Monetary Policy, Gradualism, and the Term Structure of Interest Rates

by

Brian Sack

B.A., Economics and Mathematics
University of Vermont (1992)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

© 1997 Brian Sack. All rights reserved.

The author hereby grants to Massachusetts Institute of Technology permission to reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author

Department of Economics
May 15, 1997

Certified by

Olivier Blanchard
Professor of Economics
Thesis Supervisor

Certified by

Jeremy Stein
J.C. Penney Professor of Management
Thesis Supervisor

Accepted by

Peter Temin
Elisha Gray II Professor of Economics
Chairman, Departmental Committee on Graduate Studies
Monetary Policy, Gradualism, and the Term Structure of Interest Rates

by

Brian Sack

Submitted to the Department of Economics
on May 19, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

There is a widespread belief among economists and financial market participants that the Federal Reserve conducts monetary policy in a gradual manner. The Fed appears reluctant to make aggressive changes in the federal funds rate, instead implementing monetary policy through sequences of small interest rate changes. The first two chapters of the thesis are concerned with investigating the existence of gradualism and offering an explanation for the optimality of a gradual policy rule.

The first chapter presents a model in which optimal monetary policy entails gradual adjustment of the funds rate. The model incorporates uncertainty regarding the effect of monetary policy on the economy that is endogenously determined through a process of learning. As a result of the uncertainty, the Fed dampens its initial response to a shock. It subsequently learns about policy effectiveness by observing the reaction of the economy, which resolves the uncertainty and results in additional, more aggressive interest rate movements. The optimal response to a shock therefore involves successive interest rate movements drawn out over a period of time during which the Fed is learning about the effectiveness of its policy.

The second chapter provides evidence that movements in the federal funds rate are more gradual than can be accounted for by the dynamic behavior of the fundamental variables to which the Fed may be reacting. The analysis computes the funds rate that is expected given the structural form of the economy that is estimated in a VAR. The expected policy is characterized by an immediate and strong reaction of the funds rate to shocks in the economy. The reaction of the actual funds rate is instead described by a dampened reaction to the shock followed by a gradual movement in the direction of the expected rate over the next eight to twelve months. Once uncertainty over the effect of monetary policy is incorporated, the expected policy involves gradual interest rate movements that are more consistent with the data.

The final chapter of the thesis investigates whether monetary policy has implications for the behavior of the term structure of interest rates. The analysis finds that the expectations hypothesis is strongly violated during the endogenous response of monetary policy to innovations in output and prices, as the long-term rate instead closely tracks the response of the federal funds rate. In contrast, the expectations hypothesis successfully describes interest rate responses to financial market shocks, including monetary policy shocks and innovations to longer-term interest rates.
Thesis Supervisor: Olivier Blanchard
Title: Professor of Economics

Thesis Supervisor: Jeremy Stein
Title: J.C. Penney Professor of Management
ACKNOWLEDGEMENTS

There are many people to thank for contributing in various ways to this thesis. However, there is one person without whom this would not have been possible. I will be forever grateful to my wife Amy for her continual support and her incredible patience and understanding over the past four years. She has been able to maintain my perspective about the important things in life, while at the same time constantly encouraging me to towards my goals.

I would also like to thank the other members of my family whose support I could rely on time and time again. I wish to thank my father Terry for his encouraging phone calls. One a day would have probably sustained me, but each additional one helped as well! In addition, I appreciate his hard work, as his personal accomplishments have enhanced his involvement. I would also like to thank my mother Rita and my sister Katie. It was very comforting to know that I could count on them for anything. Finally, Peter and Martha were always there for Amy and me, willing to listen intently for hours to our concerns and frustrations over Martha's apple pie or Peter's Polish pickles.

When I transferred to MIT, I underestimated the impact that my classmates would have on me. What an extraordinary group of people. My interaction with them has been one of the most beneficial aspects of my experience at MIT, and at the same time it has been a lot of fun. I would particularly like to thank Roberto Rigobon. He has been very helpful and encouraging to me as a fellow student, always available to listen to a research idea as long as I buy him a cup of coffee. He has also been a very good friend, and Amy and I appreciate the memories that we have shared with him and his family.

Finally, I cannot begin to express my gratitude to my advisors, Olivier Blanchard and Jeremy Stein, who have had the greatest impact on my development as an economist. They have invited numerous interesting conversations, during which many of the insights in this thesis were developed. In addition, I have benefited from observing their approach to economic and financial research, much of which I attempt to incorporate into my own research methods. Olivier's insight into macroeconomic issues is remarkable, and I have tried to learn from his ability to clarify opaque economic arguments and models. Jeremy's ability to focus on the essential elements of any given problem and deliver solutions that convey powerful intuition has had a substantial influence on my research. In addition to the respect that I have for them
as economists, I truly appreciate the enthusiasm, encouragement, and advice that they have expressed to me.
## Contents

1 Uncertainty and Gradual Monetary Policy ........................................ 11  
   1.1 Introduction ........................................................................ 11  
   1.2 A Model of Gradualism ......................................................... 16  
      1.2.1 The Response to Autonomous Shocks ............................... 23  
      1.2.2 The Response to Investment Demand Shocks .................... 30  
   1.3 Gradualism in the Presence of Experimentation ....................... 32  
   1.4 The Dynamics of Uncertainty ............................................... 39  
   1.5 Conclusion ....................................................................... 45  

2 Does the Fed Act Gradually? A VAR Analysis ................................. 51  
   2.1 Introduction .................................................................... 51  
   2.2 Evidence of Gradualism ...................................................... 54  
      2.2.1 Discussion ................................................................ 54  
      2.2.2 The VAR .................................................................. 56  
      2.2.3 The Analysis ............................................................. 58  
   2.3 Incorporating Uncertainty .................................................... 67  
   2.4 Conclusion .................................................................... 75  

3 The Response of Term Structure Puzzles to Monetary Policy .......... 79  
   3.1 Introduction .................................................................... 79  
   3.2 The Unconditional Behavior of the Term Structure ................ 82  
   3.3 The Reaction of the Term Structure to Monetary Policy .......... 86  
   3.4 A Decomposition of a Puzzle ............................................. 96  

7
List of Figures

1-1 Actual and Target Federal Funds Rate ........................................ 12
1-2 Actual and Expected Funds Rate Reaction ................................. 13
1-3 Potential Distributions ......................................................... 19
1-4 Potential IS Curves ............................................................... 20
1-5 The Response to an Interest Rate Change ................................. 20
1-6 Response to an Autonomous Demand Shock ............................... 26
1-7 The Model with Experimentation ............................................. 34
1-8 Deriving the Sufficient Condition for Gradualism ....................... 37
1-9 Value Functions and Policy Functions ...................................... 42
1-10 Conditional Responses to Small Output Shock ......................... 43
1-11 Conditional Responses to Large Output Shock ......................... 43
1-12 Unconditional Responses to Output Shocks .............................. 45

2-1 Daily Federal Funds Rate and Target ....................................... 52
2-2 Consecutive Target Changes with the Same Sign ....................... 53
2-3 Dynamics of the Actual and Expected Funds Rate ...................... 62
2-4 Impulse Response Functions .................................................. 63
2-5 Impulse Response Functions (cont.) ....................................... 63
2-6 Impulse Response Functions for Subsample .............................. 66
2-7 Impulse Response Functions for Subsample (cont.) ..................... 66
2-8 Impulse Response Functions Under Uncertainty ......................... 68
2-9 Impulse Response Functions Under Uncertainty (cont.) ............... 69
2-10 Weighted Variance as a Function of Interest Rate Choice ........... 71
Chapter 1

Uncertainty and Gradual Monetary Policy

1.1 Introduction

The Federal Reserve conducts monetary policy in a gradual manner, implementing movements in the federal funds rate through sequences of small interest rate changes that may take place over substantial periods of time. On several instances a directional movement in the federal funds rate target has required as many as 25 consecutive target changes. Figure 1-1 displays the target federal funds rate along with the daily funds rate over the episodes September 1974 to September 1979 and March 1984 to the present.\footnote{These are the episodes for which the target series is available. The target funds rate series is that which is reported in Rudebusch(1995) through 1992:9. I have extended the target series to the present by reading the minutes of the FOMC meetings over this period.} As evident from the figure, changes in the funds rate target predict additional changes of the same sign. In fact, 88.4\% of the target changes over the sample are “continuations,” or have the same sign as the previous change. This tendency for the Fed to move the funds rate tentatively is described in more detail in Rudebusch(1995) and is often referred to as gradualism.

While simple statistics demonstrate that the funds rate does not behave as a random walk, this may not indicate that the Fed adjusts the funds rate gradually. One potential explanation of the observed behavior of the funds rate is that changes in the fundamental variables to which
Figure 1-1: Actual and Target Federal Funds Rate
the Fed is reacting are serially correlated. Evidence that movements in the funds rate are too gradual to be accounted for by the dynamic behavior of fundamental variables is presented in a VAR analysis by Sack (1997). The analysis finds that the initial reaction of the Fed to a shock is dampened relative to the response that is expected if the Fed is concerned with stabilizing inflation, unemployment, and output growth. While the expected response is immediate, the observed funds rate response moves gradually in the direction of the expected response over the following eight to twelve months. These dynamics are demonstrated in Figure 1.2 in response to a shock that moves the expected funds rate by 100 basis points. This delayed pattern of funds rate adjustment is the gradual behavior that this paper attempts to explain.

Although the evidence of gradualism is recent, many macroeconomists have previously discussed the tendency of central banks to smooth interest rates. There have been several attempts to incorporate forms of gradualism into models of optimal monetary policy. These attempts typically proceed by appending the objective function of the Fed with a loss related to the magnitude of interest rate changes.

---

2 See Chapter 2 of this thesis.
3 A detailed description of the derivation of this figure, including the definition of the expected response, is given in Chapter 2 of this thesis. The figure is presented here for descriptive purposes only.
Goodfriend(1987) presents a model in which there is a tension between interest rate smoothing and price level stability, so that the goal of macroeconomic stability cannot explain gradualism. To account for gradualism, Goodfriend posits that the Fed smooths interest rates to protect the banking sector against financial crises. In the model the Fed willingly accepts less price stability to achieve its interest rate smoothing goal. Cukierman(1991) formalizes this idea a model of the banking sector in which it is optimal for the Fed to smooth real interest rates to insulate bank profits from shocks.

Other authors have put forward the argument that reversals in the direction of interest rate changes are costly. Meulendyke(1990) expresses the view that the Fed is reluctant to make large interest rate changes in order to avoid reversals. Goodfriend(1991) and others have described the cost of reversing the direction of target changes as "whipsawing the market." However, the meaning of this is not clear, as the cost associated with the interest rate reversal is not formally illustrated.

Other authors have investigated a forward-looking form of interest rate smoothing. In Mankiw(1987), interest rates follow a random walk in order to achieve the optimal collection of seignorage. Barro(1989) investigates the implications of interest rate targeting in a model in which the target follows a random walk, possibly for the reasons discussed in Mankiw(1987). Optimal taxation models, including those that result in variants of the Friedman Rule, tend to yield forward-looking interest rate smoothing, resulting in random-walk behavior for the interest rate. Rather than minimizing future interest rate changes, gradualism involves smoothing changes from previous interest rates, accounting for the fact that the interest rate is not a random walk.

Overall, one should be wary of the approach of amending the objective function of the Fed to include additional objectives that are not clearly justifiable. Hall and Mankiw(1994) note that price and employment stability may require large swings in interest rates and exchange rates, but they conclude that "such side effects should probably be ignored." In general, it is worthwhile to investigate whether gradual behavior can be rationalized in a model that maintains the traditional objective function of macroeconomic stability.

This paper offers an explanation of gradualism based on the idea that the Federal Reserve faces uncertainty about the effect of monetary policy on the economy. In this context, gradual
interest rate adjustment is the optimal policy given the objective of macroeconomic stability. As first analyzed by Brainard (1967), optimal policy under this type of uncertainty limits the use of the policy instrument, to an extent that is increasing with the degree of policy uncertainty. In the model presented, the degree of uncertainty is determined endogenously through a process of rational learning. As a result of the learning process, the uncertainty that the Fed faces will be affected by recent monetary policy. The amount of uncertainty in turn influences the policy that is implemented. It is this interaction between policy and uncertainty and the resulting dynamic behavior of the uncertainty that results in gradual interest rate movements in response to many shocks.

The intuition for the optimality of gradualism is as follows. The Fed faces uncertainty about the amount of output that an interest rate choice will induce. Because it has observed the response of the economy to previous interest rates, this uncertainty will be minimized around recent levels of the interest rate. As a result, the Fed will dampen the interest rate change that it implements in response to a shock. Having adjusted the funds rate, the Fed subsequently observes the reaction of the economy, allowing it to learn about the effectiveness of its policy. Since the interest rate change resolves some of the uncertainty, the Fed becomes more confident to implement additional interest rate changes. This process of learning continues, resulting in an interest rate response that may be implemented gradually over a substantial period of time.

It is obvious that the Fed does not know the precise impact of its policy decisions. What is less obvious is whether the magnitude of the uncertainty that policymakers face is sufficiently large to affect the implementation of policy. Discussions by Fed officials are suggestive that this is the case. The idea is captured nicely in the following quote by Alan Blinder while speaking on the strategy of monetary policy:4

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 ... What can you

---

4This quote is from a talk given at the Minnesota Meeting, a business forum that was held in June 1995.
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.

The paper is organized as follows. Section 1.2 presents a formal model of optimal monetary
policy in which the Fed learns passively about the effect of its policy instrument. It is shown
that the optimal reaction to a shock involves gradual funds rate adjustments, and the degree
of gradualism should vary in response to different types of shocks. The reason is that shocks
affect not only the level of output but also the amount of uncertainty that the Fed faces.
Section 1.3 contains an analysis of the optimal monetary policy when the Fed considers the
amount of information that results from its interest rate choice. In this case, the Fed engages in
experimentation, since it has an incentive to implement strong interest rate changes in order to
learn more about the effectiveness of its policy. Section 1.4 offers a numerical solution to a more
general problem with a focus on the dynamics of policy uncertainty. Section 1.5 concludes.

1.2 A Model of Gradualism

Previous attempts to explain gradualism required accepting motives for smoothing interest
rates that have not been established as part of the objective function for the Federal Reserve.
Gradualism may instead result from the uncertainty that the Fed faces regarding the effectiveness
of monetary policy. This type of uncertainty was first analyzed by Brainard(1967), and
this section begins with a brief review of his result. Consider a static model in which output is
determined by the following equation:

\[ Y_{t+1} = \phi_t \cdot m_t + \mu_{t+1} \]  (1.1)

where \( m \) is the policy instrument of the Fed and \( Y \) denotes output.\(^5\) Uncertainty over the
effectiveness of policy is captured by assuming that \( \phi \) is stochastic. Output is also affected by
an additive random variable \( \mu \). Assume that the two random variables are independent with

\(^5\)Note that I have imposed a timing assumption that the effect of the interest rate is not contemporaneous.
This corresponds to a commonly-used identification assumption in the VAR literature and is inconsequential for
the results of the model.
finite second moments. The importance of the policy uncertainty is evident in the variance of output:

\[ Var(Y_{t+1}) = \sigma_\phi^2 \cdot (m_t)^2 + \sigma_\mu^2. \]  

(1.2)

The crucial property of this model is that the variance is increasing in the magnitude of the policy instrument.

Given the simple structure of the economy in equation (1.1), the Fed chooses the policy instrument to minimize the squared deviation of output from a target next period. That is, the Fed solves the following problem:

\[ \max_{m_t} -E_t [(Y_{t+1} - Y^*)^2] \]

subject to equation (1.1). In the context of this model, Brainard(1967) demonstrates that the policymaker will not choose a level of policy that is sufficient to reach the output target in expected terms. This result follows directly from the first-order condition for the Fed. Given that expected output is \( E[Y_{t+1}] = \bar{\phi} \cdot m_t \) and the variance of output is given by equation (1.2), the optimal policy is

\[ m_t = \frac{\bar{\phi} \cdot Y^*}{\bar{\phi}^2 + \sigma_\phi^2}. \]  

(1.3)

Substituting this policy into the expression for output (1.1) and taking expectations yields the following expression for the expected output deviation next period:

\[ E[Y_{t+1}] = \frac{1}{1 + V_t} Y^*, \]

where \( V_t \) is the coefficient of variation of the policy effectiveness parameter \( \phi \).

Equation (1.3) indicates that when policy uncertainty is high, captured by a large coefficient of variation, optimal policy tends to be more inactive. As there is more uncertainty, the Fed will implement a level of its policy instrument that is closer to the variance-minimizing level, thereby not reaching the target level of output on average. The reason is that a strong policy reaction would induce a substantial amount of variance in the economy. However, it is not clear from Brainard's analysis what action results in the least uncertainty for the economy. Diamond(1985) makes this point in an argument that the variance-minimizing level of policy
does not necessarily correspond to a constant money growth rule, but instead should depend on current and lagged variables in the economy. In Brainard’s model the policy uncertainty is exogenous, with the variance-minimizing policy given by \( m = 0 \). This paper instead concentrates on endogenously determining this variance-minimizing policy through a process of rational learning. The amount of uncertainty that the Fed faces will depend on the policy that has been recently implemented. The dynamic behavior of uncertainty that results causes the interest rate to be adjusted gradually in response to many shocks.

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.\(^6\) There is a large number of investment projects described by a distribution with unknown shape over per-period returns. When the Fed chooses the interest rate, all projects with a higher return than the interest rate invest. Because there is no role for strategic reaction by investors to the Fed’s policy, the reaction of the economy is invariant to the policy rule chosen.

Caplin and Leahy(1996) argue that the gradual nature of monetary policy has implications for the degree of responsiveness of the economy. In their model individuals make an intertemporal decision of when to invest, receiving a higher reward of investing when interest rates are lower. Knowing that the Fed is acting in a gradual manner, investors have an incentive to wait for additional decreases in the interest rate. The result is that the Fed will have to decrease rates by a larger amount to stimulate investment than it would if it did not act gradually. Note that the effect described in Caplin and Leahy is not an explanation of gradualism, as the desire to move gradually is assumed in the cost structure of their model, but an explanation of the consequences of gradualism. While this is an important and interesting issue, the model presented here simplifies the reaction of the economy for the purpose of providing a tractable explanation of gradualism.

