CENTRAL BANK OF ICELAND

WORKING PAPERS No. 11

OPTIMAL MONETARY POLICY IN AN ECONOMY WITH INFLATION PERSISTENCE

by

Jón Steinsson

December 2000

CENTRAL BANK OF ICELAND

Economics Department

OPTIMAL MONETARY POLICY IN AN ECONOMY WITH INFLATION PERSISTENCE

by

Jón Steinsson*

December 2000

Abstract

This paper presents a closed economy dynamic stochastic general equilibrium model with monopolistic competition and sticky prices. 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 theoretically appropriate central bank loss function for this model is derived. This loss function depends on the rate of change of inflation squared as well as inflation squared and the output gap squared. 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. Finally the optimal Taylor rule responses to cost push supply shocks are characterized. Since the economy considered in this paper is closed the effects of international linkages on optimal monetary policy are not considered. Such effects will however be the focus of future research by the author.

Keywords: Phillips Curve, Central Bank Loss Functions, Optimal Monetary Policy, Commitment, Discretion, Taylor Rules.

JEL Classification: E31, E32, E52, E58.

[0]Research Economist, Department of Economics, Central Bank of Iceland. I would like to thank Michael Woodford for priceless advice and encouragement. I would like to thank Gauti B. Eggertsson, Marc Giannoni, Emi Nakamura and Thórarinn G. Pétursson for valuable discussions and comments; and Sif Cortes for help with the paper's layout. Special thanks to Lonn Waters for moral support. The views expressed in this paper do not necessarily reflect those of the Central Bank of Iceland.

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 π_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 "Old Keynesian" Phillips curve. [1] 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 π_{t+1} at date t. We will refer to this as the "New Keynesian" Phillips curve.[2]

Neither of these two specifications, however, seems adequate to capture the behavior of inflation in actual economies. The Old Keynesian 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.[3] On the other hand, the New Keynesian Phillips curve fails to capture the fact that significant output losses typically accompany the reduction of inflation. According to the new Keynesian Phillips curve, if a central bank announces intentions to lower the rate of inflation, this can lead to a reduction in inflationary expectations,

[1] For recent examples of papers which specify the Phillips curve in this manner, see Ball (1997) and Svensson (1997).

[2] For examples of papers which use this specification of the Phillips curve, see Roberts (1995) and Woodford (1999c).

[3] For a particularly dramatic account of this, see Thomas Sargent's essay, "The Ends of Four Big Inflations," in Sargent (1993).

and hence a reduction in inflation, without having any effect on output. The empirical literature on disinflation does not confirm this prediction of the New Keynesian Phillips curve. Ball (1994) documents that output losses seem to be associated with disinflation across countries and over time.

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

$$\pi_t = \chi_1 E_t \pi_{t+1} + \chi_2 \pi_{t-1} + k x_t. \tag{1}$$

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 χ_1 and χ_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 χ_1 and χ_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 rules for different relative weights of χ_1 and χ_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.

The paper proceeds as follows. In section 2 we incorporate the hybrid Phillips curve into a stochastic general equilibrium model with monopolistic competition and sticky prices. In so doing, we make the same assumptions as Gali and Gertler (1999) with regards to price setting behavior. We give special attention to the derivation of an appropriate central bank loss function within the model. Woodford (1999b) has shown that for the benchmark New Keynesian model with a purely forward-looking Phillips curve the appropriate loss function is

$$L_t = \pi_t^2 + \lambda_y x_t^2.$$

We apply the methods used in Woodford (1999b) to our model and find that the appropriate loss function for our model depends not only on squared deviations of inflation from steady state and the squared deviations of output from potential, but also on the rate of change of inflation squared.

In section 3 we turn to the analysis of optimal responses of the economy to cost push supply shocks. We present solutions for optimal responses both for a central bank which is able to make credible commitments, and for one that is not able to make such commitments. We calculate impulse response functions for particular examples of cost push shocks and analyze the benefits of commitment and the effect of varying the relative importance of the forward- and backward-looking terms in the Phillips curve.

In section 4 we turn to the analysis of interest rate feedback rules. We calculate the optimal Taylor rule for our model for a range of different relative values of the forward- and backward-looking terms in the Phillips curve. We also present contour plots of the welfare loss associated with different Taylor rules. These plots are particularly useful for building intuition about the nature of optimal policy. From these figures we can see that the most important aspect of choosing an appropriate Taylor rule to stabilize the economy from the effects of cost push supply shocks is the relative size of the coefficients on inflation and the output gap, not their absolute values. Our results suggest that the relative size of responses to the output gap should be substantially smaller than in Taylor's original specification.

We conclude with a brief discussion 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

The rational expectations revolution in macroeconomics has profoundly altered the study of business cycles. 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.[4]

2.1 Household Preferences and Market Structure

The economy consists of a continuum of infinitely-lived households/producers[5] of measure 1. The households all have identical preferences, represented by

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

where β is a discount factor, ξ_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

$$C_t^i = \left[\int_0^1 c_t^i(z)^{\frac{\theta-1}{\theta}} dz \right]^{\frac{\theta}{\theta-1}} \tag{3}$$

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 equation (3) implies a constant elasticity of substitution between goods, equal to $\theta > 1$. As a result of this, each household possesses a certain degree of monopoly power in the good it produces.[6]

The utility function presented in equation (2) is time separable and separable between consumption and production. Furthermore, we assume that $u(C_t^i; \xi_t)$ is increasing and strictly quasi-concave, while $v(y_t(z); \xi_t)$ is increasing and strictly 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.[7] 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,

[4] Monopolistic competition was introduced to general equilibrium macromodels in Svensson (1986), and Blanchard and Kiyotaki (1987). Recent papers in this genre include Yun (1996), Woodford (1996), Obstfeld and Rogoff (1996, chapter 10) and Rotemberg and Woodford (1998,1999).

[5] These two terms are used interchangeably in this paper.

[6] This type of market structure was first proposed by Dixit and Stiglitz (1977).

[7] See Woodford (forthcoming) for a discussion of a model of this type with firms and a labor market.

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 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. More precisely, utility derived from liquidity services should be viewed as being separable from utility derived from the consumption of goods. If we were explicit about the role of money in our model, this would result in one extra first order condition, relating equilibrium real money balances to interest rates and output (an "LM curve"). Since we will be interested in formulating monetary policy by interest rate rules, this added first order condition would play no role in our analysis of policy. It would merely determine the money supply given the evolution of other variables in the model and can therefore safely be ignored without loss of generality. [8]

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, Z_t^i. This leads to familiar expressions for the demand for each individual good

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

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

$$P_t = \left[\int_0^1 p_t(z)^{1-\theta} dz\right]^{\frac{1}{1-\theta}}. \tag{5}$$

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. The derivation of equations (4) and (5) are given in appendix A.

The reason we choose this specification for the market structure of our economy is that we want the producers in the economy to be able to set prices. In a perfectly competitive economy producers have no such ability. In a monopolistically competitive setting, like the one we are assuming, each producer's market power gives him the ability to set the price of the good he produces. While monopolistic in this sense, the market structure we assume is also competitive since the households are small. Each household's pricing and spending decisions have only an

[8] See Woodford (1999a) for a discussion of a model of this type which is explicit about the role of money.

infinitesimal effect on aggregate demand, the aggregate price level, and the demand for its own output. Each household will therefore make its decisions taking these variables as given.

Assume for a moment that prices are perfectly flexible. Since all producers are identical and therefore have the same amount of market power, they will all seek to set their price to the same level relative to the aggregate price level. In equilibrium all relative prices must therefore be equal to one. In such an equilibrium it turns out to be optimal for each producer to set his price at a constant markup over marginal costs,

$$p_t(z) = \frac{\theta}{\theta - 1} S_t(z) = \mu S_t(z), \tag{6}$$

where $S_t(z)$ denotes nominal marginal costs (see appendix A for a derivation of equation (6)). Since prices are set above marginal costs, output in a monopolistically competitive economy is lower than in a competitive economy and the price level is higher. Money is however neutral in this setting as long as prices are prefectly flexible. Additional assumptions constraining the pricing decisions of producers are needed to create real affects due to monetary policy.

2.2 The Budget Constraint

We assume that there exist complete financial markets in the economy under consideration, 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

$$P_t C_t + E_t[R_{t,t+1} B_{t+1}] \leq B_t + p_t(z) y_t(z), \tag{7}$$

where B_t is the nominal value of the household's portfolio of financial assets brought into period t, 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

$$\frac{1}{1+i_t} = E_t[R_{t,t+1}]. \tag{8}$$

In order to move from the flow budget constraints to a single intertemporal budget constraint we must make three further assumptions. First, we must specify a borrowing constraint for the households. This is done to rule out "Ponzi schemes," in which households borrow enough each period to pay off debt carried over from the previous period and, in addition, finance as much consumption as they wish. Assume that financial wealth carried into the next period, B_{t+1}, satisfies the bound

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

with certainty, that is, in each state of the world which way 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,

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

Second, 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

$$\sum_{T=t+1}^{\infty} E_{t+1}[R_{t+1,T} p_T(z) y_T(z)] < \infty \tag{10}$$

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 equation (10) is satisfied.

The third assumption that is needed for the definition of the intertemporal budget constraint is that the nominal interest rate satisfy the lower bound

$$i_t \geq 0 \tag{11}$$

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 equation (11) is satisfied 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,

$$\sum_{s=t}^{\infty} E_t[R_{t,s} P_s C_s] \leq B_t + \sum_{s=t}^{\infty} E_t[R_{t,s} p_s(z) y_s(z)]. \tag{12}$$

2.3 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. The first thing to notice is that it is optimal for the household to exhaust its intertemporal budget constraint, i.e. equation (12) must hold with equality. The household's consumption decision is then a standard constrained optimization problem where equation (2) is maximized subject to equation (12). The resulting first order conditions are

$$u_C(C_t; \xi_t) = P_t \Lambda_t, \tag{13}$$

$$R_{t,T} \Lambda_t = \beta^{T-t} \Lambda_T, \tag{14}$$

where Λ_t is the marginal utility of nominal income at time t, that is, 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.[9]

In equilibrium markets must clear. The conditions for market clearing are

$$c_t(z) = y_t(z), \quad C_t = Y_t, \quad B_t = 0, \tag{15}$$

for all t and all z, where $c_t(z)$ denotes total consumption of good z, and Y_t denotes total output. Combining equations (13), (14), (15), and (8) we get a more familiar Euler equation for household consumption

$$\beta E_t \left\{ \frac{u_C(Y_{t+1}; \xi_{t+1})}{u_C(Y_t; \xi_t)} \frac{P_t}{P_{t+1}} \right\} = \frac{1}{1 + i_t}. \tag{16}$$

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.

[9] The third first order condition which results form 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. The flexible price case is solved in appendix A since it gives a useful expression for the "natural rate" of output which we use later on in the paper.

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 change their prices in each period. More precisely, in each period each household can change its price with probability $1-\alpha$. With probability α it must keep its price unchanged. 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 call these households forward-looking. The remaining households, of measure ω, instead use a simple backward-looking rule-of-thumb when setting their prices.[10]

It follows from our assumptions that all forward-looking households who 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

$$P_t = \left[\alpha P_{t-1}^{1-\theta} + (1-\alpha)(1-\omega)p_t^{f1-\theta} + (1-\alpha)\omega p_t^{b1-\theta}\right]^{\frac{1}{1-\theta}}. \tag{17}$$

Let's 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 α; in period $t+2$ with probability α^2; and so on. It is therefore chosen to solve

$$\max_p \sum_{k=0}^{\infty} \alpha^k \left\{\Lambda_t E_t[R_{t,t+k} p y_{t+k}(p)] - \beta^k E_t[v(y_{t+k}(p))]\right\},$$

where $y_{t+k}(z)$ denotes the demand for the good at date $t+k$ as a function of its price. The marginal utility of income, Λ_t, can be treated as a constant in this calculation since the revenues from the sale of an individual good z make only an infinitesimal contribution to the representative household's total intertemporal budget constraint. Solving this optimization problem we get that

[10] 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.

the optimal price p_t^f chosen by the forward-looking households satisfies the first-order condition [11]

$$\sum_{k=0}^{\infty} \alpha^k E_t \left\{ R_{t,t+k} Y_{t+k} (\frac{p_t^f}{P_{t+k}})^{-\theta} [p_t^f - \mu S_{t+k,t}] \right\} = 0, \tag{18}$$

where μ is the households optimal markup in the case of flexible prices, which we introduced in equation (6), and

$$S_{T,t} = \frac{v_y(Y_T(p_t^f/P_T)^{-\theta}; \xi_T)}{u_C(Y_T; \xi_T)} P_T \tag{19}$$

denotes the nominal marginal cost of production at time T of a good the price of which was set at time t.

As in Gali and Gertler (1999) we assume that the backward-looking firms set their prices according to the following rule of thumb:

$$p_t^b = p_{t-1}^* \Pi_{t-1}, \tag{20}$$

where p_{t-1}^* denote an index of the prices set at date $t-1$, given by

$$\log p_{t-1}^* = (1-\omega) \log p_{t-1}^f + \omega \log p_{t-1}^b. \tag{21}$$

and $\Pi_{t-1} = P_{t-1}/P_{t-2}$. In other words, 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 rate of change of the price level they last observed, Π_{t-1}. This rule is admittedly ad hoc but it has some appealing features. As is discussed in Gali and Gertler (1999), as long as inflation is stationary there are no persistent deviations between this rule and optimal behavior and to the extent that the percent difference between the forward- and backward-looking prices is small, the loss to the household from setting its price according to this rule-of-thumb are of second order.

We have now stated the complete set of conditions which must be satisfied by a rational expectations equilibrium. Equations (15), (16), and (12) (with equality) constitute the aggregate demand block of our model; while equations (17)-(20) constitute its supply block. These two groups of equations, along with a specification of monetary policy, comprise a complete general equilibrium model.

[11] Obtaining to this expression involves some manipulation of equations. These manipulations are reported in appendix B.

2.4 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. In fact, any bounded solution is of this form, since the amplitude of the exogenous disturbances may always be reduced sufficiently to make the endogenous variables fluctuate in a small enough interval.

Let us start by log-linearizing the demand block. Equation (16) then becomes

$$(\hat{y}_t - g_t) = E_t(\hat{y}_{t+1} - g_{t+1}) - \sigma(\hat{i}_t - E_t\pi_{t+1}), \tag{22}$$

where

$$\hat{y}_t = \log(\tfrac{Y_t}{Y}); \quad \hat{i}_t = \log(\tfrac{1+i_t}{1+\bar{i}}); \quad \pi_t = \log(\tfrac{\Pi_t}{\bar{\Pi}}),$$

$\bar{\Pi} = 1$ and σ and g_t are given by

$$\sigma = -\tfrac{u_C}{u_{CC}Y} > 0; \quad g_t = \sigma\tfrac{u_{C\xi}}{u_C}\xi_t,$$

and all partial derivatives are evaluated at the steady state. We will find it convenient to express the equations of the model in terms of deviations from a time varying "natural rate" of output, y_t^n, rather than in terms of deviations from its trend path. We define such a variable, $x_t = \hat{y}_t - y_t^n$, and rewrite equation (22) as

$$x_t = E_t x_{t+1} - \sigma[(\hat{i}_t - E_t\pi_{t+1}) - r_t^n], \tag{23}$$

$$r_t^n = \sigma^{-1}E_t[(y_{t+1}^n - y_t^n) - (g_{t+1} - g_t)],$$

where r_t^n denotes the "natural real rate of interest".

