# Uncertainty, Learning, and Gradual Monetary Policy *

Brian Sack

Board of Governors of the
Federal Reserve System

July 1998

**Abstract**

This paper argues that interest-rate smoothing may be optimal when the effect of monetary policy is uncertain. A model is presented in which the Federal Reserve rationally learns about the policy multiplier by observing the reaction of the economy to recent choices of the interest rate. As a result of this learning process, the Fed faces greater uncertainty about the impact of its policy as it moves the interest rate away from its previous level. The optimal policy response to macroeconomic developments therefore involves gradual adjustment of the interest rate over a period of time during which the Fed is learning about the effect of its policy, consistent with the smoothness of interest rate movements found in estimated policy rules. The model also suggests that periods of active interest rate movements, by allowing the Fed to learn more effectively, may be followed by a more aggressive policy rule.

**Correspondence:** Brian Sack, Division of Monetary Affairs, Board of Governors of the Federal Reserve System, Washington, D.C. 20551 / phone: 202-736-5671 / e-mail: bsack@frb.gov.

*I wish to thank Olivier Blanchard, Vincent Reinhart, and Jeremy Stein for their suggestions and comments. I am also grateful for helpful discussions with Antulio Bomfim, Anil Kashyap, Athanasios Orphanides, Roberto Rigobon, and Julio Rotemberg. The opinions expressed are those of the author and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System or other members of its staff.

# 1 Introduction

There is a substantial amount of evidence suggesting that the Federal Reserve conducts monetary policy in a gradual manner. The federal funds rate appears to be adjusted at a deliberate pace, typically through long sequences of small interest rate changes in the same direction. This behavior is apparent in Figure 1, which displays the intended federal funds rate over various time periods.[1] Rudebusch(1995) provides a statistical analysis of this behavior, finding that changes in the intended funds rate are frequently followed by additional changes of the same sign, which are referred to as "continuations" (as opposed to "reversals"). In addition, he finds that continuations occur in fairly rapid succession, indicating that these changes may constitute steps within a single policy movement. In the sample plotted in Figure 1, 88% of funds rate changes are continuations, with an average of 26 days separating changes when there is a continuation, compared to the average 77 days for a reversal.

The apparent tendency to smooth interest rate changes is also evident in estimates of simple monetary policy rules. Taylor-type rules, which have been widely used to characterize recent U.S. monetary policy, typically specify a target federal funds rate $i^*$ that is a function of the output gap and the deviation of inflation from its target:[2]

$$i_t^* = r^* + \beta_y(Y_t - Y^*) + \beta_\pi(\pi_t - \pi^*), \tag{1}$$

where $r^*$ is the equilibrium real interest rate. To fit the data, though, this policy rule has often been modified to allow the funds rate to move gradually towards the target level, as in:

$$i_t = (1 - \rho) \cdot i_t^* + \rho \cdot i_{t-1}. \tag{2}$$

The results typically find that the value of the parameter $\rho$ measuring the smoothness of interest rate changes is quite high. For example, Table 1 presents estimates of the policy rule obtained by combining equations (1) and (2) using least-squares and instrumental variables estimation.[3] The IV technique would be appropriate if deviations from the policy rule affect

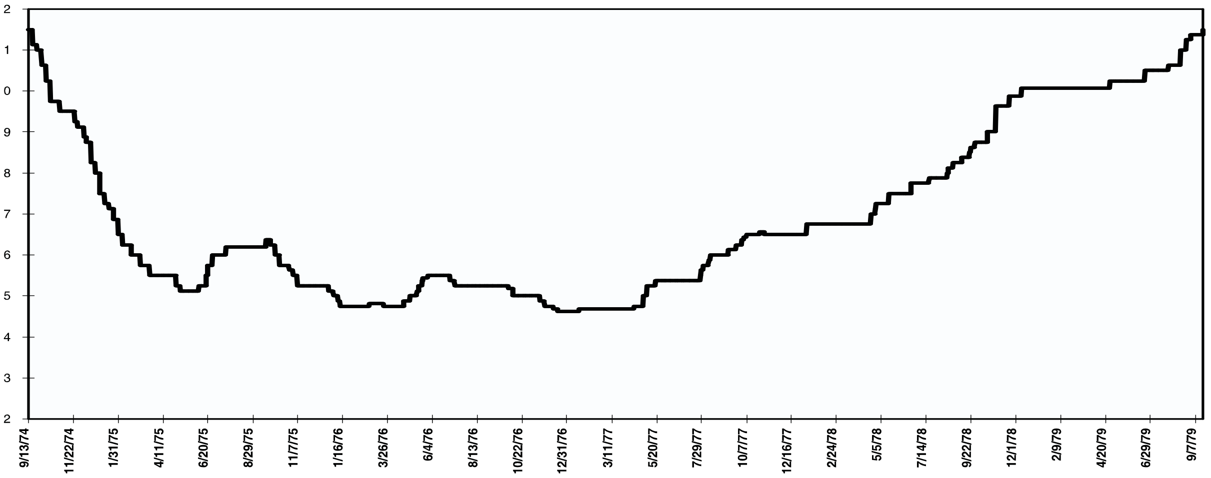
[1] The intended funds rate is the target series reported by Rudebusch(1995) for the episodes September 1974 to September 1979 and March 1984 to September 1992. The target series has been extended through December 1997 based on press releases by the FOMC announcing policy actions.

[2] Rules of this type were first examined by Bryant, Hooper, and Mann(1993) and were analyzed by Taylor(1993).

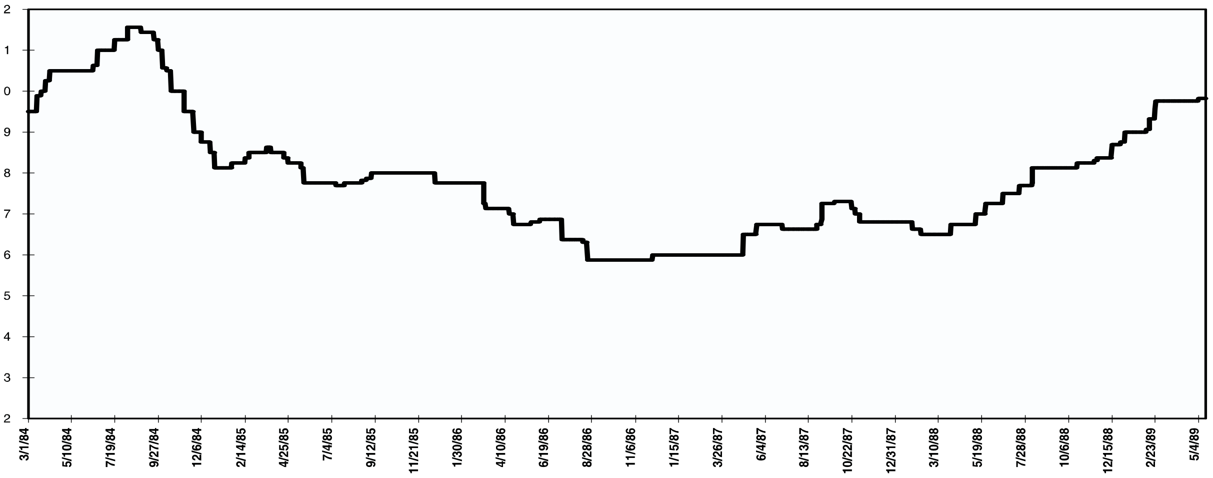
[3] The estimation is carried out using quarterly data over the period 1987:4 to 1997:4, corresponding to the tenure of Chairman Greenspan. The inflation rate is the annual percent change in the GDP deflator. The output gap is based on the potential output series constructed by the CBO as of July 1997.

# Figure 1: Intended Federal Funds Rate

September 1974 to September 1979




March 1984 to May 1989




May 1989 to December 1997

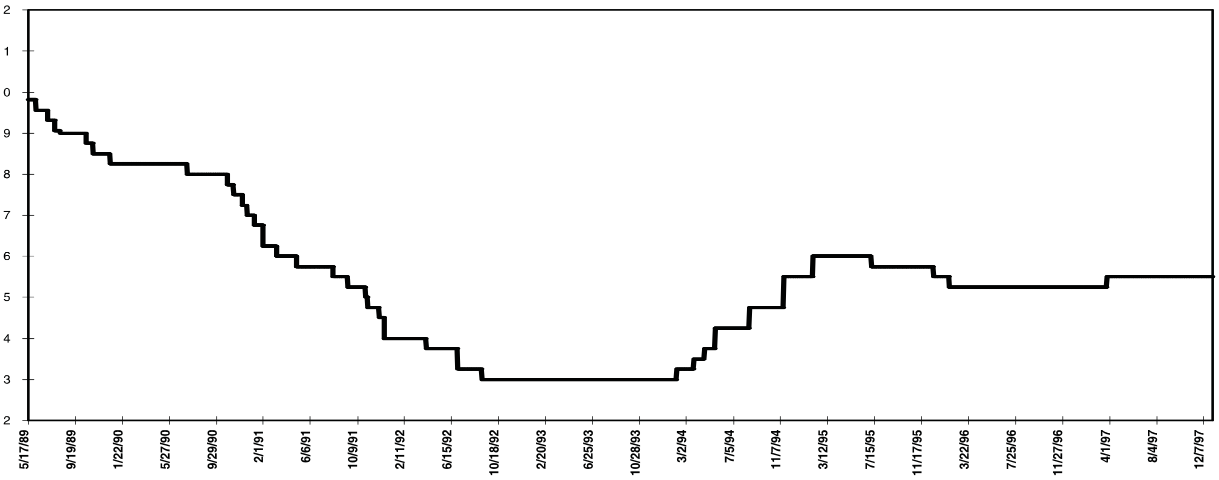


**Table 1: Estimated Policy Rule**

| | Least Squares Estimates | Instrumental Variables |
|---|---|---|
| $\beta_0$ | 1.00 (1.55) | 1.41 (1.77) |
| $\beta_y$ | 1.13 (8.14) | 1.16 (7.97) |
| $\beta_\pi$ | 1.65 (7.26) | 1.52 (5.83) |
| $\rho$ | 0.63 (7.55) | 0.65 (10.37) |
| $\overline{R}^2$ | 0.969 | 0.969 |
| $SEE$ | 0.345 | 0.347 |

Notes: Results shown are LS and IV estimates of the equation:

$$i_t = (1-\rho)\left(\beta_0 + \beta_y(Y_t - Y^*) + \beta_\pi \pi_t\right) + \rho \cdot i_{t-1}.$$

The t-statistics shown in parentheses are based on Newey-West standard errors.

Instruments include one to four lags of inflation, the output gap, and the funds rate.

the within-quarter output gap and inflation rate.[4] The results indicate that the funds rate reacts significantly to both output and inflation deviations, but the response is very gradual, as indicated by the large (and significant) coefficient $\rho$.[5] In response to a permanent change in the target funds rate, the actual funds rate responds by only 37% of the change within the quarter and takes five quarters to complete 90% of the movement.

Other studies have demonstrated that the degree of interest-rate smoothing is puzzling given the dynamic structure of the economy. An analysis by Sack(1998) indicates that movements in the funds rate are too gradual to be accounted for strictly by the dynamic behavior of macroeconomic variables estimated from a VAR. The reaction of the funds rate to various shocks in the economy is damped, as the funds rate instead moves gradually in the direction of the expected response over a period of several months.[6] Evidence of interest-rate smoothing is also found in analyses using structural models of the economy. In papers by Rotemberg and Woodford(1997) and Williams(1997), the performance of the estimated policy rule could be substantially improved by implementing more volatile movements in the funds rate.

