Expectations and the Transmission of Monetary Policy[†]

Vicente Jimenez-Gimpel
MIT

Tomás E. Caravello
MIT

November 5, 2026

Click here for the latest version.

Abstract

Monetary policy works largely through expectations: long-term yields, mortgage rates, and investment depend on expected future policy. We develop a model in which policy decisions reveal information about the central bank's imperfect and persistent real-time assessments and thus about future policy. Systematic policy determines how markets interpret decisions and hence the propagation of assessment errors into macroeconomic outcomes. We show that choosing rates freely after forming assessments (discretion) leaves the economy inefficiently exposed to assessment errors: in particular, expectations can be excessively sensitive to policy surprises. Unlike inflation or stabilization bias, this time inconsistency is rooted in the imperfect and private nature of the central bank's assessments. We quantify its aggregate cost in a New Keynesian model estimated using U.S. data, including Federal Reserve forecast revisions and inflation expectations' sensitivity to policy surprises. The optimal commitment rule—chosen before assessments are formed—substantially improves stabilization. This rule responds much less to assessments of productive capacity relative to demand. Consequently, markets attribute less of a policy surprise to persistent assessment errors, dampening revisions to expected future policy and their effects on inflation and output.

Keywords: Monetary policy, monetary transmission, expectations management, time inconsistency

JEL Codes: D83, D84, E52, E58, E61

[†]E-mails: jimenezv@mit.edu, tomasec@mit.edu. First posted version: October 9, 2026. Jimenez-Gimpel is deeply grateful to Ricardo J. Caballero, Christian K. Wolf, and Iván Werning for their continued support and guidance. We thank Daron Acemoglu, Cristina Arellano, Corina Boar, Gabriel Chodorow-Reich, Anna Cieslak, Darrell Duffie, Jesús Fernández-Villaverde, Lukas Freund, Masao Fukui, Rohan Kekre, Dmitry Mukhin, Jonathan Parker, Christina Romer, David Romer, Kunal Sangani, Karthik Sastry, Kerry Siani, David Thesmar, Jean Tirole, Adrien Verdelhan, Olivier Wang, and participants at several venues for useful comments.

1 Introduction

Central banks set policy guided by their imperfect real-time reading of economic conditions, and markets know it. Since these views are usually persistent, rate decisions help markets predict future policy by partially revealing them. In turn, expectations of future policy shape long yields, mortgage rates, asset prices, and investment decisions—some of the main channels through which monetary policy reaches the real economy. Thus, what markets infer about future policy from a given decision is not a side effect; it largely *is* monetary transmission. The same rate decision, or even the same realized path of policy rates, can have very different effects depending on how markets interpret it.

This paper develops a framework to study monetary policy when the central bank acts on imperfect real-time assessments and its decisions affect the economy partly through how the market interprets them. These assessments concern demand conditions and productive capacity, the economy's two fundamentals. The private sector observes current fundamentals, but not the central bank's assessments of them. How it interprets a policy decision shapes its beliefs about future policy and hence current spending and output. This interpretation depends on the central bank's systematic response. For example, suppose the central bank's demand assessment is fully transitory and unrelated to its persistent view of productive capacity. Consider a tightening caused by an overestimate of demand. If the central bank systematically sets the current rate using only its demand assessment, the market knows the tightening conveys no news about future policy. If its systematic response also includes productive capacity, the market can interpret that same tightening as partly reflecting persistent pessimism about capacity. Thus, the market anticipates continued tight policy and reduces current spending. Hence, the central bank's systematic response shapes how strongly expectations react to policy surprises and thus the effects of assessment errors on current spending.

We show that the interaction between imperfect assessments and endogenous market interpretation gives rise to a novel source of time inconsistency in monetary policy. Under discretion, the central bank sets its rate optimally given its assessments, accounting for the fact that market expectations react to policy but taking as given how strongly they do so. When committing to a rule before observing its assessments, instead, the central bank acknowledges how the rule itself shapes the market interpretation of each decision and thus the strength of the response of expectations to policy surprises. Thus, the optimal commitment rule balances stabilization conditional on the central bank's assessments against limiting the amplification of assessment errors through expectations. This trade-off disappears when assessments are perfect.

We extend our framework to a New Keynesian model and estimate it using U.S. data, including revisions to the Federal Reserve's real-time forecasts. We find that this mechanism entails sizable stabilization costs. The optimal commitment rule responds much less aggressively than discretion to the central bank's best real-time estimates of productive capacity (natural output). To see why, note that errors about productive capacity are slow to correct. Responding less to these assessments relative to demand makes the market attribute less of a policy surprise to persistent errors, dampening its revisions to expected future policy and thus the macroeconomic impact of these errors.

