Monetary Policy, Endogenous Inattention, and the Volatility Trade-off*

William A. Branch
U.C. Irvine

John Carlson
Federal Reserve Bank of Cleveland

George W. Evans
University of Oregon

Bruce McGough
Oregon State University

December 7, 2004

Abstract

This paper addresses the output-price volatility puzzle by studying the interaction of optimal monetary policy and agents' beliefs. We assume that agents choose their information acquisition rate by minimizing a loss function that depends on expected forecast errors and information costs. *Endogenous inattention* is a Nash equilibrium in the information processing rate. Although a decline of policy activism directly increases output volatility, it indirectly anchors expectations, which decreases output volatility. If the indirect effect dominates then the usual trade-off between output and price volatility breaks down. This provides a potential explanation for the 'Great Moderation' that began in the 1980's.

JEL Classifications: E52; E31; D83; D84

Key Words: expectations, optimal monetary policy, bounded rationality, economic stability, adaptive learning.

---

*We are indebted to Larry Ball and Ricardo Reis for their detailed comments, and for the comments received at a preliminary presentation of these results to a "brown bag" seminar at the Federal Reserve Bank of Cleveland. For helpful comments we also thank seminar participants at the European Central Bank, University of Oregon, the 10th annual International Conference on Computing in Economics and Finance, and the conference on Dynamic Models and Monetary Policymaking co-sponsored by the FRB-Cleveland, Bank of Canada, and the Swiss National Bank. Financial support from the National Science Foundation Grant is gratefully acknowledged. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve System or of the Board of Governors of the Federal Reserve System.

1 Introduction

The "Sticky-Information" model of (Mankiw and Reis 2002) and (Ball, Mankiw, and Reis 2003) has recently been proposed as an alternative to the New Keynesian Phillips curve employed, for example, by (McCallum and Nelson 1999) and (Clarida, Gali, and Gertler 1999), and developed in detail by (Woodford 2003). The New Keynesian approach, which rests on the "Calvo assumption" that only a proportion of firms each period have an opportunity to adjust their prices, delivers a forward-looking expectational Phillips curve. The sticky-information model replaces this with the assumption that each period a fixed proportion of firms update their information set, and yields a backwards-looking expectational Phillips curve arising from the slow diffusion of information through the economy. (Ball, Mankiw, and Reis 2003) argue that the sticky-information approach is more consistent with widely accepted views about inflation persistence and the effects of monetary policy, e.g. about the output costs of disinflation.[1]

Both of these approaches treat the proportion of agents that fully adjust each period as exogenous to the model. This is convenient as a simplification, but endogenizing the proportion is desirable both from a theoretical perspective and from the viewpoint of increased realism. In this paper we examine this point in detail and argue that the consequences for monetary policy can be far-reaching. We develop our analysis as an extension of the (Ball, Mankiw, and Reis 2003) model because it fits neatly with our "bounded rationality" viewpoint that the frequency with which agents update and utilize new information should depend on the benefits relative to the costs of doing so. One way to view our contribution is that we study the implications of applying the "Lucas critique" to the rate of information acquisition as well as to expectation formation.

Our approach has a number of natural applications to monetary policy, but to illustrate its potential importance we restrict attention to one: the output-price volatility trade-off that is implicit in most monetary policy models. A focus of recent research has been on the apparent change in the stance of monetary policy from the 1970's to the 1980's. A particularly striking finding is that the monetary authorities were 'passive' in reacting to inflation during the 1970's but aggressive during the 1980's and 1990's ((Clarida, Gali, and Gertler 2000), (Lubik and Schorfeide 2003), (Schorfeide 2003)). In the applied literature, there is interest in whether these findings are related to empirical evidence of a decline in output volatility in the US (McConnell and Quiros 2001) and to the finding of (Blanchard and Simon 2001) that inflation and output volatility are positively correlated. The decline in economic volatility is a finding of such paramount importance it has been given the moniker 'The Great Moderation' by (Bernanke 2004). Table 1 illustrates the decline in output and price

[1]See also (Fuhrer and Moore 1995), (Mankiw 2001) and (Mankiw and Reis 2002). Versions of the New Keynesian approach that yield inflation persistence are developed in (Woodford 2003).

volatility for the United States over 1947:1-2004:1.

INSERT TABLE 1 HERE

The cause of the Great Moderation is an important and open question.[2] Some authors have attributed the decline in economic volatility to a fundamental shift in the focus of monetary policy. (Orphanides and Williams 2003b) maintain that monetary authorities concerned themselves primarily with output stabilization ('activist policy') during the late 1960's and 1970's and then switched their emphasis to price stability in subsequent years. (Bernanke 2004) contends that monetary policy during the 1970's exhibited 'output optimism' and 'inflation pessimism'. According to Bernanke's hypothesis, an overplaced emphasis on exploiting a (perceived) Phillips curve trade-off, and a mistaken belief that monetary policy was unable to control inflation, led to higher volatility in both output and inflation – confirming the positive correlation in (Blanchard and Simon 2001).[3] Bernanke conjectures that a movement away from activist monetary policy anchored inflation expectations and produced lower volatility in *both* inflation and output.

In many models, however, there is a trade-off between inflation and output volatility: a renewed focus on inflation stabilization will lead monetary policy to produce higher output volatility.[4] Although (Bernanke 2004), (Svensson 2003), and others, conjecture that if policymakers can more tightly pin down inflation expectations then they will achieve economic stability, the specific channels for this effect are left open.

A few possible mechanisms have appeared in the literature. In (Orphanides and Williams 2003a) the trade-off disappears when agents engage in 'perpetual learning' and policymakers have the appropriate preferences on inflation and output volatility. In their model inflation expectations persistently deviate from rational expectations, becoming a source of instability and providing an additional role for monetary policy. An alternative story, given in (Clarida, Gali and Gertler 2000), retains rational expectations, but relies on multiple equilibria. In particular they suggest that inappropriate interest rate rules in the earlier part of the post-WWII period were consistent with sunspot equilibria.

We propose a complementary but distinct approach that requires neither sunspot equilibria nor persistent deviations from rational expectations and emphasizes a plausible mechanism in the spirit of (Bernanke 2004). Extending the model of (Ball,

[2]The evidence for a one-time permanent shift in monetary policy, and for a similar shift in macroeconomic volatility, is open to other interpretations. (Cogley and Sargent 2002) and (Sims and Zha 2004) present evidence of drifting and regime switching over much of the post-WWII period.

[3](Sargent 1999) develops a model where the central bank mistakenly exploits a Phillips curve even though the natural rate hypothesis holds.

[4]See (Woodford 2003) for examples.

Mankiw, and Reis 2003) to endogenize the rate at which firms update their information, the current paper develops a framework in which to study the joint determination of optimal monetary policy and private sector expectations, and the connection of this joint relationship to the Great Moderation. We study the intimate connection between optimal monetary policy and the equilibrium anchoring of price expectations that arises through the endogenous response of private sector information acquisition or 'attentiveness.' Our key insight is that if monetary authorities follow policies that stabilize the aggregate price path, then this allows firms to update information less frequently, reducing the sensitivity of the economy to exogenous shocks.[5]

Since attentive readers may have noticed that we have referred to both inflation and price volatility, before outlining our approach we comment on how we treat this distinction. There is a continuing debate among experts in monetary policy about the precise form of the price stability objective that is appropriate for policymakers to pursue.[6] Although there are several specific issues, the one that is most relevant is whether the central bank should attempt to stabilize the inflation rate or instead stabilize the path of the price level around a deterministic path. In the latter case this might be a constant growth rate price path or some more complicated trend.

This question, though of considerable importance, is essentially orthogonal to the issue under study, and we therefore take a pragmatic approach. Empirically, looking for example at the US Consumer Price Index, there was a substantial fall in the standard deviation of the quarterly inflation rate from the 1950:1 - 1983:4 period to the 1984:1 - 2003:4 period. Measuring inflation as the change in the log(CPI) the ratio of the standard deviation in the former period to the latter period is 2.45. Alternatively, if one detrends the log(CPI) using the Hodrick-Prescott filter and computes the standard deviation of the price level separately for the two periods, there is again a substantial fall: the ratio for the former compared to the latter period is 2.28.

Given the fall in aggregate output volatility over the corresponding period documented by (McConnell and Quiros 2001), it is clear that the existence of 'The Great Moderation' is robust to this issue. Because our theoretical analysis is most conveniently developed as an extension to the (Ball, Mankiw, and Reis 2003) model of "sticky information," in which the optimal monetary policy is formulated in terms of the variability of the price level around an arbitrary trend, we will develop our results in terms of price stability rather than inflation stability. However, we do not mean to imply that this issue is settled and we suspect that an alternative formulation of our ideas could be developed in terms of inflation variability.

Our principal argument is that monetary policy has both direct and indirect effects on output and price volatility: the direct effect gives the usual trade-off – by moving

[5]This is very close in spirit to the first type "stability-enhancing" change to the "economic environment ... induced by improved monetary policies" listed by (Bernanke, 2004), p. 5.

[6]See, for example, (Woodford 2002) for a discussion and references.

away from activist policy the Fed tends to increase output volatility; the indirect effect is channeled through expectation formation – policy that stabilizes price will anchor price expectations and thereby induce agents to be less reactive to intrinsic shocks, reducing both output and price variability. Thus there is a tension between the direct and indirect effects of policy; and which effect dominates determines whether a switch to 'output pessimism' and 'inflation optimism' can account for the Great Moderation. The novelty of our paper is the development of a model that can address this issue as an equilibrium response, and in particular provide conditions under which the Bernanke hypothesis is validated.

Our resolution of the policy tension begins with a relatively new approach to bounded rationality that endows agents with a correct model of the economy, but which assumes it is costly to acquire and process information.[7] Recent proponents of this approach in macroeconomics are (Sims 2003), (Mankiw and Reis 2002), and (Ball, Mankiw, and Reis 2003). These models assume that agents form conditional expectations, as in RE, but that the information set on which they condition may include only past data. (Ball, Mankiw, and Reis 2003) (hereafter BMR) assume that agents have a time-invariant probability for updating their information in any given period. The resulting model is a sticky information version of the Calvo pricing model emphasized by (Woodford 2003).[8]

In the current paper we take the BMR model as a laboratory in which to study the interaction of optimal policy, information acquisition and private sector expectations. We take their motivation of costly updating seriously and assume that agents choose the rate at which they acquire new information by minimizing a quadratic loss function. A key insight of our approach is that this loss function depends on the information updating rate of the other agents. We define *Endogenous Inattention* as a symmetric Nash equilibrium in information updating together with the associated stationary stochastic processes for aggregate price-level and output.[9] In this Nash Equilibrium we treat the monetary authorities as following the optimal monetary policy recommended by BMR, given the equilibrium updating rate.

The BMR model is a simple model of monopolistically competitive firms combined with a quantity equation aggregate demand relation. Optimal monetary policy is a stochastic process for the money supply that minimizes a second-order approximation

[7] This approach to costly expectation formation is motivated by modeling agents as rational utility maximizers in the presence of costly information acquisition or processing. This is in contrast to the full information RE approach of (Muth 1960) which has instead been called 'consistent expectations' by (Simon 1978), among others.

[8] A model of sticky information in wages is developed in (Koenig 2004). (Yetman 2003) compares symmetric Nash equilibria in sticky information and sticky price models. (Reis 2003) studies optimal endogenous inattention for consumers with an exogenous income stream. (Adam 2004) analyzes optimal monetary policy when firms have finite capacity to process information as in (Sims 2003).

[9] Some readers would find the term "endogenous attention" more natural, but the concept of "rational inattention" was introduced by Sims (2003) and used by BMR.

to the social welfare function. Because optimal policy in this model depends on the equilibrium rate of information updating, in our formulation monetary policy and the updating frequency, or 'attentiveness' of agents, are jointly determined.

The joint determination of Endogenous Inattention and optimal policy has important implications. We model the 'output optimism' of monetary policy by parameterizing policymaker preference for low price variance relative to output variance. The usual result is that as the policy authority becomes less 'activist' (i.e. places a higher weight on price variance) then the reduction in price volatility is accompanied by higher output volatility. We show that this trade-off is indeed present in the sticky information model of BMR.

Our main result is that the nature and existence of a trade-off between price and output stability depends on the joint determination of the rate of information processing and optimal policy. If policymakers are more activist, the direct effect, including the adjustment of rational expectations, is a reduction of output volatility and increased price-level volatility. However, an indirect effect on expectations arises from the increase in price level volatility which, in turn, induces agents to become more 'attentive'. This greater attentiveness tends to increase the volatility of output. Whether there is a trade-off between inflation and output volatility thus depends on whether the indirect or direct effect of policy dominates. We show that which effect dominates depends on how strongly the equilibrium level of attentiveness responds to the higher price-level volatility.

In contrast to the implications of the BMR model, we show that for relatively low costs of information accrual, the policy frontier can be non-monotonic. As the government switches from activist to less activist policy, there need be no trade-off between price and output variance – both can be lowered simultaneously.[10] However, as policy becomes increasingly vigilant against price volatility a trade-off between price and output variance can emerge. Our results, showing the possibility of a decline in both output and price volatility, provide a theoretical basis for the conjectures found in the inflation targeting literature, and, in particular, are supportive of the Bernanke hypothesis.

The organization of our paper is as follows. In Section 2 we summarize the BMR model and show how to extend it to endogenize the rate of "inattention." In Section 3 we then prove the existence of an equilibrium with endogenous inattention and explore the comparative statics results. In these sections, in which the theoretical results are presented, we treat the private agents and the policymaker as involved in a simultaneous move game and the endogenous inattention equilibrium is defined as a Nash Equilibrium.

In Section 4, we explore the policy implications of our comparative statics results. To organize our discussion we plot output and price volatility as a function of the

[10] A policy frontier is a set of inflation-output volatility pairs indexed by the activism parameter.

policymaker's preference parameter ω, and we show that it is possible for the resulting policy "frontier" to be upward sloping if the rate of information acquisition λ, i.e. the rate of inattention $1-\lambda$, is sufficiently responsive. The parameter ω measures the weight in the policymaker's loss function placed on price variance relative to output variance[11] and we refer to a high value of ω as a low degree of activism since in effect it corresponds to a reduced desire to smooth output. We then propose to interpret the "great moderation" as the result of a permanent reduction in activism beginning in the early 1980s, moving the economy down along a positively sloped frontier.

In proposing this explanation we are going beyond the formal model, and a number of specific interpretations of the shift in policy and the resulting great moderation are possible, depending on the degree of sophistication that we want to attribute to policymakers. As presented in Sections 2 and 3, the equilibrium described is the usual Nash equilibrium in a simultaneous move game. Within this setting policymakers are fully cognizant of the structure of the economy but, as in (Kydland and Prescott 1977), are condemned by the timing protocol to an inefficient equilibrium. An increase in ω leading to a simultaneous decline in output and price volatility might either be the fortuitous result of an exogenous change in policymakers preferences or a more conscious attempt to improve welfare by appointing a conservative banker, following the logic of (Rogoff 1985).

The interpretation of the great moderation just described assumed sophisticated policymakers who understood the endogeneity of the information acquisition rate λ, but were hemmed into an inefficient equilibrium by the timing protocol of the economy. If instead the timing protocol is that policymakers first choose the policy rule and that private agents then respond optimally, given this policy, then an alternative interpretation is possible. Suppose that policymakers were initially naive, believing that λ was exogenous, but that over time policymakers began to appreciate the importance of the various channels through which a more stable price level affects the economy. A growing understanding, in particular, that λ is endogenous, could eventually lead policymakers to adopt less activist policies in order to gain the additional benefits of reduced output volatility.

While both of these interpretations are viable, we prefer a third interpretation in which policymakers, as well as private agents, are neither naive nor fully informed rational, but instead are boundedly rational in the spirit of (Marcet and Sargent 1989), (Sargent 1999) and (Evans and Honkapohja 2001). In this interpretation, which we develop in Section 5, policymakers follow a policy rule of the form recommended by BMR, but instead of using fully rational forecasts to implement the policy, which would require knowledge of the full structural model, they forecast using a time-series model, updating the parameters over time using recursive least-squares. An analogous bounded rationality assumption is made for private firms, who use consultants to act

[11] It is actually the cross-sectional price variance that enters the policymakers' loss function, and we therefore examine both the cross-sectional and time-series price variance.

as information gatherers and provide firms with an estimate of their optimal frequency for information processing as well as with forecasts of the optimal prices to set. Least-squares learning allows both policymakers and firms to track changes in structural parameters that may occur for a variety of reasons.

