Monetary Policy for Inattentive Economies

Laurence Ball

Johns Hopkins University

N. Gregory Mankiw

Harvard University

Ricardo Reis[1]

Harvard University

January 2003

Abstract

This paper is a contribution to the analysis of optimal monetary policy. It begins with a critical assessment of the existing literature, arguing that most work is based on implausible models of inflation-output dynamics. It then suggests that this problem may be solved with some recent behavioral models, which assume that price setters are slow to incorporate macroeconomic information into the prices they set. A specific such model is developed and used to derive optimal policy. In response to shocks to productivity and aggregate demand, optimal policy is price level targeting. Base drift in the price level, which is implicit in the inflation targeting regimes currently used in many central banks, is not desirable in this model. When shocks to desired markups are added, optimal policy is flexible targeting of the price level. That is, the central bank should allow the price level to deviate from its target for a while in response to these supply shocks, but it should eventually return the price level to its target path. Optimal policy can also be described as an elastic price standard: the central bank allows the price level to deviate from its target when output is expected to deviate from its natural rate.

1   INTRODUCTION

What policy rule should a central bank follow? Recent years have seen a resurgence of theoretical research on this classic question. Most of this work has built on "new Keynesian" models of the output-inflation trade-off derived from forward-looking models of staggered price adjustment. Unfortunately, these models make implausible predictions about the effects of monetary policy: for example, they imply that a policy change that gradually reduces inflation causes an output boom. There is therefore good reason to be skeptical about what the literature tells us about the effects of alternative policies.

This paper tries to make progress toward determining which policies are optimal by studying this question in a model that more closely fits the facts about monetary policy. In particular, we draw on recent behavioral models of the output-inflation trade-off based on the assumption that agents are slow to incorporate information about macroeconomic conditions, even if the information is publicly available. Recent work has shown that such models capture the inertia that is central to inflation dynamics in modern economies. These models should provide more reliable insights into the policy choices facing central banks.

Section 2 of this paper briefly reviews the two literatures on which our work builds, the work on optimal monetary policy and the work on behavioral macroeconomics. Section 3 presents a specific model, which builds on the "sticky information" model of Mankiw and Reis (2002), and Section 4 discusses the determinants of welfare in this model. Sections 5, 6 and 7 derive the optimal policy rules in the model. Section 8 compares our results with those obtained from the standard new Keynesian Phillips curve. Section 9 concludes.

Our central result is that price level targeting is the optimal policy in the model. Inflation targeting – the currently popular policy of allowing base drift in the price level – is suboptimal. When the economy is hit by shocks to aggregate demand or productivity, strict price level targeting is optimal: policymakers should return the price level to a predetermined path as quickly as possible. However, if there are persistent shocks to firms' markups, the optimal rule allows temporary deviations from the long-run price target. In this way, the prescriptions of our model are similar to the practice of many central banks, which allow temporary deviations from policy rules in response to "supply shocks." One can

also describe optimal policy as the elastic price standard proposed by Hall (1984). Under this policy, the price level can deviate from target as long as output is expected to deviate from its natural rate.

2 MOTIVATION

This paper arises from two recent literatures – one on optimal rules for monetary policy and one on behavioral approaches to the Phillips curve. We believe that the second may hold the hope of remedying some deficiencies in the first. The natural place to start our analysis is with a brief overview of these two broad literatures.

2.1 The Sorry State of Monetary Policy Analysis

In recent research on policy rules, a canonical approach has emerged. The researcher sets up a model based on monopolistically competitive producers, with monetary nonneutrality arising because of some friction in price setting. Policymakers are assumed to choose policy to minimize a loss function, which is either assumed or derived as an approximation to agents' utility functions. With this framework in hand, the researcher can derive the optimal policy rule and evaluate policy proposals, such as targeting inflation or the price level. Goodfriend and King (1997) have referred to this literature as the New Neoclassical Synthesis, while Clarida, Gali, and Gertler (1999) have called it the New Keynesian Science of Monetary Policy. Whatever the label, there is no doubt that this approach to monetary-policy analysis is large and growing.

Yet policymakers should be wary of the prescriptions that this literature has yielded. The results in any such analysis depend crucially on the assumed model of the Phillips curve, which determines how inflation is related to output movements and expectations. The results should be believed only if the assumed Phillips curve is credible. Unfortunately, that is rarely the case.

The most common approach is to use the new Keynesian Phillips curve. This model is based on the assumption of time-contingent price adjustment and is derived from the seminal papers of Taylor (1980), Rotemberg (1982), and Calvo (1983). The model's appealing

microeconomic foundations have made it the workhorse of much of the modern literature on monetary policy. The papers by Goodfriend and King (1997) and Clarida, Gali, and Gertler (1999) are two examples.

There is, however, a problem with this approach: Because the new Keynesian Phillips curve lacks any source of inflation inertia, it makes absurdly counterfactual predications about the effects of monetary policy. According to this model, disinflations can result in booms rather than recessions (Ball, 1994). Similarly, the model fails to explain why monetary policy shocks have a delayed and gradual effect on inflation (Mankiw, 2001). One should be suspect of policy recommendations arising from any model that is patently inconsistent with both the econometric evidence and the views of central bankers about their influence over the economy.

An alternative approach to using the new Keynesian Phillips curve is to use a more traditional accelerationist Phillips curve. This approach, which can be justified with the assumption of backward-looking expectations, has the advantage that it is consistent with the conventional wisdom about the effects of monetary policy and with standard empirical analyses of inflation (e.g., Gordon, 1997; Staiger, Stock, and Watson, 1997). For these reasons, Ball (1999) and Svensson (1997) take an accelerationist Phillips curve as a starting point for their analyses.

Yet this approach is also problematic. Although inflation inertia is a feature of the U.S. monetary regime of the past several decades, this was not true when the economy operated under a gold standard. (Barsky, 1987; Alogoskoufis and Smith, 1991). When analyzing alternative policy rules, it is imperative that expectations be allowed to adjust to the new regime. Regardless of how well it fits the recent data, the accelerationist Phillips curve is a suspect tool for monetary-policy analysis. When analyzing alternative monetary policy rules, ignoring the Lucas critique is not just an aesthetic faux pas; it produces results that are, literally, incredible.

The problems with both the new Keynesian and the accelerationist Phillips curves are increasingly understood. This has led some authors to suggest a compromise. According to the so-called "hybrid" Phillips curve, a subset of agents have backward-looking inflation expectations and the rest have rational expectations (e.g., Gali and Gertler, 1999). Yet this

compromise may yield the worst of both worlds. Like the new Keynesian Phillips curve, the hybrid model yields an immediate jump in inflation in response to monetary policy shocks (unless all agents are backward looking). Like the accelerationist Phillips curve, the hybrid model fails to explain the absence of inflation inertia under earlier monetary regimes (unless no agents are backward looking). That is, by taking a weighted average of two flawed models, the hybrid model of the Phillips curve ends up with the flaws of each.

2.2   The Promise of Behavioral Models

Behavioral economics might offer a way out of this conundrum. Behavioral economics is, broadly defined, the growing subfield that incorporates into economic theory the flaws in human decision-making that are ignored in the standard model of rational man. It finds its roots in Herbert Simon's suggestion that people are "satisficers" rather than rational maximizers. Many economists have suggested that departures from rationality may be important for issues in macroeconomics (Akerlof, 2002; Sargent, 1993). In particular, several recent papers on the microfoundations of the Phillips curve have proposed that monetary nonneutrality arises because people are slow to process widely available macroeconomic information.

When Milton Friedman (1968) introduced the natural rate hypothesis in his AEA presidential address, he proposed a theory of monetary nonneutrality that, to modern ears, sounds remarkably behavioral. According to Friedman, when the money supply rises unexpectedly, the price level rises, pushing down the real wage. Employers hire more because the cost of labor has fallen. Employees are willing to work more because they focus on the nominal wage and infer (incorrectly) that the reward for working has risen. Friedman did not explain this asymmetry in information between firms and workers. For him, it seemed natural to assume that workers lacked full information about macroeconomic conditions. This ignorance was the centerpiece of Friedman's proposed explanation for the short-run Phillips curve.

Work that followed on Friedman's ideas tried to incorporate rational economic man into the story. The task has proved difficult. Lucas (1973) suggested a model in which producers observe the prices of what they sell but not of what they buy and, as a result, need to solve signal-extraction problems to infer relative prices. Critics of the Lucas model wondered why a rational man would fail to use the widely available information on the consumer price

index, which is published monthly. Fischer (1977) told a version of Friedman's story in which workers were locked into long-term contracts setting the nominal wage. Critics of the Fischer model wondered why a rational man would ever agree to such an inefficient contract ex ante or fail to renegotiate it ex post.

In recent years, there have been several attempts to come up with better models of inflation dynamics by returning to Friedman's idea that some people fail to incorporate all available macroeconomic information into their decision-making. According to the new Keynesian Phillips curve, prices are sticky, but inflation can respond instantly to changes in monetary policy. In practice, inflation responds sluggishly. This fact suggests that, for some reason, price setters aren't quite awake or smart or informed enough to process all available information about monetary policy and react immediately to it.

One approach to modeling this inattentiveness is to use the tools of information theory, as exposited, for instance, in the textbook by Cover and Thomas (1991). Drawing on these tools, Sims (2001) suggests modeling humans as having a limited channel for absorbing information. That is, the human brain is imperfect in the same way as a computer with a slow internet connection would be. Woodford (2001) uses this idea to build a model of inflation-output dynamics. In his model, because price setters learn about monetary policy through a limited-information channel, it is as if they observe monetary policy with a random error and have to solve a signal-extraction problem along the lines of Lucas (1973).

Ball (2000) proposes another approach to the problem. He suggests that when forming expectations of any variable, people optimally use all information in the past values of that variable, but fail to incorporate information from other variables. That is, expectations are based on optimal univariate forecasts. Ball shows that this approach can explain why the accelerationist Phillips curve fits the recent data well, while the earlier data conform more closely to a classic Phillips curve.

Mankiw and Reis (2002), like Ball and Woodford, impose a constraint on the information that people use when forming expectations. They assume that in each period there is a fixed probability that a person updates his information set; otherwise, he continues to set prices based on outdated information. Mankiw and Reis compare this sticky-information model with the standard new Keynesian Phillips curve and conclude that it is more consistent with

5

standard views about the effects of monetary policy. The Mankiw-Reis model has strong formal ties to the earlier efforts of Lucas and Fischer, but its bald informational assumption rejects the attempt to reconcile the Phillips curve with the axiom of rationality.