Log-linearization of the supply block, i.e. equations (21)-(20), gives

$$\hat{p}_t^f = (1-\alpha\beta)\sum_{k=0}^{\infty}(\alpha\beta)^k E_t\left[\hat{s}_{t+k,t} + \sum_{s=t+1}^{t+k}\pi_s\right], \tag{24}$$

$$\hat{s}_{T,t} = (\psi^{-1} + \sigma^{-1})\hat{y}_T - \frac{\theta}{w}(\hat{p}_t^f - \sum_{s=t+1}^{T}\pi_s) + (\frac{v_{y\xi}}{v_y} - \frac{u_{C\xi}}{u_C})\xi_t, \tag{25}$$

$$\pi_t = \frac{1-\alpha}{\alpha}\hat{p}_t^*, \tag{26}$$

$$\hat{p}_t^* = (1-\omega)\hat{p}_t^f + \omega\hat{p}_t^b, \tag{27}$$

$$\hat{p}_t^b = \hat{p}_{t-1}^* - \pi_t + \pi_{t-1}, \tag{28}$$

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 stationary values of one, and $\hat{s}_{T,t}$ denotes percent deviations of $S_{T,t}/P_T$ from its stationary value of $1/\mu$. The parameter ψ is defined as

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

By manipulating these five equations we are now able to derive an aggregate supply curve for our model.We start by plugging equation (85) from appendix A into equation (25):

$$\hat{s}_{T,t} = (\psi^{-1} + \sigma^{-1})(\hat{y}_T - y_t^n) - \frac{\theta}{w}(\hat{p}_t^f - \sum_{s=t+1}^{T} \pi_s). \tag{29}$$

Next we plug equation (29) into equation (24):

$$\hat{p}_t^f = \frac{\tilde{\kappa}\alpha}{1-\alpha}\sum_{k=0}^{\infty}(\alpha\beta)^k E_t x_{t+k} + \sum_{k=1}^{\infty}(\alpha\beta)^k E_t \pi_{t+k},$$

where

$$\tilde{\kappa} = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha}\frac{\psi+\sigma}{\sigma(\psi+\theta)}.$$

Since our attention is limited to bounded solutions, and since $|\alpha| < 1$, this expression can equivalently be written in quasi-differenced form as

$$\hat{p}_t^f = \alpha\beta E_t \hat{p}_{t+1}^f + \frac{\tilde{\kappa}\alpha}{1-\alpha}x_t + \alpha\beta E_t \pi_{t+1}. \tag{30}$$

Next, we combine equations (26)-(28) in order to eliminate $\hat{p}_t^b$ and $\hat{p}_t^*$. This gives

$$\left(\omega + \frac{\alpha}{1-\alpha}\right)\pi_t = (1-\omega)\hat{p}_t^f + \frac{\omega}{1-\alpha}\pi_{t-1}. \tag{31}$$

Finally, we combine equations (30) and (31), eliminating $\hat{p}_t^f$, and get

$$\pi_t = \chi^f \beta E_t \pi_{t+1} + \chi^b \pi_{t-1} + \kappa x_t, \tag{32}$$

where

$$\chi^f = \frac{\alpha}{\omega(1-\alpha+\alpha\beta)+\alpha}, \quad \chi^b = \frac{\omega}{\omega(1-\alpha+\alpha\beta)+\alpha}, \quad \kappa = \frac{\alpha(1-\omega)\tilde{\kappa}}{\omega(1-\alpha+\alpha\beta)+\alpha}.$$

This equation is valid for $0 \leq \omega < 1$. For $\omega = 1$ the derivation is incorrect, since it would involve dividing by zero. Notice that when $\omega \to 0$ equation (32) takes on the purely forward-looking

New Keynesian form, presented e.g. in Woodford (forthcoming). However, taking the limit as $\omega \to 1$, equation (32) becomes

$$\pi_t = \frac{\alpha\beta}{1+\alpha\beta} E_t \pi_{t+1} + \frac{1}{1+\alpha\beta} \pi_{t-1}. \tag{33}$$

Surprisingly, equation (32) is not reduced to the form of the Old Keynesian Phillips curve in this limit. Instead, the weight on the forward-looking term goes to $\alpha\beta/(1+\alpha\beta)$ as $\omega \to 1$. This may seem to imply that our AS curve has a non-trivial forward-looking component in this limit. However, this is an illusion. Equation (33) has a unique bounded solution in which

$$\pi_t = \pi_{t-1},$$

which is also exactly the specification of the AS curve we get for $\omega = 1$. In order to see more clearly how the forward-looking part of our AS curve falls continuously to zero when $\omega \to 1$ it is instructive to solve equation (32) for π_t. The characteristic equation of equation (32) is

$$\beta\mu^2 - \left(1 + (\beta - 1)\omega + \frac{\omega}{\alpha}\right)\mu + \frac{\omega}{\alpha} = 0.$$

Since the two roots of this characteristic equation satisfy

$$0 < \mu_1 < 1 < \mu_2,$$

equation (32) has a unique bounded solution.[12] This solution is

$$\pi_t = \frac{\omega}{\mu_2 \alpha\beta} \pi_{t-1} + \frac{(1-\omega)\tilde{\kappa}}{\mu_2 \alpha\beta} \sum_{j=0}^{\infty} \frac{1}{\mu_2^j} E_t x_{t+j}. \tag{34}$$

It is clear from equation (34) that the forward looking part of π_t goes to zero when $\omega \to 1$.

As we can see from equation (32), the aggregate supply equation of our model does not include an additive cost push shock as for instance the AS curve discussed in Clarida et al. (1999). It is possible to derive the existence of such a shock term within the framework of the model presented here by, for instance, assuming time variations in the distortions resulting from the monopoly power of the households, or by introducing time varying tax distortions. We will however not pursue this extension of our model and instead simple assume the existence of such a cost push shock term. Given this assumption our AS curve takes the following form:

$$\pi_t = \chi^f \beta E_t \pi_{t+1} + \chi^b \pi_{t-1} + \kappa x_t + \eta_t, \tag{35}$$

[12] See section 3.1 below for a more complete presentation of this solution method.

2.5 The Central Bank's Loss Function

We now turn to the derivation of a loss function for the central bank. Following Rotemberg and Woodford (1998, 1999) and Woodford (1999b) 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

$$W = E\left\{\sum_{t=0}^{\infty} \beta^t U_t\right\}, \tag{36}$$

where

$$U_t = u(Y_t;\xi_t) - \int_0^1 v(y_t(z);\xi_t)dz. \tag{37}$$

Calculating such an approximation based on our linear approximation of the structural equation is only correct if it is done around steady state values of a particular equilibrium, more specifically, an equilibrium with zero average inflation and an efficient level of output. As is shown in Woodford (1999b), at this particular point, changes to our welfare measure resulting from changes in average levels of output under alternative policy rules are at most of second order. The neglected second order terms of the structural equations therefore only contribute terms of third order to our welfare measure. In order to have a steady state with efficient output we must assume that the government subsidizes output by a small amount which counteracts the distortions caused by each household's monopoly power.

It is shown in Woodford (1999b) that under these assumptions a second order Taylor series approximation to equation (37) is

$$\begin{aligned} U_t = -\frac{\bar{Y}u_C}{2}\Big\{&(\sigma^{-1}+w^{-1})(\hat{y}_t - y_t^n)^2 \\ &+\theta(1+w^{-1}\theta)\mathrm{var}_z(\log p_t(z))\Big\} + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3), \end{aligned} \tag{38}$$

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 output gap, $\hat{y}_t - y_t^n$, and the degree of price dispersion in the economy, $\mathrm{var}_z(\log p_t(z))$. The degree of price dispersion can now be derived from our assumptions about price setting behavior.

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 α 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

$$\bar{P}_t = E_z \log p_t(z) \quad \text{and} \quad \Delta_t = \mathrm{var}_z(\log p_t(z)),$$

and observe that

$$\bar{P}_t - \bar{P}_{t-1} = E_z[\log p_t(z) - \bar{P}_{t-1}]. \tag{39}$$

Using the recursive characterization of the distribution of prices we can replace the right hand side of equation (39) with

$$\alpha E_z[\log p_{t-1}(z) - \bar{P}_{t-1}] + (1-\alpha)(1-\omega)(\log p_t^f - \bar{P}_{t-1}) + (1-\alpha)\omega(\log p_t^b - \bar{P}_{t-1}). \tag{40}$$

Noticing that the first term in equation (40) is equal to zero we get

$$\begin{aligned} \bar{P}_t - \bar{P}_{t-1} &= (1-\alpha)(1-\omega)(\log p_t^f - \bar{P}_{t-1}) + (1-\alpha)\omega(\log p_t^b - \bar{P}_{t-1}) \\ &= (1-\alpha)(\log p_t^* - \bar{P}_{t-1}). \end{aligned} \tag{41}$$

Similarly, we may derive an expression for Δ_t:

$$\begin{aligned} \Delta_t &= \mathrm{var}_z[\log p_t(z) - \bar{P}_{t-1}] \\ &= E_z\{[\log p_t(z) - \bar{P}_{t-1}]^2\} - (E_z \log p_t(z) - \bar{P}_{t-1})^2. \end{aligned} \tag{42}$$

Again, using the recursive characterization of the distribution of prices we see that the first term on the right hand side of equation (42) can be rewritten as

$$\begin{aligned} &E_z\{[\log p_t(z) - \bar{P}_{t-1}]^2\} = \\ &\quad \alpha E_z\{[\log p_{t-1}(z) - \bar{P}_{t-1}]^2\} + (1-\alpha)(1-\omega)(\log p_t^f - \bar{P}_{t-1})^2 \\ &\quad +(1-\alpha)\omega(\log p_t^b - \bar{P}_{t-1})^2. \end{aligned} \tag{43}$$

Using

$$\begin{aligned} \log p_t^* &= (1-\omega)\log p_t^f + \omega \log p_t^b, \\ \log p_t^b &= \log p_{t-1}^* + \pi_{t-1}, \\ \bar{P}_t &= \log P_t + \mathcal{O}(\|\xi\|^2), \end{aligned} \tag{44}$$

we can further develop the last two terms on the right hand side of equation (43):

$$\begin{aligned} \log p_t^b - \bar{P}_{t-1} &= \log p_{t-1}^* + \pi_{t-1} - \bar{P}_{t-1} \\ &= \log p_{t-1}^* - \bar{P}_{t-2} - (\bar{P}_{t-1} - \bar{P}_{t-2}) + \pi_{t-1} \\ &= \log p_{t-1}^* - \bar{P}_{t-2} - \pi_{t-1} + \pi_{t-1} + \mathcal{O}(\|\xi\|^2) \\ &= \log p_{t-1}^* - \bar{P}_{t-2} + \mathcal{O}(\|\xi\|^2), \end{aligned} \tag{45}$$

$$\begin{aligned}\log p_t^f - \bar{P}_{t-1} &= \frac{1}{1-\omega}\log p_t^* - \frac{\omega}{1-\omega}(\log p_{t-1}^* + \pi_{t-1}) - \bar{P}_{t-1} \\ &= \frac{1}{1-\omega}(\log p_t^* - \bar{P}_{t-1}) - \frac{\omega}{1-\omega}(\log p_{t-1}^* + \pi_{t-1} - \bar{P}_{t-1}) \\ &= \frac{1}{1-\omega}(\log p_t^* - \bar{P}_{t-1}) - \frac{\omega}{1-\omega}(\log p_{t-1}^* - \bar{P}_{t-2}) + \mathcal{O}(\|\xi\|^2). \qquad (46)\end{aligned}$$

Finally, by combining equations (41)-(46) we obtain

$$\Delta_t = \alpha\Delta_{t-1} + \frac{\alpha}{(1-\alpha)}\pi_t^2 + \frac{\omega}{(1-\alpha)(1-\omega)}\Delta\pi_t^2 + \mathcal{O}(\|\xi\|^3), \qquad (47)$$

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 equation (47) forward, starting with an initial degree of price dispersion, Δ_{-1}, in the period before the first period we get

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

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

$$\sum_{t=0}^{\infty}\beta^t\Delta_t = \frac{1}{1-\alpha\beta}\sum_{t=0}^{\infty}\beta^t\left(\frac{\alpha}{(1-\alpha)}\pi_t^2 + \frac{\omega}{(1-\alpha)(1-\omega)}\Delta\pi_t^2\right) + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3). \qquad (48)$$

Here we have used the fact that Δ_{-1} is independent of policy chosen to apply in periods $t \geq 0$. Substituting this into equation (38) we get

$$\sum_{t=0}^{\infty}\beta^t U_t = -\Omega\sum_{t=0}^{\infty}\beta^t L_t + \text{t.i.p.} + \mathcal{O}(\|\xi\|^3), \qquad (49)$$

where Ω is a constant and L_t denotes the central bank's loss function, given by

$$L_t = \pi_t^2 + \lambda_y(\hat{y}_t - y_t^n)^2 + \lambda_\Delta\Delta\pi_t^2. \qquad (50)$$

The parameters λ_y, and λ_Δ are given by

$$\lambda_y = \tfrac{\tilde{\kappa}}{\theta}, \quad \lambda_\Delta = \tfrac{\omega}{1-\omega}\tfrac{1}{\alpha}.$$

Notice that when $\omega = 0$ this loss function simplifies to

$$L_t = \pi_t^2 + \lambda_y(\hat{y}_t - y_t^n)^2,$$

which is exactly the form reported in Woodford (1999b) for the purely forward-looking New Keynesian model. The derivation of equation (50) is correct for $0 \leq \omega < 1$, but breaks down for $\omega = 1$ since in that case it would involve dividing by zero. As $\omega \to 1$ λ_Δ becomes unbounded. However, we can see from equation (32) that changes in the price level go to zero in this limit.

2.6 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,

$$x_t = E_t x_{t+1} - \sigma[(\hat{i}_t - E_t \pi_{t+1}) - r_t^n], \tag{51}$$

$$\pi_t = \chi^f \beta E_t \pi_{t+1} + \chi^b \pi_{t-1} + \kappa x_t + \eta_t, \tag{52}$$

and a welfare criterion for the central bank, given by

$$\min E \left\{ \sum_{t=0}^{\infty} \beta^t L_t \right\}, \tag{53}$$

$$L_t = \pi_t^2 + \lambda_y (\hat{y}_t - y_t^n)^2 + \lambda_\Delta \Delta \pi_t^2. \tag{54}$$

In the following analysis we will assume specific values for all free parameters in the model except ω, which we will vary from zero to one (excluding 1). In this way we nest the purely forward-looking New Keynesian model, as well as other models such as the model of Fuhrer and Moore (1995), as special cases 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, 1998) report estimates for a purely forward-looking New Keynesian model on quarterly data using structural VAR methodology. 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, 1998) estimate that $\sigma = 6.4$, $\tilde{\kappa} = 0.024$, and $\theta = 7.88$.

Gali and Gertler (1999) estimate, for quarterly data using several variations of a non-linear instrumental variables (GMM) estimator, a Phillips curve which is almost exactly the same as the one we derive here. They report estimates of α between 0.803 and 0.866, β between 0.885 and 0.957, κ between 0.015 and 0.037, and ω 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: $\beta = 0.99$, $\alpha = 0.8$, $\sigma = 5$, $\tilde{\kappa} = 0.05$ and $\theta = 5$. Table 1 reports the values of χ^f, χ^b, κ, λ_y, and λ_Δ which result from these assumptions for a range of different values of ω. Furthermore, we will assume that η_t follows an AR(1) process

$$\eta_t = \rho \eta_{t-1} + \nu_t, \tag{55}$$

where ν_t is an i.i.d. random variable with bounded support. We will study two cases. One in which $\rho = 0$ and one in which $\rho = 0.35$.

Notice that our model corresponds closely with that of Fuhrer and Moore (1995) when $\omega = 0.8$. Fuhrer and Moore's original claim was that their specification of $\chi_1 = \chi_2 = 0.5$ 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 AS curve. Fuhrer (1997) (on U.S. data) and Blake and Westaway (1996) (on U.K. data) conclude that χ_1 close to 0.2 fit their data best. Gali and Gertler (1999), using a measure of marginal costs in their AS curve instead of a measure of the output gap, however estimate χ_1 to be in the vicinity of 0.8. It is therefore evident that there is considerable uncertainty about the relative importance of the forward-looking and the backward-looking terms in the AS curve.

Our theoretical results lead us to conclude that formulations as predominantly backward-looking as those reported in Fuhrer (1997) and in Blake and Westaway (1996) are unlikely to correctly specify the structure of aggregate supply. In order to theoretically derive an AS curve that is as backward-looking as their results suggest one would have to assume even more simplistic behavior on behalf of the backward-looking households than that which we assume. There is however still a range of possible relative values that our χ parameters can take. In the next few sections we intend to analyze optimal policy for a few different cases in this range.

