Reputation for Confidence*

Laura Gáti† and Amy Handlan‡

November 11, 2024

Abstract

This paper uses computational text analysis on internal Fed statement drafts to document two novel stylized facts about Federal Reserve communication: a) the Fed drafts a varying number of *alternative statements* before releasing a policy statement; b) the similarity between alternative statements is time-varying. We explore why this is the case in a communication game between the central bank and the private sector where the central bank makes announcements about their noisy forecasts of shocks and sets policy to balance their inflation and output targets. Meanwhile, the public updates beliefs about policy conditional on (1) the bank's announcement and (2) the public's belief about the central bank's ability to accurately forecast, called *reputation for confidence*. We find that the dynamics of reputation is crucial for the model to match these data facts.

Keywords: communication, forward guidance, reputation, cheap talk, text analysis

JEL codes: E52, E58, C49

*We are grateful for the helpful comments and suggestions from Marek Jarociǹski and seminar and conference participants. Also, we thank Frank Chiu, Halleluiah Girum and Tobias Müller for their research assistance. The views expressed herein are our own and do not necessarily reflect those of the ECB or the Eurosystem.

†European Central Bank, Directorate General Research, Sonnemannstraße 20, 60314 Frankfurt am Main, Germany and CEPR. Email: laura_veronika.gati@ecb.europa.eu.

‡Brown University, Department of Economics, 64 Waterman Street, Providence, Rhode Island, USA. Email: amy_handlan@brown.edu.

1 Introduction

Monetary policy announcements attract intense attention from media and investors trying to decipher the underlying meaning of the Federal Reserve's (Fed's) statement. The public's response to the announcement depends on whether they believe the Fed is informed about the economic situation. Meanwhile, the Federal Reserve dedicates time and resources to writing announcements that balance providing detailed information to move the public's expectations about future policy without implying a strict commitment. The tension between effective forward guidance and discretionary policy-making affects the choice of monetary policy announcement language.

We approach this tension first from a data perspective. We leverage internal Fed materials that detail multiple candidate announcements, the so-called *alternative statements*, for the Fed to select from at their policy meeting. These allow us to estimate the Fed's message space at each meeting using frontier text analysis methods, and to document two novel facts about Fed communication. First, the number of alternative statements is time-varying, ranging between two and nine. Second, the similarity between alternative statements is also time-varying, so that the most extreme alternatives sometimes convey a very similar, and at other times a starkly different message.

A stylized way to think about these two features is the following. Sometimes the Fed only contemplates two possible things to say: either "the economy is too hot" or "the economy is too cool," and then depending on realized inflation, output and unemployment, they pick which of the two alternative statements to actually make. Sometimes, however, the Fed drafts many more alternatives. To carry on the example, suppose the Fed drafts the following five alternatives: "the economy is much too hot," "the economy is a little too hot," "the economy is just right", "the economy is a little too cool," and "the economy is much too cool." In this second case, not only are there more alternatives, but they are more granular, allowing the Fed to fine-tune their message. For instance, the Fed can now distinguish between two more specific degrees of the economy overheating. What our two stylized facts say is that in Fed communication, both of these dimensions, the number and similarity of alternatives, vary over time.

In the theoretical model, we propose a mechanism that can explain these facts. We set up a communication game between a central bank and the public and, through a quantitative exercise, connect the model to data with text-based measures of the Fed's realized communication strategy. In the game, the central bank sets inflation and uses communication policy to stabilize inflation and output in an economy characterized by a simple Phillips curve relationship with inflation-target shocks. The central bank only observes noisy signals of the shock, capturing the idea that central banks in practice make monetary policy decisions under uncertainty. We call the precision of the central bank's signal the central bank's *confidence*, so that a more confident central bank forms more accurate forecasts of the shock that hit the economy.

The key novelty of the theoretical setup is that we assume that the public does not observe the central bank's confidence, and instead forms beliefs over confidence from economic outcomes through Bayesian learning. We call the public's beliefs over central bank confidence the central bank's *reputation for confidence*.

To line up the model with the data moments we first documented, we entertain a Perfect Bayesian Equilibrium concept with a twist. We first solve for optimal action rules of both players (inflation expectations for the public, and inflation for the central bank) taking equilibrium communication as given. Then, as a second step, we solve for optimal communication given equilibrium action rules as the communication strategy that obeys incentive compatibility. In other words, we solve for the message space - the set of alternative statements - where the central bank has no incentive to send any other message than the equilibrium message given the optimal action rules and the signal observed by the central bank. We then explore what the equilibrium message space looks like, and what features can align it with Fed communication in the data. We devote particular attention to whether the central bank ever reveals the exact value of the signal they saw, a feature we call *point revelation*.

The analysis yields two striking results. First, assuming that the central bank perceives there to be shocks hitting the economy and has no motive to generate surprise inflation, then point revelation is possible if and only if reputation equals true confidence. What this result is saying is that if the public misperceives how confident the central bank is in their assessment of the economic outlook, then the central bank cannot exactly reveal their

information for fear of the information being misinterpreted. It is therefore optimal in such a case for the central bank to send imprecise messages. To return to the stylized example above, the central bank would in this case prefer to just say that the economy is running too hot instead of specifying how hot exactly they think the economy is.

The second key result is that in order for the model to match a varying number and similarity of alternative statements, reputation for confidence needs to evolve over time. The intuition for this is that reputation governs the public's responsiveness to central bank announcements. If the public knew the central bank's confidence, they would discount what the bank says by how confident the bank is in their assessment. But since they do not know the bank's confidence, they use their belief over true confidence, reputation. As reputation varies over time through the public learning about the central bank's confidence, the public's responsiveness to announcements becomes time-varying, confronting the central bank with a slightly different communication problem every time they are preparing to make an announcement. This generates an equilibrium message space that varies over time in terms of number and similarity of candidate statements.

The paper proceeds as follows. First, in Section 2, we provide a motivation using data on Federal Reserve communication that a time-varying message space size with varying distance between alternative messages is anecdotally a reasonable model of monetary communication. Section 3 lays out the communication game setup. Section 4 analyzes the stage game, and Section 5 the repeated game. **??** connects the model and data for the quantitative analysis. Section 6 concludes.

1.1 Related Literature

Our paper is related to three strands of literature. The first is the literature on cheap talk games in the spirit of Crawford and Sobel (1982), who were the first to show that misalignment in preferences leads a sender to send noisy signals to the receiver in the form of a partitioned message space. Our model borrows from the framework of Moscarini (2007), who in turn builds on Crawford and Sobel (1982) to investigate how the precision of a central bank's information - central bank confidence - affects the equilibrium message space. Our paper is thus related to others that explore various characteristics of the Crawford and Sobel

(1982) cheap talk game (Austen-Smith, 1990; Frankel and Kartik, 2018; Li and Madarász, 2008) or its application to central bank communication (Amador and Weill, 2010; Bassetto, 2019; Cukierman and Meltzer, 1986), as well as more general central bank communication games (Angeletos and Pavan, 2007; Chahrour, 2014; Gáti, 2023; Herbert, 2021; Ko, 2019).

Particularly related to our work are papers on reputation-building about preferences in incomplete information settings (Amador and Phelan, 2021; Benabou and Laroque, 1992; Diamond, 1989, 1991; Kreps and Wilson, 1982; Mailath and Samuelson, 2001; Milgrom and Roberts, 1982; Morris, 2001; Phelan, 2006; Sobel, 1985), and especially about the quality of information or ability to process information (Camous and Matveev, 2024; Guembel and Rossetto, 2009; Ottaviani and Sørensen, 2006a,b). Our setup differs from the latter papers in that our sender (the central bank) does not have an explicit concern for reputation a priori, nor an incentive to trick the public. Coarse communication in equilibrium does not come from preference misalignment or an inbuilt concern for reputation, but from a misperception channel, related to the main ideas in the Bayesian Persuasion literature (Kamenica and Gentzkow, 2011). This result also echoes Iovino et al. (2022)'s finding that when the central bank and the market are imperfectly informed, it is not always optimal to reveal the central bank's private information to the public.