Framework. Our baseline model isolates the informational source of time inconsistency in a New Keynesian environment with fully rigid prices. We focus on the first two periods, after which the economy returns to its steady state. Output is demand determined and spending decisions are forward looking, so the private sector's expectations of future policy affect current spending and thus output. The economy is subject to shocks to demand and productive capacity (natural output). The central bank aims to keep output close to its natural level and bases its decisions on imperfect real-time estimates of these fundamentals.

The private sector (henceforth, the market) does not observe these assessments directly, but knows how the central bank systematically responds to them. We abstract from the conventional "information effect" by assuming that the market observes contemporaneous fundamentals. Our mechanism rests on a more limited premise: the central bank's assessments of today's economy help predict how it will assess the economy tomorrow. Its current views are thus informative about future policy and, through it, future output.[1]

Market interpretation and policy design. While the central bank's real-time assessments are private, its policy decisions convey information about them to the market. A given decision, however, does not reveal the two assessments separately. For example, a surprise tightening can be caused by the central bank seeing stronger demand or lower productive capacity. The two assessments can differ both in their persistence and in how strongly policy responds to them today relative to tomorrow. Thus, for a given rate surprise, the share the market attributes to each shapes how much it revises its expectations of future output.

This attribution depends on the central bank's systematic response. Consequently, systematic policy shapes how strongly a given decision moves market expectations and, in turn,

[1] While simple, this observation is relevant for current policy debates. Federal Reserve Chair Kevin Warsh has recently called for less explicit communication about future policy, arguing that it would encourage markets to focus more on economic fundamentals (Warsh, 2026). This, however, would not change their incentive to anticipate future policy. With or without explicit guidance, there is one signal the Fed cannot avoid sending: its rate decisions.

current output (Proposition 1). Since these decisions are based on imperfect assessments, this expectations response also shapes how strongly assessment errors propagate into output.

We study two policy regimes. Under discretion, the central bank chooses its rate optimally after observing its assessments. Under commitment, it chooses a linear rule before observing them and follows it afterward.

Under discretion, the central bank uses both the direct effect of the policy rate and the response of market expectations to align expected output with its assessment of productive capacity. In equilibrium, the market interprets decisions using the central bank's systematic policy, yet the central bank takes that interpretation as given when choosing each rate (Proposition 2).

With commitment, the central bank instead accounts for how its systematic response shapes the market's interpretation of each policy decision. Thus, the optimal ex ante linear rule balances keeping expected output close to the central bank's assessment of productive capacity against limiting inefficient fluctuations caused by assessment errors (Lemma 2). We show that commitment reduces the propagation of assessment errors. Under intuitive conditions, the optimal commitment policy entails a weaker response to natural-output assessments relative to demand, making market expectations less sensitive to assessment errors (Proposition 3 and Corollary 1).

The optimal ex ante rule is *time inconsistent*. Once its assessments are realized, the central bank takes the market interpretation of policy decisions as given and wants to depart from the rule to stabilize the economy as it currently sees it. Commitment improves outcomes by shaping the market's interpretation to limit the propagation of assessment errors. This differs from the familiar benefits of correcting an inflation bias (Barro and Gordon, 1983) or committing to future policy responses to improve current outcomes (Woodford, 1999). Our baseline model isolates our mechanism since there is no inflation and later periods' policy is discretionary, even in our commitment regime. All these mechanisms, including ours, are nevertheless rooted in the forward-looking behavior of the private sector, a general insight artfully articulated by Kydland and Prescott (1977) and Calvo (1978).

Empirical analysis. We first measure how strongly market expectations respond to policy news, a central determinant of the quantitative importance of our mechanism. For this, we use high-frequency changes in one-year OIS rates around monetary-policy events to measure policy-path surprises and changes in inflation-swap rates to measure expectations revisions. We find substantial declines in expected inflation following positive policy surprises.[2]

[2]Such responses can also arise in full-information New Keynesian models following persistent exogenous shocks to a reduced-form policy rule, such as a Taylor rule. In our framework, the central bank chooses policy optimally given its real-time assessments, and policy surprises reveal information about those assessments.

