kx_t. \quad (1)$$

Fuhrer and Moore (1995) derive a Phillips curve of this type with  $\chi_1 = \chi_2 = 0.5$  from a model with two period overlapping wage contracts. They estimate this equation and conclude that it fits recent U.S. data better than either a purely forward-looking or purely backward-looking Phillips curve. Gali and Gertler (1999) derive a Phillips curve of this type from a model with staggered price setting with the additional assumption that a fraction of the producers set their prices according to a rule of thumb. They then estimate this model and report values for  $\chi_1$  and  $\chi_2$  close to 0.8 and 0.2, respectively. They are able to reject both the purely forward-looking Phillips curve and the purely backward-looking Phillips curve. Other recent papers, discussed below, come to similar conclusions, although the estimated relative values of  $\chi_1$  and  $\chi_2$  vary greatly between studies.

In light of these facts, and the importance of the Phillips curve for the conduct of monetary policy, it is surprising how little work has sought to analyze and compare optimal monetary policy for different relative weights of  $\chi_1$  and  $\chi_2$ . This is especially surprising given the current emphasis on the analysis of robustness of different types of monetary policy. Surely, variation in the relative weights on the forward- and backward-looking terms in the Phillips curve is an important dimension of such robustness analysis. The principal goal of this paper is to partially fill this hole in the literature.

Another goal of the paper is the derivation of microfoundations for several popular deviations from the benchmark new Keynesian model. We derive a hybrid Phillips curve by assuming that a fraction of the producers set their prices according to a rule of thumb. This approach to deriving a hybrid Phillips curve has recently been used by Gali and Gertler (1999). The rule of thumb we choose is however a generalization of the rule of thumb chosen by Gali and Gertler. It has the theoretically appealing property that it nests the standard new Keynesian Phillips curve and the standard acceleration Phillips curve as special (limit) cases.

A second theoretical innovation of the paper is a derivation of a “cost push” shock to the Phillips curve. Actually, we model two potential sources of such shocks: time varying income taxes and time varying monopoly power of producers. It turns out that for reasonable calibrations of our model variation in taxes results in very small shocks while reasonable variation of the monopoly power of producers is capable of creating large disturbances to the Phillips curve. Our model also includes other “supply” disturbances which do not constitute cost push shocks.<sup>4</sup> The reason is that they represent movements in the efficient level of output which monetary policy does not optimally react to.

In Section 2 we present the derivation of our model. In Section 3 we analyze optimal responses of the economy to cost push supply shocks. In Section 4 we

<sup>4</sup>These other supply shocks are shocks to agents disutility of working and shocks to the production function.

conclude with a brief discuss of the main insights that can be drawn from the analysis presented in this paper as well as a discussion of some future extensions of this research.

## 2. The model's general equilibrium foundations

In response to the Lucas (1976) critique of econometric policy evaluation, it is rapidly becoming customary within macroeconomics to conduct policy analysis with stochastic general equilibrium models in which the effects of changes in policy on the decision rules of private agents are carefully accounted for. The model studied in this paper is a stochastic general equilibrium representative household model with monopolistically competitive market structure and sluggish price adjustments. This type of model is by now quite standard. Recent papers in this genre include Yun (1996), Woodford (1996), Obstfeld and Rogoff (1996, Chapter 10) and Rotemberg and Woodford (1997, 1999).

### 2.1. Household preferences and market structure

The economy consists of a continuum of infinitely lived households/producers<sup>5</sup> of measure 1. The households all have identical preferences, represented by

$$E_t \sum_{s=t}^{\infty} \beta^s [u(C_s^i, \xi_s) - v(y_s(z); \xi_s)], \quad (2)$$

where  $\beta$  is a discount factor,  $\xi_s$  is a vector of shocks to the household's preferences and production capabilities. We assume that each household specializes in the production of one differentiated good, denoted by  $y_t(z)$ . Here,  $C_s^i$  denotes household  $i$ 's consumption of a composite consumption good. This composite consumption good takes the familiar Dixit–Stiglitz form

$$C_t^i = \left[ \int_0^1 c_t^i(z)^{(\theta_t-1)/\theta_t} dz \right]^{\theta_t/(\theta_t-1)}, \quad (3)$$

where  $c_t^i(z)$  is household  $i$ 's consumption of good  $z$  in period  $t$ . All goods enter the utility function symmetrically. The specific functional form of Eq. (3) implies a constant elasticity of substitution between goods, equal to  $\theta_t > 1$ . As a result of this, each household possesses a certain degree of monopoly power in the good it produces.

It is standard practice to assume that  $\theta_t$  is constant. We, however, assume that  $\theta_t$  is stochastic. The economic interpretation of this assumption is that the variety of goods produced in the economy and their substitutability is constantly changing. As a result of this the monopoly power of each household and therefore its desired markup over marginal costs is also constantly changing. Below we

<sup>5</sup>These two terms are used interchangeably in this paper.

will see that this turns out to be a convenient way to introduce cost push shocks into our model. Notice that we assume that the  $\theta_t$ 's of different producers in the economy bounce around in a perfectly correlated manner. This is of course not particularly realistic. A more realistic model of supply would study competitive conditions at the industry level. Any such model in which industry level shocks to competitive conditions do not cancel out at the aggregate level, either because they are correlated between industries or because most of the economy is made up of relatively few large industries, would produce aggregate fluctuations in some aggregate index of desired markups. For simplicity, we sidestep these issues by assuming that shocks to the economy's market structure are aggregate shocks.

We furthermore assume that  $u(C_t^i; \xi_t)$  is increasing and quasi-concave, while  $v(y_t(z); \xi_t)$  is increasing and convex. It is natural to interpret  $v(y_t(z); \xi_t)$  as a reduced-form representation of production costs as they would be in a model with firms and a labor market.<sup>6</sup> Under this interpretation,  $v(y_t(z); \xi_t)$  is convex because of diminishing marginal returns to labor in production, and because of increasing marginal disutility of labor supplied.

All goods produced in the economy are non-durable consumption goods, purchased and consumed immediately by households (we abstract from government purchases). Investment and capital accumulation play no role in this model. To the extent that capital is used in the production of goods, the economy is endowed with a fixed amount of non-depreciating capital, which does not change over time. The economy is closed. International trade and the price of domestic goods in terms of foreign goods therefore plays no role in the model.

We abstract from the liquidity services of real money balances. This may seem odd in a paper primarily concerned with monetary policy, but it is merely done to simplify the exposition. The model should be viewed either as a money in the utility-function model in which the household's utility function includes a third term,  $w(M_t/P_t; \xi_t)$ , representing the utility of real money balances, or as a "cash-less" limit of a monetary economy (Woodford, 1998). Since we will be interested in formulating monetary policy by interest rate rules, explicit reference to the existence of money is superfluous.<sup>7</sup>

Since we assume an economy with differentiated goods, households face a decision in each period about how much to consume of each individual good. We assume that households seek to maximize the value of the composite consumption good,  $C_t^i$ , which they can purchase given their income. This leads to familiar expressions for the demand for each individual good

$$c_t^i(z) = C_t^i \left( \frac{p_t(z)}{P_t} \right)^{-\theta_t}, \quad (4)$$

<sup>6</sup>See Woodford (2003, Chapter 3) for a discussion of a model of this type with firms and a labor market.

<sup>7</sup>See Woodford (2003, Chapter 2) for a discussion of a model of this type which is explicit about the role of money.

where  $p_t(z)$  is the price of good  $z$  in period  $t$  and,  $P_t$  is the price level in period  $t$  given by

$$P_t = \left[ \int_0^1 p_t(z)^{1-\theta_t} dz \right]^{1/(1-\theta_t)}. \quad (5)$$

This specification of the price level has the property that  $P_t C_t^i$  gives the minimum price for which an amount  $C_t^i$  of the composite consumption good can be purchased.

We assume that there exist complete financial markets in the economy. That is, a wide enough range of financial assets exist so that households can create financial portfolios with any type of return structure with regard to possible future states of the world. Households are, therefore, able to insure themselves against all types of uncertainty in the model. In particular, they can pool the risk that is associated with the constraints we will introduce on the evolution of the prices of the goods they produce. Furthermore, we assume that all households are equally well off initially in terms of their combination of financial wealth and the price of the good they produce. It follows from these assumptions that all households will consume equal amounts of the composite consumption good and equal amounts of each individual good. Thus, we can drop the superscript  $i$  on consumption variables in what follows. Each household will then face the same flow budget constraint given by

$$P_t C_t + E_t[R_{t,t+1} B_{t+1}] \leq B_t + (1 - \tau_t) p_t(z) y_t(z) + T_t, \quad (6)$$

where  $B_t$  is the nominal value of the household's portfolio of financial assets brought into period  $t$ ,  $\tau_t$  is a time-varying income tax rate levied by the government,  $T_t$  is a lump sum transfer received from the government, and  $R_{t,t+1}$  is the stochastic discount factor which determines the price to the household in period  $t$  of being able to carry a state-contingent amount  $B_{t+1}$  of wealth into period  $t+1$ . It follows from the absence of arbitrage opportunities that all assets can be priced by such a stochastic discount factor. The riskless short-term nominal interest rate,  $i_t$ , has a particularly simple representation in terms of the stochastic discount factor, namely

$$\frac{1}{1+i_t} = E_t[R_{t,t+1}]. \quad (7)$$

In order to rule out Ponzi schemes, we assume that financial wealth carried into the next period,  $B_{t+1}$ , satisfies the bound

$$B_{t+1} \geq - \sum_{T=t+1}^{\infty} E_{t+1}[R_{t+1,T} (1 - \tau_T) p_T(z) y_T(z)] \quad (8)$$

with certainty, that is, in each state of the world which may be reached in period  $t+1$ . Here  $R_{t,T}$  denotes the stochastic discount factor for discounting nominal income received in period  $T$  back to period  $t$ ,

$$R_{t,T} = \prod_{s=t+1}^T R_{s-1,s}.$$

In order for the intertemporal budget constraint to be a constraint at all, the present value of the household’s future income must be bounded, i.e.,

$$\sum_{T=t+1}^{\infty} E_{t+1}[R_{t+1,T}(1 - \tau_t)p_T(z)y_T(z)] < \infty \tag{9}$$

at all times, and in all states of the world. If this were not the case (even in some states of the world, since markets are complete) households could afford infinite consumption. This is obviously not a very interesting case. We restrict attention to the case where Eq. (9) is satisfied.

We also assume that the nominal interest rate satisfy the lower bound

$$i_t \geq 0 \tag{10}$$

at all times. If this were not the case, money would come to dominate bonds as an asset. It would then be possible to finance unbounded consumption by selling enough bonds. We restrict attention to the case where Eq. (10) is satisfied and does not bind at all times. Here, as well as elsewhere, we treat the model as a money in the utility-function model even though we are not explicit about the existence of money in our notation.

Given these three assumptions the infinite sequence of flow budget constraints of the household can be replaced by a single intertemporal constraint,

$$\sum_{s=t}^{\infty} E_t[R_{t,s}P_sC_s] \leq B_t + \sum_{s=t}^{\infty} E_t[R_{t,s}(1 - \tau_t)p_s(z)y_s(z)]. \tag{11}$$

We assume that  $\{\tau_s\}_{s=t}^{\infty}$  and  $\{T_s\}_{s=t}^{\infty}$  are exogenous processes but that the government balances its budget in each period.