The learning element of our model belongs to the Bayesian parameter learning literature, and comes from Ghofrani (2022) and Baley and Veldkamp (2022). Our work is also related to other papers where there is learning about a policymaker's type, such as King and Lu (2021) or Hansen and McMahon (2016). Thus our work also connects to the classical literature on the conduct of monetary policy when preferences are misaligned (Barro and Gordon, 1983; Clarida et al., 1999; Rogoff, 1985), with the important distinction that misalignment in preferences is not the key driver of our results. Our finding that the central bank communicates coarsely due to an evolving reputation as the market learns about central bank confidence over time corroborates the evidence presented by Kostyshyna and Petersen (2023) that central banks communicate with uncertainty, as well as that by Sastry (2022) that models seeking to rationalize disagreements in the data between markets and the central bank need to feature varying market beliefs about the precision of the bank's information.

Our empirical work relates to a growing literature seeking to measure central bank communication and its effects. In particular, papers that use text analysis to measure variation in monetary policy announcements are related. Gáti and Handlan (2023) show that the language in monetary policy announcements has a systematic relationship with the Federal Reserve's forecasts of macroeconomic variables. Recent work, like Aruoba and Drechsel (2023); Cieslak et al. (2021); Cieslak and McMahon (2023); Doh et al. (2020); Handlan (2020, 2022), use computational text analysis of the Federal Reserve's announcements and meeting materials to better understand shocks to monetary policy. In particular, Doh et al. (2020) and Handlan (2022) use alternative statements that the Federal Reserve considers at their policy meetings. Our innovation is to use computational text analysis to discipline a structural model to better understand the strategic pressures on the Federal Reserve communication decision.

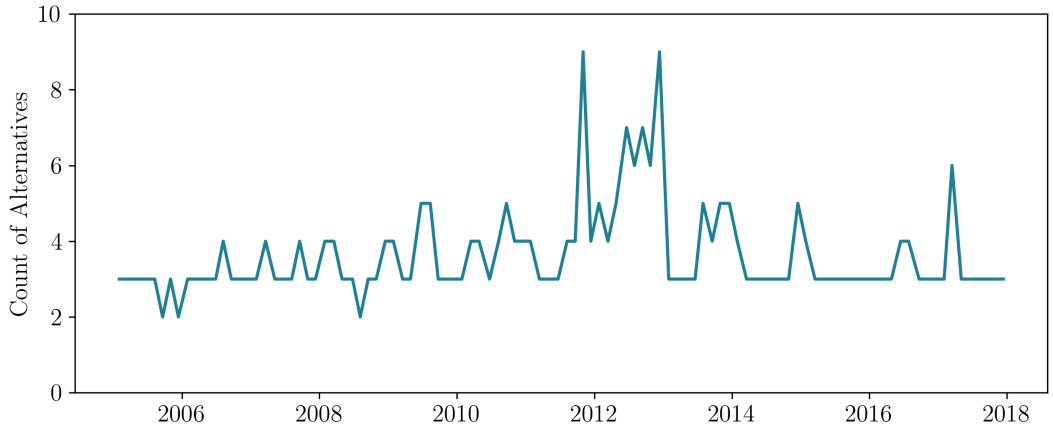
Lastly, this paper connects to the literature studying information effects of monetary policy announcements. Recent work indicates that high-frequency changes in asset prices around FOMC announcements, often called monetary surprises, are a mix of unanticipated monetary policy with other information effects or news (Bauer and Swanson, 2023; Jarociński and Karadi, 2020; Nakamura and Steinsson, 2018). In our model, the responsiveness of public beliefs, measured with asset price changes, depends on the central bank's reputation for being informed. Therefore, only looking at the response to announcements cannot disentangle reputation effects from the underlying policy shock. Understanding the strategic incentives behind monetary policy announcements is critical for quantifying policy shocks and information effects.

2 Empirical Motivation

Multiple versions of the monetary policy statement are drafted and evaluated at each FOMC policy meeting. These alternative candidate statements are in the Tealbook B (formerly the Bluebook) which broadly connects macroeconomic conditions to monetary policy action[1]. During the policy meeting, the FOMC members decide which announcement to

[1]The Tealbook B, formerly the Bluebook, contains the monetary policy recommendations for that meeting and has included clear alternate statements since 2005. Meanwhile, the Tealbook A, formerly the Greenbook,

Figure 1: Alternative FOMC Statements Over Time



Note: Federal Reserve staff members produce alternative versions of the FOMC post-meeting statement for FOMC members to consider for release. Fed staff produce a different number of alternatives depending on the economic situation and FOMC preferences. Each set of proposed wordings for the alternative statements is counted as a new alternative. The modal number of alternatives is three: a hawkish alternative (C), a dovish alternative (A), and one in between (B).

release.

Over time, there has been substantial variation in the number of alternative statements presented to the FOMC members. We tally up each different phrasing as a new alternative message to be considered for release. Figure 1 shows the number of alternatives from 2005 through 2015. For most meetings, there are three alternatives: Alternative A, Alternative B, and Alternative C. However, there are often meetings where there are many more statement versions considered, and occasionally there are only two alternatives.

The announcement language recommendations are captured in different drafts of the post-meeting statement. These alternative versions vary in phrasing, topics, and sometimes also contain different policy decisions. The span of alternative statements is commonly thought about in terms of dovishness or hawkishness. While alternative A is considered more "dovish," concerned with supporting economic growth and employment, the last letter alternative, usually C or D, is considered more "hawkish," concerned with preventing high inflation. In the middle is Alternative B, which is often called the "likely" alternative.

contains the staff forecasts and descriptions of economic conditions. Both are released to the public after a five-year lag.

To make the idea of similarity or difference between alternative announcements more concrete, Table 1 presents some examples. (The full statement texts are included in Appendix A.) The table takes the most extreme alternative statements from the 2008-12 and 2014-04 meetings, and contrasts the sections of the texts where the alternatives differ. The left column, in blue, summarizes the most dovish statement, the right column, in red, the most hawkish statement, and the middle column, in black, the middle statement. The "Topic" column shows what variable the statement summarizes.

The 2008-12 meeting featured a lot of uncertainty, which is reflected in striking differences between the announcement texts. The most hawkish statement not only contradicts the most dovish statement in terms of its assessment of inflation and the real economy, but also features a different policy decision, suggesting that the target rate should be kept at 1 percent. In contrast, the most dovish alternative is not even satisfied with suggesting that the target rate should be lowered, but suggests that since the economy is so weak, specifying a target rate is pointless. We can also see that when the alternatives feature different policy recommendations, they also back those up with differences in the assessment of the economy. This is thus an example of alternative statements being really far in meaning, which is captured by strong differences in wording.

The 2014-04 alternative statements, instead, are very close together. Across the board, they share the policy decision of keeping the target rate at 0-1/4 percent. Their assessment of the economy is also very similar, with only minor differences. One of those is word order: for example, while the most dovish alternative suggests that growth slowed, but is picking up, the hawkish alternative reverts the word order to first indicate that growth is picking up, only then to admit that there are "slowing factors." The second is the choice of slightly different qualifiers when describing the same phenomenon. To go on with the growth example, where the dovish statement simply says that growth slowed, the hawkish alternative instead talks about "slowing factors," which, moreover are also described with the word "fade," which is absent from the dovish statement. Here we can see an example of the Fed using variation in word order and word choice to experiment with statements they could make, before settling on a specific one. This experimentation appears to allow the Fed to fine-tune their message to correspond exactly to what they want to say and the markets to understand.