3 Optimal Responses to Supply Shocks

The first type of policy analysis that we will take up is the optimal response of the endogenous variables, π_t, x_t, and $\hat{i}_t$, to η_t, the cost push disturbance term in the AS 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 equations (53) and (54), in reaction to supply shocks. Questions regarding the design of optimal policy rules will be examined in the next section.

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.[13] 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 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. Barro and Gordon (1983) 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 Barro and Gordon, 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 1970's. More recently, the importance of this effect has been questioned as most OECD countries have had considerable success (in some cases too much 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

[13] See Walsh (1998, ch. 8) for a recent survey of this literature.

> 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 (1999c, d) and Clarida et al. (1999) show that 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. Woodford (1999c, d) 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.[14] 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.

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[15]

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left\{ L_t + 2\phi_t(\chi^f \beta \pi_{t+1} - \pi_t + \chi^b \pi_{t-1} + \kappa x_t) \right\} \right\}. \tag{56}$$

Notice that we have ignored equation (51). 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

[14] If the central bank doesn't carry through on earlier commitments which the private sector thought it would carry through on, private sector expectations will be systematically wrong. Something that the assumption of rational expectations rules out in equilibrium.

[15] Here we have multiplied the Lagrange multipliers by a factor of two in order to eliminate the factor two that results from differentiating the squared terms in the loss function. Notice also that the AS equation is written as a constraint without expectations since the E_0 at the front makes them unnecessary.

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 $\hat{i}_t$ it the control variable of the central bank it can be chosen freely to satisfy equation (51), given optimal paths for π_t and x_t. Equation (51) can therefore be ignored.

Differentiating equation (56) with respect to π_t and x_t we get two first order conditions which the optimal plan must satisfy,

$$(1 + \lambda_\Delta + \beta\lambda_\Delta)\pi_t - \lambda_\Delta\pi_{t-1} - \beta\lambda_\Delta E_t\pi_{t+1} + \chi^f\phi_{t-1} - \phi_t + \beta\chi^b E_t\phi_{t+1} = 0 \tag{57}$$

$$\lambda_y x_t + \kappa\phi_t = 0. \tag{58}$$

These two conditions must hold for each date $t \geq 1$, and the same conditions with

$$\phi_{-1} = 0 \tag{59}$$

must also hold for date $t = 0$. Equations (57)-(59) together with equation (52) and an initial condition π_{-1} can now be solved for the optimal paths of π_t and x_t.

Notice the ϕ_{t-1} term in equation (57). This term represents the shadow price of relaxing the time $t-1$ AS curve at time t. However, at time t the AS curve for time $t-1$ is already determined. In a normal (engineering) optimal control problem, ignoring it would therefore not adversely affect the optimization at time t. This is not the case in a forward-looking system with rational expectations. In such a system ignoring this past constraint will change the way expectations are formed, and ultimately lead to a suboptimal outcome. To attain optimality the central bank must commit itself to take proper account of this term when optimizing, that is, it must commit itself not to ignore the past. Thus, the optimal policy is history dependent.

We can now use equation (58) to substitute the Lagrange multiplier, ϕ_t out of our system and then write the remaining two equations in matrix form

$$A \begin{bmatrix} E_t z_{1t+1} \\ z_{2t} \end{bmatrix} = B \begin{bmatrix} z_{1t} \\ z_{2t-1} \end{bmatrix} + \xi_t, \tag{60}$$

where $z_{1t} = [\pi_t, x_t]'$ is a vector of forward-looking variables, $z_{2t-1} = [\pi_{t-1}, x_{t-1}]'$ is a vector of predetermined variables, A and B are 4×4 matrices whose values are given in appendix C, and $\xi_t = [\eta_t, 0, 0, 0]'$ is a vector of shocks.

Next, we premultiply equation (60) by A^{-1} and take expectations at time t

$$E_t \begin{bmatrix} z_{1t+1} \\ z_{2t} \end{bmatrix} = C \begin{bmatrix} z_{1t} \\ z_{2t-1} \end{bmatrix} + A^{-1}\xi_t. \tag{61}$$

Blanchard and Kahn (1980) show that this type of system has a unique bounded solution, if and only if, the matrix C has exactly two eigenvalues with modulus greater than one. If we assume that this is the case and that the matrix C has four distinct eigenvalues, a solution to the system can be computed using the following algorithm.[16] If μ_1 and μ_2 are the two eigenvalues of C that have modulus larger that 1 and v_1' and v_2' are the left eigenvectors associated with these eigenvalues, we premultiply equation (61) with the matrix $[v_1', v_2']'$. Using the fact that $v_1'C = \mu_1 v_1'$ we then get

$$E_t \Upsilon_{t+1} = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} \Upsilon_t + \delta_t,$$

where Υ_t and δ_t are a 2×1 vectors formed by multiplying the matrix of eigenvectors with the state vector and the vector of shocks, respectively. Notice that we have now uncoupled our system. It can now be separated in the following way

$$E_t \Upsilon_{it+1} = \mu_i \Upsilon_{it} + \delta_{it}.$$

Solving these equations forward, making use of the fact that μ_i is greater than one in modulus and that we are seeking bounded solutions we get

$$\Upsilon_{it} = -\sum_{j=0}^{\infty} \mu_i^{-(j+1)} E_t \delta_{it+j}. \tag{62}$$

Remembering that Υ_t is a linear combination of $[z_{1t}, z_{2t-1}]$ we can rewrite these equations as

$$Q z_{1t} = K z_{2t-1} - \sum_{j=0}^{\infty} \begin{bmatrix} 1/\mu_1 & 0 \\ 0 & 1/\mu_2 \end{bmatrix}^{j+1} E_t \delta_{t+j}.$$

Finally, premultiplying by Q^{-1} and using (55) we get

$$z_{1t} = Q^{-1} K z_{2t-1} - Q^{-1} \begin{bmatrix} \frac{1}{\mu_1 - \rho} & 0 \\ 0 & \frac{1}{\mu_2 - \rho} \end{bmatrix} \delta_t. \tag{63}$$

Here we have solved for the forward-looking variables in terms of the predetermined variables, i.e. we have solved for the path of inflation and output.

[16] When solving the system for particular values of ω one needs to check that these assumptions hold.

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 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.

Following Söderlind (1999) we use a dynamic programming approach to solve for optimal policy in this case. Again, we can ignore equation (51) since $\{\hat{i}_t\}$ can be chosen freely, given paths for π_t and x_t. We write equation (52) in the following matrix form[17]

$$\begin{bmatrix} z_{1t+1} \\ E_t z_{2t+1} \end{bmatrix} = A \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} + Bx_t + \begin{bmatrix} \eta'_{t+1} \\ 0 \end{bmatrix}, \tag{64}$$

where $z_{1t} = [\pi_{t-1}, \eta_t]'$, $z_{2t} = \pi_t$,

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \rho & 0 \\ \frac{-\chi^b}{\beta\chi^f} & \frac{-1}{\beta\chi^f} & \frac{1}{\beta\chi^f} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \frac{-\kappa}{\beta\chi^f} \end{bmatrix}, \quad \eta'_{t+1} = \begin{bmatrix} 0 \\ \nu_{t+1} \end{bmatrix}.$$

The problem can now be written in the form of a Bellman equation[18]

$$z'_{1t}Vz_{1t} + v = \min_{x_t}[z'_tQz_t + x'_tRx_t + \beta E_t(z'_{1t+1}Vz_{1t+1} + v)]$$

$$\text{s.t. } E_t z_{2t+1} = CE_t z_{1t+1}, (64), \text{ and } z_{1t} \text{ given.} \tag{65}$$

where $z'_{1t}Vz_{1t} + v$ is the value function which minimizes the expression on the right hand side subject to the given constraints, $z_t = [z_{1t}, z_{2t}]'$, $R = \lambda_y$, and

$$Q = \begin{bmatrix} \lambda_\Delta & 0 & \lambda_\Delta \\ 0 & 0 & 0 \\ \lambda_\Delta & 0 & 1+\lambda_\Delta \end{bmatrix}.$$

Following Söderlind (1999), we can now partition the matrices A, B and Q conformably with z_{1t} and z_{2t} and rewrite equation (65) as

$$z'_{1t}Vz_{1t} + v = \min_{x_t}[z'_{1t}Q^*z_{1t} + 2z'_{1t}U^*x_t + x'_tR^*x_t + \beta E_t(z'_{1t+1}Vz_{1t+1} + v)]$$

[17] Notice that the notation in this section is different from the notation in the last section. In this section we are following the notation of Söderlind (1999).

[18] Notice that here we are treating x_t as the control variable. The true control variable is still the $\hat{i}_t$. However, in our model the central bank chooses x_t indirectly through equation (51).

$$\text{s.t. } z_{1t+1} = A^* z_{1t} + B^* x_t + \epsilon_{t+1}, \text{ and } z_{1t} \text{ given.} \tag{66}$$

where the starred matrices are defined as

$$\begin{aligned} D &= (A_{22} - CA_{12})^{-1}(CA_{11} - A_{21}) \\ G &= (A_{22} - CA_{12})^{-1}(CB_1 - B_2) \\ A^* &= A_{11} + A_{12}D \\ B^* &= B_1 + A_{12}G \\ Q^* &= Q_{11} + Q_{12}D + D'Q_{21} + D'Q_{22}D \\ U^* &= Q_{12}G + D'Q_{22}G \\ R^* &= R + G'Q_{22}G \end{aligned} \tag{67}$$

The first order conditions of (66) are

$$x_t = -F_1 z_{1t},$$

$$F_1 = (R^* + \beta B^{*'} V B^*)^{-1}(U^* + \beta B^{*'} V A^*). \tag{68}$$

Combining equations (66) and (68) we get

$$z_{2t} = C z_{1t}$$

$$C = D - GF_1 \tag{69}$$

$$V = Q^* - U^* F_1 - F_1' U^{*'} + \beta(A^* - B^* F_1)' V (A^* - B^* F_1). \tag{70}$$

This algorithm (equations (67)-(70)) is now iterated until it converges. The initial value for V should be a symmetric positive definite matrix (In practice 0 works quite well). The initial condition for C can be chosen freely. In practice zero also works quite well for C. The solution is then given by

$$z_{1t+1} = (A_{11} + A_{12}C - B_1 F_1) z_{1t} + \eta_{t+1}', \tag{71}$$

$$x_t = -F_1 z_{1t}. \tag{72}$$

Again we have solved for the path of inflation and output given our structural model and initial conditions.

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 figure 1 through figure 27 we report such impulse response functions. More precisely, we report

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

for $\psi = \pi_{t+j}, x_{t+j}, p_{t+j}$, where p_{t+j} is the price level in period $t+j$. These impulse response functions are reported for four different values of ω. The values of ω which we choose to report are: (1) $\omega = 0.01$, which we refer to as the New Keynesian case, since in this case our AS curve corresponds closely with the purely forward-looking AS curve analyzed in Woodford (1999c,d); (2) $\omega = 0.2$, which we refer to as the Gali and Gertler case, since for this value of ω our AS curve corresponds closely with the estimated model presented in Gali and Gertler (1999); (3) $\omega = 0.5$, which we refer to as the Midway case; and (4) $\omega = 0.8$, which we refer to as the Fuhrer-Moore case, since for this value of ω our AS curve corresponds closely to the model proposed in Fuhrer and Moore (1995). In addition to this we report impulse response functions for a purely backward looking model which corresponds to choosing $\chi^f = 0$, $\chi^b = 1$, and $\kappa = 0.05$.

For each of these cases, except the purely backward-looking model, we report results for commitment and discretionary optimization, under two different assumptions about the loss function the central bank chooses to minimize, and two different assumption about the nature of the shock process. Firstly, 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 $\lambda_\Delta = 0$ in equation (54). We refer to this loss function as the traditional loss function. The two types of shocks which we report impulse response functions (IRF's) for are: η_t i.i.d., and η_t AR(1) with $\rho = 0.35$. Figure 1 through figure 12 present the i.i.d. case while figure 13 through figure 24 present the AR(1) case.

Notice that in the purely backward-looking case, presented in figures 25-27, [19] optimizing with commitment and optimizing under discretion leads to identical IRF's. 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 figures 1-24. It is evident from these

[19] These figures report the AR(1) case for the traditional ($\lambda_\Delta = 0$) loss function.

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.

Let us now turn to figures 1-24. Perhaps the most interesting feature of these IRF's is that in the commitment case it is always optimal for the central bank to induce a period of deflation following an inflationary cost push shock. In other words it is optimal in these cases to have the 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 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 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 suboptimality of discretionary optimization.

Consider next how our two different assumptions about the loss function effect the optimal responses. In figures 1-3 and 13-15 we see that in the New Keynesian case the IRF's for the two loss functions are virtually identical as we would expect. As ω rises the IRF's for the two loss functions diverge. For the other three choices of ω, the optimal IRF's for the theoretical loss function involve less initial inflation and more initial loss of output. The reason for this is that the theoretical loss function contains a third term penalizing *changes* in inflation as well as inflation itself. The difference between the two cases increases as ω increases, since λ_Δ is an increasing function of ω. The difference is greatest in the discretion case. Actually, for a central bank seeking to minimize the theoretical loss function under discretion the response of output becomes extremely explosive for large ω. The commitment case is better behaved.

As was mentioned earlier there is little consensus within the empirical literature on the relative importance of forward- versus backward-looking terms in the AS 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. 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 15% of the size of the shock, while the short term rise in the price level is between 70% and 80% of the size of the shock. In the Fuhrer-Moore case 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.

The second substantial difference between these two cases is the number of quarters that is takes a transitory shock to fully play out in case of optimal policy under commitment. In the Gali and Gertler case the economy has mostly settled into its new steady state after about six quarters. However, in the Fuhrer-Moore case it takes the economy about 12 quarters to approach its new equilibrium. In the Gali and Gertler case the price level starts to fall immediately in the period after the shock. In the Fuhrer-Moore case the price level does not start to fall until 6 periods after the shock. The gradual responses to shocks that characterize the Fuhrer-Moore case seem to correspond better to actual central bank policy then the sharp overshooting which characterizes the Gali and Gertler case.

The third difference between these cases is the enormous difference in the size of the output response. In the Gali and Gertler case the output response for all the four cases analyzed is under 10 times the size of the shock while in the Fuhrer-Moore case the output response ranges from 4 times to over 30 times the size of the shock. Admittedly, the case of discretion with the theoretical loss function is very extreme but non the less this constitutes a large difference between these two cases for all four cases.

Another interesting way to compare optimal monetary policy in the cases we have been discussing is to calculate a measure of expected welfare for different cases. In order to do this one must make a few specific assumptions about ones welfare measure. The assumptions we make are discussed in detail in section 4.1 below. We use Monte Carlo methods to estimate our welfare measures. The results are reported in Table 2. We can see that in the New Keynesian case the loss associated with economic fluctuations is 30% larger in the case of discretion than in the case of commitment. We can also see that the difference in loss shrinks as ω rises from zero to 0.5 as we would expect. This is basically due to the fact that as the Phillips curve

becomes more and more backward looking the gains from commitment become less and less important. It is however important to note that the gains from commitment remain substancial even when ω is 0.5. It is therefore somewhat surprising that the difference between commitment and discretion increases again when ω rises from 0.5 to 0.8. Although we do not claim to fully understand this we suspect that this is due to the fact that $\lambda_\Delta \to \infty$ and $\kappa \to 0$ as $\omega \to 1$. The model is therefore unrealistic in this limit.

While the analysis of these type impulse response functions is interesting and highly useful for building intuition about optimal monetary policy, a more interesting issue is to compare optimal policy rules for these different cases. We now turn to this issue.

4 Interest Rate Feedback Rules

In the last section we sought to describe the optimal responses of the economy to shocks. Optimal in the sense that they attained the lowest feasible value of the central bank's loss function. We now turn to the question of how a central bank should bring about desirable responses to shocks. A number of quite distinct approaches to this problem exist within the literature on optimal monetary policy. We will only concern ourselves with one approach here, namely the design of interest rate feedback rules.