### 2.2. Household optimization and market clearing

We now turn to household optimization. Let us begin by considering the household’s decisions regarding optimal consumption and asset holdings. This problem is a standard constrained optimization problem where Eq. (2) is maximized subject to Eq. (6) for each  $t \geq 0$  taking  $B_0$  as given. This type of problem may be solved using stochastic Lagrange multipliers. The resulting first-order conditions are

$$u_C(C_t; \xi_t) = P_t A_t, \tag{12}$$

$$R_{t,T} A_t = \beta^{T-t} A_T, \tag{13}$$

where  $A_t$  is the marginal utility of nominal income at time  $t$ , i.e., the Lagrange multiplier of the constrained optimization, and  $u_C$  denotes the partial derivative of  $u$  with respect to  $C$ . These two equations should hold for all periods  $t$  and all subsequent periods  $T$ .<sup>8</sup>

---

<sup>8</sup>The third first-order condition which results from the differentiation of our Lagrangian with respect to  $y_t(z)$  is not reported here since it only holds in the case of flexible prices. We will primarily be concerned with a sticky price version of this model.

In addition to these first-order conditions the household's choices must satisfy a transversality condition. For each feasible sequence  $\{\hat{B}_s\}_{s=t}^{\infty}$  for which the objective function is larger than that obtained with the optimal sequence,  $\{\tilde{B}_s\}_{s=t}^{\infty}$ , it must be true that

$$\limsup_{s \rightarrow \infty} \beta^s E[A_s(\tilde{B}_s - \hat{B}_s)] \leq 0. \quad (14)$$

It may be shown that a sufficient condition for Eq. (14) to hold is that  $u(C_s; \xi_s)C_s$  be bounded along  $\{\tilde{C}_s\}_{s=t}^{\infty}$ . It is rather standard to assume that  $u(C_s; \xi_s)C_s$  is bounded for  $C_s$  in any bounded subset of  $C$ 's domain (see e.g., Farmer, 1999). We shall assume that this condition holds and thereby assume that Eq. (14) holds in our model.

In equilibrium markets must clear. The conditions for market clearing are  $c_t(z) = y_t(z)$ ,  $C_t = Y_t$ ,  $B_t = 0$ , for all  $t$  and all  $z$ , where  $c_t(z)$  denotes total consumption of good  $z$ , and  $Y_t$  denotes total output. Combining these markets clearing conditions with Eqs. (12), (13), and (7) we get a more familiar Euler equation for household consumption

$$\beta E_t \left\{ \frac{u_C(Y_{t+1}; \xi_{t+1})}{u_C(Y_t; \xi_t)} \frac{P_t}{P_{t+1}} \right\} = \frac{1}{1 + i_t}. \quad (15)$$

As we can see from this equation, current consumption is determined by the current level of nominal interest rates as well as household expectations of future consumption and future inflation. Household consumption behavior is therefore forward-looking in important ways in our model.

We now turn to the pricing decisions of the households. Following Calvo (1983) we assume that a fraction  $1 - \alpha$  of the households are able to set a new price in each period. More precisely, in each period each household can set a new price with probability  $1 - \alpha$ . With probability  $\alpha$  it must let its price rise at the rate of steady-state inflation,  $\bar{\pi}$ . For each household this probability is independent of the time that has elapsed since it last changed its price, and the degree to which its price is different from the optimal price in the current period. This type of assumption turns out to be very convenient for the purpose of aggregation, since pricing decisions in period  $t$  are independent of past pricing decisions.

Until now we have assumed full rationality on behalf of all households. At this point we will deviate from this assumption and follow Gali and Gertler (1999) in assuming that there exist two types of households in the economy when it comes to pricing decisions. A fraction  $1 - \omega$  of the households behave optimally when making their pricing decisions. We refer to these households as the forward-looking households. The remaining households, of measure  $\omega$ , instead use a simple backward-looking rule-of-thumb when setting their prices.<sup>9</sup> We refer to these households as the backward-looking households.

<sup>9</sup>An earlier example of the utilization of this type of assumption in order to better explain the deviations of actual behavior from the predictions of models which assume fully rational agents is Campbell and Mankiw (1989). They use this type of assumption to explain the relation between consumption and income.



It follows from our assumptions that all forward-looking households which are able to adjust their price at date  $t$  will choose the same price. Let  $p_t^f$  denote this price. We assume that all backward-looking households who change their price at date  $t$  also set the same price. Let  $p_t^b$  denote this price. The aggregate price level will then evolve according to

$$P_t = [\alpha(\bar{\pi}P_{t-1})^{1-\theta_t} + (1-\alpha)(1-\omega)(p_t^f)^{1-\theta_t} + (1-\alpha)\omega(p_t^b)^{1-\theta_t}]^{1/(1-\theta_t)}. \tag{16}$$

Let us first consider what choice of price is optimal for a forward-looking household which is able to change its price in period  $t$ . The new price will apply with certainty in period  $t$ ; it will apply in period  $t + 1$  with probability  $\alpha$ ; in period  $t + 2$  with probability  $\alpha^2$ ; and so on. It is therefore chosen to solve

$$\max_p E_t \sum_{T=t}^{\infty} \alpha^{T-t} \{ \Lambda_t R_{t,T} (1 - \tau_t) \bar{\pi}^{T-t} p_{yT}(p) - \beta^{T-t} v_y(y_T(p)) \},$$

where  $y_T(z)$  denotes the demand for the good at date  $T$  as a function of its price. The marginal utility of income,  $\Lambda_t$ , can be treated as a constant in this calculation since risk sharing through financial markets implies that it is unaffected by the pricing decision of the household. Solving this optimization problem we get that the optimal price  $p_t^f$  chosen by the forward-looking households satisfies the first-order condition:

$$E_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left\{ u_C(Y_T; \xi_T) (1 - \tau_T) (\bar{\pi}^{T-t} p_t^f / P_T) - \frac{\theta_t}{\theta_t - 1} v_y(Y_T (p_t^f / P_T)^{-\theta_t}; \xi_T) \right\} = 0, \tag{17}$$

As in [Gali and Gertler \(1999\)](#) we assume that the backward-looking firms set their prices according to a rule of thumb. The rule of thumb we choose is, however, a slight generalization of the rule used in [Gali and Gertler \(1999\)](#). We assume that the backward-looking households set their prices according to the following rule:

$$p_t^b = p_{t-1}^* \Pi_{t-1} \left( \frac{Y_{t-1}}{Y_{t-1}^n} \right)^\delta, \tag{18}$$

where  $\Pi_{t-1} = P_{t-1}/P_{t-2}$ ,  $Y_t^n$  denotes the efficient level of output,<sup>10</sup> and  $p_{t-1}^*$  denotes an index of the prices set at date  $t - 1$ , given by

$$\log p_{t-1}^* = (1 - \omega) \log p_{t-1}^f + \omega \log p_{t-1}^b. \tag{19}$$

According to Eq. (18) the backward looking households adjust their prices to equal the geometric mean of the prices which they saw chosen in the period before,  $p_{t-1}^*$ , adjusted for the inflation rate they last observed,  $\Pi_{t-1}$ , and adjusted for a measure of the output gap they last observed,  $Y_{t-1}/Y_{t-1}^n$ . The difference between Eq. (18) and the rule of thumb [Gali and Gertler \(1999\)](#) use is that in their paper the backward-looking price setters do not take account of demand conditions when setting their prices. As we will see in the next few sections this has unappealing consequences in the limit when  $\omega \rightarrow 1$ .

<sup>10</sup>The efficient level of output is defined precisely in Section 2.3.

The main justifications for introducing irrational price setters which set prices according to a rule of thumb of this type is that it represents a simple way of introducing inflation persistence into our model. As we mentioned in the introduction, empirical studies have frequently found that inflation in the U.S. and other countries exhibits a large degree of persistence (see e.g., Gali and Gertler, 1999; Fuhrer and Moore, 1995; Fuhrer, 1997; Roberts, 2000). These studies frequently reject the purely forward-looking Phillips curve in favor of a hybrid Phillips curve and criticize proponents of optimizing models for neglecting this empirical fact. This paper analyzes optimal monetary policy in a model which is consistent with this empirical finding.

Identical Phillips curves to the one derived in this section may be derived in a variety of different ways based on the microstructure of the economy and without the introduction of irrational price setters. However, Roberts (1997) has argued that surveys of expectations of inflation can be used to distinguish between different models of inflation persistence. He finds that if survey-data on expectations of inflation are used to estimate the Phillips curve no additional lags of inflation are needed. Hence, if surveys accurately reflect expectations of inflation, this implies that imperfectly rational expectations and not the underlying structure of the economy seems to be the source of the observed persistence in inflation. Furthermore Roberts (1998) finds that survey data of expectations of inflation are well represented by a weighted average of forward-looking and backward-looking expectations.

### 2.3. Log-linearization of the model

The equations of our model are a quite complicated system of stochastic non-linear difference equations. A general solution of this type of system is beyond the scope of this paper. Instead, we will log-linearize the model around its steady state with zero inflation and study the dynamics of this approximation. We limit our attention to bounded solutions for which all the endogenous variables fluctuate in a small enough interval to make the log-linear approximation valid.

It will prove convenient to write the model in terms of the percentage difference between actual output,  $Y_t$ , and the level of output which would prevail if prices were fully flexible, markets were perfectly competitive and no distortionary taxes were levied,  $Y_t^n$ . We refer to  $Y_t^n$  as the efficient level of output and the logarithm of the ratio between actual output and the efficient level of output as the output gap,  $x_t$ .

Notice that if the steady-state tax rate is  $\bar{\tau} = -(\bar{\theta} - 1)^{-1}$  then Eq. (17) becomes  $u_C(\bar{Y}; 0) = v_y(\bar{Y}; 0)$  in the steady state. So, in this case the steady-state value of output,  $\bar{Y}$ , also defines the efficient steady-state level of output. We linearize the model around a steady-state with zero inflation and the efficient steady-state level of output. In other words, we assume that  $\bar{\tau} = -(\bar{\theta} - 1)^{-1}$  and that the monetary authorities conduct monetary policy so that prices are stable in the steady state.

Given the notation and assumptions discussed above the log-linear approximation of Eq. (15) is

$$x_t = E_t x_{t+1} - \sigma[(\hat{i}_t - E_t \pi_{t+1}) - r_t^n], \quad (20)$$

where  $\hat{i}_t = \log[1 + i_t/(1 + \bar{i})]$ ,  $\pi_t = \log(\Pi_t)$ ,  $\sigma$  and  $r_t^n$  are given by

$$\sigma = -\frac{u_C}{u_{CC}\bar{Y}} > 0; \quad r_t^n = E_t[\sigma^{-1}(\log(Y_{t+1}^n/Y_t^n) - \frac{u_{C\xi}}{u_C}(\xi_{t+1} - \xi_t))]$$

and all partial derivatives are evaluated at the steady state.