These empirical findings suggest that the Fed is reluctant to change the funds rate *from its previous level*, resulting in the inertial movements in the funds rate that are observed. For

[4]The instruments include the funds rate, inflation rate, and output gap lagged one to four quarters.

[5]The results are consistent with other estimates of this rule for the U.S. – see, for example, Orphanides(1998) and Clarida, Gali, and Gertler(1997). Also note that the similarity between the OLS and IV results suggests that simultaneity may not be a serious problem.

[6]These results refer to the case labeled Additive Uncertainty in that paper. An alternative case labeled Parameter Uncertainty investigates the effect of the imprecision of the VAR estimates on the calculated policy. See that paper for a discussion of the relation between those results and the current analysis.

this reason, attempts to explain gradual interest rate movements have typically proceeded by appending the objective function of the central bank with a loss related to the magnitude of interest rate changes. Goodfriend(1987) posits that the Fed smoothes interest rates to protect the banking sector against financial crises, an idea later formalized by Cukierman(1991), although there is no empirical support for this explanation. Meulendyke(1990) expresses the view that interest rate changes are damped in order to avoid reversals in the direction of interest rate changes, and Goodfriend(1991) and others have described the cost of reversals as "whipsawing the market," but this cost has not been theoretically developed.[7]

This paper offers an explanation of interest-rate smoothing that does not resort to amending the objective function with an ad-hoc interest-rate smoothing argument. Instead, the analysis incorporates the idea that the Federal Reserve faces uncertainty about the effect of monetary policy on the economy.[8] This type of uncertainty, first analyzed by Brainard(1967), limits movements of the policy instrument away from the level at which policymakers are most certain about its effect. Interest-rate smoothing arises from the manner in which this uncertainty is resolved through time. Importantly, in the model presented in this paper, the structure of the uncertainty is endogenously determined through a process in which the Fed rationally learns about the policy multiplier by observing the reaction of the economy to recent levels of the interest rate. As a result, not only is the policy choice affected by the uncertainty over its effect, but the dynamic evolution of the uncertainty depends on the policy that has been recently implemented. This interaction between the choice of the interest rate and the evolution of uncertainty arising from the learning process leads to gradual movements in the interest rate under the optimal policy.

Consider, for example, the policy response to developments that have moved output away from the level that is consistent with the Fed's output and inflation objectives. Because it has observed the reaction of the economy to recent choices of the funds rate, the Fed is fairly confident about the level of output that is obtained by maintaining the funds rate near its recent levels, but faces greater uncertainty about the change in output that would be induced by moving the funds rate away from these levels. To limit any undesired impact on

[7]Other authors have investigated forward-looking forms of interest-rate smoothing. In Mankiw(1987), for example, interest rates follow a random walk in order to achieve the optimal collection of seignorage. Optimal taxation models often yield forward-looking interest-rate smoothing, resulting in random-walk behavior for the interest rate. However, the empirical evidence instead suggests that the observed policy limits changes from previous interest rates rather than minimizing future interest rate changes. Explaining the gradual behavior depicted in Figure 1 entails accounting for the fact that the interest rate is *not* a random walk.

[8]This paper refers to the Federal Reserve as the monetary authority, but the theoretical reasons for interest-rate smoothing apply equally to other central banks. Foreign central banks also exhibit substantial degrees of interest rate smoothing, as described in the 68th Annual Report(1998) of the Bank for International Settlements (in particular, see the table on page 68).

the economy, the Fed damps the change in the funds rate in its policy response, and therefore output does not move fully back to its desired level. Although damped, the change in the funds rate subsequently provides information about the policy multiplier. Observing the reaction of the economy at the new level of the funds rate resolves some of the uncertainty regarding the response of output, allowing the Fed to implement additional changes in the funds rate. This process of learning continues, resulting in an funds rate response that will be implemented through a number of steps over a period of time during which the Fed is learning about the effect of its policy.

Ignorance of the effect of policy decisions and the endogenous evolution of uncertainty about the policy multiplier can therefore explain inertial movements in the funds rate. Comments by members of the FOMC suggest that this type of uncertainty is in fact a serious consideration in the formulation of policy decisions, as summarized in the following quote by Alan Blinder while speaking on the strategy of monetary policy:[9]

> Unfortunately, actually to use such a strategy in practice, you have to use forecasts, knowing that they may be wrong. You have to base your thinking on some kind of monetary theory, even though that theory might be wrong. And you have to attach numbers to that theory, knowing that your numbers might be wrong. ... We at the Fed have all these fallible tools, and no choice but to use them.

In the presence of this uncertainty, Blinder advocates the following policy approach:[10]

> What can you do to try to guard against failure? ... First of all, be cautious. Don't oversteer the ship. If you yank the steering wheel really hard, a year later you may find yourself on the rocks.

This idea of exercising caution is captured by the gradual response of the interest rate depicted in this paper.

In addition to accounting for the interest-rate smoothing observed, the model presented suggests that the amount of uncertainty about the effect of policy should vary through time. Periods of active interest rate movements provide the Fed with more informative observations about the effect of its policy, resulting in a decline in uncertainty that allows a more aggressive reaction to subsequent shocks to the economy. Some brief empirical results are presented that suggest that the reaction speed of the funds rate to inflation and the output gap is in fact significantly determined by a measure of lagged funds rate activity.

[9] This quote is from a talk given at the Minnesota Meeting, a business forum that was held in June 1995. At that time, Blinder was the Vice-Chairman of the Board of Governors of the Federal Reserve System.

[10] This view is also expressed in Blinder(1998).

The paper is organized as follows. Section 2 presents a model of optimal monetary policy in which policy effectiveness has only two potential values. This simple model clearly demonstrates the effect of uncertainty and allows a closed-form solution in which the optimal reaction to a shock involves gradual adjustment of the interest rate. Section 3 generalizes the results to a model in which the parameters determining the effect of policy follow a stochastic process. The analysis highlights the elements of the learning problem that deliver gradualism and presents a numerical solution that focuses on the dynamic behavior of policy uncertainty. Section 4 summarizes the analysis and provides some brief empirical evidence that supports the implications of the model.

# 2 A Model of Gradualism

The impact of monetary policy on the economy is not precisely known. Because of this uncertainty, the variance of targeted macroeconomic variables will be affected by the choice of the policy instrument. Since variability in the targeted variables is in general undesirable, uncertainty about the policy multiplier must be taken into consideration in determining the optimal policy choice. This effect was first analyzed by Brainard(1967), who established that parameter uncertainty reduces the responsiveness of the policy instrument. The policymaker instead implements a policy choice that is closer to the level that minimizes the variance of its effect. However, Brainard's analysis does not determine the policy action that results in the least amount of uncertainty in the economy – the policy uncertainty is exogenous.[11] The model described in this section instead endogenously determines the variance-minimizing policy through a process of rational learning. It is the dynamic evolution of the parameter uncertainty in response to recent policy choices that results in gradual movements in the interest rate under optimal policy.

## 2.1 Learning about Policy Effectiveness

Consider the following control problem in which the Fed faces uncertainty about the effect of its policy instrument, here taken to be the federal funds rate. For a given choice of the interest rate, output is determined by the following IS curve:

$$Y_{t+1} = \alpha_{t+1} - \phi \cdot i_t. \tag{3}$$

[11] Diamond(1985) makes this point in arguing that the variance-minimizing policy choice does not necessarily correspond to a constant money growth rule, but instead should depend on current and lagged variables in the economy.

The timing assumed is that the interest rate is set after observing contemporaneous output, and this rate affects output in the following period.[12] This basic equation is clearly a simplification of the relationship between monetary policy and output. One aspect of the policy multiplier that is not captured by the relationship in (3) is that the response of output may be affected by the degree of gradualism in policy, as demonstrated by Caplin and Leahy(1996).[13] Output may also be affected by expectations of future levels of output in addition to the interest rate. However, the aspects of the learning process that result in gradual movements in the interest rate apply under alternative specifications of the aggregate demand relationship.

Ignorance about the effect of policy is captured by the fact that the Fed does not know the slope or intercept of the IS curve. This section develops a tractable model of the learning problem by assuming that there are two potential values for the policy multiplier, $\phi^L$ and $\phi^H$, where $\phi^H > \phi^L$, which are referred to as low ($L$) and high ($H$) policy effectiveness. One of these values represents the actual response of the economy to monetary policy. The Fed knows the two potential values but faces uncertainty about which of the two values is the true multiplier. In addition, the intercept term $\alpha$ is stochastic and evolves through time according to the following process:

$$\alpha_{t+1} = \alpha_t + \eta_{t+1}, \tag{4}$$

which complicates the inference problem for policymakers in determining the effect of their policy.

The Fed rationally learns about policy effectiveness by observing the level of output that is realized from past choices of the interest rate. As a result, the uncertainty it faces will depend on previous policy choices, thus drawing a link between recent and future levels of the interest rate that is the basis for interest-rate smoothing. In particular, the crucial implication of this learning process is that the uncertainty regarding the effect of policy will be minimized around recent levels of the interest rate. The observed level of output $Y_t$ that results from the interest rate choice $i_{t-1}$ is consistent with either low or high effectiveness for particular values for the intercept terms of the potential IS curves, denoted $\alpha^L_{t+1}$ and $\alpha^H_{t+1}$

[12] As written, output responds to the nominal interest rate. A discussion of inflation expectations and real interest rates is offered towards the end of this section.

[13] Caplin and Leahy(1996) present a model in which acting gradually reduces the impact of interest rate changes on output, so that more interest rate changes in a particular direction are required to elicit a desired output response. Their model does not, however, provide an explanation of gradualism, as the desire to move gradually is assumed in the payoff structure of the model.

and given by:

$$\begin{aligned}\alpha_t^L &= Y_t + \phi^L i_{t-1} && (5)\\ \alpha_t^H &= Y_t + \phi^H i_{t-1}. && (6)\end{aligned}$$

Having observed the output response to the most recent choice of the interest rate, the Fed knows that the potential IS curves must intersect at this value of the interest rate, as determined by the values $\alpha_t^L$ and $\alpha_t^H$. This situation is depicted in Figure 2. The figure demonstrates that the difference in the response of output under the two possible multipliers increases as the interest rate moves away from its previous level, so that *changes* in the interest rate from recent levels increase the variance of output.

The response of output to interest rate changes will provide a signal about the true effect of policy. In general, larger interest rate changes are more informative, because the difference in the response of the economy under the two potential values of policy effectiveness increases. But because the potential IS curves shift stochastically, an interest rate change will not necessarily be fully revealing. For example, an increase in the interest rate may reveal that policy is highly effective if investment falls substantially, but the fall in output could also arise under low policy effectiveness if a negative shock to output occurred, as depicted in Figure 3.

The Fed has a belief about the true value of policy effectiveness at any point in time which is summarized by the probability that policy has low effectiveness, denoted $P_t$. Since there are only two potential values for the multiplier, the belief $P_t$ completely captures the magnitude of the uncertainty that the Fed faces, with less uncertainty as $P_t$ moves towards 0 or 1. In the model presented in this section, the problem begins with a given amount of uncertainty over the policy multiplier, denoted $P_0$, and this uncertainty is resolved through time as the interest rate reacts to shocks in the economy.[14]

The Fed rationally learns about the true state of the economy using Bayes' Rule, based on the likelihood of the observed output response under the two potential values of the multiplier. The shocks to the intercept term in the IS equation are assumed to be uniformly distributed between the values of $-\eta^*$ and $\eta^*$. This assumption is made for tractability, but the results derived in this section generalize to other stochastic processes for the intercept term. In the case of uniform shocks, policymakers either fully learn or learn nothing about

[14]A model is presented in section 3 in which policy effectiveness follows a stochastic process, so that the uncertainty regenerates through time. The variation in the degree of uncertainty that results has additional implications for interest-rate smoothing, as discussed in that section.