Even within the domain of monetary policy rules, many different approaches have been proposed in the literature. Perhaps the most important choice one needs to make, when designing a monetary policy rule, is which variable to target. The early literature on monetary policy rules was almost exclusively concerned with designing desirable rules for the evolution of the money supply. The landmark analysis of Freidman and Schwartz (1963) concluded that the most important cause of business cycles in the U.S. in the preceding century had been undesirable fluctuations in the stock of money. Armed with these results as well as one of the oldest and most respected theories in economics, the quantity theory of money, Friedman argued in the 1960's that monetary policy should be conducted by targeting a constant growth rate for the money supply.

Targets for the evolution of the money supply are however not the only possible type of monetary policy rule. Alternatively, the central bank could conduct monetary policy by committing itself to an interest rates feedback rule, that is, a rule which determines the target value of the interest rate as a function of other variables of the model. For much of the 20th century, the

question of optimal instrument choice has been an important problem within monetary theory. The literature on this issue has focused primarily on the choice between controlling the evolution of the money supply and controlling the evolution of interest rates.

Following the introduction of rational expectations to macroeconomics in the 1970's the possibility of multiple equilibria became an important issue in the design of monetary policy rules. Sargent and Wallace (1975) argued that monetary policy rules should target the money supply since interest rate rules necessarily result in price level indeterminacy. While important, this result, did not turn out to be true of interest rate rules in general, but only a special subset of such rules. McCullum (1981) showed that interest rate rules with sufficiently strong feedback from a "nominal anchor" are not subject to this defect. It has since become clear that feedback from nominal variables is not needed to resolve the problem of price level indeterminacy. Sufficiently strong feedback from endogenous variables is enough.

Money supply targeting has fallen in and out of favor with central bankers no less than with academic economists in the last few decades. The importance of the money supply in the conduct of monetary policy rose a great deal in the 1970's and reached a maximum (at least in the U.S) in the early 1980's under chairman Paul Volcher. The theoretical developments mentioned above were no doubt an important factor in this process. Another important factor, however, was the fact that the velocity of money (LM shocks) had been extremely stable and predictable in the preceding decades. As the analysis of Poole (1970) suggests, this stability made the money supply a relatively attractive target variable. However, in the early 1980's the stability of the velocity of money broke down. As a consequence money supply targeting has since become less popular among central bankers and more attention has been given to controlling interest rates.

Similarly, the recent theoretical literature on monetary policy rules has focused increasingly on interest rate feedback rules. Perhaps the most influential work in this area is Taylor (1993). Taylor showed that a particularly simple type of interest rate feedback rule characterizes the recent policy of the Federal Reserve remarkably well. The policy rule which Taylor proposed, and which now bears his name, is

$$i_t = i^* + \theta_\pi(\pi_t - \pi^*) + \theta_x x_t,$$

where $i^* = 2$, $\pi^* = 2$, $\theta_\pi = 1.5$ and $\theta_x = 0.5$. The success of monetary policy in the period when this Taylor rule fits the data well led Taylor to propose that the Federal Reserve use his rule as a reference tool for the future conduct of monetary policy. Following Taylor's work, a large

literature has developed, seeking to analyze the performance of Taylor rules and generalizations of Taylor rules when applied to various types of structural macromodels.[20] The remainder of this section is devoted to the analysis of optimal Taylor Rules for the model presented in section 2.

4.1 Optimal Taylor rules

In order to be able to ignore shocks to the IS curve, and focus on finding the Taylor rules which lead to the best responses of the economy to cost push shocks, we imagine that the central bank conducts monetary policy by setting the interest rate according to the following generalization of the Taylor rule:

$$\hat{i}_t = \theta_\pi \pi_t + \theta_x x_t - r_t^n \tag{73}$$

That is, we assume that r_t^n is perfectly observable at time t and that the central bank varies the interest rate to perfectly isolate the economy from its effect. In this section we again assume that η_t is given by equation (55). Our complete set of structural equations then consists of equations (51), (52), (73), and (55).

Ignoring equation (55) for the moment, we can write our system in matrix form as

$$A \begin{bmatrix} E_t\pi_{t+1} \\ E_t x_{t+1} \\ \pi_t \end{bmatrix} = B \begin{bmatrix} \pi_t \\ x_t \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} -\eta_t \\ 0 \\ 0 \end{bmatrix}, \tag{74}$$

where

$$A = \begin{bmatrix} \beta\chi^f & 0 & 0 \\ \sigma & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 1 & -\kappa & -\chi^b \\ \sigma\theta_\pi & 1+\sigma\theta_x & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

This system can be solved using the same methods that we used in section 3.1. It has a unique bounded solution if and only if two eigenvalues of the matrix $A^{-1}B$ have modulus larger than one. If this condition holds the evolution of the endogenous variables, π_t and x_t, can be written as

$$\pi_t = a_\pi \pi_{t-1} + b_\pi \eta_t, \tag{75}$$

$$x_t = a_x \pi_{t-1} + b_x \eta_t. \tag{76}$$

[20] A recent collection of papers in this vein is Taylor (1999).

We seek the optimal Taylor rule. However, our choice will depend on the welfare criterion we choose to use. In section 3 we sought to minimize the average expected loss of the households, derived in section 2. This welfare measure is given by

$$W = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t L_t \right\}, \tag{77}$$

where L_t is defined by equation (54). Using equations (75)-(76), we can derive an expression for W in terms of model parameters and the value of π_{-1}. Consequently, if we use equation (77) as our welfare measure, our choice of optimal Taylor rule will depend on π_{-1}. We do not find this appealing. We would like to evaluate a Taylor rule which is optimal from a "timeless perspective," that is, an optimal Taylor rule which does not depend on the state of the economy at time 0.

A more appealing welfare criterion is $E(W)$ with the additional assumption that $\pi_{-1} = 0$. Since π_t is an endogenous variable, its unconditional probability distribution depends on the chosen policy rule. However, the value of π_{-1} does not depend on the policy rule chosen for period 0 onward. It is therefore not reasonable to penalize rules that make π_t more variable by also assuming a greater degree of variability in the initial condition, π_{-1}.

The derivation of $E(W)$ is quite tedious. It is presented in appendix D. Since a_π, b_π, a_x and a_x are functions of θ_π and θ_x, we can see from equations (109)-(113) that $E(W)$ will also be a function of these two parameters. We use numerical methods to search for the vector $[\theta_\pi, \theta_x]$ which minimizes $E(W)$. The results of this analysis are presented in Table 3. For a few cases we also present contour plots of the loss associated with different Taylor rules (Figures 28-31).

A number of interesting features emerge in our analysis. Firstly, the qualitative features of the contour plots for the range of cases we study are strikingly similar. We report contour plots for four different cases. The reason we do not report this type of plot for all the other cases we have been studying is that all these plots look so similar. A unique bounded solution exists for roughly all parameter values in the first and fourth quadrant of the plane, while our system is either indeterminant (for low values of ω) or unstable (for high values of ω) for the bulk of the second and third quadrants. Notice that, these different regions correspond only approximately with the quadrants. The main deviation being that the system does not have a unique bounded solution for any value close to (0,0). These qualitative features of our contour plots are very similar to the qualitative features of the same type of plots reported in Rotemberg and Woodford (1999) for a slightly different model.

We choose to ignore the region where both θ_π and θ_x are negative since these types of policy rules do not seem to correspond in any way to actual policy. We note however that the existence of unique bounded solutions in this region is somewhat surprising and warrants future attention.

Perhaps the most important qualitative feature of our plots is the fact that the loss associated with any particular point on these plots does not depend as much on the absolute size of the coefficients as it does on their relative size. The positive-positive region in which unique bounded solutions exist forms a large "valley". The bottom of this valley forms a straight line which is slightly different for each of our different cases. These lines are reported in Table 3. We will refer to these lines as the optimal lines. The loss associated with policy rules on the optimal line seems to be constant. So, all that matters is to be on the line not *where* one is on the line.[21]

The main difference between Taylor's original specification of the Taylor rule and the Taylor rules which turn out to be optimal in our model is the relative size of θ_x. The relative size of θ_x turns out to be substantially smaller on our optimal lines. This is true for all the cases we study. The relative size of θ_x actually becomes smaller as one makes our model more backward looking (especially if one uses the loss function derived in section 2 instead of the traditional loss function). This result is perhaps somewhat surprising, especially since the simulation results of several economists, including Laurence Ball and John Williams, have indicated that raising the size of θ_x relative to the original specification by Taylor would improve performance. In contrast our results suggest that paying almost no attention to the output gap (in other words focusing almost exclusively on reacting to inflation) would be the optimal way to conduct policy.

On the other hand, one can also see from Figures 28-31 that the valley is much steeper below the optimal line than it is above the optimal line. So, monetary policy which is a little bit too aggressive in the sense that the coefficient on the output gap has been chosen somewhat too large is not as bad as monetary policy which is a little bit too passive.

5 Conclusions

In this paper we have extended the benchmark New Keynesian macromodel 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 of the producers in the economy set their

[21] This does not exactly correspond to correctly choosing the relative size of θ_π and θ_x, as our optimal lines do not pass through the origin.

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 case 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. Overshooting the inflation target is still optimal in the primarily backward-looking case, but much less overshooting.

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 1990's are a good example of this type of behavior.

Woodford (1999c) 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 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.

However, the results reported in Sections 3 and 4 must be taken with a grain of salt since the model presented in Section 2 is extremely stylized. The most immediate defect this model is the lack of microfoundations for the cost push shock. As was noted in Section 2 it is possible to derive such a shock by adding, for instance, time variations in the market power of the producers, or time varying tax distortion. However, this does not seem to be an appealing way to add cost push supply shocks to our model, at least if we believe that such shocks are an important driving force of business cycles.

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. The appropriate role of the exchange rate in monetary policy will crucially depend on whether such a loss function turns out to contain an extra term containing the exchange rate.

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 surely 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.

A Derivations Regarding the Model's Market Structure

We follow Dixit and Stiglitz (1977) in assuming that the household derives utility from the consumption of a composite consumption index defined by

$$C = \left[\int_0^1 c(z)^{\frac{\theta}{\theta-1}} dz\right]^{\frac{\theta-1}{\theta}}. \tag{78}$$

In order to consume the different differentiated goods in an efficient manner the household must solve the following maximization problem

$$\max_{c(z)} \left(\int_0^1 c(z)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}} \quad \text{s.t.} \quad Z = \int_0^1 p(z)c(z)dz$$

We form a Lagrangian

$$L = \left(\int_0^1 c(z)^{\frac{\theta-1}{\theta}} dz\right)^{\frac{\theta}{\theta-1}} + \lambda\left[Z - \int_0^1 p(z)c(z)dz\right]$$

Differentiating this we get

$$\left(\frac{C}{c(z)}\right)^{\frac{1}{\theta}} = \lambda p(z).$$

Eliminating λ we get

$$c(z) = c(z')\left(\frac{p(z')}{p(z)}\right)^{\theta}.$$

Plugging this into our constraint and rearranging we then get a demand curve for good $c(z)$ of the form

$$c(z') = \frac{p(z')^{-\theta}Z}{\int_0^1 p(z)^{1-\theta}dz}. \tag{79}$$

Next, we define the price index, P, as the minimum expenditure, Z, needed to purchase a unit of the composite consumption good, C. The highest value of C given Z is then found by plugging equation (79) into equation (78)

$$C = \left[\int_0^1 \left(\frac{p(z)^{-\theta}Z}{\int_0^1 p(z')^{1-\theta}dz'}\right)^{\frac{\theta-1}{\theta}} dz\right]^{\frac{\theta}{\theta-1}}.$$

Using the definition of P we then get

$$1 = \left[\int_0^1 \left(\frac{p(z)^{-\theta}P}{\int_0^1 p(z')^{1-\theta}dz'}\right)^{\frac{\theta-1}{\theta}} dz\right]^{\frac{\theta}{\theta-1}},$$

which may be solved for P. The resulting expression for the price index is

$$P = \left[\int_0^1 p(z)^{1-\theta}dz\right]^{\frac{1}{1-\theta}}. \tag{80}$$

Notice that it follows from our definition of P that

$$C = \frac{Z}{P}.$$

We can now use this expression and equation (80) to simplify equation (79)

$$\begin{aligned} c(z) &= \frac{p(z')^{-\theta} Z}{\int_0^1 p(z)^{1-\theta} dz} \\ &= \frac{p(z')^{-\theta} CP}{\int_0^1 p(z)^{1-\theta} dz} \\ &= \frac{p(z')^{-\theta} C \left[\int_0^1 p(z)^{1-\theta} dz\right]^{\frac{1}{1-\theta}}}{\int_0^1 p(z)^{1-\theta} dz} \\ &= \left(\frac{p(z)}{P}\right)^{-\theta} C. \end{aligned} \tag{81}$$

This last expression is the demand expression given in the main text. We also seek to derive equation (6) from the main text. Let us define $S_t(z)$ to be nominal marginal costs of production. In the flexible price case of the model presented in this paper producers will solve the following optimization problem to determine their price

$$\max_{p(z)} y_t(z) p_t(z) - S_t(z) y_t(z).$$

Substituting in for $y_t(z)$ we get

$$\max_{p(z)} \left(\frac{p(z)}{P_t}\right)^{-\theta} Y_t p(z) - S_t(z) \left(\frac{p(z)}{P_t}\right)^{-\theta} Y_t.$$

Differentiating we get the first order condition

$$(1-\theta) Y_t P_t^{\theta} p_t(z)^{-\theta} + \theta S_t(z) Y_t P_t^{\theta} p_t(z)^{-(1+\theta)} = 0,$$

which simplifies to

$$p_t(z) = \frac{\theta}{\theta - 1} S_t(z) = \mu S_t(z). \tag{82}$$

Thus, we see that in the flexible price case prices will be set as a constant markup over marginal costs. Notice that in the flexible price case the marginal cost of production is found by differentiating equation (2) of the main text with respect with $y_t(z)$. The resulting first order condition is

$$v_y(y_t(z); \xi_t) = S_t(z) \Lambda_t.$$

Combining this with equation (13) from the main text we can see that the real marginal cost of producing good z is

$$\frac{S_t(y_t(z), Y_t, \xi_t)}{P_t} = \frac{v_y(y_t(z); \xi_t)}{u_C(C_t; \xi_t)}. \tag{83}$$

In other words, real marginal costs are a function of $y_t(z)$, Y_t, ξ_t. Combining equation (82) with the fact that in a flexible price equilibrium equal amounts of each good, $y_t(z)$, are demanded we get that in equilibrium

$$\frac{S_t(Y_t^n, Y_t^n, \xi_t)}{P_t} = \mu^{-1}, \tag{84}$$

where Y_t^n denotes the equilibrium level of output in the flexible price case. We will refer to this as the "natural level" of output.