A weakness of all three of these approaches is that they fail to explain why people do not incorporate widely available news about monetary policy into their plans. In reality, it is easy to find out what the central bank is doing, but it is often hard to figure out what it means. That is, the real problem is not obtaining information but processing it. Unfortunately, economists do not have the tools to model imperfect information processing. The approaches of Woodford, Ball, and Mankiw and Reis are all based on the hope that a model of imperfect information acquisition may serve as a rough substitute.

A strength of all three approaches is that they can explain inflation inertia. In these models, when monetary policy changes, most price setters are unaware of it, so they keep marking up prices as if no policy change had occurred. That is, all three models generate inflation inertia by positing a type of inattentiveness on the part of price setters. Because these models of inattention fit the most basic facts about inflation dynamics, they hold out the greatest promise for delivering credible advice for monetary policy.

In this paper, the specific model we use is the sticky-information model proposed by Mankiw and Reis. One reason is tractability. A nice feature of the Mankiw-Reis model is that, given the dates when people get their information, they have conventional rational expectations. This feature allows the use of rational-expectations tools to solve for the resulting equilibrium. That is, we can use many standard and powerful modeling techniques, while making the behavioral assumption that price setters are inattentive.

3   THE MODEL

We begin our analysis by presenting the model, including agent's objectives, market structure, price setting, and the role of aggregate-demand policy.

3.1 Agents and Market Structure

We assume a simple variation on the standard "yeoman farmer economy" studied by authors such as Rotemberg and Woodford (1997) and Woodford (2002, chapter 6). The economy contains a continuum of agents indexed by i, with i distributed uniformly on the unit interval. Each agent uses his own labor to produce a differentiated good, sells his good to other agents, and buys goods to consume.

The utility of agent i in period t is given by

$$U_{it} = \frac{C_{it}^{1-\sigma} - 1}{1 - \sigma} - \frac{L_{it}^{1+\psi}}{1 + \psi}, \tag{1}$$

where L_{it} is the agent's labor supply in period t and C_{it} is a CES aggregator of the agent's consumption of different goods

$$C_{it} = \left[\int_0^1 \left(C_{it}^j \right)^{\frac{\gamma-1}{\gamma}} dj \right]^{\frac{\gamma}{\gamma-1}}, \tag{2}$$

where C_i^j is agent i's consumption of the good produced by agent j. The parameters measuring risk aversion (σ) and the marginal disutility of labor supply (ψ) are non-negative, while the elasticity of substitution between different goods (γ) is larger than one. The agent's production function is

$$Y_{it} = A_t L_{it}, \tag{3}$$

where Y is his level of output and A is an aggregate productivity shifter that follows an arbitrary stochastic process.

As is well-known (e.g., Woodford, 2002, chapter 2), utility maximization with this form of utility implies that the demand for each good depends on aggregate spending and the good's relative price. In logs, the demand function is

$$y_{it} = y_t - \gamma(p_{it} - p_t), \tag{4}$$

where y_i is the log of output by farmer i, y is the log of aggregate output, p_i is the log of the

price charged for good i, and p is the log of the price index for the aggregate consumption good C.

We assume that a government levies a proportional sales tax τ_t on all goods, which follows some stationary stochastic process. The tax revenues are used to finance equal lump sum rebates to all agents.

These assumptions determine the price for good i that maximizes farmer i's utility:

$$p_{it}^* = p_t + \alpha(y_t - y_t^N) + u_t, \tag{5}$$

where $\alpha = (\psi+\sigma)/(1+\gamma\psi)$, y^N is the natural level of output, and u reflects random variation in taxes. The natural level of output is defined as the level when prices are flexible (so $p_i = p_i^*$ for all i) and the tax rate is at its average level $\bar{\tau}$. It is given by:

$$y_t^N = \frac{(1+\psi)\log(A_t) - \log\left(\frac{\gamma}{(1-\gamma)(1-\bar{\tau})}\right)}{\psi + \sigma}. \tag{6}$$

Note that y^N varies with the productivity shock A. Finally, the shock u is given by $u = \log((1-\bar{\tau})/(1-\tau))/(1+\gamma\psi)$. This result is derived in the Appendix. The intuition is simple: as in standard macro models with imperfect competition (e.g., Romer, 2001, chapter 6), an agent's desired price increases one-for-one with the aggregate price level and depends positively on total spending in the economy. The price decreases with a positive shock to productivity and the natural rate (y^N) and increases with a rise in indirect taxation (u).

We introduce random taxation because it causes variation in farmer's "markups" – the gap between the desired price and the marginal cost of producing output. The key feature of these shocks is that they change the equilibrium level of output under flexible prices without changing the efficient level of output – that is, they cause variation in the level of distortions. In this way, the shocks differ from productivity shocks, which cause changes in equilibrium output that are efficient. Markup shocks are becoming a standard feature of models used to analyze monetary policy (e.g., Woodford, 2002; Clarida, Gali and Gertler, 2002; Steinsson, 2002). Shifts in markups can be justified different ways – for example, there can be shifts in the degree of collusion in an industry, or shifts in the aggressiveness of wage bargainers

in models with a labor market. We introduce variable taxation because it is the simplest way to generate time-varying markups. However, one can interpret our results as applying to markup variations arising from other sources as well.[1]

Finally, we again follow Woodford and assume the existence of complete financial markets. This implies perfect risk-sharing in consumption across all agents. The consumption level of all agents is the same (C_{it} is the same for all i), although the levels of labor supply and output can vary.

3.2 Price Setting and Sticky Information

Here we follow Mankiw and Reis (2002). Each period, a fraction λ of farmers, chosen randomly, receives complete information on the state of the economy. Prices are perfectly flexible in the sense that prices are adjusted by all farmers in each period. However, prices are set based only on the last information received by each farmer. We take a first-order approximation to optimal price-setting, which yields certainty-equivalent behavior. Thus, when farmer i sets his price in period t, he sets it equal to his expectation of the optimal price:

$$p_{it}^k = E_{t-k} p_{it}^*, \tag{7}$$

where $t - k$ is the last period when the farmer received information.

Again taking a first-order approximation, we measure the log of the price level, p, by the average of the individual prices p_i. Thus p is the average of prices set based on information at all past dates, weighted by the proportion of firms that last received information in each period. Mankiw and Reis aggregate the rule for individual price setting, (5), to derive the behavior of the price level in the absence of markup and productivity shocks. The Appendix to this paper extends their analysis to include these shocks, which is straightforward. The bottom line can be expressed as an equation either for the price level, or for aggregate

---

[1]In Ball and Mankiw (1995), for example, supply shocks arise from the interaction of menu costs and asymmetries in the distribution of relative shocks. The key effect of these shocks, however, is to change average markups, because firms with positive relative shocks respond more or less than firms with negative shocks.

inflation $\pi_t = p_t - p_{t-1}$:

$$p_t = \lambda \sum_{j=0}^{\infty}(1-\lambda)^j E_{t-j}\left[p_t + \alpha(y_t - y_t^N) + u_t\right] \tag{8}$$

$$\pi_t = \frac{\alpha\lambda}{1-\lambda}(y_t - y_t^N) + \frac{\lambda}{1-\lambda}u_t + \lambda \sum_{j=0}^{\infty}(1-\lambda)^j E_{t-1-j}\left[\pi_t + \alpha(\Delta y_t - \Delta y_t^N) + \Delta u_t)\right] \tag{9}$$

These equations show that the price level and its rate of change depend on current output and the shocks in the model, as well as past expectations of prices, output, and shocks at the various times farmers receive information. Not surprisingly, an adverse technology shock or a rise in the desired markup raises the current price level.

The second version of the equation is the sticky information Phillips curve that Mankiw and Reis propose as a replacement for the new Keynesian Phillips curve arising from the Calvo model of staggered price adjustment. Mankiw and Reis show that their model performs better at fitting the stylized facts about the output-inflation trade-off than does the new Keynesian equation. In particular, it captures the fact that a monetary contraction that gradually reduces inflation also reduces output, whereas the Calvo model implies that gradual disinflations are expansionary.

3.3 The Demand Side and Policy

The equation for the price level, (8), summarizes the supply side of the model. We close the model with the simplest possible demand side: the quantity equation. That is, we assume a cash-in-advance constraint that implies that nominal spending is proportional to the money supply, or

$$y_t = m_t - p_t, \tag{10}$$

where m is the log of the money supply.

We assume that policymakers control the money supply up to a white-noise control error.

In every period t, they choose a money-supply target $\hat{m}$, and

$$m_t = \hat{m}_t + e_t, \tag{11}$$

where e_t is a serially uncorrelated control error. $\hat{m}_t$ is set in period $t-1$, before the productivity and markup shocks for t are observed. These assumptions capture the facts that policy decisions affect the economy with a lag and that there are exogenous shifts in aggregate demand that policy cannot offset contemporaneously. We interpret the monetary control error as standing in for other kinds of demand shocks, such as shifts in fiscal policy or in confidence.

In principle, one can solve for the optimal money-supply rule in this model. A different way of analyzing the problem is more appealing, however. Using equations (8)-(11), the Appendix derives a reduced form for the price level p_t as a linear function of the target money supply, $\hat{m}_t$, variables determined before t, and shocks at t (the monetary control error and the unexpected component of the productivity and markup shocks). Since the shocks are unforecastable, this implies a linear relation between $\hat{m}_t$ and the expectation of the price level $E_{t-1}p_t$, given the predetermined variables. Thus the choice at $t-1$ of a target money supply is implicitly a choice of the expected price level at t, that is, a price-level target. We can therefore interpret $E_{t-1}p_t$ as the policy instrument, and solve for the optimal rule for this target as a function of all information through period $t-1$. One can use the result in the Appendix to find the money-supply rule that implements the optimal price-targeting rule, but we will not emphasize this form of policy.

Note that our model of the demand side is simpler than those in most recent work on policy rules. It is common to assume that policymakers control an interest rate, and to relate the interest rate to output through an IS or aggregate-spending equation. If we added such an equation to our model, we could still interpret the expected price level as the policy instrument; a rule for the price target would implicitly define an interest rate rule that implements it. We choose not to introduce an IS curve because there is currently no consensus about the right specification.[2]

───────────────

[2]The debate about the IS curve parallels the debate about the Phillips curve, with some authors ad-

4 WELFARE

The next step is to consider welfare in our model, which will provide the foundation for the policy analysis in the next section.

