The Fed, the Bond Market, and Gradualism in Monetary Policy

JEREMY C. STEIN and ADI SUNDERAM[*]

ABSTRACT

We develop a model of monetary policy with two key features: the central bank has private information about its long-run target rate and is averse to bond market volatility. In this setting, the central bank gradually impounds changes in its target into the policy rate. Such gradualism represents an attempt to not spook the bond market. However, this effort is partially undone in equilibrium, as markets rationally react more to a given move when the central bank moves more gradually. This time-consistency problem means that society would be better off if the central bank cared less about the bond market.

FED WATCHING IS SERIOUS BUSINESS for bond market investors and for the financial market press that serves these investors. Speeches and policy statements by Federal Reserve officials are dissected word by word for clues they might yield about the future direction of policy. Moreover, the interplay between the central bank and the market goes in two directions: not only is the market keenly interested in making inferences about the Fed's reaction function, the Fed also takes active steps to learn what market participants *think the Fed is thinking*. In particular, before every Federal Open Market Committee (FOMC) meeting, the Federal Reserve Bank of New York performs a detailed survey of primary dealers, asking such questions as "Of the possible outcomes below, provide the percent chance you attach to the timing of the first increase in the federal funds target rate or range. Also, provide your estimate for the most likely meeting for the first increase."[1]

In this paper, we build a model that aims to capture the main elements of this two-way interaction between the Fed and the bond market. The two

[1] This particular question appeared in the September 2015 survey, among others. All the surveys, along with a tabulation of responses, are available at https://www.newyorkfed.org/markets/primarydealer_survey_questions.html.

distinguishing features of the model are that (i) the Fed has private information about its preferred value of the target rate and (ii) the Fed is averse to bond market volatility. These assumptions yield a number of positive and normative implications for the term structure of interest rates and the conduct of monetary policy. For the sake of concreteness, and to highlight the model's empirical content, we focus most of our attention on the well-known phenomenon of gradualism in monetary policy.

As described by Bernanke (2004), gradualism is the idea that "the Federal Reserve tends to adjust interest rates incrementally, in a series of small or moderate steps in the same direction." This behavior can be represented empirically by an inertial Taylor rule, with the current level of the federal funds rate modeled as a weighted average of a target rate—which itself is a function of inflation and the output gap as in, for example, Taylor (1993)—and the lagged value of the funds rate. In this specification, the coefficient on the lagged funds rate captures the degree of inertia in policy. In recent U.S. samples, estimates of the degree of inertia are strikingly high, on the order of 0.85 in quarterly data.[2]

Several authors have proposed theories of this kind of gradualism on the part of the central bank. One influential line of thinking, due originally to Brainard (1967) and refined by Sack (1998), is that moving gradually makes sense when there is uncertainty about how the economy will respond to a change in the stance of policy. An alternative rationale, proposed by Woodford (2003), argues that committing to move gradually gives the central bank more leverage over long-term rates for a given change in the short rate, a property that is desirable in the context of his model.

In what follows, we offer a different take on gradualism. In our model, the observed degree of policy inertia is not optimal from an ex ante perspective, but rather reflects a time-consistency problem. This time-consistency problem arises from our two key assumptions. First, we assume the Fed has private information about its preferred value of the target rate. In other words, the Fed knows something about its reaction function that the market does not. Although this assumption is not standard in the literature on monetary policy, it is necessary to explain the basic observation that financial markets respond to monetary policy announcements, and that market participants devote considerable time and energy to Fed watching.[3] Notably, our basic results depend only on the Fed having a small amount of private information. The majority of the variation in the Fed's target can come from changes in publicly observed variables such as unemployment and inflation. All that we require is that *some* variation also reflects innovations to the Fed's private information.

---

[2] Coibion and Gorodnichenko (2012) provide a comprehensive recent empirical treatment; see also Rudebusch (2002, 2006). Campbell, Pflueger, and Viceira (2015) argue that the degree of inertia in Federal Reserve rate-setting became more pronounced after about 2000.

[3] Put differently, if the Fed mechanically followed a policy rule that was a function of only publicly observable variables (e.g., the current values of the inflation rate and the unemployment rate), then the market would react to news releases about movements in these variables but not to Fed policy statements.

Second, we assume that the Fed behaves as if it is averse to bond market volatility. We model this concern in reduced form by simply putting the volatility of long-term rates into the Fed's objective function. However, a preference of this sort can be rooted in an effort to deliver on the Fed's traditional dual mandate. For example, a bout of bond market volatility may be undesirable not in its own right, but rather because it is damaging to the financial system and hence to real economic activity and employment.

Nevertheless, in a world of private information and discretionary meeting-by-meeting decision making, an attempt by the Fed to moderate bond market volatility can be welfare-reducing. The logic is similar to that in signal-jamming models (Stein (1989), Holmstrom (1999)). Suppose the Fed observes a private signal that its long-run target for the funds rate has permanently increased by 100 basis points (bps). If it adjusts fully in response to this signal, raising the funds rate by 100 bps, long-term rates will move by a similar amount. If it is averse to such a large movement in long-term rates, the Fed will be tempted to announce a smaller change in the funds rate, trying to fool the market into thinking that its private signal was less dramatic. Hence, it will underadjust to its signal, raising the funds rate by perhaps only 25 bps.

However, if bond market investors understand this dynamic, the Fed's efforts to reduce volatility will be partially frustrated in equilibrium. The market will see the 25- basis point increase in the funds rate and understand that it is likely to be the first in a series of similar moves, so long-term rates will react more than one-for-one to the change in the short rate. Still, if it acts on a discretionary basis, the Fed will always try to fool the market. This is because when it decides how much to adjust the policy rate, it *takes as given the market's conjecture* about the degree of inertia in its rate-setting behavior. As a result, the Fed's behavior is inefficient from an ex ante perspective: Because in equilibrium the market understands the Fed's incentives, moving gradually has limited effectiveness in reducing bond market volatility, but causes the policy rate to be further from its long-run target than it otherwise would be.

This inefficiency reflects a commitment problem. In particular, the Fed cannot commit to not trying to smooth the private information that it communicates to the market via its changes in the policy rate.[4] One institutional solution to this problem, in the spirit of Rogoff (1985), would be to appoint a central banker who cares less about bond market volatility than the representative member of society. More broadly, appointing such a market-insensitive central banker can be thought of as a metaphor for building an institutional culture and set of norms inside the central bank such that high-frequency bond market movements are not given as much weight in policy deliberations.

---

[4] The literature on monetary policy has long recognized a different commitment problem, namely that, under discretion, the central bank will be tempted to create surprise inflation so as to lower the unemployment rate. See, for example, Kydland and Prescott (1977) and Barro and Gordon (1983). More recently, Farhi and Tirole (2012) have pointed to the time-consistency problem that arises from the central bank's ex post desire to ease monetary policy when the financial sector is in distress; their focus on the central bank's concern with financial stability is somewhat closer in spirit to ours.

We begin with a simple static model that is designed to capture the above intuition in as parsimonious a way as possible. The main result here is that in any rational expectations equilibrium, there is always underadjustment of the policy rate compared to the first-best outcome. Moreover, for some parameter values, there can be Pareto-ranked multiple equilibria with different degrees of underadjustment. The intuition for these multiple equilibria is that there is two-way feedback between the market's expectations about the degree of gradualism on the one hand and the Fed's optimal choice of gradualism on the other.[5] Specifically, if the market conjectures that the Fed is behaving in a highly inertial fashion, it will react more strongly to an observed change in the policy rate. In an inertial world, the market knows that there are further changes to come. But this strong sensitivity of long-term rates to changes in the policy rate makes the Fed all the more reluctant to move the policy rate, validating the initial conjecture of extreme inertia.

As noted above, our results in the static model generalize to the case in which, in addition to private information, there is public information about changes in the Fed's target. Strikingly, the Fed moves just as gradually with respect to this public information as it does with respect to private information. This is true independent of the relative contributions of public and private information to the total variance of the target. The logic is as follows. When public information arrives suggesting that the target has risen—for example, inflation increases—the Fed is tempted to act as if it has received dovish private information at the same time to mitigate the overall impact on the bond market. This means that it raises the funds rate by less than it otherwise would in response to the publicly observed increase in inflation. Thus, our model shows that even a small amount of private information can lead the Fed to move gradually with respect to *all* information.

We next enrich the model by adding publicly observed term-premium shocks as an additional source of variation in long-term rates. We show that, similar to the case of public information about its target, the Fed's desire to moderate bond market volatility leads it to try to offset term-premium shocks, cutting the policy rate when term premiums rise. Again, this tactic is partially undone by the market in equilibrium. Thus, the presence of term-premium shocks exacerbates the Fed's time-consistency problem.

Finally, we extend the model to an explicitly dynamic setting, which allows us to more fully characterize how a given innovation to the Fed's target works its way into the funds rate over time. These dynamic results enable us to draw a closer link between the mechanism in our model and the empirical evidence on the degree of inertia in the funds rate.

Overall, this paper carries two distinct messages, one positive and one normative. On the positive front, we argue that the Fed's private information can help explain the well-documented phenomenon of gradualism in monetary policy. To be clear, we do not claim that our private-information story by itself provides a complete quantitative explanation for the degree of gradualism observed in the

---

[5] For an informal description of this two-way feedback dynamic, see Stein (2014).

data. Other motives, such as Brainard's (1967) instrument-uncertainty mechanism, are likely to play a role as well. Nevertheless, our model shows that private information amplifies the extent of gradualism and can therefore help make sense of the empirical magnitudes. In addition, the model provides a unified explanation for both gradualism and the Fed's reaction to term-premium shocks.

On the normative front, we show that in the presence of private information, a concern on the part of the Fed about bond market volatility can lead to welfare losses in the discretionary outcome as compared to the solution with commitment—in other words, there is a time-consistency problem. This normative implication of the model is more unique and suggests that it can be socially valuable to foster a central-banking culture that leads high-frequency bond market movements to be given less attention in policy deliberations.

The remainder of the paper is organized as follows. Section I discusses motivating evidence based on readings of FOMC transcripts. Section II presents the static version of the model and summarizes our basic results on underadjustment of the policy rate, multiple equilibria, and term premiums. Section III develops a dynamic extension of the model. Section IV discusses a variety of other implications of our framework. Section V concludes.

I. Motivating Evidence from FOMC Transcripts

In their study of monetary policy inertia, Coibion and Gorodnichenko (2012) use FOMC transcripts to document two key points. First, FOMC members sometimes speak in a way that suggests a gradual adjustment model—that is, they articulate a target for the policy rate and then put forward reasons to adjust only slowly toward that target. Second, one rationale for such gradualism appears to be a desire not to create financial market instability. Coibion and Gorodnichenko (2012, p. 150) highlight the following quote from Chairman Alan Greenspan at the March 1994 FOMC meeting:

> My own view is that eventually we have to be at 4 to 4½%. The question is not whether but when. If we are to move 50 basis points, I think we would create far more instability than we realize, largely because a half-point is not enough to remove the question of where we are ultimately going. I think there is a certain advantage in doing 25 basis points . . .

In a similar spirit, at the August 2004 meeting, after the Fed had begun to raise the funds rate from the low value of 1% that had prevailed since mid-2003, Chairman Greenspan remarked:

> Consequently, the sooner we can get back to neutral, the better positioned we will be. We were required to ease very aggressively to offset the events of 2000 and 2001, and we took the funds rate down to extraordinarily low levels with the thought in the back of our minds, and often said around this table, that we could always reverse our actions. Well, reversing is not all that easy . . . We've often discussed that ideally we'd like to be in a position

where, when we move as we did on June 30 and I hope today, the markets
respond with a shrug. What that means is that the adjustment process
is gradual and does not create discontinuous problems with respect to
balance sheets and asset values.

These sorts of quotes help motivate our basic modeling approach, in which
gradualism reflects the Fed's desire to keep bond market volatility in check—
in Greenspan's words, to "not create discontinuous problems with respect to
balance sheets and asset values." This same approach may also be helpful in
thinking about changes in gradualism over time. Campbell, Pflueger, and Vi-
ceira (2015) show that inertia in Fed rate-setting behavior became significantly
more pronounced after 2000. Given the logic of our model, this heightened iner-
tia could be driven by an increase in the Fed's concern about financial markets
over time.

In a crude attempt to speak to this question, we examine all 216 FOMC
transcripts over the 27-year period from 1985 to 2011 and simply measure
the frequency of words related to financial markets.[6] Specifically, we count the
number of times the terms "financial market," "equity market," "bond market,"
"credit market," "fixed income," and "volatility" are mentioned.[7] For each year,
we aggregate this count and divide by the total number of words in that year's
transcripts.

Figure 1 displays the results from this exercise. As can be seen, there is a
strong upward trend in the data. While there is also a good deal of volatility
around the trend, a simple linear time trend captures almost 52% of the vari-
ation in the time series, with a t-statistic of 4.6.[8] Moreover, the fitted value
of the series based on this time trend goes up about fourfold over the 27-year
sample period. While extremely simplistic, this analysis does suggest an in-
creasing emphasis over time by the FOMC on financial market considerations.
If there was such a change in FOMC thinking, the model we develop below is
well suited to drawing out its implications, both for the dynamics of the policy
rate and for longer-term yields.

II. Static Model

We begin by presenting what is effectively a static version of the model, in
which the Fed adjusts the funds rate only partially in response to a one-time
innovation in its desired target rate. In Section IV, we extend the model to a
fully dynamic setting in which we can be more explicit about how an innovation
to the Fed's target rate gradually makes its way into the funds rate over time.

---

[6] Given the five-year lag in making transcripts public, 2011 is the last available year. Cieslak
and Vissing-Jorgenson (2017) perform a similar but more sophisticated exercise to understand why
the Fed responds to the stock market.

[7] We obtain similar results if we use different subsets of these terms. For instance, the results
are similar if we only count the frequency of the term "financial market."

[8] Restricting the sample to the precrisis period from 1985 to 2006, we obtain an R^2 of 49% and
a t-statistic of 4.4. So the trend in the data is not driven primarily by the post-2006 period.
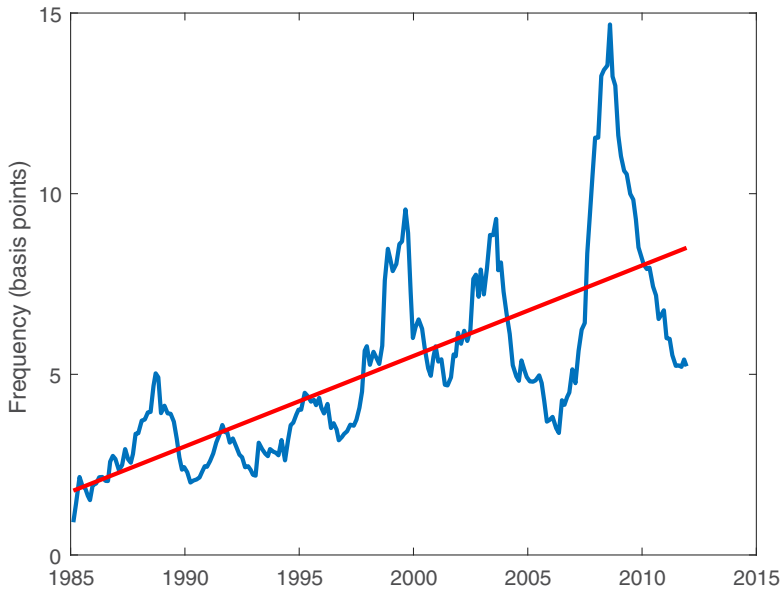
Figure 1. Frequency of terms related to financial markets in FOMC transcripts, 1985–2011. The figure plots the moving average of the frequency of terms related to financial markets over the last eight FOMC meetings. (Color figure can be viewed at wileyonlinelibrary.com)

Though it is stylized, the static model highlights the main intuition for why a desire to limit bond market volatility creates a time-consistency problem for the Fed.