Log-linearization of the supply block results in a Phillips curve of the following form:

$$\pi_t = \chi_f \beta E_t \pi_{t+1} + \chi_b \pi_{t-1} + \kappa_1 x_t + \kappa_2 x_{t-1} + \eta_t, \tag{21}$$

where  $\chi_f$ ,  $\chi_b$ ,  $\kappa_1$  and  $\kappa_2$  are parameters and  $\eta_t$  is an exogenous shock caused by changes in the tax rate  $\tau_t$  and changes in the elasticity of substitution between goods,  $\theta_t$ . A derivation of this equation is presented in Appendix A. For reasonable values of the parameters of our model it turns out that changes in the tax rate result in very small shocks to the supply curve. For the parameters chosen in Section 2.5 and  $\omega = 0$  a 1% change in the tax rate gives  $\eta_t = 0.007$ . For these same parameter values a 1% shock to  $\theta_t$  gives  $\eta_t = 6.66$ . In other words, reasonable shocks to the markets structure are able to create large supply shocks, shocks which are roughly 1000 times larger than those created by changes to the tax rate.<sup>11</sup>

Eq. (21) is valid for  $0 \leq \omega < 1$ . For  $\omega = 1$  the derivation is incorrect, since it would involve dividing by zero. It is evident from the parameters of Eq. (21) that when  $\omega \rightarrow 0$  Eq. (21) takes on the purely forward-looking new Keynesian form, presented e.g. in Woodford (2003, Chapter 3).<sup>12</sup> However, taking the limit as  $\omega \rightarrow 1$ , Eq. (21) becomes

$$\pi_t = \frac{\alpha\beta}{1 + \alpha\beta} E_t \pi_{t+1} + \frac{1}{1 + \alpha\beta} \pi_{t-1} - \frac{\alpha\beta\delta(1 - \alpha)}{1 + \alpha\beta} x_t + \frac{\delta(1 - \alpha)}{1 + \alpha\beta} x_{t-1}. \tag{22}$$

Surprisingly, Eq. (21) is not reduced to the form of the acceleration Phillips curve in this limit. Instead, the weight on the forward-looking term goes to  $\alpha\beta/(1 + \alpha\beta)$  as  $\omega \rightarrow 1$ . This may seem to imply that our Phillips curve has a non-trivial forward-looking component in this limit. However, this is an illusion. The unique bounded solution of Eq. (22) is

$$\pi_t = \pi_{t-1} + (1 - \alpha)\delta x_{t-1}. \tag{23}$$

Evidently this solution has no forward-looking component and is of the form of the acceleration Phillips curve. Furthermore, this is exactly the specification of the Phillips curve one gets for  $\omega = 1$ . If one solves Eq. (21) in a similar manner for  $\pi_t$  it is easy to show that the forward-looking component of the solution falls to zero continuously as  $\omega \rightarrow 1$ .

It is instructive to note that, if we had chosen the same specification for the rule of thumb used by the backward-looking price setters as Gali and Gertler (1999) chose, Eq. (21) would not have included an  $x_{t-1}$  term, Eq. (22) would have included neither an  $x_{t-1}$  term nor an  $x_t$  term, and the unique bounded solution of that equation

<sup>11</sup> Actually, one must be somewhat cautious when talking about reasonably sized shocks to the market structure. The size of these shocks is in the end an empirical question which has, to my knowledge, not been fully explored.

<sup>12</sup> Expressions for these parameters are presented in Appendix A.

would then have been  $\pi_t = \pi_{t-1}$ . In other words the model’s Phillips curve would not have reduced to the acceleration Phillips curve in the limit when  $\omega \rightarrow 1$  but to a much less appealing equation according to which inflation is a constant over time.

*2.4. The central bank’s loss function*

We now turn to the derivation of a loss function for the central bank. Following Rotemberg and Woodford (1997, 1999) and Woodford (2003, Chapter 6) we assume that the central bank is concerned with maximizing a quadratic Taylor series approximation of the expected utility of an equally weighted sum of the households

$$W = E \left\{ \sum_{t=0}^{\infty} \beta^t U_t \right\}, \tag{24}$$

where

$$U_t = u(Y_t; \xi_t) - \int_0^1 v(y_t(z); \xi_t) dz. \tag{25}$$

As we show in Appendix B the second-order Taylor series approximation of Eq. (25) around the steady state we log-linearize our structural equations around is

$$U_t = -\frac{\bar{Y}u_C}{2} \{(\sigma^{-1} + \psi^{-1})x_t^2 + (\bar{\theta}^{-1} + \psi^{-1})\text{var}_z \log p_t(z)\} + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3), \tag{26}$$

where the abbreviation t.i.p. stands for “terms independent of policy.” The expected utility of the households depends negatively on two terms which are influenced by policy: the square of the output gap,  $x_t^2$ , and the degree of price dispersion in the economy,  $\text{var}_z(\log p_t(z))$ .

The introduction of time varying taxes and shocks to market power turn out to have a subtle but important effect on the derivation of Eq. (26) vis a vis the corresponding derivation in, e.g., Rotemberg and Woodford (1999). In our model the monetary authorities should react to the difference between the actual level of output and the efficient level of output while in Rotemberg and Woodford (1999) the correct variable to react to was the difference between actual output and the level of output which would prevail if prices were perfectly flexible. As a result of this the monetary authorities should, in our model, react to movements in output which are caused by changes in the tax rate and shocks to the market structure while they should not react to movements in output which are caused by preference shocks or shocks to productive capabilities.

The degree of price dispersion can now be derived from our assumptions about price setting behavior. This is shown in Appendix C. The resulting welfare function is

$$\sum_{t=0}^{\infty} \beta^t U_t = -\Omega \sum_{t=0}^{\infty} \beta^t L_t + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3), \tag{27}$$

where  $\Omega$  is a constant and  $L_t$  denotes the central bank’s loss function, given by

$$L_t = \pi_t^2 + \lambda_1 x_t^2 + \lambda_2 \Delta \pi_t^2 + \lambda_3 x_{t-1}^2 + \lambda_4 \Delta \pi_t x_{t-1}. \tag{28}$$

The  $\lambda$ 's are parameters given by

$$\lambda_1 = \frac{(1 - \alpha)(1 - \alpha\beta)(\sigma - \psi)}{\alpha\sigma\bar{\theta}(\psi + \bar{\theta})}, \lambda_2 = \frac{\omega}{1 - \omega} \frac{1}{\alpha}, \lambda_3 = \frac{(1 - \alpha)^2\omega\delta^2}{\alpha(1 - \omega)}, \lambda_4 = \frac{2(1 - \alpha)\omega\delta}{\alpha(1 - \omega)}.$$

The derivation of Eq. (28) is correct for  $0 \leq \omega < 1$ , but breaks down for  $\omega = 1$  since in that case it would involve dividing by zero. Notice that when  $\omega = 0$  this loss function simplifies to

$$L_t = \pi_t^2 + \lambda_1 x_t^2,$$

which is exactly the form reported in [Woodford \(2003, Chapter 6\)](#) for the purely forward-looking new Keynesian model. As  $\omega \rightarrow 1$   $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  become unbounded. However, the relative size of these three terms remains fixed so what really happens is that the size of the first two terms in Eq. (28) shrinks relative to the last three terms as the fraction of backward-looking price setters rises.

Again, it is instructive to note the difference between the loss function we derive and the loss function we would have ended up with if we had used the same rule of thumb as [Gali and Gertler \(1999\)](#) use. In that case the loss function is identical to Eq. (28) except that  $\lambda_3 = \lambda_4 = 0$ . As a result of this the  $\Delta\pi_t$  term comes to dominate the other terms in that loss function as  $\omega \rightarrow 1$ . This means that the relative importance of output stabilization shrinks to zero in that limit. We performed a few optimal monetary policy simulations similar to those performed in Section 3 using this loss function. As one would expect, output became very explosive as  $\omega \rightarrow 1$ . In our opinion this represents a serious flaw in that model. The model used in this paper is, however, well behaved for all values of  $\omega$  as we will show in Section 3.

It should be noted that Eq. (28) only represents an accurate second-order approximation of welfare when it is used to evaluate a model around a steady state with zero inflation and the efficient level of output. As we noted in Section 2.3 our assumptions regarding the steady-state tax rate on income and the monetary policy of the monetary authority results in such an equilibrium. These two requirements on the nature of the equilibrium of course restrict the applicability of the loss function derived in this section. However, we believe that the analysis of the welfare effects of monetary policy under these rather ideal circumstances is important since in this particular case one is able to isolate the interaction of monetary policy and distortions resulting from sluggish nominal price adjustments from other distortions which result in an inefficient steady-state level of output.

### 2.5. Summary and calibration

The remainder of this paper will be concerned with policy analysis using the model just derived. The model now consists of two structural equations, Eqs. (20) and (21), and a welfare criterion for the central bank given by Eqs. (27) and (28). In this policy analysis, we will use specific values for all the parameters in the model except  $\omega$ , which we will vary between zero to one. In this way we nest the purely forward-looking new Keynesian model, the purely backward-looking acceleration model, as well as other models such as the model of [Fuhrer and Moore \(1995\)](#), as special cases

Table 1  
Values for the structural parameters of the model for four values of  $\omega$

$\omega$	0.01	0.2	0.7	0.99
$\chi_f$	1	0.7790	0.5018	0.4159
$\chi_b$	0	0.2226	0.5018	0.5882
$\kappa_1$	0.026	0.0139	-0.0016	-0.0064
$\kappa_2$	0	0.0035	0.0078	0.0092
$\lambda_1$	0.0053	0.0053	0.0053	0.0053
$\lambda_2$	0	0.3571	3.3333	141.4286
$\lambda_3$	0	$8.7 \times 10^{-5}$	0.0008	0.0344
$\lambda_4$	0	0.011	0.1040	4.4126

of our model. Since estimates do not exist for the parameters of exactly this model, we resort to choosing parameters which are close to the estimated parameters of models of similar nature.

Rotemberg and Woodford (1997) report estimates for a purely forward-looking new Keynesian model on quarterly data using a moment-matching approach. The new Keynesian model they estimate is a close relative of the benchmark new Keynesian model which we have noted is a special case of our model when  $\omega = 0$ . For this model Rotemberg and Woodford (1997) estimate that  $\sigma = 6.4$ ,  $\psi = 2.13$ , and  $\bar{\theta} = 7.88$ . They, however, do not estimate  $\alpha$ , but rather calibrate it to be  $\alpha = 0.66$  based on Blinder (1994).

Gali and Gertler (1999) estimate, for quarterly data using several variations of a non-linear instrumental variables (GMM) estimator, a Phillips curve which is a close relative of the Phillips curve derived in this paper. They report estimates of  $\alpha$  between 0.803 and 0.866,  $\beta$  between 0.885 and 0.957,  $\kappa_1$  between 0.015 and 0.037, and  $\omega$  between 0.077 and 0.522 (with 3 of their 6 estimates between 0.2 and 0.3).

We use the estimates reported in these two studies as references but choose round numbers for convenience. Our choices are:  $\beta = 0.99$ ,  $\alpha = 0.7$ ,  $\sigma = 5$ ,  $\psi = 2$  and  $\bar{\theta} = 5$ . None of the papers we have cited estimate a parameter comparable to our  $\delta$ . We set  $\delta$  so that the coefficient on  $x_{t-1}$  in Eq. (23) is equal to the coefficient on  $x_t$  in our Phillips curve when  $\omega = 0$ . Chosen in this way  $\delta = 0.052$ . Table 1 reports the values of  $\chi_f$ ,  $\chi_b$ ,  $\kappa_1$ ,  $\kappa_2$ ,  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  which result from these assumptions for several different values of  $\omega$ . We furthermore assume that  $\hat{\theta}_t$  in Eq. (A.8) is i.i.d. This assumption is not made because we think it is particularly realistic but rather because it yields simpler results which are, therefore, more useful in building intuition about the dynamics of the model.