4.1 Woodford's Approximation

We define welfare in a period as the average level of utility across all farmers. We assume the discount factor approaches one, so policymakers seek to minimize the unconditional expectation of welfare.[3] Following similar steps to Woodford (2002), the Appendix shows that a second-order approximation of farmers' utility functions is

$$-Var(y_t - y_t^N) - \chi E[Var_i(y_{it} - y_t)] + t.i.p., \tag{12}$$

where $\chi = (\psi + \gamma^{-1})/(\psi + \sigma)$ and t.i.p stands for terms independent of policy.

The two terms in this expression are intuitive. The first is the variability of output around the natural level. (Note that the natural level of output is not efficient, because of the distortion caused by the average markup. However, variability around y^N differs from variability around efficient output only by a constant independent of policy.) The second term captures the cross-sectional variability of output across different firms. Variability at the firm level is inefficient because it creates variability in labor supply around the efficient level. Throughout our analysis we define the "optimal" monetary policy as the one that minimizes the expression in equation (12).

Another way of writing the objective function is useful. The demand equation (4) implies that a farmer's output y_i is proportional to his relative price $p_i - p$. Thus the variance of y_i

---

vocating forward-looking specifications and others advocating backward-looking specifications. Gabaix and Laibson's (2001) work on lagged consumption responses to the stock market suggests that applying the sticky-information assumption to spending decisions and thus the IS curve may be a fruitful line of pursuit. In the model we consider here, however, the stickiness of information affects only the yeoman farmer's price setting decision.

[3]While the assumption that the discount factor approaches one simplifies the exposition, most of our results can be obtained with a discount factor smaller than one.

is proportional to the variance of $p_i - p$. Thus we can write welfare as

$$-Var(y_t - y_t^N) - \omega E[Var_i(p_{it} - p_t)] + t.i.p., \tag{13}$$

where $\omega = \chi\gamma^2$. That is, welfare depends on the variance of the output gap and the cross-sectional variance of relative prices.

4.2 Welfare with Mankiw-Reis Price-Setting

To analyze the model, we need to determine how the terms in the loss function depend on aggregate variables. Here is the first point at which our results for the sticky-information model depart from those for the Calvo model. In the Calvo model, Woodford (2002) shows that the cross-sectional variability of prices is determined by current and lagged values of inflation squared. In contrast, the following result gives an expression for cross-sectional price variability in terms of aggregate variables in our model:

Lemma 1

$$Var_i(p_{it} - p_t) = \sum_{j=1}^{\infty} \eta_j \left(p_t - E_{t-j}(p_t)\right)^2, \tag{14}$$

where:

$$\eta_j \equiv \left(\frac{\lambda(1-\lambda)^j}{(1-(1-\lambda)^j)(1-(1-\lambda)^{j+1})} \right)$$

Thus the variance of relative prices depends on the squared deviations of the price level from the levels expected at all past dates. The weights η_j decline as j increases.

This result is derived in the Appendix. The intuition is straightforward. Surprises in the aggregate price level create micro price variability because prices for a given period are set based on information at various times in the past. In period t, surprises since $t - j$ affect the prices of agents who have received information since then, but not the prices of other agents.

The bottom line, therefore, is that welfare is approximated by a weighted average of equation (14) and the variance of $(y - y^N)$. In the analysis below, we experiment with various weights on these two terms. Fortunately, the results are not very sensitive to the weights we assume. In particular, most of our analysis holds even if the weight on (14) is zero, so that the loss function is $Var(y - y^N)$, the simple loss function assumed in much work on optimal monetary policy.

5 OPTIMAL POLICY WITH DEMAND AND PRODUCTIVITY SHOCKS

We now turn to our central question: what is the optimal policy rule in the model? In this section, we simplify things by eliminating the markup shock u from the model. In this case, the model has two shocks, the productivity shock and the demand shock (i.e., monetary control error). The optimal policy rule turns out to be very simple. The next section discusses the complications arising from the introduction of markup shocks.

5.1 Results

Recall that we can describe policy as a rule for the price target, $E_{t-1}p_t$. In our model,

Proposition 1 *With only productivity and demand shocks, the set of optimal policies is the set of rules $E_{t-1}p_t = K_t$, where K_t follows a deterministic path known at the beginning of time.*

To interpret this result, note first that one optimal policy is the one in which K_t is a constant. In this case, policy targets a fixed price level, the purest form of "price level targeting." According to the Proposition, however, it is optimal also to allow the target to change over time, as long as the changes are deterministic. The target price level could grow at a constant rate or follow some more complicated (but predictable) path.

Proposition 1 is proved in the Appendix. We note here that it is not surprising that any deterministic path for the price level is as good as any other. In the model, the non-neutrality

14

of monetary policy arises from surprises that it creates that are not observed immediately by all price setters. A deterministic price path creates no surprises and thus has no real effects.

Having shown what policies are optimal, it is easy to point out some that are not optimal:

Corollary 1 *Inflation targeting – a policy that sets $E_{t-1}\pi_t$ equal to a constant – is suboptimal. So is nominal income targeting – a policy that sets $E_{t-1}(p_t + y_t)$ equal to a constant.*

This result follows from the facts that only price-level targeting is optimal, and that inflation and nominal-income targeting differ from price-level targeting. It is well-known that inflation targeting allows "base drift": it means that the target for the price level is adjusted one-for-one with past shocks to the price level. Thus the price target does not follow a deterministic path, as required by Proposition 1. Nominal income targeting would be equivalent to price-level targeting if there were only demand shocks. However, in the case of productivity shocks, nominal income targeting implies that the shocks have persistent effects on both the price level and output, so the expected price level is not deterministic. It is optimal for output alone to respond to a productivity shock, as this keeps output in line with its natural level y_t^N.

5.2 Discussion

These results are significant in light of the current practice of policymaking. There is a strong trend in the world towards inflation targeting, with dozens of countries adopting some version of this policy since the early 1990s. Many observers suggest that U.S. policy is close to inflation targeting, although the target is not explicit. In our model, however, this popular policy is sub-optimal. Welfare is higher if policymakers target the price level rather than the inflation rate - that is, if they eliminate base drift in prices.

To understand this result, it is useful to compare the dynamic effects of shocks under price-level and inflation targeting. We focus here on demand shocks; the story for productivity shocks is similar, since demand and productivity shocks have symmetric effects on the output gap $y - y^N$. Figure 1 shows the responses of the aggregate price level and aggregate output to a unit demand shock under the two policies, assuming $\alpha = 0.1$ and $\lambda = 0.25$, the baseline values in Mankiw and Reis. The Figure also shows the response of the

cross-sectional variance of prices. In the period when a shock occurs, all the effects are the same under price-level and inflation targeting, reflecting the fact that policy cannot respond contemporaneously to shocks. After the initial period, price targeting implies that all the effects disappear completely. In contrast, the effects on both output and relative prices die out slowly under inflation targeting. This persistence in the effects of shocks increases both components of the social loss function, the variance of the output gap and the variance of relative prices.

To understand these results, suppose the economy starts in a steady state in which all prices are zero, the output gap is zero, and this situation is expected to continue forever. A demand shock occurs in period zero; it inevitably raises output at zero and it raises the price level as well because firms who receive information in that period adjust their current prices. But these effects disappear in subsequent periods under price-level targeting. In these periods, some prices are set by firms who have not observed the shock, and therefore are fixed at zero. Price-level targeting means the monetary authority must induce informed firms to set prices of zero as well, so the aggregate price level returns to zero. In this case relative-price variability is eliminated, since informed and uninformed firms set the same prices. In addition, to induce the informed firms to choose the same price as the uninformed, monetary policy must produce an output gap of zero, since only a zero output gap implies a desired relative price of zero. Thus price targeting eliminates the output effects of the shock after period zero as well as the relative-price effects.

In contrast, inflation targeting means the price level must remain at a positive level after the shock raises it in period zero. Because the uninformed still choose prices of zero in subsequent periods, the monetary authority must induce the informed to choose positive prices to keep the aggregate price level positive. This creates variability across the prices of informed and uninformed firms, and it requires a positive output gap to induce the informed to set positive relative prices.

Underlying these results is the fact that demand shocks affect real variables in the sticky information model by creating surprises in the price level. These surprises create both relative price variability (see equation (14)) and output movements (note from (8) that the output gap is zero if the price level equals the level expected in all previous periods). To

16

stabilize the economy, the monetary authority needs to minimize price surprises. It does so by reversing deviations of the price level from its expected path as quickly as possible. It should not exacerbate uncertainty about the price level by allowing base drift.

6 OPTIMAL POLICY WITH MARKUP SHOCKS

The optimal policy is more complicated in the presence of markup shocks. In this case, strict price level targeting is no longer optimal. Instead, policymakers must make special accommodations for markup shocks, the nature of which depends on the serial correlation of the shocks.

We assume the markup shock follows an arbitrary stationary process with MA representation $u_t = \sum_{j=0}^{\infty} \rho_j \varepsilon_{t-j}$, where ε_t is a white-noise innovation. Given this notation, we have

Proposition 2 *With productivity, demand, and markup shocks, the set of optimal policies is the set of rules:*

$$E_{t-1}p_t = K_t + \sum_{j=1}^{\infty} \phi_j \varepsilon_{t-j}, \tag{15}$$

where

$$\phi_j \equiv \frac{\rho_j}{\alpha^2 \omega + \frac{(1-\lambda)^{j+1}}{1-(1-\lambda)^{j+1}}},$$

and K_t is again deterministic.

This Proposition is proved in the Appendix.

To interpret Proposition 2, consider first the case of white-noise markup shocks. In this case all the ρ_j's are zero, so all the ϕ_j's are zero: the optimal policy is strict price level targeting as before. If the markup shock is serially correlated, however, then the price target is adjusted in response to past shocks. If all the ρ_j's are positive, then an increase in the markup raises the price target for awhile. However, as long as the markup is stationary, the price level is also stationary: it eventually returns to the target path given by K_t.

17

Figure 2 illustrates this result for particular parameterizations of the model. The Figure shows how the price level responds to a unit shock to the innovation ε. In this simulation, again we set $\alpha = 0.1$ and $\lambda = 0.25$. We assume the markup shocks follow an AR(1) process with a coefficient of $\bar{\rho}$ so $\rho_j = \bar{\rho}^j$. We experiment with two values for $\bar{\rho}$, 0.8 and 0.4. Finally, we set the parameter ω equal to one, which means that output variance and relative-price variance have the same weight in the social loss function.

In Figure 2, the price level jumps up at the time of the markup shock, which is inevitable since policy cannot respond contemporaneously. In contrast to the cases of demand and productivity shocks, the price level remains high after the initial period; it either follows a hump-shaped pattern (for $\bar{\rho} = 0.8$) or falls monotonically (for $\bar{\rho} = 0.4$). In either case, the price level eventually converges back to its long-run target.