This "adaptive learning" version of the model, is explored numerically in the second part of Section 5. We first show that, with fixed parameters, under adaptive learning the economy converges over time to the equilibrium described in the earlier theoretical sections of the paper. In particular, the value of the endogenous inattention rate λ converges to its equilibrium value. We then consider a system initially in equilibrium and look at the impact of an exogenous increase in ω, i.e. a permanent decrease in policy activism, with the cost of information accrual parameter set at a moderately low level. The numerical results again track the earlier theoretical results showing that a simultaneous decline in price and output volatility is possible, but with one difference. Initially, when the new policy rule is implemented, output volatility rises in line with the "standard" view of a trade-off, reflecting the transitional period in which λ adapts over time to its new lower equilibrium level. However, in the long-run the "great moderation" emerges and output as well as price volatility decline permanently.

After reviewing the numerical results, Section 6 concludes. Our adaptive learning version of the model provides results that are more hopeful than those of (Sargent 1999) in the sense that with appropriate policy a permanent decrease in volatility is possible. However, we also end with a cautionary note: lower activism will continue to lower price volatility, but there appears to be a limit to the improvement in output volatility.

2 The Model

We begin by briefly reviewing the model developed in (Ball, Mankiw, and Reis 2003). In this review, we assume, as did BMR, that the probability of information updating, $\bar{\lambda}$, is exogenous and fixed. This allows us to use their results on optimal monetary policy to obtain equilibrium paths of price and output for a given set of structural parameters. Then, taking as given both monetary policy and the updating frequency $\bar{\lambda}$, we consider the incentive for a single agent to deviate from $\bar{\lambda}$, where this incentive is measured by expected squared forecast error plus a cost of the choice λ. An equilibrium occurs when each agent does not have an incentive to deviate from the aggregate $\bar{\lambda}$.

2.1 The Ball-Mankiw-Reis Model

The economy is populated by a continuum of yeoman farmers. Each farmer uses its own labor to produce a good to sell in a monopolistically competitive market. The instantaneous utility of agent i is given by

$$U(C_{it}, Y_{it}) = \frac{(C_{it})^{1-\sigma}}{1-\sigma} - \frac{1}{1-\sigma} - \frac{\hat{A} Y_{it}^{1+\zeta}}{1+\zeta}, \tag{1}$$

where C_{it} is the usual consumption index defined in terms of the CES aggregator:

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

The last term in (1) captures the disutility of labor. The production function is $Y = AL$ with labor L, technology A to be normalized later for convenience, and $\hat{A} = A^{-(1+\zeta)}$.

Agents choose sequences of consumption and labor in order to maximize the expected discounted utility stream subject to their budget constraint, which includes a government levied proportional sales tax τ_t assumed to follow a stationary process. The consumer problem leads to a demand that, in log form, is given by

$$y_{it} = y_t + \gamma(p_{it} - p_t) \tag{2}$$

where p_t is the log of the usual price index and $\gamma > 1$ is the elasticity of substitution between different goods. We note that to obtain this form of demand, we must assume the presence of a complete market for risk which allows agents to insure themselves against idiosyncratic information shocks, and thus allows us to identify consumption and output.

The producer's pricing problem may now be solved, taking the computed demand as given, resulting in an optimal price (given full information) of the form

$$p_{it}^* = p_t + \alpha y_t + u_t \tag{3}$$

where $\alpha = (\zeta + \sigma) / (1 + \gamma\zeta)$. We have chosen the technology constant A to normalize the log natural output level to zero.[12] u_t is a stationary stochastic process deriving its structure, for example, from the sales tax τ_t. We follow BMR by interpreting u_t as capturing mark-up shocks and take it to have an AR(1) structure: $u_t = \rho u_{t-1} + \varepsilon_t$ with $0 < \rho < 1$.[13] Mark-up (or supply) shocks are standard in the literature and are taken

[12] BMR allow A to form a stochastic process, thus allowing for drift in the natural rate as well as for the analysis of productivity shocks. We abstract from this here to focus attention on the impact of mark-up shocks, which are the usual source of volatility tradeoffs.

[13] BMR allow u_t to have general $MA(\infty)$ form.

to represent shifts in the Phillips curve; for further discussion see (Woodford 2003). The mark-up shocks u_t represent the only stochastic component to the economy. The unconditional equilibrium volatility of price and output are determined by the variance of ε_t, denoted σ_ε^2, and the way in which agents incorporate its past history into their expectations. In particular, under sticky information the persistence of shocks also depends on how frequently agents update their information sets.

Whereas the above model is fairly standard – see for example (Woodford 2003) – BMR introduce a novel information structure that fundamentally alters equilibrium outcomes. Combining the probabilistic friction of (Calvo 1983), with the limited information capacity notion of (Sims 2003) these authors assume that agents update their information with exogenous probability $0 < \lambda < 1$ each period, and each agent sets a price path optimally every period, subject to their information constraint.[14] Thus an individual who last updated information k periods ago will set price equal to $E_{t-k}p_t^*$. Equilibrium price is given by

$$p_t = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j}(p_t + \alpha y_t + u_t). \tag{4}$$

Note that this identification requires approximating the price index as an average, $p_t = \int p_{it} di$.

Equation (4) is a Phillips curve and represents the aggregate supply relationship in the economy. Aggregate demand is derived from a cash-in-advance constraint and takes the form

$$y_t = \hat{m}_t - p_t,$$

where $\hat{m}$ is the policy instrument set in time $t-1$.[15] BMR conclude with the clever observation that there is a linear relationship between $E_{t-1}p_t$, $\hat{m}_t$, and other information available at $t-1$; thus, we may assume that policymakers set $E_{t-1}p_t$.

The model is closed by specifying monetary policy, which, as we just noted, is equivalent to specifying a (stochastic) time path for $E_{t-1}p_t$. BMR assume that the preferences of policymakers are captured by a quadratic loss in output and cross-sectional relative price variance, as given by

$$\mathcal{L} = Var(y_t) + \omega E\left(Var_i(p_{it} - p_t)\right). \tag{5}$$

This equation can be derived as a second order approximation to average cross-sectional utility. When this approximation is taken seriously, the associated value

[14] This is the idiosyncratic risk mentioned earlier. An individual's income will vary with respect to average output depending on her most recent information. The complete market for risk assures the agent a yearly consumption level equal to average output, regardless of her income level.

[15] BMR allow for a noisy realization of $\hat{m}$ (i.e. aggregate demand shock), from which we abstract in much of the paper in order to focus on mark-up shocks.

of ω is $\gamma^2(\zeta+\gamma^{-1})/(\zeta+\sigma)$, though BMR consider varying values of ω for fixed structural parameters, and we will as well. We attach the interpretation of 'activism' to this parameter; as ω increases the policymaker places a higher relative loss on cross-section price variation and less on unconditional output variance. Policymakers with low values of ω are "activist" in the sense that they place a relatively high weight on reducing output volatility.[16]

Having specified the government's objective, BMR analytically solve the optimal policy problem. They show that when optimal policy is followed, the first-order condition

$$E_{t-1}p_t = -\frac{1}{\alpha\omega}E_{t-1}y_t \tag{6}$$

must be satisfied. Solving for the equilibrium paths of price and output then yields

$$p_t = \sum_{j=0}^{\infty}\phi_j\varepsilon_{t-j} \tag{7}$$

$$y_t = \sum_{j=0}^{\infty}\varphi_j\varepsilon_{t-j} \tag{8}$$

with

$$\phi_j = \frac{\rho^j}{\alpha^2\omega+\frac{(1-\lambda)^{j+1}}{1-(1-\lambda)^{j+1}}},\quad \varphi_j = -\alpha\omega\phi_j \text{ for } j>0, \text{ with} \tag{9}$$

$$\phi_0 = \frac{\lambda}{1-\lambda(1-\alpha)} \text{ and } \varphi_0 = -\phi_0. \tag{10}$$

Not surprisingly, provided $\rho \neq 0$, a decrease in activism (increase in ω) lowers the mean cross-sectional variance of prices. Equations (7) and (8) also imply the usual trade-off between σ_p^2 and σ_y^2, the unconditional (time-series) variances of price and output. This can be seen as follows. For $0<\lambda<1$, an increase in ω reduces $|\phi_j|$, for all $j>0$, and increases $|\varphi_j|$, for all $j>0$. It follows immediately that an increase in ω reduces $\sigma_p^2 = Var(\varepsilon_t)\sum_{j=0}^{\infty}\phi_j^2$ and increases $\sigma_y^2 = Var(\varepsilon_t)\sum_{j=0}^{\infty}\varphi_j^2$. (In the extreme case $\lambda=1$, σ_y^2 becomes independent of ω and the trade-off is vertical.)

Moreover, for $0<\lambda<1$, as ω continues to increase, price-level variance and output variance will converge positive, finite values. These insights will be important for discussion of the output-price volatility trade-off when λ is determined endogenously. For this reason, we summarize this discussion in the following remark.

[16] "Activism" is also sometimes used to mean a lower weight on the output gap in an interest rate rule. Because of the quantity theory form of aggregate demand used in the BMR model, there is no IS curve and consequently monetary policy is formulated in terms of $\hat{m}_t$ or $E_{t-1}p_t$ rather than an interest rate rule. In the current context our use of the term "activist policy" seems the most natural.

Remark: *Consider the BMR model with exogenous λ.*

$$\lim_{\omega\to\infty}\sigma_p^2 = Var(\varepsilon_t)\lim_{\omega\to\infty}\left(\sum\nolimits_{j=0}^{\infty}\phi_j^2\right) = \sigma_\varepsilon^2\left(\frac{\lambda}{1-\lambda(1-\alpha)}\right)^2$$

$$\lim_{\omega\to\infty}\sigma_y^2 = Var(\varepsilon_t)\lim_{\omega\to\infty}\left(\sum\nolimits_{j=0}^{\infty}\varphi_j^2\right) = \sigma_\varepsilon^2\left[\left(\frac{\lambda}{1-\lambda(1-\alpha)}\right)^2 + \frac{1}{\alpha^2}\frac{\rho^2}{1-\rho^2}\right]$$

Intuitively, in the presence of a (positive) markup shock, price will rise and output will fall due to the fact that policy is lagged one period and thus cannot respond contemporaneously to the shock. An option for policymakers is to return price to its mean the following period, but, pursuing such a policy would exacerbate the impact of the shock on output. In order for agents to lower prices in the presence of a markup shock whose influence is still felt due to serial correlation, output must fall further. The form of the government's objective function makes such a policy suboptimal as policymakers prefer to allow prices to capture some of the economy's volatility. This trade-off is consistent with the sticky-price model of (Woodford 2003) and appears in most models with mark-up or supply shocks. It is this result we seek to square with the empirical observation of a simultaneous decline in output and inflation volatility in the U.S..

The key to our results will involve the endogenous response of λ. At this stage it is therefore helpful to obtain the effects of an exogenous change in λ on σ_p^2 and σ_y^2. Incorporating also the results just stated we have:

Proposition 1 *Consider the BMR model with exogenous λ.*

1. $\omega \uparrow \Rightarrow \sigma_p^2 \downarrow,\ \sigma_y^2 \uparrow,\ E(Var_i(p-p_i)) \downarrow$.
2. $\lambda \uparrow \Rightarrow \sigma_p^2 \uparrow,\ \sigma_y^2 \uparrow$.

The effect of ω on σ_p^2 and σ_y^2 was shown above, and the impact on $E(Var_i(p-p_i))$ is shown in Appendix A. Note that the impact on the expected cross-sectional price variation of an increase in λ is ambiguous. This is intuitive as the cross-sectional variance will be zero when λ is zero or one. The second set of results, giving the impact of λ, are straightforward. Increases in λ can be seen to increase both $|\phi_j|$ and $|\varphi_j|$, for all j, and hence increase both σ_p^2 and σ_y^2. Intuitively, as λ increases there is a greater price, and hence output, response to new information.

The possibility that a reduction in activism (increased ω) could lead to greater stability in both output and prices can be seen to arise if it is accompanied by a reduction in λ. We now turn to the endogenous determination of λ in an equilibrium setting.

2.2 Endogenizing Inattention

BMR take λ as exogenous to the model. We propose extending their model by making $0 \leq \lambda \leq 1$ a choice variable. In our framework, agents choose an intensity with which to gather and analyze information and this chosen intensity yields a probability of obtaining and processing current information. To model this choice, we assume agents choose λ to minimize mean squared forecast error, as discussed below. Not surprisingly, the mean squared forecast error is decreasing in λ and so if gathering information is costless, the choice for agents is quite simple: choose $\lambda = 1$. However, we argue that information gathering and processing is not costless, and instead assume a cost function that is quadratic in λ. Purely quadratic costs allow for increasing marginal costs, with marginal cost tending to zero as $\lambda \rightarrow 0$. This implies that it is always optimal to choose a non-zero probability of updating information.

The choice of λ for a given agent depends on the equilibrium stochastic processes of price and output, which in turn depend on structural parameters, the monetary policy parameter ω, and the intensity with which other agents gather information. Given the monetary policy dictated by ω, the optimal choice of λ by private agents is interdependent. Thus the correct equilibrium concept for our model is Nash, and we focus on Nash equilibria that are symmetric with respect to the private agents. Note also that the stochastic processes for price and output depend, in turn, on the Nash equilibrium value of λ.

We need to be explicit also about the policy assumptions. As just indicated, we take ω to be exogenous, and we make the assumption that policymakers follow the optimal monetary policy recommended by BMR, so that price and output processes are given by (7)-(8) with coefficients (9) and (10). In effect, policymakers treat the equilibrium rate of information gathering by private agents as given, and thus our equilibrium value of λ is a Nash equilibrium in choices of private agents *and* the policymaker. This has important implications for comparative statics and is discussed further below.

Let $\bar{\lambda}$ be the economy-wide probability of updating information and define $p_t^*(\bar{\lambda})$ as the optimal price given the economy wide $\bar{\lambda}$, that is

$$p_t^*(\bar{\lambda}) = p_t(\bar{\lambda}) + \alpha y_t(\bar{\lambda}) + u_t,$$

where $p_t(\bar{\lambda})$ and $y_t(\bar{\lambda})$ are the equilibrium price level and output given that all agents use $\bar{\lambda}$.

Now let $\hat{p}_t(\lambda)$ be the price set by a firm at time t given that the firm updates its information with probability λ. Note, $\hat{p}_t(\lambda)$ is a random variable that depends not only on the process of markup shocks hitting the economy, but also on a process determining whether updating occurs. It may help to think of $\hat{p}_t(\lambda)$ as depending on the process s_t, which takes on the value 1 with probability λ and zero otherwise. Then

$$\hat{p}_t(\lambda) = \begin{cases} p_t^*(\bar{\lambda}) & \text{if} \quad s_t = 1 \\ E_{t-k} p_t^*(\bar{\lambda}) & \text{if} \quad s_{t-k+1}, \cdots, s_t = 0 \text{ and } s_{t-k} = 1 \end{cases} \tag{11}$$

Note also that $\hat{p}_t(\lambda)$ is firm specific.

The firm's loss function is taken to be the expected squared forecast error:

$$L(\lambda, \bar{\lambda}) = E\left(\hat{p}_t(\lambda) - p_t^*(\bar{\lambda})\right)^2. \tag{12}$$

This loss function is standard in statistical settings, but requires comment here. Private agents maximize utility by setting prices at the conditionally expected optimal level, given their information set. In principle, we could ask that agents choose the rate of information gathering λ also on the basis of expected utility maximization. Having agents instead minimize expected squared forecast error for prices has the advantage for us of technical simplicity, but it also has a natural interpretation in terms of bounded rationality. Agents are in effect splitting their decision problem into separate optimization and forecasting problems, a procedure that is often followed, for example, in the least-squares learning literature.