Notice that our model corresponds closely with that of Fuhrer and Moore (1995) when  $\omega = 0.7$ . Fuhrer and Moore's original claim was that their specification of  $\chi_1 = \chi_2 = 0.5$  in Eq. (1) matched the pattern of U.S. data much better than either a purely forward-looking or purely backward-looking model. Since then, a number of studies have taken up the issue of how much relative weight to put on the forward- and backward-looking terms in the Phillips curve. Fuhrer (1997) (on U.S. data) and Blake and Westaway (1996) (on U.K. data) conclude that  $\chi_1$  close to 0.2 fit their data

best. Gali and Gertler (1999), using a measure of marginal costs in their Phillips curve instead of a measure of the output gap, however estimate  $\chi_1$  to be in the vicinity of 0.8. Roberts (2000) estimates a hybrid Phillips curve with several different measures of inflation and the output gap, several different choices of instrument sets and several different assumptions about shifts in monetary policy. His estimates of  $\chi_1$  range from 0.15 to 1.03 with most of his estimates significantly different from both 0 and 1. It is therefore evident that there is considerable uncertainty about the relative importance of the forward-looking and the backward-looking terms in the Phillips curve. In the next Section 3 we analyze optimal policy for a range of values.

### 3. Optimal responses to supply shocks

In this section, we study the optimal response of the endogenous variables,  $\pi_t$ ,  $x_t$ , to  $\eta_t$ , the cost push disturbance term in the Phillips curve. This analysis should not be confused with the design of a policy rule which the central bank should use to bring about such an equilibrium. The aim of this section is merely to characterize the path of the endogenous variables that achieves the lowest possible value of the central bank's loss function, given by Eqs. (27) and (28), in reaction to supply shocks.

In the literature on monetary policy there are two main approaches to the type of analysis we are concerned with in this section, which correspond to two different assumptions about central bank behavior. The difference of the two approaches lies in the central bank's ability to make credible commitments about its future actions. In models with forward-looking private sector behavior, current outcomes are partly determined by the private sector's expectations about the future evolution of the economy. It turns out to be the case that in such models it can be beneficial for the central bank to make commitments about its future actions which sway these expectations in desirable directions. However, as Kydland and Prescott (1977) first pointed out, these types of commitments are not generally time consistent; that is, the type of behavior which the central bank would like to commit itself to carrying out at a future date does not generally remain optimal for the bank when that future date actually arrives.

The realization of this conflict has resulted in a large literature which asks whether it makes sense to assume that central banks are able to credibly commit themselves to follow time inconsistent policies.<sup>13</sup> The ability of a central bank to make credible commitments is intimately related to the notion of central bank reputation. Issues of central bank reputation, and especially how a central bank's reputation varies over time in reaction to the outcomes of its policy, are no doubt immensely important to the optimal conduct of monetary policy. These issues will however not be taken up in this paper. We will simply analyze the two polar cases: the full commitment case, which assumes that the central bank is able to make fully credible commitments; and

---

<sup>13</sup>See Walsh (1998, Chapter 8) for a recent survey of this literature.



the case of discretionary optimization, which assumes that it is common knowledge that the central bank is unable to follow through on time inconsistent commitments.

There are at least two distinct types of time inconsistencies that central banks are faced with. [Kydland and Prescott \(1977\)](#) show that, if the natural rate of output is inefficiently low (so that the central bank targets a rate of output above the natural rate), discretionary conduct of policy will lead to a higher average level of inflation than is optimal *without* positively effecting the average level of output. This type of inflation bias is a pure cost of not being able to make credible commitments. After the early work of Kydland and Prescott, this effect was extensively studied and was widely believed to be an important partial explanation for the relatively high average levels of inflation that many OECD countries experienced in the 1970s. More recently, the importance of this effect has been questioned as most OECD countries have had considerable success in containing inflation. Alan Blinder's recent comments on this point are characteristic of the current mood:

I can assure you that it would not surprise my central banker friends to learn that economic theories that model them as seeking to drive unemployment below the natural rate imply that their policies are too inflationary. They would no doubt reply, "Of course that would be inflationary. That's why we don't do it." ([Blinder, 1998](#))

As we saw in the previous section, a central bank which seeks to maximize the welfare of households in our model will target the natural rate of output and therefore not be subject to this effect.

The recent literature on optimal monetary policy has emphasized a second and perhaps more subtle difference between the commitment case and the discretion case. [Woodford \(1999, 2003, Chapter 7\)](#) and [Clarida et al. \(1999\)](#) discuss how a central bank which is able to make credible commitments can use this ability to influence private sector expectations in a way that leads to more favorable responses to shocks.<sup>14</sup> [Woodford \(1999, 2003, Chapter 7\)](#) shows that the optimal policy under commitment entails a certain degree of history dependence on behalf of the central bank, which is absent in the discretion case. The logic behind this history dependence is quite intuitive. In order to favorably influence private sector inflationary expectations the central bank makes commitments about its future actions. However, since private sector expectations are formed rationally, commitments by the central bank only influence these expectations if the central bank in later periods carries through on its earlier commitments. The actions of the central bank in later periods must therefore take into account the state of the economy in earlier periods (which gave rise to the bank's commitments). The analysis of this section will shed light on the implication of this type of time inconsistency for the conduct of monetary policy.

---

<sup>14</sup>This type of suboptimality is sometimes called "stabilization bias". Earlier papers which discuss this effect include [Jonsson \(1997\)](#) and [Svensson \(1997b\)](#).



### 3.1. Optimal responses to supply shocks under commitment

The analysis of optimal responses to supply shocks under commitment is simply a stochastic constrained optimization problem. We form the following Lagrangian:

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \{ L_t + 2\phi_t(\chi_f \beta \pi_{t+1} - \pi_t + \chi_b \pi_{t-1} + \kappa x_t + \eta_t) \} \right\}. \tag{29}$$

Notice that we have ignored Eq. (20). In recent years it has become customary within the theoretical literature on optimal monetary policy to assume that the central bank’s control variable is the short term nominal interest rate. This is very much in keeping with the actual conduct of monetary policy by large central banks such as the Federal Reserve, the European Central Bank, and the Bank of England. These banks have in recent years all conducted monetary policy by controlling the path of a short-term interest rate. We will follow the recent literature by assuming that the central bank controls the evolution of  $\hat{i}_t$ . Since  $\hat{i}_t$  is the control variable of the central bank it can be chosen freely to satisfy Eq. (20), given optimal paths for  $\pi_t$  and  $x_t$ . Eq. (20) can therefore be ignored.<sup>15</sup>

Differentiating Eq. (29) with respect to  $\pi_t$  and  $x_t$  we get two first-order conditions which the optimal plan must satisfy,

$$(1 + \lambda_2 + \beta \lambda_2) \pi_t - \lambda_2 \pi_{t-1} - \beta \lambda_2 E_t \pi_{t+1} - \frac{\lambda_4}{2} x_t + \frac{\lambda_4}{2} x_{t-1} + \chi_f \phi_{t-1} - \phi_t + \beta \chi_b E_t \phi_{t+1} = 0, \tag{30}$$

$$(\lambda_1 + \beta \lambda_3) x_t + \frac{\beta \lambda_4}{2} E_t \pi_{t+1} - \frac{\beta \lambda_4}{2} \pi_t + \kappa \phi_t + \beta \kappa E_t \phi_{t+1} = 0. \tag{31}$$

These two conditions must hold for each date  $t \geq 1$ , and the same conditions with  $\phi_{-1} = 0$  must also hold for date  $t = 0$ . Eqs. (21), (30) and (31) may now be written as

$$\Gamma_0 Y_t = \Gamma_1 Y_{t-1} + \Psi z_t + \Pi \varepsilon_t, \tag{32}$$

where  $Y_t = [E_t \pi_{t+1}, \pi_t, x_t, E_t \phi_{t+1}, \phi_t]'$ ,  $z_t$  is a vector of exogenous shocks,  $\varepsilon_t$  is a vector of expectation errors and  $\Gamma_0$ ,  $\Gamma_1$ ,  $\Psi$  and  $\Pi$  are matrices of coefficients. This system may be solved for the evolution of the endogenous variables using standard methods described, e.g., in Sims (2000).

### 3.2. Responses to supply shocks with discretionary optimization

As we discussed in the beginning of this section, an alternative approach to studying optimal monetary policy is to assume that the central bank is not able to commit itself to act in a time inconsistent way. A consequence of this assumption is that the central bank is not able to exert the same amount of influence over private

<sup>15</sup>Sargent and Wallace (1975) argued that monetary policy should target the money supply since interest rate rules result in price level indeterminacy. It has since become clear that their result does not hold in general. Indeterminacy may be avoided by sufficiently strong feedback from endogenous variables (see, e.g. McCullum, 1981; Svensson and Woodford, 1999).

sector expectations as in the commitment case. Instead, the central bank must take the process by which the private sector forms its expectations as given. Central banks which behave in this way are said to optimize under discretion.

Solving for the evolution of the endogenous variables under discretion is somewhat more complicated than solving the commitment case. However, Söderlind (1999) presents a method to solve this case which is quite general and easily applicable to the model discussed in this paper. We therefore used Söderlind's method to solve for the evolution of the endogenous variables under discretion.

### 3.3. Some results

The results of the analysis described above can most usefully be visualized through impulse response functions for the endogenous variables. In Figs. 1–9 we report such impulse response functions (IRF). More precisely, we report

$$\frac{\partial}{\partial \eta_t} [E_t \psi_{t+j} - E_{t-1} \psi_{t+j}]$$

for  $\psi = \pi_{t+j}, x_{t+j}$ .

These impulse response functions are reported for four different values of  $\omega$ . The values of  $\omega$  which we choose to report are: (1)  $\omega = 0.01$ , which we refer to as the new Keynesian case, since in this case our Phillips curve corresponds closely with the purely forward-looking Phillips curve analyzed in Woodford (1999, 2003, Chapter 7); (2)  $\omega = 0.2$ , which we refer to as the Gali and Gertler case, since for this value of  $\omega$  our Phillips curve corresponds closely with the estimated model presented in Gali and Gertler (1999); (3)  $\omega = 0.7$ , which we refer to as the Fuhrer–Moore case, since for this value of  $\omega$  our Phillips curve corresponds closely to the model proposed in Fuhrer and Moore (1995); and (4)  $\omega = 0.99$ , which we refer to as the acceleration case.

For each of these cases we report results for commitment and discretionary optimization, under two different assumptions about the loss function the central bank chooses to minimize. We report results assuming the central bank seeks to minimize the loss function derived in Section 2. We refer to this loss function as the theoretical loss function. We also report results for a central bank which seeks to minimize a more standard specification of the loss function, attained by setting  $\lambda_2 = \lambda_3 = \lambda_4 = 0$  and  $\lambda_1 = 0.5$  in Eq. (28). We refer to this loss function as the traditional loss function.

Notice first that in the acceleration case, presented in Figs. 7 and 8, optimal policy with commitment and under discretion leads to almost identical IRFs. This reflects the fact that optimization under discretion is truly optimal in backward-looking systems. The reason we point this out here is to contrast this with the results reported in Figs. 1–6. It is evident from these other figures that as soon as forward-looking terms are introduced the optimal responses for the commitment case and the discretion case diverge. Discretionary optimization is, therefore, truly optimal only in backward-looking systems.