The optimal response to markup shocks reflects a trade-off between the goals of stabilizing relative prices and stabilizing output. As with demand and productivity shocks, relative-price variance is minimized by a strict price level target, as this causes informed firms to set prices at the same level as uninformed firms. Such a policy requires a large fall in output, however: a higher markup means firms desire higher relative prices for a given output level, so low output is needed to induce firms to choose a zero relative price. Given this trade-off, optimal policy allows output to fall somewhat but not enough to stabilize relative prices fully. It induces informed firms to set positive prices as long as the shock persists.

Our model implies that the optimal response to markup shocks depends on the serial correlation of the shocks. Unfortunately, it is difficult to calibrate this serial correlation. Like other authors, we would like to interpret markup shifts as metaphors for the "supply shocks" that influence real-world inflation, but there is no consensus about the nature of these shocks. If we interpret the model literally, so markup shocks arise from tax changes, it is natural to interpret these changes as highly persistent. Similarly, if markup shocks are interpreted as changes in union aggressiveness, it is plausible to assume persistence in union attitudes. However, some kinds of real-world supply shocks may have more transitory effects on desired prices. We need a better understanding of supply shocks to make progress on designing optimal policies.

Our model suggests that policymakers should treat price-level movements in a special

way if they arise from certain kinds of supply shocks. In this respect, the model's prescriptions match policymaking in many countries. It is common for inflation targeters to allow deviations from their normal targets when supply shocks occur. Sometimes, this is accomplished by explicit "caveats" to policy, as in New Zealand. More often, policymakers allow temporary deviations from the target by focusing on an "underlying" or "core" measure of inflation that strips out the effects of supply shocks. Our analysis suggests that policymakers are on the right track in making exceptions to their policies for supply shocks – although the basic policy target should be the price level, not inflation as in current practice.

7 A SIMPLE DESCRIPTION OF OPTIMAL POLICY

So far, we have described optimal policy as a function of the exogenous driving variables. There is, however, a simpler description of optimal policy in terms of the endogenous variables:

Proposition 3 *For an economy that experiences shocks to demand, productivity, and markups, optimal monetary policy can be described as:*

$$E_{t-1}p_t = K_t - \frac{1}{\alpha\omega}E_{t-1}(y_t - y_t^N).$$ (16)

That is, monetary policy commits to a deterministic path for the price level, represented here by K_t, and allows the expected price level to deviate from the target path only if output is expected to deviate from its natural rate.[4]

---

[4]Clarida, Gali, and Gertler (1999, p. 1704) derive an equation for optimal policy that closely resembles this proposition in an analysis based on the Calvo model of price adjustment. This similarity of optimal policies under Calvo and Mankiw-Reis price setting is not robust, however. The result of Clarida, Gali and Gertler is based on the assumption that the monetary authority can set the money supply based on contemporaneous information. Yet if policymakers can affect the economy only with a lag (that is, if there are control errors), as we have assumed, then optimal monetary policy in the Calvo model can no longer be described by such a simple rule. In the next section, we explore more fully the similarities and differences between the two models.

This proposition encompasses our previous results. In the case of shocks to demand or productivity, discussed in Section 4, optimal policy has the property that $E_{t-1}(y_t - y_t^N) = 0$. Thus, Proposition 3 includes Proposition 1 as a special case. In the case of shocks to the markup, the expected price level varies as a complicated function of the shocks, as shown in Proposition 2, but it does so in a way that precisely parallels the expected output gap. This can be seen in Figure 2, where the paths of output and the price level after period zero are reflections of each other.

One remarkable fact about Proposition 3 is that it closely resembles the monetary policy rule proposed by Hall (1984), who dubbed it an "elastic price standard." Hall does not present a formal model to back up his proposal, but his informal and intuitive arguments are consistent with the model we have explored here. One difference is that Hall advocates a constant target (K_t), whereas in our model a rising target for the price level would serve equally well. In either case, the expected output gap determines how much the central bank can allow the price level to deviate from its target.

8   COMPARISON TO THE NEW KEYNESIAN PHILLIPS CURVE

We have argued that the sticky-information Phillips curve of Mankiw and Reis is a better tool for monetary-policy analysis than the popular new Keynesian Phillips curve, because it has more realistic implications about the interactions of output and inflation. But does the choice of a Phillips curve make a difference for our results about optimal policy? To address this issue, we now reconsider our policy problem using the new Keynesian Phillips curve - the relation derived from Calvo's (1983) model of staggered price adjustment with full information. As is well known, this Phillips curve (with no discounting) is

$$\pi_t = E_t\pi_{t+1} + \kappa(y_t - y_t^N) + u_t', \quad \kappa > 0, \tag{17}$$

where u_t' is proportional to the markup shock u_t. As discussed by Mankiw and Reis (2002), a critical difference between this equation and the Phillips curve arising from the sticky-

information model is the timing of expectations. Rather than past expectations of current inflation, it is the current expectation of next period's inflation that enters the new Keynesian specification. This feature of the model is mainly responsible for its counterfactual predictions.

8.1 Implications for the Welfare Function

Note first that the choice of a Phillips curve affects the determinants of aggregate welfare. As shown in Section 3, the canonical model of monopolistic competition yields an approximate loss function with two terms, the variance of the output gap and the variance of relative prices. This result is robust, but the determinants of the variance of relative prices depends on the specification of price adjustment. In the Mankiw-Reis model, the variance of relative prices depends on the variance of unexpected changes in the price level (see equation (14)). In the Calvo model, it is determined by current and lagged values of inflation squared. Thus the two models produce loss functions that have one term in common, the variance of the output gap, and one term that differs across models. Thus one would expect the models to deliver different results about optimal policies. However, the main differences discussed below do not depend on the difference in loss functions. The differences in results arise even when the weight on the variance of relative prices in the loss function is zero, so that the loss function is simply the variance of the output gap in both models.

8.2 Implications for Base Drift

We now consider optimal policy in a model that is identical to ours except that the Calvo sticky-price assumption replaces the Mankiw-Reis sticky-information assumption, so that the Phillips curve becomes equation (17) above. The following result shows a key difference between the normative implications of the two models:

Proposition 4 *With Calvo price adjustment, the optimal policy rule when there are demand or productivity shocks is: $E_{t-1}\pi_t = \delta\pi_{t-1}$, where $0 < \delta < 1$.*

That is, the Calvo model supports flexible inflation targeting rather than price level targeting.

Put differently, allowing base drift in the price level in response to demand and productivity shocks is optimal with Calvo price setting, while it is not with Mankiw-Reis price setting. The proof of the proposition is in the Appendix.

Proposition 4 implies not only a unit root in the price level but also persistence in inflation. In particular, the inflation rate is AR(1) with parameter δ. We believe that the following intuition explains the optimality of inflation persistence in the Calvo model: When the economy experiences a positive demand shock, output and the price level both rise. If the monetary authority responds to this shock by committing to higher inflation in the future, then price setters will respond to the shock by raising their prices more aggressively. This more aggressive price response implies a smaller initial output response, keeping the economy closer to its natural rate. In the long run, however, the only effect of inflation is relative price variability, so optimal policy always returns to zero inflation.

As in our main model, the nature of optimal policy is more complicated when there are persistent markup shocks. Just as in our model, optimal policy involves a gradual adjustment of the price level in response to a shock. Yet now the price level is non-stationary: it remains forever higher after a shock that raises prices.

The result that inflation rather than price-level targeting is optimal under Calvo pricing may appear surprising in light of the previous literature. A number of authors study models that include the Calvo Phillips curve and conclude that a price level target is optimal. The discrepancy between our results and previous ones are explained by differences in the policy problem being solved. There are two main differences.

First, some previous authors, such as King and Wolman (1999), study models in which it is feasible to stabilize the price level perfectly. That is, there is a perfect relation between a policy instrument and the current price level; there is no such thing as a demand shock that policy cannot offset contemporaneously, as with our assumption of monetary control errors. In the King-Wolman model, the policy of keeping the price level absolutely constant is optimal given the types of shocks they assume. A perfectly constant price level can be interpreted as implementing either a constant price-level target or a zero inflation target. That is, there is no meaningful distinction between price and inflation targeting if there are

no shocks that raise the possibility of base drift.[5]

Vestin (1999) compares price and inflation targeting in versions of the Calvo model in which these policies are different, and concludes that price targeting is optimal. The problem he solves, however, is that of the optimal policy rule under "discretion." That is, he asks whether it is optimal to appoint a policymaker whose objective function is the variance of prices or the variance of inflation, given that the policymaker cannot precommit to his preferred policy because it is not dynamically consistent. In contrast, we study optimal policy under commitment: we assume that policymakers can implement any rule they want and ask which rule maximizes social welfare.[6]

8.3   Implications for Trend Inflation

A final difference between our results and those for the Calvo model concern the welfare consequences of trend inflation. As shown above, trend inflation is neutral in our model, and so optimal policy can take the form of an increasing price level target rather than a constant target. In contrast, the optimal inflation target in the Calvo version of the model must be zero. That is, moving from Calvo to Mankiw-Reis price setting means introducing the possibility of trend inflation as well as eliminating base drift in the price level. Here, the difference in loss functions in the two models is crucial. In the Calvo model, welfare depends on the variance of inflation around zero rather than around some mean level, so that zero inflation is optimal. This reflects the fact that even steady inflation causes inefficient relative-price variability under Calvo's assumption of sticky prices.

9   CONCLUSION

According to the model developed here, optimal monetary policy can be described as flexible targeting of the price level. The central bank should announce a target path of the price

---

[5]By contrast, Goodfriend and King (1997) briefly consider the issue of imperfect control in a model with staggered price setting. Their analysis confirms our conclusion that base drift in the price level is optimal in this type of model.

[6]Like Vestin, Svensson's (1999) well-known paper on inflation versus price level targeting emphasizes the case of discretion rather than commitment, but with a different specification for the Phillips curve.

level and then commit itself to returning to this path in response to shocks. In response to shocks to demand or productivity, it should return the price level as quickly as possible to its target. In response to shocks to markups, which here reflect a type of supply shock, the central bank should return the price level to its target more gradually over time. Described differently, optimal policy allows the price level to deviate from its target only if output is expected to deviate from its natural rate.