We then compare expectations sensitivity and policy-surprise volatility across periods. During 2010–2019, expectations responded much more strongly per basis point than before the Global Financial Crisis or during 2022–2026. The standard deviation of policy-path surprises was also approximately 60% lower than in either comparison period. The model predicts this combination when the central bank proportionally scales down its current responses to its assessments, holding future policy responses fixed. Intuitively, in that case a given policy surprise signals a larger gap between the central bank's assessments and what the market expected, inducing larger belief revisions. A natural explanation is that the zero lower bound constrained the Fed's main policy instrument during much of this period. With less scope to adjust the federal funds rate, the Fed's overall response to its assessments may have become weaker. Importantly, this scale effect differs from the change in relative responses central to our commitment mechanism: responding less to natural output relative to demand changes which assessments the market attributes a policy surprise to.

Quantification. To assess the importance of our core mechanism, we extend our baseline framework to an otherwise standard infinite-horizon New Keynesian model in which the central bank updates its assessments each period using private noisy signals about demand, natural output, and cost-push shocks.

Since in our framework the central bank's private information concerns its own views of the economy, we can discipline the precision of those assessments using direct evidence from revisions to the Federal Reserve staff's real-time estimates of output growth and inflation. These estimates are released to the public with a five-year lag, consistent with the model's assumption that the market does not directly observe the central bank's assessments when policy is set. We further discipline the model with U.S. macroeconomic data, a model-based measure of uncertainty about natural output from Holston et al. (2023), and the inflation-swap response to policy surprises documented above. We estimate the model under discretion.

At the estimated parameters, optimal linear commitment reduces the central bank's stationary loss by 89% relative to discretion.[3] Although policy depends on history under both regimes, almost all of this improvement is captured by a simple rule responding only to assessments of current fundamentals. Both commitment policies respond much less aggressively to assessments of natural output.

[3]The precise magnitude of these gains is naturally tied to the model's structure, which abstracts from other potentially important features of monetary transmission. Furthermore, actual policy may be "less discretionary" than our estimation assumes due to reputation or institutional arrangements. Consequently, we take the size of the gains as a benchmark for the mechanism's quantitative importance, not as a literal measure of the scope for improvement in U.S. monetary policy.

The large commitment gains and the corresponding weaker response to natural output have a common source. Natural output is highly persistent and difficult to assess in real time, so errors in the central bank's assessments are corrected slowly. When the market attributes a policy surprise to such an error, it revises its expectations of policy well into the future, sharply amplifying the effects on current output and inflation. This can happen even when the surprise was caused by a short-lived error in the central bank's demand or cost-push assessments. Responding less strongly to natural output relative to demand reduces the share of a given policy surprise that the market attributes to these persistent errors, dampening the resulting fluctuations.

A counterfactual exercise illustrates the importance of market interpretation. We replace only the discretionary rule's coefficients on natural-output assessments with those of the simple commitment rule. Following an overestimate of demand pressures, both rules generate the same realized interest-rate path: the natural-output assessment is unaffected, and both respond identically to demand assessments. The market does not observe which assessment caused the tightening. Under the alternative rule, it attributes less of it to pessimism about productive capacity and revises expected future policy less sharply. Thus, while the realized policy path is the same, the future path the market anticipates after observing each decision changes. As a consequence, the corresponding output-gap contraction on impact is 81% smaller than under discretion. By contrast, in a broad class of models used in macroeconomic policy analysis, policy rules that generate the same interest-rate response to a given shock also generate the same macroeconomic response (McKay and Wolf, 2023). This difference is rooted in asymmetric information between the central bank and the market. Our analysis illustrates how abstracting from such information asymmetries—pervasive in practice—can be quantitatively consequential.

Discussion and related literature. This paper is thematically related to several strands of literature. First, the paper contributes to the literature on monetary policy and time inconsistency, originating with Kydland and Prescott (1977) and Calvo (1978) and subsequently developed by Barro and Gordon (1983). The New Keynesian literature further showed that commitment can improve stabilization even without an inflationary bias, overcoming what has come to be known as "stabilization bias" (see, e.g., Clarida et al. (1999); Woodford (1999)). Our model isolates a different role for commitment: the current policy rule shapes what the market infers about the central bank's views of the economy and thus how errors in those views affect macroeconomic variables.

This informational source of inefficiency has antecedents in the broader "signal-jamming" literature, where incentives to affect other agents' beliefs via payoff-relevant actions give rise

to inefficient allocations (for early expressions of this idea, see Holmstrom (1982); Stein (1989)). Closer to our paper, Stein and Sunderam (2018) and Jia (2023) study monetary policy problems featuring this signal-jamming logic. Stein and Sunderam (2018) study a central bank with private information about its desired target rate and a preference for low bond-market volatility. Its incentive to adjust policy rates gradually makes the market respond more strongly to each adjustment, generating an inefficiency relative to the ex-ante optimal policy. Jia (2023) shows how a perfectly informed central bank can improve stabilization by changing what its decisions reveal about natural-rate and cost-push shocks to an imperfectly informed private sector.