Ignorance about the shape of the distribution of projects captures the uncertainty over the effectiveness of policy. The Fed rationally learns about the shape of the distribution by observing the cumulative distribution at the given interest rate, which it can infer from the amount of investment. This type of signal implies that uncertainty will be minimized around recent interest

\(^6\)This paper does not discuss the optimal choice of the policy instrument. In the simple model presented, all available instruments are equivalent.
rates if there is no aggregate shock to the mean of the distribution. The cumulative distribution of investment projects around last period's interest rate is known, but as the interest rate moves away from this, the uncertainty can increase.

To form a tractable model of learning about the shape of a distribution, I assume that there are two potential distributions of projects. One of these represents the actual response of the economy to monetary policy, and the uncertainty is generated from not knowing which is the true distribution. In particular, at any point in time there are two potential uniform distributions of projects with p.d.f.'s:

\[ f_L(i) = \phi_L \text{ on the interval } i \in \left[ i^L - \frac{1}{\phi_L}, i^L \right] \]
\[ f_H(i) = \phi_H \text{ on the interval } i \in \left[ i^H - \frac{1}{\phi_H}, i^H \right] \]

where \( \phi_H > \phi_L \). I refer to these as the low-respondiveness (L) and high-respondiveness (H) states of the economy, respectively. These distributions are shown in Figure 1-3.

In this framework, the response of investment to the interest rate is simply described by two potential linear IS curves with different slopes. If the Fed has observed investment at the current interest rate but does not know the true state of the economy, the potential IS curves must intersect at the current interest rate. This situation is depicted in Figure 1-4. As the Fed changes the interest rate, the response of investment will provide a signal about the true state of the economy. The reason an interest rate change will not necessarily be fully-revealing is
that the potential IS curves may shift. The support for these distributions is stochastic, with the upper endpoint evolving according to

\[
\begin{align*}
\bar{\tau}_{t+1}^L &= \bar{i}_t^L + \epsilon_{t+1}^L \\
\bar{\tau}_{t+1}^H &= \bar{i}_t^H + \epsilon_{t+1}^H.
\end{align*}
\]

For example, an increase in the interest rate may reveal that the true state is $L$ if investment does not fall by much. However, this is also consistent with the state $H$ for a particular shift in that curve, as shown in Figure 1-5.
Suppose for now that the shocks to the location of the project distributions are uniformly-distributed, an assumption that is made for tractability. The intuition behind the results arrived at in this section generalizes to other distributions of shocks to the support of the project distributions, although the calculations are specific to this case. This simplification eliminates some interesting aspects of gradualism. These aspects are discussed and a numerical solution to a more general model is presented in section 1.4.

The distributions of these shifts are chosen so that the variance of output in the two states is equal for any given interest rate policy. That is, define the change in output induced from the shock to the support of the projects distribution as $\eta^L = \phi_L \epsilon^L_{t+1}$ and $\eta^H = \phi_H \epsilon^H_{t+1}$. These shocks are distributed as

$$\eta^L, \eta^H \sim U(-\eta^*, \eta^*)$$

This implies that if the Fed leaves the interest rate unchanged, the probability of a change in output in the two states is equal. The Fed will therefore never learn the state of the economy unless there is a change in the interest rate.\(^7\)

The Fed has a belief about which distribution is the true distribution at any point in time: $P_t = \Pr(state_t = L)$. In the model presented in this section, the problem begins with a particular level of uncertainty, which will be resolved through time as the Fed reacts to shocks. A model is presented in section 1.4 in which the uncertainty rebuilds through time, so that once the Fed stops reacting, the problem regenerates around the new interest rate.

Output in this model is given by

$$Y_{t+1} = (\bar{i}_{t+1} - i_t)\phi_n + A_{t+1} \quad (1.4)$$

for state $n = L, H$. The first component of demand, which I will call "investment" demand and denote $I_{t+1}$, directly responds to interest rate changes. The second component, which I will call "autonomous" demand, is the portion of demand that is unaffected by interest rates. Autonomous demand is assumed to follow a stochastic process that is independent of interest rate policy and the shocks to the location of the investment distributions.

\(^7\)The assumption of equal variance in the two states is appealing but is not required for the effects described in this section.
The timing assumed is that the interest rate is set after observing contemporaneous output, and this rate affects output tomorrow, with uncertainty arising from the shift in the project distribution and the realization of autonomous demand. A simplifying assumption is that the Fed can distinguish the two components of output, so given the interest rate it set in the previous period, it can infer the size of the shift required from the two distributions to generate the output level — it can infer $\eta_t^L$ and $\eta_t^H$ from observing $Y_t$. With this structure the magnitude of policy uncertainty is captured in the Fed’s belief $P_t$, with less uncertainty as $P_t$ moves towards 0 and 1.

The Fed will rationally learn about the true state of the economy using Bayes’ Rule. In the case of uniform support shocks the Fed either learns the true state or does not learn anything. In particular, for a given belief $P_t$, consider how the Fed learns in response to an increase in the interest rate. The Fed learns that the economy is in the low-responsiveness state only if the decrease in output is smaller than the smallest possible response in state $H$. In other words, the Fed learns that the economy has low responsiveness only if the support shock for state $L$ is sufficiently large: $\eta^L > \eta^* - \Delta i_t(\phi_H - \phi_L)$. Likewise, if the true state is $H$, the Fed learns completely only if $\eta^H < -\eta^* + \Delta i_t(\phi_H - \phi_L)$. The unconditional probabilities of learning are given by:

$$
\Pr(P_{t+1} = 0 \mid P_t, \Delta i_t) = \frac{1}{2\eta^*} (1 - P_t) \Delta i_{t-1} (\phi_H - \phi_L)
$$

$$
\Pr(P_{t+1} = P_t \mid P_t, \Delta i_t) = 1 - \frac{1}{2\eta^*} \Delta i_{t-1} (\phi_H - \phi_L)
$$

$$
\Pr(P_{t+1} = 1 \mid P_t, \Delta i_t) = \frac{1}{2\eta^*} P_t \Delta i_{t-1} (\phi_H - \phi_L).
$$

The control problem facing the Fed is to minimize the discounted sum of squared deviations from a target:

$$
\max_{\{u_t\}} \mathbb{E}_t \left[ \sum_{j=1}^{\infty} \beta^j (Y_{t+j} - Y^*)^2 \right] 
$$

(1.5)

There are three considerations in determining the optimal interest rate. First, the Fed wishes to respond strongly to a shock so that on average it will be close to the target level of output. Second, the Fed is reluctant to react strongly since this action would induce a large amount of variance in output as a result of its ignorance about the effect of its policy. Third, the Fed has
an incentive to react strongly because larger interest rate changes allow the Fed to learn more quickly. This last effect is often referred to as experimentation.

To develop the intuition of these effects, I first concentrate on a model in which the Fed does not consider the speed with which it will learn when choosing the optimal policy,\(^8\) deferring the case in which the Fed engages in experimentation to section 1.3. It is important to understand that 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 problem is the following:

\[ i_t = \arg \max -[(\bar{Y}_{t+1} - Y^\ast)^2 + \text{Var}(Y_{t+1})], \]

in which case an explicit solution for the Fed’s problem can be found.

The optimal policy is analyzed in response to various types of shocks. While the types of shocks to be considered is limited by the absence of prices, the model makes a distinction between two very different types of output shocks. Output shocks arise from shocks to nonresponsive demand \(A_t\) and shocks to the location of investment projects \(i_t^n\). Both types of shocks affect output for any given interest rate but differ in terms of the uncertainty that results. The response of the Fed to these shocks is compared in the next two subsections.

1.2.1 The Response to Autonomous Shocks

To begin, assume that autonomous demand follows a random walk:

\[ A_{t+1} = A_t + \mu_{t+1}. \]

This provides a useful case for characterizing the optimal policy since a shock is expected to be permanent, so that variation of the interest rate through time will not be attributed to the behavior of autonomous demand. The solution to the Fed’s optimization problem involves the

---

\(^8\)In particular, the Fed assumes that the uncertainty will be completely resolved before making its policy choice next period. An alternative case when the Fed expects the same degree of uncertainty next period regardless of its policy yields similar results.
following moments for output:

\[ E[Y_{t+1}] = P_{t+1}(\bar{\gamma}_t^L - \gamma_t)\phi_L + (1 - P_{t+1})(\bar{\gamma}_t^H - \gamma_t)\phi_H + A_t \]

\[ \text{Var}[Y_{t+1}] = P_{t+1}(1 - P_{t+1}) \left[ (\bar{\gamma}_t^L - \gamma_t)\phi_L - (\bar{\gamma}_t^H - \gamma_t)\phi_H \right]^2 + \text{Var}(\eta) + \text{Var}(\mu) \]

The role of policy uncertainty is evident in the expression for the variance. 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 \( \text{Var}(\eta) + \text{Var}(\mu), \) the portion 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 can be obtained:

\[ i_t = \frac{P_{t+1}\phi_L^2 \left( \bar{\gamma}_t^L - \frac{Y^* - A_t}{\phi_L} \right) + (1 - P_{t+1})\phi_H^2 \left( \bar{\gamma}_t^H - \frac{Y^* - A_t}{\phi_H} \right)}{P_{t+1}\phi_L^2 + (1 - P_{t+1})\phi_H^2} \]  \hspace{1cm} (1.6)

To understand this equation, define \( \bar{\gamma}_t^L = \bar{\gamma}_t^L - \frac{Y^* - A_t}{\phi_L} \) as the interest rate the Fed would set to reach the output target if it knew that the responsiveness of the economy was low. Define the analogous interest rate for the high state \( \bar{\gamma}_t^H. \) Equation (1.6) simply 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:

\[ i_t = \omega_{t+1} \cdot \bar{\gamma}_t^L + (1 - \omega_{t+1}) \cdot \bar{\gamma}_t^H \]  \hspace{1cm} (1.7)

where the weighting is given by \( \omega_{t+1} = \frac{P_{t+1}\phi_L^2}{P_{t+1}\phi_L^2 + (1 - P_{t+1})\phi_H^2}. \)

It will prove useful to compare the optimal policy to an alternative policy in which the Fed ignores the variance of output, denoted by \( i^c: \)

\[ i^c_t = \text{arg max} - (\bar{Y}_{t+1} - Y^*)^2. \]

The optimization problem in this case yields a similar weighted average rule as in (1.7) but with the weight \( \omega^c_{t+1} = \frac{P_{t+1}\phi_L}{P_{t+1}\phi_L + (1 - P_{t+1})\phi_H}. \) The entire effect of the uncertainty is therefore captured
in a change to the weights that the Fed assigns to these two interest rates. In particular, the uncertainty decreases the weight on the term that would require a larger interest rate movement ($\omega < \omega^c$).

The optimal monetary policy in this model involves a dampened response of the interest rate to a permanent shock in autonomous output. Consider an economy with an amount of investment $I^*$ such that output is equal to its target. Figure 1-6 depicts the response of this economy to a positive shock to autonomous demand of magnitude $\tau$. The response of investment in the two states of the world is shown, as well as an "expected response" schedule given prior beliefs $P_{t+1}$. The policy rule that ignores output variance, $i^c$, is the point along the line segment between $\hat{I}_t^L$ and $\hat{I}_t^H$ corresponding to the weights $\omega_{t+1}^c$. This policy is precisely the interest rate where the expected response schedule crosses $I^* - \tau$, therefore inducing the amount of investment required to reach the output target in expected terms. I refer to this policy as the fully-offsetting policy. The optimal interest rate policy (1.7) instead puts additional weight on the point $\hat{I}_t^H$, resulting in interest rate choice $i^g$. The dampening of the response of the Fed to a deviation from target is evident.

The model results in gradual interest rate movements because the level of uncertainty regarding particular interest rates is resolved as the Fed changes the interest rate. Once the Fed has reacted to the shock, it observes the amount of investment at the new interest rate. It therefore faces less uncertainty about higher interest rates, even if the true state of the economy is not revealed. As a result, the Fed is now willing to increase rates further. Since the initial reaction of the Fed is dampened, output will remain above its target level on average, making additional interest rate increases desirable. The intuition is shown graphically in Figure 4b. After reacting in period $t$ with the policy $i^g_t$, investment is expected to reach only $I^E_{t+1}$. 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. The Fed solves the same problem at time $t+1$, choosing $i^g_{t+1}$. The funds rate continues to take steps towards $i^c$, the fully-offsetting policy.

This intuition for gradualism can be found in this model by rewriting equation (1.6). Using the fact that $\hat{I}_t^L = i_{t-1} + \frac{Y_t - A_t}{\phi_L}$ and $\hat{I}_t^H = i_{t-1} + \frac{Y_t - A_t}{\phi_H}$, the optimal interest rate policy can be
Figure 1-6: Response to an Autonomous Demand Shock
expressed in differences:

$$\Delta i_t = \left( \frac{\omega_{t+1}}{\phi L} + \frac{1 - \omega_{t+1}}{\phi H} \right) (Y_t - Y^*)$$

This equation gives the slope of the line in Figure 1-6 from point \((I_t, i_{t-1})\) to \((I^* - \tau, i^g)\). 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 \(\left( \frac{\omega_{t+1}}{\phi L} + \frac{1 - \omega_{t+1}}{\phi H} \right)\). It is easily shown that the fully-offsetting policy responds more strongly to a deviation of output from its target. The reactions are equal only if \(P = 0\) or \(P = 1\), and the difference in the reaction is monotonic in \((\phi_H - \phi_L)^2\). Equation (1.8) demonstrates that the funds rate will continue to rise as long as \(Y > Y^*\).

Since the interest rate response is dampened, consecutive interest rate changes in the same direction are likely to occur. In fact, in response to a permanent autonomous demand shock, the unconditional probability of a continuation is strictly greater than the unconditional probability of a reversal as long as there is some degree of uncertainty. If all uncertainty has been resolved, continuations and reversals are equally likely. To calculate the probability of a continuation, first note that equation (1.8) implies that conditional on not learning the true state, there will be a continuation if and only if output next period remains above its target level. Combining equations (1.4) and (1.7), along with the assumed process for the dynamics of the support for the project distributions, yields an expression for output in state \(n = H, L\):

$$Y_{t+1}^n - Y^* = \eta_{t+1}^n + (Y_t - Y^*) \left[ 1 - \phi_n \left( \frac{\omega_{t+1}}{\phi L} + \frac{1 - \omega_{t+1}}{\phi H} \right) \right].$$

It can be shown that the (unconditional) probability of a continuation equals the probability that \(Y_{t+1}^n - Y^* > 0\) when the following two conditions are met:

$$\eta^* > (Y_t - Y^*) (\phi_H - \phi_L) \left( \frac{1 - \omega_{t+1}}{\phi_H} \right)$$
$$\eta^* > (Y_t - Y^*) (\phi_H - \phi_L) \left( \frac{\omega_{t+1}}{\phi_L} \right).$$

Under these conditions, the probability of a continuation becomes

$$\Pr(\Delta i_{t+1} > 0) = \sum_{n=L,H} P_n \Pr(\eta^n > (Y_t - Y^*) \left[ 1 - \phi_n \left( \frac{\omega_{t+1}}{\phi L} + \frac{1 - \omega_{t+1}}{\phi H} \right) \right] | n).$$

27
A few algebraic steps yield

\[
\Pr(\Delta t > 0) = \frac{1}{2} + \frac{(Y_t - Y^*)}{2\eta^*} \left[ \frac{P_{t+1}(1 - \omega_{t+1})}{\phi_H} - \frac{(1 - P_{t+1})\omega_{t+1}}{\phi_L} \right] (\phi_H - \phi_L). \tag{1.9}
\]

which holds for any weights given a policy rule in the form of (1.8). Since \(\Delta t > 0\) if and only if \(Y_t > Y^*\), the expression gives the probability of a continuation.

Substituting the optimal weights into (1.9) results in the expression

\[
\Pr(\Delta t > 0) = \frac{1}{2} + \frac{(Y_t - Y^*)}{2\eta^*} \left[ \frac{P_{t+1}(1 - P_{t+1})(\phi_H - \phi_L)^2}{P_{t+1}\phi_L^2 + (1 - P_{t+1})\phi_H^2} \right]. \tag{1.10}
\]

The probability of a continuation equals one half only if one of the following conditions is met: i) \(P_{t+1} = 0\), ii) \(P_{t+1} = 1\), or iii) \(\phi_H = \phi_L\). Each of these conditions implies the absence of uncertainty. In the presence of uncertainty, the probability of a continuation is strictly greater than one half. The model therefore offers an explanation of the high probability of continuations as described by Rudebusch(1995). Equation (1.10) states that even in the case where there is no serial correlation in the fundamentals, the probability of a continuation will be high because of the resolution of uncertainty. Equation (1.9) can 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 \(\omega^*\) that are obtained when the Fed ignores the variance of output.

The analysis can be carried further by describing the pattern of funds rate adjustments in response to a permanent shock. The model presented here has obviously made some simplifications that affect the behavior of the funds rate adjustment. Since all uncertainty about the true state is resolved at once, there will be a point when the funds rate jumps up or down according to the true degree of responsiveness. However, the path of the funds rate before that time can be calculated. Conditional on the event that the true state of the economy is not revealed, changes in the funds rate follow an AR(1) process in response to a deviation of output from its target. In particular, it can be shown that

\[
E_t [\Delta t_{t+1}] = \left\{ (\phi_H - \phi_L) \left( \frac{P_{t+1}(1 - \omega_{t+1})}{\phi_H} - \frac{(1 - P_{t+1})\omega_{t+1}}{\phi_L} \right) \right\} \Delta t_t.
\]

Since the probability of each state and the weights are constant, the bracketed coefficient
is constant, and changes in the funds rate are described by an AR(1) process. This result demonstrates that the uncertainty is responsible for inducing serial correlation in interest rate changes. Note that if \( P = 1 \) or \( P = 0 \) then \( E_t[\Delta_i_{t+1}] = 0 \). The funds rate will react as a random walk to shocks that are permanent if the Fed knows the effectiveness of its policy with certainty.

This simple model can therefore capture the gradual movements in the federal funds rate that are observed, even when the shocks facing the Fed are permanent. The gradualism here is attributed solely to the uncertainty over the responsiveness of the economy to monetary policy and the evolution of this uncertainty through time. The minimum-variance interest rate reacts as the Fed learns about the effectiveness of its policy. In response to the shocks considered to this point, the minimum-variance interest rate is always equal to the interest rate in the previous period since the signal obtained about the distribution from that interest rate is perfect. In reality, it is plausible that the point of minimum uncertainty moves towards the funds rate, but has some dependence on recent funds rates as well since the impact of monetary policy is not immediate. The Fed therefore responds to a permanent shock by implementing a sequence of interest rate changes, reaching the fully-offsetting interest rate level only after a period of time during which it learns about the effect of its policy.