There is a further sense in which agents minimizing (12) are boundedly rational. In principle, agents might choose a time-varying rate of information gathering that depends on their information set. In endogenizing the rate of information acquisition, we are less demanding of our agents, but in a way that we find particularly plausible. Private agents are required to choose a rate λ that minimizes the unconditional mean squared forecast error, including costs of information acquisition, given the actual stationary price process. Such a choice could plausibly arise as the outcome of a stable adaptive learning process by comparing average mean squared errors for different rates.[17]

Noting that the mean of both $p_t^*(\bar{\lambda})$ and $\hat{p}_t(\lambda)$ is zero, we see that to compute the loss value, it is sufficient to compute the variance of $p_t^*(\bar{\lambda})$, $\hat{p}_t(\lambda)$ and their covariance. Using the equilibrium price paths for p and y together with (3) we obtain

$$p_t^*(\bar{\lambda}) = \sum_{j=0}^{\infty} \bar{\theta}_j \varepsilon_{t-j}, \tag{13}$$

where $\bar{\theta}_j = \bar{\phi}_j(1 - \alpha^2\omega) + \rho^j$ if $j > 0$ and $\bar{\theta}_0 = (1 - \alpha)\bar{\phi}_0 + 1$. We use the notation $\bar{\theta}$ and $\bar{\phi}$ to emphasize that these parameters depend on the economy-wide $\bar{\lambda}$. Now set

$$\Omega(k) = \sum_{j=k}^{\infty} \bar{\theta}_j \varepsilon_{t-j}.$$

[17](Reis 2004) develops the microfoundations of endogenous inattention and appears to provide a foundation for our simpler, tractable approach.

Note $\Omega(k) = E_{t-k} p_t^*(\bar{\lambda})$. Then $Var(p_t^*(\bar{\lambda})) = Var(\Omega(0))$ and

$$Var(\hat{p}_t(\lambda)) = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j Var(\Omega(j)).$$

Also, noting

$$\begin{aligned} Cov(p_t^*(\bar{\lambda}), \hat{p}_t(\lambda)) &= \lambda \sum_{j=0}^{\infty} (1-\lambda)^j Cov(\Omega(0), \Omega(j)) \\ &= \lambda \sum_{j=0}^{\infty} (1-\lambda)^j Var(\Omega(j)) = Var(\hat{p}_t(\lambda)) \end{aligned}$$

we get that

$$L(\lambda, \bar{\lambda}) = Var(p_t^*(\bar{\lambda})) - Var(\hat{p}_t(\lambda)). \tag{14}$$

Finally, setting $\bar{\psi}_k = \sum_{j=k}^{\infty} \bar{\theta}_j^2$ we conclude with the following result:

$$L(\lambda, \bar{\lambda}) = \sigma_\varepsilon^2 \left((1-\lambda)\bar{\psi}_0 - \lambda \sum_{j=1}^{\infty} (1-\lambda)^j \bar{\psi}_j \right). \tag{15}$$

Here σ_ε^2 is the variance of ε_t. Also, it is not difficult to show all infinite sums considered are absolutely convergent, so there are no existence issues. We have the following:

Lemma 1 *The function $L(\lambda, \bar{\lambda})$ is monotonically decreasing in λ.*

This lemma follows from Lemma 2 in Appendix A since the weights on the $\bar{\psi}_j$ sum to a constant not depending on λ.

If information gathering and processing were costless then the optimal choice would be $\lambda = 1$ so that the loss would be zero. BMR motivate sticky-information by a cost to information gathering. Along these lines assume that the cost to information gathering and processing is $C\lambda^2$ where $C \geq 0$. Define the function

$$T(\bar{\lambda}) = \arg \min_{0 \leq \lambda \leq 1} \left(L(\lambda, \bar{\lambda}) + C\lambda^2 \right).$$

$T(\bar{\lambda})$ is a best-response function: for fixed $\bar{\lambda}$ and resulting equilibrium processes, $T(\bar{\lambda})$ delivers an agent's optimal choice of λ. Existence of a solution to this optimization problem is guaranteed by the compactness of the choice set, and uniqueness can be demonstrated by directly computing that $\frac{\partial^2 \hat{L}}{\partial \lambda^2} > 0$, where $\hat{L} = L + C\lambda^2$: the proof of this is contained in Appendix A. A fixed point of this map is a symmetric Nash equilibrium and is our desired notion of Endogenous Inattention.

3 Existence and Comparative Static Analysis

The previous section showed that there exists a mapping from aggregate information flows, through a loss function defined by the associated equilibrium stochastic process, into an individual 'inattentiveness' rate.

Definition. *Endogenous Inattention is the symmetric Nash equilibrium defined by the fixed point* $\lambda^* = T(\lambda^*)$.

3.1 Existence Result

Note that $T : [0,1] \to [0,1]$. Moreover, from above, it is apparent that T is a well-defined and continuous function. From Brouwer's theorem we know that a fixed point exists. The value λ^* is a symmetric Nash equilibrium in λ, taking into account the policy reaction to aggregate λ. We summarize existence as a proposition.

Proposition 2 *Endogenous Inattention exists in the BMR model.*

Some comments are in order.

1. We will say that λ^* is a stable equilibrium if $T'(\lambda^*) < 1$ since in that case if $\bar{\lambda} \neq \lambda^*$ then (locally) an individual will have an incentive to adjust λ toward λ^*. Our focus is on equilibria that are stable, but below we will highlight existence of unstable equilibria as well.

2. An increase in λ^* results in an increase in price and output variance, which may yield increased incentive for a given agent to choose a higher λ. This potentially self-fulfilling behavior suggests that multiple Nash equilibria may be present, and indeed we will see that this can arise.

3. Raising ω, and thereby decreasing the equilibrium price variance, gives an individual agent the incentive to lower her choice of λ and thus potentially reduces output variance and further reduces price variance. The usual trade-off between the price and output volatility may therefore break down.

3.2 Comparative Static Analysis

Endogenous Inattention is a fixed point of the map T, and the fixed points of this mapping depend on the deeper parameters of the model $\alpha, \rho, C, \omega, \sigma_\varepsilon$. This subsection examines how the fixed points depend on these underlying parameters. In particular, we characterize the direction in which λ^* moves for infinitesimal changes in each parameter.

It is useful to rewrite the T-map to emphasize its dependence on model parameters. Denote $\xi = (\alpha, \rho, C, \omega, \sigma_\varepsilon^2)'$. We now define the T-map to be

$$T(\bar{\lambda}; \xi) = \arg\min_{\lambda} \left(L(\lambda, \bar{\lambda}; \xi) + C\lambda^2 \right).$$

Fixed points are $\lambda^* = T(\lambda^*; \xi)$. Comparative statics require computing, for each element of ξ,

$$(T' - 1)d\lambda^* + T_{\xi_i} d\xi_i = 0$$

where $T' \equiv \partial T / \partial \bar{\lambda}$, $T_{\xi_i} \equiv \partial T / \partial \xi_i$. As mentioned above we focus on stable equilibria so that $T' < 1$. In a neighborhood of a stable fixed point, the effect of a change in one of the parameters on the fixed point is determined by $sign(T_{\xi_i})$. In particular, $sign\left(\frac{d\lambda^*}{d\xi_i}\right) = sign(T_{\xi_i})$.[18] We have the following result:

Proposition 3 *Let $\lambda^* < 1$ denote a stable symmetric Nash equilibrium. Assume $\alpha \leq 1$. For $\xi = (\alpha, \rho, C, \omega, \sigma_\varepsilon^2)'$ the effect of a change in a component of ξ on λ^* is*

$$\frac{d\lambda^*}{dC} < 0, \frac{d\lambda^*}{d\rho} > 0, \frac{d\lambda^*}{d\alpha} < 0, \frac{d\lambda^*}{d\omega} < 0, \frac{d\lambda^*}{d\sigma_\varepsilon^2} > 0.$$

The proof is contained in Appendix A. We focus on $\alpha \leq 1$, in which pricing decisions are strategic complements, because this is the case examined in the literature,[19] but extending the analysis to $\alpha > 1$ would clearly be of theoretical interest. Proposition 3 provides comparative static results for interior endogenous inattention equilibria. If $\lambda^* = 1$ then the impact on the equilibrium inattention level will either be as given in the proposition or zero, depending on the sign of the change in the parameter and on whether the associated first order condition holds with equality. The intuition behind the proposition is given below, together with graphical representations of equilibria.

To illustrate the results of this proposition, and to elaborate on the existence of equilibria, we turn to a numerical analysis. We give a graphical representation of the results, in particular, to demonstrate the possibility of multiple equilibria. Although Proposition 3 gives analytical details on comparative statics, in the policy discussion below it will be useful to have greater intuition on the comparative statics of ω and C.

We plot the T-function for various parameter values. For a vector of parameter values $(C, \omega, \rho, \alpha, \sigma_\varepsilon^2)$, we plot an agent's optimal choice of λ given that all other agents choose $\bar{\lambda}$. A few brief comments about parameter values are warranted. First, as mentioned above, we treat ω as an exogenous policy parameter and use changes

[18] Using stability in this way is closely related to the observation made in (Evans and Honkapohja 2003b) in a different context.

[19] For example, in their numerical illustrations BMR set $\alpha = 0.1$.

in its value to study the impact of the changes in policy 'activism' recently detailed in (Orphanides and Williams 2003b). An alternative interpretation, if instead ω is regarded as a function of deeper preference parameters of the agents, is that one of those preference parameters has changed.[20] However, our preferred interpretation is to view changes in ω as reflecting changing priorities of policymakers. Second, our interest is not in calibration but in the implications of the model with Endogenous inattention.

In order to conduct the numerical analysis we need a baseline parameter valuation. Our baseline parameterization sets $\alpha = .1$, $\rho = .8$, $C = 5$, $\sigma_\varepsilon^2 = .1$.[21] We choose these values as the baseline because they deliver results suitable for comparative static analysis, i.e. intermediate and not extreme results. They are not baseline in the sense of being calibrated to actual data but are consistent with the values in BMR.

Figure 1 below graphs the T-map and resulting equilibria for the baseline calibration and $\omega = 20$. Recall that the T-map takes the aggregate attentiveness parameter and maps it into an individual choice of λ. Any point on this curve that crosses the 45-degree line is a Nash equilibrium. The various comparative static results of Proposition 3 are summarized in Figure 1, which shows the way in which the T-map is altered by changing one of the parameters of the model.

INSERT FIGURE 1 HERE

Figure 1 shows that multiple equilibria can exist, though here only one equilibrium is stable.[22] In all of our numerical calculations, only one stable interior equilibrium is observed. In the baseline case there are equilibria at about .11 and at 1. The equilibrium at .11 is stable since $T' < 1$. Note that in this case full rationality – in the sense of full information, i.e. $\lambda^* = 1$ – does constitute an equilibrium. As we will see below, it is not always the case that full rationality is an equilibrium. The existence of a full-information equilibrium even though it produces higher volatility may initially seem surprising, but the result is intuitive. If all agents respond fully to contemporaneous shocks then price and output volatility will be higher. The higher volatility here reinforces agents' decisions to coordinate on full-information, making the point an equilibrium. However, $\lambda^* = 1$ is not a stable equilibrium: for values $.11 < \bar{\lambda} < 1$ agents have an incentive to reduce λ. There are parameterizations, however, in which $\lambda^* = 1$ is the only stable equilibrium; see Section 4.1 below as an example.

Having established a baseline result, we turn to comparative statics. First we alter C while holding $\rho, \alpha, \omega, \sigma_\varepsilon^2$ fixed. Figure 2 plots T-maps for various values of C. The

[20] The parameter α is also a function of deeper parameters, but there are enough degrees of freedom so that α and ω can be chosen independently.

[21] BMR use $\alpha = .1$, $\rho = .8, \omega = 1$, and implicitly $\sigma_\varepsilon^2 = 1$.

[22] The possibility of multiple equilibria with low C is related to the presence of multiple equilibria in the degree of rigidity, found in the earlier literature on nominal rigidity and coordination failures. See (Ball and Romer 1991).

arrow indicates the direction of change in the graph of the T-map, given that C is increasing. The comparative static direction is intuitive, since the optimal choice of λ, for fixed $\bar{\lambda}$, will decrease as its cost increases.

INSERT FIGURE 2 HERE

The thick horizontal line at the top of the figure is a plot of the T-map when $C = 0$. In this case, $\lambda^* = 1$ is the unique equilibrium, and it is stable. This result is as expected since whenever the cost to acquiring and processing information is sufficiently low we should expect to see full-information rational expectations arise. Figure 2 also demonstrates that as the cost increases the possibility for multiple equilibria arises. Moreover, for medium-sized costs there exists a stable interior fixed point. Clearly, for a particular value of C it is possible to generate BMR's choice of $\lambda = .25$. For very low $C > 0$ there are two stable equilibria as well as an unstable equilibrium.[23] As C continues to rise, the full-information equilibrium disappears and the only equilibrium is the stable sticky information equilibrium.

Figure 3 plots the comparative statics for varying the policy parameter ω. According to Figure 3, for fixed $\bar{\lambda}$, as ω rises firms have less incentive to update their information since the higher ω is associated with a monetary policy that decreases price volatility and, as a result, reduces the value of new information.

INSERT FIGURE 3 HERE

Low values of ω imply a unique full information equilibrium. As ω increases, the full information equilibrium becomes unstable and a stable interior equilibrium emerges. As already shown, further increases in ω lead to lower rates of information processing.

3.3 Demand Shocks

¿From an aggregate welfare perspective, it is optimal for agents to coordinate on an information acquisition rate of zero. In this case, equilibrium price and output would be zero, so that the government's loss function, which approximates aggregate welfare, would be zero. The apparently paradoxical result that less information is welfare enhancing stems from the fact that the mark-up shocks, which are the only shocks in the model, are distortionary in nature and so, on aggregate, are best ignored.

It is natural, then, to wonder whether incorporating non-distortionary shocks into the model, say, in the form of demand shocks, would alter our results. To address

[23] Note that the unstable equilibrium has counterintuitive comparative statics. For example, higher costs will increase the value of λ^* in the unstable equilibrium.

this question, following BMR, we amend aggregate demand to include a white noise shock:

$$AD : y_t = \hat{m}_t - p_t + e_t. \tag{16}$$

In the presence of aggregate demand shocks, it is no longer socially optimal for agents to set $\lambda = 0$, for if they do, while the price level will be set to zero, output will follow a white noise process. It is welfare improving for agents to observe this process with positive frequency and therefore have price movements capture some of the variance in the economy.

With the demand curve (16) we can solve the model precisely as before, and, with one exception, obtain the comparative static results listed in Proposition 3: for details on how aggregate demand shocks impact the model, see Appendix B. The exception is the sign of $\partial\lambda^*/\partial\alpha$, which becomes ambiguous in the presence of demand shocks. We also find that $\partial\lambda^*/\partial\sigma_e^2 > 0$, and the intuition for this is straightforward: as the variance of the demand shocks increases, agents have increased incentive to gather information, thus the graph of the T-map is shifted upward.

Finally, the numerical analysis presented above is not qualitatively altered. Because of this, and to promote simplicity, for the remainder of the paper we assume that the demand shock is equal to zero.

4 Policy Implications

The previous section on comparative statics revealed that the number and nature of the equilibria in our model is strongly impacted by parameter values. We turn now to our central interest, which is how the relationship between output and price volatility depends, through endogenous changes in λ, on the activism of optimal policy. The framework in this paper is the first to allow for an equilibrium study of this issue. The novel implication of our approach is that policy 'activism' has both direct and indirect effects on unconditional price and output variance. Above we noted that the Bernanke Hypothesis is a conjecture on the tension between these effects. This Section examines this relationship.

4.1 Policy Implication Results

Result one of Proposition 1 obtained the usual trade-off between σ_p^2 and σ_y^2 in the BMR model with exogenous λ. Increasing ω leads policy to reduce price variation. Because λ has not changed, the real mark-up shocks are observed with the same regularity, and if prices do not move to accommodate them, then output must. Combining both results of Proposition 1 with the result for $\frac{d\lambda^*}{d\omega}$ in Proposition 3, indicates the potential shape of the trade-off in case of endogenous inattention. For an interior equilibrium we know that $\frac{d\lambda^*}{d\omega} < 0$. It is thus unambiguous that an increase in ω will reduce price volatility. However, while for fixed λ, increasing ω directly increases output volatility, raising ω indirectly decreases output volatility as a result of the equilibrium reduction in λ. Thus the effect of an increase in ω on output volatility in the case of endogenous inattention is ambiguous. The results of Proposition 1 and Proposition 3 therefore suggest that the usual trade-off between output and price volatility may not always obtain. In this section we investigate this issue numerically and show that it is indeed possible, over at least part of the range of ω, for the usual trade-off to disappear, and that decreased policy activism may lead to a decline in both price and output volatility. However, our numerical results also indicate that an output-price volatility trade-off will emerge for sufficiently high ω.