**Figure 2: Potential IS Curves**

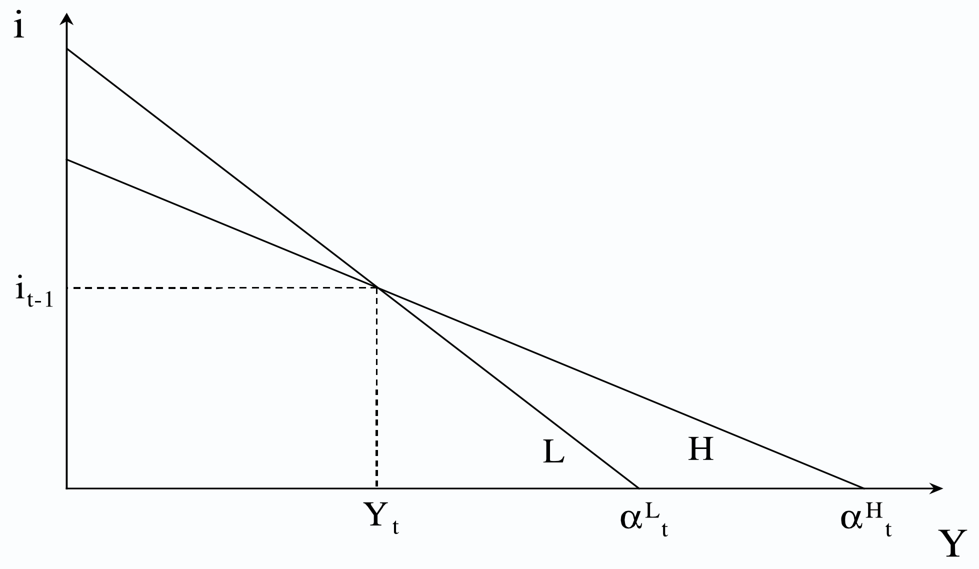
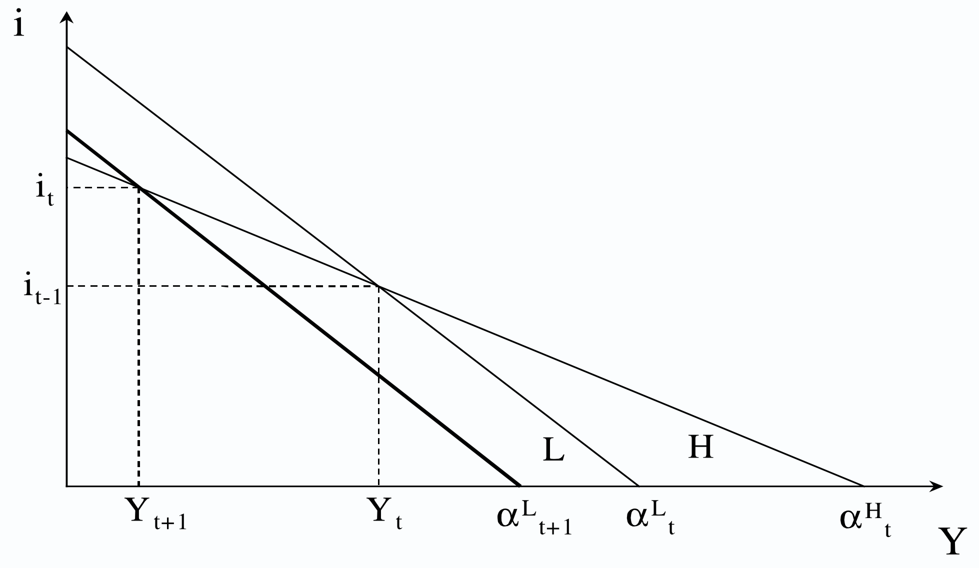



**Figure 3: The Response to an Interest Rate Change**



the true value of the multiplier.[15] For a given belief $P_t$, the high value of the multiplier ($\phi^H$) is revealed if, for example, an increase in the interest rate reduces output by more than the largest possible decrease under the low multiplier ($\phi^L$), which occurs only if the output shock is sufficiently small: $\eta < -\eta^* + \Delta i_t(\phi^H - \phi^L)$. Likewise, if the true multiplier is $\phi^H$, it is revealed only if $\eta > \eta^* - \Delta i_t(\phi^H - \phi^L)$. The unconditional probabilities of learning are given by:

$$\begin{aligned}
\Pr\left(P_{t+1} = 0 \mid P_t, \Delta i_t\right) &= \frac{1}{2\eta^*}\left(1 - P_t\right)\Delta i_t\left(\phi^H - \phi^L\right) \\
\Pr\left(P_{t+1} = P_t \mid P_t, \Delta i_t\right) &= 1 - \frac{1}{2\eta^*}\Delta i_t\left(\phi^H - \phi^L\right) \\
\Pr\left(P_{t+1} = 1 \mid P_t, \Delta i_t\right) &= \frac{1}{2\eta^*}P_t\Delta i_t\left(\phi^H - \phi^L\right).
\end{aligned}$$

Given this learning process, the optimal policy rule can be calculated.

## 2.2 Optimal Policy under Passive Learning

The control problem facing the Fed at time $t$ is to minimize the discounted sum of squared output deviations from a target:

$$\max_{\{i_t\}} -E_t\left[\sum_{j=1}^{\infty}\beta^{j-1}(Y_{t+j} - Y^*)^2\right] \tag{7}$$

subject to the output equation (3). Note that the per-period loss function can be rewritten as $E_t\left[(Y_{t+j} - Y^*)^2\right] = (E_t[Y_{t+j}] - Y^*)^2 + Var_t(Y_{t+j})$. Under this objective function, the Fed wishes to adjust the interest rate strongly so that on average it will be close to the target level of output, but it is reluctant to do so since aggressive policy actions would induce a large amount of variance in output as a result of the uncertainty over the effect of its policy.

The Fed has an additional incentive to implement strong policy moves because larger interest rate changes provide more information about the policy multiplier. In other words, optimal policy will involve some experimentation, since this will allow more successful policy to be implemented in the future. In general, the optimal policy under experimentation will involve larger movements in the interest rate. However, the incentive to learn will not necessarily outweigh the uncertainty generated from the more aggressive policy. An analysis of the effect of experimentation, including conditions under which policy may remain gradual, is offered in the appendix. In the analysis that follows, a model is developed in which the

[15] This simplification is also eliminated in the more general model in section 3.

Fed does not consider the speed with which it will learn when choosing the optimal policy.

This simplification does not imply that learning is irrelevant in determining monetary policy. On the contrary, learning is valuable and does influence policy, but the Fed does not take into account the effect of its policy choice on the amount of information produced. Under this simplification the optimal policy is given by:

$$i_t = \arg\max -[(E_t[Y_{t+1}] - Y^*)^2 + Var_t(Y_{t+1})],$$

for which an explicit solution can be found. The solution to the optimization problem involves the following moments for output:

$$\begin{aligned} E_t[Y_{t+1}] &= P_t(\alpha_t^L - \phi^L \cdot i_t) + (1 - P_t)(\alpha_t^H - \phi^H \cdot i_t) \\ Var_t[Y_{t+1}] &= P_t(1 - P_t)\left[(\alpha_t^L - \alpha_t^H) + (\phi^H - \phi^L)i_t\right]^2 + Var(\eta). \end{aligned}$$

The role of policy uncertainty is evident in the expression for the variance of output. When there is no uncertainty about policy effectiveness, which occurs when either $P = 1$, $P = 0$, or $\phi^L = \phi^H$, the variance is given by $Var(\eta)$, the component that is unavoidable given the stochastic nature of the economy. Uncertainty over the true state of policy effectiveness increases the variance. In fact, the variance of output is strictly concave over $P \in [0, 1]$ and is maximized at $P = \frac{1}{2}$.

Given these moments, the optimal interest rate rule can be obtained:

$$i_t = \frac{P_t\phi^L(\alpha_t^L - Y^*) + (1 - P_t)\phi^H(\alpha_t^H - Y^*)}{P_t(\phi^L)^2 + (1 - P_t)(\phi^H)^2}. \tag{8}$$

This equation states that the Fed sets the interest rate to a weighted average of the interest rates it would set in the two states of the world without uncertainty. To see this, define $\widehat{i}_t^L = \frac{\alpha_t^L - Y^*}{\phi^L}$ as the interest rate that would be set to reach the output target under low responsiveness, and define an analogous interest rate $\widehat{i}_t^H$ as the target rate under high responsiveness. Equation (8) can be rewritten as follows:

$$i_t = \omega_t \cdot \widehat{i}_t^L + (1 - \omega_t) \cdot \widehat{i}_t^H, \tag{9}$$

where the weighting is given by $\omega_t = \frac{P_t(\phi^L)^2}{P_t(\phi^L)^2 + (1-P_t)(\phi^H)^2}$.

It will prove useful to compare the optimal policy to the "completely-offsetting" policy, denoted $i^c$, which ignores the variance of output and minimizes the expected deviation of output from its target:

$$i_t^c = \arg\max -(E_t[Y_{t+1}] - Y^*)^2.$$

The optimization problem in this case yields a similar weighted average rule as in (9) but with the weight $\omega_t^c = \frac{P_t\phi^L}{P_t\phi^L+(1-P_t)\phi^H}$. The entire effect of the uncertainty is therefore captured in a change to the weights assigned to these two interest rates. In particular, the uncertainty decreases the weight on the term that would require a larger interest rate movement, or $\omega < \omega^c$.

The optimal monetary policy given in equation (9) involves a gradual interest rate response to a deviation of output from its target. Consider an economy that begins with output equal to its target level but experiences a positive shock to output (to $\alpha$) of magnitude $\tau$, as shown in Figure 4. The Fed will react to the deviation by increasing the interest rate. The top panel of the figure depicts the response of the economy under the two potential values for the multiplier, as well as an expected response schedule given the belief $P_t$. The interest rates $\hat{i}_t^L$ and $\hat{i}_t^H$ are those at which output is expected to return to $Y^*$ under the respective multipliers. The completely-offsetting policy rule $i_t^c$ is found at the point along the line segment between $\hat{i}_t^L$ and $\hat{i}_t^H$ corresponding to the weights $\omega_t^c$, precisely where the expected response schedule crosses $Y^*$. The optimal interest rate policy (9) instead puts additional weight on the point $\hat{i}_t^H$, resulting in the interest rate choice denoted $i_t^g$.