A. Model Setup

We begin by assuming that, at any time t, the Fed has a target rate based on its traditional dual-mandate objectives. This target rate is the Fed's best estimate of the value of the federal funds rate that keeps inflation and unemployment as close as possible to their desired levels. For tractability, we assume that the target rate, denoted by i_t^*, follows a random walk, so that

$$i_t^* = i_{t-1}^* + \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2 \equiv \frac{1}{\tau_\varepsilon})$ is normally distributed. Our key assumption is that i_t^* is private information of the Fed and is unknown to market participants before the Fed acts at time t. One can think of the private information embodied in i_t^* as arising from the Fed's attempts to follow something akin to a Taylor rule, where it has private information about the appropriate coefficients to use in the rule (i.e., its reaction function) or about its own forecasts of future inflation or unemployment. As noted above, an assumption of private information along these lines is necessary if one wants to understand why asset prices respond to Fed policy announcements.

Once it knows the value of i_t^*, the Fed acts to incorporate some of its new private information ε_t into the federal funds rate i_t, which is observable to the market. We assume that the Fed picks i_t on a discretionary period-by-period basis to minimize the loss function L_t, given by

$$L_t = \left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2, \tag{2}$$

where i_t^∞ is the infinite-horizon forward rate. Thus, the Fed has the usual concerns about inflation and unemployment, as captured in reduced form by a desire to keep i_t^* close to i_t. However, when $\theta > 0$, the Fed also cares about the volatility of long-term bond yields, as measured by the squared change in the infinite-horizon forward rate.

For simplicity, we start by assuming that the expectations hypothesis holds, so there is no time-variation in the term premium. This implies that the infinite-horizon forward rate is equal to the expected value of the funds rate that will prevail in the distant future. Because the target rate i_t^* follows a random walk, the infinite-horizon forward rate is then given by the market's best estimate of the *current target i_t^** at time t. Thus, we have $i_t^\infty = E_t[i_t^*]$ so long as we are in an equilibrium where rates eventually adjust toward the target, no matter how slowly. In Section III.E.2, we relax the assumption that the expectations hypothesis holds so that we can also consider the Fed's reaction to term-premium shocks.

Several features of the Fed's loss function are worth discussing. First, in our simple formulation, θ reflects the degree to which the Fed is concerned about bond market volatility, over and above its desire to keep i_t close to i_t^*. To be clear, this loss function need not imply that the Fed cares about asset prices for their own sake. An alternative interpretation is that volatility in financial market conditions can affect the real economy and hence the Fed's ability to satisfy its traditional dual mandate. This linkage is not modeled explicitly here, but as one example of what we have in mind, the Fed might believe that a spike in bond market volatility could damage highly levered financial intermediaries and interfere with the credit supply process. With this stipulation in mind, we take as given that θ reflects the socially "correct" objective function—in other words, it is exactly the value that a well-intentioned social planner would choose. We then ask whether there is a time-consistency problem when the Fed tries to optimize this objective function period by period, in the absence of a commitment technology.

Second, note that the Fed's target in the first term of equation (2) is the *short-term* policy rate, whereas the bond market volatility that it is concerned about in the second term of equation (2) refers to the variance of *long-term* market-determined rates. This short versus long divergence is crucial for our results on time-consistency. To see why, consider two alternative loss functions:

$$L_t = \left(i_t^* - i_i\right)^2 + \theta(\Delta i_i)^2, \tag{2'}$$

$$L_t = \left(i_t^{\infty*} - i_t^{\infty}\right)^2 + \theta\left(\Delta i_t^{\infty}\right)^2. \tag{2''}$$

In equation (2′), the Fed cares about the volatility of the funds rate, rather than the volatility of the long-term rate. This objective function mechanically produces gradual adjustment of the funds rate, but because there is no forward-looking long rate, there is no issue of the Fed trying to manipulate market expectations and hence no time-consistency problem. Thus, in positive terms, a model of this sort delivers gradualism, but in normative terms this gradualism is entirely optimal from the Fed's perspective. However, by emphasizing only short-rate volatility, this formulation arguably fails to capture the financial stability goal articulated by Greenspan, namely, to "not create discontinuous problems with respect to balance sheets and asset values." We believe that putting the volatility of the long-term rate directly in the objective function, as in equation (2), does a better job in this regard.

In equation (2″), the Fed pursues its dual-mandate objective not by having a target for the funds rate, but instead by explicitly targeting the long rate of $i_t^{\infty*}$. In this case, given that the same rate i_t^{∞} appears in both parts of the objective function, there is again no time-consistency problem.[9] Moreover, one might think that the first term in equation (2″) is a reasonable reduced-form way to model the Fed's efforts to achieve its mandate. In particular, in standard New Keynesian models, it is long-term real rates, not short rates, that matter for inflation and output stabilization.

Nevertheless, our formulation of the Fed's objective function in equation (2) can be motivated in a couple of ways. First, equation (2) is arguably more realistic than equation (2″) as a description of how the Fed actually behaves, and—importantly for our purposes—communicates about its future intentions. For example, in recent statements, the Fed has argued that one motive for adjusting policy gradually is the fact that *r**, defined as the equilibrium (or neutral) value of the short rate, is itself slow-moving.[10] But the concept of a slow-moving *r** can only make logical sense if the short rate matters directly for real outcomes. If instead real outcomes were entirely a function of long rates, the near-term speed of adjustment of the short rate would be irrelevant, holding fixed the total expected adjustment.

Second, in many other models of the monetary transmission mechanism outside of the New Keynesian genre, the short rate does have an important independent effect on economic activity. For instance, the literature on the bank lending channel, including Bernanke and Blinder (1992), Kashyap and Stein (2000), and Drechsler, Savov, and Schnabl (2017), finds that bank loan

---

[9] We are grateful to Mike Woodford for emphasizing this point to us and for providing a simple proof.

[10] In Chair Janet Yellen's press conference of December 16, 2015, she said: "This expectation (of gradual rate increases) is consistent with the view that the neutral nominal federal funds rate—defined as the value of the federal funds rate that would be neither expansionary nor contractionary if the economy were operating near potential—is currently low by historical standards and is likely to rise only gradually over time." See http://www.federalreserve.gov/mediacenter/files/FOMCpresconf20151216.pdf.

supply is directly influenced by changes in the short rate because the short rate effectively governs the availability of low-cost deposit funding. Similarly, recent papers on "reaching for yield," including Gertler and Karadi (2015), Jimenez et al. (2014), and Dell'Ariccia, Laeven, and Suarez (2013), argue that risk premiums and ultimately real activity respond to the level of the short rate. Both of these channels help explain why it can make sense for the Fed to target the short rate per se.

The assumption that the Fed cares about the volatility of the infinite-horizon forward rate i_t^∞, along with the fact that the target rate i_t^* follows a random walk, makes for a convenient simplification. Because i_t^* follows a random walk, all of the new private information ε_t will, in expectation, eventually be incorporated into the future short rate. Focusing on the reaction of the infinite-horizon forward rate reflects this revision in expectations and allows us to abstract from the exact dynamic path that the Fed follows in ultimately incorporating its new information into the funds rate. In contrast, revisions in finite-horizon rates depend on the exact path of the short rate. For example, suppose i_t^* were public information and changed from 1% to 2%. The infinite-horizon forward rate would immediately jump to 2%, while the two-year forward rate would move less if the market expected a slow tightening. However, for our key conclusions, it is not necessary that the Fed care about the volatility of the infinite-horizon forward rate; as we demonstrate below, a time-consistency problem still emerges if the Fed cares instead about the volatility of a finite-horizon rate that is effectively a weighted average of short-term and infinite-horizon rates.[11]

Finally, our formulation in equation (2) assumes that the Fed optimizes a myopic period-by-period objective function, so that when it acts at time t, it cares only about the consequences for bond market volatility at t, and not about volatility in future periods. However, in Sections III.D.1 and IV.A, we show that we obtain very similar results when the Fed has a forward-looking objective function that incorporates a concern about both future as well as current volatility.

B. Equilibrium

With the Fed's objective function in place, we are now ready to describe the nature of equilibrium in the static model. Once it observes i_t^*, we assume that the Fed sets the federal funds rate i_t by following a partial adjustment rule of the form

$$i_t = i_{t-1} + k\left(i_t^* - i_{t-1}\right) + u_t, \qquad (3)$$

where $u_t \sim N(0, \sigma_u^2 \equiv \frac{1}{\tau_u})$ is normally distributed noise that is overlaid onto the rate-setting process. The noise u_t is a modeling device the usefulness of which

---

[11] The time-consistency problem arises because of the forward-looking nature of longer term rates, and the resulting incentives that the Fed has to manipulate market expectations. As long as the Fed seeks to moderate the volatility of a finite-horizon rate that has some forward-looking component, an element of the time-consistency problem will remain.

will become clear shortly. Loosely speaking, this noise, which can be thought of as a "tremble" in the Fed's otherwise optimally chosen value of i_t, ensures that the Fed's actions cannot be perfectly inverted to fully recover its private information i_t^*. As will be seen, this imperfect inversion feature helps avoid some degenerate equilibrium outcomes. For a concrete interpretation, one can think of u_t as coming from the Fed's use of round numbers (typically in 25 bps increments) for the funds rate settings that it communicates to the market, while its underlying private information about i_t^* is presumably continuous.

The market tries to infer the Fed's private information i_t^* based on its observation of the funds rate i_t. To do so, it conjectures that the Fed follows a rule given by

$$i_t = i_{t-1} + \kappa \left(i_t^* - i_{t-1} \right) + u_t. \tag{4}$$

Thus, the market correctly conjectures the form of the Fed's smoothing rule but, crucially, does not directly observe the Fed's smoothing parameter k; rather, it has to make a guess κ as to the value of this parameter. In a rational expectations equilibrium, this guess will turn out to be correct, and we will have $\kappa = k$. However, the key piece of intuition is that when the Fed chooses k, it takes the market's conjecture κ as a fixed parameter and does not impose $\kappa = k$. The equilibrium concept is thus of the signal-jamming type introduced by Holmstrom (1999): The Fed, taking the market's estimate of κ as fixed, tries to fool the market into thinking that i_t^* has moved by less than it actually has, in an effort to reduce the volatility of long-term rates. In equilibrium, the market will rationally unwind the Fed's actions and not be misled, but the Fed cannot resist the temptation to try.[12]

Suppose that the economy was previously in steady state at time $t - 1$, with $i_{t-1} = i_{t-1}^*$. Given the Fed's adjustment rule, the funds rate at time t satisfies

$$i_t = i_{t-1} + k \left(i_t^* - i_{t-1} \right) + u_t = i_{t-1} + k \varepsilon_t + u_t. \tag{5}$$

Based on its conjecture about the Fed's adjustment rule in equation (4), the market tries to back out i_t^* from its observation of i_t. Because both shocks ε_t and u_t are normally distributed, the market's expectation is given by

$$E\left[i_t^* | i_t \right] = i_{t-1} + \frac{\kappa \tau_u}{\tau_\varepsilon + \kappa^2 \tau_u} \left(i_t - i_{t-1} \right) = i_{t-1} + \chi \left(i_t - i_{t-1} \right), \tag{6}$$

where $\chi = \frac{\kappa \tau_u}{\tau_\varepsilon + \kappa^2 \tau_u}$.

The less noise there is in the Fed's adjustment rule, the higher is τ_u and the more the market reacts to the change in the rate i_t.

---

[12] There is a close analogy to models in which corporate managers with private information pump up their reported earnings in an effort to impress the stock market. See, for example, Stein (1989).

In light of equations (5) and (6) and the random-walk property that $i_t^\infty = E_t[i_t^*]$, the Fed's loss function, taking expectations over the realization of the noise, can be written as

$$L_t = E_{u_t}\left[\left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2\right] = (1-k)^2\varepsilon_t^2 + \sigma_u^2 + \theta\chi^2\left(k^2\varepsilon_t^2 + \sigma_u^2\right). \quad (7)$$

The Fed then minimizes this loss function by picking the optimal value of k. Again, we emphasize that, in doing so, it takes the market's conjectures about its behavior, and hence the parameter χ, as fixed. The first-order condition with respect to k then yields

$$k = \frac{1}{1 + \theta\chi^2}. \quad (8)$$

In rational expectations equilibrium, the market's conjecture turns out to be correct, so we have $\kappa = k$. Imposing this condition, we have that, in equilibrium, the Fed's adjustment rule satisfies

$$k = \frac{\left(\tau_\varepsilon + k^2\tau_u\right)^2}{\left(\tau_\varepsilon + k^2\tau_u\right)^2 + \theta(k\tau_u)^2}. \quad (9)$$

It follows immediately from equation (9) that the Fed follows a partial adjustment rule in any rational expectations equilibrium.

PROPOSITION 1: *In any rational expectations equilibrium, the Fed's adjustment to a change in its target rate is partial: we have $k < 1$ so long as $\theta > 0$.*

B.1. Equilibrium with No Noise in the Rate-Setting Process

To build some intuition for the Fed's behavior, let us begin by considering the simple limiting case in which there is no noise in the rate-setting process: $\sigma_u^2 = 0$ (i.e., $\tau_u \to \infty$). In this case, the market's inference of i_t^* is simply

$$E\left[i_t^*|i_t\right] = i_{t-1} + \frac{(i_t - i_{t-1})}{\kappa} = i_{t-1} + \frac{k}{\kappa}\varepsilon_t, \quad (10)$$

and the Fed's loss function is

$$L_t = \left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2 = ((1-k)\varepsilon_t)^2 + \theta\left(\frac{k}{\kappa}\varepsilon_t\right)^2. \quad (11)$$

When the Fed considers lowering k, it trades off being further away from the optimal i_t^* against the fact that it believes it can reduce bond market volatility

by moving more slowly. In rational expectations equilibrium, the optimal k now satisfies:[13]

$$k = \frac{k^2}{k^2 + \theta}.$$ (12)

First note that when $\theta = 0$, the only solution to equation (12) is given by $k = 1$. When the Fed does not care about bond market volatility, it fully adjusts to its new private information ε_t. By contrast, when $\theta > 0$, there may be more than one solution to equation (12), but in any equilibrium it must be the case that $k < 1$, and the Fed only partially adjusts.

Moreover, when $\theta > 0$, it is always the case that $k = 0$ satisfies equation (12). Thus, there is a degenerate solution in which *the Fed does not adjust the funds rate at all*. This somewhat unnatural outcome is a function of the extreme feedback that arises between the Fed's adjustment rule and the market's conjecture when there is no noise in the rate-setting process. Specifically, when the market conjectures that the Fed never moves the funds rate at all (i.e., that $\kappa = 0$), then even a tiny out-of-equilibrium move by the Fed leads to an infinitely large change in the infinite-horizon forward rate. Thus, the Fed validates the market's conjecture by not moving at all, that is, by choosing $k = 0$.

However, as soon as there is even an infinitesimal amount of noise u_t in the rate-setting process, this extreme $k = 0$ equilibrium is ruled out, as small changes in the funds rate now lead to bounded market reactions. This explains our motivation for keeping rate-setting noise in the more general model.

In addition to $k = 0$, for $0 < \theta < 0.25$, there are two other solutions to equation (12), given by

$$k = \frac{1 \pm \sqrt{(1 - 4\theta)}}{2}.$$ (13)

Of these, the larger of the two values also represents a stable equilibrium outcome. Thus, as long as θ is not too big, the model also admits a nondegenerate equilibrium, with $\frac{1}{2} < k < 1$, even with $\sigma_u^2 = 0$. Within this region, higher values of θ lead to lower values of k. In other words, the more intense the Fed's concern about bond market volatility, the more gradually it adjusts the funds rate.

B.2. Equilibrium Outcomes across the Parameter Space

Now we return to the general case in which $\sigma_u^2 > 0$ and explore the range of outcomes produced by the model for different parameter values. In each panel of Figure 2, we plot the Fed's best response k as a function of the market's conjecture κ. Any rational expectations equilibrium must lie on the 45° line where $k = \kappa$.