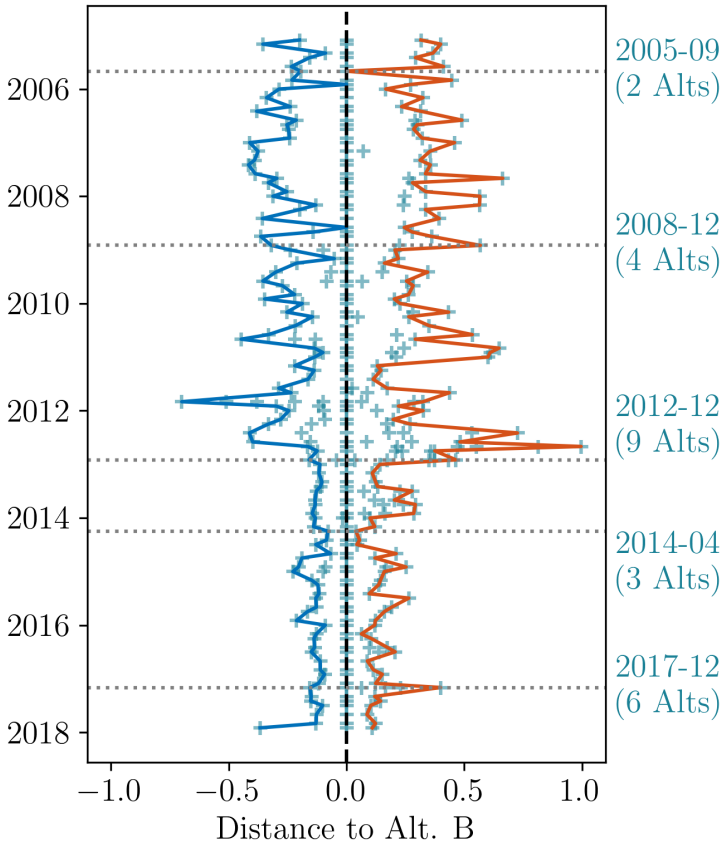
We can quantify this notion of semantic similarity or difference between alternative statements using a frontier language model. More precisely, we use a BERT sentence encoder which takes FOMC statement text from words to a 700-dimension vector, called an embedding, that encodes the meaning behind the joint occurrence of words[2]. We can then measure the differences between statements by measuring the differences between the statement vectors using euclidean distance[3]. Figure 2 plots the alternative messages in relation to each other, normalized around Alt. B. The orange line shows the relative position of the most hawkish alternative (Alt. C, or the last available letter), and the blue line shows the relative position of the most dovish alternative (Alt. A). We order them such that hawkishness, or greater preference for price stability and higher interest rates, is considered more positive. As the orange and blue lines deviate more from zero, it represents that the Fed is considering a wider range of messages at that meeting.

These two measures of the alternatives statements - the number of alternatives and the distance between the alternatives - capture different variations which are useful to study the monetary policy communication strategy. When alternative statements are farther apart then that implies the Federal Reserve is choosing from more opaque statements. Conversely, when the alternative statements are closer together, it reflects that policymakers are trying to make a more nuanced point. When the Federal Reserve staff provide many alternatives it could reflect either uncertainty over the economic situation or uncertainty over how the public will respond to their words. We can think of the set of alternative statements for the Federal Reserve as a function of those two dimensions.

[2] Appendix A includes more background information on the alternative policy statements and the full text of alternative statements from the December 2008 FOMC meeting. Appendix B includes a detailed discussion of text processing and analysis.

[3] We also provide 1-cosine similarity versions of the graphs in Appendix B. The measures differ in magnitude, but qualitatively are consistent.

Figure 2: Distance Between Alternative Statements Over Time



Note: The graph shows the distance in meaning between alternative messages at each FOMC meeting and that meeting's baseline statement, Alt. B (black dashed line). We use a pretrained sentence BERT language model to generate text embedding vectors for comparison. Distance is measured as Euclidean distance between embedding vectors for different statements. The most hawkish statement (orange line) is the last letter (usually Alt C) in the Tealbook/Bluebook. The most dovish statement (blue line) is the first, Alt A. Crosses represent the other alternative statements in the message space for that meeting relative to that meeting's alternative B.

Table 1: Comparison of Statement Alternatives from Two FOMC Meetings

		Information Across Statement Alternatives		
Meeting	**Topic**	**Most Dovish (A)**	**Middle/Likely (B)**	**Most Hawkish (C/D)**
2008-12	Inflation	Below target	Below target	Consistent with target
	Real Economy	Bad and further weakening	Bad and further weakening	Bad but responding to policy
	Target rate	Lower without a new target	Lower to 0-1/4 percent	Keep at 1 percent
2014-04	Inflation	Below target, less confident will return to target over the medium run	Below target, and will monitor developments	Below target, but anticipates a return to target over the medium run
	Real Economy	Growth slowed, but picking up	Picked up recently, after having slowed	Growth picking up as transitory slowing factors fade
	Target rate	Keep at 0-1/4 percent	Keep at 0-1/4 percent	Keep at 0-1/4 percent

Note: This table highlights key features of FOMC statement alternatives from two policy meetings: 2008-12 and 2014-04. The most dovish alternative (usually Alt A) is in the left column, the most hawkish alternative (usually Alt C/D) is in the right column, and the middle or "likely" alternative (Alt B) is in black in the middle. Full versions of the statements are sourced form the Tealbooks/Bluebooks on the Federal Reserve Board website. The most dovish and most hawkish alternatives from 2008-12 and 2014-04 are in the appendix.

3 The Communication Game Setup

There are two players, a central bank (CB) and a public (P). The economic environment consists of a Phillips curve that relates output (y) to inflation (π) and to the public's inflation expectations (x) as

$$y_t = s(\pi_t - x_t), \tag{1}$$

where $s > 0$ is the slope of the Phillips curve, and $x_t \equiv \mathbb{E}_t^P[\pi_t]$ denotes the inflation expectations of the public. The superscript $i = \{CB, P\}$ on the expectation operator indicates whose information set, $\mathcal{I}_t^i$, the expectations are conditioned on. In other words, $\mathbb{E}_t^i[\cdot] \equiv \mathbb{E}[\cdot|\mathcal{I}_t^i]$. The only shock in this simple environment is an inflation target shock, which is a reduced-form stand-in for demand shocks:

$$\omega_t \sim \mathcal{N}(0, 1). \tag{2}$$

The information structure is as follows. The shock ω_t is not revealed to the players, but the central bank observes a noisy signal of it

$$\theta_t = \omega_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \qquad \sigma_\varepsilon^2 \in [0, \infty). \tag{3}$$

Here the CB's signal precision is

$$H \equiv (1 + \sigma_\varepsilon^2)^{-1} \in (0, 1], \tag{4}$$

which we refer to as the central bank's *confidence*. This captures the idea that monetary policy is generally conducted in an environment of uncertainty, where policymakers try to form an idea of the shocks hitting the economy in real time. If the signal precision is perfect, the CB is completely confident in their assessment of the economy ($H \to 1$), while when the signal is infinitely noisy, the CB is completely uncertain about the shock after observing the signal ($H \to 0$).