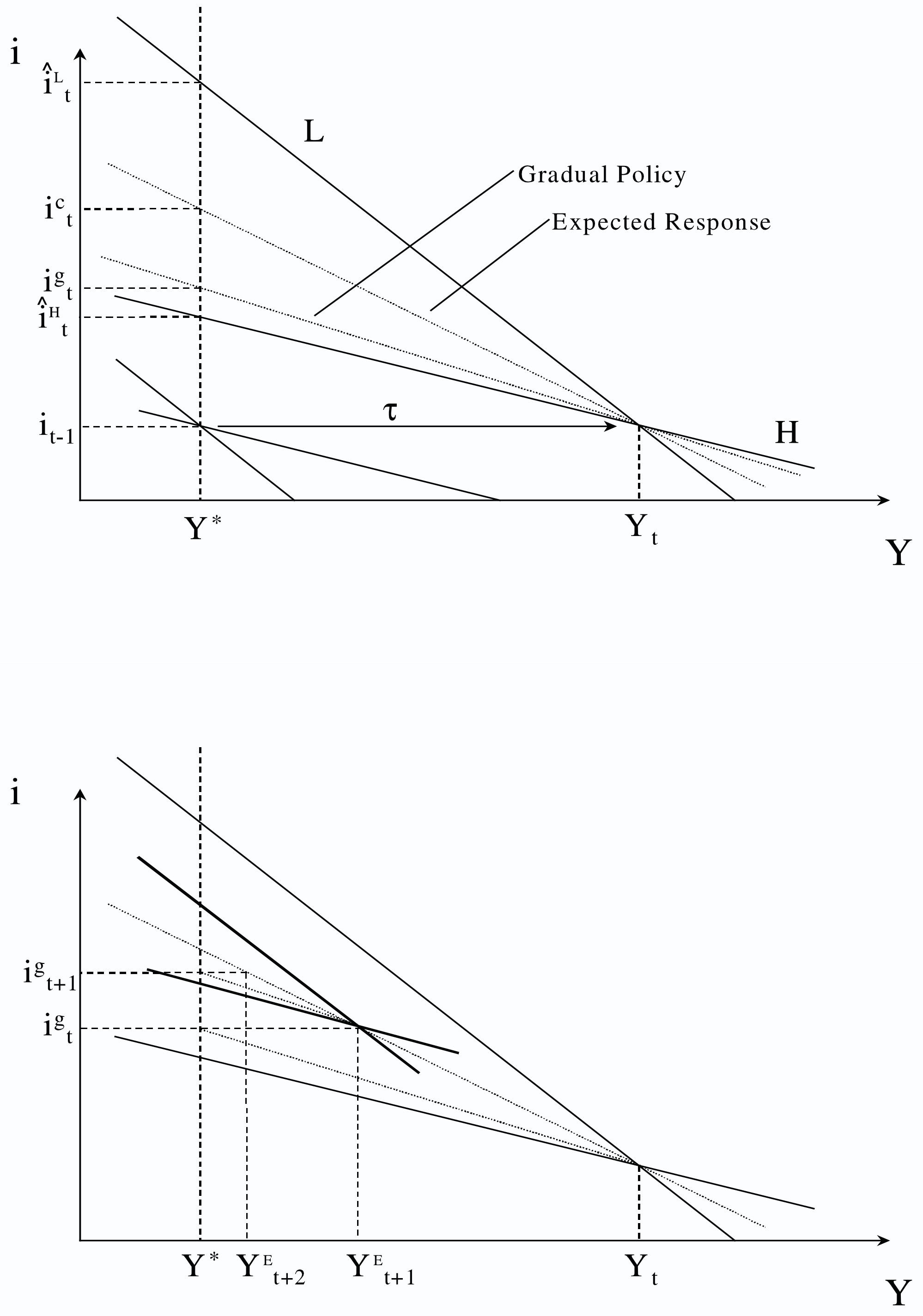
Uncertainty about the effect of monetary policy therefore damps the change in the interest rate that is implemented in response to a deviation of output from target. Using equations (5) and (6), the optimal interest rate policy rule (9) can be expressed in differences:

$$\Delta i_t = \left(\frac{\omega_t}{\phi^L} + \frac{1-\omega_t}{\phi^H}\right)(Y_t - Y^*). \tag{10}$$

This equation gives the slope of the line in Figure 4 from the point $(Y_t, i_{t-1})$ to $(Y^*, i_t^g)$. The degree of gradualism is determined by the difference between the slope of this line and the slope of the expected response line, which is $\left(\frac{\omega_t^c}{\phi^L} + \frac{1-\omega_t^c}{\phi^H}\right)$. It is easily shown that the completely-offsetting policy responds more strongly to a deviation of output from its target. The reactions are equal only if $P = 0$ or $P = 1$, in which case there is no uncertainty, and the difference in the reaction is monotonic in $(\phi^H - \phi^L)^2$, a measure of the amount of uncertainty generated from ignorance about the true multiplier.

Although the initial policy response may be damped, this increase in the interest rate allows the Fed to learn about the policy multiplier. Even if the true value of policy effectiveness is not revealed, observing the reaction of the economy to the initial interest rate change will reduce the amount of uncertainty that the Fed faces about implementing higher interest

## Figure 4: The Response to an Output Shock



rates, thereby permitting additional interest rate increases. As shown in the bottom panel of Figure 4, following the policy choice $i_t^g$, output is expected to reach $Y_{t+1}^E$. The damping of the initial interest rate change causes output to remain above its target level on average. But since the Fed has observed the reaction to this interest rate change, it faces less uncertainty about additional interest rate changes, as demonstrated by the narrower band of potential reactions (in bold). The Fed solves the same problem at time $t+1$, choosing $i_{t+1}^g$, which will in turn further reduce the uncertainty. The interest rate continues to increase on average until it reaches $i^c$, the rate at which output finally reaches its target. The model can therefore explain gradual movements in the interest rate because of the manner that uncertainty regarding the effect of policy is resolved as the Fed changes the interest rate.

The damping of interest rate changes and the dynamic pattern of uncertainty arising from the learning process can account for the high frequency of continuations observed in the data. It can be shown that the probability of an interest rate increase next period is given by the following expression:

$$\Pr(\Delta i_{t+1} > 0) = \frac{1}{2} + \frac{(Y_t - Y^*)}{2\eta^*}\left[\frac{P_t(1-\omega_t)}{\phi^H} - \frac{(1-P_t)\omega_t}{\phi^L}\right](\phi^H - \phi^L). \tag{11}$$

Since $\Delta i_t > 0$ if and only if $Y_t > Y^*$, the expression gives the probability of a continuation when $Y_t > Y^*$ and a reversal otherwise. Replacing the output deviation by its absolute value therefore gives the probability of a continuation, regardless of the sign of the deviation.

The probability given in (11) is valid for policy rules of the form (9) with any set of weights. Equation (11) can therefore be used to solve for the weights that would yield an equal probability of a continuation and a reversal. As expected, this calculation yields the weights $\omega^c$ that are obtained when the Fed ignores the variance of output. The probability of a continuation under the optimal policy rule is found by substituting the optimal weights into (11), which results in the following expression:

$$\Pr(continuation) = \frac{1}{2} + \frac{\mid Y_t - Y^* \mid}{2\eta^*}\left[\frac{P_{t+1}(1-P_{t+1})(\phi^H - \phi^L)^2}{P_{t+1}(\phi^L)^2 + (1-P_{t+1})(\phi^H)^2}\right]. \tag{12}$$

The probability given in (12) is strictly greater than one half unless $P_{t+1} = 0$, $P_{t+1} = 1$, or $\phi^H = \phi^L$, in which case the probability equals one half. In other words, the probability of a continuation is strictly greater than the probability of a reversal as long as there is some degree of uncertainty.

The analysis can be carried further by describing the pattern of interest rate adjustments in response to an output deviation. Because of the simplification that all uncertainty about the true state is resolved at once, there will be a point when the interest rate jumps up or down according to which value of the multiplier is revealed. However, the dynamic behavior of the interest rate before that time is given by

$$E_t\left[\Delta i_{t+1}\right] = \left[\left(\phi^H - \phi^L\right)\left(\frac{P_t(1-\omega_t)}{\phi^H} - \frac{(1-P_t)\omega_t}{\phi^L}\right)\right]\Delta i_t. \tag{13}$$

Since the probability of each state and the weights are constant, the bracketed coefficient is constant. Conditional on the event that the true state of the economy is not revealed, changes in the interest rate follow an AR(1) process in response to a deviation of output from its target. Note that $E_t\left[\Delta i_{t+1}\right] = 0$ when $P = 1$ or $P = 0$. Absent uncertainty over the effect of policy, the initial response of the interest rate is expected to fully offset a deviation, and the interest rate follows a random walk process.

In the above results all of the serial correlation of interest rate changes result from the resolution of uncertainty. However, output itself is likely to be serially correlated, which might result in additional serial correlation in interest rate changes. Suppose that output has an observable component that does not respond to monetary policy, referred to as autonomous demand $a_t$, so that output is given by:

$$Y_{t+1} = \alpha_{t+1} - \phi \cdot i_t + a_{t+1}. \tag{14}$$

The above results describe the interest rate response to a permanent shock in autonomous demand, for example if autonomous demand follows a random walk, in which case a change in $a_t$ results in an output gap that must be completely offset by monetary policy. A similar solution arises under more general processes for observable output. For any process with a predictable component $\{\widehat{a}_{t+i}\}$ for $i > 0$, the interest rate change is given by

$$\Delta i_t = \left(\frac{\omega_t}{\phi^L} + \frac{(1-\omega_t)}{\phi^H}\right)\left(Y_t - Y^* + \widehat{a}_{t+1} - a_t\right) \tag{15}$$

rather than equation (10). This expression indicates that some portion of the serial correlation in interest rate changes may be generated by serially-correlated movements in the fundamentals. However, the resolution of uncertainty will cause the interest rate to be adjusted more gradually than expected simply from the dynamic behavior of the fundamentals, which is consistent with the evidence discussed in the introduction.

Overall, this model captures the gradual movements in the federal funds rate that are observed. The tentative reaction of the interest rate to output deviations is attributed to the uncertainty over the response of the economy to monetary policy. The manner in which this uncertainty is resolved through time can account for the large number of continuations in the data (equation (12)) and the related serial correlation of interest rate changes (equation (13)). The crucial aspect of the learning process is that the variance-minimizing choice of the interest rate reacts to recent policy movements. In the model, the variance-minimizing interest rate is always equal to the interest rate in the previous period since the effect of that policy choice is observed perfectly. In reality, policymakers observe the impact of a policy decision imprecisely and with some delay. Although the model simplifies these aspects of learning process, it highlights the elements of learning that would lead to interest-rate smoothing in general. To the extent that the Fed learns about policy effectiveness by observing the reaction of the economy to recent levels of the interest rate, uncertainty will limit the willingness to move the interest rate aggressively from those levels.

## 2.3 Incorporating Aggregate Supply

The model to this point determines output through a simple IS curve, without any consideration of aggregate supply and the determination of prices. This subsection presents a brief analysis of the model when aggregate supply is incorporated, which highlights two important issues. The first is whether the model predicts nominal or real interest-rate smoothing. The second is a clarification of the interpretation of the output target in the previous model.

Aggregate demand is again determined by the IS schedule with the same uncertainty structure assumed, only now explicitly considering inflation expectations in determining the real interest rate:

$$Y_{t+1} = \alpha_{t+1} - \phi \cdot (i_t - \pi^e_{t+1}). \tag{16}$$

This modification to the IS schedule immediately raises the issue of whether the model predicts nominal or real interest-rate smoothing. For a given output target $Y^*$, an observable shock to $\pi^e_{t+1}$ warrants a one-to-one reaction of the nominal interest rate. The obvious reason is that the nominal interest rate that completely offsets the shock is also the variance-minimizing rate – both equal the level that keeps the real interest rate constant. The uncertainty about the policy multiplier in this case does not reduce the reaction of the nominal rate, and the model instead involves real interest-rate smoothing.

However, the model also presents an argument for nominal interest-rate smoothing. Interest-rate smoothing arises because policymakers learn from observing the output response to the interest rate, but the interest rate that is observed is the nominal interest rate, not $i_t - \pi^e_{t+1}$. Uncertainty about the real interest rate resulting from ignorance about inflation expectations has the same implications for monetary policy as the other components of uncertainty about policy effectiveness. In particular, mistakes about the magnitude of the real interest rate will be offset by mistakes about the policy multiplier, and uncertainty will tend to be minimized around lagged values of the nominal interest rate. Therefore, to the extent that the ex-ante real interest rate is unknown, the model leads to nominal interest-rate smoothing. Given these considerations, optimal monetary policy will involve some blend of real and nominal interest-rate smoothing as determined by the predictability of changes in inflation expectations. In addition, there may not be a large empirical distinction between the two types of smoothing due to the sluggishness of inflation expectations.