Finally we will find it useful to notice that log-linearization of equation (83) gives us

$$(\frac{v_{y\xi}}{v_y} - \frac{u_{C\xi}}{u_C})\xi_t = -(\psi^{-1} + \sigma^{-1})y_t^n. \tag{85}$$

B Manipulations Leading to Equation (18)

Here we are interested in solving

$$\max_p \sum_{k=0}^{\infty} \alpha^k \left\{ \Lambda_t E_t[R_{t,t+k} p y_{t+k}(p)] - \beta^k E_t[v(y_{t+k}(p))] \right\}. \tag{86}$$

In doing this we will make use of the following equations of our model:

$$y_{t+k} = \left(\frac{p_t(z)}{P_t}\right)^{-\theta} Y_t, \tag{87}$$

$$u_C(Y_t, \xi_t) = P_t \Lambda_t, \tag{88}$$

$$R_{t,T}\Lambda_t = \beta^{T-t}\Lambda_T. \tag{89}$$

From equations (88) and (89) we get

$$\beta^k = \frac{\Lambda_t P_{t+k}}{u_C(Y_{t+k}, \xi_t)} R_{t,t+k}. \tag{90}$$

Plugging equation (87) into equation (86) we get

$$\max_p \sum_{k=0}^{\infty} \alpha^k \left\{ \Lambda_t E_t[R_{t,t+k} p \left(\frac{p}{P_{t+k}}\right)^{-\theta} Y_{t+k}] - \beta^k E_t[v(\left(\frac{p}{P_{t+k}}\right)^{-\theta} Y_{t+k})] \right\}. \tag{91}$$

Differentiating equation (91) and setting the result equal to zero we get the following first order conditions for optimality

$$\sum_{k=0}^{\infty} \alpha^k \left\{ \Lambda_t E_t \left[R_{t,t+k}(1-\theta) \left(\frac{p_t^f}{P_{t+k}}\right)^{-\theta} Y_{t+k} \right] \right.$$
$$\left. + \beta^k E_t \left[v_y(\left(\frac{p_t^f}{P_{t+k}}\right)^{-\theta} Y_{t+k}) \theta \frac{p_t^{f-(1+\theta)}}{P_{t+k}^{-\theta}} Y_{t+k} \right] \right\} = 0. \tag{92}$$

$$\tag{93}$$

Using equation (90) and rearranging we get

$$\sum_{k=0}^{\infty} \alpha^k E_t \left\{ R_{t,t+k} Y_{t+k} (\frac{p_t^f}{P_{t+k}})^{-\theta} [p_t^f - \mu S_{t+k,t}] \right\} = 0,$$

where $S_{T,t}$ is given by

$$S_{T,t} = \frac{v_y(Y_T(p_t^f/P_T)^{-\theta}; \xi_T)}{u_C(Y_T; \xi_T)} P_T.$$

These two equations are the same as those reported in equations (18) and (19) in the text.

C Values of Matrices A and B in Section 3.2

$$A = \begin{bmatrix} \beta\chi^f & 0 & 0 & 0 \\ \beta\lambda_\Delta & \beta\lambda_y\chi^b/\kappa & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -\kappa & -\chi^b & 0 \\ (1+\lambda_\Delta+\beta\lambda_\Delta) & \lambda_y/\kappa & -\lambda_\Delta & -\lambda_y\chi^f/\kappa \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

D Derivation of $E(W)$ for Non-Inertial Case with an AR(1) Shock Process

When monetary policy is conducted in a non-history dependent way, π_t and x_t can be expressed as time-invariant functions of the cost push shock, η_t, and lagged inflation, π_{t-1}:

$$\pi_t = a_\pi \pi_{t-1} + b_\pi \eta_t, \tag{94}$$

$$x_t = a_x \pi_{t-1} + b_x \eta_t. \tag{95}$$

Let's assume that the cost push shock is given by an AR(1) process

$$\eta_t = \rho \eta_{t-1} + \nu_t. \tag{96}$$

Solving equations (94)-(96) backwards we get

$$\pi_t = b_\pi \sum_{j=0}^{t} a_\pi^j \eta_{t-j} + a_\pi^{t+1} \pi_{-1}, \tag{97}$$

$$x_t = b_x \eta_t + b_\pi a_x \sum_{j=0}^{t-1} a_\pi^j \eta_{t-1-j} + a_x a_\pi^t \pi_{-1}, \tag{98}$$

$$\eta_t = \sum_{j=0}^{t} \rho^j \nu_{t-j} + \rho^{t+1} \eta_{-1}. \tag{99}$$

Plugging equation (99) into equations (97) and (98) we get

$$\pi_t = b_\pi \sum_{j=0}^{t} \sum_{i=0}^{j} \rho^i a_\pi^{j-i} \nu_{t-j} + b_\pi \sum_{j=1}^{t} \rho^{j+1} a_\pi^{t-j} \eta_{-1} + a_\pi^{t+1} \pi_{-1} \tag{100}$$

$$\begin{aligned} x_t = & b_x \sum_{j=0}^{t} \rho^j \nu_{t-j} + b_x \rho^{t+1} \eta_{-1} \\ & + b_\pi a_x \sum_{j=0}^{t-1} \sum_{i=0}^{j} \rho^i a_\pi^{j-i} \nu_{t-1-j} + b_\pi a_x \rho \sum_{j=0}^{t-1} \rho^j a_\pi^{t-j-1} \eta_{-1} + a_x a_\pi^t \pi_{-1} \end{aligned} \tag{101}$$

Equations (100) and (101) can be simplified down to

$$\pi_t = b_\pi \sum_{j=0}^{t} \frac{a_\pi^{j+1} - \rho^{j+1}}{a_\pi - \rho} \nu_{t-j} + b_\pi \rho \frac{a_\pi^{t+1} - \rho^{t+1}}{a_\pi - \rho} \eta_{-1} + a_\pi^{t+1} \pi_{-1}, \tag{102}$$

$$\begin{aligned} x_t = & b_x \sum_{j=0}^{t} \rho^j \nu_{t-j} + b_x \rho^{t+1} \eta_{-1} \\ & + b_\pi a_x \sum_{j=0}^{t-1} \frac{a_\pi^{j+1} - \rho^{j+1}}{a_\pi - \rho} \nu_{t-1-j} + b_\pi a_x \rho \frac{a_\pi^t - \rho^t}{a_\pi - \rho} \eta_{-1} + a_x a_\pi^t \pi_{-1}. \end{aligned} \tag{103}$$

We seek to compute

$$\begin{aligned} E(W) &= E\left\{ \sum_{t=0}^{\infty} \beta^t L_t \right\} \\ &= E\left\{ \sum_{t=0}^{\infty} \beta^t (\pi_t^2 + \lambda_y x_t^2 + \lambda_\Delta (\pi_t^2 - 2\pi_t \pi_{t-1} + \pi_{t-1}^2)) \right\} \\ &= \sum_{t=0}^{\infty} \beta^t ((1+\lambda_\Delta) E\pi_t^2 + \lambda_y E x_t^2 - 2\lambda_\Delta E\pi_t \pi_{t-1} + \lambda_\Delta E\pi_{t-1}^2), \end{aligned} \tag{104}$$

given the initial condition $\pi_{-1} = 0$. In order to compute equation (104) we therefore first need to compute expressions for $E\pi_t^2$, $E\pi_t \pi_{t-1}$, $E\pi_{t-1}^2$, and Ex_t^2. Given our initial condition and the fact that ν_t is an i.i.d. random variable we get that

$$E\pi_t^2 = b_\pi^2 \sum_{j=0}^{t} \Gamma_{j+1}^2 \mathrm{var}(\nu_t) + b_\pi^2 \rho^2 \Gamma_{t+1}^2 \mathrm{var}(\eta_t), \tag{105}$$

$$E\pi_t \pi_{t-1} = b_\pi^2 \sum_{j=0}^{t-1} \Gamma_{j+2} \Gamma_{j+1} \mathrm{var}(\nu_t) + b_\pi^2 \rho^2 \Gamma_{t+1} \Gamma_t \mathrm{var}(\eta_t), \tag{106}$$

$$Ex_t^2 = b_x^2 \mathrm{var}(\nu_t) + \sum_{j=0}^{t-1} (b_x \rho^{j+1} + b_\pi a_x \Gamma_{j+1})^2 \mathrm{var}(\nu_t) + \rho^2 (b_x \rho^t + b_\pi a_x \Gamma_t)^2 \mathrm{var}(\eta_t), \tag{107}$$

where

$$\Gamma_j = \frac{a_\pi^j - \rho^j}{a_\pi - \rho}. \tag{108}$$

From this it follows that

$$\begin{aligned}\sum_{t=0}^{\infty}\beta^t E\pi_t^2 &= b_\pi^2\sum_{t=0}^{\infty}\beta^t\sum_{j=0}^{\infty}\beta^j\Gamma_{j+1}^2\mathrm{var}(\nu_t)+b_\pi^2\rho^2\sum_{t=0}^{\infty}\beta^t\Gamma_{t+1}^2\mathrm{var}(\eta_t)\\ &= \frac{b_\pi^2(1+\beta a_\pi\rho)\mathrm{var}(\nu_t)}{D(1-\beta)}+\frac{b_\pi^2\rho^2(1+\beta a_\pi\rho)\mathrm{var}(\eta_t)}{D},\end{aligned}\tag{109}$$

$$\begin{aligned}\sum_{t=0}^{\infty}\beta^t E\pi_t\pi_{t-1} &= b_\pi^2\sum_{t=0}^{\infty}\beta^t\sum_{j=1}^{\infty}\beta^j\Gamma_{j+1}\Gamma_j\mathrm{var}(\nu_t)+b_\pi^2\rho^2\sum_{t=1}^{\infty}\beta^t\Gamma_{t+1}\Gamma_t\mathrm{var}(\eta_t)\\ &= \frac{b_\pi^2\beta(a_\pi+\rho)\mathrm{var}(\nu_t)}{D(1-\beta)}+\frac{b_\pi^2\beta\rho^2(a_\pi+\rho)\mathrm{var}(\eta_t)}{D},\end{aligned}\tag{110}$$

$$\sum_{t=0}^{\infty}\beta^t Ex_t^2=\frac{b_x^2\mathrm{var}(\nu_t)}{1-\beta}+\frac{K\mathrm{var}(\nu_t)}{D(1-\beta)}+\frac{\rho^2K\mathrm{var}(\eta_t)}{D},\tag{111}$$

where

$$D=(1-\beta a_\pi^2)(1-\beta a_\pi\rho)(1-\beta\rho^2),$$

$$\begin{aligned}K=&\beta a_\pi^2b_\pi^2a_x^2+2\beta\rho a_\pi b_\pi^2a_x^2-\beta^2\rho a_\pi^3b_\pi^2a_x^2+2\beta\rho^2a_\pi b_xb_\pi a_x-2\beta^2\rho^2a_\pi^3b_xb_\pi a_x+\beta\rho^2b_\pi^2a_x^2\\ &-\beta^2\rho^2a_\pi^2b_\pi^2a_x^2+2\beta\rho^3b_xb_\pi a_x-2\beta^2\rho^3a_\pi^2b_xb_\pi a_x-\beta^2\rho^3a_\pi b_\pi^2a_x^2+\beta^3\rho^3a_\pi^3b_\pi^2a_x^2\\ &+\beta\rho^4b_x^2-\beta\rho^4a_\pi^2b_x^2-2\beta^2\rho^4a_\pi b_xb_\pi a_x+2\beta^3\rho^4a_\pi^3b_xb_\pi a_x-\beta^2\rho^5a_\pi b_x^2+\beta^3\rho^5a_\pi^3b_x^2.\end{aligned}$$

Taking the variance of equation (96) we find that

$$\mathrm{var}(\eta_t)=\frac{\mathrm{var}(\nu_t)}{1-\rho^2}.\tag{112}$$

Furthermore, notice that

$$\sum_{t=0}^{\infty}\beta^t E\pi_{t-1}^2=\beta\sum_{t=0}^{\infty}\beta^t E\pi_t^2.\tag{113}$$

Equations (109)-(113) can now be used to evaluate equation (104) without reference to the $\mathrm{var}(\eta_t)$ and $\mathrm{var}(\nu_t)$.

References

[1] Ball, Laurence (1994): "What Determines the Sacrifice Ratio," in N. Gregory Mankiw, ed., *Monetary Policy,* Cambridge, Mass.: MIT Press.

[2] Ball, Laurence (1997): "Efficient Rules for Monetary Policy," NBER Working Paper No. 5952, March 1997.

[3] Ball, Laurence (1999): "Policy Rules for Open Economies," in John B. Taylor, ed., *Monetary Policy Rules,* 127-144, Chicago: University of Chicago Press.

[4] Barro, Robert J. and David B. Gordon (1983): "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy,* 91: 589-610.

[5] Blake. A. P. and P. F. Westaway (1996): "Credibility and the Effectiveness of Inflation Targeting Regimes," *Manchester School,* 64: 28-50.

[6] Blanchard, Olivier J. and Charles Kahn (1980): "The Solution of Linear Difference Equations under Rational Expectations," *Econometrica,* 48: 1305-1311.

[7] Blanchard, Olivier J. and Nobu Kiyotaki (1987): "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review,* 77: 647-666.

[8] Blinder, Alan S. (1998): *Central Banking in Theory and Practice,* Cambridge Mass.: MIT Press.

[9] Calvo Guillermo A. (1983): "Staggered Prices in a Utility-Maximizing Framwork," *Journal of Monetary Economics,* 12: 383-398.

[10] Campbell, John Y. and N. Gregory Mankiw (1989): "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," in Olivier J. Blanchard, ed., *NBER Macroeconomic Annual 1989,* 185-216. Cambridge, Mass.: MIT Press.

[11] Clarita, Richard, Jordi Gali and Mark Gertler (1999): "The Science of Monetary Policy: A New Keynesian Perspective," NBER Working Paper No. 7147, May 1999.

[12] Dixit, Avinash K., and Joseph E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review,* 67: 297-308.

[13] Friedman, Milton (1968): "The Role of Monetary Policy," *American Economic Review,* 58: 1-19.

[14] Friedman, Milton (1969): "The Optimum Quantity of Money," in *The Optimum Quantity of Money and Other Essays,* 1-50. Chicago: Aldine Publishing.

[15] Friedman, Milton and Anna J. Schwarts (1963): *A Monetary History of the United States,* Princeton N.J.: Princeton University Press.

[16] Fuhrer, Jeffrey C. (1997): "The (Un)importance of Forward-Looking Behavior in Price Specifications," *Journal of Money, Credit and Banking,* 29: 338-350.

[17] Fuhrer, Jeffery C. and George R. Moore (1995): "Inflation Persistence," *Quarterly Journal of Economics,* 109: 197-230.

[18] Gali, Jordi and Mark Gertler (1999), " Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics,* 195-222.

[19] Kydland, Finn E. and Edward C. Prescott (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy,* 85: 473-491.

[20] Levin, Andrew, Volker Wieland, and John C. Williams (1999): "Robustness of Simple Monetary Policy Rules under Model Uncertainty," in John B. Taylor, ed., *Monetary Policy Rules,* 263-316. Chicago: University of Chicago Press.

[21] Lucas, Robert E., Jr. (1976): "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy,* 1: 19-46.

[22] McCullum, Bennett T. (1981): "Price Level Determinacy with an Interest Rate Policy Rule and Rational Expectations," *Journal of Monetary Economics,* 8: 319-329.

[23] Obstfeld, Maurice and Kenneth Rogoff (1996): *Foundations of International Macroeconomics.* Cambridge, Mass.: MIT Press.

[24] Phelps, Edmund S. (1967): "Phillips Curves, Expectations of Inflation and Optimal Unemployment over Time," *Economica,* 34: 254-281.

[25] 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.

[26] Poole, William (1970): "Optimal Choice of Monetary Policy Instrument in a Simple Stochastic Macro Model," *Quarterly Journal of Economics,* 84: 197-216.

[27] Roberts, John M. (1995): "New Keynesian Economics and the Phillips Curve," *Journal of Money, Credit and Banking,* 27: 975-984.

[28] Rotemberg, Julio J. and Michael Woodford (1997): "An Optimizing-Based Econometric Framework for the Evaluation of Monetary Policy," in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER Macroeconomics Annual 1997,* 297-346. Cambridge, Mass.: MIT Press.

[29] Rotemberg, Julio J. and Michael Woodford (1998): "An Optimizing-Based Econometric Framework for the Evaluation of Monetary Policy," NBER Technical Working Paper No. 233, May 1998.

[30] Rotemberg, Julio J. and Micheal Woodford (1999): "Interest Rate Rules in an Estimated Sticky Price Model," in John B. Taylor, ed., *Monetary Policy Rules,* 57-119. Chicago: University of Chicago Press, 1999.

[31] Sargent, Thomas J. (1993): *Rational Expectations and Inflation,* 2nd ed., 40-119. New York: HarperCollins Publishers.

[32] Sargent, Thomas J. and Neil Wallace (1975): "Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy,* 83: 241-254.

[33] Svensson, Lars E.O. (1986): "Sticky Goods Prices, Flexible Asset Prices, Monopolistic Competition and Monetary Policy," *Review of Economic Studies,* 53: 385-405.

[34] Svensson, Lars E.O. (1997): "Inflation Forcast Targeting: Implementing and Monitoring Inflation Targets," *European Economic Review,* 41: 1111-1146.