The solution to the model under more general processes for autonomous demand is similar. Equation (1.6) holds with predictable autonomous output \( \hat{A}_{t+1} \) replacing current autonomous output. Interest rate changes can now be described by

\[
\Delta_i_t = \left( \frac{\omega_{t+1}}{\phi_L} + \frac{1-\omega_{t+1}}{\phi_H} \right) (Y_t - Y^* + \hat{A}_{t+1} - A_t). 
\]

(1.11)

Changes in the funds rate will now be affected by the serial properties of autonomous demand, but there will be additional positive serial correlation as a result of the resolution of uncertainty as shown above. This is a useful observation because it is doubtful that the resolution of uncertainty could explain the use of 25 steps to implement a policy action as observed in several policy episodes (see Figure 1-1). Much of the serial correlation in funds rate changes is generated by movements in the fundamentals. Equation (1.11) allows for this but claims that there is an additional effect that cannot be accounted for by the behavior of the fundamentals.
This is consistent with the fact that the funds rate exhibits gradual behavior beyond what can be accounted for by the fundamentals as reported in Figure 1.2.

1.2.2 The Response to Investment Demand Shocks

The effect of shocks to investment demand is not as straightforward. The Fed’s reaction depends on the exact form of the information that the Fed has about the shock. In particular, there is an important distinction between a shock that increases the return of all projects by a particular amount and a shock that increases the amount of investment at any given interest rate by a particular amount.

Consider a shock that increases the return on every project by \( \lambda \) basis points. If autonomous output is zero, expected output next period conditional on state \( n \) is given by

\[
Y_{t+1} = (i_t^m - i_t)\phi_n + \lambda\phi_n
\]

which results in an optimal interest rate of

\[
i_t = \lambda + \omega_{t+1} \left( \frac{\gamma_L}{i_t} - \frac{Y^*}{\phi_L} \right) + (1 - \omega_{t+1}) \left( \frac{\gamma_H}{i_t} - \frac{Y^*}{\phi_L} \right).
\]

The shock has a one-to-one effect on the optimal interest rate — there is no gradualism in the response to this shock. In this case the minimum-variance interest rate and the fully-offsetting interest rate coincide. The optimal policy of the Fed is then perfectly clear — the interest rate should be raised by the magnitude of the shock.

Suppose instead that the Fed observes an unexpected increase in investment today. The distinction from the previous case is that the Fed no longer knows the magnitude of the shock in terms of basis point increase per project. The observed shock could have been either a large shift in project returns in the low response state, or a small shift in the high response state. Specifically, the Fed observes an increase of \( \lambda_t \) which is equal to either \( \lambda_t\phi_L \) or \( \lambda_t\phi_H \). If the shock process is a random walk, then the amount of investment that must be offset is also a random walk. The difference from the previous case is that the minimum-variance interest rate and the fully-offsetting interest rate diverge. If the Fed leaves interest rates unchanged, it knows the amount of output next period. For any change in the interest rate, including that
required to offset the shock, the Fed begins to induce uncertainty into output. The optimal policy is given by

\[ i_t = \omega_{t+1} \left( i_t^L - \frac{(Y^* - \lambda^*_t)}{\phi_L} \right) + (1 - \omega_{t+1}) \left( i_t^H - \frac{(Y^* - \lambda^*_t)}{\phi_H} \right), \]

(1.13)

which is completely analogous to equation (1.6). Therefore, in this case the degree of gradualism is equal to that in response to an autonomous demand shock.

This last observation, however, results from the simplifying assumption in the model that the Fed immediately learns about the location of the project distributions. In reality, the learning will not be immediate, so that even when the Fed observes the unexpected increase in investment, the minimum-variance interest rate will change. A shock that affects interest-rate sensitive activity will therefore warrant a swifter reaction by the Fed, with the precise degree of gradualism determined by the nature of the shock and the extent of knowledge that the Fed has regarding the shock.

In conclusion, it has been shown that shocks to autonomous demand should elicit the most gradual response from the Fed. Since the minimum-variance interest rate does not respond to these shocks,\(^9\) the Fed chooses to draw out its response over a period of time during which it learns about the effectiveness of its policy. The result is an interest rate movement that is implemented through a number of small steps. On the contrary, investment demand shocks are met with a more rapid interest rate response. Depending on the nature of the investment shock, the reaction could be immediate and complete, with no additional interest rate changes expected. While shocks in the economy do not fit perfectly into the autonomous and investment categories defined here, some shocks are certainly closer in nature to one of the two categories. For example, offsetting the effects of a government expenditure shock would correspond closely to the analysis of an autonomous shock. The change in activity would be unrelated to interest-rate sensitive output, and the reaction of the Fed would be quite gradual. An example of a shock corresponding to an investment demand shock is a change in the investment tax credit. Suppose the tax code were changed to make all investment projects more profitable. In this case,

---

\(^9\)This is not exactly true since there would be general equilibrium effects that in fact would shift the return of investment projects. In any case, one would expect a smaller shift in the minimum-variance interest rate than found in an investment shock.
the variance-minimizing monetary no longer corresponds to leaving the funds rate unchanged, but will instead involve an increase in the funds rate to offset the increase in investment. The reaction to this shock would therefore be far less gradual.

In the model presented, the parameter space characterizing policy effectiveness is quite limited, with only two potential values. One could envision a more general model in which output is determined by an IS curve $Y_t = \alpha - \beta \cdot r_t$ with a distribution of potential values for the intercept $\alpha$ and the policy effectiveness parameter $\beta$. Problems of learning about two unknown parameters in the context of monetary policy have been explored by Weiland (1996a, 1996b). However, the key element of the learning problem presented in this paper is the interaction of the uncertainty about $\alpha$ and $\beta$ that arises from observing $Y_t$. Because there are only two potential values of $\beta$, observing $Y_t$ yields only two potential values of $\alpha$. Given these values, a mistake about the slope of the IS curve is compensated for by a mistake about the intercept coefficient. As a result, uncertainty is minimized around recent levels of the interest rate, where the mistakes most effectively offset one another. Balvers and Cosimano (1994) discuss a similar "hedging benefit" in the context of the initiation of a disinflationary policy, arguing that it may be optimal to disinflate in a gradual manner, since in the short run this takes advantage of the fact that mistakes in the expected value of $\beta$ are negatively correlated with mistakes in the expected value of $\alpha$.

1.3 Gradualism in the Presence of Experimentation

In the model of the previous section, 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. The optimal policy rule 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 learning passively. Because larger interest rate changes provide a better signal about the true state of the economy, the incentive to learn leads to a more aggressive interest rate policy. This point has been made in a similar context by Bertocchi and Spagat (1993), who note that the change in the policy instrument in the optimal policy will always be as large as that in the myopic case.
In this section I explore the implications of active rather than passive learning for the results regarding gradual monetary policy. The simple form of the uncertainty assumed in the model 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. Furthermore, once the true state is learned, the Fed can exercise very effective control over the economy at any interest rate, since the distribution of projects is uniform. One could imagine instead project distributions with nontrivial shapes, where an increase in the interest rate change would at best reveal information about the local distribution around the new interest rate. While the model may overstate the benefits of experimentation, the simplicity of the model allows some analytical results to be obtained, and so I maintain the assumptions of the model, understanding that the potential for gradualism may be limited.

There is no closed-form solution for the case with experimentation. However, some interesting properties of the Fed's policy can be inferred from the optimization problem. The Fed solves the problem described in equation (1.5), only now accounting for the effect of its interest rate choice on the probability of learning the state of the economy. For simplicity, assume that autonomous demand follows a deterministic process. The solution that follows describes the response of the Fed to a positive output deviation, which could arise from a permanent shock to autonomous output.

The solution of this problem is characterized by the following Bellman equation:

\[
V(Y_t - Y^*) = \max_{\Delta t} \left\{ \begin{array}{l}
- (E_t[Y_{t+1}] - Y^*)^2 - \text{Var}(Y_{t+1}) \\
\quad - \Delta t_i (\phi_H - \phi_L) \frac{\beta}{1-\beta} \eta^* \\
\quad + P \int_{-\Delta t_i (\phi_H - \phi_L)}^{\eta^*} \frac{\beta}{2\eta} V(Y_{t+1}(\eta) - Y^*) d\eta \\
\quad + (1 - P) \int_{-\Delta t_i (\phi_H - \phi_L)}^{\eta^*} \frac{\beta}{2\eta} V(Y_{t+1}(\eta) - Y^*) d\eta 
\end{array} \right. 
\]

(1.14)

where \( Y_{t+1}^n(\eta) \) is the output in state \( n \) given realization \( \eta \). The first line on the right-hand side is familiar from the model in section 1.2. The second line gives the value to the Fed of learning the true state (regardless of the state), and the last two lines give the value in the two states associated with not learning, all weighted by the appropriate probabilities. For the purpose of discussion, I decompose the right-hand side into two parts. I will refer to the effect in the first line as the cost of an interest rate change, to capture the idea that additional interest rate
changes induce more variance in output. I will refer to the rest of the expression as the benefit of an interest rate change, since this term captures the value of additional information that the Fed learns.

These two components are graphed as a function of the interest rate change in Figure 1-7. The cost function, denoted by \( C(\Delta i_t) \), is convex in the interest rate change. This function is minimized at the interest rate policy computed in the section 1.2.1, which I denote \( \Delta i^g \). The diagram also indicates the location of the fully-offsetting policy, denoted \( \Delta i^c \), which (as shown in section 1.2.1) involves a larger interest rate change than \( \Delta i^g \). The magnitude of the gradual policy and the fully-offsetting policy will be determined by the output deviation.

The benefit function, denoted by \( B(\Delta i_t) \), captures the informational component of the problem that was previously ignored. Since the probability of learning the true state is linear in \( \Delta i_t \), there is a large enough interest rate change for which the Fed always learns the true state of the economy, given by \( \Delta i^f = \frac{2\eta^*}{\phi_U - \phi_L} \). The benefit is constant at interest rate changes greater than this fully-revealing level. Unlike \( \Delta i^g \) and \( \Delta i^c \), the fully-revealing policy \( \Delta i^f \) is independent of the deviation of output from target. It is instead determined by the magnitude of the shifts in the project distributions. 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. The optimal policy under experimentation, which will be denoted \( \Delta i^c \), is found at the point where the slopes of the two schedules are equal.

Much of the analysis that follows involves characterizing the shape of the benefit function. In particular, the benefit function is concave and has a continuous first derivative over a range of interest rate changes that includes the values \( \Delta i^g \), \( \Delta i^c \), and \( \Delta i^f \). To see this, note that the problem characterized by the Bellman equation (1.14) can be rewritten with expected output as the choice variable:

\[
V(E[Y_t], \eta_t) = \max_{E[Y_{t+1}]} \{ \Pi(E[Y_{t+1}]) + E_\eta[V(E[Y_{t+1}], \eta_{t+1})] \}
\]

where \( \eta \) is the realization of the stochastic terms and \( \Pi(E[Y_{t+1}]) \) is derived from the per-period cost function. Because the potential shift in the project distributions is bounded, the analysis can be limited to interest rate changes between the smallest change that induces zero investment and the largest change that induces all projects to be undertaken. Note that this range contains the non-experimenting policy \( \Delta i^g \), the fully-offsetting policy \( \Delta i^c \), and the fully revealing policy \( \Delta i^f \) by definition. This set of possible interest rate choices is a closed interval of the real line, which is a nonempty, compact, and convex Borel set. The form of the cost function implies that \( \Pi(E[Y_{t+1}]) \) is bounded, continuous, and concave in this region. These conditions imply that the value function will be concave in this region (see Theorem 9.3 of Stokey, Lucas, and Prescott). Given the concavity of the value function, it can easily be shown that the benefit function is concave in the change in interest rates over the relevant region. In addition, since the function \( \Pi(E[Y_{t+1}]) \) has a continuous first derivative in this region, the benefit function will also have a continuous derivative with respect to the interest rate change (see Theorem 9.10 of Stokey, Lucas, and Prescott). Simple differentiation demonstrates that the benefit function is non-decreasing over this range.

Because the benefit function has a slope of zero to the right of \( \Delta i^f \), the derivative at the fully revealing interest rate is zero: \( B'(\Delta i^f) = 0 \). This yields some immediate results. Consider the case in which the interest rate policy under passive learning does not involve learning the true state with probability one, or \( \Delta i^g < \Delta i^f \). In this case the optimal policy with experimentation will involve an interest rate change strictly greater than the interest rate change under passive
learning but strictly less than the fully-revealing policy: $\Delta i^c$ will be contained in the interval $(\Delta i^g, \Delta i^f)$. The policy involves a larger movement than $\Delta i^g$, the point made by Bertocchi and Spagat, since the cost of a marginal increase is second-order while the benefit from the additional information is first-order. Similarly, the policy will not involve learning the true state of the economy with certainty, since this is very costly in terms of the variance in output it induces while the informational benefit is second-order.

In the other case, when the shock to autonomous output is large relative to the shocks to the distributions, or $\Delta i^g > \Delta i^f$, the gradual policy solution from section 1.2.1 is the optimal solution with experimentation. This is a less-interesting case in which the gradual policy entails learning the true state of the economy with probability one, so that experimentation offers no additional incentive to change interest rates. While it is true that the interest rate response is dampened from the uncertainty, the probability of a continuation is simply equal to the probability of the low-responsiveness state.\(^\text{10}\) From this point on, it is assumed that $\Delta i^g < \Delta i^f$.

The fact that the optimal policy does not involve full learning with certainty does not ensure that the Fed continues to act gradually under active learning. In particular, we are interested in conditions under which $\Delta i^c < \Delta i^c$, which occurs when the marginal cost at $\Delta i^c$ exceeds the marginal benefit. The marginal cost at the fully-offsetting policy can be directly calculated:

$$
MC(\Delta i) \equiv \frac{dC(\Delta i)}{d\Delta i} = -2(Y_{t+1} - Y^*) (\phi_L + (1 - P)\phi_H) + 2P(1 - P)(\phi_H - \phi_L)^2 (\Delta i),
$$

so that the marginal cost at the fully-offsetting policy is

$$
MC(\Delta i^c) = \frac{2P(1 - P)(\phi_H - \phi_L)^2}{(P\phi_L + (1 - P)\phi_H)} (Y_t - Y^*).
$$

As previously argued, this cost is a function of the output deviation and is independent of $\eta^*$.

The marginal benefit cannot be directly calculated without knowing the value function. However, an upper limit on the marginal benefit can be found using the fact that $B(\Delta i)$ is increasing and concave over the range $[0, \Delta i^f]$. This limit is derived using Figure 1-8. First,

---

\(^\text{10}\)In a more general model where the effectiveness takes on more than two values, the dampened response would make the probability of a continuation higher than it would be without the uncertainty.
Figure 1-8: Deriving the Sufficient Condition for Gradualism

A lower bound $B$ can be determined for the benefit function evaluated at zero interest rate change, $B(0)$. Consider a policy that sets the change in the interest rate to zero today and for every period in the future. This policy is sub-optimal to a policy of setting $\Delta i_t = 0$ today and the optimal interest rate change for all periods $s > t$. Therefore, the value of $B(0)$ is bounded from below by the value of the sub-optimal policy $B$, which can be calculated directly as

$$B = -\frac{\beta}{1 - \beta} (Y_t - Y^*)^2 - \frac{\beta}{1 - \beta} \left(\eta^*\right)^2 .$$

The value of the benefit at $\Delta i^f$ is the expected value function given that the true state of the economy is known, which can be calculated as

$$\bar{B} = -\frac{\beta}{1 - \beta} \left(\eta^*\right)^2 .$$

The point $C$ is an upper bound for the value of the benefit function at $\Delta i^c$ since the benefit function is non-decreasing and must equal $\bar{B}$ at $\Delta i^f$. The bounds given by $B$ at $\Delta i = 0$ and the point $C$, along with the fact that the benefit function is concave, provide an upper bound.
on the marginal benefit at $\Delta i^c$ given by the slope of the line $LL$. This bound is given by

$$\overline{MB} (\Delta i^c) = \frac{\beta}{1 - \beta} (P\phi_L + (1 - P)\phi_H) (Y_L - Y^*) \cdot$$  \hspace{1cm} (1.16)

A sufficient condition for gradualism in the presence of active learning is $\overline{MB} (\Delta i^c) < MC (\Delta i^c)$. Using equations (1.15) and (1.16), this condition reduces to the following expression:

$$\frac{(P\phi_L + (1 - P)\phi_H)^2}{P(1 - P)(\phi_H - \phi_L)^2} < \frac{2(1 - \beta)}{\beta} \cdot$$  \hspace{1cm} (1.17)

Equation (1.17) has a very intuitive interpretation. The left-hand side of this condition is the ratio of the squared expected value of policy effectiveness to the variance of policy effectiveness. The condition is more likely to hold when the coefficient of variation of policy effectiveness is high, in particular higher than some threshold that depends only on the rate of discounting. If there is complete discounting, when $\beta = 0$, the solution to this problem is equivalent to the case with passive learning. Equation (1.17) indicates that there is gradualism in this case, which is known from the previous analysis since $\Delta i^e = \Delta i^g$ and $\Delta i^g < \Delta i^c$. At the other extreme, when there is no discounting of the future, equation (1.17) will not hold. Equation (1.17) therefore provides a limit on the amount of discounting that is sufficient for gradualism.

Condition (1.17) is a sufficient condition for gradualism, but it is likely to be a much stronger condition than required. One dimension in which this condition is too restrictive is that it is independent of $\eta^*$, which determines the ability of the Fed to learn about its policy effectiveness. As $\eta^*$ increases, the signal that the Fed obtains from a given $\Delta i$ becomes more noisy, and therefore the probability that the Fed learns the value of policy effectiveness declines. This will likely decrease the marginal benefit associated with policy $\Delta i^c$. In Figure 1-8, $\Delta i^f$ moves out to the right, while $\underline{B}$ and $\overline{B}$ move down by the same magnitude. This implies that $\overline{MB} (\Delta i^c)$ will tend to fall, while $MC (\Delta i^c)$ is unaffected by the change in $\eta^*$. Although condition (1.17) is sufficient for any value of $\eta^*$, it does not take advantage of the effect of $\eta^*$. Therefore, it is likely that gradualism can occur under much weaker conditions than equation (1.17) when $\eta^*$ is large. An interesting observation is that there will also be gradualism when $\eta^*$ is very small relative to the output deviation. It can be shown that the following condition is also sufficient
for gradualism:

$$\eta^* \leq \frac{\phi_H - \phi_L}{P\phi_L + (1 - P)\phi_H} \frac{Y_t - Y^*}{2}.$$ (1.18)

This condition implies that $\Delta i^f \leq \Delta i^c$, which trivially implies the existence of gradual behavior.