The model incorporates an aggregate supply relationship in which inflation is determined by an expectational Phillips Curve:

$$\pi_{t+1} = \pi^e_{t+1} + \gamma \cdot (Y_{t+1} - Y^p) + u_{t+1}, \tag{17}$$

where $Y^p$ denotes potential output. By assumption, there is no uncertainty about the aggregate supply relationship other than the additive disturbance $u_{t+1}$.[16] A common objective function assumed for the Fed involves losses from deviations of both output from its potential level and inflation from a target level:

$$\max_{i_t} -E_t \left[ (\pi_{t+1} - \pi^*)^2 + \lambda (Y_{t+1} - Y^p)^2 \right]. \tag{18}$$

For a known value of $\pi^e_{t+1}$, it can be shown that optimal policy is determined by the following equation:

$$i_t = \pi^e_{t+1} + \omega_t \left( \frac{\alpha^L_t - Y^*}{\phi^L} \right) + (1 - \omega_t) \left( \frac{\alpha^H_t - Y^*}{\phi^H} \right). \tag{19}$$

The policy rule therefore is of the same form as equation (9). Incorporating the aggregate supply relationship only affects the level of output that is targeted by monetary policy, which is given by

$$Y^* = \frac{\lambda}{\gamma^2 + \lambda} Y^p + \frac{\gamma^2}{\gamma^2 + \lambda} Y^{\pi^*}, \tag{20}$$

where $Y^{\pi^*} = Y^p - \frac{\pi^e_{t+1} - \pi^*}{\gamma}$ is the level of output such that inflation is expected to exactly reach its target. Policymakers therefore target a weighted average of potential output and the level of output required to reach the inflation target. The weight on potential output increases with the relative importance of the output gap in the Fed's preferences, and the

[16] This paper concentrates entirely on uncertainty over the effect of monetary policy in determining output. For an analysis that focuses on uncertainty in the aggregate supply relationship, see Wieland(1998).

weight on inflation increases as the ability to affect the inflation rate increases, as reflected in the aggregate supply parameter $\gamma$. The process of learning about the policy multiplier therefore results in the gradual movements in the interest rate demonstrated in the previous section, given the more explicit definition of the output target in equation (20).

# 3 Stochastic Policy Effectiveness and the Dynamics of Uncertainty

This section investigates the effect of uncertainty in a model in which the policy multiplier is not limited to two potential values, but instead follows a stochastic process. Under this modification, the analysis generalizes the findings of section 2 to the case where the learning process is described by a Kalman filter, so that the beliefs about the policy multiplier are updated at each point in time. In addition, while movements in the interest rate allow the Fed to learn about the policy multiplier, the uncertainty now rebuilds through time because the effect of monetary policy is stochastic. The degree of uncertainty will therefore vary in a manner determined by recent policy actions, which has additional implications for the gradual behavior of monetary policy that are discussed below.

Output is determined by an IS equation as in the previous section as follows:

$$Y_{t+1} = \alpha_{t+1} - \phi_{t+1} \cdot i_t. \tag{21}$$

Unlike the previous model, both parameters governing this equation are assumed to evolve through time according to the following equations:

$$(\alpha_{t+1} - \overline{\alpha}) = \delta_\alpha \cdot (\alpha_t - \overline{\alpha}) + v_{t+1}^{\alpha} \tag{22}$$

$$(\phi_{t+1} - \overline{\phi}) = \delta_\phi \cdot (\phi_t - \overline{\phi}) + v_{t+1}^{\phi}, \tag{23}$$

where the vector of disturbances $v = \left[ \begin{array}{cc} v^\alpha & v^\phi \end{array} \right]'$ is normally distributed with mean zero and a variance matrix given by $E[v \cdot v'] = Q = \left[ \begin{array}{cc} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\phi^2 \end{array} \right]$.

By assumption, the Fed knows the mean value of the parameters governing output but does not observe the deviations of these variables from their means. It must learn about the parameters by observing the level of output that results from the choice of the interest rate. The Fed's beliefs about the parameters in equations (22) and (23) are summarized by $\widehat{\alpha}_{t+1|t}$ and $\widehat{\phi}_{t+1|t}$, the expected value of the time $t+1$ parameters based on information at time $t$. These beliefs are updated based on the signal obtained from equation (21). Because these parameters are normally distributed, the beliefs about the two parameters will be normally distributed about their means, with the variance of the time $t+1$ parameters based on time $t$ information denoted $P_{t+1|t} = E_t\left(\begin{bmatrix} (\alpha_{t+1} - \widehat{\alpha}_{t+1|t}) \\ (\phi_{t+1} - \widehat{\phi}_{t+1|t}) \end{bmatrix}\begin{bmatrix} (\alpha_{t+1} - \widehat{\alpha}_{t+1|t}) \\ (\phi_{t+1} - \widehat{\phi}_{t+1|t}) \end{bmatrix}'\right)$. This variance matrix measures the uncertainty about the policy multiplier that faces the Fed when determining the interest rate at time $t$. The expected values of the parameters and the uncertainty associated with those beliefs evolve according to the following equations:

$$\begin{bmatrix} \widehat{\alpha}_{t|t} \\ \widehat{\phi}_{t|t} \end{bmatrix} = \begin{bmatrix} \widehat{\alpha}_{t|t-1} \\ \widehat{\phi}_{t|t-1} \end{bmatrix} + P_{t|t-1}H_t\left(H_t'P_{t|t-1}H_t\right)^{-1}\left(Y_t - H_t'\begin{bmatrix} \widehat{\alpha}_{t|t-1} \\ \widehat{\phi}_{t|t-1} \end{bmatrix}\right) \tag{24}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H_t\left(H_t'P_{t|t-1}H_t\right)^{-1}H_t'P_{t|t-1} \tag{25}$$

$$\begin{bmatrix} (\widehat{\alpha}_{t+1|t} - \overline{\alpha}) \\ (\widehat{\phi}_{t+1|t} - \overline{\phi}) \end{bmatrix} = F \cdot \begin{bmatrix} (\widehat{\alpha}_{t|t} - \overline{\alpha}) \\ (\widehat{\phi}_{t|t} - \overline{\phi}) \end{bmatrix} \tag{26}$$

$$P_{t+1|t} = F \cdot P_{t|t} \cdot F' + Q, \tag{27}$$

where the vector $H_t' = \begin{bmatrix} 1 & -i_{t-1} \end{bmatrix}$ and $F = \begin{bmatrix} \delta_\alpha & 0 \\ 0 & \delta_\phi \end{bmatrix}$.

These equations describe a standard Kalman filter in which there is no noise in the observation equation. The critical feature of the model, as in section 2, is that although the Fed faces uncertainty about the parameters determining output, it observes the level of output that is realized from its choice of the policy instrument. The interaction of the uncertainty about $\alpha$ and $\phi$ that arises from observing $Y_t$ in the learning process results in interest-rate smoothing under the optimal policy. In the learning process described by equations (24) to (27), uncertainty tends to be minimized around recent levels of the interest rate because any error in the estimated slope of the IS curve is offset by a mistake about the intercept coefficient. To demonstrate this, consider the uncertainty that policymakers face after observing output at time $t$, assuming that the parameters of the IS equation will remain unchanged next period. If the interest rate is set to $i^*$, the variance of output $Y_{t+1}$ will be given by the non-negative quadratic expression $\begin{bmatrix} 1 & -i^* \end{bmatrix} P_{t|t} \begin{bmatrix} 1 \\ -i^* \end{bmatrix}$. However, pre- and post-multiplying equation (25) by $H_t$ indicates that $\begin{bmatrix} 1 & -i_{t-1} \end{bmatrix}' P_{t|t} \begin{bmatrix} 1 \\ -i_{t-1} \end{bmatrix} = 0$. There is no uncertainty associated with the choice of the interest rate $i_{t-1}$. This is because even though the parameters have positive variance, so that $P_{t|t}$ is positive definite, the response to last period's choice of the interest rate has been observed. All other choices of the interest rate would involve positive variance of output today arising from the uncertainty over the two parameters. However, at $i^* = i_{t-1}$, any mistake about the parameter $\alpha$ is perfectly offset by a mistake about the parameter $\phi$.

To further analyze the structure of the parameter uncertainty, note that equation (25) also implies that $H_t' P_{t|t} = 0$. In order to satisfy this equation, the matrix $P_{t|t}$ must have the following form:

$$P_{t|t} = \begin{bmatrix} i_{t-1}^2 \Sigma_t & i_{t-1}\Sigma_t \\ i_{t-1}\Sigma_t & \Sigma_t \end{bmatrix}, \tag{28}$$

where $\Sigma_t$ is defined as the variance of the variance of $\phi_{t+1}$ as of time $t$. The parameter $\Sigma_t$ and the lagged interest rate $i_{t-1}$ therefore completely determine the structure of the uncertainty in this model. Equation (28) implies that the correlation between $\alpha_t$ and $\phi_t$ based on information at time $t$ equals 1. It is this perfect correlation between the two parameters that minimizes the uncertainty around lagged values of the interest rate. This effect is the "hedging benefit" that Balvers and Cosimano(1994) discuss in the context of disinflation policy, arguing that it may be optimal to disinflate in a gradual manner, since it reduces the uncertainty associated with the disinflationary policy. Their analysis arrives at the restrictions on the structure of parameter uncertainty that are summarized in equation (28).

The policy choice at time $t$ will be a function of the parameter estimates $\widehat{\alpha}_{t+1|t}$ and $\widehat{\phi}_{t+1|t}$ as well as the uncertainty associated with those estimates $P_{t+1|t}$. For a particular choice of the interest rate today, the expected level and variance of output are determined by the following equations:

$$\begin{aligned} E_t[Y_{t+1}] &= \widehat{\alpha}_{t+1|t} - \widehat{\phi}_{t+1|t} \cdot i_t & (29) \\ Var_t[Y_{t+1}] &= (\delta_\alpha \cdot i_{t-1} - \delta_\phi \cdot i_t)^2 \Sigma_t + \sigma_\alpha^2 + i_t^2 \cdot \sigma_\phi^2. & (30) \end{aligned}$$

The interest-rate smoothing effect is evident from equation (30). If $\delta_\alpha = \delta_\phi \equiv \delta$, the variance is given by $\delta^2 (\Delta i_t)^2 \Sigma_t + \sigma_\alpha^2 + i_t^2 \cdot \sigma_\phi^2$. The first term indicates that large interest rate *changes* result in a higher variance associated with output. The variance is also associated with the interest rate level only because the slope of the IS schedule is affected by a stochastic shock.