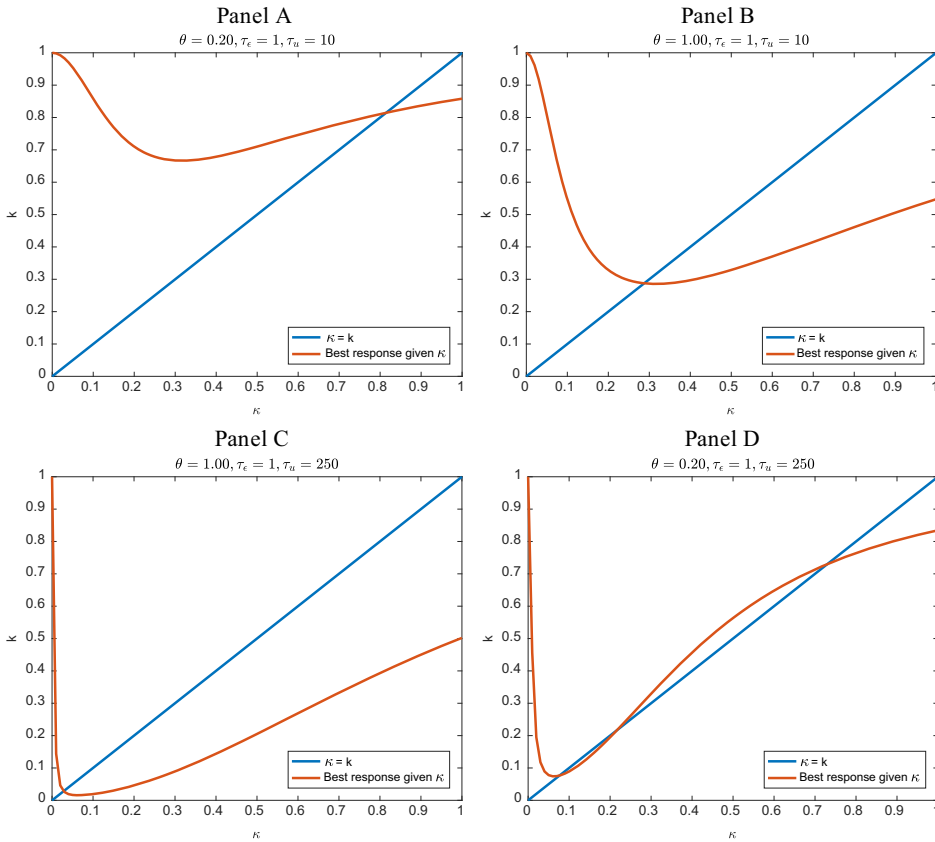[13] Equation (12) can be derived from equation (9) by taking the limit as τ_u goes to infinity and applying L'Hopital's Rule.

Figure 2. Equilibria of the static model for different parameter values. θ indexes the Fed's aversion to volatility, τ_ε is the precision of innovations to the target rate, τ_u is the precision of rate-setting noise, κ is the market's conjecture about the Fed's degree of adjustment, k is the Fed's best response degree of adjustment. Panel A: $\theta = 0.2$, $\tau_\varepsilon = 1$, and $\tau_u = 10$. Panel B: $\theta = 1.0$, $\tau_\varepsilon = 1$, and $\tau_u = 10$. Panel C: $\theta = 1.0$, $\tau_\varepsilon = 1$, and $\tau_u = 250$. Panel D: $\theta = 0.2$, $\tau_\varepsilon = 1$, and $\tau_u = 250$. (Color figure can be viewed at wileyonlinelibrary.com)

In Panel A of the figure, we begin with a relatively low value of θ, namely, 0.2, and set $\sigma_\varepsilon^2/\sigma_u^2 = 10$, so that the variance of rate-setting noise is one-tenth the variance of innovations to the target rate. As can be seen, this leads to a unique equilibrium with a high value of k of 0.81. When the Fed cares only a bit about bond market volatility, it adjusts rates fairly aggressively because it does not want to deviate too far from its target rate i_t^*.

In Panel B, we raise θ to 1.0, while keeping $\sigma_\varepsilon^2/\sigma_u^2 = 10$ as in Panel A. In this high-θ case, the unique equilibrium involves a lower value of k of 0.29. Thus, when the Fed cares a lot about bond market volatility, the funds rate significantly underadjusts to new information.

In Panel C, we keep $\theta = 1.0$ but decrease the rate-setting noise, so that $\sigma_\varepsilon^2/\sigma_u^2 = 250$. Now the equilibrium involves an extremely low value of k of just

0.03. The difference between Panels B and C is the degree of rate-setting noise. The greater amount of noise in Panel B allows the Fed to hide the information content of its actions, so that it can be more responsive to changes in its target without changing the market's inference about its private information by as much.[14]

Finally, in Panel D, we set $\theta = 0.2$ and $\sigma_\varepsilon^2 / \sigma_u^2 = 250$. Here, we have multiple equilibria: the Fed's best response crosses the 45° line in three places. Of these three crossing points, the two outer ones (at $k = 0.08$ and 0.73) correspond to stable equilibria, in the sense that if the market's initial conjecture κ takes an out-of-equilibrium value, the Fed's best response to that conjecture will tend to drive the outcome toward one of these two extreme crossing points.

The existence of multiple equilibria highlights an essential feature of the model: the potential for market beliefs about Fed behavior to become self-validating. If the market conjectures that the Fed adjusts rates only very gradually, then even small changes are heavily freighted with informational content about the Fed's reaction function. Given this strong market sensitivity and its desire not to create too much volatility, the Fed may then choose to move very carefully, even at the cost of accepting a funds rate that is quite far from its target. Conversely, if the market conjectures that the Fed is more aggressive in its adjustment of rates, it reacts less to any given movement, which frees the Fed up to track the target rate more closely.

When there are multiple equilibria, the one with the higher value of k typically leads to a better outcome from the perspective of the Fed's objective function: With a higher k, i_t is closer to i_t^*, but in equilibrium, bond market volatility is not much greater. Thus, if we are in a range of the parameter space where multiple equilibria are possible, it is important for the Fed to avoid doing anything in its communications that tends to lead the market to conjecture too low a value of κ. Even keeping its own preferences fixed, fostering an impression of strong gradualism among market participants can lead to an undesirable outcome in which the Fed gets stuck in the low-k equilibrium.

C. The Time-Consistency Problem

A central property of our model is that there is a time-consistency problem: The Fed would do better in terms of optimizing its own objective function if it were able to commit to behaving as if it had a lower value of θ than it actually does. This is because while the Fed is always tempted to move the funds rate gradually so as to reduce bond market volatility, this desire is partially undone in equilibrium: The more gradually the Fed acts, the more the market responds to any change in rates, and yet the Fed is still left with a funds rate that is further from target on average.

This point is most transparent if we consider the limiting case in which there is no noise in the Fed's rate-setting process. In this case, once we impose the

---

[14] This is similar to the intuition in Kyle (1985). When there are more noise traders, the insider can trade more aggressively on his private information without as much market impact.

rational expectations assumption that $\kappa = k$, the Fed's attempts at smoothing have no effect on bond market volatility in equilibrium:

$$\left(\Delta i_t^\infty\right)^2 = \left(\frac{k}{\kappa}\varepsilon_t\right)^2 = \varepsilon_t^2. \tag{14}$$

Thus, the value of the Fed's loss function is

$$L_t = \left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2 = ((1-k)\varepsilon_t)^2 + \theta\varepsilon_t^2, \tag{15}$$

which is decreasing in k for $k < 1$.

To the extent that the target rate i_t^* is nonverifiable private information, it is hard to think of a contracting technology that can readily implement the first-best outcome under commitment: How does one write an enforceable rule that says that the Fed must always react fully to its private information? Thus, discretionary monetary policy will inevitably be associated with some degree of inefficiency. However, even in the absence of a binding commitment technology, there may still be scope for improving on the fully discretionary outcome. One possible approach follows in the spirit of Rogoff (1985), who argues that society should appoint a central banker who is more hawkish on inflation than is society itself. The analogy in the current context is that society should appoint a central banker who cares less about financial market volatility (i.e., has a lower value of θ) than society as a whole. Put differently, society—and the central bank itself—should foster an institutional culture and set of norms that discourage the monetary policy committee from being overly attentive to short-term market volatility considerations.

To see this, consider the problem of a social planner choosing a central banker whose concern about financial market volatility is given by θ_c. This central banker will implement the rational expectations adjustment rule $k(\theta_c)$, where k is given by equation (9), replacing θ with θ_c. In the rational expectations equilibrium, this will result in bond market volatility

$$\mathrm{Var}\left[\chi\left(k\left(\theta_c\right)\varepsilon_t + u_t\right)\right] = \frac{k(\theta_c)^2\tau_u}{\tau_\varepsilon\left(\tau_\varepsilon + k(\theta_c)^2\tau_u\right)}. \tag{16}$$

In contrast to the no-noise case, here bond market volatility is not invariant to the choice of k. Thus, the planner's ex ante problem is to pick θ_c to minimize its ex ante loss, recognizing that its own concern about financial market volatility is given by θ:

$$E_{\varepsilon_t}\left[\left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2\right] = E\left[((1-k(\theta_c))\varepsilon_t + u_t)^2 + \theta(\chi\left(k\left(\theta_c\right)\varepsilon_t + u_t\right))^2\right]. \tag{17}$$

The following proposition characterizes the optimal θ_c.

PROPOSITION 2: *In the presence of rate-setting noise, it is ex ante optimal to appoint a central banker with $\theta_c < \theta$. In the absence of rate-setting noise, it is ex ante optimal to appoint a central banker with $\theta_c = 0$, so that $k(\theta_c) = 1$.*

Table I

Characterization of the Static Model with Noise

This table characterizes the equilibrium in the static model with noise for different parameter values. θ indexes the Fed's aversion to volatility, τ_u/τ_ε is the ratio of the variance of innovations to the target rate to the variance of rate-setting noise, k is the equilibrium degree of adjustment without commitment, θ_c is the volatility aversion of the optimal central banker under commitment, L is the expected value of the Fed's loss function without commitment, and L_c is the value of the loss function under commitment.

Θ	τ_u/τ_ε	k	θ_c	$(L - L_c)/L$
0.2	10	0.81	0.02	0.09
1.0	10	0.29	0.13	0.07
1.0	250	0.03	0.004	0.11

When there is no rate-setting noise, the Fed's attempts to fool the market are completely frustrated in equilibrium. Therefore, it is optimal to set $\theta_c = 0$ and appoint a central banker who does not care about bond market volatility at all. When there is rate-setting noise, it is no longer true that the Fed would like to commit to behaving as if θ were exactly equal to zero. However, it would still like to commit to behaving as if θ were considerably smaller than its actual value.

Table I illustrates characterizing the optimal θ_c and the magnitude of the time-consistency problem for the same sets of parameter values as in the first three panels of Figure 1. In the first row, we set $\theta = 0.20$ and $\sigma_\varepsilon^2/\sigma_u^2 = 10$. In this case, the optimally chosen central banker has $\theta_c = 0.02$, and there is a 9% improvement in the Fed's loss function by appointing such a banker. The second row shows that, with $\theta = 1.0$ and $\sigma_\varepsilon^2/\sigma_u^2 = 10$, the optimally chosen central banker has $\theta_c = 0.13$ and brings a 7% improvement in the loss function. In the third row, where we keep $\theta = 1.0$ but set $\sigma_\varepsilon^2/\sigma_u^2 = 250$, θ_c declines to 0.004. Intuitively, with very little rate-setting noise in this last case, a time-inconsistent Fed would be tempted to behave in an extremely gradualist fashion, so there is more value in appointing a central banker who cares almost not at all about the bond market.[15] Overall, the table shows that for our parameter values it is optimal to set θ_c much lower than θ, often by an order of magnitude or more, and that this can achieve significant welfare gains.

D. Changing the Fed's Objective Function

We next consider two variations on the Fed's objective function and show that the general properties of the model—namely, gradual adjustment and the existence of a time-consistency problem—are qualitatively unchanged.

---

[15] We omit the fourth set of parameters that we used in Figure 1 because they generate multiple equilibria, which are not central to the rest of the discussion.

D.1. Forward-Looking Objective: A Concern with Future Volatility

We first consider what happens when the Fed has a forward-looking objective rather than a myopic period-by-period one. Note that in the loss function specified in equation (2), when the Fed acts at time t, it worries about its impact on bond market volatility at t but does not take into account the consequences of its decision for volatility in future periods. One might suspect that if it did, its temptation to act gradually would be reduced, since by doing so it simply defers any adjustment of prices to the future.

To address this issue, suppose that there is a single realization of private information ε_t at time t and no further shocks after that. For simplicity, suppose further that the Fed follows the partial adjustment rule in equation (3) at time t and then fully impounds ε_t into the funds rate at time $t + 1$. That is, $i_{t+1} = i_{t+1}^* = i_{t-1}^* + \varepsilon_t$. Finally, assume that the Fed is forward-looking and takes time $t + 1$ into account when it picks its partial adjustment speed at time t. Specifically, it has the following loss function:

$$L_t = \left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2 + \left(i_{t+1}^* - i_{t+1}\right)^2 + \theta\left(\Delta i_{t+1}^\infty\right)^2. \tag{18}$$

In the Appendix, we prove the following proposition, which states that both gradualism and the time-consistency problem remain in this setting.

PROPOSITION 3: *With a forward-looking objective function in the static model with noise, both partial adjustment and the time-consistency problem remain: $k < 1$ so long as $\theta > 0$, and it is ex ante optimal to appoint a central banker with $\theta_c = 0$.*

The intuition for the proposition is as follows. On the one hand, taking the market's conjectures about its behavior as given, when the Fed tries to reduce the size of a move in bond prices at time t by adjusting the funds rate gradually, it recognizes that, if it is successful, this will lead to a larger move at time $t + 1$—because eventually its private information must come out. On the other hand, because its loss function is convex in the size of price moves in each period, distributing the price change over time is still attractive. A motive to act gradually therefore remains.[16]

An interesting feature of this variation is that the time-consistency problem is starker, in the sense that, even with nonzero rate-setting noise, it is now always optimal to appoint a central banker who does not care at all about the bond market, that is, who has $\theta_c = 0$. With a myopic loss function, a Fed that acts gradually in the presence of rate-setting noise partially succeeds in reducing the size of this period's move in the infinite-horizon forward rate.

---

[16] In a more fully dynamic setup, the Fed might act as if private information suppressed at time t would not come out all at once at $t+1$, but would make its way into prices more slowly over several subsequent periods. This case actually behaves more like the baseline one with the myopic objective function, again due to the convex nature of the loss function—the more future periods the remainder of a shock can be spread over, the less the volatility incurred in future periods matters to the Fed. So arguably, the objective functions in equations (2) and (18) span the full range of possibilities.

Table II

Characterization of the Static Model with Noise When the Fed Has a Forward-Looking Objective

This table characterizes the equilibrium in the static model with noise when the Fed has a forward-looking objective for different parameter values. θ indexes the Fed's aversion to volatility, τ_u/τ_ε is the ratio of the variance of innovations to the target rate to the variance of rate-setting noise, $k_{forward}$ is the equilibrium degree of adjustment without commitment, θ_c is the volatility aversion of the optimal central banker under commitment, L is the expected value of the Fed's loss function without commitment, and L_c is the value of the loss function under commitment.

θ	τ_u/τ_ε	$k_{forward}$	θ_c	$(L - L_c)/L$
0.2	10	0.84	0	0.08
1.0	10	0.46	0	0.21
1.0	250	0.07	0	0.46

Given this partial success, a social planner does not want to totally eliminate the gradualism motive, and hence leaves θ_c slightly positive. By contrast, with a forward-looking objective function, the same rate-setting noise that tempers the time-t price move means that there is more left to be revealed at time $t + 1$. So when one takes both periods into account, efforts at gradualism once again look totally fruitless, as they do in the no-noise case with a myopic objective function. Therefore, it is now optimal to set $\theta_c = 0$, independent of the level of rate-setting noise.

Table II describes outcomes in this version of the model for the same parameter values used in Table I. As expected, the Fed moves less gradually with a forward-looking objective than it does with a myopic one, though the differences in speed of adjustment are generally quite modest. Moreover, for the reasons described just above, the gains from appointing a central banker with the optimal θ_c (which is now always zero) can be larger with a forward-looking objective. For instance, when $\theta = 1.0$ and $\sigma_\varepsilon^2/\sigma_u^2 = 250$, the gain from commitment is 11% with the myopic objective and 46% with the forward-looking objective.