Our paper identifies a different time inconsistency rooted in the central bank's imperfect and private real-time assessments of the state of the economy. We show that it arises in a model with standard central-bank stabilization objectives and assess its quantitative importance by embedding our mechanism in an otherwise standard New Keynesian model. Since the central bank's private information regards its own real-time assessments, revisions to Federal Reserve staff forecasts provide direct evidence on the errors underlying this inefficiency. The estimated model shows that market expectations amplify the macroeconomic effects of assessment errors substantially more under discretion than under the optimal ex ante rule.

Second, this paper contributes to the literature on the signaling role of monetary policy. Related work studies how policy actions reveal fundamentals or policy objectives, shaping firms' pricing and inflation expectations (see, e.g., Baeriswyl and Cornand (2010); Mertens (2016); Melosi (2017)). Our framework assumes no superior central-bank information about fundamentals: the central bank knows its own real-time views of the economy, which the market must infer from policy decisions.[4] Relatedly, Caballero and Simsek (2022) study monetary policy when the central bank and the market disagree about the economic outlook. They also extend their analysis to allow the market to infer the central bank's views from its policy decisions. There, however, systematic policy does not affect how the market attributes policy surprises to different components of the central bank's outlook. In our model, the policy rule shapes that attribution and thus the propagation of central-bank assessment errors. This dependence is crucial for the time inconsistency at the core of our analysis.

Third, this paper relates to previous work on optimal monetary policy under imperfect central-bank information (e.g., Orphanides (2001); Aoki (2003); Svensson and Woodford (2003, 2004)). In these analyses, however, imperfect central-bank information does not create

[4] Bauer and Swanson (2023) report survey evidence that some forecasters view policy surprises as news about future policy, not about fundamentals (since they do not believe the Fed to have superior information). Relatedly, Bauer et al. (2024) document revisions in perceptions of future policy following policy surprises.

a persistent inference problem for the market.[5] In our model, by contrast, the central bank's assessments remain private and persistent, and the market must infer them from policy decisions. In turn, systematic policy determines how informative interest rates are and thus affects the strength of monetary transmission itself. This provides a distinct rationale for the caution toward noisy real-time estimates emphasized by Orphanides (2001, 2003): weaker responses to optimally formed assessments can improve stabilization by changing market interpretation and thus limiting the propagation of assessment errors.

Relatedly, our quantitative analysis has a precedent in Aoki (2006), which measures the gains from commitment in the U.S. using a forward-looking model with imperfect central-bank information. Since that model does not feature private central-bank information, its source of time inconsistency is different. Unlike our paper, it finds small commitment gains.[6] Lippi and Neri (2007) also estimate a model with imperfect information using euro-area macroeconomic data, assuming that the central bank and the private sector share the same information. Information asymmetry is central to our mechanism: the market must infer the central bank's assessments from its policy decisions. Furthermore, our estimation uses revisions to Federal Reserve real-time forecasts to discipline the precision of those assessments.

Roadmap. The remainder of the paper is organized as follows. Section 2 introduces the environment and beliefs. Section 3 studies how market inference shapes transmission and characterizes discretion and commitment. Section 4 presents the empirical evidence. Section 5 develops and estimates the extended New Keynesian model and quantifies the mechanism and commitment gains. Section 6 concludes. The appendices provide the baseline model's foundations, proofs, additional results, and supporting material for the empirical and quantitative analyses.

2 Model

This section lays out the baseline model. First, we describe the physical environment: a very simple economy with rigid prices. Then, we discuss the features of the information structure that underpin the analysis.

[5] In Svensson and Woodford (2004), contemporaneously observed endogenous variables reveal the central bank's information relevant for the equilibrium. In Aoki (2003), the market does not directly observe the central bank's noisy indicators, but the private component of its assessment is transitory and has no implications for future policy. For more extended discussions of these, see Appendix G.1.

[6] In Aoki's model, the central bank observes fundamentals perfectly with a one-period delay. This can naturally limit the consequences of the imperfect information to a significant extent.

2.1 Environment

Consider a linear economy where prices are fully rigid and thus output is demand-determined.[7] In particular, output evolves as

$$y_t = \mathbb{E}_t^M y_{t+1} - \gamma(r_t - \rho_t), \tag{1}$$

where $\gamma > 0$ is the elasticity of intertemporal substitution, r_t the interest rate, ρ_t a demand shifter, and $\mathbb{E}_t^M$ the market's period-t expectation.