The optimal policy can be solved numerically for the infinite-horizon objective function given in equation (7). Given the structure of $P_{t+1|t}$, there will be only four state variables for the problem: $\widehat{\alpha}_{t+1|t}$, $\widehat{\phi}_{t+1|t}$, $i_{t-1}$, and $\Sigma_t$. The optimal policy will be the solution to the following Bellman equation:

$$V\left(\widehat{\alpha}_{t+1|t}, \widehat{\phi}_{t+1|t}, i_{t-1}, \Sigma_t\right) = \max_{\{i_t\}} -(E_t[Y_{t+1}] - Y^*)^2 - Var_t(Y_{t+1}) \\ +\beta \cdot E_t\left[V\left(\widehat{\alpha}_{t+2|t+1}, \widehat{\phi}_{t+2|t+1}, i_t, \Sigma_{t+1}\right)\right] \tag{31}$$

subject to equations (29) and (30) and the dynamic behavior described by equations (24) to (27). It can be shown that the dynamic behavior of the parameter uncertainty is governed by the following equation:

$$\Sigma_{t+1} = \frac{\Sigma_t(\delta_\phi^2\sigma_\alpha^2 + \delta_\alpha^2\sigma_\phi^2 i_{t-1}^2) + \sigma_\alpha^2\sigma_\phi^2}{\Sigma_t(\delta_\phi i_t - \delta_\alpha i_{t-1})^2 + \sigma_\alpha^2 + i_t^2\sigma_\phi^2}, \tag{32}$$

which can potentially range between zero and $\frac{\sigma_\phi^2}{1-\delta_\phi^2}$. The parameterization of the problem is chosen to effectively demonstrate the impact of the parameter uncertainty on the optimal behavior of the interest rate, with the following values: $\delta_\alpha = \delta_\phi = .95$, $\overline{\alpha} = 0$, $\overline{\phi} = 8$, $\sigma_\alpha^2 = 0.5$, and $\sigma_\phi^2 = 1$.

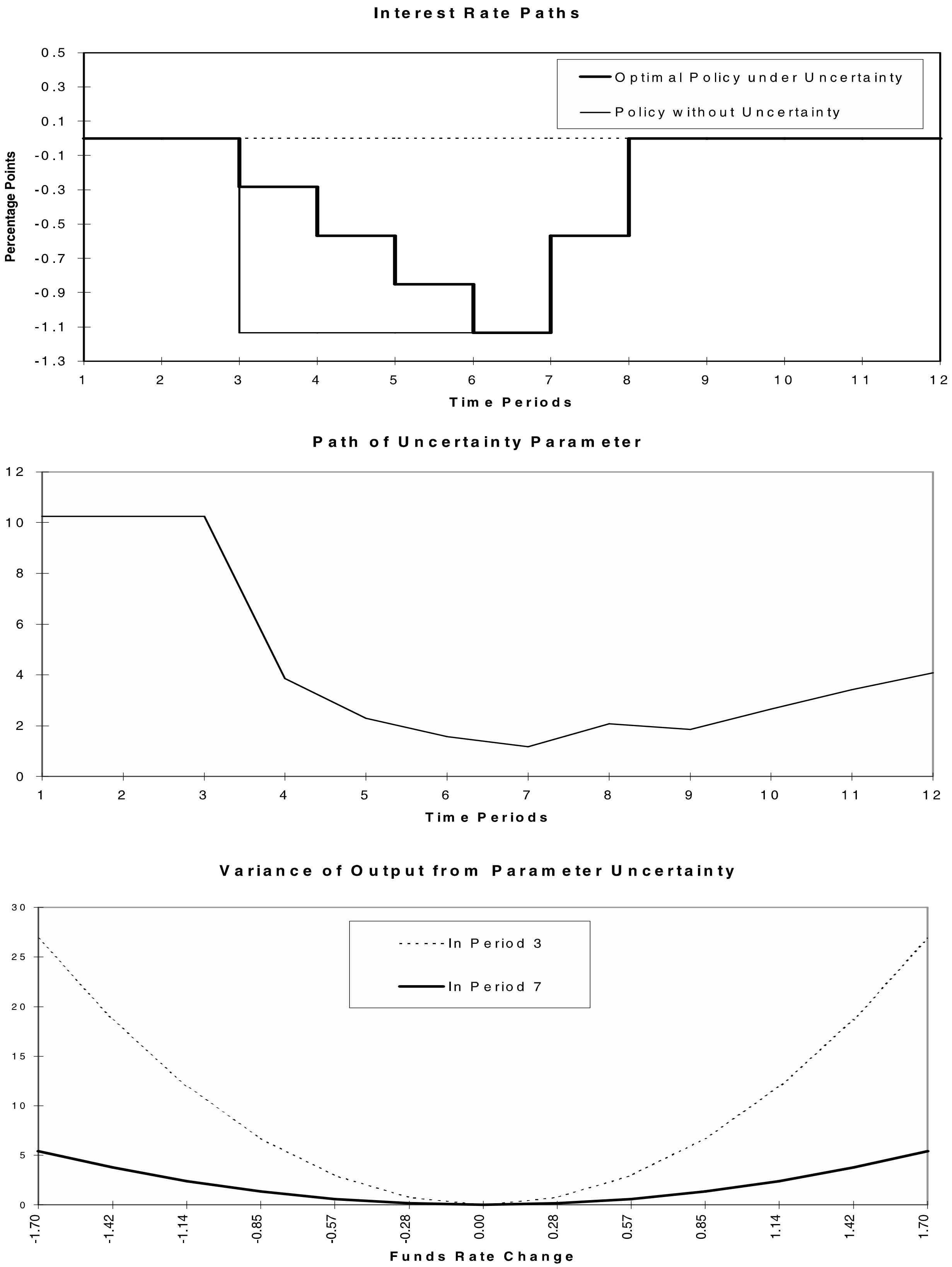
Figure 5 displays the dynamic response of the interest rate and the uncertainty over policy effectiveness in response to two successive output shocks. The economy begins in an initial state that is known by the Fed, with $\alpha$ at its steady-state value $\widehat{\alpha} = \alpha = 0$ and $\phi$ at a low value $\widehat{\phi} = \phi = 1.59$. It is assumed that the interest rate has been inactive, holding at its equilibrium level of 0 at which output equals its target. The first shock decreases $\alpha$ by a magnitude of 2.26 at time period 3. If the Fed were not concerned about its ignorance over the policy multiplier, it would decrease the interest rate immediately by 1.14%, as indicated by the thin line in the top panel of Figure 5.[17]

The effect of uncertainty is evident in the behavior of the optimal policy under uncertainty, given by the thicker line in the figure. Because of the inactivity of the interest rate before the shock, equation (32) indicates that the uncertainty about the policy multiplier will settle to its maximum level $\Sigma = \frac{\sigma_\phi^2}{1-\delta_\phi^2}$, or $\Sigma = 10.26$ under the chosen parameters. The high level of uncertainty makes the Fed unwilling to implement the aggressive decrease in the interest rate required to offset the shock. This is shown in the bottom panel of the figure, which demonstrates that the variance of output is minimized at the lagged value of the interest rate and accelerates in the magnitude of the interest rate change.[18] Because the

[17]This policy is not exactly the completely-offsetting policy previously discussed, which ignores all variance in output. This policy takes into consideration that the level of the interest rate affects the variance of output because of the stochastic shock $v^\phi$. It ignores the variance that otherwise arises from the Fed's ignorance.

[18]This is true for the portion of the variance of output arising from the Fed's ignorance about the parameters, shown in the figure, ignoring the portion arising from the variance of the shocks $v^\alpha$ and $v^\phi$.

# Figure 5: Policy Response to Two Shocks



aggressive interest rate movement would induce a substantial amount of variance in output, the optimal policy limits the interest rate change and instead gradually decreases the interest rate in four successive changes of about 28 basis points each.[19]

The middle panel of Figure 5 displays the dynamic behavior of uncertainty that results from the process of learning. Observing the reaction of the economy to the four changes in the interest rate following the first shock provides the Fed with information about the policy multiplier, so that the measure of uncertainty $\Sigma$ falls from 10.26 to 1.18 at time period 7. As a result, the Fed is willing to implement more aggressive interest rate changes at the time of the second shock. The second shock takes place in period 7 and returns the parameter $\alpha$ to its original value of 0. Unlike the response to the first shock, the increase in the interest rate under the optimal policy is aggressive, coinciding with the policy that ignores parameter uncertainty.[20] The uncertainty over the policy multiplier has decreased enough that it does not temper the policy reaction to the second shock. The difference in the circumstances is evident in the bottom panel, which shows that the variance of output does not increase as rapidly in the change in the interest rate at the time of the second shock.

The model indicates that periods of active interest rate movements reduce the uncertainty that the Fed faces regarding the policy multiplier, which leads to more aggressive responses of the interest rate. However, once the interest rate has fully reacted to the shocks and again becomes inactive, no additional information is being received about the policy multiplier. Accordingly, the uncertainty begins to increase over time, with the $\Sigma$ parameter reaching 4.07 by period 12 and eventually returning to its maximum level of 10.26. As a result, just as periods of active interest rate movements lead to more aggressive policy movements, periods of inactive interest rate policy may damp future policy movements by limiting the opportunity to learn about the policy multiplier.

# 4 Conclusion

The tendency to implement gradual movements in the federal funds rate does not necessarily indicate that the Fed has an explicit interest rate smoothing objective. Uncertainty over the effect of monetary policy on the economy can account for gradual interest rate movements under a traditional objective function concerned only with the deviations of output and inflation from their respective targets. The reason is that the Fed must learn about the

[19] The interest rate changes are of exactly the same magnitude only because of the discrete grid used in obtaining the solution.

[20] The reaction occurs in two steps only because of the path of the Fed's beliefs in learning the source of the shock.

policy multiplier by observing the reaction of the economy to previous choices of the interest rate. Because the structure of the economy varies stochastically, recent observations are particularly informative. As a result, the Fed faces more uncertainty about the reaction of the economy as it moves the funds rate away from its recent levels, which limits its willingness to implement aggressive interest rate changes.

The learning process leads to gradual movements in the funds rate because the evolution of uncertainty is determined by the policy choices that have been recently implemented. Observing the output response to an interest rate change provides the Fed with information about the impact of its policy. This information arrives in many forms and at various speeds. Some of the transmission will arise through the uncertain reaction of asset markets, which the Fed may observe immediately, while the actual response of output is observed only after a substantial amount of time. As information about the policy multiplier is revealed, the uncertainty associated with movements in the interest rate diminishes, thereby permitting additional interest rate changes. This gradual adjustment of the interest rate continues until output reaches its desired level.

Because active interest rate changes allow the Fed to learn more effectively about the policy multiplier, the degree of uncertainty may vary through time. Periods of active interest rate movements should be followed by a more aggressive policy rule in which the Fed engages in less interest-rate smoothing. To investigate whether this is indeed the case, the Taylor-type policy rule (2) estimated in the introduction can be modified to allow the reaction speed of the funds rate to depend on a measure of funds rate activity. Since monetary policy affects output with a substantial delay, the information that the Fed currently receives about the policy multiplier will depend on lagged funds rate movements. Accordingly, a measure of funds rate activity is adopted that involves the absolute magnitude of quarterly funds rate changes over a year, with a lag of three quarters, given by $\left(\sum_{j=3}^{6} |\Delta i_{t-j}|\right)$.[21] The degree of interest-rate smoothing is allowed to react to this measure by modifying the parameter $\rho$ in policy rule (2) as follows:

$$\rho = \rho_0 + \rho_1 \cdot \left( \sum_{j=3}^{6} |\Delta i_{t-j}| \right). \tag{33}$$

Results from estimating this policy rule are presented in Table 2.

The results indicate that the degree of interest-rate smoothing varies significantly with the measure of lagged funds rate activity, while the coefficients on the inflation and output deviations remain plausible and significant. The Fed implements a more aggressive

[21]The timing is consistent with the finding from numerous VAR studies that a large portion of the effect of a funds rate shock on output occurs between six months and two years following the shock.