D.2. The Fed Cares about the Volatility of Finite-Horizon Rates

The time-consistency problem is especially pronounced in our baseline setting because the Fed is assumed to care about the volatility of the infinite-horizon rate. In this case, the only reason for the Fed to adjust partially to new private information is to manipulate the market's inference about that information. To the extent that the market undoes that manipulation in the rational expectations equilibrium, the Fed's efforts are unsuccessful—hence the time-consistency problem.

By contrast, if we instead assume that the Fed cares about the volatility of a finite-horizon rate (e.g., the 10-year rate), there is still a time-consistency problem, but it is attenuated. If the expectations hypothesis holds, then the finite-horizon rate is given by the expected path of the funds rate over that

finite horizon. The finite-horizon rate therefore responds to both information about the Fed's ultimate destination and the particular path that the Fed chooses to get there. By moving gradually, the Fed can succeed in reducing the volatility of the realized path. However, its attempts to manipulate the market's inference about its ultimate destination will still be partially undone in equilibrium, leaving some element of time-inconsistency.

To develop this intuition, we observe that, in our one-factor model of the term structure, any finite-horizon rate can be approximated by a weighted average of the current funds rate i_t and the infinite-horizon forward rate i_t^∞, as long as we pick the weights correctly. Using this approximation, we establish the following proposition, which is proven in the Appendix.

PROPOSITION 4: *Suppose that $\theta > 0$ and that the Fed cares about the volatility of a finite-horizon rate, approximated as a weighted average of the short rate and the infinite-horizon rate. Then it is ex ante optimal to appoint a central banker with θ_c being an increasing function of the weight on the short rate.*

Under discretion, the Fed has two motives when it moves gradually. First, it aims to reduce the volatility of the component of the finite-horizon rate that is related to the current short rate i_t. Second, it hopes to reduce the volatility of the component of the finite-horizon rate that is related to the infinite-horizon forward rate i_t^∞ by fooling the market about its private information ε_t. The first goal can be successfully attained in rational expectations equilibrium, but the second will be partially undone, as in the baseline model.

Under commitment, the second motive is weakened, and thus it is less appealing to move gradually than it would be under discretion. Thus, if society is appointing a central banker whose concern about financial market volatility is given by θ_c, it would like to appoint one with $\theta_c < \theta$. In other words, there is still a time-consistency problem. However, it is not as extreme as the case in which the social planner cares about the volatility of the infinite-horizon forward rate.

Table III demonstrates this intuition numerically. The table replicates Table I, but now considers different finite-horizon rates, which put weight α on the funds rate and weight $(1 - \alpha)$ on the infinite-horizon rate. In Panel A of Table III we set $\alpha = 0.15$, and in Panel B we set $\alpha = 0.30$. The table shows that as α increases, it is optimal to appoint a central banker with a higher value of θ_c. Still, there are significant improvements in the Fed's loss function under commitment.

We have also tried combining the assumption that the Fed cares about a finite-horizon rate with the assumption that is has a forward-looking objective function, as in Table II. The results (not tabulated) are as one would expect based on the two mechanisms operating in isolation. That is, there are no particularly surprising interaction effects. Relative to Table III, making the objective function forward-looking leads to somewhat less gradualism, but also

Table III
Characterization of the Static Model with Noise When the Fed Cares About a Finite-Horizon Rate

This table characterizes the equilibrium in the static model with noise when the Fed cares about a finite-horizon rate for different parameter values. θ indexes the Fed's aversion to volatility, τ_u/τ_ε is the ratio of the variance of innovations to the target rate to the variance of rate-setting noise, α is the weight put on the short rate and $(1 - \alpha)$ is the weight on the infinite-horizon rate, k is the equilibrium degree of adjustment without commitment, θ_c is the volatility aversion of the optimal central banker under commitment, L is the expected value of the Fed's loss function without commitment, and L_c is the value of the loss function under commitment.

Panel A: $\alpha = 0.15$				
θ	τ_u/τ_ε	k	θ_c	$(L - L_c)/L$
0.2	10	0.82	0.05	0.06
1.0	10	0.31	0.30	0.03
1.0	250	0.03	0.14	0.11

Panel B: $\alpha = 0.30$				
θ	τ_u/τ_ε	k	θ_c	$(L - L_c)/L$
0.2	10	0.82	0.08	0.04
1.0	10	0.34	0.46	0.03
1.0	250	0.04	0.24	0.14

tends to increase the welfare gains from commitment, particularly when rate-setting noise is relatively low.

E. Additional Shocks

Next, we enrich the model by adding additional shocks beyond innovations to the Fed's private information about its target rate. For simplicity, we work with the model without rate-setting noise.

E.1. Public Information about the Fed's Target

Thus far, we have assumed that the Fed's target value of the short rate is entirely private information. A more empirically realistic assumption is that the target depends on both public and private information, where the former might include current values of inflation, the unemployment rate, and other macroeconomic variables. It turns out that our basic results carry over to this setting. That is, the Fed adjusts gradually to *both* publicly and privately observed innovations to its target.[17] To see this, suppose that the target rate follows the process

$$i_t^* = i_{t-1}^* + \varepsilon_t + \nu_t. \tag{19}$$

---

[17] We thank David Romer for pointing out this generalization of the model to us.

As before, ε_t is the Fed's private information. However, ν_t is publicly observed by both the Fed and the bond market.[18] Given the more complicated nature of the setting, the Fed's optimal choice regarding how much to adjust the short rate no longer depends only on its private information ε_t. It also depends on the public information ν_t. To allow for a general treatment, we posit that this adjustment can be described by a potentially nonlinear function of the two variables and that there is no noise in the Fed's adjustment rule ($\sigma_u^2 = 0$):

$$i_t = i_{t-1} + f\left(\varepsilon_t; \nu_t\right). \tag{20}$$

Moreover, we assume that the market conjectures that the Fed is following a potentially nonlinear rule:

$$i_t = i_{t-1} + \phi\left(\varepsilon_t; \nu_t\right). \tag{21}$$

In the Appendix, we use a calculus-of-variations type of argument to establish that, in a rational expectations equilibrium in which $f(\cdot; \cdot) = \phi(\cdot; \cdot)$, the Fed's adjustment rule is given by

$$i_t = i_{t-1} + k_\varepsilon \varepsilon_t + k_\nu \nu_t, \tag{22}$$

where

$$k_\nu = k_\varepsilon \quad \text{and} \quad k_\varepsilon = k_\varepsilon^2 + \theta. \tag{23}$$

The following proposition summarizes the key properties of the equilibrium.

PROPOSITION 5: *The Fed responds as gradually to public information about changes in the target rate as it does to private information. This is true regardless of the relative contributions of public and private information to the total variance of the target. As the Fed's concern about bond market volatility θ increases, both k_ε and k_ν fall.*

Proposition 5 is striking and may at first glance seem counterintuitive. Given our previous results, one might think that there is no reason for the Fed to move gradually with respect to public information. However, while this is correct in the limit case where there is no private information whatsoever, it turns out to be wrong as soon as we introduce a small amount of private information. The logic is as follows. Suppose there is a piece of public information that suggests that the funds rate should rise by 100 bps, for example, there is a sharp increase in the inflation rate. This news, if released on its own, would tend to also create a spike in long-term bond yields. In an effort to mitigate this spike, the Fed is tempted to show a more dovish hand than it had previously,

---

[18] We assume that the Fed is equally averse to volatility in the infinite-horizon rate induced by either ε_t or ν_t. For example, suppose that the Fed's aversion to volatility is rooted in the recognition that (i) sharp bond market movements can affect the solvency of financial institutions and (ii) distressed financial institutions can adversely affect the real economy. If this is the case, the Fed will want to spread volatility out over time, regardless of its source.

that is, to act as if it has simultaneously received a negative innovation to the privately observed component of its target. To do so, it raises the funds rate by less than it otherwise would. In an out-of-equilibrium sense, this is an attempt to convey that it has offsetting private information.

As before in the no-noise case, this effort to fool the market is not successful in equilibrium, but the Fed cannot resist the temptation to try. And as long as there is just a small amount of private information, the temptation always exists because, at the margin, the existence of private information leads the Fed to act as if it can manipulate market beliefs. Hence, even if private information does not represent a large fraction of the total variance of the target, the degree of underadjustment predicted by the model is the same as in an all-private-information setting.

E.2. Term-Premium Shocks

We next enrich the model in another direction, to consider how the Fed behaves when financial market conditions are not purely a function of the expected path of interest rates. Here we relax the assumption that the expectations hypothesis holds and instead assume that the infinite-horizon forward rate consists of both the expected future short rate and an exogenous term-premium component r_t:

$$i_t^\infty = E_t\left[i_t^*\right] + r_t. \tag{24}$$

The term premium is assumed to be public information, observed simultaneously by market participants and the Fed. We allow the term premium to follow an arbitrary process and let η_t denote the innovation in the term premium:

$$\eta_t = r_t - E_{t-1}\left[r_t\right]. \tag{25}$$

The solution method is similar to that in the preceding section. Specifically, we assume that there is only private information about the Fed's target, there is no noise in the Fed's rate-setting rule ($\sigma_u^2 = 0$), and the rule can be an arbitrary nonlinear function of both the new private information ε_t that the Fed learns at time t as well as the term-premium shock η_t, which is publicly observable.

In the Appendix, we show that in equilibrium, the Fed's adjustment rule is given by

$$i_t = i_{t-1} + k_\varepsilon \varepsilon_t + k_\eta \eta_t, \tag{26}$$

where

$$\begin{aligned} k_\varepsilon &= k_\varepsilon^2 + \theta \qquad \text{and} \\ k_\eta &= -\tfrac{\theta}{k_\varepsilon}. \end{aligned} \tag{27}$$

The following proposition summarizes the key properties of the equilibrium.

PROPOSITION 6: *The Fed acts to offset publicly observable term-premium shocks, lowering the funds rate when the term-premium shock is positive and raising*

it when the term-premium shock is negative. As the Fed's concern with bond market volatility θ increases, k_ε falls and k_η increases in absolute magnitude. Thus, when it cares more about the bond market, the Fed reacts more gradually to changes in its private information about its target rate, but more aggressively to changes in term premiums.

The intuition here is similar to that for why the Fed underreacts to public information about its target. When the term premium spikes up, the Fed is unhappy about the prospective increase in the volatility of long rates. So even if its private information about i_t^* has not changed, it would like the market to think it has become more dovish. It therefore cuts the short rate to create this impression. Again, in a no-noise equilibrium, this attempt to fool the market is not successful, but taking the market's conjectures at any point in time as fixed, the Fed is always tempted to try.

To see why the equilibrium must involve the Fed reacting to term-premium shocks, think about what happens if we try to sustain an equilibrium in which it does not. That is, consider what happens if we try to sustain an equilibrium in which $k_\eta = 0$. In such a hypothetical equilibrium, when the market sees any movement in the funds rate, it attributes that movement entirely to changes in the Fed's private information ε_t about its target rate. But if this is the case, then the Fed can indeed offset movements in term premiums by changing the short rate, thereby contradicting the assumption that $k_\eta = 0$. Hence, $k_\eta = 0$ cannot be an equilibrium.

A noteworthy feature of the equilibrium is that the absolute magnitude of k_η increases as θ rises and as k_ε decreases: When it cares more about bond market volatility, the Fed's responsiveness to term-premium shocks becomes more aggressive even as its adjustment to new information about its target becomes more gradual. In particular, because we are restricting ourselves to the region of the parameter space over which the simple no-noise model yields a nondegenerate equilibrium for k_ε, this means that we must have $0 < \theta < \frac{1}{4}$ from equation (13) above. As θ moves from the lower to the upper end of this range, k_ε declines monotonically from 1 to ½, and k_η increases in absolute magnitude from 0 to $-\frac{1}{2}$.

This property yields a sharp testable empirical implication. As noted earlier, Campbell, Pflueger, and Viceira (2015) show that the Fed's behavior has become significantly more inertial in recent years. We might be tempted to use the logic of the model to claim that this is a result of the Fed placing increasing weight over time on the bond market, that is, having a higher value of θ than it used to. While the evidence on financial market mentions in the FOMC transcripts that we plot in Figure 1 is loosely consistent with this hypothesis, it is obviously far from being a decisive test. However, with Proposition 6 in hand, if we want to attribute a decline in k_ε to an increase in θ, then we also have to confront the additional prediction that we should observe the Fed responding more forcefully over time to term-premium shocks. That is, the absolute value of k_η must have gone up. If this is not the case, it would represent a rejection of the hypothesis.

III. Dynamic Model

In the static model considered thus far, the phenomenon of gradualism is really nothing more than underreaction of the policy rate to a one-time shock to the Fed's target. This leaves open the important question of dynamic adjustment: If a private-information shock of ε_t is only partially incorporated into the funds rate at time t, how long before it is fully impounded? As we show below, in the absence of any rate-setting noise ($\sigma_u^2 = 0$), the adjustment process is trivially fast —ε_t is fully reflected in the funds rate one period later, by time $t + 1$. In this case, one might question whether the model captures economically meaningful effects, given that the FOMC meets twice per quarter.

Things become more interesting in the case with nonzero rate-setting noise ($\sigma_u^2 > 0$). Here, the dynamic adjustment process is more protracted, and the positive and normative implications of the model correspondingly more substantial. However, this case is technically challenging to analyze in its full generality because, in the presence of noise, the fully rational solution to the market's inference problem becomes extremely complex. Loosely speaking, at each point in time t, in order to estimate the current Fed target i_t^*, the market needs to have a separate running estimate of each of the past innovations in the Fed's private information, ε_{t-j}, each of which it then updates using a separate filtering process over all of the past realizations of the funds rate.

To attack this difficult problem, we proceed in two steps. We first solve a fully rational three-period version of the model. Here, the filtering problem is easy enough to handle, and we can use this setup to show that when there is nonzero rate-setting noise, the dynamic adjustment process is no longer trivial. That is, a private-information shock ε_t can remain substantially underreflected in the funds rate not only at time t, but at time $t + 1$ as well. This suggests, albeit only qualitatively, that the positive and normative implications of the model may be more economically interesting, even when Fed meetings are spaced relatively closely together.

We next turn to an approximate but fully dynamic version of the model with an infinite horizon. This allows us to compare our model's quantitative predictions more directly with the existing empirical evidence on gradualism. To make things tractable, we model the market's inferences about the Fed's target rate using a heuristic, near-rational approximation of the optimal Bayesian filtering process. As can be seen by comparison to the fully rational three-period model, our heuristic arguably allows us to capture the broad spirit of how updating about i_t^* works when there is rate-setting noise, but greatly simplifies the analytics.

A. Three-Period Model

We begin with a fully rational, three-period model with rate-setting noise. Suppose we start at time 0 in steady state with $i_0 = i_0^*$. At time 1, the Fed receives new private information ε_1, and there are no further innovations to the target rate afterward, so $i_1^* = i_2^* = i_0^* + \varepsilon_1$. The Fed follows partial adjustment

rules with noise at both times 1 and 2, and then fully incorporates its private information into the funds rate at time 3. Thus, we have

$$
\begin{aligned}
i_1 &= i_0 + k_1\varepsilon_1 + u_1 \\
i_2 &= i_0 + k_2\varepsilon_1 + u_2 \\
i_3 &= i_0 + \varepsilon_1.
\end{aligned}
\tag{28}
$$

We assume that the noise shocks at times 1 and 2 are independent, and that both u_1 and u_2 have variance σ_u^2. The market conjectures that the Fed follows the adjustment rules

$$
\begin{aligned}
i_1 &= i_0 + \kappa_1\varepsilon_1 + u_1 \\
i_2 &= i_0 + \kappa_2\varepsilon_1 + u_2 \\
i_3 &= i_0 + \varepsilon_1.
\end{aligned}
\tag{29}
$$

Finally, the Fed's loss function is given by

$$
\left(i_1 - i_1^*\right)^2 + \theta\left(\Delta i_1^{\infty}\right)^2 + \left(i_2 - i_2^*\right)^2 + \theta\left(\Delta i_2^{\infty}\right)^2 + \left(i_3 - i_3^*\right)^2 + \theta\left(\Delta i_3^{\infty}\right)^2.
\tag{30}
$$

The Fed picks the optimal k_1 and k_2 with this forward-looking objective, but without the ability to commit to moving at a particular speed. In other words, k_2 is set on a discretionary basis at time 2 and cannot be locked in at time 1. In the Appendix, we show that the rational expectations equilibrium in this version of the model has the following properties.