This information structure implies that the conditional distribution of the shock given

the CB's signal is

$$\omega_t|\theta_t \sim \mathcal{N}(H\theta_t, 1-H). \tag{5}$$

This simply says that, given the signal θ_t, the CB's point forecast of ω_t is $H\theta_t$, with a variance of $1-H$.

The novelty of our paper is to assume that the public does not observe confidence. This allows us to introduce our notion of *reputation for confidence*: the public's belief over confidence, which we denote by $\bar{H}_t \equiv \mathbb{E}_t^P[H]$. The key question we will ask is how the evolution of reputation affects central bank communication, and how communication affects reputation.

We assume that reputation evolves through Bayesian parameter learning. Note that, from the perspective of the public, the CB's signal is normally distributed, $\theta_t \sim \mathcal{N}(0, H^{-1})$, so that our notion of reputation corresponds to learning about true confidence, H, the second moment of the signal distribution. Following Baley and Veldkamp (2022) and Ghofrani (2022), we postulate that the public entertains a Gamma-distribution for true confidence: $H \sim \Gamma(\alpha, \beta)$. Reputation then evolves as P updates beliefs on H as

$$\bar{H}_t = \frac{\alpha_t}{\beta_t}, \tag{6}$$

where

$$\alpha_t = \alpha_{t-1} + \frac{1}{2}, \tag{7}$$

$$\beta_t = \beta_{t-1} + \frac{(k_t - \mu_k)^2}{2}. \tag{8}$$

Here k_t is the observable the public learns from, and μ_k is its average. As we will see later, the public will observe the CB's forecast of ω_t, $H\theta_t$, at the end of period t, when the learning takes place. But because the public cannot separately observe H and θ_t, the observable $k_t = \mathbb{E}_t^P[\theta_t|H\theta_t] = H\theta_t/\bar{H}_{t-1}$, and $\mu_k = 0$.

The CB has two policy tools: setting inflation, π_t, and making a public announcement. Based on the announcement, the public takes a single action, which is to form (rational) expectations of inflation. Concretely, the CB makes an announcement $A_t(\theta_t) \in \mathbb{R}$ of the

signal θ_t, such that $\theta_t \in A_t$. For instance, if the CB observes a signal of $\theta_t = 2$, then a possible example announcement is the interval $[1, 3]$. Importantly, the infinitely fine message "2" is also a possible announcement, as well as the infinitely coarse message $(-\infty, \infty)$. Note that in this way, we are restricting central bank announcements to truthtelling in the sense that no announcement is permitted to convey a lie, i.e. an interval which does not contain θ_t. Therefore, unlike many papers on communication games that concentrate on the notion of truthtelling, we instead focus on the width of the announcement interval, which captures whether announcements are coarse or fine. For example, $A_t(\theta_t) = \mathbb{R}$ is the coarsest candidate message, and $A_t(\theta) = \theta_t$ is the finest, a situation we will refer to as *point revelation*.

Conditional on the announcement and reputation, the public forms expectations of the signal the CB saw

$$\bar{\theta}_t \equiv \mathbb{E}_t^P[\theta_t | A_t, \bar{H}_{t-1}], \tag{9}$$

which they use to form inflation expectations $x_t(\bar{\theta}_t(A_t, \bar{H}_{t-1}))$. Note that at any given period, the central bank can affect the private sector's inflation expectations by making an announcement because, due to the truthtelling restriction, any announcement will contribute weakly useful information. Thus, a key object of interest is the message space, denoted by $\mathcal{M}_t$, which maps each signal θ_t that the CB observes into an announcement $A_t(\theta_t)$. In other words, $\mathcal{M}_t$ is a partition of the real line such that $\mathcal{M}_t : \ \theta_t \to A_t(\theta_t)$. Importantly, the message space will be an equilibrium object, as we outline below.

The public's loss is a quadratic in current inflation

$$\mathcal{L}^P(x_t, \pi_t) = \mathbb{E}_t^P[(x_t - \pi_t)^2], \tag{10}$$

while the period loss of the central bank is given by

$$\mathcal{L}^{CB}(x_t, \pi_t, \theta_t) = \mathbb{E}^{CB}[(y_t - b)^2 + \lambda(\pi_t - \pi^* - \omega_t)^2]. \tag{11}$$

Here $\pi^* > 0$ is the inflation target, $\lambda > 0$ is the weight on inflation stabilization in the loss function and $b \geq 0$ is an output bias that introduces an incentive for the CB to induce

surprise inflation.[4]

Figure 3: Timing

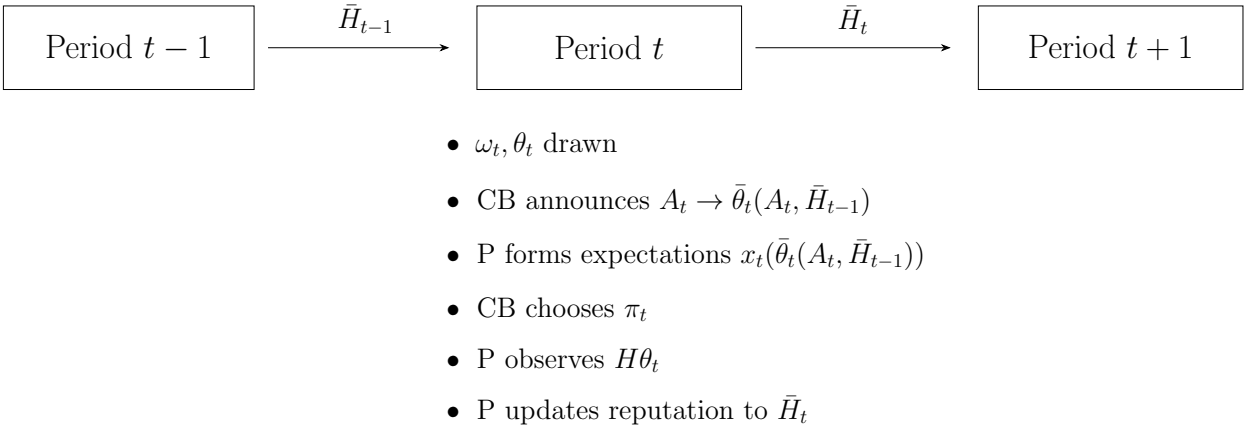


Figure 3 summarizes the timing. In the beginning of period t, reputation is inherited from period $t-1$. The shock ω_t is drawn, and the CB gets a noisy signal of it, θ_t. Given their measurement of the state of the economy (θ_t), the CB makes an announcement A_t to the public about θ_t. The public uses the announcement to form their inflation expectations x_t, taking into account the CB's reputation for confidence in making inferences from the announcement. Then the CB chooses inflation, and thus output y_t is also realized through the Phillips curve. Given the realizations of the observables, the public observes the CB's forecast of ω_t, $H\theta_t$, and uses it to update reputation.

We proceed by analyzing the communication problem in two stages. First, we investigate how an exogenous reputation affects communication in the stage game. In the language of Figure 3, we zoom in on period t, taking the prior on reputation, $\bar{H}_{t-1}$, as given. We then consider a repeated game setting, in which in each period, the same stage game as in period t is played. The only difference is that we now allow reputation to evolve, so that each new period inherits a new reputation.

[4] A large literature in the wake of Crawford and Sobel (1982) and Barro and Gordon (1983) has focused on misaligned preferences between the CB and the public as a mechanism that leads to coarse communication in equilibrium. Here, instead, the mere presence of evolving reputation will enable coarse equilibrium communication, and we can set $b = 0$.