Policy in this model is pinned down by the bank's objective function. We alter policy by varying the relative weight ω in the central bank's preferences. For each chosen value of ω, we compute the unconditional equilibrium output and price variance and plot the relationship between σ_p^2 and σ_y^2. This relationship is a "policy frontier" in the sense that it describes the equilibrium outcome for each level of policy activism.

By way of comparison, we present the policy frontier first for the BMR model with exogenous λ and then for our model which endogenizes λ. We choose the parameters as $\alpha = .1$, $\rho = .85$, $C = 5$, $\sigma_\varepsilon^2 = .1$, which are close to our benchmark values.[24] Figure 4 sets $\lambda = .25$ and thus provides an illustration of the BMR model with exogenous λ. The figure contains four panels describing, for $3 \leq \omega \leq 30$, (clockwise, starting from the NW corner) the frontier, the exogenous value of λ, and the values of σ_y^2 and σ_p^2 as ω varies. The arrow indicates the direction of motion along the frontier as ω is increased. The downward sloping nature of the frontier represents the usual trade-off between output and price variance. As ω is increased, policy is chosen to reduce price variance, and the equilibrium response is to increase output variance.

INSERT FIGURE 4 HERE

In Figure 5 we consider the impact of increasing ω when λ is chosen endogenously as in our model. For each value of ω we compute the associated stable fixed point of the T-map and the resulting equilibrium variances. The frontier is described in the northwest panel of Figure 5. The arrow indicates the direction of motion along the frontier as $13 \leq \omega \leq 30$ is increased. For $\omega < 13$ the shape of the frontier becomes quite steep and so, except for the λ panel, we omit this range for clarity of presentation.[25]

INSERT FIGURE 5 HERE

[24] The value for α is the one used by BMR. The values of ρ and σ_ε^2 are chosen to roughly match observed values of σ_p^2 and σ_y^2 for our choice of C.

[25] Under this parameterization there exists a unique stable equilibrium.

Unlike when λ is fixed exogenously, the frontier in the case of endogenous inattention is non-monotonic and takes the shape of a 'nose'. The usual trade-off between price and output variance exists for sufficiently large ω but, most interestingly, the 'nose' implies that for some range of ω the output-price variance trade-off is eliminated. In particular, we find that in this range a decrease in activism reduces both output variance and price variance. When the policymaker's preferences shift toward lower activism, the unconditional variance of price will decline accordingly. For fixed λ, this would increase output volatility. However, the decrease in price level volatility lowers the firms' incentive to pay for information and decreases λ^*, as is seen in the northeast panel of Figure 5.[26]

The decrease in equilibrium λ^* associated with this range of ω acts to decrease output volatility. The northeast and southwest panels illustrate that for $13 \leq \omega \leq 39$ the indirect effect – whose strength is measured by the responsiveness of λ^* to changes in ω – is greater than the direct effect and so output variance falls sharply. As ω increases beyond 39, the associated point on the frontier moves onto the downward sloping portion corresponding to the usual trade-off. As the northeast panel clearly demonstrates this occurs when λ^* adjusts slowly to its lower bound. At this point, the direct effect of ω on output variance outweighs the indirect expectation formation effect; hence, the southwest panel indicates an increase in output variance. We conclude that by decreasing policy activism, the central bank may be able to *jointly* lower the volatility of the price level and output. The gain to a conservative central banker is not without bounds, however, as eventually a volatility trade-off emerges. Below we present further discussion of the implications of Figure 5 for government policy and social welfare.

The intuition behind the results above suggest that, depending on the responsiveness of λ^* to changes in ω, the slope of the frontier could be positive or negative. While the frontier is upward sloping for $C/\sigma_\varepsilon^2 = 50$ and sufficiently small ω, for a sufficiently high C agents will be less likely to increase their rate of information accrual and the frontier will be everywhere downward sloping. This conjecture is verified in Figure 6 which takes the same parameter values as Figure 5 except that it sets $C/\sigma_\varepsilon^2 = 200$.[27] By increasing the relative costs of updating by a factor of four, the usual trade-off exists over the entire range.[28] Figure 6 illustrates that if the marginal cost of information acquisition increases sufficiently rapidly in λ then the results are close to the BMR case of exogenous λ.

INSERT FIGURE 6 HERE

[26]The results of this section, and in particular, the presence of a 'nose-shaped' frontier, are not qualitatively altered by assuming non-zero aggregate demand shocks.

[27]In the figure we have scaled up σ_ε^2 in order to roughly match observed price and output variances.

[28]In this case there are two stable equilibria for low values of ω: $\lambda^* = 1$ and $0 < \lambda^* < 1$. In Figure 6 plot the results for the choice of the stable interior equilibrium. Choosing $\lambda^* < 1$ is in the spirit of BMR and, thereby, appropriate for examining the policy implications of endogenous inattention.

Non-monotonic policy frontiers exist also in (Orphanides and Williams 2003a). In their model, private agents forecast inflation using a constant gain version of recursive least squares (RLS). The constant gain learning produces greater persistence to exogenous shocks. (Orphanides and Williams 2003a) study the implications of this greater persistence for the conduct of optimal monetary policy. They find that optimal policy should be more vigilant against inflation when agents engage in least-squares learning. In our model, the response to mark-up shocks depends on the equilibrium value of λ which depends on the activism of policy. More active policy lowers the optimal attentiveness of agents and consequently can lower economic volatility as observed in the 'Great Moderation'. However, our results in Figure 5 caution policymakers that there may be a limit to the reduced output volatility, resulting from heightened vigilance against price volatility, since eventually a trade-off may emerge.

The key intuition to this cautionary insight is the effect ω has on the equilibrium value of λ^*. Successively higher values of ω will decrease λ^*, as detailed in Proposition 3. It can be shown that, as $\omega \to \infty$, $\lambda^*(\omega)$ converges to a positive value. One might therefore expect that the direct effect of ω on σ_y^2 will dominate for sufficiently large ω, leading to an eventual trade-off.[29]

4.2 Discussion of Policy Implications

The results illustrated in Figures 5-6 are new and important. Previous work on optimal monetary policy has either not taken into account the costs of processing and collecting information or has ignored the endogenous feedback between policy and the degree of inattention. Our results show that if the policy authority decreases its output activism it induces agents to reduce the rate at which they gather new information. This has the effect of lowering the unconditional variance of the economy. This result is at variance with what is generally found in the literature, but is consistent with the empirical evidence of the 'Great Moderation.'

Whether the policy frontier is upward or downward sloping depends crucially on the costs of updating and processing information. We have shown that for relatively low costs the usual trade-off between price level and output volatility gives way to an upward sloping frontier over a wide range of the policy parameter ω. However, for sufficiently high ω it appears the policy frontier is always eventually downward sloping.

It may appear odd that the mechanism for the reduction in economic volatility is a reduction in the rate of information acquisition by private agents. Intuitively, one might expect a higher rate of information gathering to be socially optimal. This

[29] Our numerical investigations suggest that there is always a trade-off for sufficiently high ω (at least for $\alpha < 1$). Investigating this issue theoretically is not straightforward since $\partial\sigma_y^2/\partial\omega$ vanishes as $\omega \to \infty$.

is not the case in our set-up since the mark-up shocks are distortionary. It can therefore be welfare improving to reduce the effect of these shocks on the pricing and output decisions of firms. By reducing the price volatility associated with these shocks, private agents are induced to reduce their intensity of information acquisition and diminish their response to the distortionary shocks.[30]

Our finding that a stronger response to markup shocks not only lowers price variance, but also provides an incentive for agents to update their information less often, is related to (Svensson's 2003) hypothesis about inflation targeting. Svensson argues that by targeting an inflation rate agents' expectations will be anchored and economic volatility reduced. In our model, the policymaker becomes less 'active' and as a result the *equilibrium* outcome is that agents' expectations are anchored. This is an intuitively appealing result as it supports the inflation/price targeting hypothesis by exhibiting it as the equilibrium response in a model with information updating costs.

We note that it is not the case that the policy frontier is analogous to a production possibilities frontier or a budget constraint. The points on the frontier are equilibrium outcomes resulting from the joint determination of optimal monetary policy and Endogenous Inattention. The possibility of a positively sloped policy frontier does, however, raise the possibility that there may be gains to commitment analogous to the gains in other set-ups from appointing a conservative central banker.

To pursue this line of thought, imagine that the government evaluates outcomes according to the loss function in (5) with weighting parameter ω^* which is not necessarily equal to the parameter ω used to set policy. In other words, the government hires a central banker with activism parameter ω so that the resulting equilibrium outcomes $\sigma_y^2, Var(p_i - p)$ minimize their loss with preferences ω^*. Is appointing a central banker with $\omega > \omega^*$ socially preferable? For the economy illustrated in Figure 5, for all realistic ω^* there is an unambiguous welfare gain to choosing a central banker that moves along the policy frontier, past the 'nose', and onto the usual trade-off portion of the curve. In this case the loss-minimizing policy parameter ω is greater than ω^*. This example suggests that appointing a more conservative central banker and placing the economy along the usual trade-off is socially optimal.[31]

A conservative bias is not a fully general result, however. In Figure 6 there is always a trade-off between output and price volatility. In this case, the socially optimal point on the frontier depends critically on ω^*. For $\omega^* < \hat{\omega} \backsimeq 20$, our numerical results indicate that the government should choose a more conservative (i.e. $\omega >$

[30]As discussed in Section 3.3, the inclusion of aggregate demand shocks would modify this result. Introducing idiosyncratic shocks would also give firms have a motive to gather information that is socially beneficial.

[31]We remark that the government's loss function could be adjusted to include costs of information gathering by private agents. This would strengthen the argument for a conservative central banker (and would weaken the counter-example given in the following paragraph).

ω^*) central banker, while for $\omega^* > \hat{\omega}$ the government benefits by choosing a less conservative central banker. Since this issue is not central to the current paper we reserve further investigation for future work.

5 Adaptive Learning and the Great Moderation

As noted in the introduction, the result of the policy experiment discussed in the previous section depends on the timing of the "game" between policymakers and private agents. The structure of our model assumes a simultaneous move game with ω parameterizing the preferences of the government. This timing assumption results in a prisoner's dilemma and the economy can be trapped in an inefficient outcome. The Great Moderation obtains in this game given an exogenous increase in preferences ω. As an alternative, we could specify the structure as a Stackleberg game with the government as the large player who moves first. In such a setting, policymakers could announce a policy consistent with preferences less activist than their own, and thereby choose their preferred point on the frontier. The Great Moderation could then be explained by assuming that until 1983 policymakers believed λ was exogenous, as in the BMR model, and that after 1983 the government became aware of the endogeneity of λ and exercised its first-mover status.

While both of these timing structures are consistent with the model in this paper, and both are capable in theory of explaining the Great Moderation, they are clearly implausible as an historical account, since they require not only that our extension of BMR provides a valid model of the economy, but also that policymakers understood the BMR model, in either its original form or our extension, decades before the BMR model was published, and followed the optimal policy rule dictated by the model.

Our own view is that it is more plausible to extend the bounded rationality viewpoint to policymakers, as well as private agents, and to think in terms of an evolution and improvement over time in the exercise of monetary policy. This view is in line with Bernanke's hypothesis and seems implicit in the discussions of monetary policy in both Svensson (2003) and McCallum (2000), as well as the earlier cited papers providing evidence of a change in the stance of US monetary policy from the 1980s. In this Section we will therefore assume that policymakers over time arrive at the view that decreased policy activism can improve economic performance, without their necessarily fully understanding all of the relevant mechanisms.

In this setting, we assume that policymakers choose their monetary policy rule to take the same form as that given by BMR's first-order condition for optimal policy (6), namely

$$E_{t-1}y_t = -kE_{t-1}p_t. \tag{17}$$

In BMR optimal policy is given by $k = \alpha\omega$ but we can also treat k, in a more boundedly rational (and arguably a more historically plausible) way, as an (inverse) measure of the degree of activism believed appropriate by policymakers. We then study the impact of an increase in k, representing the decline in activism starting in the 1980s.

We introduce at the same time an adaptive learning framework for both policymakers and private agents that is consistent with agents acquiring rational expectations, but which does not require them to know the full structural model and all of its parameters. More specifically we assume that policymakers use least squares to compute the forecasts required for implementation of the policy and that private agents use least-squares updating to compute the forecasts required for their decisions. Thus the expectations of both policymakers and private agents are in this Section modeled using the adaptive learning approach described in (Evans and Honkapohja, 2001, 2003a). Least squares learning allows policymakers and private agents to learn how to make optimal forecasts, given their information sets, without knowing structural parameters, and also allows them to appropriately track structural change.

This approach makes policy and λ time-dependent. A natural question is: to what type of equilibrium will this adaptive version of the economy converge (if any)? If, after removing the strategic interaction of the model, the economy converges to the Nash equilibrium/endogenous inattention outcome, then this provides additional support for our model. This approach also allows us to consider the Great Moderation in terms of stability under adaptation.[32] If the relevant equilibria are stable then an exogenous change in policymaker preferences could cause the economy to move to a lower point on the upward sloping section of the policy frontier, thus resulting in reduced volatility in both prices and output.

5.1 Real-time Learning Version of the Model

We now develop in detail the adaptive learning version of the model just described. We begin by describing the behavior of each of the three types of agents: policymakers; private-sector firms; and consultants.

5.1.1 Policymakers

In BMR, and in our extension, it is assumed that the government minimizes a loss-function given the laws of motion for output and the price-level. As just discussed, BMR show that optimal policy must satisfy the following first-order condition (17) with $k = \alpha\omega$. As noted by BMR, the condition (17) is also consistent with the optimal policy condition based on the Calvo model of policy adjustment, and with the "elastic

[32]We now use the term "stability" to refer to the *a priori* stronger condition that the economy under adaptation converges to the Nash equilibria.

price standard" rule proposed by Hall (1984). Thus this characterization of policy has an appeal that goes well beyond the specific model at hand.[33]

Furthermore, given the simple aggregate demand structure assumed, policy taking the form (17) can be equivalently described by the rule

$$\hat{m}_t = (1-k)E_{t-1}p_t, \tag{18}$$

for setting the policy instrument $\hat{m}_t$, at time $t-1$. Such a rule is close to the one studied, for example, by (Taylor 1980), who refers to $1-k$ as the "degree of accommodation" (to price shocks). Here $k>0$ but $1-k$ can be positive or negative and, of course, the case $1-k=0$ corresponds to rules based on fixed money supply targets. In the numerical simulations below it is convenient to report the effects of the policy shift in terms of an increase in ω, but the policy change can equivalently be interpreted in terms of a reduction in activism or in the degree of accommodation.

In this section it is convenient to assume that policymakers aim to implement (18). Implementation of this rule still requires forecasts of prices. Since we do not want to assume full knowledge of the structure by policymakers we replace $E_{t-1}p_t$ by an econometric forecast $\hat{E}_{t-1}p_t$ based on a reduced form time-series model.[34] In equilibrium, the price process is $MA(\infty)$ and it is natural to assume that policymakers approximate this process using an ARMA(r,q) specification. In addition we assume that the exogenous shocks ε_t are observable at t, so that policymakers can use recursive least squares (RLS) to update the estimates of their ARMA model's parameters.[35] Policymakers thus set $\hat{m}_t$ according to (18) with $E_{t-1}p_t$ replaced by $\hat{E}_{t-1}p_t$.

A couple of comments are warranted. Since the ARMA specification is an approximation, the model will, at best, converge to an approximate equilibrium. Also, since policy is set at time $t-1$, as in BMR, policymakers are unable to incorporate the time t shock ε_t into $\hat{m}_t$. Since in BMR, and in our model above, some firms do observe ε_t at t, we allow our consultants, below, to observe contemporaneous ε_t.