**Table 2: Estimated Policy Rule with Varying Adjustment Speed**

| | Least Squares Estimates | Instrumental Variables |
|---|---|---|
| $\beta_0$ | 1.97 (2.56) | 3.27 (3.29) |
| $\beta_y$ | 1.22 (8.32) | 1.33 (7.31) |
| $\beta_\pi$ | 1.38 (5.74) | 0.99 (3.36) |
| $\rho_0$ | 0.87 (6.69) | 0.96 (8.33) |
| $\rho_1$ | -0.13 (-1.71) | -0.18 (-2.21) |
| $\overline{R}^2$ | 0.970 | 0.968 |
| $SEE$ | 0.341 | 0.350 |

Notes: Results shown are LS and IV estimates of the equation:

$$i_t = (1-\rho)\left(\beta_0 + \beta_y(Y_t - Y^*) + \beta_\pi \pi_t\right) + \rho \cdot i_{t-1}$$

where $\rho = \rho_0 + \rho_1 \cdot \left(\sum_{j=3}^{6} |\Delta i_{t-j}|\right)$.

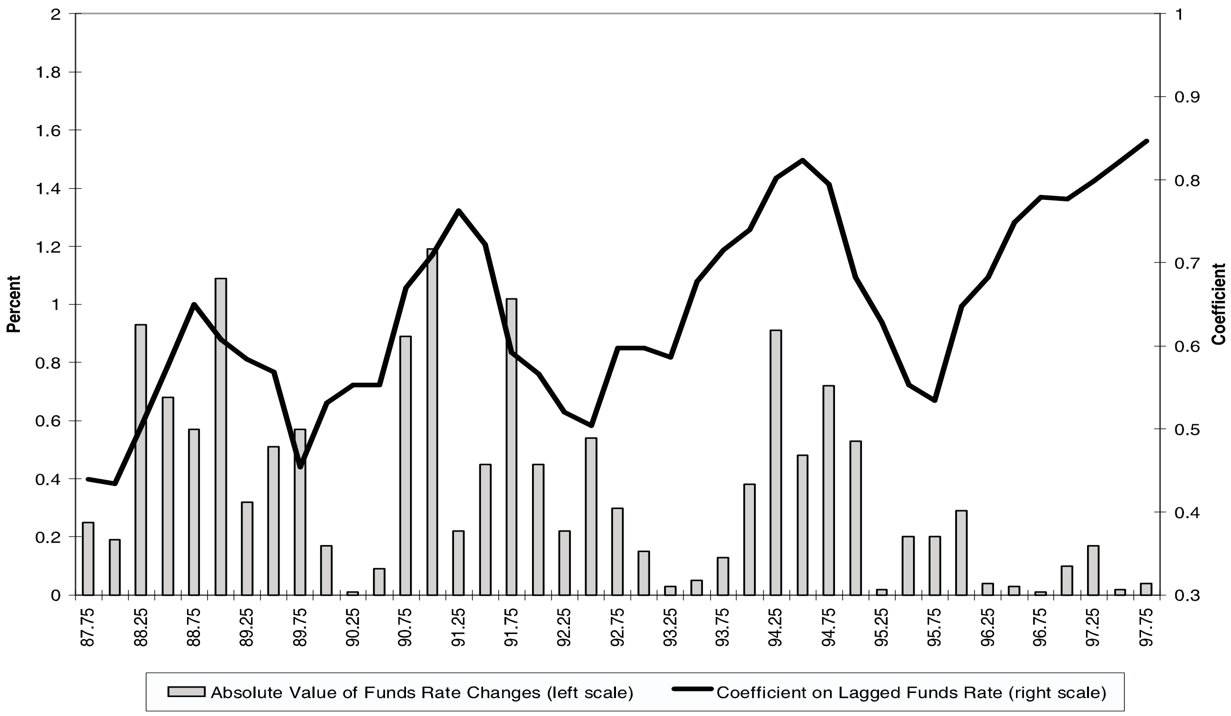
The t-statistics shown in parentheses are based on Newey-West standard errors.

Instruments include one to four lags of inflation, the output gap, and the funds rate.

policy rule following active movements in the funds rate. The $\rho$ parameter implied by the least-squares estimates has an average value of 0.64 and a standard deviation of 0.11. The parameter varies from a minimum of 0.43, when the funds rate is responding to output and inflation deviations very quickly, to a maximum of 0.85, when the adjustment speed is quite low.

The variation in the adjustment speed over the sample is displayed in Figure 6, along with the absolute value of changes in the funds rate. As can be seen in the figure, periods during which the funds rate is actively changed are followed by a decline in the interest-rate smoothing parameter, consistent with the resolution of uncertainty described in the model. Similarly, the degree of interest-rate smoothing increases during periods of low funds rate activity, such as those preceding the funds rate reversals in February 1994 and March 1997. This finding indicates that the low funds rate activity during these periods reflects in part that the inactivity itself causes the Fed to react more tentatively to macroeconomic developments. According to the model, these are periods of heightened uncertainty, as the inactivity limits the flow of information regarding the policy multiplier. This evidence, while brief and therefore only suggestive, indicates that uncertainty may be an important determinant in the formulation of monetary policy, one that warrants further research.

**Figure 6: Variation in the Degree of Gradualism**




# References


[1] Balvers, Ronald J. and Thomas F. Cosimano. Inflation Variability and Gradualist Monetary Policy. *Review of Economic Studies* 61 (1994), 721-738.

[2] Bank for International Settlements. 68th Annual Report (1998).

[3] Barro, Robert J. Interest-Rate Smoothing. NBER Working Paper #2581 (1988).

[4] Bertocchi, Graziella and Michael Spagat, Learning, Experimentation, and Monetary Policy. *Journal of Monetary Economics* 32 (1993), 169-183.

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

[6] Brainard, William. Uncertainty and the Effectiveness of Policy. *American Economic Review, Papers and Proceedings* 57 (1967), 411-425.

[7] Bryant, Ralph C., Peter Hooper, and Catherine Mann, eds., *Evaluating Policy Regimes: New Research in Empirical Macroeconomics.* Brookings: Washington, D.C. (1993).

[8] Caplin, Andrew and John Leahy. Monetary Policy as a Process of Search. *American Economic Review* 86 (1996), 689-702.

[9] Clarida, Richard, Jordi Gali, and Mark Gertler. Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. Mimeo (1997).

[10] Cukierman, Alex. Why Does the Fed Smooth Interest Rates? in Michael T. Belognia, ed., *Monetary Policy on the 75th Anniversary of the Federal Reserve System* (1991), 111-147.

[11] Diamond, Peter. Ignorance and Monetary Policy. Mimeo (1985).

[12] Goodfriend, Marvin. Interest Rate Smoothing and Price Level Trend-Stationarity. *Journal of Monetary Economics* 19 (1987), 335-348.

[13] Goodfriend, Marvin. Interest Rates and the Conduct of Monetary Policy. *Carnegie-Rochester Conference Series on Public Policy* 34 (1991), 7-30.

[14] Mankiw, N. Gregory. The Optimal Collection of Seignorage: Theory and Evidence. *Journal of Monetary Economics* 20 (1987), 327-341.

[15] Mankiw, N. Gregory and Jeffrey A. Miron. Should the Fed Smooth Interest Rates? The Case of Seasonal Monetary Policy. *Carnegie-Rochester Series on Public Policy* 34 (1991), 41-70.

[16] Meulendyke, *U.S. Monetary Policy and Financial Markets*. New York: Federal Reserve Bank of New York (1990).

[17] Orphanides, Athanasios. Monetary Policy Rules Based on Real-Time Data. Finance and Economics Discussion Series, 1998-3. Board of Governors of the Federal Reserve System (1998).

[18] Rotemberg, Julio and Michael Woodford. An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy. Mimeo (1997).

[19] Rudebusch, Glenn D. Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure. *Journal of Monetary Economics* 35 (1995), 245-274.

[20] Sack, Brian. Does the Fed Act Gradually? A VAR Analysis. Finance and Economics Discussion Series, 1998-17, Board of Governors of the Federal Reserve System (1998).

[21] Stokey, Nancy, Robert Lucas, and Edward Prescott. *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press (1989).

[22] Svensson, Lars E. O. Inflation Targeting: Some Extensions. NBER Working Paper No. 5962 (1997).

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

[24] Wieland, Volker. Monetary Policy and Uncertainty about the Natural Unemployment Rate. Finance and Economics Discussion Series, 1998-22, Board of Governors of the Federal Reserve System (1998).

[25] Williams, John C. Simple Rules for Monetary Policy. Mimeo (1997).

# Appendix: Experimentation and Optimal Policy

In the model presented in section 2, the Fed ignores the fact that the policy rule chosen affects the speed at which it will learn about the effectiveness of its policy instrument. Optimal monetary policy will instead be influenced by the incentive to obtain more information, since this will allow the Fed to implement more successful policy in the future. In other words, the Fed will engage in experimentation rather than learn passively. This subsection explores the implications of active rather than passive learning for the results regarding gradual monetary policy.

The simple form of the uncertainty assumed in section 2 provides a framework in which experimentation is very beneficial. In particular, the probability of learning the true state of the economy is linear in the magnitude of the interest rate change. Once the true state is revealed, the Fed can exercise very effective control over the economy, since the slope of the IS schedule is constant. In a more general structure an increase in the interest rate might only reveal information about the effect of policy around the new interest rate. While the model may overstate the benefits of experimentation, this structure does simplify the analysis and is therefore maintained.

The Fed solves the problem described in equation (7), only now accounting for the effect of its interest rate choice on the probability of learning the state of the economy. The solution to the optimization problem is characterized by the following Bellman equation:

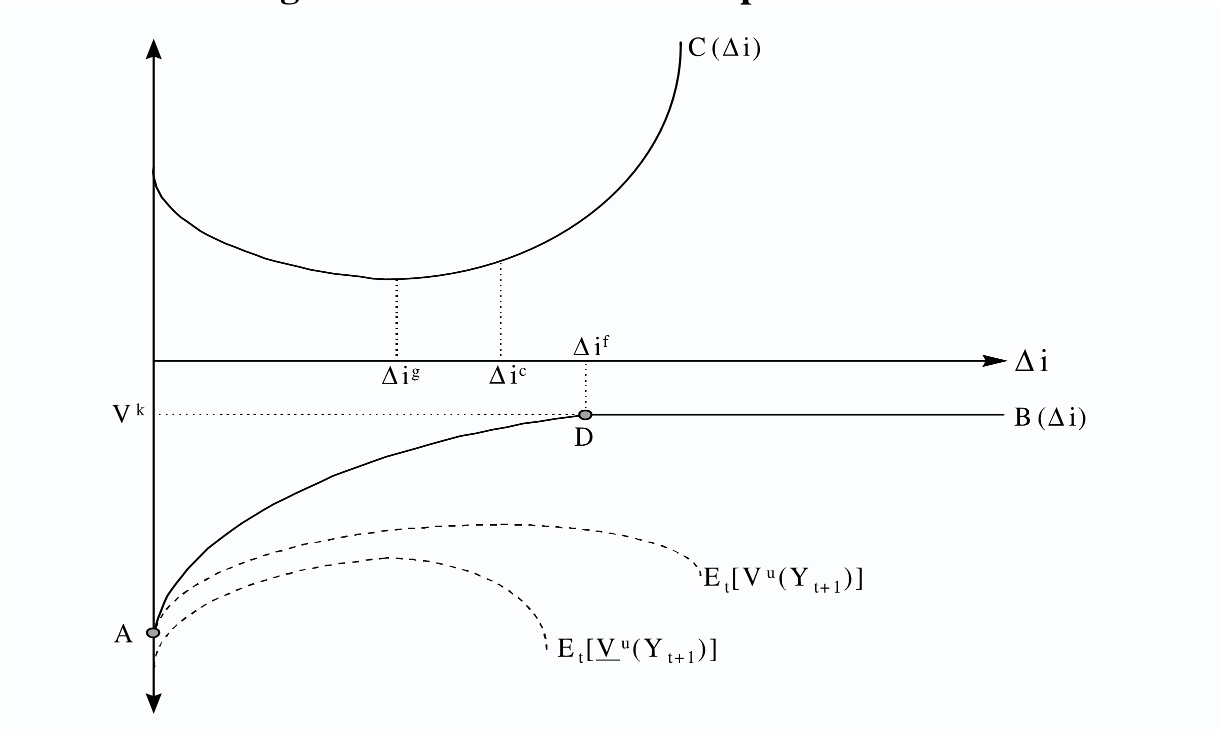
$$V^u(Y_t) = \max_{\Delta i_t} \left\{ -\left(E_t[Y_{t+1}] - Y^*\right)^2 - Var_t(Y_{t+1}) + \beta \cdot E_t\left[V(Y_{t+1})\right] \right\}, \tag{34}$$

where $V^u(Y_t)$ is the value of the problem given that the policy multiplier is unknown. No closed-form solution for the case with experimentation exists, making it difficult deriving conditions under which policy will be gradual. However, the analysis that follows characterizes the various components of the problem to provide some insight into the optimal policy under experimentation.