In summary, the desire to learn about the effectiveness of monetary policy results in larger changes in the funds rate than in the case without experimentation. However, it is not necessarily the case that this effect outweighs the uncertainty effect that resulted in gradualism, and some conditions can be derived to ensure that this will not be the case. Gradualism appears to be more likely in two circumstances. One occurs when the ability of the Fed to learn is so high that it does not lose much information from acting gradually. The other occurs when it is so difficult to learn that the benefit from more aggressive policy is small. For cases in-between, the presence of gradualism will depend on the strength of opposing forces as determined by the parameter values assumed.

1.4 The Dynamics of Uncertainty

This section presents a numerical solution to a version of the model that incorporates a more interesting and realistic learning process for the Federal Reserve. The model removes the trivial form of the learning process in which the true state was either fully revealed or fully concealed. Instead, the degree of knowledge that the Fed has regarding its policy effectiveness will vary through time, which has implications for the gradual behavior of policy. In addition, policy effectiveness is stochastic in this model. As in the previous model, the Fed will learn about the true state as it adjusts the funds rate in response to a shock. However, while before the Fed began with a particular degree of uncertainty that was resolved through time, with stochastic policy effectiveness the uncertainty will rebuild as the Fed completes the interest rate response. Finally, this section characterizes the optimal solution that involves active learning as discussed in section 1.3, which is useful for understanding the relative importance of the uncertainty and experimentation effects.

The setup of the problem is identical to that in sections 1.2 and 1.3 with the following
differences. The shocks to the supports of the project distributions are normally distributed:

\[ \eta^L, \eta^H \sim N(0, \sigma^2). \]

Whereas in the previous model Bayes Theorem implied trivial updating of the probability of the low state, the posterior distribution is now given by:

\[
\tilde{P}_t = \frac{P_t \cdot \phi\left(\Delta i^L_t / \sigma\right)}{P_t \cdot \phi\left(\Delta i^L_t / \sigma\right) + (1 - P_t) \cdot \phi\left(\Delta i^H_t / \sigma\right)},
\]

where \( P_t \) is the prior belief about the \( L \) state, \( \Delta i^n_t \) is the shift in the project distribution required to yield \( \Delta Y_t \) given the state \( n \), and \( \phi(.) \) is the p.d.f. of the standard normal distribution. The true state of the economy is given by a Markov process with transition matrix

\[
\begin{bmatrix}
q_L & 1 - q_H \\
1 - q_L & q_H
\end{bmatrix}
\]

where \( q_L, q_H < 1 \) are the probabilities of remaining in the respective states. The posterior distribution of the true state after observing output at time \( t \) therefore yields the following prior distribution for the next period:

\[
P_{t+1} = q_L \cdot \tilde{P}_t + (1 - q_H) \left(1 - \tilde{P}_t\right).
\]

With these assumptions, the optimal policy rule is numerically calculated. In the exercises that follow, the parameters of the model are set to the following values: \( q_L = q_H = .9, \phi_L = 2, \phi_H = 5, \sigma = 1.5 \), and \( \beta = .9 \).

This economy has two state variables, the output deviation and the current knowledge about the true state. The value function associated with this problem is displayed in panel (A) of Figure 1-9.\(^{11}\) For a given value of \( P \), the value function is maximized at zero deviation and declines symmetrically from this value. The decline is more pronounced for intermediate values of \( P \), or when ignorance about policy is high. Alternatively, for a given level of the deviation, the value function is maximized at the extreme values of \( P = .1 \) and \( P = .9 \), since the Fed can

\(^{11}\)To emphasize the effect of the uncertainty on the shape of the value and policy functions of Figure 1-9, the per-period objective function is rewritten as \( E[(Y - Y^*)^2] = (\bar{Y} - Y^*)^2 + \lambda \cdot \text{Var}(Y) \) with \( \lambda > 1 \) rather than \( \lambda = 1 \). This is only to assist in the interpretation of the diagrams. In the simulations that follow, the standard objective function with \( \lambda = 1 \) is used.
respond to the deviation more effectively when it is more certain about the effect of its policy. This pattern across values of $P$ is more pronounced the larger the deviation, indicating that ignorance is less costly when policy does not have to be used strongly. Panel (B) of the figure presents the value function when the Fed does not give consideration to the variance induced by the policy uncertainty, an alternative policy that is useful for the exposition below.\textsuperscript{12}

The optimal policy rule will determine the change in the funds rate for a given output deviation from target and a given belief about the true state of the world. This policy function is shown in panel (C) of Figure 1-9. To assist the interpretation of this function, panel (D) of the figure displays the policy that fully offsets the deviation in expected terms, previously denoted $\Delta i^e$. The fully-offsetting policy function involves larger interest rate changes as the Fed increases its belief that its policy is not very effective (as the value of $P$ increases). Instead, the optimal interest rate strategy actually involves smaller interest rate changes as $P$ increases from .1 to .5. While the decrease in expected effectiveness warrants a more aggressive policy, the degree of uncertainty is also increasing, so that the Fed is becoming reluctant to use its policy aggressively. As $P$ increases above .5, policy unambiguously becomes more aggressive. In fact, the two policies are very similar when the Fed is confident about the true state of the economy, both at $P = .1$ and $P = .9$ (and in fact they would be identical if $P$ could reach 0 and 1). It is when the Fed faces the most uncertainty about the effectiveness of its policy, or $P$ is closer to 0.5, that the two policies diverge, with the optimal policy involving dampened interest rate changes.

The evolution of the policy uncertainty and the dynamic behavior of the interest rate can be understood most clearly by considering the optimal policy reaction to a shock that causes a deviation in output, beginning from the belief $P = 0.5$. The reaction to a negative output shock is shown in Figure 1-10\textsuperscript{13}, which displays the response of the interest rate conditional on the two states of the world, as well as the evolution of the belief about the true state.

The panels displaying the evolution of $P$ indicate that the learning process now involves interesting dynamics. In particular, the Fed begins with a large amount of ignorance about its policy, and as a result the initial response of the funds rate is dampened. The small magnitude

\textsuperscript{12}This policy corresponds to $\lambda=0$.

\textsuperscript{13}The results in Figures 1-10 and 1-11 report the average of 501 simulated responses in the stochastic economy.
Figure 1-9: Value Functions and Policy Functions
Figure 1-10: Conditional Responses to Small Output Shock

Figure 1-11: Conditional Responses to Large Output Shock
of the initial response limits the amount of information received from observing the reaction to the policy choice. The value of $P$ begins to move towards its true value of either zero or one, but by a limited amount. Recall that in the previous model, $P$ would either remain unchanged or jump to its true value. Following the initial interest rate response, there is some reduction in the uncertainty that it faces, so that the Fed is now willing to act more aggressively in the following period. This results in a funds rate change that reveals even more information about the true state of policy effectiveness, as $P$ moves closer to its true value, causing subsequent changes in the funds rate to be even more aggressive.

The degree of gradualism and the speed of learning are intimately related. The Fed will tend to initially react very weakly in response to a shock. As the response is implemented, the level of confidence that the Fed has about the use of its policy will increase, and the degree of gradualism will decline. This effect reinforces itself, as the more aggressive policy allows the Fed to learn quickly. This variation in the degree of gradualism was absent in the model with uniformly-distributed shocks. It is consistent with the notion that the Fed tends to “fall behind” immediately following a shock and then will “catch up” by responding more strongly over time. In this model, this is the optimal strategy given the intertemporal behavior of uncertainty.

This process involving the interaction of funds rate changes and the degree of learning causes the Fed to continue adjusting the funds rate in the direction of its initial response until output has moved close to its target. Of course, the Fed will continue to adjust the funds rate by more when it is learning that its policy is not very effective, as demonstrated in the top row of Figure 1-10. The unconditional response is shown in the top panel of Figure 1-12, demonstrating that on average the optimal reaction involves gradual movements of the funds rate. As the size of the output deviation becomes small, the magnitude of funds rate changes declines. This implies that the Fed is no longer learning very much from observing the economy, and because the state of the economy is stochastic, $P$ slowly returns to 0.5. Therefore, once the policy reaction to a shock has been completed, the uncertainty will regenerate around the new interest rate.

Figure 1-11 repeats this exercise in response to a much larger output shock. While the reaction again entails gradual adjustment of the funds rate, the reason for this result is somewhat different. The magnitude of this shock warrants a large funds rate reaction so that the Fed immediately learns a great deal about the effectiveness of its policy. However, the Fed can
never become fully confident due to the stochastic nature of policy effectiveness. Because the deviation remains large following the immediate reaction, the Fed is unwilling to fully offset the shock despite its high level of confidence, since this policy would require a substantial interest rate change. Therefore, as shown in the bottom panel of Figure 1-12, this case is also characterized by gradualism, despite the fact that the Fed learns about its policy quite quickly.

1.5 Conclusion

This paper offers an explanation of gradual interest rate movements based on the presence of uncertainty regarding the effectiveness of monetary policy. The tendency to limit fluctuations in the funds rate does not indicate that the Fed has additional objectives. The traditional objective function of macroeconomic stability provides an incentive for gradual behavior once policy uncertainty is introduced. Uncertainty over the effect of its policy is certainty a realistic feature of the problem facing the Fed and one that has been openly discussed by members of the Open Market Committee.

The fact that this form of uncertainty leads to dampened policy reactions has been known
for some time. By deriving this uncertainty endogenously, the model presented above highlights the evolution of the uncertainty over time. The dynamic behavior of the uncertainty is affected by the policy that is implemented, and it is this interaction between policy and uncertainty that results in gradualism.

There are times at which monetary policy has been very inactive. Given the stochastic nature of the reaction of the economy to policy changes, these are times at which the Fed likely perceives large amounts of uncertainty. A good example of such times may be the most recent funds rate change on March 25, 1997, which took place after the funds rate remained unchanged at 5.25% for 418 days. This paper argues that these are times when the Fed is more likely to "fall behind" in its reaction to a shock, acting more gradually as a result of the high uncertainty. While it is reluctant to act, the Fed will eventually adjust the funds rate once the expected output deviation is large enough. Although limited in size by the uncertainty, the initial funds rate movements provide some information to the Fed about the reaction to its policy. This information comes in many forms and at different rates. The uncertain reaction of long-term interest rates may be resolved quickly, while observing the reaction of production may take a substantial amount of time. But as any information is revealed, the uncertainty perceived by the Fed begins to diminish. This is the effect that results in gradualism. As the knowledge of the Fed about the reaction to its policy increases, it becomes optimal for the Fed to implement additional interest rate changes, into regions that were previously not optimal because of the higher uncertainty. The subsequent movements in the funds rate become stronger, and this effect is reinforcing, since more information is then revealed, warranting even stronger movements. This is the phase during which the Fed appears to be "catching up." Finally, as the output deviation is reduced through the reaction of the funds rate, funds rate changes diminish, reducing the flow of information to the Fed and allowing the uncertainty to rebuild.

This argument derived from the model presented captures many realistic features of the actual policy implemented by the Fed, despite the fact that the model has some over-simplified components. There are some important elements of the problem confronting policymakers excluded from the model. This reflects the fact that these issues are not relevant for the optimality of gradual policy. One component of the problem faced by the Fed is determining
both whether a shock has occurred and the type of shock that has occurred. Uncertainty over the characteristics of (additive) shocks cannot produce unconditional underreaction by the Fed. For example, consider a situation where the Fed does not know if an output deviation is the result of a temporary or permanent shock. It is true that conditional on a shock being permanent, the reaction of the Fed to the shock is too gradually. However, it would be a sub-optimal policy for the Fed to underreact unconditionally. On average, the Fed will respond correctly, which entails underreaction to permanent shocks and overreaction to transitory shocks.

While other types of uncertainty may deliver conditional gradualism, uncertainty over policy effectiveness results in unconditional gradualism. The fact that unconditional gradualism is an evident and widely-perceived characteristic of monetary policy indicates that policy uncertainty is an important element of the problem faced by policymakers.
Bibliography


48


Chapter 2

Does the Fed Act Gradually? A VAR Analysis

2.1 Introduction

Once the Fed starts raising rates, multiple rate hikes are par for the course.¹

There is a widespread belief among financial market participants as well as academics that the Federal Reserve conducts monetary policy in a gradual manner. This belief has originated in the fact the Fed has historically implemented movements in the federal funds rate using sequences of small interest rate changes. This fact is evident in Figure 2-1, which displays the target federal funds rate series for the episode May 1989 to October 1996.² As can be seen, the target series has a tendency to move in one direction over substantial periods of time. Figure 2-2 captures this behavior by reporting the number of consecutive changes in the funds rate target in a particular direction over the episodes September 1974 to September 1979 and March 1984 to the present.³ The figure demonstrates that directional changes in the federal funds rate target have on several instances required as many as 25 steps. The deliberate pace of funds

¹This comment by Larry J. Wipf, vice president of Norwest Corporation, appeared in the New York Times following the increase in the federal funds rate by the Federal Reserve on March 25, 1997.
²Chapter 1 of this thesis displays the same graph over a larger time period.
³These are the episodes for which the target series is available. The funds rate series is that which is reported in Rudebusch(1995) through 1992:9. I have extended the target series to the present by reading the minutes of the FOMC meetings since this date.
rate movements captured in these figures is often referred to as gradualism.

This description of gradualism corresponds to the simplest definition, that the federal funds rate is not a random walk. In particular, gradualism under this definition implies that changes in the funds rate predict additional funds rate changes with the same sign. Defined as such, there is substantial evidence of gradualism in a paper by Rudebusch (1995). Rudebusch focuses on understanding the variation in the predictive power of interest rates of different maturities, arguing that the restrained pace of monetary policy can account for the strong predictive power of short-term interest rates. He convincingly demonstrates that a target rate change predicts additional changes in the same direction over the next several weeks. This fact can be conveyed with some simple statistics. In particular, 88.4% of the target changes in the sample are “continuations,” or have the same sign as the previous change. The probability of a continuation does not appear to depend on the direction of the interest rate movement. The proportion of continuations is 88.6% following a funds rate target increase and 88.2% following a target decrease.

In addition, the length of time between continuations is much shorter than that between reversals.\(^4\) The average length of time between target changes when the next change is a con-

\(^4\)A reversal is defined analogously to a continuation. It is a funds rate target change with the opposite sign of the last change.
tinuation is 26 days, compared to 63 days when the next change is a reversal. This is consistent with the interpretation that continuations often constitute steps within a directional movement in the funds rate target, while reversals may instead describe the reaction to a new shock.\(^5\)

There is also a momentum effect, in that the second step also takes a longer time to occur, averaging 41 days since the first step, after which the steps occur quite rapidly. Goodhart (1996) presents similar evidence suggesting that the central banks of other industrialized countries act gradually (under this definition) as well.

While it is clearly the case that the funds rate is not a random walk, it is less clear that this is a satisfactory definition of gradualism. The fact that the fundamental variables to which the Fed reacts are serially correlated could explain why the funds rate does not behave as a random walk.\(^6\) In addition, lagged interest rates are expected to enter the reaction function of the Fed to the extent that the economy responds to monetary policy with delay.

This paper provides evidence of gradualism under a more stringent definition by demonstrating that funds rate changes exhibit a higher level of positive serial correlation that can be

---

\(^5\) Rudebusch conveys this fact by estimating non-parametric hazard rates showing that in the first several weeks following a target change, a continuation is much more likely than a reversal, while once four weeks have elapsed since the last target change, the target behaves much more like a random walk.

\(^6\) This of course depends on the model of monetary policy that is being employed. In many models, the monetary instrument will have the same stochastic properties of at least one of the fundamental variables.
accounted for by the dynamic behavior of fundamental variables. Using the structural form of the economy estimated in a VAR, I calculate the funds rate policy that is expected if the Fed is optimizing a traditional objective function involving the deviation of output growth, unemployment, and inflation from their respective targets, assuming that all of the uncertainty in the economy is additive. This policy is then compared to the actual behavior of the funds rate. The analysis finds that movements in the funds rate are too gradual to be explained by the dynamic behavior of fundamental variables, taking into consideration the delayed effect of monetary policy. The funds rate response to most shocks involves a dampened initial reaction relative to the expected response, followed by eight to twelve months of interest rate changes toward the expected level.

One potential explanation for the gradual behavior is the presence of uncertainty about the effectiveness of monetary policy, as argued by Sack (1997).\(^7\) When the effect of policy on the economy is uncertain, aggressive responses of the funds rate induce a large amount of expected variance in the fundamentals. Because of this, the Fed will optimally choose to respond to shocks gradually over a period of time during which it can learn about the effect of its policy. The analysis is extended to incorporate this uncertainty in section 2.3. In particular, uncertainty over policy effectiveness is determined by the precision of the VAR estimates, since this precision yields a confidence interval for the reaction of the economy to monetary policy. The expected funds rate that takes into consideration this imprecision exhibits gradual movements that more closely resemble movements in the actual funds rate.

2.2 Evidence of Gradualism

2.2.1 Discussion

The fact that changes in the funds rate predict future funds rate changes in the same direction is not necessarily indicative of gradual behavior by the Federal Reserve. As discussed in the introduction, this behavior could arise simply from the serial correlation of the fundamental variables to which the Fed is reacting. Furthermore, lagged interest rates are expected to enter the reaction function of the Fed to the extent that the economy responds to monetary policy.

\(^7\)See Chapter 1 of this thesis.
with delay.

The analysis in this section provides evidence of gradualism that is robust to these considerations. Estimates from a VAR describe the dynamic reaction of non-policy variables in the economy to monetary policy as well as the policy rule employed by the Fed. The approach of the following analysis is to use the dynamic structure estimated for the non-policy variables to solve for the interest rate policy that would be expected if the Fed were concerned with the traditional goal of macroeconomic stability. The funds rate policy that results from this optimization problem is a function of all current and lagged variables in the economy, so that the expected funds rate policy will depend on lagged funds rate movements to an extent determined by the intertemporal structure of the economy. The behavior of this expected policy is then compared to the actual policy behavior estimated in the VAR to investigate whether actual funds rate movements are excessively gradual.\(^8\) As will be seen, the analysis indicates that given the dynamic structure of the economy, the Fed should be more aggressive in its interest rate policy.