5.1.2 Firms

As before, firms are price setters and would prefer to set price to

$$p_t^* = p_t + \alpha y_t + u_t$$

[33]The condition (17) is a "specific targeting rule" of the type advocated and discussed in detail by (Svensson 2003). As stressed by (Svensson 2003), one of the advantages of this type of rule is that its specification does not require knowledge of the full structure of the economy.

[34]Other implementations of bounded rationality are possible in which policymakers make use of their knowledge of the structure. For a discussion of optimal monetary policy with structural parameter learning see (Evans and Honkapohja 2003a). The key qualitative results of the current paper are unlikely to depend on the detailed implementation of learning.

[35]See (Evans and Honkapohja 2001) for a detailed discussion of least-squares learning in dynamic macroeconomics.

each period, but, as before, there is a cost to processing new information. We assume that firms do not know the full economic structure and are thus unable to form fully rational expectations or to compute the optimal λ, given their costs. Instead, firms hire consultants to provide real-time estimates of both optimal price forecasts and of λ_t, given the costs to the firm of updating prices at frequency λ. We think this set-up is a reasonable stylized description of actual agent behavior. For example, (Carroll 2003) provides evidence that consumer expectations follow a distributed lag of professional forecasters.

We emphasize that, as in BMR, we interpret these costs broadly, not as just the literal cost of obtaining information, but rather as including all the internal costs of processing and utilizing the information. Note that the estimates are interdependent over time, i.e. the estimate of the optimal price process parameters will feed back into the optimal choice of λ and vice-versa.[36] For simplicity we assume that all agents learn in precisely the same way: consultants provide identical forecasts of p_t^* and the optimal λ_t to each firm.

5.1.3 Consultants

Consultants act as information gatherers, providing to firms forecasts of future optimal prices as well as the optimal rate of information processing. We assume that consultants, like the policymakers, do not know the full structure of the economy. Each period consultants forecast the value of p_t^* using an ARMA(r, q) specification, with ε_t observable.[37] As before, the ARMA(r, q) may be estimated using RLS.

We can think of consultants either as private organizations or as public servants, using the most recent information to provide firms with the optimal pattern of behavior for given information processing costs.[38] In particular, consultants are willing to provide $\hat{E}_t p_{t+k}^*$, for $k = 0, 1, 2, \ldots$, either free of charge or for a fixed fee willingly paid by all firms; however, the consultants are aware that firms incur a cost of information processing. The consultants provide the additional service of computing the optimal rate of information accrual, λ_t, by solving the firm's optimization problem. Recall, that in Section 2.2, firms computed their mean-square forecast error given the $\bar{\lambda}$ chosen by all other agents and given optimal monetary policy. Computing the associated loss function assumed that firms are able to compute the unconditional

[36] The simultaneous estimation of forecasting model and real-time choice of forecasting model was considered in a different context by (Branch and Evans 2004).

[37] We could instead assume that policymakers forecast with an ARMA(r', q'). with (r', q') possibly different from (r, q). However, this would not change the results below. Similarly, we could instead have the consultants forecast p_t, y_t and u_t separately ,and then combine them to construct the forecast of p_t^*. The impact on our results of this alternative set-up would be minimal.

[38] The notion of a consultant is a descriptive device designed to remove the explicit strategic interaction between agents. Some of the roles of consultant could be served by newspapers, business publications or the forecasting community.

variance of their optimal price-level. In a real-time setting this involves knowing the actual price process, and keeping memory of all past rates of information accrual $\{\lambda_t\}$. However, we assume that none of the agents know the structural equations. So instead, the consultants compute an estimate of the mean-square forecast error given their estimated ARMA(r,q) process for p_t^*. This can be done by following the same steps as before, but with the $\bar{\theta}_j$ given by the consultants' time $t-1$ estimate of the distribution of p^*. Specifically, if the time $t-1$ estimate of p_t^* is given by $\Psi_{t-1}(L)p_t^* = \Phi_{t-1}(L)\varepsilon_t$ then the $\bar{\theta}_j$ are determined by the polynomial division

$$\frac{\Phi_{t-1}(L)}{\Psi_{t-1}(L)} = \sum_{j=0}^{\infty} \bar{\theta}_j L^j.$$

The consultants then compute the λ_t that will minimize the usual loss function given this estimated sequence of $\bar{\theta}_j$.

As mentioned above, exact convergence to a stable Nash equilibrium of the non-adaptive model would require that the ARMA processes for both the consultants and the policymakers to be exact. Whether the equilibrium price process in Nash equilibrium is representable as an ARMA solution is unknown, so if we find that λ_t numerically converges to a particular value, the most we can really say is that value is an approximate equilibrium. However, we will see that numerically it appears to converge almost exactly to the Nash equilibrium. Also, although the consultants know the value of λ_t and have memory of the conditional forecasts $\hat{E}_{t-j}p_t^*$, the consultants do not know the full structural equations and so do not know how this translates into actual prices and, hence, actual optimal prices. This learning set-up is constructed specifically so that none of the agents know how λ_t affects the actual dynamics. Convergence to a Nash Equilibrium then provides strong support for our equilibrium concept and 'Great Moderation' explanation.

We remark that other interpretations of our dynamic system under learning are possible. For example, we could assume that firms use the consultants' estimate of the p_t^* process to compute λ_t at each t, provided the costs to the firm of this computation are regarded as negligible relative to the costs of processing and utilizing the information about p_t^*. Similarly, the role of the consultants in computing econometric forecasts could instead be provided by Central Bank forecasts of prices and output or by the "forecasting community" more generally.

5.1.4 Dynamic System

The following system, written in recursive causal ordering, describes the evolution of the economy under adaptive learning (and summarizes the

$$\hat{E}_{t-1}p_t = \{\text{ ARMA(r,q) Policy Maker Forecast}\}$$

$$\hat{m}_t = (1-k)\hat{E}_{t-1}p_t, \text{ where } k = \alpha\omega.$$
$$\hat{E}_t p^*_{t+k},\ k = 0,1,\ldots = \{\text{ ARMA(r, q) Consultant Forecast}\}$$
$$\lambda_t = \{\text{ Consultant Computed}\}$$
$$p_t = \sum_{j=0}^{\infty} \lambda_{t-j} \prod_{i=0}^{j-1} (1-\lambda_{t-i})\hat{E}_{t-j}p^*_t$$
$$p^*_t = \alpha\hat{m}_t + (1-\alpha)p_t + u_t$$

where the last equation is obtained using the AD relation and the definition of p^*_t.

We now address two questions:

1. Will this economy converge to the equilibrium associated with a stable Nash equilibrium of the non-adaptive model?
2. Suppose that ω increases exogenously. Will the economy converge to a new, more "moderate" Nash equilibrium and thus reproduce the Great Moderation?

5.2 Numerical Results

In this section, we present numerical results from simulations of the dynamic system presented in the previous subsection. Our numerical procedure is straightforward. For any given simulation, we run the model for up to 2200 periods, with an initial transient segment of 500 periods. At the beginning of each simulation we initialize the vectors of ARMA parameters by randomly drawing from a uniform distribution. We also randomly draw the initial conditions for p_t, p^*_t, λ_t.[39] The simulation then follows the recursive ordering above.

5.2.1 Stability of Endogenous Inattention

To check the stability of the Endogenous Inattention we set the model parameters to the two cases presented in Section 3. We first start with what we termed the benchmark case. We set $\alpha = .1, \sigma^2_\varepsilon = .1, C = 5, \rho = .85$. We set the ARMA parameters to $r = 1, q = 5$ as these provide a good approximation to the actual stochastic process. As we saw in Figure 5, this case induces the 'nose'-shaped frontier. Figure 7 illustrates the results from a typical simulation when $\omega = 15$.

INSERT FIGURE 7 HERE

[39]The initial conditions for these variables are vectors of length $r+q+1$.

As indicated by Figure 7, λ_t converges to its Nash equilibrium value, marked by the horizontal line in the top panel. Notice that for the purposes of illustrating convergence, the values of λ_t are plotted during the transient period. In the bottom two panels, the time t estimates of the unconditional variances of price and output are plotted. These estimates were obtained using a moving average with window length 500; thus the horizontal scales in these figures do not include the transient period. The horizontal lines in these panels correspond to the theoretical variances of output and price at the associated Nash equilibrium.

The results of Figure 7 strongly suggest that the Nash outcome is stable under our adaptive model. The intuition for this stability is as follows. As was previously mentioned, for fixed λ, the ARMA models are approximations to the true MA(∞) equilibrium price process. Then, in a sense, the ARMA model is an underparameterized forecasting model. Since the true process depends on the underparameterized ARMA models– through policy and λ_t– the equilibrium here is closely related to the Restricted Perceptions Equilibrium (RPE) defined in (Evans and Honkapohja 2001). Moreover, the RPE in models with an expectational structure similar to the one presented here are stable under adaptive learning. Further, for a fixed price process we restrict attention to Nash equilibria which are stable fixed points of our T-map. When combined, it is not surprising (though not obvious) that these two stable mechanisms imply convergence.

5.2.2 The Great Moderation in Real-Time

We now turn to examining the Great Moderation in real-time. Figure 6 illustrated the possibility of a Great Moderation as the resolution of a tension between the direct and indirect effect of changes in policy activism, ω. As policy becomes less activist, there is a tendency for output variance to increase and equilibrium attentiveness λ to decrease, which induces lower output variance. We showed that along the 'nose' it is possible for the indirect effect to outweigh the direct effect and so that declines in policy activism could lead to decreases in both output and price volatility. We now examine this hypothesis under real-time learning by running simulations as above but assuming that during the simulation there is an exogenous increase in ω. This increase in ω could be due a shift in policy stance accompanying the appointment of a conservative central banker, which either can be thought of as exogenous or as response by the government to a series of adverse price shocks.

Figure 8 illustrates the results of this experiment. Initially, (after a transient period of length 500), the economy is near the equilibrium corresponding to $\omega = 15$. At time $t = 800$, ω increases abruptly from 15 to 30.[40] Immediately following this change, price volatility plummets as predicted, but output volatility rises. This

[40]To compute price and output variance after the shock to ω, the window length is shorted to 20, and then allowed to rise back to 500.

reflects the fact that λ_t is falling from its pre-shock level, but has not yet reached its new equilibrium level; thus, temporarily, the usual trade-off exists. As λ_t gets close to its new equilibrium level, both volatility time-series converge to levels lower than those of the pre-shock equilibrium, thus representing a Great Moderation.

INSERT FIGURE 8 HERE

We emphasize that we have not attempted to calibrate our model to provide a description of the actual historical experience. Any serious exercise along these lines would require significantly more elaborate detail for both the aggregate demand and the aggregate supply sides of the model. Nonetheless, the finding that a Great Moderation is possible in the model with a boundedly rational policymaker and adaptive learning by all agents is significant.

A principal result of this paper is that whether or not a decline in policy activism leads to lower output volatility depends on the joint equilibrium determination of optimal policy and attentiveness. That the dynamics in an adaptive learning version of the model converge to Endogenous Inattention equilibria, and can exhibit the Great Moderation in real-time, supports our equilibrium interpretation of the decline in output and price volatility. The additional finding, that in the adaptive learning formulation there is a transitional period of higher output volatility, in our view strengthens the plausibility of this account.

6 Conclusion

This paper has studied the implications for monetary policy of an economy in which agents endogenously choose the rate at which they update their information. Following (Ball, Mankiw, and Reis 2003) we assume that it is costly for agents to update their information sets each period. We extend their model, however, by explicitly modeling the choice of the rate at which they acquire information. We assume that agents choose the frequency with which they update their information sets by minimizing a quadratic loss function that depends on the costs of updating and forecast errors. The aggregate rate at which agents update their information is determined in a Nash equilibrium, among the private agents as well as the policymaker, in which policy is set optimally given the equilibrium rate.

We characterize the set of equilibria and use an adaptive version of this framework to provide an equilibrium explanation for the Great Moderation. (Bernanke 2004) conjectures that a fundamental shift in Federal Reserve objectives led to an anchoring of expectations and a reduction in economic volatility. This paper provides a systematic account of this hypothesis. A primary insight of this paper is to elucidate the important interactions between monetary policy and the degree of private agent attentiveness, which in turn determines the relationship between price and output volatility.

Previous studies have emphasized that price and output variance move in opposite directions when a policymaker becomes less activist. We argue that these results are inconsistent with the empirical evidence that as the Federal Reserve became more aggressive in fighting inflation both output and price volatility declined. Our model can explain these features of the data by showing that the reduction in price volatility can make it unnecessary for agents to update information as quickly, leading in turn to a reduction in output volatility. This finding is closely related to the idea of inflation targeting advocated by (Svensson 2003) and others. At the same time, we show that there is a tension between the direct effect of a policy rule and its indirect effect on the equilibrium attentiveness of agents. At a sufficiently low level of activism, the direct effect can dominate so that a volatility trade-off reappears.

Appendix A

Proof of Proposition 1. These results follow immediately from the equilibrium descriptions of the price and output processes (7), (8), and from the definitions of ϕ and φ, with the exception of the result concerning cross-sectional variance. Here, we require a result from BMR; they determine that

$$Var_i(p - p_i) = \sum_{j \geq 1} \eta_j \left(p_t - E_{t-j}(p_t)\right)^2$$

where

$$\eta_j = \frac{\lambda(1-\lambda)^j}{(1-(1-\lambda)^j)(1-(1-\lambda)^{j+1})}.$$

Substituting into this expression the equilibrium price path in (7), it follows that

$$Var_i(p - p_i) = \sum_{j \geq 1} \eta_j \left(\sum_{k=0}^{j-1} \phi_k \varepsilon_{t-k} \right)^2$$

Taking unconditional expectations leads to,

$$EVar_i(p - p_i) = \sigma_\varepsilon^2 \sum_{j \geq 1} \eta_j \hat{\phi}_j,$$

where

$$\hat{\phi}_j = \sum_{k=0}^{j-1} \phi_k^2.$$

The result then follows from the fact that $\frac{\partial \phi_k}{\partial \omega} < 0$.

Proof of Proposition 3. Recall $T(\bar{\lambda}, \xi) = \arg\min_\lambda \hat{L}(\lambda, \bar{\lambda}, \xi)$, where $\hat{L} = L + C\lambda^2$, ξ is the vector of model parameters and $\bar{\lambda}$ is the economy wide value of λ, which is taken as given by individual agents. The equilibrium λ^* is defined by $T(\lambda^*, \xi) = \lambda^*$, so that by the implicit function theorem,

$$\frac{\partial \lambda^*}{\partial \xi_i} = \frac{T_{\xi_i}}{1 - T_{\bar{\lambda}}}.$$

Stability then implies that $\text{sign}\left(\frac{\partial \lambda^*}{\partial \xi_i}\right) = \text{sign}\left(T_{\xi_i}\right)$. To compute T_{ξ_i}, we note that T is defined by the first order condition $\hat{L}_\lambda(T(\bar{\lambda}, \xi), \bar{\lambda}, \xi) = 0$. Again we may apply the implicit function theorem to obtain

$$T_{\xi_i} = -\frac{\hat{L}_{\lambda \xi_i}}{\hat{L}_{\lambda\lambda}}. \tag{19}$$

We will show that $\hat{L}_{\lambda\lambda} > 0$, so that $\text{sign}\left(T_{\xi_i}\right) = -\text{sign}\left(\hat{L}_{\lambda\xi_i}\right)$. Thus it remains to compute the relevant second partials of $\hat{L}$.