The optimality of targeting the price level rather than the inflation rate is a common theme in the recent literature on monetary policy. For example, Hall (1984) and Hall and Mankiw (1994) argue that a price level target would aid personal financial planning by making the cost of living far in the future more predictable. Svensson (1999) and Vestin (1999) argue for price level targeting on the grounds that it would help solve some of the time-inconsistency problems associated with discretionary monetary policy. Cover and Pecorino (2001) claim that a price target is stabilizing because any inflationary shock automatically causes a decrease in expected inflation and thus an increase in the real interest rate. The analysis in this paper is very different. Our households have no money illusion in financial planning, our monetary policymaker can commit to a policy rule, and we omit the effect of expected inflation on aggregate demand. Nonetheless, the bottom line for policy is the same: central banks should target the price level, not the inflation rate.

These results leave open an intriguing question: Why have central banks adopted inflation targeting if price level targeting has all these desirable properties? One possible answer is that central bankers know something about the world that is missing in these theories of optimal policy. But another possibility is that central bankers have been misled by assuming, incorrectly, that some features of the recent monetary regime are structural, while in fact these features would change if policy did.

Price dynamics are the natural place to look for such a mistaken assumption. The recent data are well described by a backward-looking Phillips curve. If this equation for price dynamics were invariant to policy, then inflation targeting would be optimal, and price level targeting would be unattractive. (Ball, 1999). But a radical change in the monetary policy rule, such as a target for the price level, would most likely alter the reduced-form equation for the Phillips curve. Whether the particular behavioral model we have examined in this paper

correctly captures the shift that would occur is open to debate. But the results presented here suggest the issue is well worth pursuing, for the implications for monetary policy could not be more profound.

Appendix

In this appendix, we derive the many results alluded to in the main text.

The model

Our notation and approach builds on Woodford (2002). Each agent's problem is to maximize the expected value of the sum of discounted utility:

$$E_0 \left((1 - \beta) \sum_{t=0}^{\infty} \beta^t U_t \right) \tag{A1}$$

where the utility function U_t is defined in (1) and β is the discount factor. (Later we will focus on the case when β goes to 1). This must be subject to the budget constraint:

$$\int_0^1 C_{it}^j P_{jt} dj + M_{it} + E_t \left[Q_{t+1} V_{it+1} \right] = (1 - \tau_t) P_{it} Y_{it} + T_t + M_{it-1} + V_{it}.$$

Agents spend their wealth purchasing the different varieties C_i^j for consumption, and acquiring money balances M_t. They receive income from selling their product of specialization Y_i for the price P_i, paying a share τ of sales revenue to the government. Tax revenues are instantly returned to consumers by lump-sum transfers T, which also include monetary injections into the economy. M_{t-1} units of money are brought forward from the last period. The assumption of complete financial markets means agents can insure themselves against idiosyncratic uncertainty. We let V_t stand for the total nominal value of the portfolio of financial assets brought into period t. In order to acquire a bond portfolio at t which returns V_{t+1} in the next period, the agent must expend an amount given by the conditional expectation $E_t \left[Q_{t+1} V_{it+1} \right]$, where Q_{t+1} is the stochastic discount factor between t and $t + 1$.

The agent faces two further constraints: the production function in (3); and a cash in advance constraint: $P_t C_{it} \leq M_{it}$.

Optimization with respect to C_i^j leads to the demand function in (4), where the aggregate price index is defined as:

$$P_t = \left[\int_0^1 P_{it}^{1-\gamma} di \right]^{\frac{1}{1-\gamma}}. \tag{A2}$$

Since agents are all ex ante identical and there are complete financial markets, they will all have the same ex post wealth. Thus, they will make the same consumption decisions and the common level of consumption is denoted by C_t. An Euler equation links the stochastic discount factor to consumption. We assume that $E_t(Q_{t+1})$ is always strictly smaller than one, which can be ensured by for example having a sufficiently high deterministic growth in output. This ensures that money has a lower return than a riskless bond, which in turn leads to the cash in advance constraint holding as an equality as another optimality condition.

Finally, agents decide what price they desire to charge P_{it}^* given the demand function in (4). The first-order condition is:

$$\frac{P_{it}^*}{P_t} = \frac{\gamma}{(\gamma-1)(1-\tau_t)} \frac{C_t^\sigma L_{it}^\psi}{A_t}. \tag{A3}$$

The first fraction on the right hand side is the markup, given by the standard monopoly markup $\gamma/(\gamma-1)$ divided by the tax wedge $(1-\tau)$. The second fraction is the real marginal cost of production, given by the marginal rate of substitution between consumption and labor (the shadow cost of labor supply) $C^\sigma L_i^\psi$, divided by the marginal product of labor A.

Having established all the optimality conditions, we determine the equilibrium by imposing the market clearing conditions: $C = Y$, $C_i = Y_i$. Substituting out for L_{it} in (A3) using the production function (3), and then using the demand function (4) to substitute out for Y_i, we obtain (after taking logs):

$$p_{it}^* = p_t + \alpha y_t - \frac{(1+\psi)\log(A_t)}{1+\gamma\psi} - \frac{\log(1-\tau_t)}{1+\gamma\psi} + \frac{\log(\gamma/(1-\gamma))}{1+\gamma\psi} \tag{A4}$$

where $\alpha = (\psi+\sigma)/(1+\gamma\psi)$.

The natural level of output is defined in the text by the conditions: (i) $p_i = p_i^*$ for all i (prices are flexible) and (ii) $\tau = \bar{\tau}$ (no markup shocks). Since the right hand side of (A4) is the same for all i, $p_i = p$. These conditions and (A4) lead to the natural level of output in (6). Using this result to substitute A out from (A4) gives equation (5).

We work with a linearized version of the price index in (A2). A first-order Taylor approximation around the point where all products' prices are the same, gives the approximate

price index, as the simple arithmetic average: $p_t = \int_0^1 p_{it} di$. Given the arrival rate of λ of new information for each firm, there is a share $\lambda(1 - \lambda)^k$ of firms that last adjusted their price plan k periods ago. The price index therefore becomes:

$$p_t = \lambda \sum_{k=0}^{\infty} (1 - \lambda)^k E_{t-k} \left(p_{it}^* \right)$$

Using the expression for p_{it}^* in (5), leads to equation (8) in the text.

Equation (8) can be re-written as:

$$p_t = \lambda(p_t + \alpha(y_t - y_t^N) + u_t) + \lambda \sum_{j=0}^{\infty} (1 - \lambda)^{j+1} E_{t-1-j} \left[p_t + \alpha(y_t - y_t^N) + u_t \right]. \qquad (A5)$$

Subtracting the analogue of equation (8) for p_{t-1} from this expression above, we obtain an expression for the inflation rate:

$$\begin{aligned} \pi_t &= \lambda(p_t + \alpha(y_t - y_t^N) + u_t) + \lambda \sum_{j=0}^{\infty} (1 - \lambda)^j E_{t-1-j} \left[\Delta p_t + \alpha(\Delta y_t - \Delta y_t^N) + \Delta u_t \right] \\ &\quad - \lambda^2 \sum_{j=0}^{\infty} (1 - \lambda)^j E_{t-1-j} \left[p_t + \alpha(y_t - y_t^N) + u_t \right]. \end{aligned}$$

Now, equation (A5) can be rearranged to show that:

$$\lambda p_t - \frac{\lambda^2}{1 - \lambda}(\alpha(y_t - y_t^N) + u_t) = \lambda^2 \sum_{j=0}^{\infty} (1 - \lambda)^j E_{t-1-j} \left[p_t + \alpha(y_t - y_t^N) + u_t \right]$$

Using this to replace for the last term in the previous expression, and rearranging gives the Phillips curve in equation (9).

Since from optimality the cash in advance constraint holds as an equality, taking logs and using the market clearing condition to replace c by y, we obtain equation (10).

Replacing for y in the equation for the price level (8) using the demand side equation

(10), we obtain an equation for the price level:

$$
\begin{aligned}
(1 - \lambda(1 - \alpha))\, p_t \;=\;& \alpha\lambda(2 - \lambda)\hat{m}_t + \lambda(\alpha e_t - \alpha y_t^N + u_t) \\
&+ \lambda(1 - \lambda)\left[(1 - \alpha)E_{t-1}p_t + E_{t-1}u_t - \alpha E_{t-1}y_t^N\right] \\
&+ \lambda\sum_{j=2}^{\infty}(1 - \lambda)^j E_{t-j}\left[(1 - \alpha)p_t + \alpha(\hat{m}_t - y_t^N) + u_t\right].
\end{aligned}
$$

Taking expectations of this expression at time $t - 1$ gives:

$$
\begin{aligned}
(1 - \lambda(2 - \lambda)(1 - \alpha))\, E_{t-1}p_t \;=\;& \alpha\lambda(2 - \lambda)\hat{m}_t + \lambda(2 - \lambda)(E_{t-1}u_t - \alpha E_{t-1}y_t^N) \\
&+ \lambda\sum_{j=2}^{\infty}(1 - \lambda)^j E_{t-j}\left[(1 - \alpha)p_t + \alpha(\hat{m}_t - y_t^N) + u_t\right].
\end{aligned}
$$

The policymaker controls the first term in the right hand side, and observes all the other terms in the right hand side. Therefore, it can set $E_{t-1}p_t$ at whatever level it desires. Conversely, for any given value of $E_{t-1}p_t$, the equation above determines the corresponding $\hat{m}_t$.

The welfare approximation

We approximate the utility function around the equilibrium point when there are no real disturbances, so we set all shocks at their means. As Woodford (2002) emphasizes, it is important for the accuracy of the approximation that this equilibrium is close to being efficient. One assumption that ensures it is $E(\frac{\gamma}{\gamma-1}(1 - \tau)) = 1$, so the expected markup is 1. The point of linearization is then:

$$
\bar{y}_t = \frac{(1 + \psi)\bar{a}_t}{\psi + \sigma},
$$

where $\bar{a}l = \log(\bar{A}_t)$.