I refer to the outcome of the optimization problem as the expected policy rather than the optimal policy because the calculation of this policy is subject to the Lucas(1976) critique. The expected policy is derived assuming that the structure of the economy is invariant to the policy chosen by the Fed. While the Lucas critique applies to all optimal policy analysis, the criticism may be more poignant in the context of a VAR. The VAR is a reduced form in which the parameters reflect structural relationships as well as the formation of expectations. Because expectations in particular will be very sensitive to policy rules, the parameters of the VAR will certainly depend on the policy that is implemented.\(^9\)

To better address the Lucas critique, there have been several recent attempts at optimal policy analysis using structural models of the economy, including McCallum and Nelson(1997), Fuhrer(1997), Fuhrer and Moore(1995), and Rotemberg and Woodford(1997). Sims(1996) offers a discussion on the relative merit of the VAR-based approach as opposed to a rational

\(^8\)One could address this question with a more structural approach to describe optimal policy, such as a forward-looking Taylor equation. To my knowledge there have been no attempts to provide evidence of gradualism using a Taylor equation. However, the strong significance of lagged interest rates found in these equations is suggestive of gradualism. See, for example, Clarida and Gertler(1995) for an application to the Bundesbank.

\(^9\)The deviation of the optimal policy from the actual policy is not extreme, so the significance of the criticism is not clear. This type of analysis has been conducted for much larger deviations in policy, for example in a paper by Furman and Leahy(1996) on the transmission of U.S. monetary policy in Canada.
expectations model, arguing that the VAR approach may be more appropriate since it matches the data more accurately. Regardless of this debate, it is reassuring that the structural approach has in general found that optimal policy rules are characterized by volatile interest rate behavior, which is consistent with the results of this paper.

In light of the Lucas critique, the VAR estimates must be interpreted as describing the responsiveness of the economy to monetary policy, given the expectation that the Fed is following the observed policy rule. While it cannot be claimed that the aggressive policy is optimal, it may be claimed that conditional on current expectations about the policy rule, the Fed will not want to maintain the observed policy. In other words, without some form of commitment, the observed reaction of the economy and the observed policy rule do not constitute a Nash equilibrium – the Fed would want to deviate to a more aggressive policy rule. The exercise in this paper ultimately argues that the expected policy may be very similar to the observed policy, in particular once uncertainty is incorporated in section 2.3. A discussion of whether this policy can then be interpreted as optimal is deferred to a later section.

2.2.2 The VAR

Suppose the economy is described by the following linear structural model:

\[
\begin{align*}
Z_t &= \sum_{i=0}^{T} A_i Z_{t-i} + \sum_{i=0}^{T} b_i i_{t-i} + v_t^Z \\
i_t &= \sum_{i=0}^{T} c_i' Z_{t-i} + \sum_{i=1}^{T} d_i i_{t-i} + u_t^i
\end{align*}
\]

where \(Z_t\) is an \(n \times 1\) vector of non-policy variables, \(i_t\) is the policy variable, and boldface denotes matrices or vectors. This is a system of simultaneous equations in which each variable is allowed to depend on the current and lagged values of all other variables in the system. The \(v_t^Z\) and \(u_t^i\) terms are the uncorrelated structural disturbances to the system.

The non-policy variables in the VAR include a commodity price index, industrial production, the unemployment rate, and the CPI (excluding shelter), while the policy variable is the federal funds rate. All variables are entered in log differences except the interest rate and the unemployment rate, which are entered in levels. While this procedure may disregard infor-
formation contained in the cointegrating relationships of the variables, performing the VAR in differences allows for a tractable solution to the optimization problem that follows.

The reduced form of this system can be estimated and the structural coefficients recovered with a sufficient number of identification assumptions. Identification is achieved through a Choleski factorization of the covariance of residuals. In particular, given the following order of the non-policy variables, \( Z_t = \{ c_t, y_t, u_t, \pi_t \} \), where \( c_t \) denotes commodity price inflation, \( y_t \) production growth, \( u_t \) the unemployment rate, and \( \pi_t \) inflation, the identification assumption is that \( A_0 \) is lower-triangular and \( b_0 = 0 \). This imposes a timing assumption that the Fed can respond to contemporaneous variables in the economy when setting the funds rate, but the funds rate does not have a contemporaneous impact on the economy, which is reasonable for the U.S. given that the data is monthly.

One might ask if this simple, non-structural approach is appropriate for the issue being addressed. Bernanke and Blinder(1992) demonstrate that the funds rate provides an effective measure of monetary policy. However, many recent VAR studies have proposed alternative measures based on structural models of the market for reserves. Bernanke and Mihov(1995) offer a review of these alternatives and demonstrate that the Fed’s operation in the market for reserves has varied over time. In particular, monetary policy is accurately characterized by movements in the federal funds rate over most of the sample. However, between 1979:9 and 1982:10, Fed policy appears to be better described by non-borrowed reserves targeting. The VAR is estimated over the period 1972:11 to 1995:3, so that the use of the funds rate as a policy measure is not unreasonable.\(^{10}\) However, it is of some concern that the sample includes three years during which this may be a poor measure of policy. To account for this concern as well as concerns about structural stability of the rest of the model, the VAR is also estimated over the subperiod 1982:10 to 1995:3. A likelihood ratio test for the difference of the parameter estimates between this subperiod and the full-sample suggests differences that are significant at the .05 but not the .01 level. Because of this moderate difference, the robustness of the results is verified by repeating the analysis for this subperiod.

The structure of the VAR for the full sample is restricted to include lags 1 to 6, 9,

\(^{10}\)The starting date of the sample is chosen to coincide with the beginning of funds rate targeting by the Fed, as discussed in Strongin(1995).
and 12, which achieves a balance between allowing sufficient dynamics and avoiding over-parameterization. A likelihood ratio test using the modification suggested in Sims(1980) does not reject this lag structure at the .01 significance level, although it marginally rejects it at the .05 level. I therefore maintain this structure but repeat the analysis with lags 1 to 12, finding similar results that are not reported. A lag structure of 1 to 6 is not rejected in the subsample and is therefore maintained since the decrease in the sample size causes over-parameterization to become more of a concern.\footnote{The data used is seasonally-adjusted, so that the absence of a twelfth lag is not a problem.}

2.2.3 The Analysis

The VAR describes both the structural form of the non-policy variables in the economy and the reaction function of the Federal Reserve. The exercise that follows is to calculate the expected funds rate policy taking as given the structural form of the economy estimated from the VAR (excluding the reaction function of the Fed). A traditional objective function is assumed for the Fed:

$$-\frac{1}{2} \sum_{i=1}^{\infty} \beta^i \left\{ (\pi_{t+i} - \pi^*)^2 + \lambda_u (u_{t+i} - u^*)^2 + \lambda_y (y_{t+i} - y^*)^2 \right\}$$

where the weights $\lambda_u$ and $\lambda_y$ determine the relative importance of the deviation of output growth, unemployment, and inflation from their respective targets. The analysis is to solve for the policy rule that maximizes this objective subject to the dynamics implied by the VAR estimates.

To solve this problem, define a state vector as

$$X_t = \{ c_t, c_{t-1}, \ldots, c_{t-12}, y_t, y_{t-1}, \ldots, y_{t-12}, u_t, u_{t-1}, \ldots, u_{t-12}, \pi_t, \pi_{t-1}, \ldots, \pi_{t-12}, i_t, i_{t-1}, \ldots, i_{t-12} \}.$$ 

The expected policy will be a solution to the following Bellman equation:

$$V(X_t) = \max_{\pi_t} \left\{ -(X_t - X^*)'G(X_t - X^*) + \beta E_t [V(X_{t+1})] \right\}.$$
subject to

\[ X_{t+1} = F \cdot X_t + H \cdot i_t + J + \mu_{t+1} \]

where coefficient matrices \( F \) and \( H \) and the constant vector \( J \) are composed of parameters from the VAR, and the only uncertainty enters through an additive stochastic vector \( \mu_{t+1} \). \( G \) is a matrix of zeros except for the diagonal elements corresponding to contemporaneous output growth, unemployment, and inflation, which contain the relative weights. Since the per-period payout is quadratic, the value function will have the form:

\[ V(X) = X'\Lambda X + 2X'\omega + \rho. \]

It can be shown that the solution to this problem is given by

\[ i_t^* = -(H'\Lambda H)^{-1}(H'\Lambda F \cdot X_t + H'\Lambda J + H'\omega) \tag{2.1} \]

where \( \Lambda \) must satisfy the Riccati equation

\[ \Lambda = -G + \beta F'\Lambda F - \beta F'\Lambda H(H'\Lambda H)^{-1}H'\Lambda F \tag{2.2} \]

and \( \omega \) satisfies

\[ \omega = (I - \beta F'(I - \Lambda H(H'\Lambda H)^{-1}H'))^{-1} \cdot \]

\[ (GX^* + \beta F'\Lambda(I - H(H'\Lambda H)^{-1}H'\Lambda)J). \tag{2.3} \]

In deriving this policy, it is assumed that the structure of the economy is known with certainty by the Fed. In particular, the Fed does not give any consideration to the precision of the estimates in the VAR, instead taking the point estimates to truly describe the dynamic structure of the economy. In section 2.3, the variance of the VAR estimates is considered in formulating the expected policy in order to gauge whether the uncertainty found in the VAR is sufficient to account for the degree of gradualism observed.

The behavior of the calculated expected policy will depend on the six parameters in the problem: \( \beta, \lambda_u, \lambda_Y, y^*, \pi^* \), and \( u^* \). The solution offered in equations (2.1) through (2.3)
holds when these parameters are constant.\textsuperscript{12} In the results that follow, I impose the following parameter values: $\beta = .996$, $y^* = 2.5$, $\pi^* = 3.0$, and $u^* = 6.0$, where the output growth and inflation rates are annualized. The target values are chosen a priori but are not far from the sample averages, which are 2.46, 5.71, and 6.87, respectively. The results are not significantly affected by other (reasonable) choices of the discount factor.

The weights in the objective function $\lambda_u$ and $\lambda_y$ will be estimated from the actual behavior of the federal funds rate. For any pair of weights, the expected policy rule can be calculated. The expected funds rate is computed by substituting the vector of state variables into the derived rule period-by-period. The estimates of the weights that minimize the sum of squared deviations between the expected policy and the actual funds rate are $\lambda_u = .85$ and $\lambda_y = .30$. By using these weights, the analysis that follows allows for the most successful description of monetary policy within this class of quadratic objective functions.

The calculated policy is successful at characterizing the behavior of the actual funds rate measured at low frequencies. However, the month-to-month dynamics of the two series are quite different. The expected policy explains only 31\% of the monthly movements in the funds rate. Furthermore, a likelihood ratio test rejects the coefficients implied by the expected policy at the .000 significance level.

The expected rate is more volatile than the actual rate, with a standard deviation in changes of 1.42 compared to .80. Moreover, the funds rate tends to move gradually towards the expected rate. The dynamics of the two can be described by the following equation:

$$\Delta i_t = \alpha_1 (i_t^* - i_{t-1}) + \alpha_2 \Delta i_{t-1},$$  \hspace{1cm} (2.4)

where $i_t^*$ is the expected policy. Equation (2.4) is a partial adjustment model amended with a second term to capture inertia in funds rate changes. Estimates from this specification are presented in Table 2-1, indicating that both of these terms are very significant. The coefficient on the partial adjustment term indicates that absent previous funds rate changes, only 9\% of the deviation from target is offset in a given period. The inertia effect demonstrates that changes in the funds rate are strongly dependent on lagged funds rate changes, even when controlling

\textsuperscript{12}In particular, the solution above assumes a value function with constant parameters, which requires that the targets do not vary through time.
for the deviation from an expected level.

<table>
<thead>
<tr>
<th>Table 2-1: Estimated Funds Rate Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Equation: $\Delta i_t = \alpha_0 + \alpha_1 \Delta i_{t-1} + \alpha_2 [i_t^* - i_{t-1}]$</td>
</tr>
<tr>
<td>Sample: 1973:11 to 1995:3</td>
</tr>
<tr>
<td>$R^2 = .225$</td>
</tr>
<tr>
<td>Durbin Watson: 1.871</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta i_{t-1}$</td>
<td>0.373</td>
<td>0.055</td>
<td>.000</td>
</tr>
<tr>
<td>$[i_t^* - i_{t-1}]$</td>
<td>0.088</td>
<td>0.018</td>
<td>.000</td>
</tr>
</tbody>
</table>

The dynamics that are produced from the estimates of equation (2.4) are shown in Figure 2-3. To be certain that these dynamics are not simply a result of the structure imposed in the equation, the figure also presents the unrestricted dynamics from the funds rate estimated on eight lags of itself and the expected rate. The similarity of the results indicate that the specification in equation (2.4) is reasonable. The figure can be interpreted as the response to an unconditional shock to the fundamentals that moves the expected funds rate by one standard deviation.13 The response of the expected policy is immediate, reaching a permanent new level very quickly. On the contrary, the actual funds rate responds very weakly to a change in the expected rate. The funds rate moves gradually, reaching the level that fundamental variables necessitate only after a substantial period of time.14

Figures 2-4 and 2-5 displays the impulse response functions of the actual funds rate and the expected funds rate in response to a shock to each of the five variables in the VAR. In each case the expected policy involves a strong, immediate reaction of the funds rate, after which the interest rate moves quickly back to its previous level. Without uncertainty about the effect of its policy, the Fed finds it optimal to offset the shock with a strong reaction without delay. On the other hand, the actual response of the funds rate in most cases involves a gradual, persistent

---

13 The dynamics given for the expected funds rate are from a univariate autoregressive process.
14 Some of the gradual behavior estimated in the partial-adjustment equation could be attributed to mismeasurement of the expected funds rate, which would bias the coefficient $\alpha_2$ downward. However, this equation is later repeated with much success, indicating that this is not the problem. In addition, the response functions from the VAR strongly demonstrate this pattern of underreaction as well.
movement in the funds rate. The difference between the two responses provides evidence of
gradualism. The evidence is impressive considering that this difference is found despite the
choice of weights that will tend to make the two responses as similar as possible.\textsuperscript{15}

To characterize the degree of gradualism observed in these response functions, Table 2-2
reports the amount of time for the funds rate to reach its peak response and return one quarter
of the way to its previous level. The response to an unemployment shock remains near its
peak reaction 14 months after the shock, whereas the expected policy instead returns from its
peak a full year earlier. The extent of gradualism to an output shock is very similar to that
of an unemployment shock. The response to an inflation shock has an unexpected movement
in the first two periods, reflecting the “price puzzle” that is commonly found. The funds rate
then sharply increases, but this response and the VAR response are estimated imprecisely. A
commodity price shock results in the most gradual funds rate movement, with the Fed taking
several years to implement its response. The most interesting case is that of a shock to the
federal funds rate. In the VAR, the reaction function to this shock is markedly different from
the others, involving a strong movement in the actual funds rate in response to the shock. The

\textsuperscript{15}The finding that the expected policy is too aggressive is robust to other criteria for choosing the weights in
the objective function. For example, minimizing the variance of the expected funds rate results in very different
weights ($\lambda_u = 3.0$, $\lambda_y = 0.5$) but a similar strong initial reaction of the expected rate.
Figure 2-4: Impulse Response Functions

Figure 2-5: Impulse Response Functions (cont.)
table shows that the reaction to a funds rate shock is short-lived for both the actual and the expected funds rate.

<table>
<thead>
<tr>
<th>Shock Type</th>
<th>VAR Response</th>
<th>Expected Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Output</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Inflation</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Commodity Price</td>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td>Funds Rate</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

As discussed in the introduction, the gradual response of the funds rate that is observed in response to most shocks is consistent with the expected behavior in the presence of policy uncertainty. Interestingly, under this explanation of gradualism, a shock to the federal funds rate is the strongest candidate for a rapid reaction by the Fed. Recall that according to this theory, the Fed responds to shocks to fundamental variables in the economy with gradual funds rate movements since it does not know precisely the effect of its policy. However, a shock to the federal funds rate may be different. The interest rate predicted by the VAR describes the policy chosen by the Fed taking into account the uncertainty it faces. Suppose that the current funds rate deviates from this, for example because of institutional reasons. This deviation does not substantially affect the uncertainty that the Fed faces at the predicted rate, so there is no reason that the deviation should not be offset immediately. The differences among the impulse response functions therefore offer some preliminary support for the hypothesis that uncertainty can account for the gradual behavior of the funds rate. Additional evidence will be presented in section 2.3.

A qualification of the results is that the estimated expected policy is very volatile. The figures display 90% confidence intervals for the reaction of a particular policy rule. However, the policy rule changes dramatically with variation in the coefficients of the VAR. This imprecision

---

\[16\] This evidence, while suggestive, is far from conclusive because the interpretation of the funds rate shock is not clear. Under alternative interpretations, one might expect gradualism to remain.
results from the strength of the response of the expected funds rate policy, which dramatically changes when the estimates of the structure of the economy change slightly. Many of the expected responses also contain peculiar movements in the funds rate for some period after the shock. This again is a result of the strength of the funds rate reaction. The VAR may not precisely estimate the longer lags of the effect of the funds rate, and the expected policy reacts strongly to any mismeasurement in this effect, leading to erratic movements.

Repeating this analysis on the subsample indicates that the evidence of gradualism is robust, although with qualitative differences. The estimated weights for the policy rule during this period differ from the full sample, given by $\lambda_u = .25$ and $\lambda_y = .20$. This indicates that the Fed has placed relatively more weight on the inflation target in this later period.\textsuperscript{17} In addition, the dynamic structure of the economy appears to be different, now warranting a somewhat gradual reaction. The reaction functions to various shocks in the economy are presented in Figures 2-6 and 2-7.\textsuperscript{18} Although the expected responses are now more gradual, the observed responses involve the same pattern of a dampened initial reaction followed by a more gradual, persistent movement than can be explained. Another difference is that the response to a funds rate shock appears more gradual in this subsample, although the confidence intervals are fairly large. It may be that the reaction function for the funds rate in the whole sample is strongly influenced by the volatile interest rate environment of the late 1970's and early 1980's.