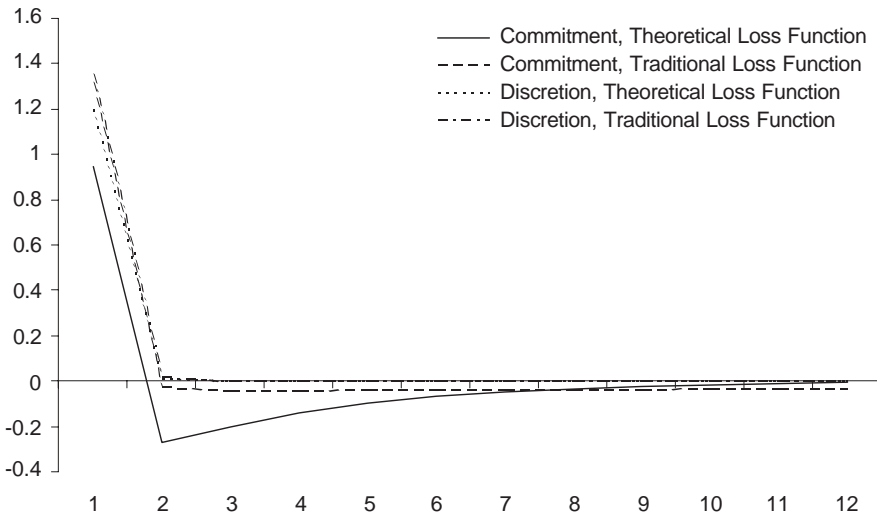


Fig. 1. The response of inflation to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.01$  under four different types of monetary policy.

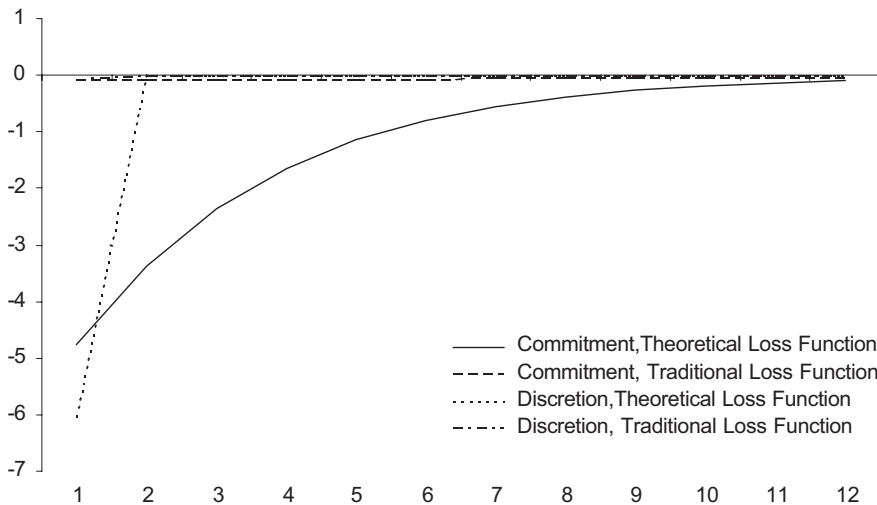


Fig. 2. The response of output to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.01$  under four different types of monetary policy.

Consider next how our two different assumptions about the loss function effect the optimal responses. The main difference in the optimal responses is that in the case of the traditional loss function output is stabilized to a much greater extent than in the case of the theoretical loss function. This is not surprising given that the biggest difference between these two loss functions is that the traditional loss function puts a much higher weight on output stabilization. This large emphasis on

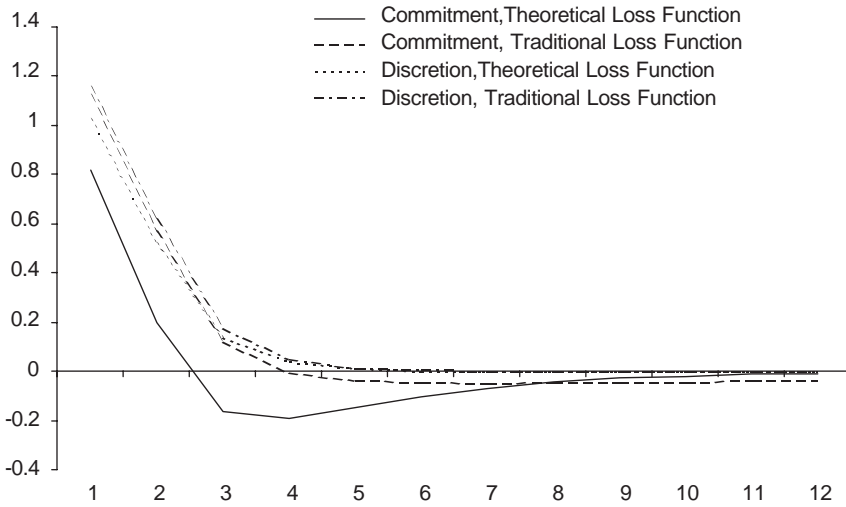


Fig. 3. The response of inflation to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.2$  under four different types of monetary policy.

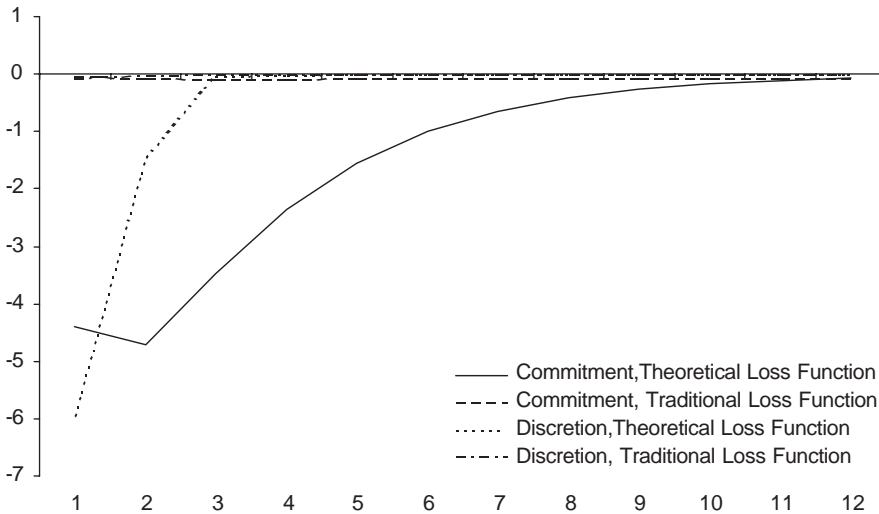


Fig. 4. The response of output to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.2$  under four different types of monetary policy.

output stabilization not only means that inflation is higher in the period of the shock but also that the inflation is much more persistent (see Figs. 5 and 6).

An interesting feature of these IRFs is that as long as the Phillips curve has a forward looking component it is optimal for a central bank that is committed to the theoretical loss function to induce a period of deflation following an inflationary cost push shock (see Fig. 9). In other words it is optimal in these cases to have the

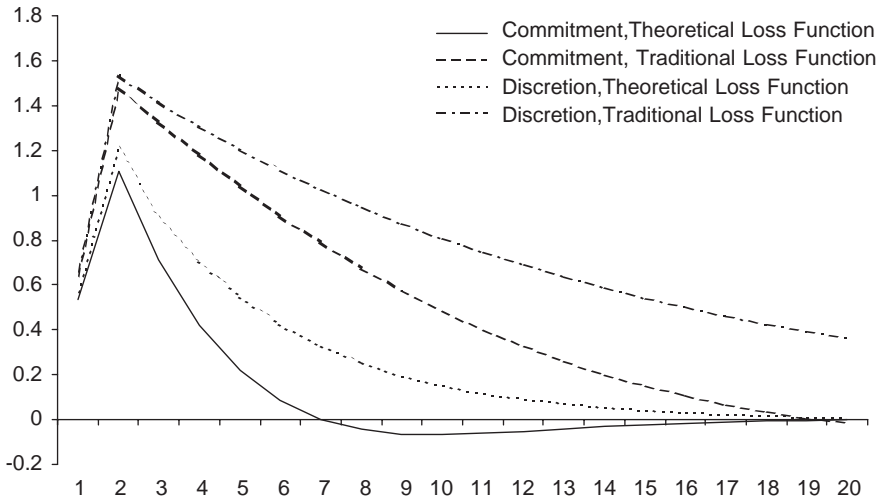


Fig. 5. The response of inflation to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.7$  under four different types of monetary policy.

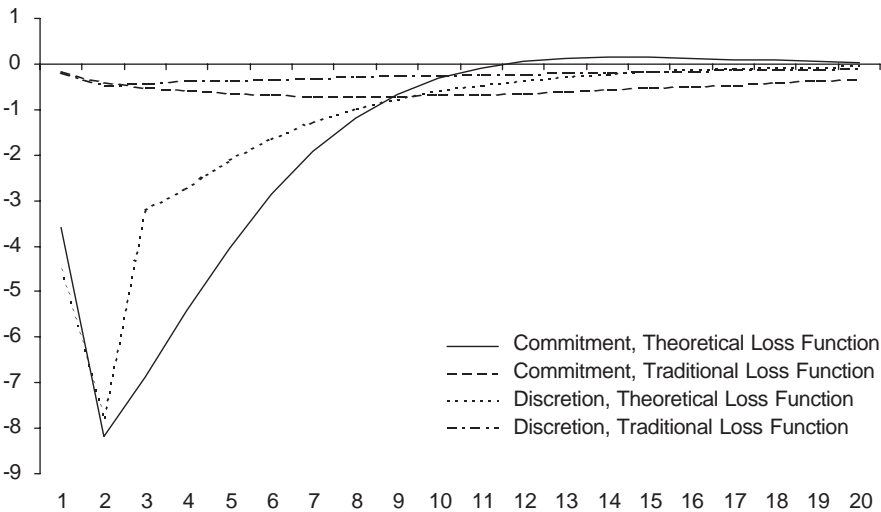


Fig. 6. The response of output to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.7$  under four different types of monetary policy.

inflation rate overshoot its target. The intuition behind this result is quite simple. If the private sector understands that the central bank will act in this way, the future deflation will be incorporated into the private sector's current inflationary expectations. The response of private sector inflationary expectation to an inflationary cost push shock will therefore be less violent than it otherwise would

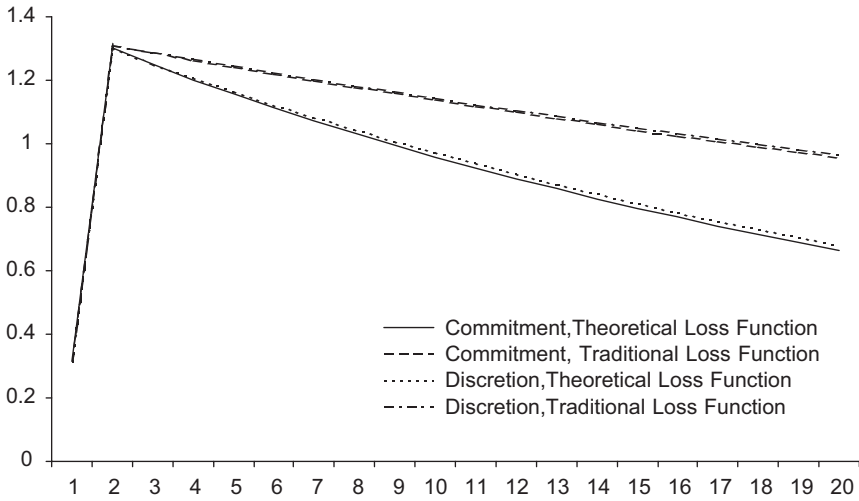


Fig. 7. The response of inflation to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.99$  under four different types of monetary policy.