We approximate:

$$
U_t = \frac{e^{(1-\sigma)y_t}}{1 - \sigma} - \int_0^1 \frac{e^{(1+\psi)(y_{it} - a_t)}}{1 + \psi}\, di,
$$

where the production function was used to replace for L_{it} and market clearing to replace consumption by output. A second order Taylor approximation gives:

$$Ul \approx e^{(1-\sigma)\bar{y}_t}\left(\hat{y}_t + \frac{1-\sigma}{2}\hat{y}_t^2\right) - e^{(1+\psi)(\bar{y}_t-\bar{a}_t)}\int_0^1\left(\hat{y}_{it} + \frac{1+\psi}{2}\hat{y}_{it}^2 - (1+\psi)\hat{y}_{it}\hat{a}_t\right)di + t.i.p,$$

where $\hat{y}_t = y_t - \bar{y}_t$, $\hat{a}_t = a_t - \bar{a}_t$, $\hat{y}_{it} = y_{it} - \bar{y}_t$ and additive terms that are independent of policy are collected in $t.i.p$. Defining the cross-sectional mean as $E_i(\hat{y}_{it}) = \int\hat{y}_{it}di$ and the cross-sectional variance as $Var_i(\hat{y}_{it}) = E_i(\hat{y}_{it}^2) - E_i(\hat{y}_{it})^2$, we can re-write this expression as:

$$e^{(1-\sigma)\bar{y}_t}\left(\hat{y}_t + \frac{1-\sigma}{2}\hat{y}_t^2\right) - e^{(1+\psi)(\bar{y}_t-\bar{a}_t)}\left(E_i(\hat{y}_{it}) + \frac{1+\psi}{2}\left(Var_i(\hat{y}_{it}) + E_i(\hat{y}_{it})^2\right) - (1+\psi)\hat{a}_t E_i(\hat{y}_{it})\right) + t.i.p.$$

A second order approximation of the CES aggregator in equation (2), around $\bar{y}_t$ yields:

$$\hat{y}_t \approx E_i(\hat{y}_{it}) - \frac{1-\gamma^{-1}}{2}Var_i(\hat{y}_{it}).$$

Using this expression to substitute for $E_i(\hat{y}_{it})$ in the previous expression, and dropping third or higher stochastic order terms, we obtain:

$$U_t \approx e^{(1-\sigma)\bar{y}_t}\left(\hat{y}_t + \frac{1-\sigma}{2}\hat{y}_t^2\right) - e^{(1+\psi)(\bar{y}_t-\bar{a}_t)}\left(\hat{y}_t + \frac{1+\psi}{2}\hat{y}_t^2 + \frac{\gamma^{-1}+\psi}{2}Var_i(\hat{y}_{it}) - (1+\psi)\hat{a}_t\hat{y}_t\right) + t.i.p.$$

Now, realize that from the definition of $\bar{y}_t$: $(1-\sigma)\bar{y}_t = (1+\psi)(\bar{y}_t - \bar{a}_t)$. Combining the two terms in brackets we obtain:

$$U_t \approx -e^{(1-\sigma)\bar{y}_t}\frac{\sigma+\psi}{2}\left(\hat{y}_t^2 - 2\frac{(1+\psi)}{\sigma+\psi}\hat{a}_t\hat{y}_t + \frac{\gamma^{-1}+\psi}{\sigma+\psi}Var_i(\hat{y}_{it})\right) + t.i.p.$$

Next, recall the definition of the natural rate in (6) to replace for the second term inside the brackets, and add a term in $\left(\hat{y}_t^N\right)^2$ to the expression which is valid since this is a $t.i.p.$

Dropping the proportionality factor, which is beyond the control of the policymaker, gives:

$$-U_t \approx (\hat{y}_t - \hat{y}_t^N)^2 + \frac{\gamma^{-1} + \psi}{\sigma + \psi} Var_i(\hat{y}_{it} - \hat{y}_t) + t.i.p.$$

We include $\hat{y}_t$ inside the cross-sectional variance since this term does not vary with i. Next, note that we can drop all the hats from the output variables as they all share the same point of approximation $\bar{y}_t$. Finally, note that given our assumption that the expected markup is one, to a first-order approximation $E(y_t) \approx E(y_t^N)$, so we can re-write the expression as:

$$-Ul \approx (y_t - y_t^N - E(y_t - y_t^N))^2 + \frac{\gamma^{-1} + \psi}{\sigma + \psi} Var_i(y_{it} - y_t) + t.i.p.$$

Using this expression, we can take the unconditional expectation of (A1) to obtain equation (12) in the text. If the discount factor β approaches one, the expectation of the sum of discounted utility over time equals this expression.

Proof of the lemma

We start the proof by introducing some useful notation. We use Λ^k to denote the fraction of price setters that has updated their information by k periods after an event:

$$\Lambda^k = \lambda \sum_{j=0}^{k} (1 - \lambda)^j. \tag{A6}$$

We denote the deviation of the price level from its expected value j periods ago by: $p_{t|j} = p_t - E_{t-j}(p_t)$. Finally, we will denote the deviation of the optimal price set by group j of price-setters relative to the overall price index in the economy by: $p_{t|j}^* = E_{t-j}(p_t^*) - p_t$.

We now prove the main lemma through a sequence of other lemmas.

Lemma A.1:

$$p_{t|i-1}^* = \frac{p_{t|i} - p_{t|i-1}}{\Lambda^{i-1}} + p_{t|i}^*$$

Proof. From the definition of $p_{t|i}$:

$$
\begin{aligned}
p_{t|i} &= p_t - E_{t-i}(p_t) \\
&= p_{t|i-1} + E_{t-i+1}(p_t) - E_{t-i}(p_t).
\end{aligned}
\tag{A7}
$$

From the supply equation in (8) and using the law of iterated expectations:

$$
E_{t-i}(p_t) = \Lambda^{i-1} E_{t-i}(p_t^*) + \lambda \sum_{j=i}^{\infty} (1-\lambda)^j E_{t-j}(p_t^*).
\tag{A8}
$$

Using this and its equivalent for i-1, we obtain:

$$
\begin{aligned}
E_{t-i}(p_t) - E_{t-i+1}(p_t) &= \Lambda^{i-1}(E_{t-i}(p_t^*) - E_{t-i+1}(p_t^*)) \\
&= \Lambda^{i-1}(p_{t|i}^* - p_{t|i-1}^*).
\end{aligned}
$$

Replacing this result in (A8) gives the desired result. ∎

Lemma A.2:

$$
Var_i(p_{it}) = \sum_{i=1}^{T} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) p_{t|i}^2 + \frac{p_{t|T+1}^2}{\Lambda^T} + 2p_{t|T+1}^* p_{t|T+1} + \Lambda^{T+1} p_{t|T+1}^{*2} + \lambda \sum_{j=T+2}^{\infty} (1-\lambda)^j p_{t|j}^{*2},
$$

for all $T \geq 1$.

Proof. From the definition of the cross-sectional variance:

$$
Var_i(p_{it}) = \lambda \sum_{j=0}^{T+1} (1-\lambda)^j p_{t|j}^{*2} + \lambda \sum_{j=T+2}^{\infty} (1-\lambda)^j p_{t|j}^{*2},
$$

so all we have to show is that:

$$
\sum_{i=1}^{T} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) p_{t|i}^2 + \frac{p_{t|T+1}^2}{\Lambda^T} + 2p_{t|T+1}^* p_{t|T+1} + \Lambda^{T+1} p_{t|T+1}^{*2} = \lambda \sum_{j=0}^{T+1} (1-\lambda)^j p_{t|j}^{*2}.
\tag{A9}
$$

We prove that this is true for any T≥1 by induction.

32

For T=1, note that the right hand side of (A9) equals:

$$\lambda p_{t|0}^{*2} + \lambda(1-\lambda)p_{t|1}^{*2} + \lambda(1-\lambda)^2 p_{t|2}^{*2}$$

$$= \lambda\left(\frac{p_{t|1}}{\Lambda^0} + p_{t|1}^*\right)^2 + \lambda(1-\lambda)p_{t|1}^{*2} + \lambda(1-\lambda)^2 p_{t|2}^{*2}$$

where the second line comes from using Lemma A.1 to replace for $p_{t|0}^*$. Calculating the square, and collecting terms this equals:

$$\frac{p_{t|1}^2}{\Lambda^0} + 2p_{t|1}^* p_{t|1} + \Lambda^1 p_{t|1}^{*2} + \lambda(1-\lambda)^2 p_{t|2}^{*2}$$

Next, use Lemma A.1 again to now replace for $p_{t|1}^*$, which leads to, after taking squares and cancelling some terms:

$$\frac{p_{t|1}^2}{\Lambda^0} + 2\left(\frac{p_{t|2} - p_{t|1}}{\Lambda^1}\right)p_{t|1} + \frac{(p_{t|2} - p_{t|1})^2}{\Lambda^1} + 2p_{t|2}^* p_{t|2} + \Lambda^2 p_{t|2}^{*2}$$

Combining the second and third term gives:

$$\left(\frac{1}{\Lambda^0} - \frac{1}{\Lambda^1}\right)p_{t|1}^2 + \frac{p_{t|2}^2}{\Lambda^1} + 2p_{t|2}^* p_{t|2} + \Lambda^2 p_{t|2}^{*2}$$

which shows (A9) holds when T=1.

Continuing the proof by induction, we now show that if (A9) holds for T-1, then it will also hold for T. From the definition of the cross-sectional variance and since (A9) holds for T-1, we know that:

$$\lambda\sum_{j=0}^{T}(1-\lambda)^j p_{t|j}^{*2} = \sum_{i=1}^{T-1}\left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i}\right)p_{t|i}^2 + \frac{p_{t|T}^2}{\Lambda^{T-1}} + 2p_{t|T}^* p_{t|T} + \Lambda^T p_{t|T}^{*2},$$

which implies that:

$$\sum_{i=1}^{T}\left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i}\right)p_{t|i}^2 + \frac{p_{t|T+1}^2}{\Lambda^T} + 2p_{t|T+1}^* p_{t|T+1} + \Lambda^{T+1} p_{t|T+1}^{*2} + \lambda\sum_{j=T+2}^{\infty}(1-\lambda)^j p_{t|j}^{*2} - Var_i(p_{it}) =$$

$$-\lambda(1-\lambda)^{T+1}p_{t|T+1}^{*2} + \frac{p_{t|T+1}^2 - p_{t|T}^2}{\Lambda^T} + 2p_{t|T+1}^* p_{t|T+1} - 2p_{t|T}^* p_{t|T} + \Lambda^{T+1}p_{t|T+1}^{*2} - \Lambda^T p_{t|T}^{*2}.$$

$$(A.10)$$

Now, from lemma A.1:

$$\frac{p_{t|T+1}^2 - p_{t|T}^2}{\Lambda^T} = \left(p_{t|T}^* - p_{t|T+1}^*\right)\left(p_{t|T+1} + p_{t|T}\right),$$

so the right hand side of (A.10) becomes:

$$-\lambda(1-\lambda)^{T+1}p_{t|T+1}^{*2} + \left(p_{t|T}^* - p_{t|T+1}^*\right)\left(p_{t|T+1} + p_{t|T}\right) + 2p_{t|T+1}^* p_{t|T+1} - 2p_{t|T}^* p_{t|T} + \Lambda^{T+1}p_{t|T+1}^{*2} - \Lambda^T p_{t|T}^{*2}.$$