The sum of the first two terms of equation (34) equals the (negative) current loss function from the policy choice, which is denoted $C(\Delta i_t) \equiv \left(E_t[Y_{t+1}] - Y^*\right)^2 + Var_t(Y_{t+1})$. This loss can be expressed as a function of current output and the change in the interest rate as follows:

$$\begin{aligned} C(\Delta i_t) = \;& (Y_t - Y^*)^2 - 2(Y_t - Y^*)(P\phi^L + (1-P)\phi^H)\Delta i_t + \\ & (P(\phi^L)^2 + (1-P)(\phi^H)^2)\Delta i_t^2 + \eta^2/3. \end{aligned} \tag{35}$$

**Figure 7: The Model with Experimentation**




This function is referred to as the cost function and is displayed in Figure 7 in response to a positive output deviation. Equation (35) indicates that the cost function is convex in the interest rate change. In addition, since the problem in section 2 is concerned only with minimizing this cost, the function reaches its minimum at the optimal policy rule given by equation (8), denoted in the figure by $\Delta i^g$. The diagram indicates the location of the completely-offsetting policy, denoted $\Delta i^c$, which (as shown in section 2) involves a larger interest rate change than $\Delta i^g$.

The figure also displays the continuation value of the problem, given by the last term in equation (34). Because it measures the value of additional information that the Fed learns, this component is referred to as the benefit function, denoted by $B(\Delta i_t) \equiv E_t\left[V(Y_{t+1})\right]$. The benefit function captures the informational component of the problem that was ignored under the policy $\Delta i^g$. While the benefit function cannot be directly calculated (since it depends on the value function), the shape of the function can be characterized using the figure. The expected value function next period is given by the following weighted sum:

$$E_t\left[V(Y_{t+1})\right] = q(\Delta i_t)V^k + (1 - q(\Delta i_t))\, E_t\left[V^u(Y_{t+1})\right], \tag{36}$$

where $V^k$ is the value function conditional on knowing the policy multiplier and $q(\Delta i_t)$ is the probability of learning the value of the multiplier. The value $V^k$ is independent of the level of output or the value of the multiplier. Once the policy multiplier is revealed, the value function equals the negative discounted sum of the variance of output arising from the stochastic shocks only, so that $V^k = -\frac{1}{1-\beta}\frac{(\eta^*)^2}{3}$.

For a given change in the interest rate, the probability of learning the true policy multiplier, $q(\Delta i_t)$, increases according to $q(\Delta i_t) = \frac{(\phi^H - \phi^L)\Delta i_t}{2\eta^*}$. Since the probability of learning the true state is linear in $\Delta i_t$, there is a large enough interest rate change, $\Delta i^f \equiv \frac{2\eta^*}{\phi^H - \phi^L}$, for which the Fed learns the true state of the economy with certainty. The benefit function is therefore constant and equal to $V^k$ for interest rate changes equal to or greater than $\Delta i^f$. Whereas the policies $\Delta i^g$ and $\Delta i^c$ depend on the deviation of output from target, the fully-revealing policy $\Delta i^f$ is instead determined by the magnitude of the shifts in the IS curves. In particular, an increase in $\eta^*$ limits the ability of the Fed to distinguish the two states and therefore requires a larger interest rate change to learn with certainty. For interest rate changes smaller than $\Delta i^f$, the benefit function is given by the weighted average in (36). As shown in Figure 7, this schedule equals $E_t\left[V^u(Y_{t+1})\right]$ when there is no interest rate change, indicated by point $A$. As the interest rate change increases, the weighting shifts from $E_t\left[V^u(Y_{t+1})\right]$ to $V^k$, placing full weight on $V^k$ when $\Delta i = \Delta i^f$ at point $D$.

The optimal policy under experimentation, denoted $\Delta i^e$, considers both the cost and benefit of an interest rate change as measured in these two components. In the case where $\Delta i^g > \Delta i^f$, the policy under experimentation will involve interest rate changes no greater than the gradual policy, since $B'(\Delta i) = 0$ and $C'(\Delta i) > 0$ for all $\Delta i > \Delta i^g$. This is a less-interesting case in which the gradual policy entails learning the true state of the economy with certainty, so that experimentation offers no additional incentive to change interest rates. While it is true that the interest rate response is damped from the uncertainty, the probability of a continuation is simply equal to the probability of the low-responsiveness state. It is therefore assumed from this point on that $\Delta i^g < \Delta i^f$, as depicted in Figure 7. A sufficient condition for gradualism in this case is that $\beta \cdot B'(\Delta i_t) < C'(\Delta i_t)$ for all $\Delta i_t > \Delta i^c$. Policy will not fully offset an output deviation, or $\Delta i^e < \Delta i^c$, if the marginal cost of increasing the interest rate more than $\Delta i^c$ outweighs the benefit.

Given the form of the benefit function in (36), the marginal benefit is given by $B'(\Delta i_t) = q'(\Delta i_t)\left(V^k - E_t\left[V^u(Y_{t+1})\right]\right) + (1 - q(\Delta i_t))\frac{dE_t[V^u(Y_{t+1})]}{d\Delta i_t}$. The effect of learning is measured by the term $q'(\Delta i_t)\left(V^k - E_t\left[V^u(Y_{t+1})\right]\right)$, which depends on the marginal probability of learning and the value of information, $\left(V^k - E_t\left[V^u(Y_{t+1})\right]\right)$. Although it cannot be solved analytically, an upper bound can be calculated for the value of information. Consider a policy that sets the change in the interest rate to $\Delta i^g_{t+s}$ for all periods $s \geq 1$. This policy is sub-optimal, and therefore the value function under the optimal policy will be greater than the value realized from the alternative policy. It can be shown that under the alternative policy, the expected squared deviation of future output will be given by

$$E_{t+1}\left[(Y_{t+n}-Y^*)^2\right]=\Lambda^{n-1}(Y_{t+1}-Y^*)^2+\frac{(\eta^*)^2}{3}\left[\frac{1-\Lambda^{n-1}}{1-\Lambda}\right], \tag{37}$$

where $\Lambda=\frac{P(1-P)(\phi^H-\phi^L)^2}{P(\phi^L)^2+(1-P)(\phi^H)^2}$. Using this expression, and assuming that the true multiplier is never revealed, a lower bound for the value function can be calculated as the discounted sum of the terms in (37):

$$\underline{V}^u(Y_{t+1})=-\frac{\Lambda}{1-\beta\Lambda}(Y_{t+1}-Y^*)^2-\frac{(\eta^*)^2}{3}\frac{1}{(1-\beta)(1-\beta\Lambda)}. \tag{38}$$

Because it lies below the value function for every value of $Y_{t+1}$, the expected value function next period $E_t[V^u(Y_{t+1})]$ is bounded from below by $E_t[\underline{V}^u(Y_{t+1})]$, as shown in Figure 7. This lower bound for the value function places an upper bound on the value of information:

$$\left(V^k-E_t\left[V^u(Y_{t+1})\right]\right)\leq\frac{\Lambda}{1-\beta\Lambda}E_t\left[(Y_{t+1}-Y^*)^2\right]+\frac{(\eta^*)^2}{3}\left[\frac{\beta\Lambda}{(1-\beta)(1-\beta\Lambda)}\right]. \tag{39}$$

The increasing probability of learning and the positive value of information tend to make the benefit function upward sloping in $\Delta i$. The incentive to learn leads to a more aggressive interest rate policy because larger interest rate changes provide a better signal about the true multiplier, a point made in a similar context by Bertocchi and Spagat(1993).[22] However, the incentive to learn will not necessarily outweigh the uncertainty that is generated from more aggressive interest rate changes. In fact, consider the marginal incentive to learn and the marginal cost associated with an interest rate change equal to $\Delta i^c$. It can be shown that $\beta\cdot q'(\Delta i^c)\left(V^k-E_t\left[V^u(Y_{t+1})\right]\right)<C'(\Delta i^c)$ if the following condition holds:

$$\frac{\beta\Lambda}{(1-\beta)(1-\beta\Lambda)}\frac{(\eta^*)^2}{3}<\left(2\Delta i^f-\frac{\beta\Lambda}{1-\beta\Lambda}\right)\frac{\Lambda}{1-\Lambda}(Y_t-Y^*)^2. \tag{40}$$

The fraction $\Lambda$, as defined above, is equal to $\frac{CV}{1-CV}$, where $CV\equiv\frac{P(1-P)(\phi^H-\phi^L)^2}{(P\phi^L+(1-P)\phi^H)^2}$ is the coefficient of variation of the policy multiplier. As the variance of the effect of policy relative to the mean effect of policy increases from zero to infinity, the fraction $\Lambda$ increases from 0 to

[22]In their model, the change in the policy instrument under the optimal policy will always be at least as large as that in the myopic case. A similar definitive result cannot be obtained here because the choice of the policy instrument also affects the value function when the multiplier is not revealed.

1. As $\Lambda$ approaches 1, the left-hand side of equation (40) approaches a finite limit, while the right-hand side has no limit. Equation (40) therefore implies that there exists a coefficient of variation sufficiently large such that the benefit of learning at the completely-offsetting policy will be below the cost of implementing that policy, as long as $\Delta i^f$ is large enough. Furthermore, it can be shown that the marginal incentive to learn will be below the marginal cost of policy for all interest rate changes $\Delta i \geq \Delta i^c$ under the conditions $\Delta i^f > \frac{1}{2}\frac{\beta}{1-\beta}$ and $\beta < \frac{2}{3}$.

These conditions are sufficient conditions for the marginal value of learning to fall below the marginal cost for interest rate changes large enough to reach the desired level of output. These conditions may be much stronger than required because as they are derived using an upper bound that may imprecisely reflect the benefit of learning. However, these conditions reflect the type of restrictions that may in general make gradual interest rate movements more desirable — large amounts of uncertainty over the effect of policy, a high rate of discounting, and an environment in which learning is difficult (a large $\Delta i^f$). In addition, note that these conditions only ensure that the marginal incentive to learn is below the cost of interest rate changes. The marginal benefit of an interest rate change has an additional component involving the effect of the interest rate change when the Fed does not learn the multiplier, $\frac{dE[V_{t+1}]}{\Delta i_t}$. The fact that interest rate changes greater that $\Delta i^c$ are expected to move output farther away from the output target will tend to make this component negative. However, no analytical result can be obtained without knowing the shape of the value function, since the convexity of the value function must also be considered.