Overall, the reaction functions provide strong evidence of gradualism. Interest rate reactions that take years to implement cannot be rationalized from the structure of lags in the economy, as the reaction of the expected policy demonstrates. The structure of the economy calls for a strong, short-lived reaction by the Fed to all shocks. The simple optimization problem in this section therefore does not adequately describe interest rate behavior. The failure of this model could be attributed to the form of the objective function. This paper will concentrate instead on a different deficiency, the absence of uncertainty about the effectiveness of policy. The difference in the reaction speeds across shock types in the full sample is supportive of an explanation based on policy uncertainty. The analysis that follows therefore investigates whether uncertainty can account for the gradual funds rate behavior that has been described.

\textsuperscript{17}This increase in the emphasis on inflation in the later sample is also found in Fuhrer and Moore(1995).

\textsuperscript{18}The response to an inflation shock is imprecise and nonsensicle, as the price puzzle worsens in this subsample.
Figure 2-6: Impulse Response Functions for Subsample

Figure 2-7: Impulse Response Functions for Subsample (cont.)
2.3 Incorporating Uncertainty

This section investigates whether uncertainty regarding the effects of monetary policy can account for the gradual behavior of the funds rate documented in section 2.2. In particular, the following analysis examines whether this uncertainty has the characteristics and the magnitude required to explain the degree of gradualism that is observed.

The interesting aspect of the exercise is that the VAR provides estimates of the type of uncertainty that may explain the gradual behavior. In particular, uncertainty over policy effectiveness is determined by the precision of the VAR estimates, since this precision yields a confidence interval for the reaction of the economy to monetary policy. The results from section 2.2 ignored the fact that the dynamics of the economy are estimated imprecisely, using only the point estimates of the coefficients in deriving the expected policy. This section accepts that the variance-covariance matrix of the point estimates also contains important information that is required to describe the expected policy. The expected policy calculation that follows takes into account the uncertainty generated by the imprecision in the estimates of the structural form of the economy.

The optimization problem that incorporates this uncertainty is simplified by redefining the state variable of the problem. Because the per-period payout is a function of the expected value of the variables in the economy, the state variable will be defined as \( \hat{X}_t = E_{t-1}[X_t] \). The solution to the optimization problem will then satisfy the following Bellman equation:

\[
V(\hat{X}_t) = \max_{\hat{X}_t} \left\{ -(\hat{X}_t - X^*)'G(\hat{X}_t - X^*) - Var_{t-1}(\hat{X}_t'G\hat{X}_t) + \beta \left[ V(\hat{X}_{t+1}) \right] \right\}.
\] (2.5)

The problem in section 2.2 assumes that all uncertainty in the economy is captured by an additive stochastic term affecting the dynamics of the state variables. With the quadratic objective function, additive uncertainty has no effect on the optimal policy. Equation (2.5) instead explicitly accounts for the uncertainty over the matrices governing the state variable dynamics. The difference in the two cases is captured by the second term on the right-hand side of equation (2.5), which measures the amount of uncertainty that is associated with a given expected value of the state variable.

Incorporating this term causes the expected variance to depend on the policy implemented
by the Fed. The variance term will be a function of the covariance matrix of the coefficients as estimated in the VAR. In particular, the variance term can be written as

\[ \text{Var}_{t-1}(\tilde{X}_t'G\tilde{X}_t) = \tilde{X}_t'K\tilde{X}_t. \]

The matrix K is given by \( K = \Sigma_{\beta(n)} + \lambda_u \Sigma_{\beta(u)} + \lambda_p \Sigma_{\beta(p)} \) where \( \Sigma_{\beta(n)} \) represents the covariance matrix of the coefficients in the equation describing variable n, modified to account for the timing of the problem. The interest rate choice affects the expected state variable next period, which has implications for the variance of the target variables as determined by the covariance matrix of the parameter estimates in the VAR.

One can verify that the solution to this problem is similar to that of the previous section, with the difference arising in the parameters of the value function. The expected policy will again be given by equation (2.1), and the parameter \( \omega \) remains given by equation (2.3). The only difference from the previous case is a modification in the equation determining \( \Lambda \), which is now given by

\[ \Lambda = -G - K + \beta F'\Delta F - \beta F'\Delta H (H'\Delta H)^{-1} H'\Delta F \]  

(2.6)

rather than equation (2.2).
The difference in the two problems appears modest, captured by the presence of the $K$ matrix in equation (2.6). However, the impact of this modification on the expected policy is substantial. Using the same criterion as before, the estimated objective function now involves the weights $\lambda_u = .10$ and $\lambda_y = .45$, a reversal of the relative importance of unemployment versus output growth. Furthermore, the expected funds rate is more successful in describing monetary policy, now accounting for 79% of the variation of the funds rate (compared to 31%). The reason for the improvement is that the expected reaction of the Fed no longer involves the (excessively) aggressive interest rate responses previously found. The responses of the expected policy to various shocks in the economy are displayed in Figures 2-8 and 2-9. The immediate reaction of the expected policy is very consistent with the behavior of the funds rate. Consistent with the dampening of these responses, the standard deviation of the expected policy falls from 1.42 to .81, much closer to the .80 value for the funds rate. However, this accuracy does not extend to the entire funds rate response. The actual policy still entails funds rate movements that are more persistent than can be accounted for in the expected policy, a point that is returned to below.

The overall improvement in the model can be demonstrated using the partial adjustment specification in equation (2.4). The estimates from this equation are presented in Table 2-3. The dynamic behavior resulting from these estimates implies that the funds rate would offset
80% of a deviation from the expected rate in less than six months. The estimates in the case without uncertainty (Table 2-1) indicate that it would take more than ten months for the funds rate to respond by this amount. The improvement arises from the fact that the expected rate is no longer characterized by dramatic funds rate movements, instead exhibiting dynamic behavior more consistent with the observed funds rate.

<table>
<thead>
<tr>
<th>Table 2-3: Estimated Funds Rate Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Equation: ( \Delta i_t = \alpha_0 + \alpha_1 \Delta i_{t-1} + \alpha_2 [i^*<em>t - i</em>{t-1}] )</td>
</tr>
<tr>
<td>Sample: 1973:11 to 1995:3</td>
</tr>
<tr>
<td>Uncertainty Case:</td>
</tr>
<tr>
<td>( R^2 = .227 )</td>
</tr>
<tr>
<td>Durbin Watson: 1.857</td>
</tr>
<tr>
<td>( [i^*<em>t - i</em>{t-1}] )</td>
</tr>
<tr>
<td>( \Delta i_{t-1} )</td>
</tr>
</tbody>
</table>

The reason for the dampened behavior of the expected policy is that the aggressive response in the case without policy uncertainty induces a substantial amount of variance in the targeted variables. The covariance matrix of the VAR parameter estimates captures the uncertainty associated with large movements in the funds rate. The amount of variance associated with various interest rate choices is demonstrated in Figure 2-10. The figure displays the weighted average of the variance as a function of the state variable and the current interest rate choice. The state variables are set to their mean values with the exception of the funds rate. The twelve lags of the funds rate are set equal to the rate given on one of the horizontal axes, while the current funds rate choice is varied along the other axis. A movement away from the 45-degree line of the horizontal axes therefore represents a deviation of the interest rate from recent levels. As evident in the figure, the variance is strongly associated with the deviation of the interest rate from previous levels rather than the interest rate level. This characteristic of the variance is the source of dampened funds rate movements. Attempting to reach the optimal level of the expected value of the fundamentals leads to aggressive changes in the funds rate, but these changes are dampened to avoid the high variance that would result from uncertainty about the
structure of the economy.\textsuperscript{19}

This last point can be emphasized by comparing the performance of the two monetary rules that have been derived to the performance of the observed funds rate policy. The comparison is conducted using 500 simulations of the behavior of the economy first under the assumption that the coefficient matrix is constant, so that there are additive disturbances only, and then under the assumption that the coefficients are stochastic, drawn from the distribution estimated in the VAR. In both cases, the performance of the VAR reaction function is compared to the "aggressive" policy derived in section 2.2 and the "tentative" policy derived in this section, where the performance is measured according to the objective function with the (case-specific) estimated weights. In the case without multiplicative uncertainty, the VAR policy performs 28\% worse than the aggressive policy. By offsetting deviations more rapidly, the aggressive policy results in a substantial improvement in performance. Surprisingly, despite the lack of similar aggressive movements, the tentative policy also dominates the actual policy, with the

\textsuperscript{19}This characteristic of the variance is slightly modified when the state vector is away from its mean. In that case, there is a tendency for the variance-minimizing interest rate to move in the direction that would return the state vector to its mean. For example, when unemployment has been high, the variance minimizing interest rate is found to be somewhat below the previous interest rate.
VAR policy performing 25% worse.\textsuperscript{20} This poor performance however depends critically on the form of the uncertainty in the economy. In the case with stochastic coefficients, the aggressive policy leaves the policymaker 22% worse off than the VAR policy. This reversal is of course attributed to the variance induced by moving aggressively as compared to a policy that moves very cautiously. The tentative policy in this case will accept more variance than the observed policy in order to move the fundamentals closer to the target, choosing the optimal trade-off between the two costs. This policy outperforms the VAR policy by an 8% margin.

<table>
<thead>
<tr>
<th>Shock Type</th>
<th>Periods to Return 75% of Peak Response</th>
<th>Expected Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Output</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Inflation</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Commodity Price</td>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td>Funds Rate</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Despite being successful in dampening the reaction of the funds rate, the expected policy still falls short of describing the persistence of the funds rate response. Table 2-4 summarizes the responses in this case, labeled "With Uncertainty," and presents the policy derived without uncertainty for comparison. As indicated in the table, there is a limited amount of improvement in explaining the amount of time it takes the Fed to implement its peak response.

It is interesting to note that the persistence observed may in fact be consistent with uncertainty over monetary policy, only with an objective function that places more weight on variance. In fact, objective functions that involve only the variance are not uncommon in the monetary policy literature. While it is unclear how the policymaker should emphasize the variance relative to expected movements, the form of the objective function assumes a very particular weighting. Suppose instead each term of the objective function were rewritten as

\textsuperscript{20}Note that the aggressive policy and the tentative policy cannot be directly compared because of the difference in the objective functions.
\(- (\bar{X} - X^*)^2 - \phi \cdot \text{Var}(X)\), where \(X\) represents output growth, inflation, and unemployment. The previous case assumes that \(\phi = 1\). In fact, the expected policy can describe actual funds rate policy much more accurately as \(\phi\) increases.

Figures 2-11 and 2-12 displays the response of the expected funds rate policy to various shocks for the parameter value \(\phi = 10\).\(^{21}\) The expected policy is characterized by gradual interest rate adjustments that are very similar to the observed behavior of the federal funds rate. The incentive to smooth the funds rate in order to reduce the variance implies that the funds rate will return to its previous level more slowly. The persistence of these responses is described in the column of Table 2-4 labeled “Emphasized Uncertainty.” The success of the expected policy can also be gauged using the partial adjustment model, the estimates from which are presented in Table 2-5. According to these estimates, the funds rate would offset 80% of a deviation from the expected rate within two months, a significant improvement over the six months required in the previous case and the ten months in the case without uncertainty.

<table>
<thead>
<tr>
<th>Table 2-5: Estimated Funds Rate Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Equation: (\Delta i_t = \alpha_0 + \alpha_1 \Delta i_{t-1} + \alpha_2 [i^*<em>t - i</em>{t-1}])</td>
</tr>
<tr>
<td>Sample: 1973:11 to 1995:3</td>
</tr>
<tr>
<td>Uncertainty Case: (\phi = 10)</td>
</tr>
<tr>
<td>(R^2 = .304)</td>
</tr>
<tr>
<td>Durbin Watson: 1.830</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>([i^*<em>t - i</em>{t-1}])</td>
<td>0.322</td>
<td>0.053</td>
<td>.000</td>
</tr>
<tr>
<td>(\Delta i_{t-1})</td>
<td>0.447</td>
<td>0.060</td>
<td>.000</td>
</tr>
</tbody>
</table>

The results from repeating this analysis on the subsample since 1982 are even more successful. Figure 2-13 demonstrates this success by comparing the actual funds rate to the expected funds rate derived from the optimization problem with parameter value \(\phi = 10\).\(^{22}\) The expected funds rate describes the dynamic movements of the funds rate very accurately. The tendency to under-predict the level of the funds rate may be a result of the parameter values that were imposed rather than estimated, in particular the target variables.

\(^{21}\)The values of the weights in the objective function are maintained.
\(^{22}\)In this case the weights in the objective function are estimated to be \(\lambda_u = 1.43\) and \(\lambda_v = .14\).
Figure 2-11: Impulse Response Functions with Emphasized Uncertainty

Figure 2-12: Impulse Response Functions with Emphasized Uncertainty (cont.)
Figure 2-13: Actual and Expected Funds Rate

Overall, the estimated structure of the uncertainty can account for the degree of gradualism that we observe with remarkable success. While the traditional objective function performs quite well, the behavior of the Fed is better described by an objective function that places additional weight on the variance of the variables in the economy.

2.4 Conclusion

The VAR analysis presented in this paper demonstrates that in the absence of uncertainty over policy effectiveness, the extent of gradualism observed in the dynamic behavior of the funds rate is puzzling. The policy is expected to react more aggressively in response to shocks in the economy. Understanding the behavior of the Fed requires not only the VAR estimates of the dynamics structure of the economy, but the estimates of the precision of that dynamic structure as well. The exercise demonstrates that uncertainty about the structure of the economy can account for the gradual interest rate policy implemented by the Fed. Faced with the uncertainty estimated from the VAR, the Fed will optimally choose to act gradually, so that the expected policy becomes more consistent with the observed policy. The success at describing funds rate
dynamics improves as the emphasis on the variance of the target variables is increased.

Since uncertainty over the dynamic structure of the economy can possibly account for the gradualism that we observe, it is not necessary to postulate additional motives for interest rate smoothing. Note that the optimality of a gradual policy rule has been demonstrated conditional on the expectations that the Fed will act gradually. However, the conclusion falls short of claiming that the gradual policy is in fact optimal. For example, it may still be the case be that the Fed acts too gradually, if a more aggressive policy would alter the reaction of the economy in such a way to make the Fed better off by implementing this policy. However, this seems unlikely, as the shift to an aggressive policy would have to substantially reduce the uncertainty in the response of the economy. Furthermore, even if this were the case, the transition between policy rules may be very costly. Reaching the optimal policy requires that individuals learn that the Fed is employing a new policy rule rather than reacting to disturbances under the current rule. The above analysis demonstrates that the cost of the aggressive policy is very high when individuals do not alter their expectations. Therefore, if individuals learn about the transition slowly, the cost of the transition may be sufficiently high that the Fed may choose to continue using the gradual policy rule, even if it believes that it may be dominated by an alternative policy rule.
Bibliography


Chapter 3

The Response of Term Structure
Puzzles to Monetary Policy

3.1 Introduction

Economists have been largely unsuccessful at understanding the dynamic behavior of the term structure of interest rates. The most widely applied theory of the term structure, the expectations hypothesis, claims that the expected return to holding bonds of various maturities should be equal to a constant, maturity-specific risk premium. There have been numerous empirical studies demonstrating the statistical failure of the expectations hypothesis in different forms. One approach, adopted by Campbell and Shiller (1984) among others, estimates an autoregressive univariate process for the short-term interest rate. The dynamic response of the short-term interest rate to an innovation determines the expected response of the long-term interest rate which can be compared to the actual response of the long-term rate. A similar approach that is employed by Campbell and Shiller (1987) estimates a VAR of interest rate variables to project the expected path of the short-term rate, allowing a calculation of the expected interest rate spread at any time. The correlation of this expected spread is compared to the actual spread, as well as the variance of the two spreads. Other papers, the most comprehensive of which is Campbell and Shiller (1991), have tested the forecasting ability of interest rate spreads, since the spread between long-term and short-term interest rates must predict a particular path for future changes in these variables to be consistent with the expectations hypothesis.
In each of these approaches, there is a lack of focus on the source of the movement in the short-term interest rate or spread.\textsuperscript{1} In general, this literature has documented the existence of term structure puzzles in response to unconditional movements in the short term interest rate or in interest rate spreads. While this approach has been useful in describing anomalies in interest rate behavior, investigating the response of the term structure to different types of movements in the short-term interest rate may provide additional information regarding the source of these puzzles.

This paper realizes that monetary policy is an important determinant of short-term interest rates and may therefore have implications for the behavior of the term structure of interest rates. Accordingly, interest rate movements are decomposed into different components of monetary policy. In particular, interest rate changes can arise from the predictable reaction of monetary policy to current innovations in the economy, from unpredictable policy shocks, and from movements in interest rates that are orthogonal to fundamental variables and monetary policy. The exercise that follows investigates whether the source of a movement in the short-term interest rate has implications for the behavior of the term structure, and in particular the success or failure of the expectations hypothesis.

The results indicate that the reaction of long-term interest rates varies substantially across different sources of movements in the short-term interest rate. Long-term rates severely underreact to predictable movements in the short-term rate in response to output innovations in the economy, leading to significant arbitrage opportunities. On the other hand, long-term rates tend to overreact to innovations in the inflation rate. The difference in the behavior of long-term interest rates conditional on different types of movements in the short-term rate indicates that the unconditional evidence must be interpreted with caution, since this evidence combines dramatically different behavior of the term structure. As demonstrated by the results found below, the combination of overreaction and underreaction in response to different shocks may allow the unconditional evidence to appear more consistent with the expectations hypothesis than warranted.

The conditional responses of the term structure may clarify the failure of the expectations hypothesis that has been previously documented. For example, it is not the case that the

\textsuperscript{1}Exceptions are discussed in Section 3.3.
long-term interest rate pervasively underreacts to movements in the short-term rate, which has become one popular interpretation of the unconditional evidence. In general, the conditional evidence suggests that deviations from the expectations hypothesis arise because the long-term interest rate tends to mimic the behavior of the federal funds rate too closely. This finding may be suggestive of some explanations that are discussed in the final section.

The conditional evidence also presents some additional aspects of term structure puzzles. The most significant finding is that there is a sharp difference in the success of the expectations hypothesis in describing the term structure response to innovations in fundamental variables (output and prices) versus innovations arising in financial markets. As previously mentioned, interest rate movements determined by the endogenous response of monetary policy to innovations in fundamental variables lead to significant term structure puzzles. In contrast, interest rate movements that are orthogonal to movements in fundamental variables appear consistent with the expectations hypothesis. This latter category includes innovations to the federal funds rate, often interpreted as policy shocks, and innovations in interest rates with longer maturities that are unrelated to policy shocks.