We require the following result:

Lemma 2 *Suppose β_i is a decreasing positive sequence and for each real number ν, $\gamma_i(\nu)$ is a sequence with $\sum \gamma_i(\nu) = M$, and $V(\nu) = \sum \gamma_i(\nu)\beta_i < \infty$. If there exists $N(\nu)$ so that $\frac{\partial \gamma_i}{\partial \nu} > 0 \Leftrightarrow i < N(\nu)$ then $V_\nu > 0$.*

Proof. The idea is simple: increase the values of γ_i corresponding to larger weights, and decrease the values corresponding to lower weights. Formally, we have

$$\begin{aligned} V_\nu &= \sum_{i\in\mathbb{N}} \frac{\partial \gamma_i}{\partial \nu}\beta_i = \sum_{i<N(\nu)} \frac{\partial \gamma_i}{\partial \nu}\beta_i + \sum_{i\geq N(\nu)} \frac{\partial \gamma_i}{\partial \nu}\beta_i \\ &> \sum_{i<N(\nu)} \frac{\partial \gamma_i}{\partial \nu}(\beta_i - \beta_{N(\nu)}) + \beta_{N(\nu)} \sum_{i\in\mathbb{N}} \frac{\partial \gamma_i}{\partial \nu} \\ &= \sum_{i<N(\nu)} \frac{\partial \gamma_i}{\partial \nu}(\beta_i - \beta_{N(\nu)}) > 0, \end{aligned}$$

where the last *equality* follows from the fact that $\sum \gamma_i(\nu) = M$ implies the sum of partials equals zero. ■

Now define the following notation:

$$f(\lambda, j) = \frac{(1-\lambda)^{j+1}}{1-(1-\lambda)^{j+1}} \quad \text{and} \quad g(\lambda, j) = \lambda(1-\lambda)^j.$$

Then

$$\bar{\theta}_j = \begin{cases} \frac{1}{1-(1-\alpha)\bar{\lambda}} & j = 0 \\ \frac{(1+f(\bar{\lambda},j))\rho^j}{\alpha^2\omega + f(\bar{\lambda},j)} & j > 0 \end{cases}.$$

and

$$\hat{L} = \sigma_\varepsilon^2 \bar{\psi}_0 - \sigma_\varepsilon^2 \sum_{j=0}^{\infty} g(j,\lambda)\bar{\psi}_j + C\lambda^2.$$

The partials we are to compute are then given by

$$\begin{aligned} \hat{L}_\lambda &= -\sigma_\varepsilon^2 \sum_{j=0}^{\infty} g_\lambda \bar{\psi}_j + 2C\lambda \\ \hat{L}_{\lambda\lambda} &= -\sigma_\varepsilon^2 \sum_{j=0}^{\infty} g_{\lambda\lambda} \bar{\psi}_j + 2C, \\ \hat{L}_{\lambda\xi} &= -\sigma_\varepsilon^2 \sum_{j=0}^{\infty} g_\lambda \frac{\partial \bar{\psi}_j}{\partial \xi_i}. \end{aligned}$$

We now proceed to prove the proposition in a series of steps.

Step 1. $\hat{L}_{\lambda\lambda} > 0$.

First notice that for all λ, $\sum g(\lambda, i) = 1$ so that $\sum g_\lambda(\lambda, i) = 0$. We may compute

$$\begin{aligned} g_\lambda &= (1-\lambda)^{j-1}(1-(j+1)\lambda) \\ g_{\lambda\lambda} &= -(1+j)(1-\lambda)^{j-1} - (j-1)(1-\lambda)^{j-2}(1-(j+1)\lambda). \end{aligned} \tag{20}$$

We find that

$$g_{\lambda\lambda} < 0 \Leftrightarrow \frac{j+1}{j-1} > \frac{(1+j)\lambda - 1}{1-\lambda},$$

thus implying the existence of $N(\lambda)$ so that $j < N(\lambda) \Leftrightarrow g_{\lambda\lambda} < 0$. Applying the lemma with $V = \hat{L}_\lambda - 2C\lambda$ and $\gamma_i(\nu) = -g_\lambda(\lambda, i)$ yields the result.[41]

Before moving on to the remaining steps, we show the following:

$$\text{sign}\left(\hat{L}_{\lambda\xi_i}\right) = -\text{sign}\left(\frac{\partial\bar{\theta}_j}{\partial\xi_i}\right), \tag{21}$$

provided the sign of $\frac{\partial\bar{\theta}_j}{\partial\xi_i}$ is independent of j and $\alpha \leq 1$. [42] Indeed, notice

$$\frac{\partial\bar{\psi}_k}{\partial\xi_j} = \sum_{j\geq k} 2\bar{\theta}_j \frac{\partial\bar{\theta}_j}{\partial\xi_i},$$

Assume for the moment that $\frac{\partial\bar{\theta}_j}{\partial\xi_i} < 0$. Then $\beta_j \equiv -\frac{\partial\bar{\psi}_k}{\partial\xi_i}$ form a decreasing positive sequence. Also notice that, from (20), there is a $M(\lambda)$ so that $g_\lambda(\lambda, i) > 0 \Leftrightarrow i < M(\lambda)$. Thus we may apply the Lemma above to $\hat{L}_{\xi_j} = \sum g(\lambda, j)\beta_j$ to get $\hat{L}_{\xi_j\lambda} > 0$. A similar argument applies in case $\frac{\partial\bar{\theta}_j}{\partial\xi_i} > 0$.

To complete the proof of the proposition, we simply compute the sign of $\frac{\partial\bar{\theta}_j}{\partial\xi_i}$, and then appeal to (21).

Step 2. $\hat{L}_{\lambda\alpha} > 0$.

For $j = 0$ computing the sign of $\frac{\partial\bar{\theta}_j}{\partial\alpha}$ to be negative is straightforward. Let $B_j = \alpha^2\omega + f(\bar{\lambda}, j)$. For $j > 0$ we compute

$$\begin{aligned} \frac{\partial\bar{\phi}_j}{\partial\alpha} &= \frac{-2\rho^j\alpha\omega}{B_j^2} < 0 \\ \frac{\partial\theta_j^*}{\partial\alpha} &= -2\alpha\omega\bar{\phi}_j + (1-\alpha^2\omega)\frac{\partial\bar{\phi}_j}{\partial\alpha}. \end{aligned}$$

[41] Note that $g_{\lambda\lambda}(\lambda, 0) = 0$, so that the premise of the Lemma is not precisely met. However, it is trivial to modify the proof of the Lemma to account for this minor generalization: just have the premise read $i < N(\nu) \Rightarrow \frac{\partial\gamma_i}{\partial\nu} \geq 0$ with at least one strict inequality, and $i \geq N(\nu) \Rightarrow \frac{\partial\gamma_i}{\partial\nu} \leq 0$, and notice the proof goes through unchanged.

[42] It may be the case that $\frac{\partial\bar{\theta}_0}{\partial\xi_i} = 0$, but this does not impact the result.

Combining these two equations with the definition of $\bar{\phi}_j$ in terms of B_j yields

$$\frac{\partial \bar{\theta}_j}{\partial \alpha} = 1/B_j^2 \left(-2\alpha\omega\rho^j B_j - 2\alpha\omega\rho^j (1 - \alpha^2\omega) \right).$$

Finally, recognizing $1 - \alpha^2\omega = 1 + f(\bar{\lambda}, j) - B_j$ yields

$$\frac{\partial \bar{\theta}_j}{\partial \alpha} = -\frac{2\alpha\omega\rho^j (1 + f(\bar{\lambda}, j))}{B_j^2} < 0.$$

Step 3. $\hat{L}_{\lambda C} > 0$

This follows easily from the fact that $\hat{L}_C = \lambda^2$.

Step 4. $\hat{L}_{\lambda\rho} < 0$.

Note that

$$\frac{\partial \bar{\theta}_j}{\partial \rho} = \frac{j(1 + f(\bar{\lambda}, j))\rho^{j-1}}{\alpha^2\omega + f(\bar{\lambda}, j)} > 0,$$

for $j > 0$.

Step 5. $\hat{L}_{\lambda\omega} > 0$.

Simply notice

$$\frac{\partial \bar{\theta}_j}{\partial \omega} = -\frac{\alpha^2 (1 + f(\bar{\lambda}, j))\rho^j}{(\alpha^2\omega + f(\bar{\lambda}, j))^2} < 0$$

for $j > 0$.

Step 6 To show that $d\lambda^*/d\sigma_\varepsilon^2 > 0$ note that

$$\begin{aligned} T(\bar{\lambda}) &= \arg\min_{\lambda} \left(L(\lambda, \bar{\lambda}) + C\lambda^2 \right) \\ &= \arg\min_{\lambda} \left(\sigma_\varepsilon^2 \left((1 - \lambda)\bar{\psi}_0 - \lambda \sum_{j=1}^{\infty} (1 - \lambda)^j \bar{\psi}_j \right) + C\lambda^2 \right) \\ &= \arg\min_{\lambda} \left(\frac{\sigma_\varepsilon^2}{C} \left((1 - \lambda)\bar{\psi}_0 - \lambda \sum_{j=1}^{\infty} (1 - \lambda)^j \bar{\psi}_j \right) + \lambda^2 \right). \end{aligned}$$

Thus, $T(\bar{\lambda})$ depends only on the ratio $\frac{\sigma_\varepsilon^2}{C}$. The result then follows from the fact $d\lambda^*/dC < 0$.

Appendix B

In this appendix, we show how to modify our model to incorporate aggregate demand shocks. Following BMR, we assume aggregate demand is given by

$$AD : y_t = \hat{m}_t - p_t + e_t, \tag{22}$$

and that the AS relation is unchanged. Here, $\hat{m}_t$ is the policy instrument formed at time $t-1$, and e_t is the white noise demand shock.

BMR show that in the presence of mark-up and demand shocks, optimal policy may be written in one of two equivalent ways:

$$E_{t-1}p_t = \sum_{j=1}^{\infty} \phi_j \varepsilon_{t-j} \tag{23}$$

$$E_{t-1}p_t = -\frac{1}{\alpha\omega} E_{t-1}y_t \tag{24}$$

where the ϕ_i are defined as before, see (9). Combining the AS relation with (22) yields

$$p_t - E_{t-1}p_t = \frac{\lambda}{1-\lambda(1-\alpha)}(\varepsilon_t + \alpha e_t).$$

Using (23), we obtain the following equilibrium price process:

$$p_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j} + \alpha\phi_0 e_t.$$

To obtain equilibrium output, use (22) to write

$$y_t - E_{t-1}y_t = (1-\alpha\phi_0)e_t - \phi_0\varepsilon_t.$$

By (24) it follows that

$$y_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j} + (1-\alpha\phi_0)e_t,$$

where the φ_i are defined as before, see (9). Recall $p_t^*(\bar{\lambda})$ is the optimal price given the economy wide $\bar{\lambda}$, that is

$$p_t^*(\bar{\lambda}) = p_t(\bar{\lambda}) + \alpha y_t(\bar{\lambda}) + u_t.$$

Notice then that we may write

$$p_t^*(\bar{\lambda}) = \sum_{j=0}^{\infty} \bar{\theta}_j \varepsilon_{t-j} + A(\alpha, \bar{\lambda})e_t,$$

where

$$A(\alpha, \bar{\lambda}) = \alpha((1-\alpha)\phi_0 + 1),$$

and $\bar{\theta}_j$ are defined as before.

Now set

$$\Omega(k) = \begin{cases} \sum_{j=k}^{\infty} \bar{\theta}_j \varepsilon_{t-j} & \text{if} \quad k \geq 1 \\ \\ \sum_{j=k}^{\infty} \bar{\theta}_j \varepsilon_{t-j} + Ae_t & \text{if} \quad k = 0. \end{cases}$$

Using this altered definition of Ω allows the derivation of the loss function to go through, just as in the paper, and we get

$$L(\lambda, \bar{\lambda}) = Var(p_t^*(\bar{\lambda})) - Var(\hat{p}_t(\lambda)).$$

Just as we altered Ω, we must also alter ψ. Further, the presence of two variance terms requires a slight modification of the way the ψ's enter the loss function. We require

$$\bar{\psi}_k = \begin{cases} \sigma_\varepsilon^2 \sum_{j=k}^{\infty} \bar{\theta}_j^2 & \text{if} \quad k \geq 1 \\ \\ \sigma_\varepsilon^2 \sum_{j=k}^{\infty} \bar{\theta}_j^2 + A^2\sigma_e^2 & \text{if} \quad k = 0. \end{cases}$$

Notice, contrary to the old version, we have σ_ε^2 embedded within the formula for ψ_k. Finally, we have:

$$L^N(\lambda, \bar{\lambda}) = (1-\lambda)\bar{\psi}_0 - \lambda \sum_{j=1}^{\infty} (1-\lambda)^j \bar{\psi}_j,$$

where the superscript N is meant to capture the fact that this is the "new" loss function, that is, the loss function written using the new definition of ψ.

For the purposes of linking demand shocks to the analysis already obtained, it is helpful to write the new loss function in terms of the old loss function. So let L represent the old loss function. Then we have

$$L^N(\lambda, \bar{\lambda}) = L(\lambda, \bar{\lambda}) + (1-\lambda)A(\alpha, \bar{\lambda})^2\sigma_e^2,$$

and notice that $L^N \to L$ as $\sigma_e^2 \to 0$.

We may use the already established properties of L to analyze the properties of L^N. We obtain

1. L^N is decreasing in λ.
2. $L^N_{\lambda\lambda} > 0$ (so the relevant minimum is unique).

Furthermore, if we define $\hat{L}^N = L^N + C\lambda^2$, we obtain

1. $\hat{L}^N_{\lambda C} > 0$,

2. $\hat{L}^N_{\lambda\rho} < 0$,

3. $\hat{L}^N_{\lambda\omega} > 0$.

These results, together with the proof presented in Appendix A, show that the comparative statics obtained for parameters $C, \rho, \sigma^2_\varepsilon$, and ω are still valid: see Proposition 3. However, the comparative statics with respect to α are no longer clear because α enters into the parametric dependence of A.

Finally, we may compute the comparative statics for σ^2_e. First, notice $\hat{L}_{\lambda\sigma^2_e} = 0$ and

$$\frac{\partial^2 \left((1-\lambda)\, A\left(\alpha, \bar{\lambda}\right)^2 \sigma^2_e \right)}{\partial\lambda\partial\sigma^2_e} = -A\left(\alpha, \bar{\lambda}\right)^2 < 0.$$

This shows $\hat{L}_{\lambda\sigma^2_e} < 0$, so that from the proof in Appendix A, we get $\frac{\partial\lambda^*}{\partial\sigma^2_e} > 0$. And the intuition for this is straightforward: as the variance of the demand shocks increases, agents have increased incentive to gather information, thus the graph of the T-map is shifted upward.

References

[1] Adam, Klaus, 2004, "Optimal Monetary Policy when the Private Sector has Limited Capacity to Process Information," mimeo.

[2] Ball, Laurence, and David Romer, 1991, "Sticky Prices as Coordination Failure," *American Economic Review*, 81, 539-552.

[3] Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis, 2003, "Monetary Policy for Inattentive Economies," forthcoming *Journal of Monetary Economics.*

[4] Bernanke, Ben S., 2004, "The Great Moderation," Remarks by Governor Ben S. Bernanke at the meetings of the Eastern Economic Association, Washington, DC, February 20, 2004.

[5] Blanchard, Olivier, and John Simon, 2001, "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity,* 1, 135-64.

[6] Branch, William A., and George W. Evans, 2004, "Intrinsic Heterogeneity in Expectation Formation," University of Oregon Economics Department Working Paper 2003-32, forthcoming *Journal of Economic Theory.*

[7] Calvo, Guillermo, 1983, "Staggered Pricing in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12, 383-398.

[8] Carroll, Christopher D., 2003, "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of Economics*, 118, 1, 269-298.

[9] Clarida, Richard, Jordi Gali, and Mark Gertler, 1999, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature,* 37.

[10] Clarida, Richard, Jordi Gali, and Mark Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147-180.

[11] Cogley, Timothy and Thomas J. Sargent, 2002, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," mimeo.

[12] Evans, George W., and Seppo Honkapohja, 2001, *Learning and Expectations in Macroeconomics,* Princeton University Press, Princeton NJ.

[13] Evans, George W., and Seppo Honkapohja, 2003a, "Adaptive Learning and Monetary Policy Design," *Journal of Money Credit and Banking*, 35, 1045-1072.

[14] Evans, George W., and Seppo Honkapohja, 2003b, "The E-Correspondence Principle," University of Oregon Economics Department Working Paper 2003-27.

[15] Evans, George W., and Garey Ramey, 2003, "Adaptive Expectations, Underparameterization, and the Lucas Critique," University of Oregon Economics Department Working Paper 2003-2, forthcoming *Journal of Monetary Economics*.

[16] Fuhrer, Jeff and George Moore, 1995,"Inflation Persistence," *Quarterly Journal of Economics*, 110, 127-159.

[17] 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*.

[18] Koenig, Evan F., 2004, "Optimal Monetary Policy in Economies with 'Sticky-Information' Wages," mimeo., Federal Reserve Bank of Dallas.