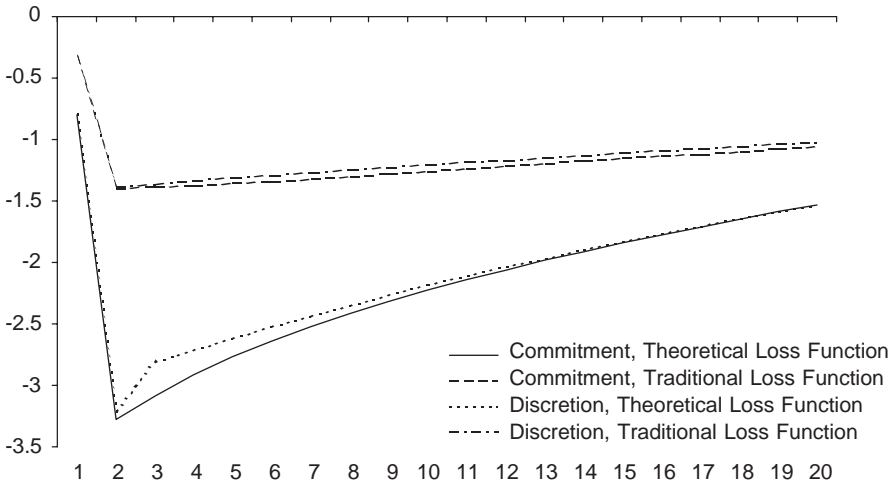


Fig. 8. The response of output to a  $\hat{\theta} = 0.2$  shock when  $\omega = 0.99$  under four different types of monetary policy.

have been, which in turn yields less actual inflation in the period of the shock. The benefits of lower inflation in the period of the shock turn out to outweigh the loss associated with subsequently carrying out the deflationary period.

In contrast, a central bank which optimizes under discretion cannot take advantage of this effect, since the private sector understands that it will renege on

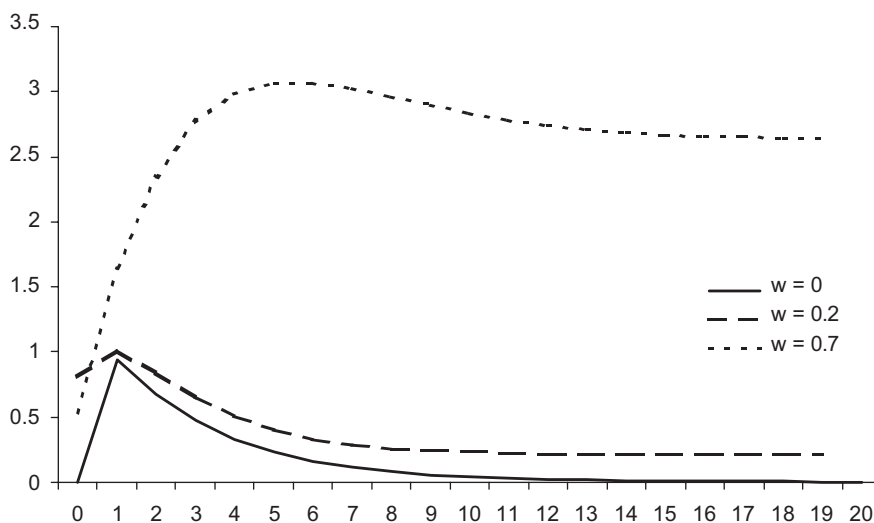


Fig. 9. The response of the price level to a  $\hat{\theta} = 0.2$  shock under optimal monetary policy with commitment using the theoretical loss function for three values of  $\omega$ .

carrying out the deflationary period once the initial reaction to the shock has passed. This inability to carry out earlier commitments lies at the heart of the sub-optimality of discretionary optimization.

Vestin (2000) shows that in a forward-looking model that corresponds closely to our new Keynesian case a central bank that pursues price-level targeting without being able to commit to time inconsistent policies is able to replicate the optimal policy under commitment. In this purely forward-looking model the price-level target turns out to force the central bank to take account of past inflation in exactly the same way a central bank optimizing the theoretical loss function under commitment would. It is evident from Fig. 9 that price-level targeting will no longer be optimal in a model with both forward-looking and backward-looking price setters. The optimal policy is in some sense “between” a price-level target where all movements in the price level are fully reversed and an inflation target where bygones are treated as bygones.

As was mentioned earlier there is little consensus within the empirical literature on the relative importance of forward- versus backward-looking terms in the Phillips curve. However, most of the literature is able to reject both the purely forward-looking and purely backward-looking cases. Thus, it is especially interesting to compare optimal responses for the intermediate cases.

Comparing the Gali and Gertler case with the Fuhrer–Moore case we can see at least three substantial differences. First, the size of the deflation which it is optimal to induce relative to the initial spurt of inflation is much greater in the Gali and Gertler case. As we can see from Fig. 9, in the Gali and Gertler case the long run increase in the price level resulting from a purely transitory cost push shock is only about 20% of the increase of the price level over the first two periods. In the Fuhrer–Moore case

Table 2  
Welfare comparisons

$w$	Theoretical Loss Function			Traditional Loss Function		
	Commitment	Discretion	Difference (%)	Commitment	Discretion	Difference (%)
0.01	1	1.26	26.3	1	1.03	3.2
0.1	1.13	1.43	26.6	0.94	0.97	3.5
0.2	1.30	1.65	26.7	0.98	1.02	4.3
0.3	1.55	1.96	26.4	1.15	1.21	5.5
0.4	1.95	2.44	24.9	1.51	1.62	7.5
0.5	2.60	3.16	21.5	2.21	2.47	11.6
0.6	3.66	4.25	16.2	3.72	4.49	20.8
0.7	5.46	6.01	10.1	7.25	9.46	30.4
0.8	8.85	9.28	4.8	14.2	16.8	18.5
0.9	17.9	18.1	1.3	24.3	25.5	5.2
0.99	145	145	0.0	36.4	36.5	0.1

The welfare measure has been normalized to one in the new Keynesian case under commitment for both loss functions.

the deflation is minimal. Only about 20% of the short term rise in the price level resulting from an inflationary cost push shock is reversed in the long run.

Second, the relative size of the output and inflation responses are very different in these two cases. In the Fuhrer–Moore case it is optimal to endure a much more severe contraction of output in order to avoid getting too much inflation into the system since inflation is persistent and therefore very costly to get rid of.

Third, in the Fuhrer–Moore case the effect of the shock is felt more strongly in the period after the shock than in the period of the shock. The explanation of this is that the  $\hat{\theta}_1 = 0.2$  shock we are assuming translates into  $\eta_1 = 0.35$  and  $\eta_2 = 0.50$  in the Fuhrer–Moore case while it translates into  $\eta_1 = 0.90$ ,  $\eta_2 = 0.22$  in the Galí and Gertler case (see Eq. (A.8)). As a consequence of this a substantial part of the difference in the impulse responses in these two case results from the difference in the way the  $\theta_1$  shock hits the economy as opposed to a difference in the policy response.

Another interesting way to compare optimal monetary policy in the cases we have been discussing is to calculate the loss incurred by the economy from a cost push shock in the different cases. Table 2 reports this loss for the shock considered in the figures as well as the difference in loss between commitment and discretion. A priori one might think that the benefit of being able to commit would fall monotonically as the proportion of backward-looking agents in the economy rises. Surprisingly this does not turn out to be the case in general. In the case of the theoretical loss function this difference is 26.3% when  $\omega = 0.01$ ; it rises modestly to 26.7% when  $\omega = 0.2$  and then falls to zero. In the case of the traditional loss function, however, the gains from commitment are modest for small values of  $\omega$ , but rise to 30.5% when  $\omega = 0.7$  and then fall to zero.

To understand this result notice that in the case of the theoretical loss functions for small  $\omega$  the optimal response of output to the shock is close to zero with and



without commitment (see Figs. 1–4). So there is as sense in which the optimal responses are a “corner solution”. We calculated the difference in loss between commitment and discretion for even higher values of  $\lambda_1$  and found that as  $\lambda_1 \rightarrow \infty$  the response of output and the difference in the loss goes to zero for all values of  $\omega$ .<sup>16</sup> In this limit the costs of even a mild recession are not worth bearing so the central bank lets the entire shock pass into the price level. Since there is only one way to do this there is no difference between the central bank’s response with and without commitment.

Evidently the gains from commitment are not just a peculiarity of the purely forward-looking case. For a large range of relative weights of inflation and output in the objective function of the central bank the gains from commitment are substantial even when a large fraction of the agents in the economy set prices in a backward-looking way.

#### 4. Conclusions

In this paper, we have extended the benchmark new Keynesian macro-model by making the Phillips curve a convex combination of a forward-looking term and a backward-looking term. This was motivated by our assumption that a fraction  $\omega$  of the producers in the economy set their prices according to a rule of thumb. We have seen that the main features of optimal policy in the purely forward looking case, such as the importance of commitment, carry over to this hybrid case. However, we have also seen that some features of the solution change in important ways. In our primarily backward looking cases it is optimal to bring inflation back down to zero in a gradual manner instead of the immediate overshooting that characterizes the purely forward-looking case. Also, in the backward-looking cases it is optimal to endure a much larger contraction of output in order to avoid getting too much inflation into the system.

These features of our hybrid cases seem to correspond quite well with actual central bank policy. The sharp overshooting of inflation in the period immediately following a supply shock which is optimal in the purely forward looking case does not seem to correspond to the way actual central banks react to supply shocks. Quite to the contrary, actual central banks often seek to gradually bring inflation back in line with their target. The policies of both the Bundesbank and the Federal Reserve in the early 1990s are a good example of this type of behavior.

Woodford (1999, Chapter 7) notes that an important feature of recent monetary policy by central banks such as the Federal Reserve is a high degree of interest rate inertia. Woodford argues that this type of behavior can be explained as being a feature of the optimal response of a central bank optimizing under commitment in a purely forward-looking economy. This explanation is however not consistent with the behavior of inflation in actual economies. Such behavior within a purely

---

<sup>16</sup>The shape of the difference in loss as a function of  $\omega$  stays the same as  $\lambda_1$  rises, i.e., increasing up to a point and then decreasing to zero. The value of  $\omega$  at which the difference is maximized goes to one as  $\lambda_1$  rises but the size of the maximum difference falls to zero.

forward-looking model should produce sharp overshooting of inflation. The actual behavior of inflation is more in line with the responses produced by a central bank optimizing with commitment in our hybrid cases.

The theoretical discussion in Section 2 also brought some interesting points to light. We added a time varying income tax to the model and made the monopoly power of the producers in the economy stochastic. We then showed that both of these extensions of the benchmark model resulted in a cost push supply shocks being added to the Phillips curve. However, we also found that reasonably sized tax shocks resulted in only miniscule cost push shocks while reasonably sized shocks to the market power of producers resulted in large shocks. We were actually quite surprised to see how much instability such shocks could create.

The model studied in this paper is a closed economy model. In future research, we are particularly interested in extending the model to a small open economy setting. Surprisingly little work to date has sought to analyze the optimal role of the exchange rate in the monetary policy of small open economies. A particularly interesting feature of this extension is the derivation of an appropriate central bank loss function in such a setting. Another important extension would be to derive a fully general second-order approximation to welfare and analyze optimal policy given this loss function and a second-order approximation to the structural equations of our model.

A feature of the our model that we are particularly uneasy about is the extremely low weight which the central bank loss function we derive puts on deviations of output from potential. As a result of this low value the optimal tradeoff between stabilizing output and stabilizing inflation is seriously skewed towards the stabilization of inflation, much more so than we think is reasonable. It is possible that this is due to the fact that all the frictions that we have introduced in our model are frictions to price adjustments. If frictions were introduced evenhandedly to every part of the model this would probably raise the relative weight on the output gap in the central bank loss function. The following two types of frictions, for instance, seem likely to become important pieces of more realistic models in this genre: rule-of-thumb consumers such as the ones introduced by [Campbell and Mankiw \(1989\)](#) and labor market frictions which result in persistence in the level of unemployment. Both of these types of frictions would most likely raise the relative weight on the output gap in the central bank loss function. Hopefully, we will be able to shed some light on these issues in future research.