4 Stage Game

How does reputation affect communication in a one-shot game? To analyze this, we first zoom in on period t from Figure 3, and drop time indices for simplicity (we will reintroduce them in Section 5). Our equilibrium concept is that of a Perfect Bayesian Equilibrium with Communication, as spelled out below.

Definition 1 *(Perfect Bayesian Equilibrium with Communication (PBEC))*
Given reputation, a PBEC is π_t, x_t *and message space* $\mathcal{M}_t$ *(a partition of* $\mathbb{R}$*) such that*

- $\pi_t = argmin\ \mathcal{L}^{CB}(x_t, \pi_t, \theta_t)$
- $x_t = argmin\ \mathcal{L}^{P}(x_t, \bar{H}_{t-1}, \bar{\theta}_t)$
- $\mathcal{M}_t$ *induces CB to make an announcement* $A_t \in \mathcal{M}_t$ *such that* $\theta_t \in A_t$ *and*
  - *P believes the announcement so* $\bar{\theta}_t = \mathbb{E}^P[\theta_t | \theta_t \in A_t, \bar{H}_{t-1}]$
  - *CB prefers* A_t *that induces* $\bar{\theta}_t$ *to any alternative* $\bar{\theta}'_t$

$$\mathcal{L}^{CB}(\cdot, \bar{\theta}_t) \leq \mathcal{L}^{CB}(\cdot, \bar{\theta}'_t) \qquad (Incentive\ Compatibility)$$

Using the PBEC equilibrium concept, we now solve the stage game backwards, first computing the best-response inflation rate that minimizes the period loss of Equation 11

$$\pi^{BR} = \frac{1}{s^2 + \lambda}(s^2 x_t + \lambda \pi^* + \lambda H \theta_t + sb). \tag{12}$$

Postulating rational expectations on the part of the public and conditioning on a central bank announcement $A_t(\theta_t)$, this gives us the best-response inflation expectations

$$x_t^{BR} = \pi^* + \bar{H}_{t-1}\bar{\theta}_t + \frac{sb}{\lambda}. \tag{13}$$

Given the two best-responses, equilibrium inflation is

$$\pi_t^{eqb} = \pi^* + \frac{sb}{\lambda} + \frac{s^2 \bar{H}_{t-1}\bar{\theta}_t + \lambda H \theta_t}{s^2 + \lambda}, \tag{14}$$

and equilibrium output is given by the Phillips curve in Equation 1.

Now we derive equilibrium communication given optimal inflation and inflation expectations, suppressing the details to Appendix C. Substituting π^{eqb} and x^{BR} into the period loss of Equation 11, and postulating that the equilibrium announcement should fulfill the incentive compatibility from the definition of equilibrium, we arrive at the following characterization.

Proposition 1 *(Point Revelation)*
A point revealing equilibrium (where the CB sends the singleton message $\bar{\theta}_t = \theta_t$) exists if and only if

- $b = 0$, *and*
- $\bar{H}_{t-1} = H$ *or* $\theta_t = 0$.

Proof: See Appendix C.

Proposition 1 gives conditions for when the CB point reveals their signal, i.e. makes an announcement that exactly corresponds to the number θ_t the CB observed. The proposition is saying that point revelation is only possible if there is no output bias and either reputation equals true confidence, or the CB's signal corresponds to the unconditional expectation of the shock, zero.[5]

The requirement that $b = 0$ is familiar from the literature on communication games which emphasize the importance of aligning the CB's preferences with the public in order to get rid of incentives for the CB to trick the public (Barro and Gordon, 1983; Clarida et al., 1999; Crawford and Sobel, 1982; Rogoff, 1985). When trying to rationalize coarse communication in equilibrium, papers either modify the central bank's loss function to include an explicit concern for reputation (Ottaviani and Sørensen, 2006a,b), or set $b > 0$ as in Moscarini (2007).

Here, instead, $b = 0$ is not sufficient for point revelation. The first of the two possible additional requirements, $\theta_t = 0$, is trivial. It simply says that if the CB thinks that no shock is hitting the economy, then they anticipate not to have to take any inflation-stabilizing

[5]The existence of point revealing equilibria does not close out the existence of other equilibria that involve coarse communication.

action, and therefore point revealing this information leaves the economy ex ante in the steady state, as private expectations remain at their steady state value. Clearly, this is not an interesting case in practice.

The crucial novel result here is the other possible requirement, namely that $\bar{H}_{t-1} = H$. The interpretation is that if the CB thinks that shocks are hitting the economy ($\theta_t \neq 0$), then not only does $b = 0$ need to hold, but the private sector also needs to believe that the CB is exactly as confident in the shocks as they truly are. Conversely, if the public is mistaken in their assessment of the CB's confidence, then even without an inflation surprise motive, point revelation cannot happen in equilibrium.

The intuition behind this result is that if the private sector misperceives the central bank's confidence, then they either overreact (if reputation is higher than actual confidence), or underreact (if reputation is lower than confidence) to the same announcement. Therefore the central bank cannot point reveal their signal, but needs to adjust their announcement to reflect the over- or underreaction of the public. Varying the width of the announcement interval is a way of doing so.

5 Repeated Game

In this section we explore the repeated game in order to see how different shocks and the reputation for confidence affect the equilibrium message space, in both number of messages and the announcement interval width.

We use the calibration outlined in Table 2. π^* is set to correspond to the Federal Reserve's inflation target. λ is set to match the ratio of weights on inflation versus output stabilization. Following Rotemberg and Woodford (1997)'s estimate of a 0.05 weight on output stabilization when setting the weight on inflation to 1, this implies a 20:1 ratio. Lastly, we set s by observing that we can rewrite our Phillips curve of Equation 1 to express inflation as a function of output and inflation expectations, implying that $1/s$ corresponds to the New Keynesian slope of the Phillips curve. Setting the latter to match an expected price duration of two quarters, we get a number of 0.0850, which we invert to get 11.7647, and then round

up.[6] We look at both a zero and a positive output bias, for the latter case following Moscarini (2007)'s calibration.

Table 2: Calibration

Parameter	Value	Target
π^*	2	Federal Reserve inflation target
s	12	Expected price duration of two quarters
λ	20	Rotemberg and Woodford (1997)
b	(0,0.02)	Moscarini (2007)

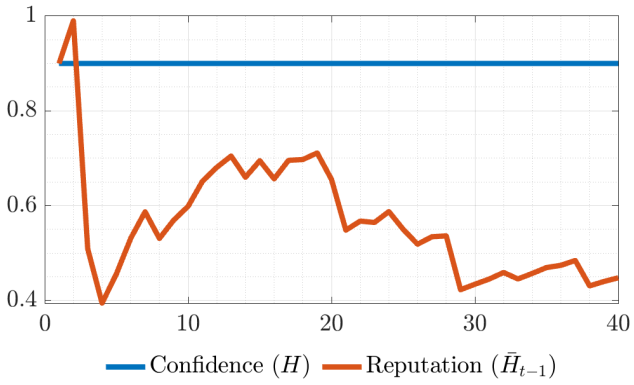
We then simulate a sequence of shocks, and initializing reputation at true confidence ($\bar{H}_0 = H$), we solve for the equilibrium message space. To do this, we employ a so-called shooting algorithm, outlined in Appendix D. The idea of the algorithm is to guess a sequence of cutoffs $\{\theta^k\}_{k=1}^{K-1}$ for a given K, compute the associated conditional expectations, $\{\bar{\theta}_k\}_{k=1}^{K}$, and adjust the cutoffs to minimize errors to incentive compatibility conditions, subject to the constraint that the cutoff and conditional expectations sequences are increasing sequences. We start the algorithm with $K = K_0 = 12$, and if no equilibrium is found, we decrease K by 1.[7] We keep going until $K = 2$, which is the most coarse message space. If no equilibrium message space is found, we set $K = 1$, designating a babbling equilibrium where the public's post-announcement expectation is just the unconditional mean of θ, zero.