PROPOSITION 7: *In the three-period model with noise, there is partial adjustment at both times* 1 *and* 2 *:* $k_1 < 1$ *and* $E[k_2] < 1$ *if and only if* $\sigma_u^2 > 0$. *In addition, the time-consistency problem remains, and it is ex ante optimal to appoint a central banker with* $\theta_c = 0$.

As before, we have partial adjustment at time 1. If there is no rate-setting noise, then the underadjustment is short-lived: In the next period, the Fed fully impounds what was left of the time-t innovation into the rate. In other words, $k_2 = 1$ if $\sigma_u^2 = 0$. Intuitively, with no noise, investors have already figured out all of the Fed's time-1 private information by time 2. Given that the Fed can no longer fool investors about this old information, there is no remaining motive for it to continue to incorporate it into rates slowly.[19]

However, if there is noise in the rate-setting process so that $\sigma_u^2 > 0$, then the market still has some residual uncertainty about the Fed's time-1 private information ε_1 at time 2. By moving gradually once again at time 2, the Fed hopes to keep this information from hitting the market all at once in the next go-round.[20] As in the static model with noise, this hope is partially frustrated in

[19] Why is it that the Fed underreacts to public news about the target v_t in Proposition 5 but does not continue to underreact to what is now (by time 2) public information about ε_1 in Proposition 7? The key difference is timing. In Proposition 5, we have two simultaneous innovations to the target rate, so the Fed can try to offset the public news by pretending it has offsetting private news. In Proposition 7, the bond market volatility due to ε_1 has already been realized at time 1. So being more dovish at time 2 does not help reduce volatility at time 2.

[20] The actual value of k_2 depends on the realizations of ε_1 and u_1. For this reason, Proposition 7 characterizes $E[k_2]$.

Table IV
Characterization of the Three-Period Model with Noise

This table characterizes the equilibrium in the three-period model with noise for different parameter values. θ indexes the Fed's aversion to volatility, τ_u/τ_ε is the ratio of the variance of innovations to the target rate to the variance of rate-setting noise, k_1 and k_2 are the equilibrium degrees of adjustment without commitment at times 1 and 2, respectively, θ_c is the volatility aversion of the optimal central banker under commitment, L is the expected value of the Fed's loss function without commitment, and L_c is the value of the loss function under commitment.

θ	τ_u/τ_ε	k_1	k_2	θ_c	$(L - L_c)/L$
0.20	10	0.83	0.99	0	0.07
1.0	10	0.40	0.85	0	0.24
1.0	250	0.05	0.11	0	0.63

rational expectations equilibrium, and the time-consistency problem persists. Indeed, given the forward-looking nature of the objective function, a generalization of Proposition 3 applies, and we can show that it is ex ante optimal to appoint a central banker with $\theta_c = 0$.

Table IV illustrates the outcomes in the three-period model for different parameter values. The first row shows that when $\theta = 0.20$ and $\sigma_\varepsilon^2/\sigma_u^2 = 10$, there is partial adjustment at time 1 but near-full adjustment by time 2. Nonetheless, committing to a central banker with $\theta_c = 0$ results in a 7% improvement in the Fed's loss function. The second row shows that as we increase the Fed's concern about bond market volatility ($\theta = 1.0$), we get slower adjustment at both times 1 and 2, and there is significant underincorporation of the Fed's private information into the funds rate even as of time 2, with $k_2 = 0.85$. Here, committing to a central banker with $\theta_c = 0$ would result in a 23% improvement in the Fed's loss function. Finally, the third row shows that, with $\theta = 1.0$ and relatively little rate-setting noise ($\sigma_\varepsilon^2/\sigma_u^2 = 250$), the Fed can get stuck in a very gradual equilibrium at both times 1 and 2. In this case, k_2 falls all the way to 0.11, so that most of the time-1 private information still has not been impounded into the funds rate as of time 2. With so little noise, the Fed has almost no cover to adjust the funds rate at either date if it does not want to change the market's inference much. Here, committing to a central banker with $\theta_c = 0$ would result in a 63% improvement in the Fed's loss function, so the time-consistency problem is quite severe.

Overall, the results of the three-period model suggest that the key implications of our model survive in a dynamic setting. Partial adjustment can persist over multiple periods, amplifying the welfare impact of the time-consistency problem.

B. Near-Rational Dynamic Model

To more satisfactorily explore the positive properties of the model, we need a more fully articulated dynamic framework. Moreover, if we want to directly compare the model to empirical evidence, we need to consider the case in which

there are both public- and private-information shocks to the Fed's target rate. Accordingly, we assume that

$$i_t^* = i_{t-1}^* + \varepsilon_t + v_t, \tag{31}$$

where ε_t is the Fed's private information and the v_t is public information—for example, news about inflation or unemployment.

Suppose that we enter period t with a preexisting gap between the time $t-1$ target rate and the time $t-1$ federal funds rate of

$$i_{t-1}^* - i_{t-1} \equiv X_{t-1} + Y_{t-1}, \tag{32}$$

where X_{t-1} is the private-information component of the gap and Y_{t-1} is the public-information component. Even in the presence of rate-setting noise, Y_{t-1} is known by the market with certainty. In contrast, the market must infer X_{t-1} from the Fed's previous actions, and with rate-setting noise may still have an imperfect estimate as of time t. As mentioned above, modeling the fully rational Bayesian filtering problem for estimating X_{t-1} is extremely complicated in this setting. Therefore, to maintain tractability, we adopt a simpler "near-rational" representation of the effects of rate-setting noise and assume that just after time $t-1$, but before time t, investors observe a noisy signal of X_{t-1}:

$$s_{t-1} = X_{t-1} + z_{t-1}, \tag{33}$$

where $z_t \sim N(0, \sigma_z^2)$.

The market's expectation of X_{t-1} given s_{t-1} is therefore impounded into the infinite-horizon forward rate before the Fed sets the funds rate at time t. Thus, the change in the infinite-horizon forward rate at time t is the revision in the market's expectations about the target rate given the change in the federal funds rate:

$$
\begin{aligned}
\Delta i_t^\infty &= E\left[i_t^* | s_{t-1}, Y_{t-1}, \Delta i_t, v_t\right] - E\left[i_{t-1}^* | s_{t-1}, Y_{t-1}\right] \\
&= E\left[X_{t-1} + Y_{t-1} + \varepsilon_t + v_t | s_{t-1}, Y_{t-1}, \Delta i_t, v_t\right] - E\left[X_{t-1} + Y_{t-1} | s_{t-1}, Y_{t-1}\right].
\end{aligned}
\tag{34}
$$

We assume that the Fed follows the most general possible rule, according to which it can adjust to each of the four components of the target, X_{t-1}, Y_{t-1}, ε_t, and v_t, at different speeds. We further assume that the Fed conditions on X_{t-1} but not on the market's signal s_{t-1}, which it cannot directly observe.[21]

---

[21] If there are no term-premium shocks, the Fed could in principle infer s_{t-1} by looking at the long rate. However, if we were to add term-premium shocks that are not directly observable, these would complicate the inference problem and make it impossible to uniquely invert s_{t-1} from the long-term rate. Recall that, in Section III.E.2, we assumed that term-premium shocks are publicly observed. However, in the setting of that section, it would have been equivalent to assume that term-premium shocks cannot be seen directly but rather are perfectly inferred from observation of the long rate, given knowledge of the expected future short rate. In other words, there need be no

<div align="center">

Table V

Characterization of the Near-Rational Dynamic Model with Noise

</div>

This table characterizes the equilibrium in the near-rational dynamic model with noise for different parameter values. θ indexes the Fed's aversion to volatility, τ_z/τ_ε is the ratio of the variance of private-information innovations to the target rate to the variance of rate-setting noise, τ_v/τ_ε is the ratio of the variance of private-information innovations to the target rate to the variance of public-information innovations, $k_\varepsilon = k_v$ is the equilibrium degree of adjustment to new information, both public and private, k_X is the degree of adjustment to the existing private-information gap, and k_Y is the degree of the adjustment to the existing public-information gap. b_i and b_V are coefficients from the quarterly regression $i_t = a + b_i i_{t-1} + b_V V_t + e_t$, where $V_t = \sum v_t$ is the "public information target rate"—the sum of all of the public-information innovations to the Fed's target up to time t. $\rho(\Delta i_t, \Delta i_{t-1})$ and $\rho(\Delta i_t, \Delta i_{t-2})$ are the first and second autocorrelations of quarterly changes in the fed funds rate. The regressions and autocorrelations are calculated in simulated data, where each simulation is 100 quarters with two FOMC meetings per quarter. We run 500 simulations and report the average results across simulations.

θ	τ_z/τ_ε	τ_v/τ_ε	$k_\varepsilon = k_v$	k_X	k_Y	b_i	b_V	$\rho(\Delta i_t, \Delta i_{t-1})$	$\rho(\Delta i_t, \Delta i_{t-2})$
0.2	10	1	0.73	0.83	1	0.85	0.21	0.12	−0.01
0.2	10	10	0.73	0.83	1	0.94	0.14	0.14	0.01
1.0	10	1	0.04	0.41	1	0.87	0.18	0.15	0.04
1.0	10	10	0.04	0.41	1	0.96	0.11	0.39	0.12

With these assumptions in place, we are able to establish the following proposition.[22]

PROPOSITION 8: *In the near-rational dynamic model with noise, the Fed partially impounds X_{t-1}, ε_t, and v_t into the funds rate at time t. By contrast, it fully impounds Y_{t-1} at time t.*

The proposition essentially combines the intuitions from several of our previous results. As in Proposition 5, the Fed moves gradually with respect to contemporaneous public- and private-information innovations to its target, ε_t and v_t. As in Proposition 7, the Fed also moves gradually with respect to lagged private information where, thanks to the rate-setting noise, the market still retains some uncertainty (X_{t-1}). However, it fully impounds lagged public information that the market observes perfectly (Y_{t-1}).

Table V illustrates the behavior of the near-rational dynamic model for different parameter values. Each row of the table displays the speed of adjustment to new information, both public and private, as well as to the existing private-information gap X_{t-1} and the existing public-information gap Y_{t-1}. In addition, the table characterizes the dynamic properties of the funds rate in simulated data. To conduct the simulations, we assume that there are two FOMC meetings per quarter. Then, for each FOMC meeting, we draw random realizations

---

logical tension between our approach in Section III.E.2 and our approach here of taking s_{t-1} to be unobservable by the Fed.

[22] As we discuss further in the Appendix, to simplify the problem, we also assume that at any time t, the Fed picks its adjustment rule taking X_{t-1} and Y_{t-1} as given, but before knowing the realizations of ε_t and v_t. This timing convention is purely a technical trick that makes the problem more tractable without changing anything of economic substance.

1044 *The Journal of Finance*®

of ε_t and v_t and compute the resulting funds rate path using all of the speed-of-adjustment coefficients generated by the model.

We next use the simulated data to conduct two exercises. First, we estimate an inertial Taylor rule. Specifically, we run the following quarterly regression on the simulated data:

$$i_t = a + b_i i_{t-1} + b_V V_t + e_t, \tag{35}$$

where $V_t = \sum v_t$ is the "public information target rate," that is, the sum of all of the public information innovations to the Fed's target up to time t. We think of this variable as the analog of the target rate that is used in the empirical literature, which is typically based on the output gap and inflation. In this literature, a typical finding in quarterly data is that $i_t = 0.85 \times i_{t-1} + 0.15 \times V_t$ (Coibion and Gorodnichenko (2012)). Second, we estimate the first and second quarterly autocorrelations of changes in the funds rate in the simulated data, $\rho(\Delta i_t, \Delta i_{t-1})$ and $\rho(\Delta i_t, \Delta i_{t-2})$. In our data, we find that the first and second quarterly autocorrelations are both around 0.30. Each simulation covers 100 quarters with two FOMC meetings per quarter. In other words, each simulation has 200 periods. For each set of parameter values in the table, we run 500 simulations and report the average moments across these simulations.

In the first row of Table V, the Fed cares modestly about bond market volatility ($\theta = 0.20$), and public and private shocks to its target have equal variance ($\sigma_\varepsilon^2/\sigma_v^2 = 1$). There is partial adjustment to new public and private information ($k_\varepsilon = k_v = 0.73$), slightly more adjustment to the existing private-information gap X_{t-1} ($k_X = 0.83$), and full adjustment to the public-information gap Y_{t-1} ($k_Y = 1$). The table also shows that we do well in terms of replicating the inertial Taylor rule regression, with an inertia coefficient b_i of 0.85, very close to that reported by Coibion and Gorodnichenko (2012). This is true despite the fact that there is full adjustment to the existing public information gap Y_{t-1} ($k_Y = 1$). The reason is that both the lagged funds rate and the public data V_t are informative about the public-information part of the Fed's target, but only the lagged funds rate is informative about the private-information part. Thus, if there is any private information, the only way the regression can load on it is by loading on the lagged funds rate. And given that the lagged funds rate is also informative about the public-information part of the Fed's target, there is little reason for the regression to incrementally load on V_t.

The first row of Table V also shows that we generate a positive first autocorrelation of 0.12 for quarterly changes in the funds rate with these parameter values, but essentially a zero second autocorrelation. Because there is partial adjustment to both public- and private-information shocks, both contribute to the first autocorrelation. The parts of ε_t and v_t that are not impounded into the funds rate at time t will continue to be impounded at time $t+1$. However, only private information contributes to the second autocorrelation, since $k_Y = 1$ and therefore the public shock v_t is fully incorporated after time $t+1$. In contrast, because $k_X < 1$, the private shock ε_t continues to contribute to changes in the fed funds rate at times $t+2$ and beyond. In this case, however, the contribution

is relatively small, and therefore the second autocorrelation is low. There are two reasons for this. First, there is a lot of public information, which contributes to the variance but not to the second autocorrelation. Second, the relatively low value of θ means that even private information is impounded into the funds rate fairly quickly.

The remaining rows of Table V show that we need both a lot of private information relative to public information and a high value of θ to meaningfully increase the second autocorrelation of the funds rate in our simulated data. In the second row, we increase the variance of private information relative to public information ($\sigma_\varepsilon^2/\sigma_v^2 = 10$) but maintain $\theta = 0.20$. In the third row, we instead keep equal the variance of public and private information ($\sigma_\varepsilon^2/\sigma_v^2 = 1$) but set $\theta = 1.0$. As with the first set of parameters above, in these rows we do relatively well in terms of replicating the inertial Taylor rule, but we do not obtain significant second autocorrelations. By contrast, the final row of the table shows that when we combine the assumptions that $\theta = 1.0$ and $\sigma_\varepsilon^2/\sigma_v^2 = 10$, we do well on both dimensions. The inertial Taylor rule regressions put a lot of weight on the lagged funds rate, with $b_i = 0.96$, and both the first and second autocorrelations of changes in the funds rate are noticeably higher, at 0.39 and 0.12, respectively.

In sum, our limited experimentation with the near-rational dynamic model suggests that it can help explain an economically meaningful portion of the inertia in the funds rate that is observed in the data. At the same time, we reemphasize a point made earlier: We do not mean to suggest that *all* of the gradualism that we see in the data is driven by the mechanisms in our model. Some of it is almost certainly a product of other considerations, such as Brainard's (1967) instrument-uncertainty principle. Thus, even if we could mimic all of the moments of the data by choosing the right set of parameter values for our simulations, we do not think that this would be a particularly informative exercise.

IV. Discussion

Although we have focused most of our attention on the model's implications for gradualism, our basic framework can be used to think about a number of related issues in monetary policy communication. Here we give two brief examples. The first makes the point that private information, combined with an aversion to bond market volatility, can lead to other distortions besides gradualism. In particular, the Fed can find itself led by the market to be overly reactive to recent news releases. The second example explores an approach to setting expectations about the path of policy that might be helpful in addressing such distortions.