## **Appendix A. Log-Linearization of the supply block**

In this appendix, we present details of the log-linearization of the supply block of our model which leads to Eq. (21) in the main text. We begin by log-linearizing Eqs. (16)–(19):

$$\pi_t = \frac{1 - \alpha}{\alpha} ((1 - \omega)\hat{p}_t^f + \omega\hat{p}_t^i) - \frac{\bar{\theta}}{\alpha} \hat{\theta}_t, \quad (\text{A.1})$$

$$E_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left\{ (\sigma^{-1} + \psi^{-1})x_T - \frac{\bar{\tau}}{1-\bar{\tau}}\hat{\tau}_T - \frac{1 + (\bar{\theta} - 1)\bar{\theta}^2\psi^{-1}}{\bar{\theta} - 1}\hat{\theta}_T - (1 + \psi^{-1}\bar{\theta})(\hat{p}_t^f - \frac{\alpha\beta\pi_{T+1}}{1-\alpha\beta}) \right\} = 0, \tag{A.2}$$

$$\hat{p}_t^b = \hat{p}_{t-1}^* - \pi_t + \pi_{t-1} + \delta x_{t-1}, \tag{A.3}$$

$$\hat{p}_t^* = (1 - \omega)\hat{p}_t^f + \omega\hat{p}_t^b, \tag{A.4}$$

where  $\hat{p}_t^f$ ,  $\hat{p}_t^b$  and  $\hat{p}_t^*$  denote percent deviations of  $p_t^f/P_t$ ,  $p_t^b/P_t$  and  $p_t^*/P_t$ , respectively, from their steady-state values of one,  $\hat{\tau}_t$  denotes percent deviations of the tax rate from its steady-state value,  $\bar{\tau}$ , and the parameter  $\psi$  is defined as

$$\psi = \frac{v_y}{v_{yy}\bar{Y}}.$$

In Eq. (A.2) we use the fact that

$$(\sigma^{-1} + \psi^{-1})\log\left(\frac{Y_t^n}{\bar{Y}}\right) = \left(\frac{u_{C\xi}}{u_C} - \frac{v_{y\xi}}{v_y}\right)\xi_t.$$

By manipulating Eqs. (A.1)–(A.4) we are now able to derive the Phillips curve of our model. First we notice that since our attention is limited to bounded solutions, and since  $|a| < 1$ , Eq. (A.2) can equivalently be written in quasi-differenced form as

$$\hat{p}_t^f = \alpha\beta E_t \hat{p}_{t+1}^f + \frac{1 - \alpha\beta}{1 + \psi^{-1}\bar{\theta}} \left( (\sigma^{-1} + \psi^{-1})x_t - \frac{\bar{\tau}}{1-\bar{\tau}}\hat{\tau}_t - \frac{1 + (\bar{\theta} - 1)\bar{\theta}^2\psi^{-1}}{\bar{\theta} - 1}\hat{\theta}_t \right) + \alpha\beta E_t \pi_{t+1}. \tag{A.5}$$

Next, we combine Eqs. (A.1), (A.3) and (A.4) in order to eliminate  $\hat{p}_t^b$  and  $\hat{p}_t^*$ . This gives

$$\left(\omega + \frac{\alpha}{1-\alpha}\right)\pi_t = (1 - \omega)\hat{p}_t^f + \frac{\omega}{1-\alpha}\pi_{t-1} + \omega\delta x_{t-1} - \frac{\bar{\theta}}{1-\alpha}\hat{\theta}_t + \frac{\omega\bar{\theta}}{1-\alpha}\hat{\theta}_{t-1}. \tag{A.6}$$

Finally, we combine Eqs. (A.5) and (A.6), eliminating  $\hat{p}_t^f$ , and get

$$\pi_t = \lambda_f \beta E_t \pi_{t+1} + \lambda_b \pi_{t-1} + \kappa_1 x_t + \kappa_2 x_{t-1} + \eta_t, \tag{A.7}$$

where

$$\lambda_f = \frac{\alpha}{\omega(1 - \alpha + \alpha\beta) + \alpha}, \quad \lambda_b = \frac{\omega}{\omega(1 - \alpha + \alpha\beta) + \alpha},$$

$$\begin{aligned} \kappa_1 &= \frac{(1 - \alpha)(1 - \alpha\beta)(1 - \omega)(\sigma + \psi)}{\sigma(\omega(1 - \alpha + \alpha\beta) + \alpha)(\psi + \bar{\theta})} - \frac{\alpha\beta\omega\delta(1 - \alpha)}{\omega(1 - \alpha + \alpha\beta) + \alpha} \\ \kappa_2 &= \frac{\omega\delta(1 - \alpha)}{\omega(1 - \alpha + \alpha\beta) + \alpha} \\ \eta_t &= \frac{(1 - \alpha)(1 - \alpha\beta)(1 - \omega)}{(\omega(1 - \alpha + \alpha\beta) + \alpha)(1 + \psi^{-1}\bar{\theta})\bar{\theta}} \hat{\tau}_t + \frac{\alpha\beta\bar{\theta}}{\omega(1 - \alpha + \alpha\beta) + \alpha} E_t \hat{\theta}_{t+1} \\ &\quad + \left( \frac{(1 - \alpha\beta\omega)\bar{\theta}}{\omega(1 - \alpha + \alpha\beta) + \alpha} - \frac{(1 - \alpha)(1 - \alpha\beta)(1 - \omega)(\psi + (\bar{\theta} - 1)\bar{\theta}^2)}{(\omega(1 - \alpha + \alpha\beta) + \alpha)(\bar{\theta} - 1)(\psi + \bar{\theta})} \right) \hat{\theta}_t \\ &\quad + \frac{\omega\bar{\theta}}{\omega(1 - \alpha + \alpha\beta) + \alpha} \hat{\theta}_{t-1}. \end{aligned} \tag{A.8}$$

**Appendix B. Derivation of Eq. (26)**

In this appendix, we present details of the derivation of Eqs. (26) from Eqs. (24) and (25). First we take a second-order Taylor approximation of the first term in Eq. (25):

$$u(Y_t, \xi_t) = \bar{u} + u_C \hat{Y}_t + u_\xi \xi_t + \frac{1}{2} u_{CC} \hat{Y}_t^2 + u_{C\xi} \xi_t \hat{Y}_t + \frac{1}{2} \xi_t' u_{\xi\xi} \xi_t + \mathcal{O}(\|\xi\|^3), \tag{B.1}$$

where  $\bar{u} = u(\bar{Y}; 0)$ , and  $\hat{Y}_t = Y_t - \bar{Y}$ . Next we notice that the second-order Taylor approximation of  $\hat{Y}_t = \log(Y_t/\bar{Y})$  is

$$Y_t = \bar{Y}(1 + \hat{Y}_t + \frac{1}{2} \hat{Y}_t^2) + \mathcal{O}(\|\xi\|^3). \tag{B.2}$$

We use this expression to write Eq. (B.1) in terms of  $\hat{Y}_t$  as

$$u(Y_t; \xi_t) = \bar{Y}u_C \left\{ \hat{Y}_t + \frac{1}{2}(1 - \sigma^{-1})\hat{Y}_t^2 + \frac{u_{C\xi}}{u_C} \xi_t \hat{Y}_t \right\} + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3). \tag{B.3}$$

Next we take a second-order Taylor approximation of  $v_y(v_t(z); \xi_t)$  and with similar manipulations get that

$$\begin{aligned} v(y_t(z); \xi_t) &= \bar{Y}v_y \left\{ \hat{y}_t(z) + \frac{1}{2}(1 - \psi^{-1})\hat{y}_t(z)^2 + \frac{v_{y\xi}}{v_y} \xi_t \hat{y}_t(z) \right\} \\ &\quad + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3), \end{aligned} \tag{B.4}$$

where  $\hat{y}_t(z) = \log(y_t(z)/\bar{Y})$ . Integrating this over  $z$  and using the fact that  $u_C = v_y$  in the steady state we get that

$$\begin{aligned} \int_0^1 v(y_t(z); \xi_t) dz &= \bar{Y}u_C \left\{ E_z \hat{y}_t(z) + \frac{1}{2}(1 + \psi^{-1})[(E_z \hat{y}_t(z))^2 + \text{var}_z \hat{y}_t(z)] \right. \\ &\quad \left. + \frac{v_{y\xi}}{v_y} \xi_t E_z \hat{y}_t(z) \right\} + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3), \end{aligned} \tag{B.5}$$

where  $E_z \hat{y}_t(z)$  denotes the mean value of  $\hat{y}_t(z)$  over  $z$ , and  $\text{var}_z \hat{y}_t(z)$  denotes the variance of  $\hat{y}_t(z)$  over  $z$ . In order to eliminate  $E_z \hat{y}_t(z)$  from the above expression, we

use the second-order approximation of Eq. (3). Since  $\hat{Y}_t - E_z \hat{y}_t(z) = \mathcal{O}(\|\xi\|^2)$  this can be written simply as

$$\hat{Y}_t = E_z \hat{y}_t(z) + \frac{1}{2}(1 - \bar{\theta}^{-1})\text{var}_z \hat{y}_t(z) + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3). \tag{B.6}$$

Combining equations (B.5) and (B.6) we get that

$$\int_0^1 v(y_t(z); \xi_t) dz = \bar{Y}u_C \left\{ \hat{Y}_t + \frac{1}{2}(1 + \psi^{-1})\hat{Y}_t^2 + \frac{1}{2}(\bar{\theta}^{-1} + \psi^{-1})\text{var}_z \hat{y}_t(z) + \frac{v_{y\xi}}{v_y} \xi_t \hat{Y}_t \right\} + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3). \tag{B.7}$$

Combining Eqs. (B.3) and (B.7) we get that

$$U_t = -\frac{\bar{Y}u_C}{2} \{(\sigma^{-1} + \psi^{-1})x_t + (\bar{\theta}^{-1} + \psi^{-1})\text{var}_z \hat{y}_t(z)\} + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3). \tag{B.8}$$

Finally, we notice that it follows from Eq. (4) that

$$\text{var}_z \hat{y}_t(z) = \bar{\theta}^2 \text{var}_z p_t(z) + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3).$$

Combining this equation with Eq. (B.8) we get Eq. (26) in the main text.