[35] Söderlind, Paul (1999): "Solution and Estimation of RE Macromodels with Optimal Policy," *European Economic Review,* 43: 813-823.

[36] Taylor, John B. (1993): "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy,* 39: 195-214.

[37] Taylor, John B. (ed) (1999): *Monetary Policy Rules,* Chicago: University of Chicago Press, 1999.

[38] Walsh, Carl E. (1998): *Monetary Theory and Policy,* Cambridge Mass.: MIT Press.

[39] Woodford, Michael (1996): "Control of the Public Debt: A Requirement for Price Stability?," NBER Working Paper No. 5684, July 1996.

[40] Woodford, Michael (1999a): "Price-Level Determination Under Interest-Rate Rules," unpublished manuscript, Princeton University, April 1999.

[41] Woodford, Michael (1999b): "Inflation Stabilization and Welfare," unpublished manuscript, Princeton University, August 1999.

[42] Woodford, Michael (1999c): "Optimal Monetary Policy Inertia," NBER Working Paper No. 7261, August 1999.

[43] Woodford, Michael (1999d): "Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?," unpublished, Princeton University, September 1999.

[44] Woodford, Michael (forthcoming): "Optimal Models with Nominal Rigidities," unpublished manuscript, Princeton University.

[45] Yun, Tack (1996): "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles," *Journal of Monetary Economics,* 37: 345-370.

Table 1: Calibrated parameters for different chioces of ω

ω	χ^f	χ^b	κ	λ_y	λ_Δ
0	1	0	0.05	0.01	0
0.1	0.890	0.111	0.0400	0.01	0.139
0.2	0.801	0.200	0.0321	0.01	0.313
0.3	0.729	0.273	0.0255	0.01	0.536
0.5	0.617	0.386	0.0154	0.01	1.25
0.8	0.502	0.502	0.0050	0.01	5.00
0.9	0.473	0.532	0.0024	0.01	11.25
0.99	0.449	0.556	0.0002	0.01	123.75

Table 2: A comparison of welfare under commitment and discretion with the theoretical loss function

	Welfare		
ω	Commitment	Discretion	Difference
0.01	63.5	82.5	30%
	(0.1)	(0.1)	
0.2	145.1	169.6	17%
	(0.2)	(0.2)	
0.5	470	518	10%
	(0.5)	(0.6)	
0.8	2023	2480	23%
	(2)	(3)	

The standard deviation of the welfare estimate is reported in apprentices.

Table 3: Optimal Taylor Rules

Optimal Line:		$\theta_x = \alpha + \beta\theta_\pi$	
ω	Loss Function	α	β
0.01	Traditional	-0.401	0.200
0.2	Theoretical	-0.188	0.141
0.8	Traditional	-0.144	0.126
0.8	Theoretical	-0.101	0.082