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

[20] Lubik, Thomas A., and Frank Schorfeide, 2003, "Testing for Indeterminacy: An Application to U.S. Monetary Policy," forthcoming *American Economic Review*.

[21] Mankiw, N. Gregory, 2001, "The Inexorable and Mysterious Trade-off between Inflation and Unemployment," *Economic Journal*, 111, C45-C61.

[22] 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*, 117, 4.

[23] Marcet, Albert and Thomas J. Sargent, 1989, "Convergence of Least-Squares Learning Mechanisms in Self-Referential Linear Stochastic Models," *Journal of Economic Theory*, 48, 337-368.

[24] McCallum, Bennett T., 2000, "The Present and Future of Monetary Policy Rules," *International Finance*, 3, 273-286.

[25] McCallum, Bennett T. and Edward Nelson, 1999, "An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis," *Journal of Money Credit and Banking*, 31(3), part 1, 296-316.

[26] Muth, John F., 1961, "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29, 315-335.

[27] Orphanides, Athanasios, and John C. Williams, 2003a, "Imperfect Knowledge, Inflation Expectations, and Monetary Policy," forthcoming *Inflation Targeting*, eds. Ben Bernanke and Michael Woodford, University of Chicago Press.

[28] Orphanides, Athanasios, and John C. Williams, 2003b, "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations," mimeo.

[29] Reis, Ricardo, 2003, "Inattentive Consumers," mimeo.

[30] Reis, Ricardo, 2004, "Inattentive Producers," mimeo.

[31] Rogoff, Kenneth, 1985, "The Optimal Degree of Commitment to an Intermediate Monetary Target," *Quarterly Journal of Economics*, 100, 1169-1189.

[32] Sargent, Thomas J., 1999, *The Conquest of American Inflation*, Princeton University Press, Princeton NJ.

[33] Simon, Herbert A, 1978, "Rationality as Process and the Product of Thought," *American Economic Review*, 68, 1-16.

[34] Sims, Christopher A., 2003, "Implications of Rational Inattention," *Journal of Monetary Economics,* 50, 3.

[35] Sims, Christopher A. and Zha, 2004, "Were there Regime Switches in US Monetary Policy," mimeo.

[36] Schorfeide, Frank, 2003, "Learning and Monetary Policy Shifts," mimeo.

[37] Svensson, Lars E., 2003, "What is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules," *Journal of Economic Literature*, 41, 426-477.

[38] Taylor, John B., 1980, "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 88, 1-23.

[39] Taylor, John B., 1999, *Monetary Policy Rules*, ed., University of Chicago Press.

[40] Woodford, Michael, 2002, "Inflation Stabilization and Welfare," *Contributions to Macroeconomics*, 2, 1, 1-51.

[41] Woodford, Michael, 2003, *Interest and Prices,* Princeton University Press, Princeton NJ.

[42] Yetman, James, 2003, "Fixed Prices versus Predetermined Prices and the Equilibrium Probability of Price Adjustment," *Economics Letters*, 80, 3.

Table 1.

	Standard Deviation in %		
	1947:1-2004:1	1947:1-1983:4	1984:1-2004:1
y	1.70	2.00	0.95
p	0.98	1.16	0.48

Note: Standard deviation in percent of log real GDP, y, and implicit price deflator p. Data have been HP-detrended.

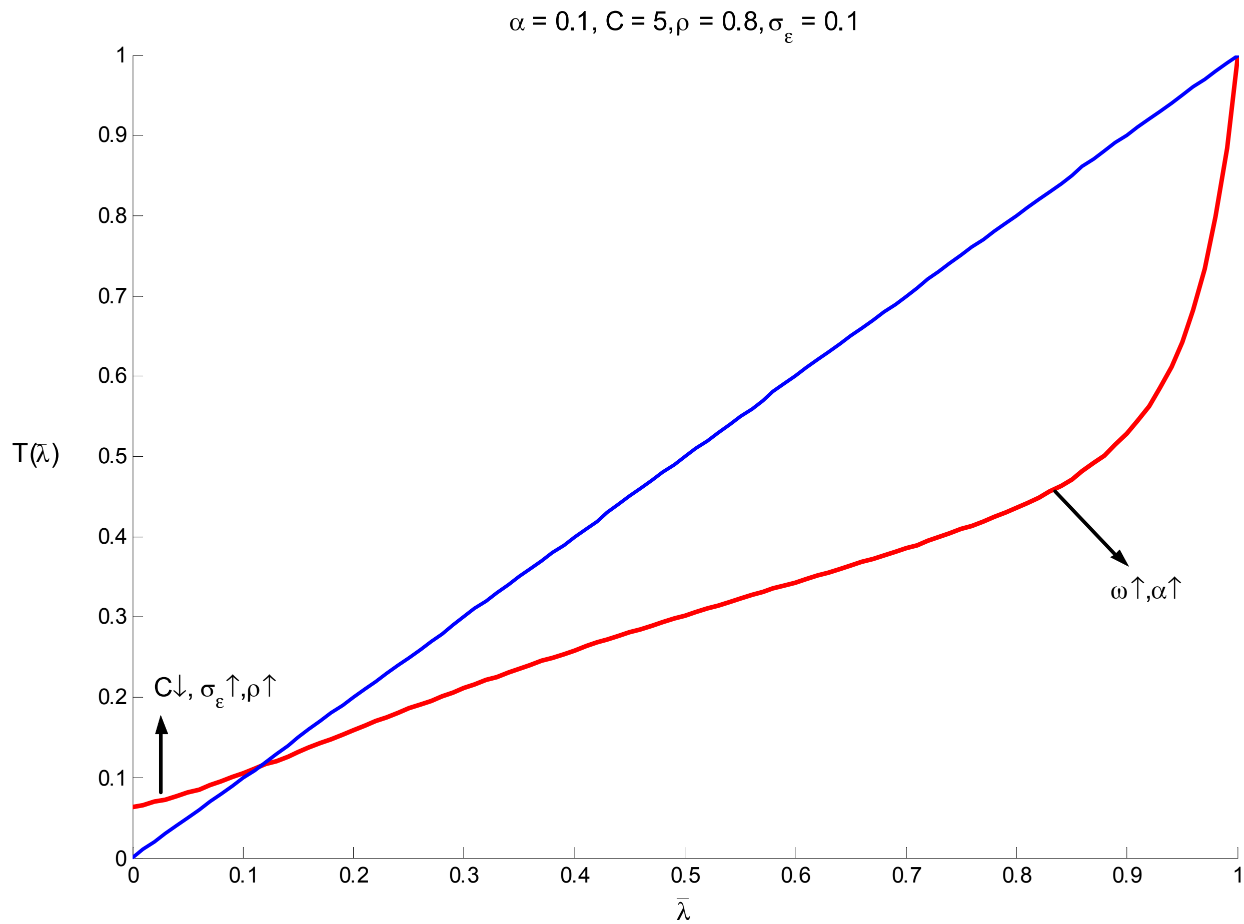

Figure 1. T-map under baseline parameterization. Arrows indicate the way in which T-map is altered by changing a given parameter value.

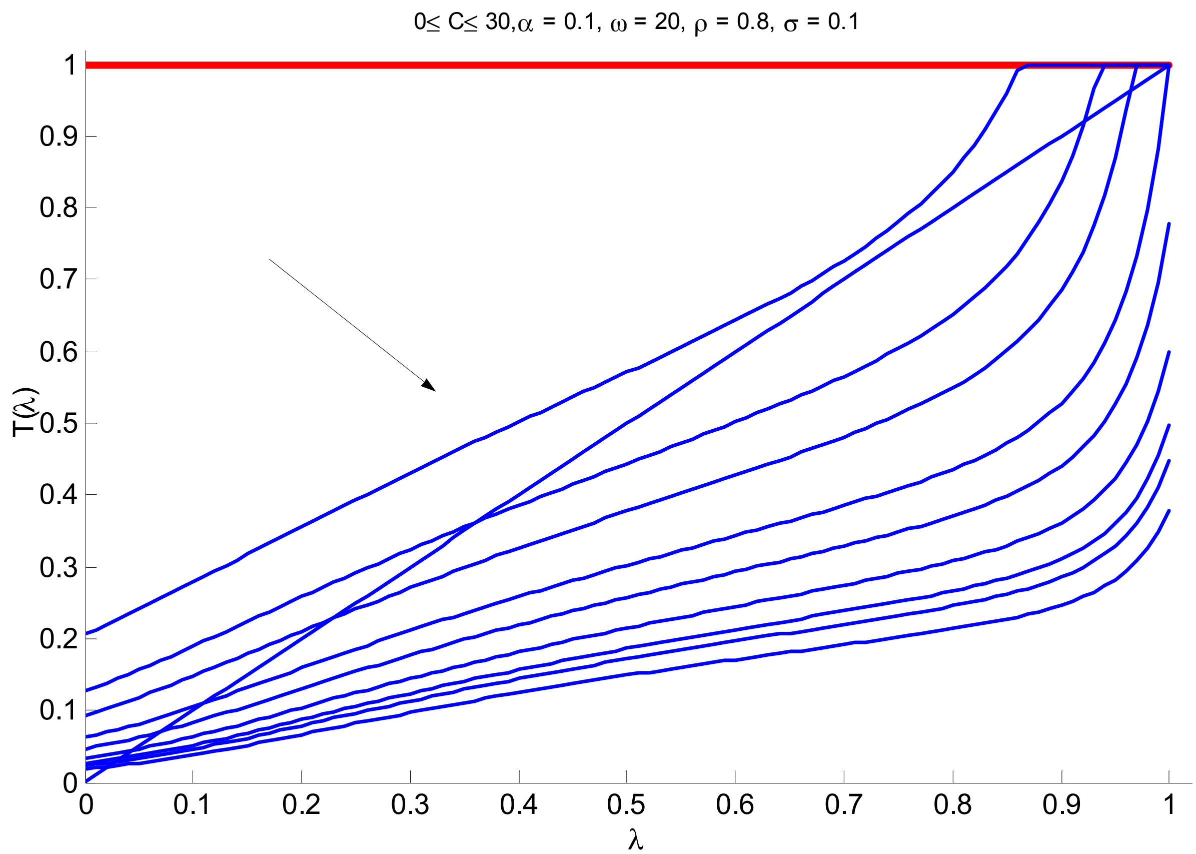

Figure 2. Comparative Statics for $0 \le C \le 30$.

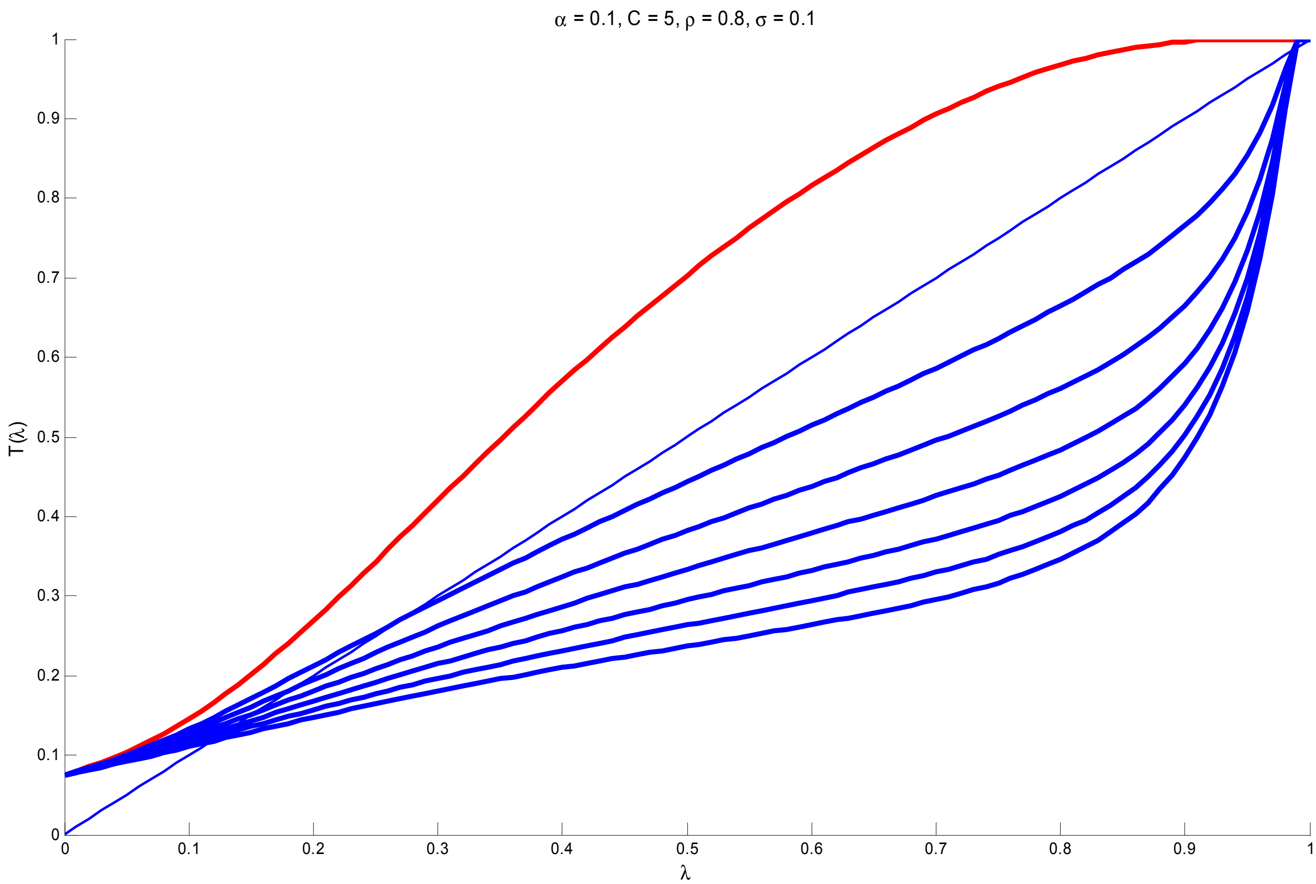

Figure 3. Comparative Statics for $0 \leq \omega \leq 40$.

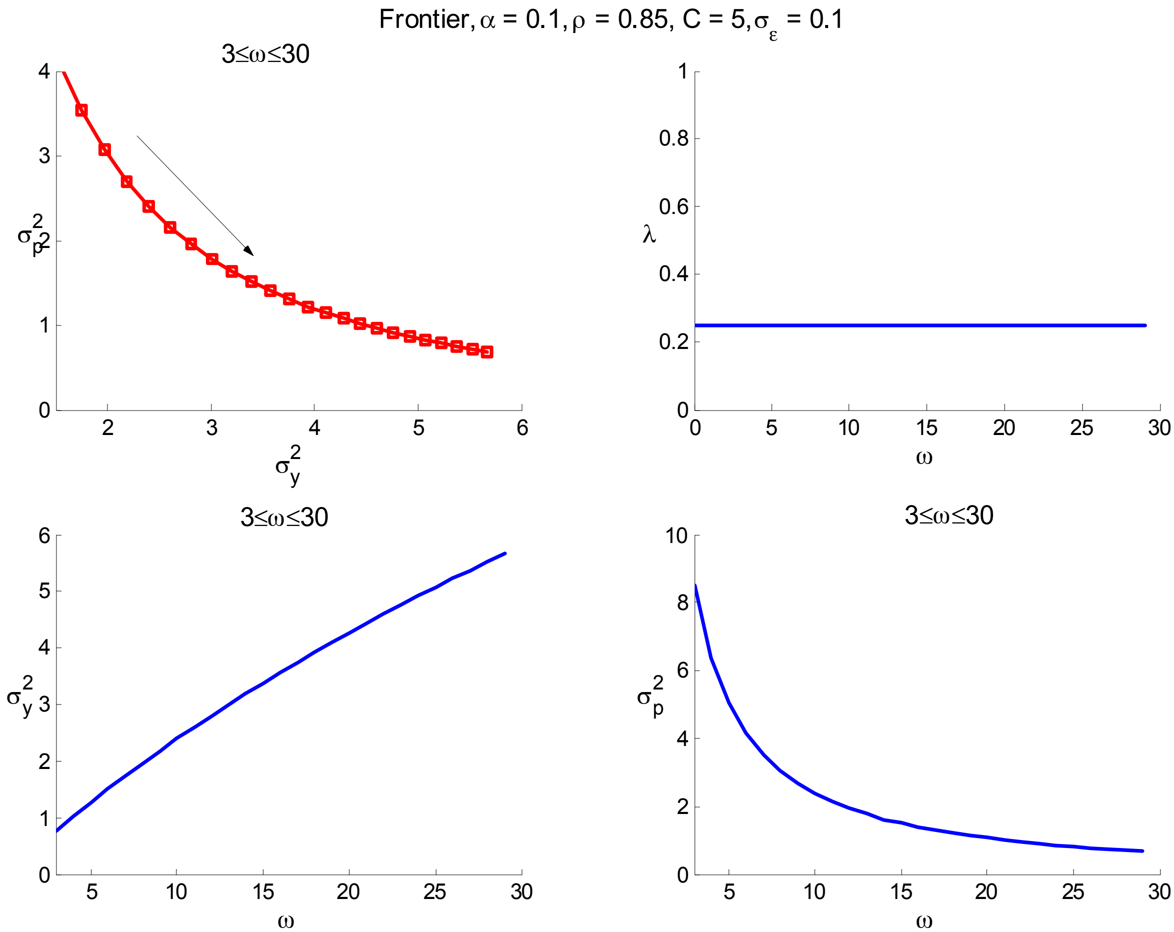

Figure 4. Frontier with fixed $\lambda = .25$.