Cancelling terms and rearranging, this equals:

$$(p_{t|T+1} - p_{t|T})(p_{t|T}^* + p_{t|T+1}^*) - \Lambda^T(p_{t|T}^{*2} - p_{t|T+1}^{*2}).$$

But using Lemma A.1 for i=T+1, this equals zero. Looking back at (A.10), this shows that (A9) holds for T, which completes the proof by induction, and also the proof of the lemma. ∎

Lemma A.3:

$$Var_i(p_{it}) = \sum_{i=1}^{\infty}\left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i}\right)p_{t|i}^2$$

Proof. Since lemma A.2 holds for all T, we can take limits as T→ ∞, to obtain:

$$Var_i(p_{it}) = \sum_{i=1}^{\infty}\left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i}\right)p_{t|i}^2 + \frac{p_{t|\infty}^2}{\Lambda^{\infty}} + 2p_{t|\infty}^* p_{t|\infty} + \Lambda^{\infty}p_{t|\infty}^{*2}.$$

Next, realize that $\Lambda^{\infty} = \lambda\sum_{j=0}^{\infty}(1-\lambda)^i = 1$. Also, from (A8) at i= ∞, $E_{t-\infty}(p_t) = E_{t-\infty}(p_t^*)$, so it follows that $p_{t|\infty}^* = -p_{t|\infty}$. Using these results in the expression above, proves the lemma. ∎

Finally, starting from lemma A.3 and using the definition of $p_{t|i}$ and the fact that from the properties of geometric series $\Lambda^i = 1 - (1-\lambda)^{i+1}$ proves the main lemma.

Time 0 impact of shocks

34

Since policy is pre-determined at the time of a shock, its contemporaneous impact is the same regardless of policy. This is determined by the system of two equations in the deviations of p and y from their pre-shock levels (denoted by a $\tilde{}$ over the variable):

$$\tilde{p}_t = \lambda\left(\tilde{p}_t + \alpha(\tilde{y}_t - \tilde{y}_t^N) + \tilde{u}_t\right)$$

$$\tilde{p}_t + \tilde{y}_t = \tilde{e}_t.$$

Solving this gives the time 0 impact of a unit shock to demand or to the (negative of the) natural rate: $\tilde{p}_t = \alpha\lambda/(1 - \lambda(1-\alpha))$ and $\tilde{y}_t - \tilde{y}_t^N = (1-\lambda)/(1 - \lambda(1-\alpha))$, and the impact of a unit shock to markups: $\tilde{p}_t = -\tilde{y}_t = \lambda/(1 - \lambda(1-\alpha))$.

Proof of Proposition 1

Under the proposed policy, since K_t is deterministic, $E_{t-j}p_t = K_t$ for all $j \geq 1$. Taking expectations at $t-1$ of the equation for the price level (8) we obtain:

$$0 = E_{t-1}(y_t - y_t^N) + \sum_{j=1}^{\infty}(1-\lambda)^j E_{t-j}\left(y_t - y_t^N\right)$$

This difference equation has the solution $E_{t-1}(y_t - y_t^N) = 0$, so any shocks at $t-1$ or before have no effect on the first term in the welfare function. Policy is therefore minimizing the variance of $(y_t - y_t^N)$, which varies only with shocks at time t, which policy can do nothing about.

From lemma 1 and given the proposed policy:

$$Var_i(p_{it} - p_t) = \sum_{j=1}^{\infty} \eta_j \left(p_t - K_t\right)^2$$

$$= (p_t - E_{t-1}p_t)^2 \left(\sum_{j=1}^{\infty} \eta_j\right).$$

It then follows that the cross-sectional dispersion of prices is minimized, since unanticipated changes in the price level occur only with respect to innovations at time t, and policy, which is set at $t-1$, can do nothing about these. Thus with respect to the second term in the

35

welfare function, policy is also reaching the optimum. Thus, K_t minimizes the full welfare function and so it is the optimal policy.

Proof of Propositions 2 and 3

The proof of propositions 2 and 3 with respect to demand and natural rate shocks follows immediately from proposition 1. With markup shocks, we represent policy in reaction to a shock as the set of coefficients ϕ_j in the general representation: $E_{t-1}(p_t) = \sum_{j=1}^{\infty} \phi_j \varepsilon_{t-j}$. Similarly, we represent output by the set of φ_j in $y_t = \sum_{j=1}^{\infty} \varphi_j \varepsilon_{t-j}$. From the supply curve in (8), then:

$$\sum_{j=1}^{\infty} \phi_j \varepsilon_{t-j} = \sum_{j=1}^{\infty} \Lambda^j \left(\phi_j \varepsilon_{t-j} + \alpha \varphi_j \varepsilon_{t-j} + \rho_j \varepsilon_{t-j} \right)$$

Since this expression must hold for all possible realizations of ε_{t-j}, it then follows that:

$$\varphi_j = \frac{1}{\alpha} \left(\frac{1 - \Lambda^j}{\Lambda^j} \phi_j - \rho_j \right). \tag{A11}$$

Noting that $p_t - E_{t-i}(p_t) = \sum_{j=0}^{i-1} \phi_j \varepsilon_{t-j}$ and using lemma A.3, we can then write the objective function as:

$$\left[\sum_{j=1}^{\infty} \varphi_j^2 + \omega \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=0}^{i-1} \phi_j^2 \right] \sigma^2, \tag{A12}$$

where σ^2 is the variance of ε_t. Minimizing (A12) subject to (A11), gives the optimal policy

$$\phi_j = \frac{1}{\alpha^2 \omega + \frac{(1-\lambda)^{j+1}}{1-(1-\lambda)^{j+1}}} \rho_j.$$

(A key step is to realize that $\sum_{i=j+1}^{\infty} \left(1/\Lambda^{i-1} - 1/\Lambda^i \right) = 1/\Lambda^j - 1$.) This proves proposition 2. As for proposition 3, we can use the solution above to solve for the output coefficients using (A11). We find that the relation between output and the price level coefficients is given by $\varphi_j = -\alpha \omega \phi_j$, which proves proposition 3 when there are markup shocks.

Proof of Proposition 4

Here we solve for the optimal policy with Calvo pricing, when there are only demand shocks. The problem with natural rate shocks is precisely the same, so we set the y_t^N to zero for the proof. The problem facing the policymaker is to minimize:

$$E\left\{(1-\beta)\sum_{t=0}^{\infty}\beta^t\left[y_t^2+\omega\pi_t^2\right]\right\},$$

where for now we allow for discounting ($\beta \leq 1$), but will then take the limit as β approaches 1. The minimization is subject to the constraints given by the quantity theory (10) and the Phillips Curve:

$$\pi_t = \beta E_t\pi_{t+1} + \kappa(y_t - y_t^N) + u_t \tag{A.13}$$

which corresponds to equation (17). As in our analysis with the sticky information Phillips curve, we can see policy as perfectly choosing $E_{t-1}p_t$, or equivalently $E_{t-1}\pi_t$, by using the quantity theory equation.

Following the approach in Aoki (2002), we can find the optimal policy under commitment by minimizing the Lagrangian function:

$$E\left\{(1-\beta)\sum_{t=0}^{\infty}\beta^t\left[(y_t^2+\omega\pi_t^2)+2\psi_t(\pi_t-\beta\pi_{t+1}-\kappa y_t)\right]\right\},$$

where the ψ_t are Lagrange multipliers. The necessary conditions describing the optimum (taking into account that π_t is set conditional on $t-1$ information) are:

$$y_t = \kappa\psi_t$$
$$\omega\pi_{t|t-1} = \psi_{t-1} - \psi_{t|t-1},$$

for all $t \geq 0$, with the initial condition $\psi_{-1} = 0$, and introducing the notation: $\pi_{t|t-1} = E_{t-1}\pi_t$. Combining the two first-order conditions we obtain:

$$\kappa\omega\pi_{t+1|t} = y_t - y_{t+1|t}. \tag{A.14}$$

This gives us an expression for expected inflation, which we can use in the Phillips curve in (A.13) to get:

$$\kappa \omega \pi_t = \beta(y_t - y_{t+1|t}) + \kappa^2 \omega y_t.$$

Taking expectations of this at $t-1$, using (A.14) lagged one period, and rearranging yields:

$$\beta y_{t+1|t-1} - (1 + \beta + \kappa^2 \omega) y_{t|t-1} + y_{t-1} = 0.$$

This stochastic difference equation can be shown to have one root:

$$\delta = \frac{1 + \beta + \kappa^2 \omega - \sqrt{(1 + \beta + \kappa^2 \omega)^2 - 4\beta}}{2\beta},$$

which is positive and smaller than one, and another root given by $1/\delta\beta$ which is larger than one. An explosive representation for output cannot be optimal, since then the variance of output would be infinity. Thus we can exclude the root which is larger than one. The solution to the difference equation then is:

$$y_{t+j|t-1} = \delta^{j+1} y_{t-1}, \tag{A.15}$$

for any $j \geq 0$.

Combining this result with (A.14) gives the equation:

$$\kappa \omega \pi_{t+1|t} = (1 - \delta) y_t. \tag{A.16}$$

We can use this in the Phillips curve (A.13) to obtain:

$$\pi_t = \frac{\beta(1 - \delta) + \kappa^2 \omega}{\kappa \omega} y_t. \tag{A.17}$$

38

Using the quantity theory equation (10) we can derive:

$$y_t - y_{t|t-1} = e_t - \pi_t + \pi_{t|t-1}.$$

Combining the results in (A.15) and (A.16) with this last expression gives:

$$y_t - \delta y_{t-1} = e_t - \pi_t + \frac{1-\delta}{\kappa\omega} y_{t-1}.$$

Finally, we can replace for all the terms in output using (A.17) and rearrange to get the final solution:

$$\pi_t = \delta\pi_{t-1} + \frac{\beta(1-\delta) + \kappa^2\omega}{\beta(1-\delta) + \kappa\omega(1+\kappa)} e_t.$$

Taking expectations of this expression conditional on $t-1$ information and taking the limit as β approaches 1, proves Proposition 4.

References

[1] Akerlof, George (2002) "Behavioral Macroeconomics and Macroeconomic Behavior," *American Economic Review*, vol. 92 (3), pp. 411-433.