A. Excess Sensitivity to Recent Data

The Fed often stresses that its policy decisions are "data-dependent." While this is certainly a desirable principle in the abstract, one might wonder whether

in practice there is a risk that policy may become *too sensitive* to the most salient recent data releases, especially when these releases on their own have relatively little incremental information content.

Consider the following hypothetical example. There is an upcoming FOMC meeting, and the Fed faces an important binary decision, for instance, whether to taper the pace of its asset purchases for the first time. The Friday before the meeting, the monthly payroll report will be released. Market participants anticipate that the policy decision will be data-dependent, and they naturally focus attention on the payroll report, even though it is common knowledge that this report only contains a small amount of new information about the underlying state of the economy. Suppose that investors develop an initially arbitrary conjecture that the Fed will taper if and only if the payroll number exceeds 200,000 new jobs. There is a sense in which this conjecture can be partially self-validating.

To see why, imagine that the payroll number turns out to be only 160,000. Given their conjecture, investors sharply revise downward their assessment of the likelihood of the Fed tapering. This means that if the Fed was indeed to follow through with the taper, it would be seen as a hawkish surprise. The market would infer that the innovation in the Fed's private information about its target was unusually high, and bond yields would spike upward accordingly. Given an aversion to such volatility, the Fed might decide to sit tight, thereby validating the market's initial conjecture. This could happen even if the Fed fully recognizes that there is little meaningful difference between a payroll number of 200,000 and 160,000.

As a rough sketch of how one might capture this intuition more formally, assume that the Fed can either taper (Y) or not taper (N). The Fed has private information ε about its long-run target for the funds rate, and this private information is distributed symmetrically on some interval $[-H, H]$. Initially, before the arrival of the news release, the optimal decision if the Fed ignored bond market volatility is to taper if $\varepsilon > 0$. Thus, the prior probability of a Y decision is ½. If the market holds such priors, then no matter how concerned the Fed is with bond market volatility, it will follow through with the taper if and only if $\varepsilon > 0$. This is because a Y decision and an N decision are seen as equally likely ex ante, and so the amount of information about ε that is conveyed is the same either way. The Fed cannot affect market volatility by choosing one option or the other.

Now suppose instead that there is a small bit of bad public news right before the meeting, for example, a disappointing payroll number. Further suppose that the market interprets this news as being both quite informative and pessimistic, so the market now believes that the Fed will only taper if $\varepsilon > X$ for some positive X. In other words, the market thinks the likelihood of a Y decision has declined. Given this revision in market beliefs, a concern with bond market volatility will now make the Fed less inclined to taper, even if it believes that the news release was totally uninformative. The reason is that if the market thinks that a Y decision implies $\varepsilon > X$, then a Y decision conveys more information about the Fed's long-run target, and hence creates more

movement in long-term rates, than does an N decision. Hence, the market's belief that the news release influences Fed behavior is at least partially fulfilled in equilibrium.[23]

B. *The Usefulness of Default Presumptions*

The fundamental tension in our model arises because the Fed's decisions convey its private information and, all else equal, the Fed would prefer that the market not update as strongly about this private information. So it is natural to think about ways to coarsen the mapping from the Fed's private information to its actions. For example, taking the model literally, if one could somehow exogenously increase the rate-setting noise σ_u^2, this would reduce the equilibrium degree of gradualism, as can be seen from a comparison of Panels B and C of Figure 2, and might improve overall efficiency. The intuition is that when there is more rate-setting noise, observed changes in the funds rate are less informative about changes in the Fed's long-run target, and hence have less impact on long-term rates. Thus, the Fed faces less pressure to smooth the funds rate.

A real-world example of such a coarsening effect comes from the period subsequent to the Fed's initial taper of asset purchases in December of 2013. The first \$10 billion reduction in the pace of purchases was viewed as a significant change in the stance of policy, and the run-up to it was associated with a great deal of market volatility. Things changed dramatically, however, in the meetings following the first reduction. The Fed continued to cut asset purchases by \$10 billion more at each meeting, and it quickly became the consensus market view that this mechanically decreasing pattern was the Fed's default behavior.[24] Given this belief on the part of the market, it was easy for the Fed to follow through with the remaining \$10 billion cuts, as the informational content had been largely stripped from them.

To be a bit more precise, a default in our setting can be thought of like an (S, s) policy, a range over which the Fed takes the same action independent of local variation in its private information. As with rate-setting noise, the benefit is a coarsening of the mapping from private information to actions. A given action becomes less freighted with informational content, and hence easier for the Fed to take.

Generalizing from this example, and from the (S, s) intuition, it is unclear that the Fed should always aim to respond sensitively to small changes in public or private information. Of course, policy must necessarily be data-dependent

---

[23] We are not claiming that this is a full rational expectations equilibrium, in the sense that the market's conjecture X about the revision in the Fed's threshold is matched exactly one-for-one by a change in the Fed's actual threshold. Rather, we are only making the weaker statement that a change in the former induces a positively correlated change in the latter.

[24] This occurred despite the fact that the Fed explicitly vowed to remain data-dependent. Each FOMC statement during this period contained a line saying that "...asset purchases are not on a preset course, and the Committee's decisions about their pace will remain contingent on the Committee's outlook for the labor market and inflation...."

in some global sense—substantial changes in the economic environment must be met with commensurate changes in the stance of policy. At the same time, however, it may be better to foster a market expectation that the default policy path will be followed so long as things do not change too much. For example, in the midst of a tightening cycle, there may be some virtue to establishing a default presumption that the funds rate will be mechanically raised by 25 bps at every other FOMC meeting, so long as the economy continues to perform broadly as expected. At the margin, a default of this sort sacrifices some fine-tuning of policy to economic conditions, but has the advantage of making each move in rates less of an incremental informational event.

Admittedly, setting defaults in this manner may be complicated in practice. In the model, changes in the funds rate are the only thing that reveals the Fed's private information to the market. In reality, FOMC statements, speeches, and testimony also transmit such information. One might hypothesize that something of a substitution effect could come into play—to the extent that the Fed makes changes in the funds rate less informative, market participants may begin to scrutinize other forms of communication more intensely.[25] Nevertheless, even if crafting such purely verbal communication remains a challenge, it might have less consequence for economic efficiency. Attempting to manage bond market volatility with words is not the same thing as managing it by distorting the short rate away from the ideal target level. Analyzing a richer version of the model in which both the Fed's words and its actions convey some of its private information is an interesting avenue for future work.

V. Conclusion

This paper examines the complicated strategic interplay between the Fed and the bond market that arises when the Fed has some private information about its long-run target for the policy rate and is averse to bond market volatility. From a positive-economics perspective, we argue that a model with these ingredients can help explain the phenomenon of gradualism in monetary policy, and can also shed light on how the Fed responds to other shocks, such as changes in market risk premiums and salient data releases. From a normative perspective, we emphasize that, in our setting, it can be valuable for a central bank to develop an institutional culture such that a concern with bond market volatility does not play an outsized role in policy deliberations. In other words, it can be useful for policy makers to build a reputation for not caring too much about the bond market.

We close by highlighting an important caveat with respect to this reputational argument. As we have framed it, our key point rests on a comparison

---

[25] Casual observation suggests that in the period when policy was stuck at the zero lower bound, so there was nothing to be learned from rate changes, there was enormous focus on the wording of FOMC statements, with small adjustments to the statement often met with strong market reactions.

of equilibrium outcomes with different values of the parameter θ_c, which measures the appointed central banker's concern about financial market volatility. But crucially, an implicit assumption in making this comparison is that, in any given equilibrium, the market has come to fully know the true value of θ_c. We have not addressed the more difficult question of out-of-equilibrium dynamics, that is of how the market learns about θ_c, from the central bank's observed behavior and other forms of communication.

For this reason, our model offers no guidance on the best way to make the transition to a less bond market sensitive equilibrium, in which there is less inertia in the policy rate and the market eventually comes to understand that θ_c has declined. At any point in time, taking market conjectures as fixed, a "cold turkey" approach that involves an unexpectedly sharp adjustment in the path of the policy rate, relative to prevailing expectations of gradualism, is likely to cause significant market volatility and perhaps some collateral costs for the real economy. Nothing in our analysis should be taken as advocating such an approach. To the extent that an institution-building effort of the sort we have in mind is at all useful, it may be something that is better undertaken over a longer horizon, such that market expectations regarding Fed behavior are given more of a chance to adjust.

Appendix: Proofs of Propositions

A. *Appendix Proof of Proposition 2*

Society's ex ante problem is to choose a central banker with concern about market volatility θ_c. In the absence of noise, this central banker will implement the rational expectations equilibrium $k(\theta_c)$ given by equation (12) replacing θ with θ_c, in which case society's ex ante loss function is given by

$$\left((1 - k(\theta_c))^2 + \theta\right)\sigma_\varepsilon^2,$$

which is minimized by setting $\theta_c = 0$ so that $k(\theta_c) = 1$.

In the presence of noise, society's ex ante loss function is given by

$$L = E\left[\left((1-k)\varepsilon_t + u_t\right)^2 + \theta\left(\chi\left(k\varepsilon_t + u_t\right)\right)^2\right] = (1-k)^2\sigma_\varepsilon^2 + \sigma_u^2 + \theta\chi^2\left(k^2\sigma_\varepsilon^2 + \sigma_u^2\right).$$

Differentiating with respect to θ_c and recognizing that χ and k depend on θ_c yields

$$\frac{\partial L}{\partial \theta_c} = 2\left[\left(-1 + k\left(1 + \theta\chi^2\right)\right)\sigma_\varepsilon^2\frac{\partial k}{\partial \theta_c} + \theta\chi\frac{\partial \chi}{\partial k}\frac{\partial k}{\partial \theta_c}\left(k^2\sigma_\varepsilon^2 + \sigma_u^2\right)\right].$$

Evaluating this expression at $\theta_c = \theta$, the first term is zero by the definition of $k(\theta)$. Thus, we have

$$\frac{\partial L}{\partial \theta_c}|_{\theta_c = \theta} = 2\theta \chi \left(k^2 \sigma_\varepsilon^2 + \sigma_u^2\right) \frac{\partial \chi}{\partial k} \frac{\partial k}{\partial \theta_c}.$$

Further, we have

$$\frac{\partial \chi}{\partial k} = -\tau_u \frac{\left(k^2 \tau_u - \tau_\varepsilon\right)}{\left(\tau_\varepsilon + k^2 \tau_u\right)^2}$$

and

$$\frac{\partial k}{\partial \theta_c} = -\left(1 - \frac{2\theta_c k \tau_u^2 \left(\tau_\varepsilon + k^2 \tau_u\right) \left(k^2 \tau_u - \tau_\varepsilon\right)}{\left(\left(\tau_\varepsilon + k^2 \tau_u\right)^2 + \theta_c \left(k\tau_u\right)^2\right)^2}\right)^{-1} \frac{\left(\tau_\varepsilon + k^2 \tau_u\right)^2 \left(k\tau_u\right)^2}{\left(\left(\tau_\varepsilon + k^2 \tau_u\right)^2 + \theta_c \left(k\tau_u\right)^2\right)^2}.$$

Finally, note that when we write k as a function of κ in equation 8, this function is minimized at $\kappa = \sqrt{\tau_\varepsilon/\tau_u}$. This means that the upper stable equilibrium in any rational expectations equilibrium will have $k > \sqrt{\tau_\varepsilon/\tau_u}$. This implies that $k^2\tau_u - \tau_\varepsilon > 0$ so that $\frac{\partial \chi}{\partial k} < 0$, $\frac{\partial k}{\partial \theta_c} < 0$, and $\frac{\partial L}{\partial \theta_c}|_{\theta_c = \theta} > 0$.

B. Appendix Proof of Proposition 3

As before, the market's inference about the Fed's private information at time t is

$$\widetilde{\varepsilon}_t = \chi \left(k\varepsilon_t + u_t\right),$$

where

$$\chi = \frac{\kappa \tau_u}{\tau_\varepsilon + \kappa^2 \tau_u}.$$

Now, however, the Fed understands that there will be a time $t + 1$ surprise

$$\varepsilon_{t+1} - \widetilde{\varepsilon}_t = (1 - k\chi)\varepsilon_t - \chi u_t.$$

Plugging this into the Fed's objective function, we have

$$E\left[\left((1 - k)\varepsilon_t + u_t\right)^2 + \theta \left(\chi \left(k\varepsilon_t + u_t\right)\right)^2 + 0 + \theta \left((1 - k\chi)\varepsilon_t - \chi u_t\right)^2\right]$$

$$= (1 - k)^2 \varepsilon_t^2 + \sigma_u^2 + \theta \chi^2 k^2 \varepsilon_t^2 + \theta \chi^2 \sigma_u^2 + \theta(1 - k\chi)^2 \varepsilon_t^2 + \theta \chi^2 \sigma_u^2.$$

The first-order condition with respect to k is

$$k = \frac{1 + \theta \chi}{1 + 2\theta \chi^2}.$$

Thus, we have $k < 1$ whenever $\theta > 0$ since $\chi > 0$.

Society's ex ante loss function is given by

$$L = (1-k)^2 \sigma_\varepsilon^2 + \sigma_u^2 + \theta \chi^2 k^2 \sigma_\varepsilon^2 + \theta \chi^2 \sigma_u^2 + \theta (1-k\chi)^2 \sigma_\varepsilon^2 + \theta \chi^2 \sigma_u^2.$$

Differentiating with respect to θ_c and recognizing that χ and k depend on θ_c, we have

$$\frac{dL}{d\theta_c} = 2\left[\left(-1 + k + 2\theta \chi \frac{d\chi}{dk} k^2 + 2\theta \chi^2 k - \theta k \frac{d\chi}{dk} - \theta \chi\right)\sigma_\varepsilon^2 + 2\theta \chi \frac{d\chi}{dk} \sigma_u^2\right] \frac{dk}{d\theta_c}.$$

Evaluating this at $\theta_c = 0$, which implies $k = 1$, $\chi = \tau_u/(\tau_\varepsilon + \tau_u)$, and $d\chi/dk = (\tau_\varepsilon - \tau_u)\tau_u/(\tau_\varepsilon + \tau_u)^2$, yields zero. The second-order condition is positive at $\theta_c = 0$, ensuring that this is a minimum.

C. Appendix Proof of Proposition 4

Suppose there is rate-setting noise and the Fed cares about the volatility of a finite-horizon rate given by

$$\Delta i_t^{finite} = \alpha \Delta i_t + (1-\alpha)\Delta i_t^\infty.$$

The Fed continues to have adjustment rule

$$i_t = i_{t-1} + k\varepsilon_t + u_t,$$

and the market conjectures the adjustment rule is

$$i_t = i_{t-1} + \kappa \varepsilon_t + u_t.$$

As in the case with noise in the paper, the market's inference is that

$$\widetilde{\varepsilon}_t = \chi \Delta i_t = \chi (k\varepsilon_t + u_t),$$

where

$$\chi = \frac{\kappa \tau_u}{\tau_\varepsilon + \kappa^2 \tau_u}.$$

Thus, the change in the finite-horizon yield is

$$\Delta i_t^{finite} = \alpha \Delta i_t + (1-\alpha)\widetilde{\varepsilon}_t$$
$$= (\alpha + (1-\alpha)\chi_1) \Delta i_t.$$

The Fed picks k to minimize the loss function

$$E_{u_t}\left[(i_t^* - i_t)^2 + \theta\left(\Delta i_t^{finite}\right)^2\right] = E_{u_t}\left[((1-k)\varepsilon_t + u_t)^2 + \theta\left((\alpha + (1-\alpha)\chi)(k\varepsilon_t + u_t)\right)^2\right]$$
$$= (1-k)^2 \varepsilon_t^2 + \sigma_u^2 + \theta\left(\alpha + (1-\alpha)\chi\right)^2 k^2 \varepsilon_t^2$$
$$+ \theta\left(\alpha + (1-\alpha)\chi\right)^2 \sigma_u^2.$$

Differentiating with respect to k yields

$$k = \frac{1}{1 + \theta \left(\alpha + (1 - \alpha)\chi\right)^2}.$$

Again, we have $k < 1$ so long as $\theta > 0$.