The central bank chooses the policy stance r_t to stabilize output around its natural level y_t^n,

$$\min_{\{r_t\}} \quad \mathbb{E}^C \sum_{t=0}^{\infty} \beta^t (y_t - y_t^n)^2, \tag{2}$$

where $\beta \in (0,1)$ is the discount factor and $\mathbb{E}^C$ the central bank's ex ante expectation. The two fundamentals, natural output y_t^n and the demand shifter ρ_t, evolve as

$$y_1^n = \varphi_n y_0^n + \varepsilon_n, \qquad y_t^n = 0 \text{ for } t \geq 2, \tag{3}$$

and

$$\rho_1 = \bar{\rho} + \varphi_\rho(\rho_0 - \bar{\rho}) + \varepsilon_\rho, \qquad \rho_t = \bar{\rho} \text{ for } t \geq 2, \tag{4}$$

with $\varphi_n \in (0,1]$, $\varphi_\rho \in [0,1)$, and $\varphi_n \geq \varphi_\rho$: natural output is at least as persistent as demand. The $t = 1$ innovations $(\varepsilon_n, \varepsilon_\rho)$ are mean-zero, Gaussian, and mutually independent.

2.2 Beliefs

This subsection presents the relevant features of the information structure. We specify the central bank's and market's conditional beliefs directly. A detailed heterogeneous-prior foundation and discussion of what a common prior would imply are provided in Appendix B.

Since it is common knowledge that $(y_t^n, \rho_t) = (0, \bar{\rho})$ in periods $t \geq 2$, we focus on periods where fundamentals are stochastic, $t \in \{0, 1\}$.

[7] The model can be obtained as an approximation of a standard nonlinear economy with households, firms and a monetary authority (see Appendix A for details).

Central bank beliefs. The central bank's $t = 0$ information is summarized by a vector of imperfect assessments of the state of the economy, $(\widehat{y}_0^n, \widehat{\rho}_0)$. The central bank's conditional distribution of the fundamentals is given by

$$\begin{pmatrix} y_0^n \\ \rho_0 \end{pmatrix} \,\Big|\, \begin{pmatrix} \widehat{y}_0^n \\ \widehat{\rho}_0 \end{pmatrix} \overset{C}{\sim} \mathcal{N}\left(\begin{pmatrix} \widehat{y}_0^n \\ \widehat{\rho}_0 \end{pmatrix}, \begin{pmatrix} \sigma_{\eta,C}^2 & 0 \\ 0 & \sigma_{\varepsilon,C}^2 \end{pmatrix} \right), \tag{5}$$

so the assessments correspond to the central bank's best estimates of the fundamentals. The parameters $\sigma_{\eta,C}^2, \sigma_{\varepsilon,C}^2 > 0$ denote the corresponding (conditional) variances. For later use, we define the central bank's assessment errors as

$$\eta_0 \equiv y_0^n - \widehat{y}_0^n, \qquad \varepsilon_0 \equiv \rho_0 - \widehat{\rho}_0.$$

We also state the central bank's ex ante beliefs (i.e., before observing its $t = 0$ assessments), since they are relevant for the policy problem in Section 3.4. Specifically,

$$\begin{pmatrix} \widehat{y}_0^n \\ \widehat{\rho}_0 \end{pmatrix} \overset{C}{\sim} \mathcal{N}\left(\begin{pmatrix} 0 \\ \bar{\rho} \end{pmatrix}, \begin{pmatrix} \sigma_{n,C}^2 & 0 \\ 0 & \sigma_{\rho,C}^2 \end{pmatrix} \right), \tag{6}$$

where $\sigma_{n,C}^2, \sigma_{\rho,C}^2 > 0$. Note that the two assessments are independent under the central bank's ex ante beliefs.

We assume the central bank gets no additional information between its policy decisions at $t = 0$ and $t = 1$.[8] This implies

$$\widehat{y}_1^n \equiv \mathbb{E}_1^C[y_1^n] = \varphi_n \widehat{y}_0^n, \qquad \widehat{\rho}_1 \equiv \mathbb{E}_1^C[\rho_1] = \bar{\rho} + \varphi_\rho(\widehat{\rho}_0 - \bar{\rho}), \tag{7}$$

so the parameters φ_n and φ_ρ also denote the persistence of the central bank's assessments.