**Appendix C. Derivation of the degree of price dispersion**

In this appendix, we present details of the derivation of the degree of price dispersion in the economy which is used in the main text to derive Eqs. (27) and (28). We have assumed that a fraction  $1 - \alpha$  of the households in the economy are able to change their prices in each period. Consequently, the distribution of prices,  $\{p_t(z)\}$ , at time  $t$  consists of  $\alpha$  times the distribution of prices at time  $t - 1$ , plus two atoms of size  $(1 - \alpha)(1 - \omega)$  and  $(1 - \alpha)\omega$  at the two new prices,  $p_t^f$  and  $p_t^b$ , respectively. Define

$$\bar{P}_t = E_z \log p_t(z) \text{ and } A_t = \text{var}_z(\log p_t(z))$$

and observe that

$$\bar{P}_t - \bar{P}_{t-1} = E_z[\log p_t(z) - \bar{P}_{t-1}]. \tag{C.1}$$

Using the recursive characterization of the distribution of prices we can replace the right-hand side of Eq. (C.1) with

$$\alpha E_z[\log p_{t-1}(z) - \bar{P}_{t-1}] + (1 - \alpha)(1 - \omega)(\log p_t^f - \bar{P}_{t-1}) + (1 - \alpha)\omega(\log p_t^b - \bar{P}_{t-1}). \tag{C.2}$$

Noticing that the first term in Eq. (C.2) is equal to zero we get

$$\begin{aligned} \bar{P}_t - \bar{P}_{t-1} &= (1 - \alpha)(1 - \omega)(\log p_t^f - \bar{P}_{t-1}) + (1 - \alpha)\omega(\log p_t^b - \bar{P}_{t-1}) \\ &= (1 - \alpha)(\log p_t^* - \bar{P}_{t-1}). \end{aligned} \tag{C.3}$$

Similarly, we may derive an expression for  $\Delta_t$ :

$$\begin{aligned}\Delta_t &= \text{var}_z[\log p_t(z) - \bar{P}_{t-1}] \\ &= E_z\{[\log p_t(z) - \bar{P}_{t-1}]^2\} - (E_z \log p_t(z) - \bar{P}_{t-1})^2.\end{aligned}\quad (\text{C.4})$$

Again, using the recursive characterization of the distribution of prices we see that the first term on the right-hand side of Eq. (C.4) can be rewritten as

$$\begin{aligned}E_z\{[\log p_t(z) - \bar{P}_{t-1}]^2\} &= \alpha E_z\{[\log p_{t-1}(z) - \bar{P}_{t-1}]^2\} + (1-\alpha)(1-\omega)(\log p_t^f - \bar{P}_{t-1})^2 \\ &\quad + (1-\alpha)\omega(\log p_t^b - \bar{P}_{t-1})^2.\end{aligned}\quad (\text{C.5})$$

Using

$$\log p_t^* = (1-\omega)\log p_t^f + \omega \log p_t^b,$$

$$\log p_t^b = \log p_{t-1}^* + \pi_{t-1} + \delta x_{t-1},$$

$$\bar{P}_t = \log P_t + \mathcal{O}(\|\xi\|^2), \quad (\text{C.6})$$

we can further develop the last two terms on the right-hand side of equation (C.5):

$$\begin{aligned}\log p_t^b - \bar{P}_{t-1} &= \log p_{t-1}^* + \pi_{t-1} + \delta x_{t-1} - \bar{P}_{t-1} \\ &= \log p_{t-1}^* - \bar{P}_{t-2} - (\bar{P}_{t-1} - \bar{P}_{t-2}) + \pi_{t-1} + \delta x_{t-1} \\ &= \log p_{t-1}^* - \bar{P}_{t-2} - \pi_{t-1} + \pi_{t-1} + \delta x_{t-1} + \mathcal{O}(\|\xi\|^2) \\ &= \log p_{t-1}^* - \bar{P}_{t-2} + \delta x_{t-1} + \mathcal{O}(\|\xi\|^2),\end{aligned}\quad (\text{C.7})$$

$$\begin{aligned}\log p_t^f - \bar{P}_{t-1} &= \frac{1}{1-\omega} \log p_t^* - \frac{\omega}{1-\omega} (\log p_{t-1}^* + \pi_{t-1} + \delta x_{t-1}) - \bar{P}_{t-1} \\ &= \frac{1}{1-\omega} (\log p_t^* - \bar{P}_{t-1}) - \frac{\omega}{1-\omega} (\log p_{t-1}^* + \pi_{t-1} + \delta x_{t-1} - \bar{P}_{t-1}) \\ &= \frac{1}{1-\omega} (\log p_t^* - \bar{P}_{t-1}) - \frac{\omega}{1-\omega} (\log p_{t-1}^* - \bar{P}_{t-2} + \delta x_{t-1}) \\ &\quad + \mathcal{O}(\|\xi\|^2).\end{aligned}\quad (\text{C.8})$$

Finally, by combining Eqs. (C.3)–(C.8) we obtain

$$\begin{aligned}\Delta_t &= \alpha \Delta_{t-1} + \frac{\alpha}{(1-\alpha)} \pi_t^2 + \frac{\omega}{(1-\alpha)(1-\omega)} \Delta \pi_t^2 + \frac{(1-\alpha)\omega\delta^2}{1-\omega} x_{t-1}^2 \\ &\quad + \frac{2\omega\delta}{1-\omega} \Delta \pi_t x_{t-1} + \mathcal{O}(\|\xi\|^3),\end{aligned}\quad (\text{C.9})$$

where  $\Delta \pi_t = \pi_t - \pi_{t-1}$ , i.e.  $\Delta \pi_t$  is the acceleration of the price level at time  $t$ . Solving Eq. (C.9) forward, starting with an initial degree of price dispersion,  $\Delta_{-1}$ , in the

period before the first period we get

$$\Delta_t = \alpha^{t+1} \Delta_{-1} + \sum_{s=0}^t \alpha^{t-s} \left( \frac{\alpha}{(1-\alpha)} \pi_t^2 + \frac{\omega}{(1-\alpha)(1-\omega)} \Delta \pi_t^2 + \frac{(1-\alpha)\omega\delta^2}{1-\omega} x_{t-1}^2 + \frac{2\omega\delta}{1-\omega} \Delta \pi_t x_{t-1} \right) + \mathcal{O}(\|\xi\|^3).$$

We can now take the discounted present value of these terms for all periods  $t \geq 0$ :

$$\sum_{t=0}^{\infty} \beta^t \Delta_t = \frac{1}{1-\alpha\beta} \sum_{t=0}^{\infty} \beta^t \left( \frac{\alpha}{(1-\alpha)} \pi_t^2 + \frac{\omega}{(1-\alpha)(1-\omega)} \Delta \pi_t^2 + \frac{(1-\alpha)\omega\delta^2}{1-\omega} x_{t-1}^2 + \frac{2\omega\delta}{1-\omega} \Delta \pi_t x_{t-1} \right) + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3). \tag{C.10}$$

Here, we have used the fact that  $\Delta_{-1}$  is independent of policy chosen to apply in periods  $t \geq 0$ . We can now Substitute this equation into Eq. (26) in the main text. This gives us Eqs. (27) and (28).

## References

- Ball, L., 1997. Efficient rules for monetary policy. NBER Working Paper No. 5952, March 1997.
- Blake, A.P., Westaway, P.F., 1996. Credibility and the effectiveness of inflation targeting regimes. *Manchester School* 64, 28–50.
- Blinder, A.S., 1994. On sticky prices: academic theories meet the real world. In: Mankiw, N.G. (Ed.), *Monetary Policy*. University of Chicago Press, IL, pp. 117–150.
- Blinder, A.S., 1998. *Central Banking in Theory and Practice*. MIT Press, Cambridge, MA.
- Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12, 383–398.
- Campbell, J.Y., Mankiw, N.G., 1989. Consumption, income, and interest rates: reinterpreting the time series evidence. In: Blanchard, O.J. (Ed.), *NBER Macroeconomic Annual 1989*. MIT Press, Cambridge, MA, pp. 185–216.
- Clarida, R., Gali, J., Gertler, M., 1999. The science of monetary policy: a new Keynesian perspective. *Journal of Economic Literature* 37, 1661–1707.
- Farmer, R.E.A., 1999. *Macroeconomics of Self-Fulfilling Prophecies*. MIT Press, Cambridge, MA.
- Friedman, M., 1968. The role of monetary policy. *American Economic Review* 58, 1–19.
- Fuhrer, J.C., 1997. The (un)importance of forward-looking behavior in price specifications. *Journal of Money, Credit and Banking* 29, 338–350.
- Fuhrer, J.C., Moore, G.R., 1995. Inflation persistence. *Quarterly Journal of Economics* 109, 197–230.
- Gali, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. *Journal of Monetary Economics* 44, 195–222.
- Jonsson, G., 1997. Monetary politics and unemployment persistence. *Journal of Monetary Economics* 39, 303–325.
- Kydland, F.E., Prescott, E.C., 1977. Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy* 85, 473–491.
- Lucas, R.E., 1976. Econometric policy evaluation: a critique. *Carnegie–Rochester Conference Series on Public Policy* 1, 19–46.
- McCullum, B.T., 1981. Price level determinacy with an interest rate policy rule and rational expectations. *Journal of Monetary Economics* 8, 319–329.

- Obstfeld, M., Rogoff, K., 1996. *Foundations of International Macroeconomics*. MIT Press, Cambridge, MA.
- Phelps, E.S., 1967. Phillips curves, expectations of inflation and optimal unemployment over time. *Economica* 34, 254–281.
- Phillips, A.W., 1958. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Economica* 25, 283–299.
- Roberts, J.M., 1995. New Keynesian economics and the Phillips curve. *Journal of Money, Credit and Banking* 27, 975–984.
- Roberts, J.M., 1997. Is inflation sticky? *Journal of Monetary Economics* 39, 173–196.
- Roberts, J.M., 1998. Inflation expectations and the transmission of monetary policy. *Finance and Economics Discussion Series*, Federal Reserve Board, October 1998.
- Roberts, J.M., 2000. How well does the new Keynesian sticky-price model fit the data? *Finance and Economics Discussion Series*, Federal Reserve Board, February 2000.
- Rotemberg, J.J., Woodford, M., 1997. An optimizing-based econometric framework for the evaluation of monetary policy. In: Bernanke, B.S., Rotemberg, J.J. (Eds.), *NBER Macroeconomics Annual 1997*. MIT Press, Cambridge, MA, pp. 297–346.
- Rotemberg, J.J., Woodford, M., 1999. Interest rate rules in an estimated sticky price model. In: Taylor, J.B. (Ed.), *Monetary Policy Rules*. University of Chicago Press, Chicago, IL, pp. 57–119.
- Sargent, T.J., 1993. *Rational Expectations and Inflation*, 2nd Edition. HarperCollins Publishers, New York, pp. 40–119.
- Sargent, T.J., Wallace, N., 1975. Rational expectations, the optimal monetary instrument, and the optimal money supply rule. *Journal of Political Economy* 83, 241–254.
- Sims, C.A., 2000. Solving rational expectations models. Unpublished, Princeton University, January 2000.
- Söderlind, P., 1999. Solution and estimation of RE macromodels with optimal policy. *European Economic Review* 43, 813–823.
- Svensson, L.E.O., 1997a. Inflation forecast targeting: implementing and monitoring inflation targets. *European Economic Review* 41, 1111–1146.
- Svensson, L.E.O., 1997b. Optimal inflation targets, ‘conservative’ central banks, and linear inflation contracts. *American Economic Review* 87, 98–114.
- Svensson, L.E.O., Woodford, M., 1999. Implementing optimal policy through inflation-forecast targeting. Unpublished, Princeton University, November 1999.
- Vestin, D., 2000. Price-level targeting versus inflation targeting in a forward-looking model. *Sveriges Riksbank Working Paper No. 106*, June 2000.
- Walsh, C.E., 1998. *Monetary Theory and Policy*. MIT Press, Cambridge, MA.
- Woodford, M., 1996. Control of the public debt: a requirement for price stability? *NBER Working Paper No. 5684*, July 1996.
- Woodford, M., 1998. Doing without money: controlling inflation in a post-monetary world. *Review of Economic Dynamics* 1, 173–219.
- Woodford, M., 1999. Commentary: how should monetary policy be conducted in an era of price stability? In: *New Challenges for Monetary Policy*, Federal Reserve Bank of Kansas City, Kansas, pp. 277–316.
- Woodford, M., 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, NJ.
- Yun, T., 1996. Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics* 37, 345–370.