There are two things to note here. First, there is a multiplicity of equilibria. For a particular set of parameter values, there are different coarse message spaces, ranging from $K = 1$ (a babbling equilibrium) to some initially unknown $K = K^{max} \in \mathbb{N}^+$ (which may be 1). In what follows, from among the multiple coarse equilibrium message spaces, we will focus on the finest message space; that is, the highest finite K.

[6] The slope of the Phillips curve in the standard New Keynesian model is $\kappa = \frac{\omega+1/\sigma}{1+\omega\theta}\frac{(1-\alpha)(1-\alpha\beta)}{\alpha}$, where the parameters $(\omega, \sigma, \theta, \beta, \alpha)$ are the elasticity of marginal cost to output (including the Frisch elasticity), the intertemporal elasticity of substitution, the price elasticity of demand, the discount factor and the probability that a firm is stuck with their price, respectively. We set α to match an expected price duration of two quarters, and the other parameters are set to the values of Chari et al. (2000), which is the calibration followed in Woodford (2003).

[7] K_0 is selected by experimentation. One way to calibrate it would be to set it equal to the largest number of alternative FOMC statements the Fed has entertained, which is nine.

Figure 4: Reputation and Confidence

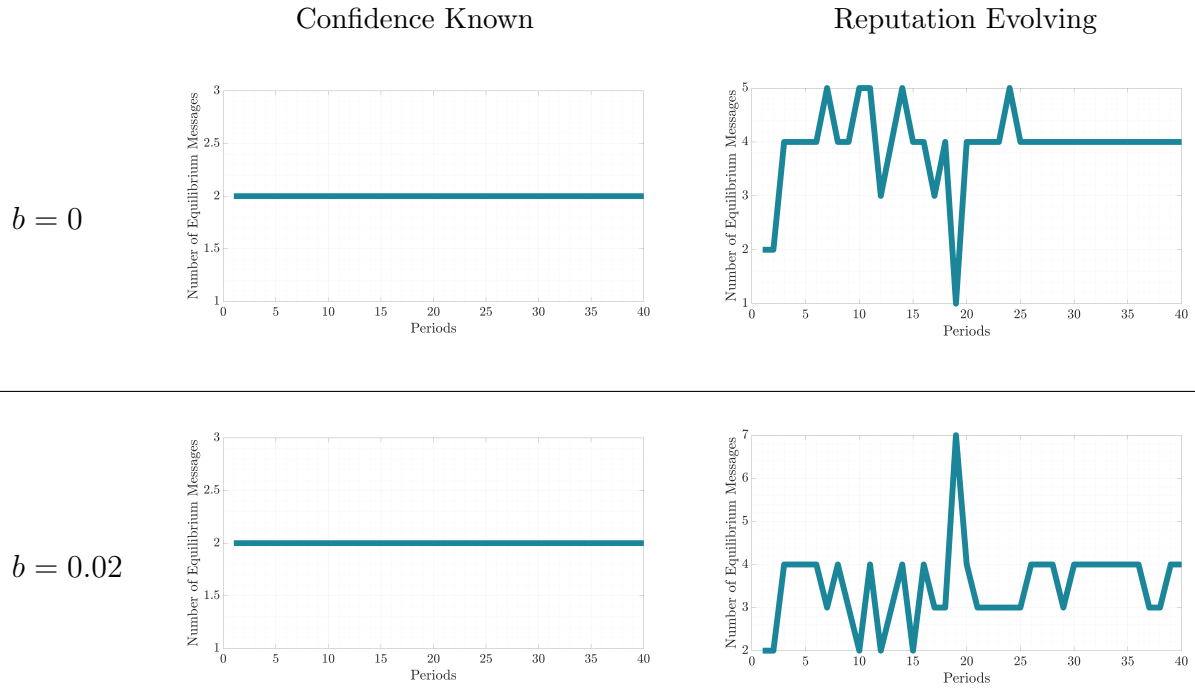


Second, even when the conditions of Proposition 1 are met, and an infinitely fine message space exists, coarse message spaces also exist. Since the algorithm is designed to find coarse message spaces, in this case it will find message spaces with a finite K. In the following plots, to aid interpretation, in cases where both coarse and fine message spaces exist, we focus on the finest coarse message space.

Throughout this section, we employ the same simulated sequence of shocks, and compare two versions of the repeated game. In the first, designated as "Confidence Known" in the plots and depicted in the left columns, we assume that $\bar{H}_t = H \ \forall t$, so that the public knows the CB's true confidence and so there is no learning, and reputation plays no rule. In the second case, the public learns about true confidence and so reputation evolves according to Bayesian learning. We show this case in the right column of the plots, under the label "Reputation Evolving." We also distinguish between the case where there is no output bias ($b = 0$, top row), and when there is a positive bias ($b = 0.02$, bottom row).

Before comparing the two repeated game versions, let us look at the evolution of reputation in this simulation (when we let the public learn about confidence). Figure 4 shows true confidence in blue and reputation in red. As we can see, we set confidence to a fairly high value, 0.9, which is also where we initialize reputation. For this sequence of shocks, reputation goes up in the first period, only to drop excessively in the next few periods, and to hover between 0.4 and about 0.7 for the rest of the sample. This sequence of shocks is interesting because even though we have not built in any concern on the part of the central bank for reputation-building, the sample features not only periods of dropping reputation,

Figure 5: Number of Alternative Statements



Note: The figure compares the number of alternative statements the central bank prepares in the data with two versions of the repeated game: one in which confidence is known (reputation equals confidence), and the other where reputation evolves as the public is learning about true confidence. The top row shows the case without output bias ($b = 0$), while the bottom row depicts the case with positive bias ($b = 0.02$).

but also a period of accumulating reputation (approximately between period five and 19). This suggests that even if the CB does not take the effect of their own actions on their reputation into account, they may get "lucky" in the sense that a sequence of shocks hits the economy that happen to contribute to improving the CB's reputation.

Figure 5 plots the number of alternative statements that the CB drafts over time. When the CB's confidence is known (left column), the exact same stage game is replayed in every period, and thus the equilibrium number of alternative messages is the same in all periods. If $b = 0$, this number is two, which displays a case where even as the conditions of Proposition 1 are fulfilled, there also exists a coarse communication equilibrium. Thus, next to the point revelation equilibrium, the finest coarse message space that exists is two-message equilibrium. When $b = 0.02$, the equilibrium number of alternative messages continues to be always two. Thus, interestingly, even though Proposition 1 is now violated, these two equilibria look