Market beliefs. The market observes the value of contemporaneous fundamentals (y_t^n, ρ_t) and knows their laws of motion, but it does *not* observe the central bank's estimates. Under the market's $t = 0$ beliefs (i.e., conditional on $t = 0$ fundamentals), the estimate vector

[8] This assumption has two elements. First, the fact that no additional *exogenous* signals are available to the central bank is a harmless simplification that avoids introducing additional parameters that do not change the qualitative insights of the paper (this is relaxed in the quantitative model where the central bank gets signals each period). Second, the fact that the central bank does not observe output y_0 is more substantial. Appendix B.1 discusses it in detail.

$(\widehat{y}_0^n, \widehat{y}_1^n, \widehat{\rho}_0, \widehat{\rho}_1)$ is jointly Gaussian such that $(\widehat{y}_0^n, \widehat{y}_1^n) \perp (\widehat{\rho}_0, \widehat{\rho}_1)$, consistent with the independent nature of $t = 0$ fundamentals and (7). Concretely, market beliefs regarding the central bank's $t = 0$ assessments are such that[9]

$$\begin{pmatrix} \widehat{y}_0^n \\ \widehat{\rho}_0 \end{pmatrix} \overset{M}{\sim} \mathcal{N}\left(\begin{pmatrix} \lambda_n y_0^n \\ \rho_0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & s \end{pmatrix} \right), \tag{8}$$

where $\lambda_n \in [0, 1)$ and $s > 0$. The fundamentals' $t = 1$ innovations $(\varepsilon_n, \varepsilon_\rho)$ are independent of the central bank's $t = 0$ assessments under the market's beliefs.

Note that the above entails an asymmetry, since

$$\mathbb{E}_0^M[\widehat{y}_0^n] = \lambda_n y_0^n, \qquad \mathbb{E}_0^M[\widehat{\rho}_0] = \rho_0. \tag{9}$$

Thus, the market expects the central bank's demand estimate to be centered on realized demand and its natural-output estimate to be a weighted average of realized natural output and its unconditional mean, zero.[10]

The parameter s introduced in (8) represents the market's *relative* uncertainty about the two estimates:

$$s \equiv \frac{\mathrm{Var}_0^M(\widehat{\rho}_0)}{\mathrm{Var}_0^M(\widehat{y}_0^n)} \in (0, \infty), \tag{10}$$

which we call the *interpretation state*. As will be apparent in the analysis, this parameter governs the signal-extraction problem a rate surprise poses: a higher s leads the market to read a given surprise more as news about the central bank's demand assessment and less as news about its natural-output assessment. Throughout, s is fixed and common knowledge.[11]

We assume the market observes contemporaneous fundamentals to isolate learning about the central bank's assessments, which matter in their own right because they determine future policy. This assumption is stronger than needed: if the central bank observes noisy versions of the market's signals and both agents regard the market's remaining estimation

[9] The normalization $\mathrm{Var}_0^M(\widehat{y}_0^n) = 1$ is innocuous, since only the *relative* uncertainty about the two assessments plays a role in the model.

[10] The demand centering is a knife-edge case of the more general belief structure in Appendix B. We adopt it for analytical simplicity: it delivers a unit discretionary coefficient on demand assessments. The appendix also explains why a standard proper common-prior information structure instead produces interior centerings (i.e., the loading of $\mathbb{E}_0^M[\widehat{\rho}_0]$ on ρ_0 is strictly between 0 and 1). The quantitative model instead derives interior centerings from its common-prior Gaussian signal structure.

[11] Jimenez-Gimpel (2026) studies the policy implications of central bank uncertainty about s, and hence about how markets interpret its decisions. There, it is shown that this form of transmission uncertainty induces caution in the discretionary equilibrium. Since systematic policy shapes the extent of uncertainty, this caution entails additional margins of inefficiency.

errors as independent (and mean zero), the analysis carries over with market estimates replacing fundamentals. If, on the other hand, policy decisions also reveal information about fundamentals to the markets, the systematic rule shapes that inference as well, so our mechanism of interest would still be at play.

3 Equilibrium

This section shows how systematic policy shapes the market interpretation of policy decisions and why this generates a time inconsistency: a central bank that sets policy meeting by meeting leaves the economy too exposed to volatility driven by its assessment errors. First, we characterize the equilibria at $t \geq 1$, showing that the central bank's persistent assessments shape y_1. Then, we turn to linear equilibria at $t = 0$ and present the key mechanism of the model: how policy decisions affect market beliefs and thus monetary transmission. We then characterize the discretionary equilibrium and the optimal ex ante commitment policy, highlighting the key difference: discretion takes market interpretation as given, whereas commitment internalizes its dependence on policy and thus reduces the propagation of assessment errors and, under conditions characterized below, the sensitivity of market beliefs.