iid case

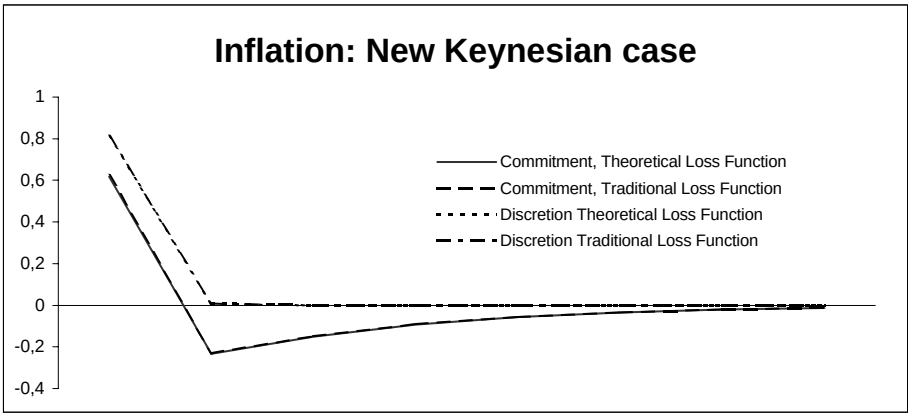

Figure 1

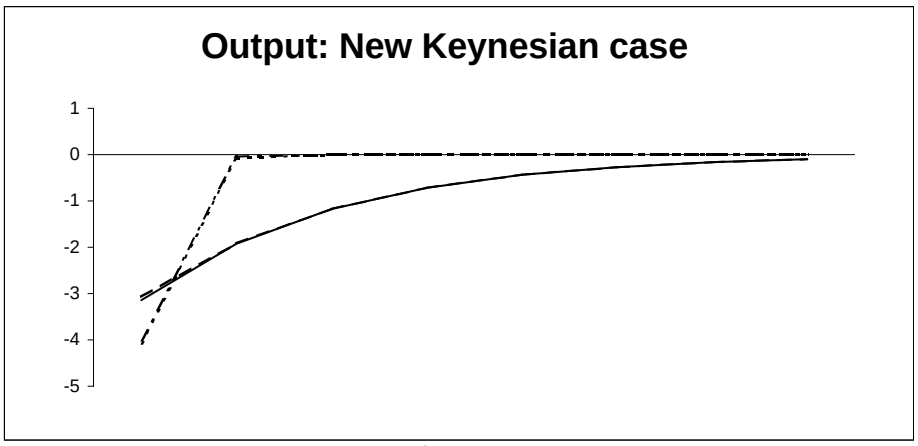

Figure 2

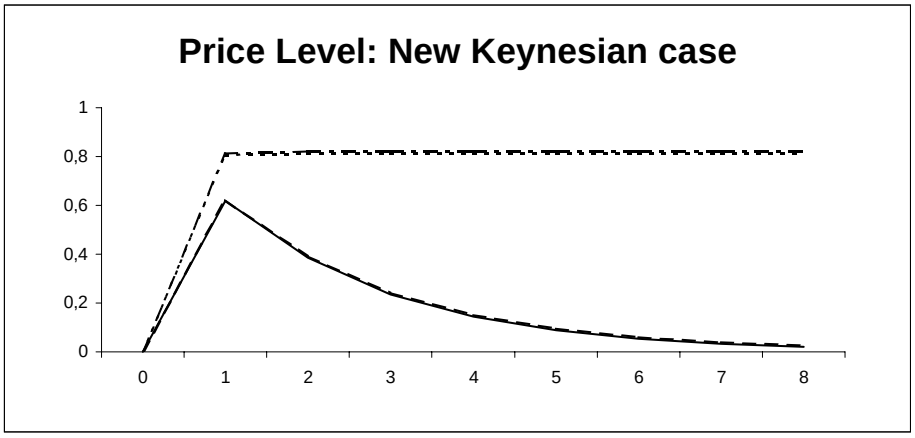

Figure 3

iid case

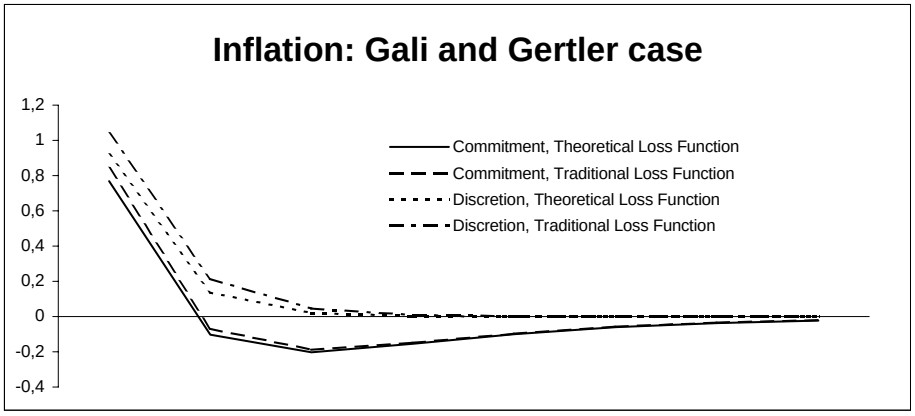

Figure 4

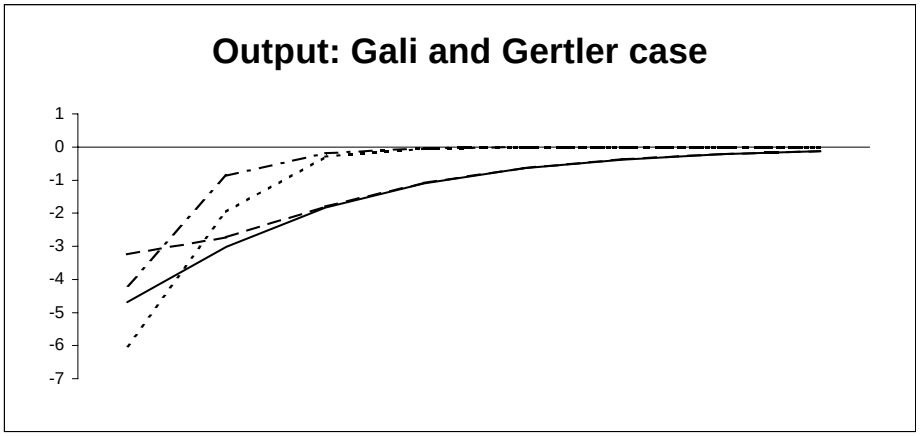

Figure 5

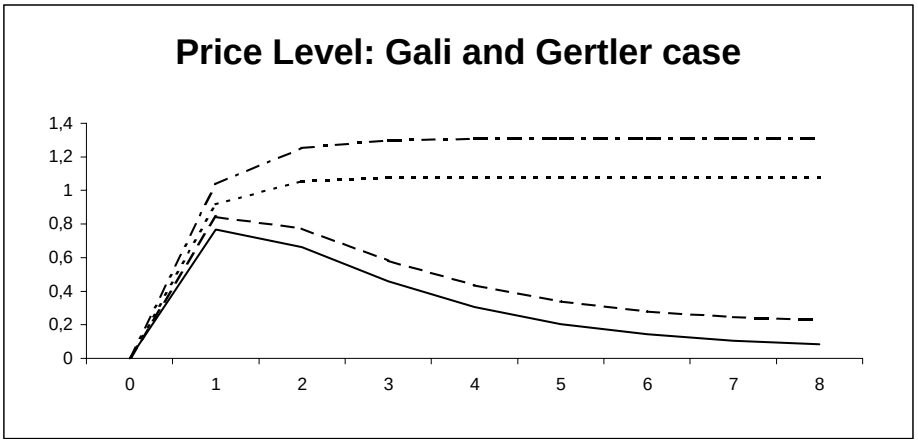

Figure 6

iid case

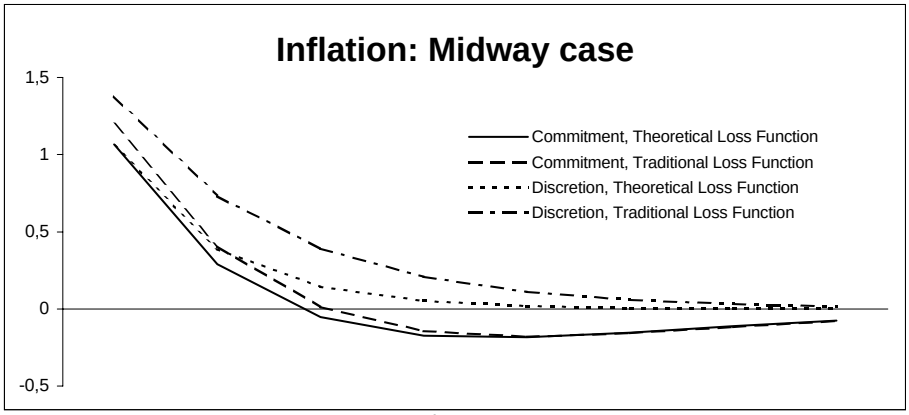

Figure 7

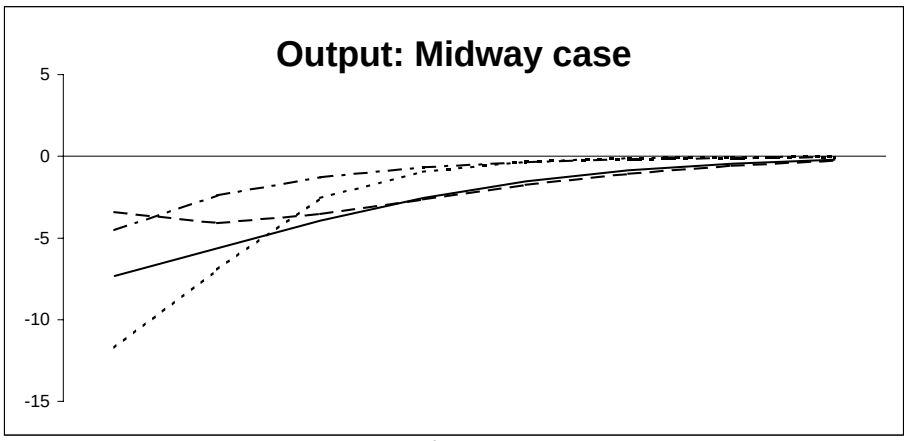

Figure 8

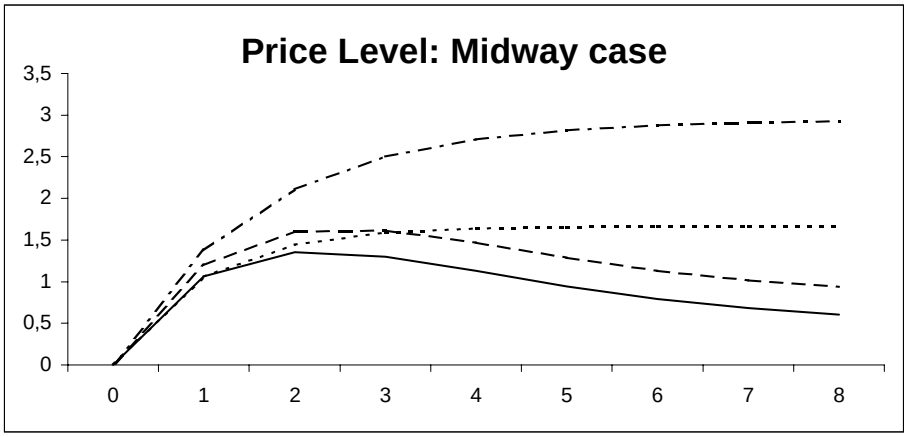

Figure 9

iid case

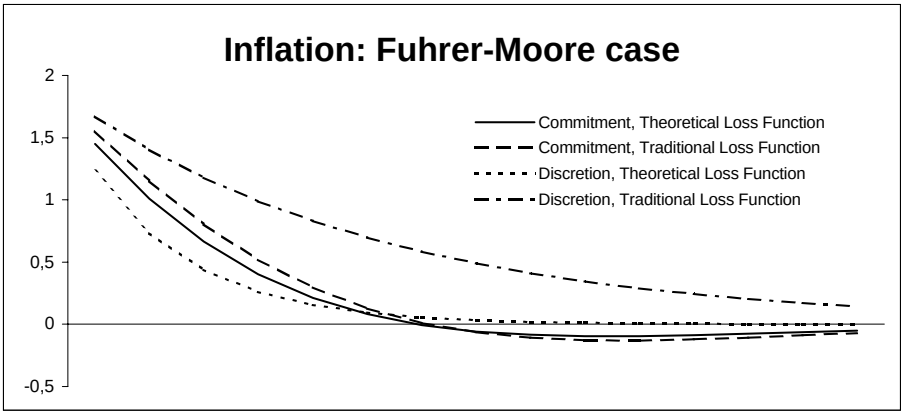

Figure 10

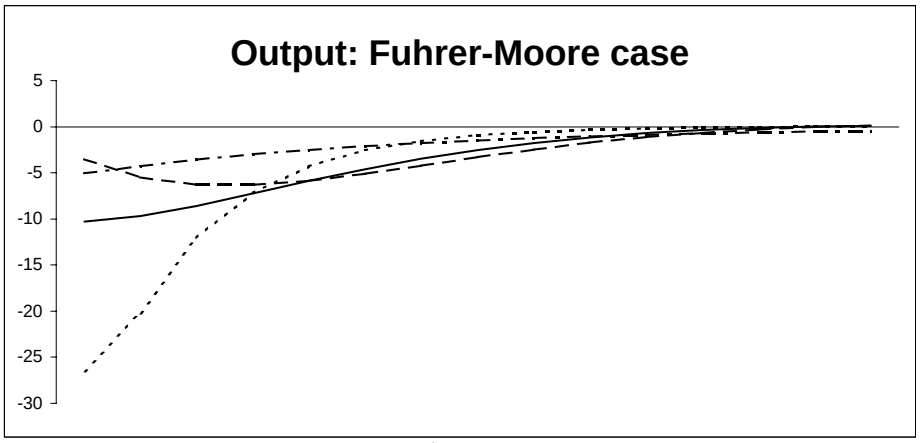

Figure 11

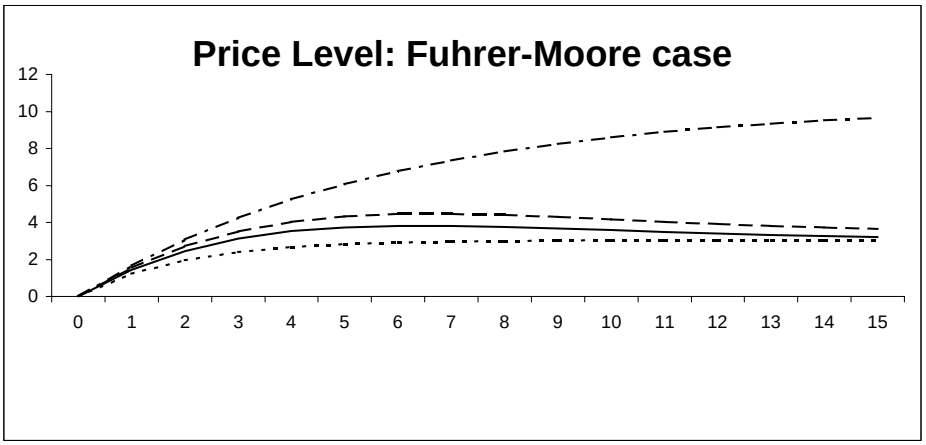

Figure 12

AR(1) case

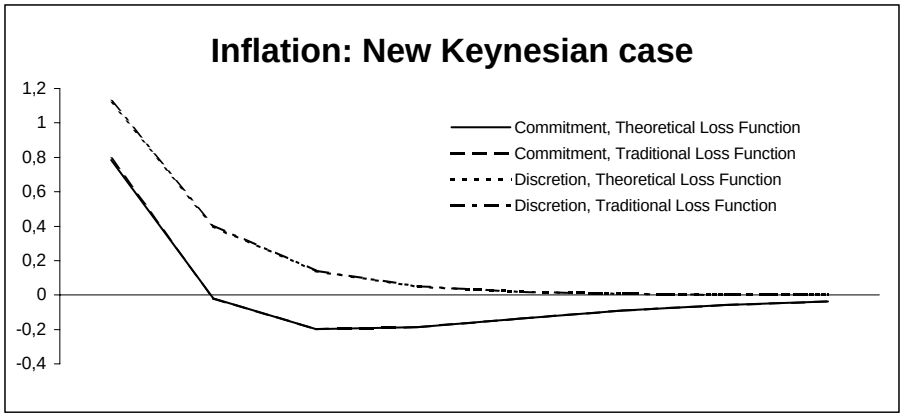

Figure 13

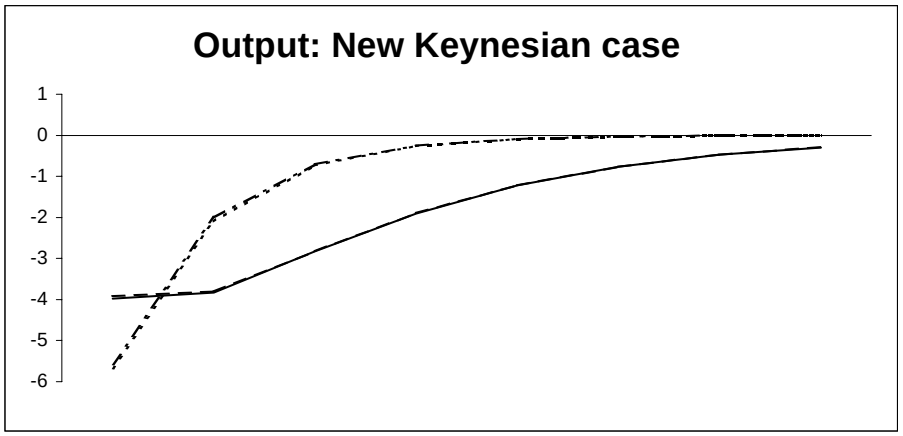

Figure 14

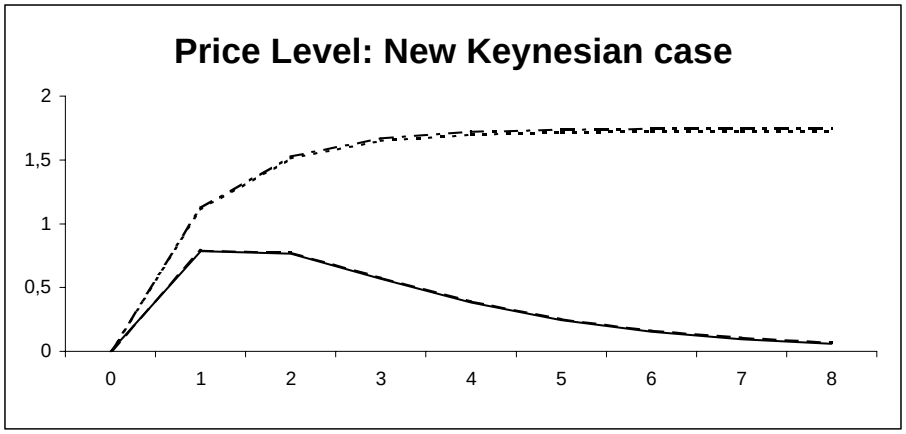

Figure 15

AR(1) case

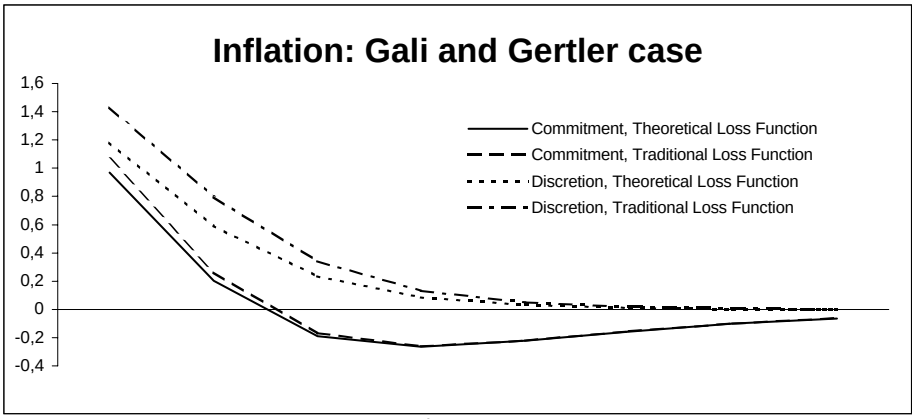

Figure 16

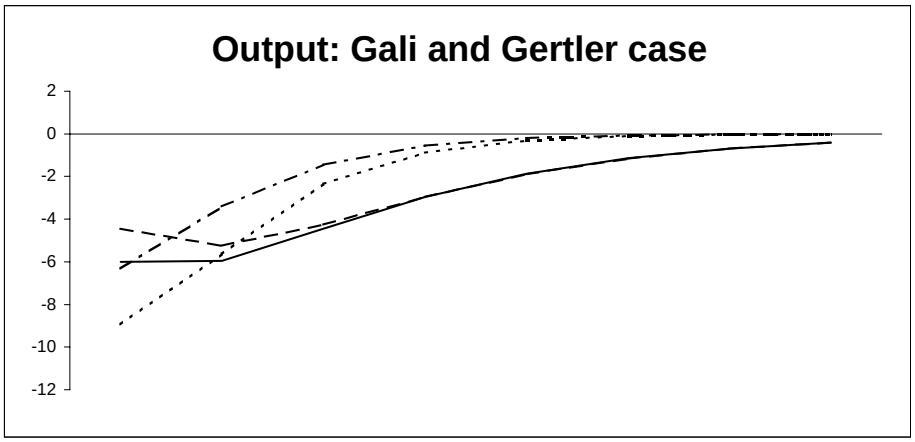

Figure 17

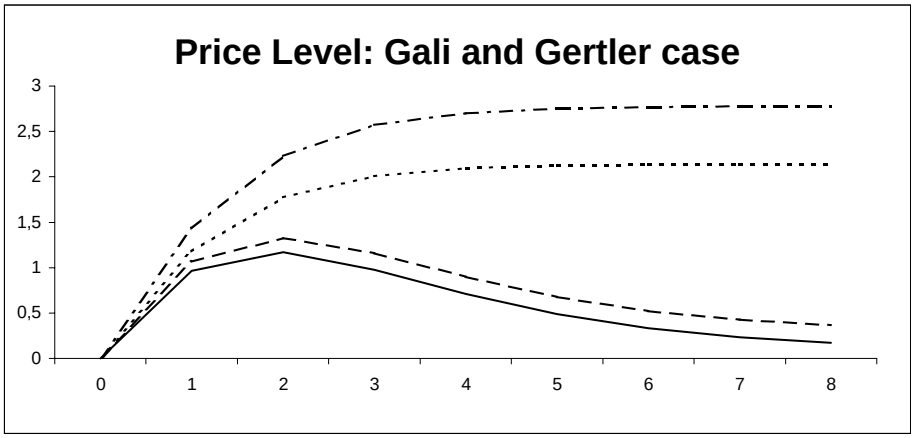

Figure 18

AR(1) case

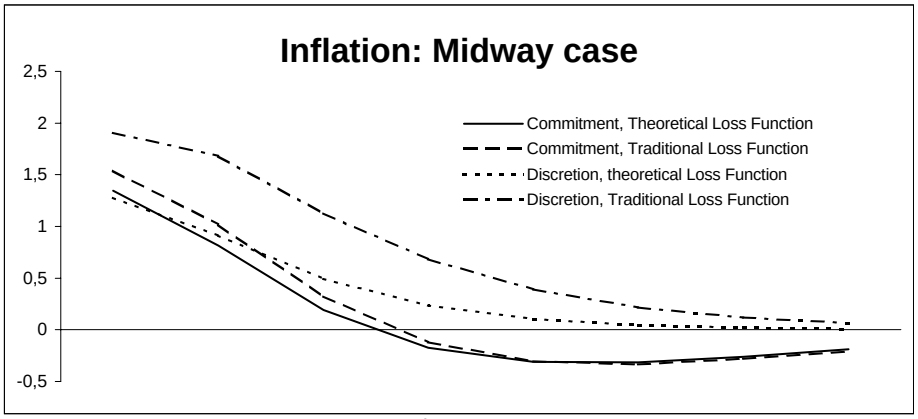