The difference in the conditional success of the expectations hypothesis can be emphasized by considering the short-run behavior of the three-month interest rate compared to the federal funds rate. Monetary policy shocks account for 94% of the innovation to the federal funds rate and nearly 40% of the innovation to the three-month interest rate in the month following the shock. However, policy shocks explain less than 1% of the (positive or negative) excess returns to holding the three-month T-bill over the federal funds rate! On the other hand, the response of monetary policy to innovations in output accounts for only 2.5% of the movement in the federal funds rate and even less of the movement in the three-month interest rate, but accounts for more than 50% of the excess returns. This sharp difference indicates the importance of distinguishing between endogenous policy movements and exogenous policy shocks as a conditioning factor.

The choice of the current decomposition is admittedly somewhat arbitrary, although it is based on the idea that the primary determinant of the short-term interest rate is the behavior of the Federal Reserve. Within this framework, the current analysis places a large emphasis on interest rate movements arising from the established reaction of policy to innovations in fundamental variables as compared to exogenous policy innovations in the interest rate. This
approach places less emphasis on structural interpretation of the innovation to the fundamentals. Alternative choices could have instead emphasized the correct decomposition of these innovations, for example into supply and demand disturbances as in Blanchard and Quah (1989) or technological and non-technology disturbances as described in Gali (1996). However, the results that follow indicate that it is the distinction between endogenous policy reactions and policy innovations that is critical to term structure behavior.

The remainder of the paper is structured as follows. Since the VAR approach adopted here is less common than a regression analysis on interest rate spreads, section 3.2 investigates the unconditional behavior of the term structure, comparing the results of a VAR to some familiar regression results. Section 3.3 explains the methodology used in this paper and analyzes the response of the term structure to different components of the movement in short-term interest rates. Section 3.4 discusses the relative importance of the various components of short-term interest rate movements for describing interest rate movements and the violation of the expectations hypothesis. Section 3.5 offers additional discussion and concluding remarks.

3.2 The Unconditional Behavior of the Term Structure

The expectations theory of the term structure of interest rates claims that the expected holding returns on bonds with different maturities are equalized up to a constant risk premium. Denote the yield to maturity on a zero-coupon bond with maturity \( n \) months at time \( t \) as \( R^n_t \). While the expectations hypothesis can be written in several potential forms, as discussed in Shiller, Campbell, and Schoenholtz (1983), the most commonly applied form requires the following relation between bond returns:

\[
R^n_t = \frac{1}{k} \sum_{i=0}^{k-1} E_t R^m_{t+i} + c^{n,m} \quad (3.1)
\]

where \( k = n/m \). The risk premium \( c^{n,m} \) may vary across different maturities but is independent of time.

As described in the introduction, the failure of the expectations hypothesis has been extensively documented. One approach for documenting this failure is based on a regression analysis of the predictive power of interest rate spreads. Campbell and Shiller (1991) offer a comprehen-
sive set of results in this direction. The following analysis replicates some of their key findings to demonstrate the robustness of the term structure puzzles to the sample considered in this paper, 1982:11 to 1991:2. (A description of the sample choice and the data used is deferred to section 3.3.) In addition, developing these results will allow a clear comparison to the VAR method employed in this paper.

According to the expectations hypothesis, the spread between the current returns on long-term and short-term bonds can exist only if one expects an increase in the long-term interest rate over the maturity of the short-term bond. Formally, equation (3.1) implies that the spread between the return on an \( n \)-period bond and the return on an \( m \)-period bond, \( s_{t}^{n,m} = R_{t}^{n} - R_{t}^{m} \), must satisfy the following relationship:

\[
\frac{m}{n - m} s_{t}^{n,m} = E_{t} R_{t+m}^{n-m} - R_{t}^{n}.
\] (3.2)

This can be tested by a regression of the change in the long rate, \( R_{t+m}^{n-m} - R_{t}^{n} \), on the (adjusted) interest rate spread at time \( t \). A rejection of the expectations hypothesis is indicated by a coefficient that is significantly different from one. While Campbell and Shiller perform this test for all possible pairs of maturities, the analysis here maintains the short-term interest rate \( (m) \) as the federal funds rate.\(^2\)

Campbell and Shiller find that a positive spread between long and short-term interest rates predicts a falling long-term interest rate, therefore indicating the failure of the expectations hypothesis.\(^3\) The results from the empirical test of equation (3.2) for the sample in this paper are presented in Table 3-1, demonstrating a similar failure. These results indicate that there are predictable excess returns to holding long-term bonds that are positively correlated with the spread between the current returns on long-term and short-term bonds.

---

\(^2\)This is because the emphasis of the analysis in the sections that follow is on monetary policy, and the funds rate is the policy instrument of the Federal Reserve.

\(^3\)In addition, the authors find that the increase in the short-term rate in response to an interest rate spread is in general too small to be consistent with the expectations hypothesis. I will instead emphasize the results for the long-term interest rate.
As an alternative to the regression analysis on interest rate spreads, this paper investigates the dynamic behavior of interest rates with different maturities using a VAR. Consider the following VAR:

\[ X_t = \sum_{i=0}^{T} C_i X_{t-i} + \mu_t, \]

where \( X_t = [FF_t, ST_t, LT_t]' \) is a vector of interest rates that includes the federal funds rate, a short-term interest rate, and a long-term interest rate. Figure 3-1 depicts the response of the short-term and long-term interest rates to an innovation to the federal funds rate. This may be interpreted as an unconditional response of the term structure to a change in the funds rate under the identification assumption that the funds rate does not react to current innovations in the other interest rates.\(^4\) Each row of the figure corresponds to a separate VAR in which the long-term interest rate is changed, with maturity increasing moving down the page. The short-term interest rate is maintained as the three-month rate in each VAR.\(^5\) All figures report one-standard error bands computed using the bootstrap method described in Blanchard and Quah(1989).

The striking feature of the response functions is the apparent success of the expectations hypothesis. The left-hand column presents the reaction of the indicated interest rate along with the reaction of the federal funds rate. The right-hand column presents the reaction of the indicated interest rate along with its expected reaction, calculated by summing over the expected funds rate response according to equation (3.1). The reaction of the federal funds rate

---

\(^4\)More extensive discussion of this interpretation is offered in section 3.3.

\(^5\)The three month rate is included in each VAR because it may be an important part of the Fed's reaction function, a point discussed in section 3.3. The first row of the figure presents the reaction of the three month rate in a VAR that also includes the two-year interest rate. Each VAR delivers a very similar response of the three-month rate.
Figure 3-1: Unconditional Response (dotted line: response of indicated interest rate; solid line: funds rate reaction (left) or expected interest rate reaction (right))
is transitory, so that the longer-term interest rate should react by less to an extent determined by its maturity. This is indeed the case. There is a slight overreaction for two-year and five-year interest rates, but the difference is short-lived.

The success of the expectations hypothesis depicted in Figure 3-1 can be reconciled with the failure of the regression analysis if there are large transitory movements in excess returns. Campbell and Shiller(1987) conduct a VAR exercise, with particular focus on the cointegrating relationship between short-term and long-term interest rates, and find a strong relationship between the actual spread and that predicted by future changes in the short-term interest rate. The correlation between the two is positive and large, although the actual interest rate spread is more volatile than the spread that is predicted by theory. They conclude that the ability of the expectations hypothesis to explain much of the observed movements in interest rates is obscured by tests such as the regression analysis above. In particular, tests of the predictability of returns based on interest rate spreads may be very sensitive to transitory deviations from the expectations hypothesis.

This paper will analyze the success of the expectations hypothesis by focusing on the response of interest rates as described in a VAR. In this sense, the unconditional behavior of the term structure is a very successful benchmark for comparison with the conditional responses developed in the following section. The implications of the conditional responses for the regression analysis on interest rate spreads is discussed in section 3.5.

3.3 The Reaction of the Term Structure to Monetary Policy

This paper is not the first to realize that monetary policy may have critical implications for the behavior of the term structure of interest rates. Other authors have connected the expectations hypothesis with various components of monetary policy, although none have investigated term structure puzzles conditional on different components of policy.

There have been several papers investigating whether the infrequent and gradual behavior of monetary policy can account for the failure of spread regressions, including Rudebusch(1995), Cook and Hahn(1990), Balduzzi, Bertola, and Foresi(1993), and Dotsey and Otrok(1995). This direction of research demonstrates that the coefficient in spread regressions will be biased ac-
cording to the predictability of the short-term interest rate. Since policy is conducted infre-
quently and gradually, there may be variation in this uncertainty over different maturities that
can explain some of the patterns of the failure that we observe in spread regressions. Other
authors have also concentrated on "microstructure" considerations of monetary policy. Roley
and Sellon(1996) and Reinhart and Simin(1996) investigate the behavior of the term structure
conditional on the occurrence of a policy action by the Fed.

There has also been some work on whether variation in the operating procedure of the Fed
can account for variation in the behavior of the term structure. Fuhrer(1996) demonstrates
that reasonable variation in the coefficients of the reaction function of the Federal Reserve can
account for the failure of the expectations hypothesis. Slight variation in the policy rule can
cause large differences in the path of the short-term rate over the life of the long-term bond
and can therefore account for large deviations between the short and long rates. His analysis
most closely resembles this paper since it explicitly considers the effect of the monetary policy
reaction function on the future behavior of interest rates and therefore also on long-term interest
rates. Also in this direction, Roberds, Runkle, and Whiteman(1996) examine the predictive
power of the term structure over different policy regimes for the Fed.

Finally, some authors have noted that the Fed may react to long-term interest rates in its
conduct of monetary policy. Goodfriend(1993) describes several episodes where a rise in the
spread between long-term and short-term rates may have led to an increase in the federal funds
rate since it was a signal of expected future inflation. McCallum(1994) demonstrates that this
endogenous reaction can lead to the appearance of the failure of the expectations hypothesis
even under full rationality.

This paper instead uses our understanding of monetary policy to distinguish between different
types of movements in the short-term interest rate. In particular, the distinction is drawn
between the endogenous response to movements in fundamental variables in the economy ac-
cording to an estimated reaction function, and exogenous policy shocks that are not explained
by the reaction function. In addition, the analysis also allows for a component of interest rate
movements that is orthogonal to both innovations in fundamentals and policy shocks.

This decomposition is implemented by the following VAR. Suppose the economy is described
by the following linear structural model:

\[ Z_t = \sum_{i=0}^{T} A_i Z_{t-i} + \nu_t, \]

where \( Z_t \) is the following vector of variables: \( Z_t = [Y_t, P_t, FF_t, ST_t, LT_t]' \) and the variance-covariance matrix of \( \nu \) is given by the diagonal matrix \( \Sigma \). \( Y_t \) denotes industrial production, \( P_t \) the CPI, \( FF_t \) the federal funds rate, \( ST_t \) a short-term interest rate, and \( LT_t \) a long-term interest rate. The contemporaneous movement of these variables is given by the matrix \( A_0 \), which has the following form:

\[ A_0 = \begin{bmatrix} A^1 & A^3 \\ A^2 & A^4 \end{bmatrix}, \]

where \( A^1 \) is a three-by-three matrix corresponding to output, prices, and the funds rate and \( A^4 \) is a two-by-two matrix describing the contemporaneous relationship between short-term and long-term interest rates.

Full identification of this VAR is a difficult task since it requires imposing restrictions on the contemporaneous relationship between interest rate variables, \( A^4 \). Since both interest rates are likely to respond to any type of disturbance to financial markets, the restrictions required to recover \( A^4 \) are unclear. As a result, the strategy employed here will be to conduct the analysis using only an unidentified VAR in which \( A^4 \) is not determined. In particular, I impose restrictions on two of the blocks of this matrix. The first is that \( A^1 \) is lower-triangular. This restriction imposes the common assumption that the federal funds rate does not have a contemporaneous impact on the economy, which is reasonable given that the data is monthly. The second restriction is that \( A^3 = 0 \), or that innovations to the short-term and long-term interest rates do not have a contemporaneous effect on the other variables in the economy. Restrictions on the effect on output and the price level are reasonable given that similar restrictions are imposed for the funds rate. The assumption that the Fed does not respond to interest rate innovations is more concerning. There could be innovations in these rates (orthogonal to innovations in the fundamental variables) to which the Fed is reacting which are being captured as part of the policy innovation. In general, there is little evidence in favor of or against the role of innovations in financial variables in the reaction function of the Fed. Most VAR studies
in fact do not include long-term interest rates as a result of the same difficulty regarding the identification of $A^4$. In doing so, these papers are implicitly making the same identification assumption for policy shocks that is made here.

Although including these additional interest rates in the VAR causes the identification to be difficult, this procedure has advantages as well. Rudebusch (1996) presents evidence that the policy innovations from standard VAR equations contain a component that is anticipated at least one period in advance by financial markets. This finding indicates that the reaction function may be improperly specified and should include lags of short-term interest rates. Retaining both a short-term and a long-term interest rate in the VAR in addition to the federal funds rate corrects this misspecification. In all of the results that follow, the short-term interest rate is given by the three-month interest rate since this rate may be the strongest candidate for inclusion in the reaction function of the Fed.

The assumptions mentioned leave the VAR unidentified, requiring one additional restriction for full identification. However, the restrictions are sufficient to identify the reaction of the economy to shocks to output, the price level, and the federal funds rate. Rewrite the VAR in its reduced form as

$$\mathbf{Z}_t = \sum_{i=1}^{T} \mathbf{B}_i \mathbf{Z}_{t-i} + \mathbf{u}_t.$$  

The reduced form errors are correlated, with the variance-covariance matrix of $\mathbf{u}$ given by $\Omega$. The relation between the reduced form innovations and the structural disturbances is

$$\mathbf{u}_t = \mathbf{B}_0 \mathbf{v}_t,$$  

where as before $\mathbf{B}_0$ can be segmented as

$$\mathbf{B}_0 = \begin{bmatrix} \mathbf{B}^1 & \mathbf{B}^3 \\ \mathbf{B}^2 & \mathbf{B}^4 \end{bmatrix}$$

where $\mathbf{B}^1$ is a three-by-three matrix and $\mathbf{B}^4$ is a two-by-two matrix. The assumption that $\mathbf{A}^1$ is

---

6One of the few papers that does allow a reaction to long-term interest rates is and Leeper (1994). However, in this paper the authors only identify this effect by assuming that long-term rates do not react to innovations in the funds rate, which seems more unacceptable than the identification proposed here.
lower-triangular identifies the matrix $B^1$. The assumption that $A^3 = 0$ allows identification of the matrix $B^2$ without identifying its components $A^2$ and $A^4$. The impulse response functions to output, price level, and funds rate shocks can be recovered once $B^1$ and $B^2$ are known.

To summarize, without making implausible identification assumptions in order to fully identify the system, the analysis can recover the response of interest rates to three types of disturbances. The first two affect the funds rate through the reaction to an innovation in fundamental variables in the economy, the level of output and the price level. The last is a movement in the funds rate that is orthogonal to any disturbances in the economy, or a policy shock. The response of the two interest rates is completely unrestricted. These interest rates will certainly respond to the innovations in the level of output, prices, and the funds rate. In addition, the interest rates also respond to contemporaneous movements in the other rate, but it is unnecessary to recover this interdependence in order to observe the reaction of the two rates to the shocks mentioned.

The analysis begins by investigating the behavior of the term structure in response to these types of shocks. To do so, the VAR described is implemented using monthly data over the period 1982:11 to 1991:2. The term structure data used is that calculated by McCulloch and is described in McCulloch(1990). The interest rates reported are the continuously-compounded yield to maturity on a zero-coupon bond, so that the duration equals the maturity of the bond. This data is constructed using portfolios of coupon bonds similar to those used to create stripped treasury securities. The short sample period is chosen to ensure the stability of the monetary policy reaction function. Bernanke and Mihov(1995) find that 1982:11 corresponds to a shift to a funds rate targeting procedure for the Fed, so that a reaction function based for the federal funds rate offers a good description of policy. To extend the sample earlier, one must deal with the difficult issue of defining a measure of monetary policy shocks that is robust across various operating procedures for the Fed. The ending date of the sample is determined by the availability of the interest rate data. A lag structure of six lags is not rejected by a likelihood ratio test and is therefore utilized.

The response of the term structure to various changes in the short-term interest rate are presented in Figures 3-2 through 3-4, which are structured in the same manner as Figure 3-1. Figure 3-2 depicts the interest rate response to a shock in output. In this case the responses of
Figure 3-2: Response to an Output Shock (dotted line: response of indicated interest rate; solid line: funds rate reaction (left) or expected interest rate reaction (right))
Figure 3-3: Response to a Price Level Shock (dotted line: response of indicated interest rate; solid line: funds rate reaction (left) or expected interest rate reaction (right))
Figure 3-4: Response to a Monetary Policy Shock (dotted line: response of indicated interest rate; solid line: funds rate reaction (left) or expected interest rate reaction (right))
interest rates with longer maturities strongly violate the expectations hypothesis. The initial funds rate response to this shock is very limited, but the response increases through time, reaching a peak response after a year at which time it begins to decline. Remarkably, even for long maturities, the longer-term interest rate mimics this behavior closely. There is a strong underreaction of the long-term rate following this shock, followed by a period of overreaction as the funds rate decreases. For example, holding the funds rate over the year following this shock yields an excess return of 2.5% over a strategy of holding a one-year bond.

Figure 3-3 depicts the response to an inflation shock. Again, the longer-term interest rate appears to follow the funds rate too closely. Since the funds rate response is declining, there is an overreaction of the longer-term interest rate during the year following the shock. However, as will be seen in the next section, inflation shocks account for very little of the movement in the funds rate.

Figure 3-4 presents the reaction of the funds rate to a monetary policy shock. The behavior of the long-term interest rate is dramatically different from that depicted in Figures 3-2 and 3-3. In the year following the policy shock, the long-term interest rate reaction is consistent with the expectations hypothesis. The response of longer-term interest rates is more dampened, consistent with the transitory movement in the funds rate. The only violation of the expectations hypothesis occurs for long-term interest rates after a year has elapsed since the shock, at which time the two rates follow each other too closely.