3.1 Continuation Equilibrium

We first characterize allocations from $t = 1$ onward, when the central bank sets policy under discretion: at each date, it chooses the rate that minimizes its expected loss given its information. This continuation equilibrium is common to both policy regimes considered below. The following lemma characterizes the resulting policy and allocation.

Lemma 1 (Continuation equilibrium)**.** *For all $t \geq 2$, the central bank implements its first best, i.e., $y_t = y_t^n = 0$. At $t = 1$, it sets $r_1 = \widehat{\rho}_1 - \widehat{y}_1^n/\gamma$ to achieve*

$$y_1 = \widehat{y}_1^n + \gamma(\rho_1 - \widehat{\rho}_1). \tag{11}$$

Because all uncertainty is resolved by period 2, and the central bank has one tool (r_t) for one target (y_t) each period, its problem is straightforward and it achieves its first best for $t \geq 2$.

On the other hand, as is standard in linear-quadratic environments, optimal discretionary policy satisfies certainty equivalence conditional on the central bank's estimates. Thus, at

$t = 1$ the central bank simply implements its target output *on average*, i.e., it chooses policy r_1 to set $\mathbb{E}_1^C[y_1] = \widehat{y}_1^n$.

Crucially, the market anticipates this when forming its expectations at $t = 0$: it knows that at $t = 1$, output will be determined by both the actual demand shifter and the central bank's $t = 1$ assessments through the policy decision. Since the market does not observe those assessments and they are persistent, the market uses the $t = 0$ policy decision to learn about the central bank's contemporaneous assessments and thus about $t = 1$ output.

These revisions in expected y_1 affect current output through the IS equation, making the market's anticipation of future policy part of current monetary transmission. The next subsection develops this idea.

3.2 The Role of Market Expectations: Policy as a Signal

Timing. Figure 1 lays out the sequence of events at $t = 0$. At the start of the period, the fundamentals (y_0^n, ρ_0) and the central bank's assessments $(\widehat{y}_0^n, \widehat{\rho}_0)$ are realized. The market observes the fundamentals but not the assessments, while the central bank observes only its assessments. The central bank then sets the policy rate r_0. Observing it, the market updates its beliefs about the central bank's assessments—and hence about next period's output—and y_0 is determined.

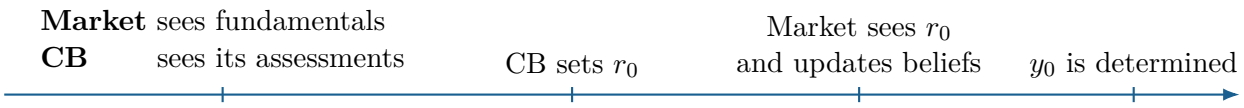

Figure 1: Timeline of events at $t = 0$.

Linear policy rules and policy transmission. We consider policy rules that are linear in the central bank's assessments,[12]

$$r_0 = \bar{\rho} + a\Big(\widehat{\rho}_0 - \bar{\rho}\Big) - K\widehat{y}_0^n, \tag{12}$$

for some $(a, K) \in \mathbb{R}^2$. For interpretation purposes, it is useful to reparameterize each rule with $a \neq 0$ by writing $K = ab$ and thus

$$r_0 = \bar{\rho} + a\Big(\widehat{\rho}_0 - \bar{\rho} - b\widehat{y}_0^n\Big), \qquad a \in \mathbb{R} \setminus \{0\}, \quad b \in \mathbb{R}. \tag{13}$$

[12]The choice of the intercept is without loss of generality, since both in the discretionary equilibrium and the optimal commitment it equals $\bar{\rho}$. This is discussed in more detail in Remark 4.

The *scale* parameter a determines the overall aggressiveness of the policy response, while the *composition* parameter b pins down the relative weight the policy rule places on the natural-output assessment. The case $a = 0$ is handled separately: if $K \neq 0$, the central bank responds only to its natural-output assessment, while if $K = 0$, the interest rate is completely unresponsive to the central bank's assessments.

As shown in (11) and mentioned above, $t = 1$ output is a function of the central bank's $t = 1$ assessments. Since the market anticipates this and assessments are persistent, the market learns about future output from the $t = 0$ policy decision, r_0. The following proposition characterizes the market's inference.