[2] Alogoskoufis, George and Ron Smith (1991) "The Phillips Curve, the Persistence of Inflation and the Lucas Critique: Evidence from Exchange Rate Regimes," *American Economic Review*, vol. 81 (5), pp. 1254-1275.

[3] Aoki, Kosuke (2002) "Optimal Commitment Policy Under Noisy Information," CEPR Discussion Paper No. 3370.

[4] Ball, Laurence (1994) "Credible Disinflation with Staggered Price Setting," *American Economic Review*, vol. 84, pp. 282-289.

[5] Ball, Laurence (1999) "Efficient Rules for Monetary Policy," *International Finance*, vol. 2 (1) , pp. 63-83.

[6] Ball, Laurence (2000) "Near Rationality and Inflation in Two Monetary Regimes," NBER Working Paper No. 7988.

[7] Ball, Laurence and N. Gregory Mankiw (1995) "Relative-Price Changes as Aggregate Supply Shocks," *Quarterly Journal of Economics*, vol. 110 (1), pp. 161-193.

[8] Barsky, Robert (1987) "The Fisher hypothesis and the forecastability and persistence of inflation," *Journal of Monetary Economics*, vol. 19, pp. 3-24.

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

[10] Clarida, Richard, Jordi Gali, and Mark Gertler (1999) "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature,* vol. 37, pp. 1661-1707.

[11] Clarida, Richard, Jordi Gali and Mark Gertler (2002) "A Simple Framework for International Monetary Policy Analysis," *Journal of Monetary Economics*, vol. 49 (5), pp. 879-904.

[12] Cover, James P. and Paul Pecorino (2001) "Price and Output Stability Under Price Level Targeting," University of Alabama working paper WP01-05-01.

[13] Cover, Thomas M., and Joy A. Thomas (1991) *Elements of Information Theory,* New York: John Wiley and Sons.

[14] Fischer, Stanley (1977) "Long-term Contracts, Rational Expectations, and the Optimal Money Supply Rule," *Journal of Political Economy,* vol. 85 (1), pp.191-205.

[15] Friedman, Milton (1968) "The Role of Monetary Policy," *American Economic Review,* vol. 58, pp. 1-17.

[16] Gabaix, Xavier and David Laibson (2001) "The 6D Bias and the Equity Premium Puzzle," *NBER Macroeconomics Annual.*

[17] Gali, Jordi and Mark Gertler (1999) "Inflation Dynamics: A Structural Econometric Model," *Journal of Monetary Economy*, vol. 44 (2), pp. 195-222.

[18] Goodfriend, Marvin and Robert King (1997) "The New Neoclassical Synthesis and the Role of Monetary Policy," *NBER Macroeconomics Annual*, pp. 231-283.

[19] Gordon, Robert (1997) "The Time-Varying NAIRU and Its Implications for Monetary Policy," *Journal of Economic Perspectives*, Winter, pp. 11-32.

[20] Hall, Robert E. (1984) "Monetary Strategy with an Elastic Price Standard," in *Price Stability and Public Policy: A Symposium Sponsored by the Federal Reserve Bank of Kansas City.*

[21] Hall, Robert E. and N. Gregory Mankiw (1994) "Nominal Income Targeting," in N. G. Mankiw, ed., *Monetary Policy*, Chicago University Press, Chicago.

[22] King, Robert G. And Alexander L. Wolman (1999) "What Should Monetary Policy Do If Prices Are Sticky?", pp. 349-404 in John B. Taylor, ed., *Monetary Policy Rules*, University of Chicago Press, Cambridge, MA.

[23] Lucas, Robert E. (1973) "Some International Evidence on Output-Inflation Trade-offs," *American Economic Review,* vol. 63, pp. 326-334.

[24] Mankiw, N. Gregory (2001) "The Inexorable and Mysterious Tradeoff Between Inflation and Unemployment," *Economic Journal*, col. 111, pp. C45-C61.

[25] Mankiw, N. Gregory and Ricardo Reis (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, vol 117 (4).

[26] Romer, David (2001) *Advanced Macroeconomics.* Second edition. McGraw-Hill.

[27] Rotemberg, Julio (1982) "Monopolistic Price Adjustment and Aggregate Output," *Review of Economic Studies*, vol. 44, pp. 517-531.

[28] Rotemberg, Julio J. and Michael Woodford (1997) "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual*, pp. 297-346.

[29] Sargent, Thomas (1993) *Bounded Rationality in Macroeconomics*, Oxford: Oxford University Press.

[30] Sims, Christopher A. (2002) "Implications of Rational Inattention," *Journal of Monetary Economics*, forthcoming.

[31] Staiger, Douglas, James Stock and Mark Watson (1997) "The NAIRU, Unemployment, and Monetary Policy," *Journal of Economic Perspectives*, vol. 11 Winter 1997, pp. 33-51.

[32] Steinsson, Jon (2002) "Optimal Monetary Policy in an Economy with Inflation Persistence," *Journal of Monetary Economics*, forthcoming.

[33] Svenssson, Lars E. O. (1997) "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets," *European Economic Review*, vol. 41, pp. 1111-1146.

[34] Svensson, Lars E. O. (1999) "Price Level Targeting vs. Inflation Targeting: A Free Lunch?", *Journal of Money, Credit and Banking*, vol. 31, pp. 277-295.

[35] Taylor, John B. (1980) "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, vol 88, pp. 1-22.

[36] Vestin, David (1999) "Price Level Targeting versus Inflation Targeting in a Forward Looking Model", unpublished, Stockholm University.

[37] Woodford, Michael (2001) "Imperfect Common Knowledge and the Effects of Monetary Policy," Princeton University, forthcoming in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds.

[38] Woodford, Michael (2002) *Interest and Prices*, in preparation for Princeton University Press.

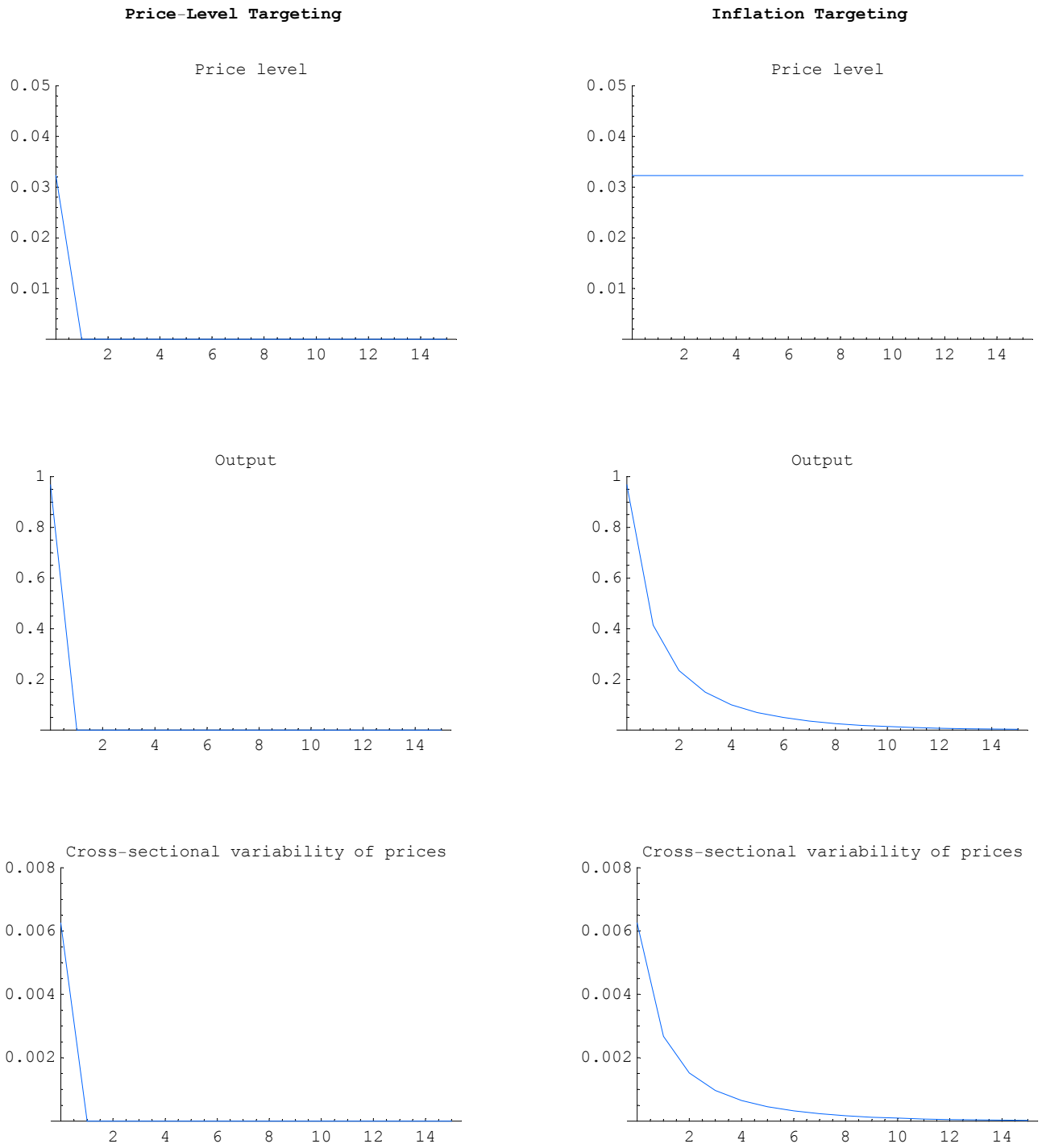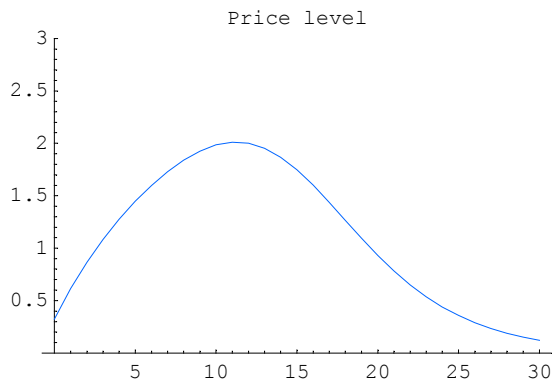# Figure 1. Impact of a demand shock with price-level or inflation targeting

Price-Level Targeting

Inflation Targeting

Price level

Price level

Output

Output

Cross-sectional variability of prices

Cross-sectional variability of prices

Figure 2. Impact of a markup shock under the optimal policy

If serial correlation of shock is 0.8   **If serial correlation of shock is 0.4**

Price level