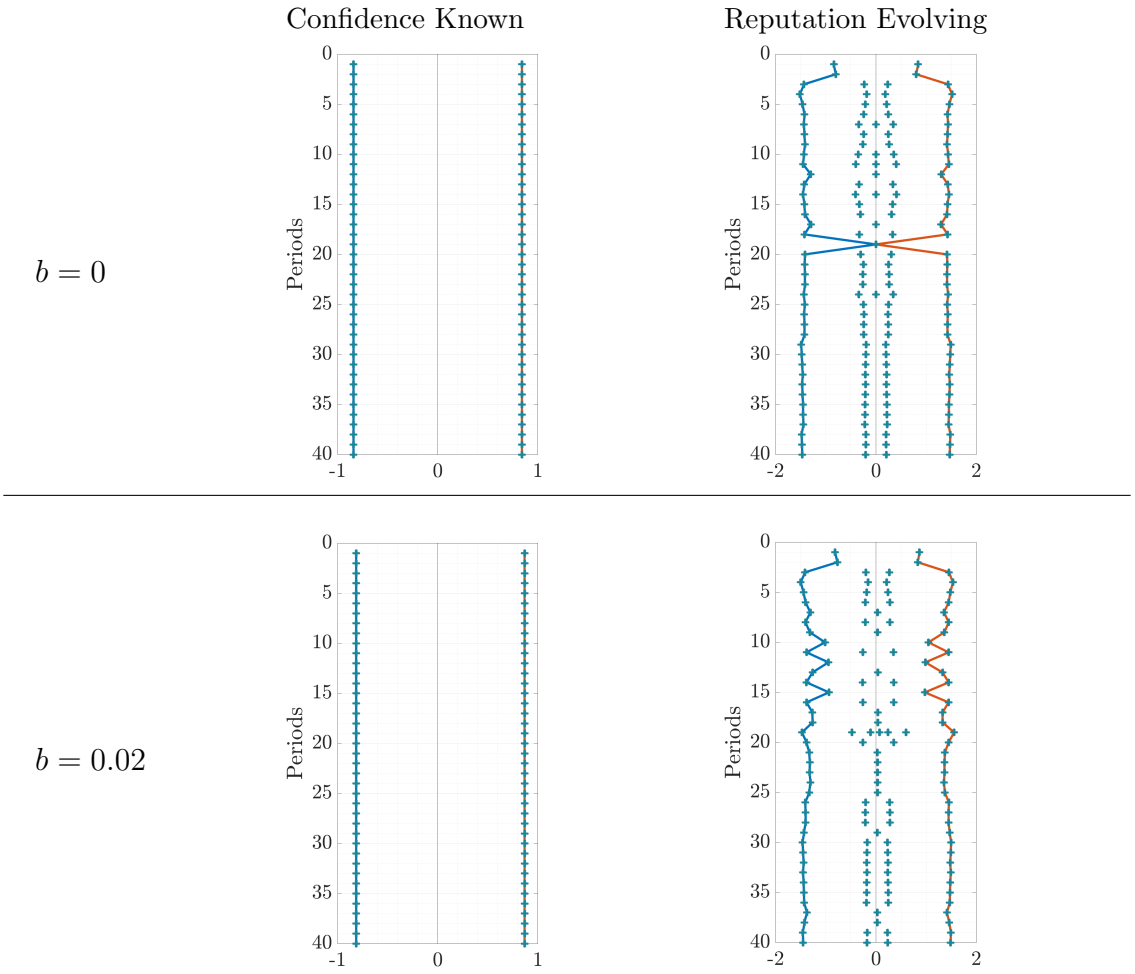
remarkably similar. In both cases, one can think of the CB as alternating between two announcement drafts: they either announce that "the economy is too cool" when they see a negative signal, or that "the economy is too hot" when they see a positive signal.

But, as we saw on Figure 1, the number of alternative statements the Fed drafts varies over time. When looking at the right column of Figure 5, we see that what one needs to match this pattern in the data is for reputation for confidence to evolve over time. When the CB has an incentive to induce an inflation surprise ($b > 0$), as in the bottom row, the model produces anything between two and seven alternative statements, and the number of alternatives goes up and down in a pattern that is remarkably close to the data. But even without an output bias ($b = 0$), the model is able to replicate the time-varying nature of alternative statements, albeit with somewhat more symmetry.

Why is it crucial for reputation to evolve in order to replicate the pattern that the CB drafts a varying number of alternative statements over time? The reason is that the private sector's optimal response to a CB announcement is to discount it by the CB's confidence. However, since the public does not know the true level of confidence, they rely on their perception of it: the CB's reputation. In this way, reputation modulates the response of private sector expectations to CB announcements - a feature the CB takes into account when drafting potential statements. A high-reputation situation does not require finely nuanced alternative statements because a given statement will be given more weight in the private sector's expectations. When instead the CB has a bad reputation, they need to craft very fine-tuned alternatives in order to end up releasing the statement that will move expectations by exactly the right amount and in the right direction.

Similarly to the number of statements, also the distance between the most extreme alternatives is time-varying in the data, as we saw on Figure 2. Figure 6 shows that for the model to reproduce this feature, we again need reputation to be evolving over time. When the public knows how confident the CB is, the public's responsiveness to the CB's words is constant over time, and thus the CB can just pick between the same alternative statements every period, depending on the signal they observe. Given that the alternatives are the same, they also span the same space. Instead, evolving reputation and thus evolving responsiveness of the public's expectations leads to a varying number of potential messages

Figure 6: Distance Between Alternatives



Note: The figure compares the distance in meaning between the alternative statements in the data with two versions of the repeated game: one in which confidence is known (reputation equals confidence), and the other where reputation evolves as the public is learning about true confidence. The messages are normalized with respect to the 2% inflation target. The top row shows the case without output bias ($b = 0$), while the bottom row depicts the case with positive bias ($b = 0.02$).

which span spaces of varying sizes. Thus the model can reproduce messages that are as far apart as the alternatives in 2008-12, as well as the nearly indistinguishable alternatives in 2014-04.

One question that one may have is whether, instead of time-varying reputation, time-varying confidence would suffice to match a time-varying number and similarity of alternative statements? The answer depends on whether reputation equals confidence or not. Suppose that $b = 0$ and $\theta_t \neq 0 \; \forall t$. In this case, what determines whether Proposition 1 holds or not is whether $\bar{H}_t = H_t$, where we are now including a t-subscript on confidence to indicate that it may be time-varying.

Now suppose that confidence evolves over time, but this is perfectly observed by the public. In this case, $\bar{H}_t = H_t \; \forall t$, and thus Proposition 1 indicates that point revelation can be an equilibrium. Under these assumptions, then, the model does not guarantee that alternative statements are drafted at all. In this case, then, time-varying confidence is not sufficient to match the data.

Thus, if we want time-varying confidence to drive a varying number and similarity of messages, we need reputation to be constant in order to guarantee that the equilibrium message space is coarse. Under the assumption of constant reputation, then, variation in how confident the central bank is in their own economic assessments introduces time-variation in the incentive to provide granular information, and is thus sufficient to match the data.

This analysis reveals that what is really key for the model to match the data is to have a distinction between the CB's true confidence and the market's perception of confidence. Once that is true, time-variation in either of the variables will do the job. In practice, however, we think that it is very likely that both confidence and reputation for confidence vary over time.

6 Conclusion

When a central bank makes announcements, not only do markets pay attention to the words, but the public's belief at how good the central bank is at doing its job - the reputation for confidence - matters for how markets respond to announcements. In this paper, we model

this interaction, and explore how it affects central bank communication and monetary policy.

We find that even absent a concern for reputation - so that the central bank ignores that their actions may affect the evolution of reputation - the model generates a central bank communication policy that is in line with how the Fed communicates. This involves drafting a set of candidate messages and picking from this set a single announcement that is released to the public. Importantly, we find that the evolution of reputation for confidence drives the central bank to consider a varying number of alternative statements that are varying in their similarity to each other.

The reason this happens is that the central bank creates their announcements by anticipating how the public will receive that information and balancing that against their incentives to adjust public expectations. We show that this strategic tradeoff measurably affects the message space the FOMC considers, and thus ultimately the announcements they choose. With the growing literature studying information effects, it is important to take these strategic language motives into account.

References

Amador, Manuel and Christopher Phelan (2021) "Reputation and sovereign default," *Econometrica*, 89 (4), 1979–2010.