The Fed's ex ante loss function is given by

$$E_t \left[\left(i_t^* - i_t\right)^2 + \theta \left(\Delta i_t^{finite}\right)^2\right] = (1 - k)^2 \sigma_\varepsilon^2 + \sigma_u^2 + \theta \left(\alpha + (1 - \alpha)\chi\right)^2 k^2 \sigma_\varepsilon^2$$

$$+ \theta \left(\alpha + (1 - \alpha)\chi\right)^2 \sigma_u^2.$$

Differentiating with respect to θ_c and recognizing that χ and k depend on θ_c yields

$$\frac{dL}{d\theta_c} = 2 \left[\left(-(1 - k) + \theta k \left(\alpha + (1 - \alpha)\chi\right)^2\right) \sigma_\varepsilon^2 \frac{\partial k}{\partial \theta_c} + \theta \left(\alpha + (1 - \alpha)\chi\right) \left(k^2 \sigma_\varepsilon^2 + \sigma_u^2\right) \frac{\partial \chi}{\partial k} \frac{\partial k}{\partial \theta_c}\right].$$

Evaluating this expression at $\theta_c = \theta$, the first term is zero by the definition of k. Thus, we have

$$\frac{dL}{d\theta_c}|_{\theta_c = \theta} = 2\theta \left(\alpha + (1 - \alpha)\chi\right) \left(k^2 \sigma_\varepsilon^2 + \sigma_u^2\right) \frac{\partial \chi}{\partial k} \frac{\partial k}{\partial \theta_c}.$$

As before, since $\frac{\partial \chi}{\partial k} < 0$ and $\frac{\partial k}{\partial \theta_c} < 0$, we have $\frac{dL}{d\theta_c}|_{\theta_c = \theta} > 0$. Note that $\frac{d^2 L}{d\theta_c d\alpha}|_{\theta_c = \theta}$ is proportional to $1 - \chi$, which is negative. Thus, θ_c is increasing in α.

D. Appendix Proof of Proposition 5

Given its conjecture about the rule the Fed is following, $\phi(\varepsilon_t; \nu_t)$, the market's conjecture about ε_t is

$$\widetilde{\varepsilon}_t = \phi^{-1} \left(f\left(\varepsilon_t; \nu_t\right); \nu_t\right)$$

and the Fed's loss function is

$$(\varepsilon_t + \nu_t - f\left(\varepsilon_t; \nu_t\right))^2 + \theta \left(\phi^{-1}\left(f\left(\varepsilon_t; \nu_t\right); \nu_t\right) + \nu_t\right)^2.$$

Consider the effect on the value of the loss function of a small perturbation in the value of $f(\varepsilon_t; \nu_t)$, df. The effect of this perturbation is zero at the optimal $f(\cdot; \cdot)$, so we have

$$-\left(\varepsilon_t + \nu_t - f\left(\varepsilon_t; \nu_t\right)\right) + \theta \left(\phi^{-1}\left(f\left(\varepsilon_t; \nu_t\right); \nu_t\right) + \nu_t\right) \frac{\partial \phi^{-1}}{\partial i}|_{i = f(\varepsilon_t; \nu_t)} = 0. \qquad (A1)$$

Since $\phi(\phi^{-1}(x)) = x$, we have

$$\frac{\partial \phi^{-1}}{\partial i}|_{i = f(\varepsilon_t; \nu_t)} = \frac{1}{\frac{\partial \phi}{\partial i}|_{\varepsilon = \phi^{-1}(f(\varepsilon_t; \nu_t); \nu_t)}}.$$

Substituting into (A1) gives

$$- (\varepsilon_t + v_t - f(\varepsilon_t; v_t)) + \theta \left(\phi^{-1}(f(\varepsilon_t; v_t); v_t) + v_t \right) \frac{1}{\frac{\partial \phi}{\partial i}|_{\varepsilon = \phi^{-1}(f(\varepsilon_t; v_t); v_t)}} = 0.$$

Imposing rational expectations, we have $\phi = f$, so that this reduces to the differential equation

$$\frac{\partial f}{\partial \varepsilon}|_{\varepsilon = \varepsilon_t} (\varepsilon_t + v_t - f(\varepsilon_t; v_t)) = \theta (\varepsilon_t + v_t),$$

which the optimal $f(\cdot; \cdot)$ must satisfy.

Now conjecture that $f = k_\varepsilon \varepsilon_t + c$. In this case the differential equation reduces to

$$k_\varepsilon (\varepsilon_t + v_t - k_\varepsilon \varepsilon_t - c) = \theta (\varepsilon_t + v_t),$$

or

$$k_\varepsilon (1 - k_\varepsilon) \varepsilon_t + k_\varepsilon (v_t - c) = \theta \varepsilon_t + \theta v_t.$$

Matching coefficients yields

$$k_\varepsilon (1 - k_\varepsilon) = \theta \quad \text{and}$$
$$c = v_t \left(1 - \frac{\theta}{k_\varepsilon} \right).$$

Thus, we can write the optimal $f(\cdot; \cdot)$ as

$$f = k_\varepsilon \varepsilon_t + k_v v_t,$$

where $k_\varepsilon (1 - k_\varepsilon) = \theta$ and $k_v = 1 - \theta / k_\varepsilon$. From the definition of k_ε, we have

$$k_v = 1 - \frac{\theta}{k_\varepsilon}$$
$$= 1 - (1 - k_\varepsilon) = k_\varepsilon.$$

E. Appendix Proof of Proposition 6

Given its conjecture about the rule the Fed is following, $\phi(\varepsilon_t; \eta_t)$, the market's conjecture about ε_t is

$$\widetilde{\varepsilon}_t = \phi^{-1}(f(\varepsilon_t; \eta_t); \eta_t)$$

and the Fed's loss function is

$$(\varepsilon_t - f(\varepsilon_t; \eta_t))^2 + \theta \left(\phi^{-1}(f(\varepsilon_t; \eta_t); \eta_t) + \eta_t \right)^2.$$

Consider the effect on the value of the loss function of a small perturbation in the value of $f(\varepsilon_t; \eta_t)$, df. The effect of this perturbation is zero at the optimal $f(\cdot; \cdot)$, so we have

$$-(\varepsilon_t - f(\varepsilon_t; \eta_t)) + \theta \left(\phi^{-1}(f(\varepsilon_t; \eta_t); \eta_t) + \eta_t\right) \frac{\partial \phi^{-1}}{\partial i}|_{i=f(\varepsilon_t; \eta_t)} = 0. \qquad \text{(A2)}$$

Since $\phi(\phi^{-1}(x)) = x$, we have

$$\frac{\partial \phi^{-1}}{\partial i}|_{i=f(\varepsilon_t; \eta_t)} = \frac{1}{\frac{\partial \phi}{\partial i}|_{\varepsilon=\phi^{-1}(f(\varepsilon_t; \eta_t); \eta_t)}}.$$

Substituting into (A2) gives

$$-(\varepsilon_t - f(\varepsilon_t; \eta_t)) + \theta \left(\phi^{-1}(f(\varepsilon_t; \eta_t); \eta_t) + \eta_t\right) \frac{1}{\frac{\partial \phi}{\partial i}|_{\varepsilon=\phi^{-1}(f(\varepsilon_t; \eta_t); \eta_t)}} = 0.$$

Imposing rational expectations, we have $\phi = f$, so that this reduces to the differential equation

$$\frac{\partial f}{\partial \varepsilon}|_{\varepsilon=\varepsilon_t} (\varepsilon_t - f(\varepsilon_t; \eta_t)) = \theta(\varepsilon_t + \eta_t),$$

which the optimal $f(\cdot; \cdot)$ must satisfy.

Now conjecture that $f = k_\varepsilon \varepsilon_t + c$. In this case the differential equation reduces to

$$k_\varepsilon(\varepsilon_t - k_\varepsilon \varepsilon_t - c) = \theta(\varepsilon_t + \eta_t),$$

or

$$k_\varepsilon(1 - k_\varepsilon)\varepsilon_t - k_\varepsilon c = \theta \varepsilon_t + \theta \eta_t.$$

Matching coefficients yields

$$k_\varepsilon(1 - k_\varepsilon) = \theta \quad \text{and}$$

$$c = -\frac{\theta}{k_\varepsilon} \eta_t.$$

Thus, we can write the optimal $f(\cdot; \cdot)$ as

$$f = k_\varepsilon \varepsilon_t + k_\eta \eta_t,$$

where $k_\varepsilon(1 - k_\varepsilon) = \theta$ and $k_\eta = -\theta/k_\varepsilon$.

F. Appendix Proof of Proposition 7

We start at time 2 and work backwards. Assume that the market's conjecture about ε_1 entering the period is given by

$$\Delta y_1 = \widetilde{\varepsilon_{1,1}} = \chi_{11}(i_1 - i_0) = \chi_{11}k_1\varepsilon_1 + \chi_{11}u_1,$$

where

$$\chi_{11} = \frac{\tau_u \kappa_1}{\tau_\varepsilon + \tau_u \kappa_1^2}.$$

After observing i_2, the market will believe that ε_1 is given by

$$\widetilde{\varepsilon_{2,1}} = \chi_{21}(i_1 - i_0) + \chi_{22}(i_2 - i_0)$$

$$= \chi_{21}(k_1 \varepsilon_1 + u_1) + \chi_{22}(k_2 \varepsilon_1 + u_2),$$

where

$$\chi_{21} = \frac{\tau_u \kappa_1}{\tau_\varepsilon + \tau_u \left(\kappa_1^2 + \kappa_2^2\right)}$$

$$\chi_{22} = \frac{\tau_u \kappa_2}{\tau_\varepsilon + \tau_u \left(\kappa_1^2 + \kappa_2^2\right)}.$$

Thus, the change in the forward rate at time 2 will be

$$\Delta y_2 = \widetilde{\varepsilon_{2,1}} - \widetilde{\varepsilon_{1,1}} = (\chi_{21} - \chi_{11})(k_1 \varepsilon_1 + u_1) + \chi_{22}(k_2 \varepsilon_1 + u_2),$$

and the change in the forward rate at time 3 will be

$$\Delta y_3 = \varepsilon_1 - \widetilde{\varepsilon_{2,1}} = (1 - \chi_{21} k_1 - \chi_{22} k_2)\,\varepsilon_1 - \chi_{21} u_1 - \chi_{22} u_2.$$

The Fed's objective at time 2 is

$$E\left[\left(i_2 - i_2^*\right)^2 + \theta(\Delta y_2)^2 + \left(i_3 - i_3^*\right)^2 + \theta\left(\Delta y_3\right)^2\right] = (1 - k_2)^2 \varepsilon_1^2 + \sigma_u^2$$

$$+ \theta \left(\begin{array}{c} ((\chi_{21} - \chi_{11})\, k_1 + \chi_{22} k_2)^2\, \varepsilon_1^2 + (\chi_{21} - \chi_{11})^2\, u_1^2 + \chi_{22}^2 \sigma_u^2 \\ + 2\,((\chi_{21} - \chi_{11})\, k_1 + \chi_{22} k_2)(\chi_{21} - \chi_{11})\, \varepsilon_1 u_1 \end{array} \right)$$

$$+ \theta \left(\begin{array}{c} (1 - \chi_{21} k_1 - \chi_{22} k_2)^2\, \varepsilon_1^2 + \chi_{21}^2 u_1^2 + \chi_{22}^2 \sigma_u^2 \\ - 2\,(1 - \chi_{21} k_1 - \chi_{22} k_2)\, \chi_{21} \varepsilon_1 u_1 \end{array} \right).$$

The first-order condition with respect to k_2 is

$$k_2 = \frac{1 + \theta \chi_{22}}{1 + 2\theta \chi_{22}^2} + \frac{\theta \chi_{22}(\chi_{11} - 2\chi_{21})}{1 + 2\theta \chi_{22}^2} k_1 + \frac{\theta \chi_{22}(\chi_{11} - 2\chi_{21})}{1 + 2\theta \chi_{22}^2} \frac{u_1}{\varepsilon_1}.$$

Since $\chi_{21} < \chi_{11}$ and $k_1 < 1$, we have $E[k_2] < 1$ when $\sigma_u^2 > 0$ and $\theta > 0$.

We next fold this back to time 1. For simplicity assume that the Fed acts as though k_2 takes its expected value:

$$E[k_2] = \frac{1 + \theta \chi_{22}\left(1 + (\chi_{11} - 2\chi_{21})\, k_1\right)}{1 + 2\theta \chi_{22}^2}.$$

Let

$$\delta = \frac{dk_2}{dk_1} = \frac{\theta \chi_{22}(\chi_{11} - 2\chi_{21})}{1 + 2\theta \chi_{22}^2}.$$

Folding back to time 1, we have the Fed's objective at time 1:

$$E\left[\left(i_1 - i_1^*\right)^2 + \theta\left(\Delta y_1\right)^2 + \left(i_2 - i_2^*\right)^2 + \theta\left(\Delta y_2\right)^2 + \left(i_3 - i_3^*\right)^2 + \theta(\Delta y_3)^2\right]$$

$$= (1-k_1)^2\varepsilon_1^2 + \sigma_u^2 + \theta\chi_{11}^2 k_1^2\varepsilon_1^2 + \theta\chi_{11}^2\sigma_u^2 + (1-k_{12})^2\varepsilon_1^2 + \sigma_u^2$$

$$+ \theta\left((\chi_{21} - \chi_{11})k_1 + \chi_{22}k_{12}\right)^2\varepsilon_1^2 + \theta\left(\chi_{21} - \chi_{11}\right)^2\sigma_u^2 + \theta\chi_{22}^2\sigma_u^2$$

$$+ \theta\left(1 - \chi_{21}k_1 - \chi_{22}k_2\right)^2\varepsilon_1^2 + \theta\chi_{21}^2\sigma_u^2 + \theta\chi_{22}^2\sigma_u^2.$$

The first-order condition with respect to k_1 is

$$k_1 = \frac{1 + \delta + \theta\left(\chi_{21} + \chi_{22}\delta\right)}{1 + \theta\chi_{11}\left(2\chi_{11} - \chi_{21}\right) + \theta\left(2\chi_{21} - \chi_{11}\right)\left(\chi_{21} + \chi_{22}\delta\right)}$$

$$- \frac{\delta + \theta\chi_{22}\left(2\left(\chi_{21} + \chi_{22}\delta\right) - \chi_{11}\right)}{1 + \theta\chi_{11}\left(2\chi_{11} - \chi_{21}\right) + \theta\left(2\chi_{21} - \chi_{11}\right)\left(\chi_{21} + \chi_{22}\delta\right)}k_{12}.$$

The Fed's ex ante loss function is

$$L = E\left[\begin{array}{c} \left((1-k_1)\varepsilon_1 + u_1\right)^2 + \theta\left(\chi_{11}k_1\varepsilon_1 + \chi_{11}u_1\right)^2 + \left((1-k_2)\varepsilon_1 + u_1\right)^2 \\ + \theta\left(((\chi_{21} - \chi_{11})k_1 + \chi_{22}k_2)\varepsilon_1 + (\chi_{21} - \chi_{11})u_1 + \chi_{22}u_2\right)^2 + 0 \\ + \theta\left((1 - \chi_{21}k_1 - \chi_{22}k_2)\varepsilon_1 - \chi_{21}u_1 - \chi_{22}u_2\right)^2 \end{array}\right]$$

$$= (1-k_1)^2\sigma_\varepsilon^2 + \sigma_u^2 + \theta\chi_{11}^2 k_1^2\sigma_\varepsilon^2 + \theta\chi_{11}^2\sigma_u^2 + (1-k_2)^2\sigma_\varepsilon^2 + \sigma_u^2$$

$$+ \theta\left((\chi_{21} - \chi_{11})k_1 + \chi_{22}k_2\right)^2\sigma_\varepsilon^2 + \theta\left(\chi_{21} - \chi_{11}\right)^2\sigma_u^2 + \theta\chi_{22}^2\sigma_u^2$$

$$+ \theta\left(1 - \chi_{21}k_1 - \chi_{22}k_2\right)^2\sigma_\varepsilon^2 + \theta\chi_{21}^2\sigma_u^2 + \theta\chi_{22}^2\sigma_u^2.$$