The figures presented indicate that the tendency of the longer-term rate to mimic the funds rate in the short run is limited to instances where the Fed is reacting to innovations to fundamental variables in the economy. This can be demonstrated statistically by considering the null hypothesis that the reaction of the two interest rates is equal at various maturities. The test investigates whether the reaction functions shown in the figures are significantly different from one another, giving consideration to the variance in the funds rate reaction (not shown in the figures) and the covariance between the two. The significance of this test based on the bootstrapping method is presented in Table 3-2 for the response of interest rates of different maturities. The results confirm that there is no significant difference in the reaction of any of the longer-term interest rates and the federal funds rate in response to output and inflation shocks during the three months following the shock. In contrast, there is a significant difference
in the reaction to the funds rate shock, with the exception of the month during which the shock occurs.

<table>
<thead>
<tr>
<th>Shock to:</th>
<th>$Y$</th>
<th>$P$</th>
<th>$FF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizon (mo.):</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6 mo. rate</td>
<td>.105</td>
<td>.345</td>
<td>.400</td>
</tr>
<tr>
<td>12 mo. rate</td>
<td>.125</td>
<td>.470</td>
<td>.490</td>
</tr>
<tr>
<td>24 mo. rate</td>
<td>.170</td>
<td>.255</td>
<td>.220</td>
</tr>
<tr>
<td>60 mo. rate</td>
<td>.175</td>
<td>.300</td>
<td>.185</td>
</tr>
</tbody>
</table>

Figures 3-2 to 3-4 are useful for observing the dynamic pattern of interest rate movements. The comparison of the actual and expected policy responses demonstrate the arbitrage opportunity available over the life of the longer-term bond. But since this horizon varies, it is difficult to compare the reaction of different interest rates following a shock. It would be useful to describe the dynamic reaction of period-by-period excess holding returns for bonds of various maturities.

Since these are discount bonds, the holding return for a bond with maturity $n$ at time $t$ is approximately given by

$$H_t^n = nR_t^n - (n - 1)R_{t+1}^{n-1}.$$

Define the excess holding return as the difference between the holding return and the funds rate at that time, or

$$ER_t^n = nR_t^n - (n - 1)R_{t+1}^{n-1} - FF_t.$$

The difficulty is that the response of $R_t^{n-1}$ is not observed in the VAR. However, the $n-1$-period rate can be replicated by a portfolio of the $n$-period bond and future funds rate. In particular, holding 1 unit of the $n-1$-period bond (with face value $1$) at $t+1$ is equivalent to the strategy of holding $(1 + FF_{t+n-1})$ units of the $n$-period bond at time $t + 1$ and going short $1$ of federal funds at time $t + n - 1$. The excess holding return of this $n$-period bond strategy is then

$$ER_t^n = nR_t^n - nR_{t+1}^n + (FF_{t+n-1} - FF_t).$$
The dynamic responses of the excess holding returns of the long-term interest rate are given in Figures 3-5 and 3-6. Figure 3-5 depicts the response to the endogenous policy reaction to a shock to output. Because the response to an output shock involves a gradual funds rate movement, in which the initial reaction is small followed by additional movement in the same direction, the long-term rate is expected to react more strongly than the funds rate. The funds rate reaction to an inflation shock is also gradual but reverts more quickly, resulting in an expected interest rate reaction that is less than the response of the funds rate. As a result, the tendency for the longer-term rate to mimic the federal funds rate discussed above corresponds to an underreaction to an output innovation and an overreaction to an inflation innovation, which is shown in the figures. Figure 3-5 demonstrates that the underreaction of longer-term rates results in excess returns of holding federal funds as opposed to longer-term bonds following an output shock. The magnitude of the underreaction is large, resulting in a 1% to 2% per month difference in annualized returns over the first several months for the two-year bond.

Figure 3-6 depicts the response to an exogenous policy shock, indicating that the excess returns to the longer-term interest rate in response to a monetary policy shock are in general insignificant. As shown above, the policy shock is not followed by the tendency for longer-term rates to mimic the funds rate that occurs in response to the other shocks. If there is any effect, there is a short-lived overreaction of interest rates with maturities of two and five years, but the excess returns are small relative to the underreaction shown in Figure 3-5.

3.4 A Decomposition of a Puzzle

The results from the previous section describe the response of the term structure puzzle to different types of innovations in interest rates. However, the analysis falls short of demonstrating the importance of the various types of innovations in describing the puzzle. It is a standard procedure in VAR analysis to decompose the variance of interest rate movements into the amount accounted for by different innovations. The exercise that follows is to conduct a similar decomposition on the portion of interest rate movements that is inconsistent with the expectations hypothesis.

Because the VAR that is being used is underidentified by one restriction, a complete variance
Figure 3-5: Excess Returns in Response to an Output Shock
Figure 3-6: Excess Returns in Response to a Monetary Policy Shock
decomposition is not possible. In particular, one cannot disentangle the variance attributed to short-term interest rate innovations and long-term interest rate innovations separately. However, the variance can be broken down into that resulting from output shocks, inflation shocks, monetary policy shocks, and shocks originating in the financial markets. The exercise requires the following decomposition of the variance of the reduced-form innovations:

\[
\Omega = b_1 b'_1 Var(z_i^p) + b_2 b'_2 Var(z_i^p) + b_3 b'_3 Var(z_i^{lf}) + b_4 b'_4 Var(z_i^{st}) + b_5 b'_5 Var(z_i^{lt}),
\]

where \( b_i \) denotes the \( i \)th column of matrix \( B_0 \). The elements of the first line are determined by the partial identification. The sum of the terms on the second line can then be determined by the difference, although the two components cannot be separately calculated.

The other element required for the variance decomposition is the moving-average from of the vector autoregression, or equivalently the impulse response function. Since the impulse response function for excess returns to the long-term bond has been calculated, the variance of the excess returns shown in Figures 3-5 and 3-6 can be decomposed into the fraction arising from the various innovations.

Tables 3-3 through 3-5 present the results of the decomposition of interest rate movements and the excess returns of interest rates over the horizons of one month, six months, and three years. This exercise was conducted for all maturities, and the tables report the results for the three-month, 12-month, and five-year interest rates. One-standard deviation confidence intervals are presented in parenthesis.
<table>
<thead>
<tr>
<th></th>
<th>% of $FF$ explained by</th>
<th>% of $i_t$ explained by</th>
<th>% of $ER^*$ explained by</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Y$</td>
<td>$P$</td>
<td>$FF$</td>
</tr>
<tr>
<td>1</td>
<td>.023</td>
<td>.035</td>
<td><strong>.942</strong></td>
</tr>
<tr>
<td></td>
<td>(.015,</td>
<td>(.023</td>
<td>(.888</td>
</tr>
<tr>
<td></td>
<td>(.066)</td>
<td>(.069)</td>
<td>(.957)</td>
</tr>
<tr>
<td></td>
<td>(.205</td>
<td>(.018</td>
<td>(.291</td>
</tr>
<tr>
<td></td>
<td>(.411)</td>
<td>(.069)</td>
<td>(.401)</td>
</tr>
<tr>
<td>36</td>
<td>.564</td>
<td>.060</td>
<td>.142</td>
</tr>
<tr>
<td></td>
<td>(.424</td>
<td>(.046</td>
<td>(.107</td>
</tr>
<tr>
<td></td>
<td>(.618)</td>
<td>(.147)</td>
<td>(.186)</td>
</tr>
</tbody>
</table>

* $ER^*$ denotes excess holding return of $i_t$ over $FF$ rate between $T$ and $T+1$

The striking feature of this decomposition occurs in the periods immediately following the innovation. Consider the results of Table 3-3. The monetary policy shock dominates interest rate movements in the short run, accounting for 94% of the innovation in the funds rate and 40% of the innovation in the three-month rate. However, the majority of these interest rate changes is consistent with the expectations hypothesis. Despite the strong influence on interest rates, the monetary policy shock explains less than one percent of the violation of the expectations hypothesis in the short run!

The puzzling behavior of interest rates in the short run can instead be attributed to movements in the funds rate in response to innovations in output and prices. The endogenous response to an innovation in output is responsible for less than 3% of the movement in the funds rate and less than 1% of the movement in the three-month rate. However, as described in section 3.3, these movements severely violate the expectations hypothesis, accounting for 51% of the excess returns in the short run. In fact, it is the underreaction of the long-term interest rate that limits the importance of this response in describing overall interest rate movements.

The short run responses of the funds rate and the three-month interest rate emphasize the importance of considering conditional term structure puzzles. The performance of the expectations hypothesis certainly depends on the source of the movement in interest rates, so
that considering only the response to unconditional movements in the interest rate may be terribly misleading.

| *Table 3-4: Decomposition: \(i_t = 12\)-Month Interest Rate* |
|---|---|---|---|---|---|---|---|---|
| \(T\) | % of \(FF\) explained by | % of \(i_t\) explained by | % of \(ER\) explained by |
| | \(Y\) | \(P\) | \(FF\) | \(i_t\) | \(Y\) | \(P\) | \(FF\) | \(i_t\) | \(Y\) | \(P\) | \(FF\) | \(i_t\) |
| 1 | .022 | .036 | .942 | .000 | .001 | .003 | .301 | .694 | .399 | .364 | .072 | .165 |
| | (.015) | (.023) | (.890) | (.000) | (.000) | (.002) | (.256) | (.627) | (.284) | (.249) | (.047) | (.122) |
| | (.064) | (.084) | (.956) | (.000) | (.016) | (.017) | (.367) | (.736) | (.554) | (.425) | (.232) | (.306) |
| 6 | .294 | .019 | .366 | .321 | .276 | .066 | .154 | .504 | .527 | .191 | .126 | .156 |
| | (.221) | (.014) | (.285) | (.260) | (.203) | (.046) | (.117) | (.426) | (.418) | (.145) | (.098) | (.154) |
| | (.400) | (.057) | (.433) | (.397) | (.365) | (.089) | (.219) | (.585) | (.571) | (.238) | (.172) | (.257) |
| 36 | .615 | .042 | .154 | .189 | .620 | .067 | .090 | .223 | .523 | .198 | .120 | .159 |
| | (.473) | (.035) | (.111) | (.148) | (.472) | (.052) | (.062) | (.184) | (.422) | (.159) | (.097) | (.159) |
| | (.674) | (.094) | (.226) | (.290) | (.637) | (.118) | (.153) | (.338) | (.523) | (.222) | (.174) | (.245) |

* \(ER\) denotes excess holding return of \(i_t\) over \(FF\) rate between \(T\) and \(T+1\)*

Results for the three-month interest rate at longer horizons indicate that the endogenous response of monetary policy to output innovations becomes more important to describing interest rate movements beyond the short run. This is consistent with the delayed reaction of the funds rate in response to this shock shown in Figure 3-2. The fraction of the puzzling behavior of the longer-term interest rate remains significant as the time horizon increases. In the long run, the endogenous response of monetary policy to output innovations accounts for 35% of the variance of interest rate movements that violate the expectations hypothesis.

Price level shocks are nearly irrelevant for explaining interest rate movements even at longer horizons. However, the small portion that these shocks do explain are inconsistent with the expectations hypothesis. This is in stark contrast to monetary policy shocks, which never explain more than 5% of term structure puzzles at any horizon, even though their influence on interest rate movements is often significant. These differences highlight the idea that the distinction that appears to be relevant in describing term structure puzzles is between the
endogenous response of monetary policy to developments in the economy and policy shocks that are unrelated to the state of the economy.

Financial market shocks appear to be quite important for describing interest rate behavior. Much of the short-term movement in the three-month interest rate, nearly 60%, is attributed to movements that are orthogonal to endogenous and exogenous changes in monetary policy. These movements appear largely consistent with the expectations hypothesis in the short run, accounting for only 15% of the short-term puzzle. In the long run, these innovations explain a large portion of funds rate movements as well as a large portion of the term structure puzzle in the long run.

<table>
<thead>
<tr>
<th>Table 3-5: Decomposition: $i_t = 5$-Year Interest Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of $FF$ explained by</td>
</tr>
<tr>
<td>Y</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>36</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* $ER$ denotes excess holding return of $i_t$ over $FF$ rate between $T$ and $T+1$

The results using interest rates with longer maturities shown in Tables 3-4 and 3-5 convey the major points discussed from Table 3-3. The tables also demonstrate that orthogonal interest rate innovations are an important component of interest rate movements, including the federal funds rate. This indicates that identifying these innovations may be critical for understanding the term structure. While we cannot identify the reaction function to these shocks, the variance decompositions indicate that these innovations may result in interest rate movements that are consistent with the expectations hypothesis, similar to the innovations in the federal funds rate.
rate. The fraction of excess returns explained by these shocks is limited compared to the magnitude of interest rate movements for which they account. Tables 3-3 through 3-5 report the fraction of the overall variance that is attributed to each source of innovations. An alternative measure that is of interest is the amount of variance in excess returns as a fraction of the overall amount of variance in interest rate movements in response to each type of shock. One potential normalization is to divide the variance of excess returns between two interest rates by the sum of the variances of the two interest rates. By normalizing this way, one may ask whether the proportion of interest rate movements that is in accordance with or violation of the expectations hypothesis is similar for different shocks, even when these shocks account for very different amounts of interest rate movements.

The results from such a test are presented in Table 3-6. The table reports the significance level under the null hypothesis that the ratio of the variance of excess returns to the sum of the variance of interest rate movements is equal in response to different types of shocks. The statistic is calculated using the bootstrap method previously described. The first column of data indicates that in the period following the shock, the fraction of interest rate movements that are inconsistent with the expectations hypothesis is not significantly different in response to an output or an inflation shock. Both of these shocks account for interest rate movements that are mostly puzzling. However, there is a significant difference from the fraction of interest rate movements that are puzzling in response to a funds rate shocks, as shown in the second column. This result is consistent with the differences in the variance decompositions reported in Tables 3-3 to 3-5 that were previously discussed.

The interesting component of Table 3-6 is that one cannot reject the similarity between funds rate shocks and innovations in other interest rates, as shown in the fourth column. At the same time, the similarity of the response to interest rate innovations and output innovations is strongly rejected in the third column. Although the exercise does not identify the reaction function, the behavior of the term structure in response to these shocks is more closely related to the reaction to funds rate innovations rather than the endogenous policy response to innovations in fundamentals.

Note also that in the long run (i.e., T=36), one cannot reject the similarity of all sources of shocks. Each shock explains a portion of the excess returns found in the data, and to some
extent each does so proportionally to the amount of overall interest rate movements that it explains.

<table>
<thead>
<tr>
<th>Table 3-6: Test of Proportion of Puzzling Interest Rate Movements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Horizon:</strong></td>
</tr>
<tr>
<td>Shocks Compared:</td>
</tr>
<tr>
<td>Interest Rate:</td>
</tr>
<tr>
<td>6-Month</td>
</tr>
<tr>
<td>12-Month</td>
</tr>
<tr>
<td>24-Month</td>
</tr>
<tr>
<td>60-Month</td>
</tr>
</tbody>
</table>

### 3.5 Discussion

Unconditionally, the expectations hypothesis performs very well in the VAR analysis presented here, as demonstrated in section 3.2. However, this finding masks the fact that the long-term interest rate underreacts to some movements in the short-term interest and overreacts to other components. In particular, there is a strong underreaction of long-term interest rates conditional on the endogenous response of monetary policy to output innovations. On the other hand, the endogenous reaction to price level shocks is characterized by overreaction of long-term interest rates.

In general, the puzzling behavior of the term structure is best described by the tendency of long-term interest rates to follow the short-term rate too closely. This is inconsistent with the expectations hypothesis since the future path of the funds rate varies in response to the different types of shocks. The expected reaction of the long-term interest rate is to generate a negative spread in response to price level and policy shocks and a positive spread (at least initially) in response to an output shock. The long-term rate instead generates a limited spread in reaction to both of these shocks, delivering the pattern of overreaction and underreaction.

104
described.

Recall that while the unconditional VAR evidence is broadly consistent with the expectations hypothesis, the failure of interest rate spread regressions may reflect transitory deviations from the expectations hypothesis. The conditional evidence indicates that transitory deviations arise in response to various shocks, so that the failure of the regression analysis is not inconsistent with the VAR analysis. In addition, the fact that the violation of the expectations hypothesis is characterized by an insufficient interest rate spread may have interesting implications for the spread regression approach described in section 3.2. If the most severe violations of the expectations hypothesis occur when the spread is limited, this approach may understate the magnitude of the violation of the expectations hypothesis.

The conditional responses may provide additional information about the unconditional failures of the expectations hypothesis. It has been suggested by Campbell and Shiller (1984, 1991) that the unconditional evidence may be interpreted as an underreaction of long-term interest rates to movements in the short-term rate. The above analysis indicates that it is not the case that the long-term interest rate pervasively underreacts to the short-term interest rate, although this is the case in response to particular components of monetary policy.

Perhaps the most intriguing finding in the conditional evidence is the sharp distinction in the success of the expectations hypothesis between the endogenous response of monetary policy to innovations in fundamental variables, and the response to financial market innovations that are orthogonal to innovations in fundamentals. Explanations for the failure of the expectations hypothesis must therefore describe the tendency for the long-term interest rate to mimic the funds rate in response to movement in prices and output, while at the same time explaining the absence of this tendency when movements originate in the financial markets.

One potential explanation is that bond holders have difficulty distinguishing the sources that is driving interest rate changes. The state of the economy may be difficult for the bond holders to observe, since there are many stochastic indicators for output and prices. Therefore, the bond holder may observe an endogenous movement in the federal funds rate without exactly knowing the state of the economy. Observing a movement in the funds rate, the bond trader is uncertain about whether this is the beginning of a sustained movement in the funds rate.
(an output shock) or a movement that will begin to revert more quickly (a price level shock). This permits a transitory underreaction to the output shock and a transitory overreaction to the price level shock for the period during which he learns about the type of the shock, which is consistent with the responses observed in the data. At the same time, it would be irrational for the trader to make a mistake on average, which delivers the observation that the expectations hypothesis appears to hold unconditionally. Financial market shocks are different, though. Innovations in longer-term interest rates reflect expectations about future movements in the funds rate. In this case bond holders have already established an expectation for the state of the economy rather than inferring it from the observed movement in the fundamentals. Similarly, a monetary policy shock may be more transparent since it is orthogonal to changes in the state of the economy. Interest rate movements are observed frequently and precisely and are therefore not subject to the same type of uncertainty found in movements in output and the price level. As a result, innovations that originate in the financial markets are not subject to the same learning dynamics and therefore do not exhibit transitory deviations from the expectations hypothesis.

\footnote{This would be even more likely if the reaction of the Fed to fundamentals is uncertain or stochastic, a feature of the reaction function discussed in Fuhrer (1996).}
Bibliography