Amador, Manuel and Pierre-Olivier Weill (2010) "Learning from Prices: Public Communication and Welfare," *Journal of Political Economy*, 118 (5), 866–907, http://www.jstor.org/stable/10.1086/657923.

Angeletos, George-Marios and Alessandro Pavan (2007) "Efficient Use of Information and Social Value of Information," *Econometrica*, 75 (4), 1103–1142, 10.1111/j.1468-0262.2007.00783.x.

Aruoba, S. Borağan and Thomas Drechsel (2023) "Identifying Monetary Policy Shocks: A Natural Language Approach," *Working Paper*.

Austen-Smith, David (1990) "Information transmission in debate," *American Journal of Political Science*, 34.

Baley, Isaac and Laura Veldkamp (2022) "Bayesian Learning," in *Handbook of Economic Expectations*: Elsevier.

Barro, Robert J and David B Gordon (1983) "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, 12 (1), 101–121.

Bassetto, Marco (2019) "Forward guidance: Communication, commitment, or both?" *Journal of Monetary Economics*, 108, 69–86, 10.1016/j.jmoneco.2019.08.015.

Bauer, Michael and Eric Swanson (2023) "An Alternative Explanation for the "Fed Information Effect"," *American Economic Review*, 113 (3), 664–700.

Benabou, Roland and Guy Laroque (1992) "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *The Quarterly Journal of Economics*, 107 (3), 921–958, http://www.jstor.org/stable/2118369.

Camous, Antoine and Dmitru Matveev (2024) "Central Bank Strategic Communication and the Dynamics of Reputation," *Working Paper*.

Chahrour, Ryan (2014) "Public Communication and Information Acquisition," *American Economic Journal: Macroeconomics*, 6 (3), 73–101, 10.1257/mac.6.3.73.

Chari, Varadarajan, Patrick Kehoe, and Ellen McGrattan (2000) "Sticky price models of the business cycle: can the contract multiplier solve the persistence problem?" *Econometrica*, 68 (5), 1151–1179.

Cieslak, Anna, Stephen Hansen, Michael McMahon, and Song Xiao (2021) "Policymakers' Uncertainty," October, 10.2139/ssrn.3936999.

Cieslak, Anna and Michael McMahon (2023) "Tough Talk: The Fed and the Risk Premium," *Working Paper*, 10.2139/ssrn.4560220.

Clarida, Richard, Jordi Galí, and Mark Gertler (1999) "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37 (4), 1661–1707, http://www.jstor.org/stable/2565488.

Crawford, Vincent P. and Joel Sobel (1982) "Strategic Information Transmission," *Econometrica*, 50 (6), 1431–1451, 10.2307/1913390.

Cukierman, Alex and Allan H. Meltzer (1986) "A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information," *Econometrica*, 54 (5), 1099–1128, 10.2307/1912324.

Diamond, Douglas (1989) "Reputation Acquisition in Debt Markets," *Journal of Political Economy*, 97 (4), 828–862.

——— (1991) "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt," *Journal of Political Economy*, 99 (4), 689–721.

Doh, Taeyoung, Dongho Song, and Shu-Kuei Yang (2020) "Deciphering Federal Reserve Communication via Text Analysis of Alternative FOMC Statements," *The Federal Reserve Bank of Kansas City Research Working Papers*, 10.18651/RWP2020-14.

Frankel, Alex and Navin Kartik (2018) "What kind of central bank competence?" *Theoretical Economics*, 13 (2), 697–727.

Gáti, Laura (2023) "Talking Over Time – Dynamic Central Bank Communication," *Journal of Money, Credit and Banking*, 55 (5), 1147–1176.

Gáti, Laura and Amy Handlan (2023) "Monetary Communication Rules," *Working Paper*.

Ghofrani, Erfan (2022) "Learning Uncertainty," *Working Paper*.

Guembel, Alexander and Silvia Rossetto (2009) "Reputational cheap talk with misunderstanding," *Games and Economic Behavior*, 67 (2), 736–744.

Handlan, Amy (2020) "FedSpeak Matters: FOMC Statements and Monetary Policy Expectations," *Published Manuscript*.

——— (2022) "Text shocks and monetary surprises: Text analysis of fomc statements with machine learning," *Published Manuscript*.

Hansen, Stephen and Michael McMahon (2016) "First Impressions Matter: Signalling as a Source of Policy Dynamics," *Review of Economic Studies*, 83 (4), 1645–1672.

Herbert, Sylverie (2021) "State-Dependent Central Bank Communication with Heterogeneous Beliefs,"Technical report.

Iovino, Luigi, Jennifer La'O, and Rui Mascarenhas (2022) "Optimal monetary policy and disclosure with an informationally-constrained central banker," *Journal of Monetary Economics*, 125, 151–172.

Jarociński, Marek and Peter Karadi (2020) "Deconstructing Monetary Policy Surprises—The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12 (2), 1–43.

Kamenica, Emir and Matthew Gentzkow (2011) "Bayesian Persuasion," *American Economic Review*, 101 (6), 2590–2615.

King, Robert G and Yang K Lu (2021) "The rise, fall and stabilization of us inflation: Shifting regimes and evolving reputation,"Technical report, National Bureau of Economic Research.

Ko, Eunmi (2019) "Optimal Central Bank Forward Guidance,"Technical report.

Kostyshyna, Olena and Luba Petersen (2023) "The Effects of Communicating Inflation Uncertainty on Household Expectations," *Bank of Canada Working Paper*.

Kreps, David M and Robert Wilson (1982) "Reputation and imperfect information," *Journal of Economic Theory*, 27 (2), 253–279.

Li, Ming and Kristóf Madarász (2008) "When mandatory disclosure hurts: Expert advice and conflicting interests," *Journal of Economic Theory*, 139 (1), 47–74.

Mailath, George and Larry Samuelson (2001) "Who Wants a Good Reputation?" *The Review of Economic Studies*, 68 (2), 415–441.

Milgrom, Paul and John Roberts (1982) "Predation, reputation, and entry deterrence," *Journal of Economic Theory*, 27 (2), 280–312.

Morris, Stephen (2001) "Political Correctness," *Journal of Political Economy*, 109 (2), 231–265, http://www.jstor.org/stable/10.1086/319554.

Moscarini, Giuseppe (2007) "Competence Implies Credibility," *American Economic Review*, 97 (1), 37–63, 10.1257/aer.97.1.37.

Nakamura, Emi and Jón Steinsson (2018) "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect," *Quarterly Journal of Economics*, 133 (3), 1283–1330.

Ottaviani, Marco and Peter Norman Sørensen (2006a) "Professional Advice," *Journal of Economic Theory*, 126 (1), 120–142.

——— (2006b) "Reputational Cheap Talk," *The RAND Journal of Economics*, 37 (1), 155–175, http://www.jstor.org/stable/25046233.

Phelan, Christopher (2006) "Public trust and government betrayal," *Journal of Economic Theory*, 130 (1), 27–43.

Rogoff, Kenneth (1985) "The optimal degree of commitment to an intermediate monetary target," *The Quarterly Journal of Economics*, 100 (4), 1169–1189.

Rotemberg, Julio and Michael Woodford (1997) "An optimization-based econometric framework for the evaluation of monetary policy," *NBER Macroeconomics Annual*, 12, 297–346.

Sastry, Karthik (2022) "Disagreement about monetary policy," *Available at SSRN 3421723*.

Sobel, Joel (1985) "A Theory of Credibility," *The Review of Economic Studies*, 52 (4), 557–573, http://www.jstor.org/stable/2297732.

Woodford, Michael (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*: Princeton University Press.