Figure 19

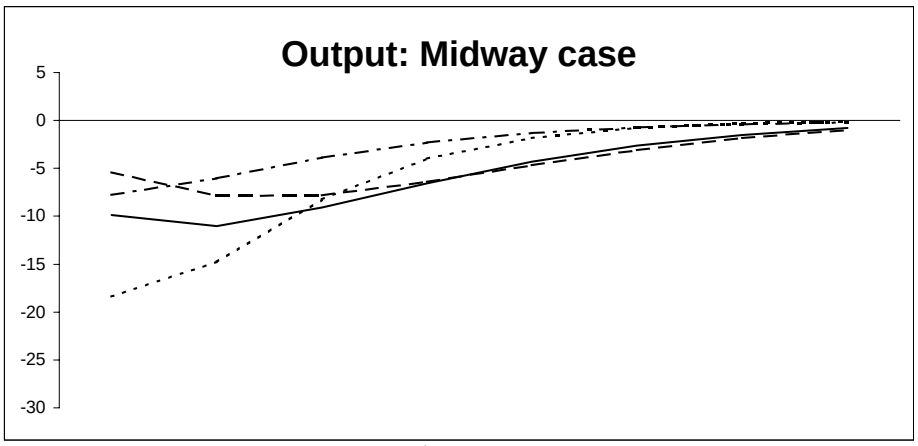

Figure 20

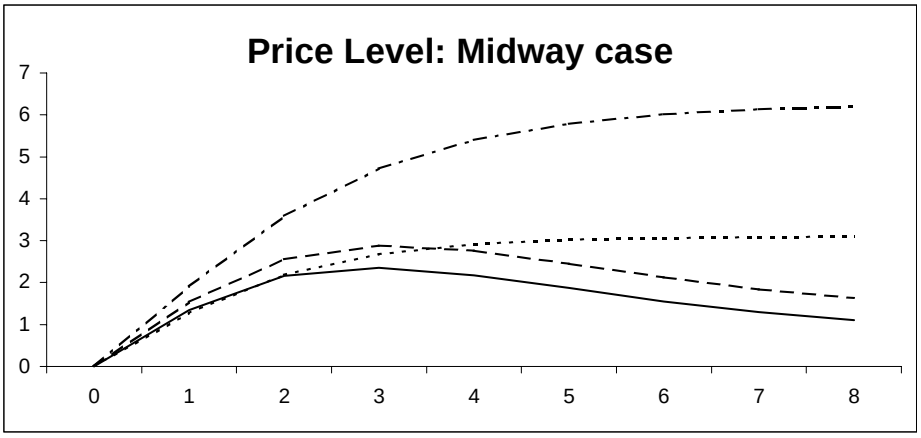

Figure 21

AR(1) case

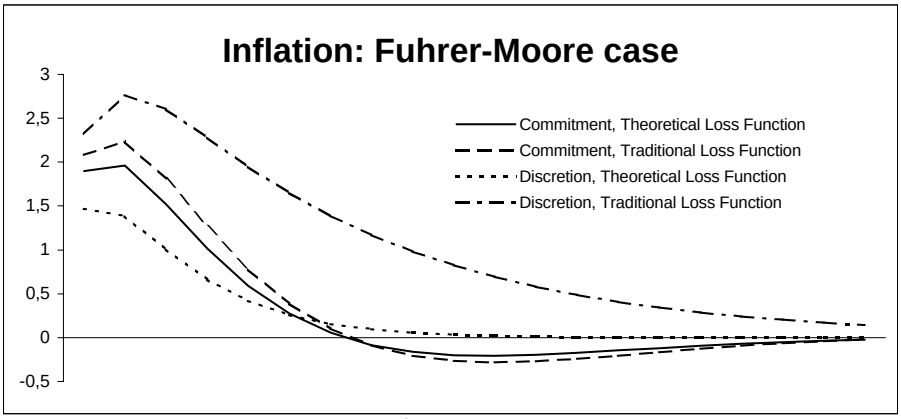

Figure 22

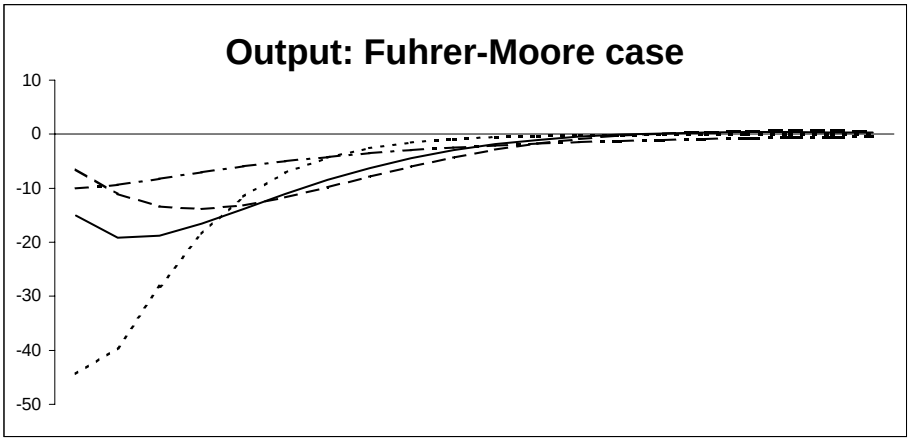

Figure 23

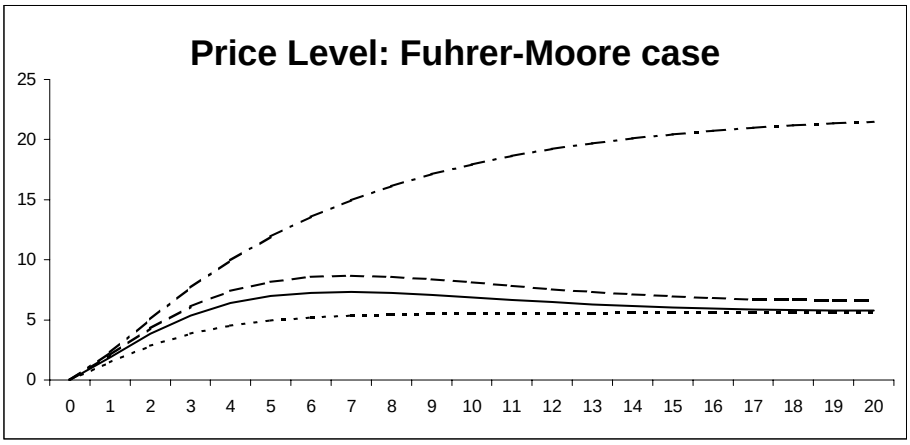

Figure 24

AR(1) case, Traditional Loss Function

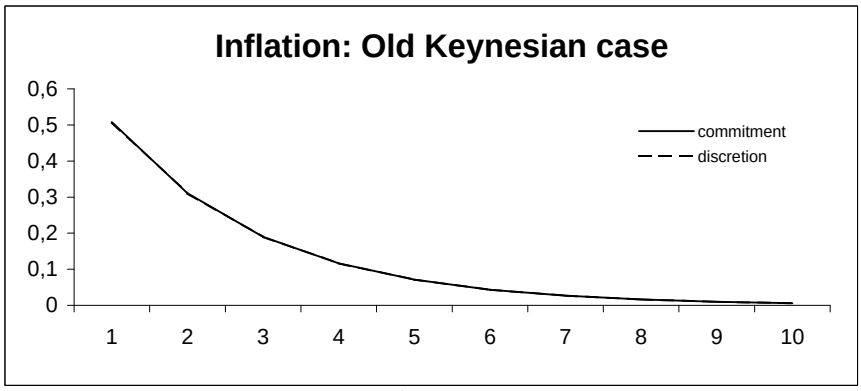

Figure 25

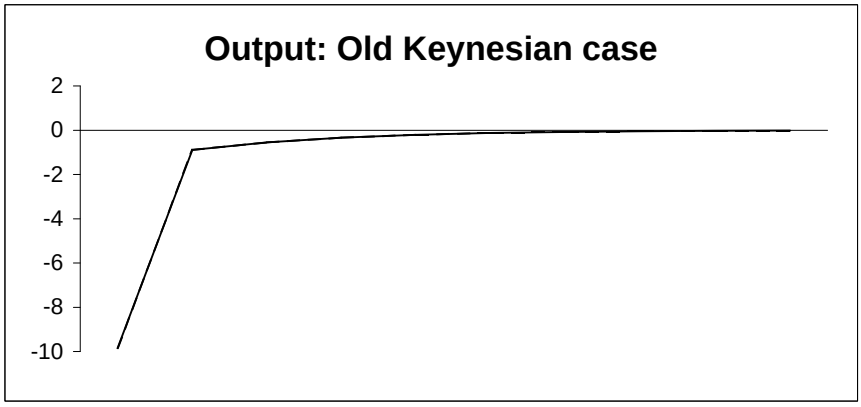

Figure 26

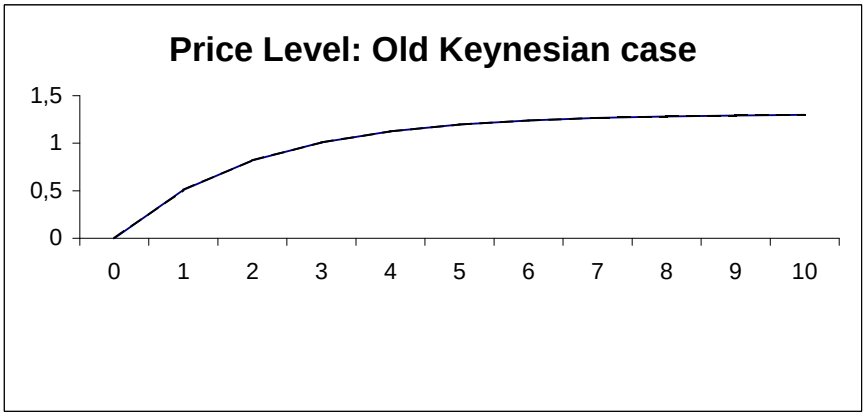

Figure 27

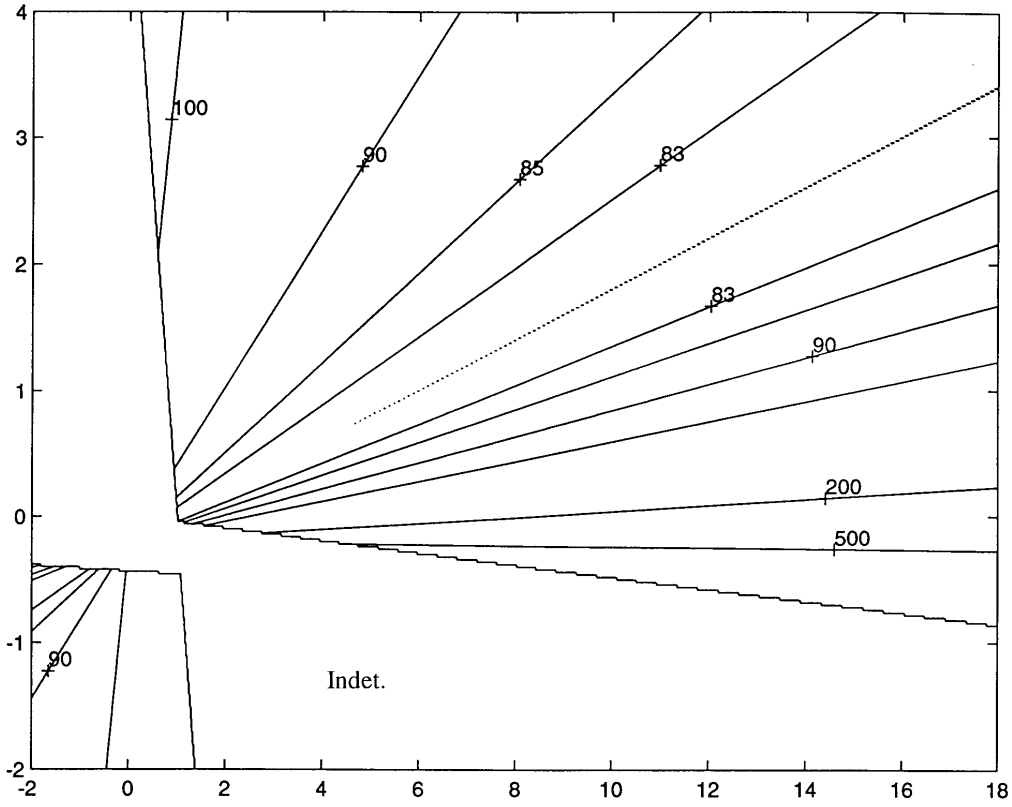

Figure 28: Loss Associated with Taylor Rules for $\omega = 0.01$ and the Traditional Loss Function

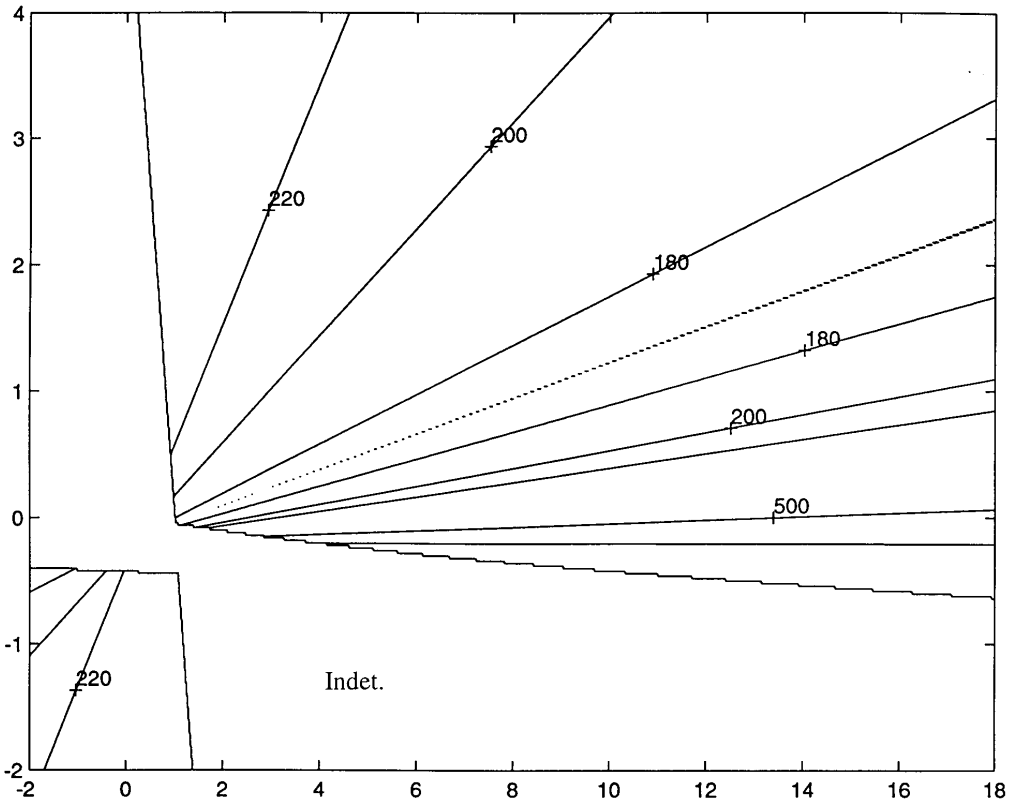

Figure 29: Loss Associated with Taylor Rules for $\omega = 0.2$ and the Theoretical Loss Function

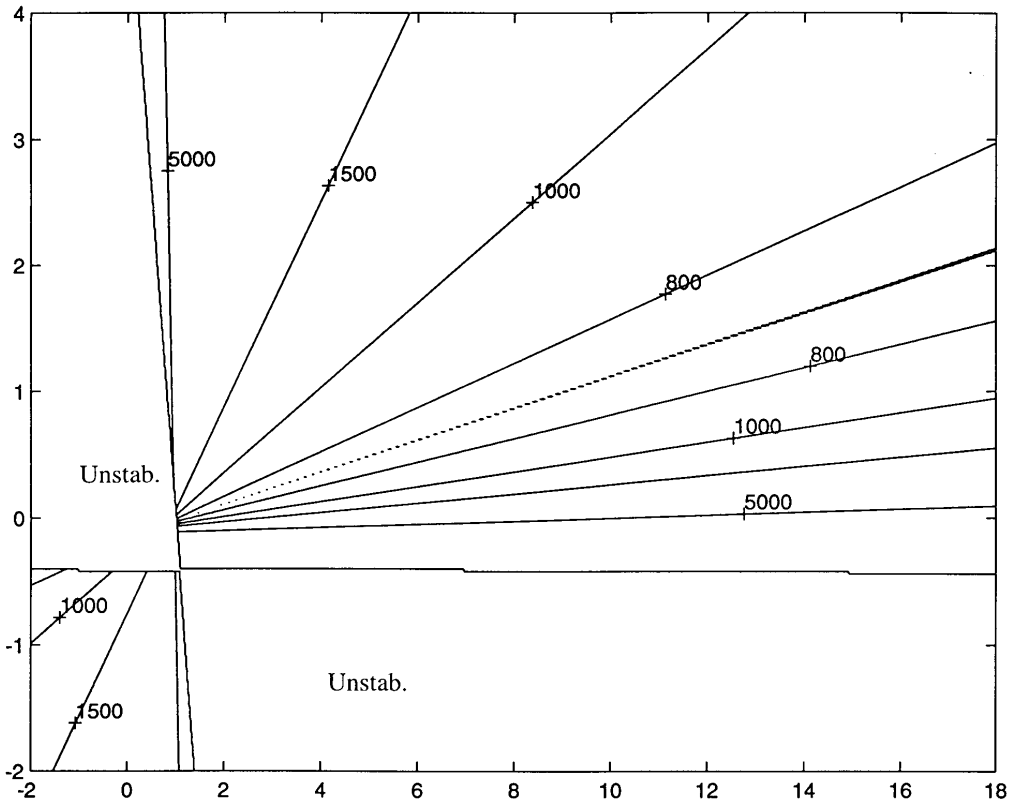

Figure 30: Loss Associated with Taylor Rules for $\omega = 0.8$ and the Traditional Loss Function

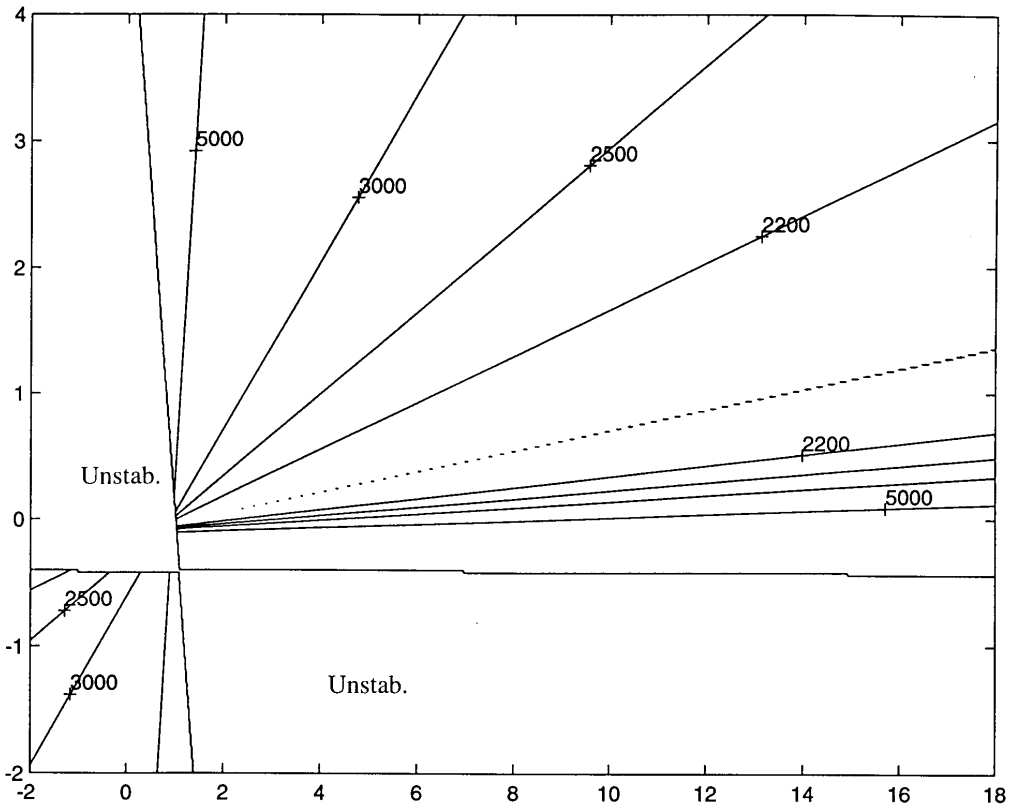

Figure 31: Loss Associated with Taylor Rules for $\omega = 0.8$ and the Theoretical Loss Function