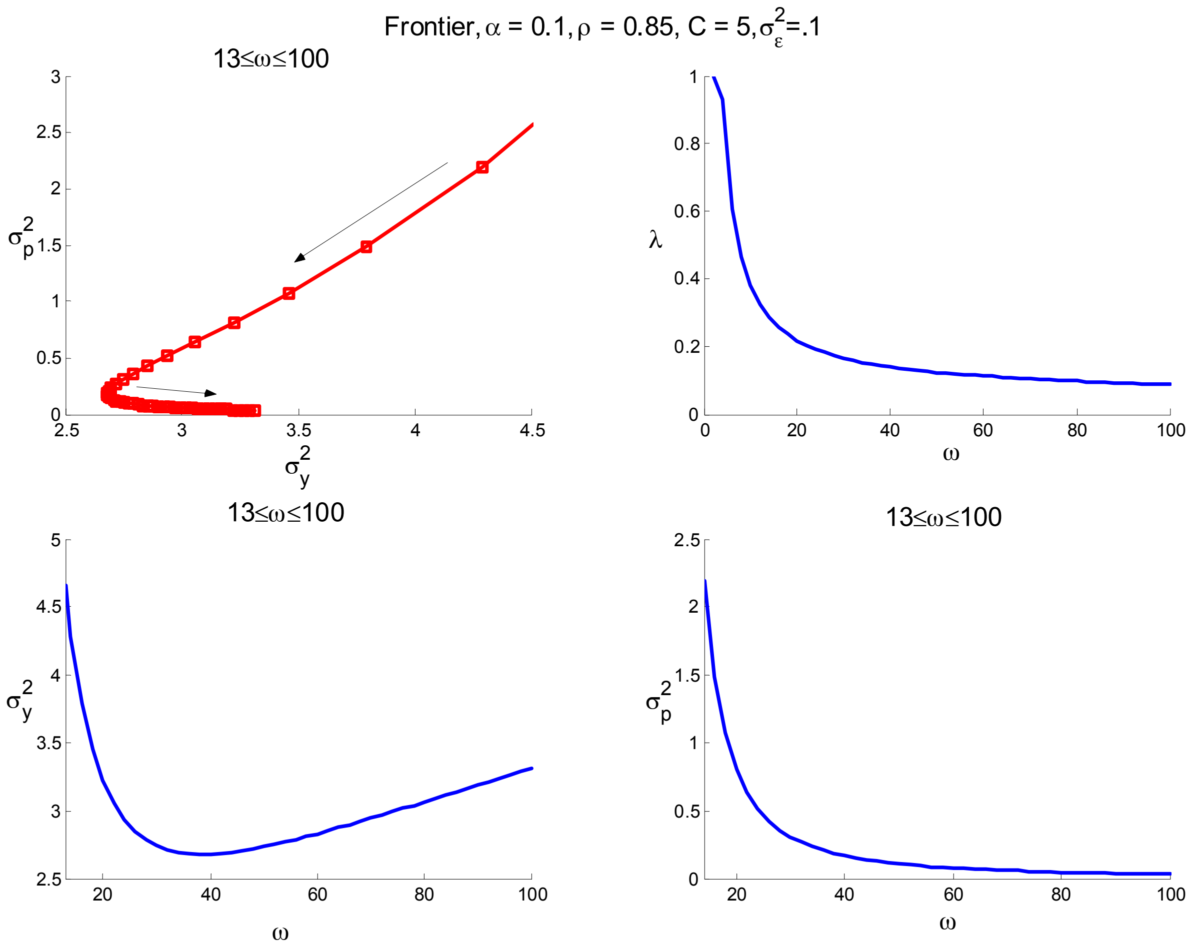

Figure 5. Policy Frontier with Endogenous Inattention and low costs. Northwest, Southwest, and Southeast panels plot $13 \le \omega \le 100$. Northeast panel plots $0 \le \omega \le 100$.

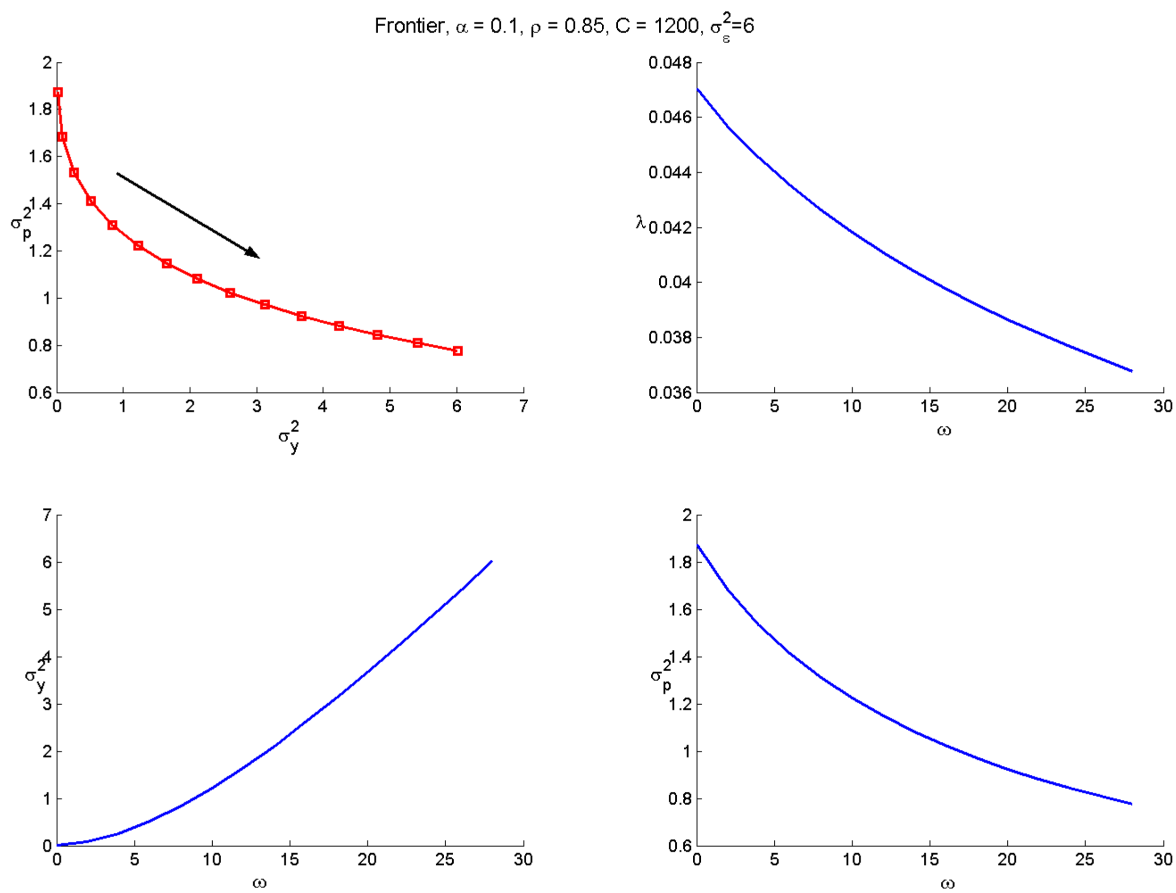

Figure 6. Policy Frontier with Endogenous Inattention and high costs.

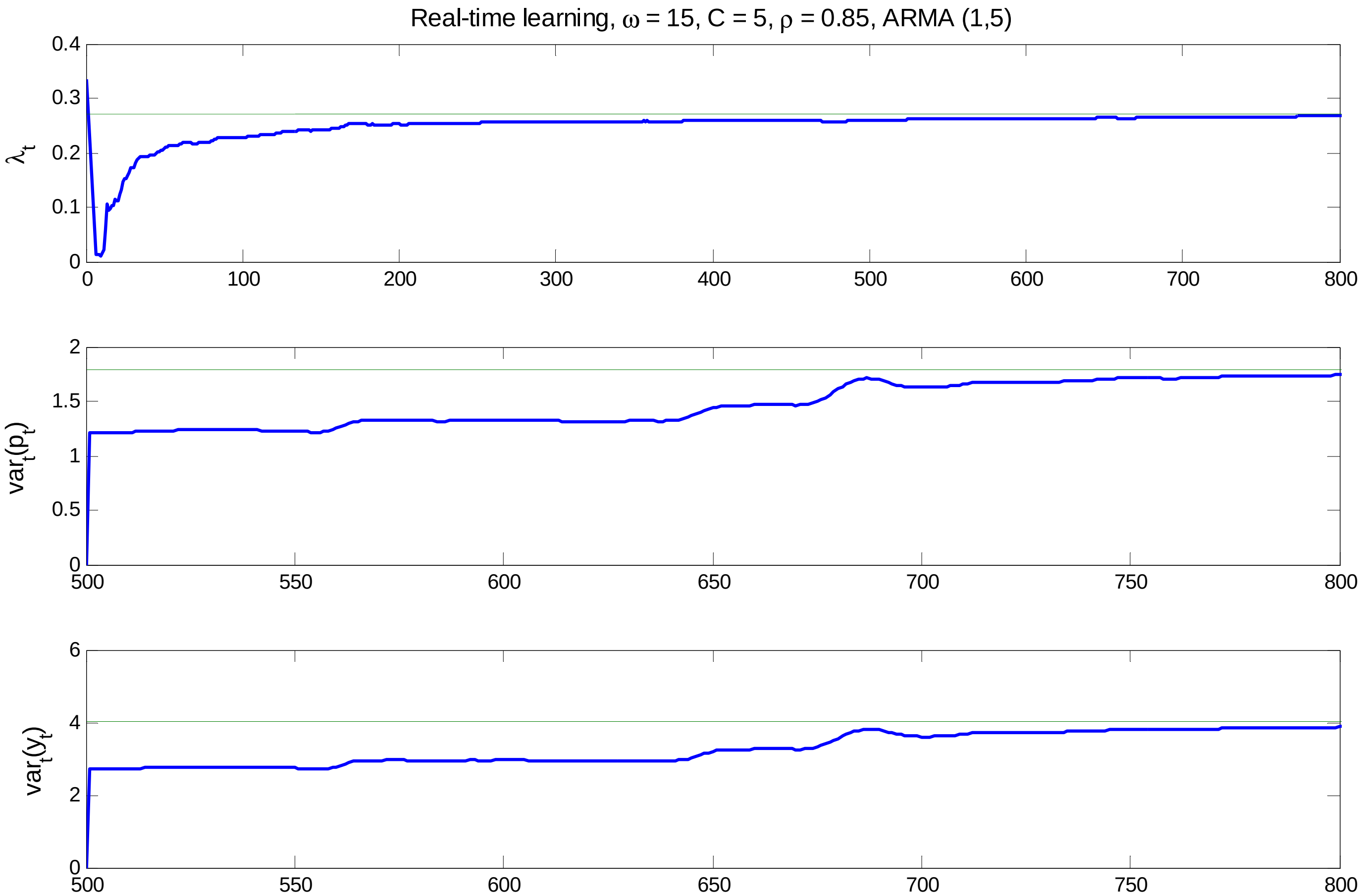

Figure 7. Stability of Endogenous Inattention under Adaptive Learning.

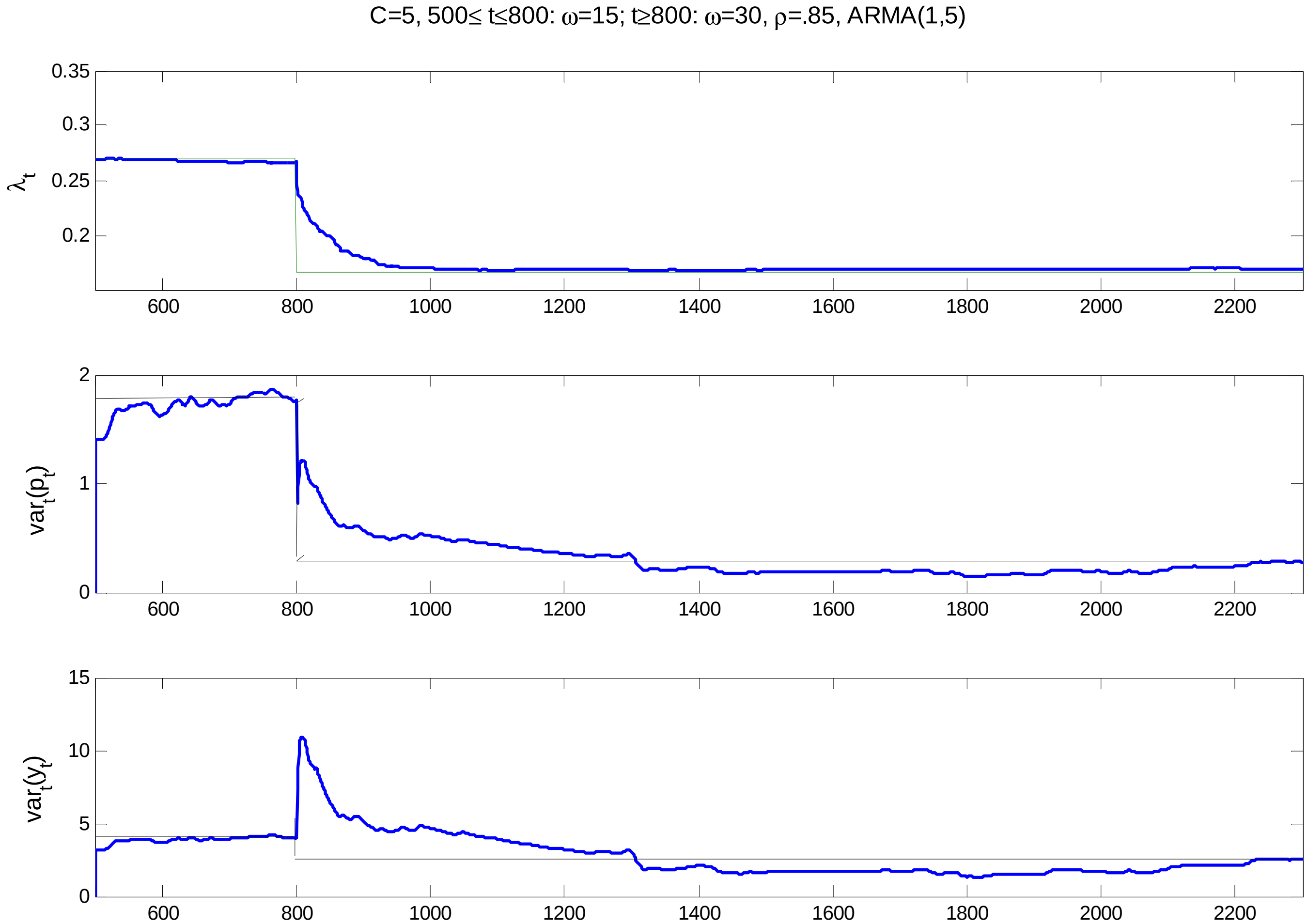

Figure 8. The Great Moderation in Real-time. At time 800 ω increases to 30.