Differentiating with respect to θ_c and recognizing that $k_1, k_2, \chi_{11}, \chi_{21}$, and χ_{22} all depend on θ_c we have

$$\frac{dL}{d\theta_c} = -2(1-k_1)\frac{dk_1}{d\theta_c}\sigma_\varepsilon^2 + 2\theta\chi_{11}^2 k_1\frac{dk_1}{d\theta_c}\sigma_\varepsilon^2 + 2\theta\chi_{11}k_1^2\frac{d\chi_{11}}{dk_1}\frac{dk_1}{d\theta_c}\sigma_\varepsilon^2$$

$$+ 2\theta\chi_{11}\frac{d\chi_{11}}{dk_1}\frac{dk_1}{d\theta_c}\sigma_u^2 - 2(1-k_2)\frac{dk_2}{d\theta_c}\sigma_\varepsilon^2 + 2\theta\left((\chi_{21} - \chi_{11})k_1 + \chi_{22}k_2\right)$$

$$\times \left(\begin{array}{c} \left(\frac{d\chi_{21}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{21}}{dk_2}\frac{dk_2}{d\theta_c} - \frac{d\chi_{11}}{dk_1}\frac{dk_1}{d\theta_c}\right)k_1 \\ + (\chi_{21} - \chi_{11})\frac{dk_1}{d\theta_c} + \left(\frac{d\chi_{22}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{22}}{dk_2}\frac{dk_2}{d\theta_c}\right)k_2 + \chi_{22}\frac{dk_2}{d\theta_c} \end{array}\right)\sigma_\varepsilon^2$$

$$+ 2\theta\left(\chi_{21} - \chi_{11}\right)\left(\frac{d\chi_{21}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{21}}{dk_2}\frac{dk_2}{d\theta_c} - \frac{d\chi_{11}}{dk_1}\frac{dk_1}{d\theta_c}\right)\sigma_u^2$$

$$+ 2\theta\chi_{22}\left(\frac{d\chi_{22}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{22}}{dk_2}\frac{dk_2}{d\theta_c}\right)\sigma_u^2 - 2\theta\left(1 - \chi_{21}k_1 - \chi_{22}k_2\right)$$

$$\times \left(\left(\frac{d\chi_{21}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{21}}{dk_2}\frac{dk_2}{d\theta_c}\right)k_1 + \chi_{21}\frac{dk_1}{d\theta_c} + \left(\frac{d\chi_{22}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{22}}{dk_2}\frac{dk_2}{d\theta_c}\right)k_2 + \chi_{22}\frac{dk_2}{d\theta_c}\right)\sigma_\varepsilon^2$$

$$+ 2\theta\chi_{21}\left(\frac{d\chi_{21}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{21}}{dk_2}\frac{dk_2}{d\theta_c}\right)\sigma_u^2 + 2\theta\chi_{22}\left(\frac{d\chi_{22}}{dk_1}\frac{dk_1}{d\theta_c} + \frac{d\chi_{22}}{dk_2}\frac{dk_2}{d\theta_c}\right)\sigma_u^2.$$

At $\theta_c = 0$ we have $k_1 = k_2 = 1$. In addition, we have $\chi_{11} = \frac{\tau_u}{\tau_\varepsilon + \tau_u}$, $\chi_{21} = \chi_{22} = \frac{\tau_u}{\tau_\varepsilon + 2\tau_u}$, $\frac{d\chi_{11}}{dk_1} = \frac{\tau_u(\tau_\varepsilon - \tau_u)}{(\tau_\varepsilon + \tau_u)^2}$, $\frac{d\chi_{21}}{dk_1} = \frac{d\chi_{22}}{dk_2} = \frac{\tau_\varepsilon \tau_u}{(\tau_\varepsilon + 2\tau_u)^2}$, and $\frac{d\chi_{21}}{dk_2} = \frac{d\chi_{22}}{dk_1} = \frac{-2\tau_u^2}{(\tau_\varepsilon + 2\tau_u)^2}$. Recognizing this, the derivative simplifies to

$$\frac{dL}{d\theta_c} = \tau_u \tau_\varepsilon \frac{\sigma_u^2 \tau_u - 1}{(\tau_u + \tau_\varepsilon)^3} \frac{dk_1}{d\theta_c} = 0.$$

G. Appendix Proof of Proposition 8

Assume we enter period t with a preexisting gap between the time $t-1$ target and the time $t-1$ rate of

$$i_{t-1}^* - i_{t-1} = X_{t-1} + Y_{t-1},$$

where X_{t-1} is the private-information part of the gap and Y_{t-1} is the public-information part. The innovation in the target rate has both public and private components

$$i_t^* = i_{t-1}^* + \varepsilon_t + v_t,$$

where ε_t is new private information and v_t is new public information. Assume that the Fed follows an adjustment rule of the form

$$i_t = i_{t-1} + k_X X_{t-1} + k_Y Y_{t-1} + f(\varepsilon_t; v_t),$$

and the market conjectures that the Fed is following a rule of the form

$$i_t = i_{t-1} + \kappa_X X_{t-1} + \kappa_Y Y_{t-1} + \phi(\varepsilon_t; v_t).$$

We assume the following timing convention. The Fed picks k_X, k_Y, and $f(\cdot; \cdot)$ knowing X_{t-1} and Y_{t-1}, but not knowing the realizations of ε_t and v_t. This timing convention is purely a technical trick that makes the problem more tractable without really changing anything of economic substance. Without it, the Fed's adjustment rule would turn out to depend on the realization of cross-products like $X_{t-1}\varepsilon_t$. With the timing trick, what matters instead is the expectation of the product, which is zero. In addition, when choosing $f(\cdot; \cdot)$, the Fed still understands it will know ε_t and v_t at the same time— this is important to get partial adjustment to v_t. We should emphasize that, even with this somewhat strained intra meeting timing, the Fed still behaves on a discretionary basis from one meeting to the next. Thus, while it agrees to values of k_X, k_Y, and $f(\cdot; \cdot)$ in the first part of the time-t meeting, it has no ability to bind itself to those values across meetings. Hence, the basic commitment problem remains.

We assume that the market gets a noisy signal of X_{t-1}

$$s_{t-1} = X_{t-1} + z_{t-1},$$

where $z_t \sim N(0, \sigma_z^2)$. The change in the infinite horizon forward rate is then given by

$$\Delta i_t^\infty = E\left[i_t^*|s_{t-1}, \Delta i_t, Y_{t-1}, v_t\right] - E\left[i_{t-1}^*|s_{t-1}, Y_{t-1}\right]$$

$$= E\left[X_{t-1} + Y_{t-1} + \varepsilon_t + v_t|s_{t-1}, \Delta i_t, Y_{t-1}, v_t\right] - E\left[X_{t-1} + Y_{t-1}|s_{t-1}, Y_{t-1}\right].$$

For the second term, we have

$$E\left[X_{t-1} + Y_{t-1}, |s_{t-1}, Y_{t-1}\right] = Y_{t-1} + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_z^2} s_{t-1}.$$

For the first term, we have

$$E\left[X_{t-1} + Y_{t-1} + \varepsilon_t + v_t|s_{t-1}, \Delta i_t, Y_{t-1}, v_t\right] = E\left[X_{t-1} + \varepsilon_t|s_{t-1}, \Delta i_t, Y_{t-1}, v_t\right]$$

$$+ Y_{t-1} + v_t.$$

Using Stein's Lemma, we have

$$E\left[X_{t-1} + \varepsilon_t|s_{t-1}, \Delta i_t, Y_{t-1}, v_t\right] = \frac{1}{E\left[\phi'\right]^2 \sigma_\varepsilon^2 \left(\sigma_X^2 + \sigma_z^2\right) + \kappa_X^2 \sigma_X^2 \sigma_z^2}$$

$$\times \left(\begin{array}{c} \left(E\left[\phi'\right] \sigma_\varepsilon^2 \left(\sigma_X^2 + \sigma_z^2\right) + \kappa_X \sigma_X^2 \sigma_z^2\right) \left(\Delta i_t - \kappa_Y Y_{t-1}\right) \\ + E\left[\phi'\right] \left(E\left[\phi'\right] - \kappa_X\right) \sigma_\varepsilon^2 \sigma_X^2 s_t \end{array} \right).$$

Thus, we have

$$\Delta i_t^\infty = \chi_i \Delta i_t - \chi_i \kappa_Y Y_{t-1} + v_t - \chi_s s_{t-1},$$

where

$$\chi_i = \frac{E\left[\phi'\right] \sigma_\varepsilon^2 \left(\sigma_X^2 + \sigma_z^2\right) + \kappa_X \sigma_X^2 \sigma_z^2}{E\left[\phi'\right]^2 \sigma_\varepsilon^2 \left(\sigma_X^2 + \sigma_z^2\right) + \kappa_X^2 \sigma_X^2 \sigma_z^2}$$

and

$$\chi_s = \frac{E\left[\phi'\right] \left(\kappa_X - E\left[\phi'\right]\right) \sigma_\varepsilon^2 \sigma_X^2}{E\left[\phi'\right]^2 \sigma_\varepsilon^2 \left(\sigma_X^2 + \sigma_z^2\right) + \kappa_X^2 \sigma_X^2 \sigma_z^2} + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_z^2}.$$

The Fed's objective function is then

$$E_t\left[\left(i_t^* - i_t\right)^2 + \theta\left(\Delta i_t^\infty\right)^2\right] = E_t\left[\begin{array}{c} \left(i_{t-1}^* + \varepsilon_t + v_t - i_{t-1} - k_X X_{t-1} - k_Y Y_{t-1} - f\left(\varepsilon_t; v_t\right)\right)^2 \\ + \theta\left(\chi_i \Delta i_t - \chi_i \kappa_Y Y_{t-1} - \chi_s s_{t-1} + v_t\right)^2 \end{array} \right]$$

$$= \left(1 - k_X\right)^2 X_{t-1}^2 + \left(1 - k_Y\right)^2 Y_{t-1}^2 + 2\left(1 - k_X\right)\left(1 - k_Y\right) X_{t-1} Y_{t-1}$$

$$+ \left(\varepsilon_t + v_t - f\left(\varepsilon_t; v_t\right)\right)^2$$

$$+ \theta\left(\begin{array}{c} \left(\chi_i k_X - \chi_s\right)^2 X_{t-1}^2 + \chi_i^2 \left(k_Y - \kappa_Y\right)^2 Y_{t-1}^2 + \left(\chi_i f\left(\varepsilon_t; v_t\right)\right)^2 + \chi_s^2 \sigma_z^2 + \sigma_v^2 \\ + 2\left(\chi_i k_X - \chi_s\right) \chi_i \left(k_Y - \kappa_Y\right) X_{t-1} Y_{t-1} \end{array} \right).$$

Differentiating with respect to k_Y we have

$$-2(1 - k_Y)Y_{t-1}^2 - 2(1 - k_X)X_{t-1}Y_{t-1} + 2\theta \chi_i^2(k_Y - \kappa_Y)Y_{t-1}^2$$
$$+ 2\theta (\chi_i k_X - \chi_s) \chi_i X_{t-1}Y_{t-1} = 0.$$

Differentiating with respect to k_X we have

$$-2(1 - k_X)X_{t-1}^2 - 2(1 - k_Y)X_{t-1}Y_{t-1} + 2\theta (\chi_i k_X - \chi_s) \chi_i X_{t-1}^2$$
$$+ 2\theta \chi_i^2(k_Y - \kappa_Y)X_{t-1}Y_{t-1} = 0.$$

If we have $k_Y = \kappa_Y = 1$ in a rational expectations equilibrium, these two equations reduce to

$$-(1 - k_X) + \theta (\chi_i k_X - \chi_s) \chi_i = 0$$

and

$$-(1 - k_X) + \theta (\chi_i k_X - \chi_s) \chi_i = 0,$$

or

$$k_X = \frac{1 + \theta \chi_i \chi_s}{1 + \theta \chi_i^2}.$$

Now consider a perturbation of f around the optimum. The corresponding first-order condition is

$$f(\varepsilon_t; \nu_t) = \frac{1}{1 + \theta \chi_i^2}\varepsilon_t + \frac{1}{1 + \theta \chi_i^2}\nu_t.$$

Finally, to close the loop, we note that $X_t = (1 - k_X)X_{t-1} + (1 - k_\varepsilon)\varepsilon_t$ implies

$$\sigma_X^2 = \frac{(1 - k_\varepsilon)^2}{1 - (1 - k_X)^2}\sigma_\varepsilon^2.$$

REFERENCES

Barro, Robert J., and David B. Gordon, 1983, Rules, discretion and reputation in a model of monetary policy, *Journal of Monetary Economics* 12, 101–121.

Bernanke, Ben, 2004, Gradualism, Remarks at an economics luncheon co-sponsored by the Federal Reserve Bank of San Francisco and the University of Washington, Seattle, WA. Available at https://www.federalreserve.gov/boarddocs/speeches/2004/200405202/default.htm.

Bernanke, Ben, and Alan Blinder, 1992, The federal funds rate and the channels of monetary transmission, *American Economic Review* 82, 901–921.

Brainard, William, 1967, Uncertainty and the effectiveness of policy, *American Economic Review* 57, 411–425.

Campbell, John Y., Carolin Pflueger, and Luis Viceira, 2015, Monetary policy drivers of bond and equity risks, Working paper, Harvard University.

Cieslak, Anna, and Annette Vissing-Jorgenson, 2017, The economics of the Fed put, Working paper, UC Berkeley.

Coibion, Olivier, and Yuriy Gorodnichenko, 2012, Why are target interest rate changes so persistent? *American Economic Journal: Macroeconomics* 4, 126–162.

Dell'Ariccia, Giovanni, Luc Laeven, and Gustavo Suarez, 2013, Bank leverage and monetary policy's risk-taking channel: Evidence from the United States, Working paper, International Monetary Fund.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2017, The deposits channel of monetary policy, *Quarterly Journal of Economics* 132, 1819–1876.

Farhi, Emmanuel, and Jean Tirole, 2012, Collective moral hazard, maturity mismatch, and systemic bailouts, *American Economic Review* 102, 60–93.

Gertler, Mark, and Peter Karadi, 2015, Monetary policy surprises, credit costs, and economic activity, *American Economic Journal: Macroeconomics* 7, 44–76.

Holmstrom, Bengt, 1999, Managerial incentive problems: A dynamic perspective, *Review of Economic Studies* 66, 169–182.

Jimenez, Gabriel, Steven Ongena, Jose-Luis Peydro, and Jesus Saurina, 2014, Hazardous times for monetary policy: What do 23 million loans say about the impact of monetary policy on credit risk-taking? *Econometrica* 82, 463–505.

Kashyap, Anil K., and Jeremy C. Stein, 2000, What do a million observations on banks say about the transmission of monetary policy? *American Economic Review* 90, 407–428.

Kydland, Finn, and Edward Prescott, 1977, Rules rather than discretion: The inconsistency of optimal plans, *Journal of Political Economy* 85, 473–491.

Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1336.

Rogoff, Kenneth, 1985, The optimal degree of commitment to an intermediate monetary target, *Quarterly Journal of Economics* 100, 1169–1189.

Rudebusch, Glenn D., 2002, Term structure evidence on interest rate smoothing and monetary policy inertia, *Journal of Monetary Economics* 49, 1161–1187.

Rudebusch, Glenn D., 2006, Monetary policy inertia: Fact or fiction? *International Journal of Central Banking* 2, 85–135.

Sack, Brian, 1998, Uncertainty, learning, and gradual monetary policy, FEDS Working Paper Series no. 98-34, Board of Governors of the Federal Reserve System.

Stein, Jeremy C., 1989, Efficient capital markets, inefficient firms: A model of myopic corporate behavior, *Quarterly Journal of Economics* 104, 655–669.

Stein, Jeremy C., 2014, Challenges for monetary policy communication, Remarks at the Money Marketeers of New York University, New York. Available at https://www.federalreserve.gov/newsevents/speech/stein20140506a.htm.

Taylor, John, 1993, Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.

Woodford, Michael, 2003, Optimal interest-rate smoothing, *Review of Economic Studies* 70, 861–886.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Replication Code.