Proposition 1 (Belief sensitivity and policy transmission)**.** *For every policy rule* (a, K) *with* $b \equiv K/a$ *whenever* $a \neq 0$*, the market's beliefs following a policy decision* r_0 *are such that*[13]

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \begin{cases} \varphi_n \lambda_n y_0^n - \chi(b,s)\left(\dfrac{r - \mathbb{E}_0^M[r_0]}{a}\right), & \text{if } a \neq 0, \\ \varphi_n \lambda_n y_0^n - \left(\dfrac{\varphi_n}{K}\right)(r - \mathbb{E}_0^M[r_0]), & \text{if } a = 0, K \neq 0, \\ \varphi_n \lambda_n y_0^n, & \text{if } a = K = 0, \end{cases} \tag{14}$$

where

$$\chi(b,s) = \frac{\varphi_n b + \gamma \varphi_\rho s}{b^2 + s}. \tag{15}$$

Consequently, given a rate decision $r_0 = r$*, output is given by*

$$y_0 = \begin{cases} \varphi_n \lambda_n y_0^n - \gamma(r - \rho_0) - \chi(b,s)\left(\dfrac{r - \mathbb{E}_0^M[r_0]}{a}\right), & \text{if } a \neq 0, \\ \varphi_n \lambda_n y_0^n - \gamma(r - \rho_0) - \left(\dfrac{\varphi_n}{K}\right)(r - \mathbb{E}_0^M[r_0]), & \text{if } a = 0, K \neq 0, \\ \varphi_n \lambda_n y_0^n - \gamma(r - \rho_0), & \text{if } a = K = 0. \end{cases} \tag{16}$$

First, note that if $a = 0$ and $K \neq 0$, the policy decision fully reveals the central bank's natural output assessments $\widehat{y}_0^n$, while providing no information about its demand assessments $\widehat{\rho}_0$. In this case, the market's forecast of $t = 1$ output equals the central bank's $t = 1$ natural-output assessment, since the remaining demand term has conditional mean zero. To recover that, the market simply subtracts the common-knowledge component of the rate, $\bar{\rho}$, and scales down the remainder using the conjectured rule. If instead $(a, K) = (0, 0)$, the policy rate is constant and conveys no information. Thus, the market's forecast then remains equal to its prior.

[13] Under the proper information structure in Appendix B.1, the market's beliefs are shaped by signals and thus the expectations should be interpreted as conditional on those as well as the policy rate.

Now consider $a \neq 0$. Then, the policy rate reveals the combination $\widehat{\rho}_0 - b\widehat{y}_0^n$. When $b \neq 0$, the market cannot recover either assessment separately, and it effectively solves a signal-extraction problem to improve its estimates of the central bank's $t = 1$ assessments (and thus its forecast of y_1) from the policy surprise $r_0 - \mathbb{E}_0^M[r_0]$.[14] Due to the Gaussian nature of all the random variables involved, the revision in expected output is linear in the policy surprise, with slope $-\chi(b, s)/a$. We call $\chi(b, s)$ the *sensitivity of market beliefs*. It represents the market's revision of expected future output in response to a policy surprise normalized by the scale a of the policy rule. Throughout the theoretical analysis, "belief sensitivity" refers to this normalized coefficient. Note that $\chi(b, s)$ is independent of the scale a, since the market scales its updating to exactly offset the overall size of any given policy reaction. For a given realization of assessments, changing the (nonzero) scale a simply rescales the rate signal without changing the information it conveys. It does depend, however, on the composition b and the interpretation state s. Both elements are crucial, since they shape how much the market learns about both central bank assessments, and thus determine the overall impact of policy on output. This is the central object of the paper.

This distinction yields a simple empirical prediction. Holding policy composition b and the parameters fixed, a lower positive scale a reduces the size of policy surprises while increasing the magnitude of the belief response per basis point, $|\chi(b, s)|/a$.[15] Intuitively, when policy responds (systematically) less to the central bank's assessments, the same rate surprise induces a larger revision in market beliefs about those assessments. A lower policy scale thus generates smaller surprises and larger belief revisions per basis point, even though the information conveyed by policy is unchanged. The empirical analysis in the next section examines this pattern.

To anchor intuition, consider the case where demand assessments are fully transitory, $\varphi_\rho = 0$. Then,

$$\chi(b, s) = \frac{\varphi_n b}{b^2 + s}.$$

First, note that χ is decreasing in s in the empirically relevant case where $b > 0$. This is because the greater the share of the market's prior uncertainty that concerns demand estimates, the more it attributes a given rate surprise to news about them, which is fully transitory and thus uninformative about future output. In contrast, χ is hump-shaped in b. The reason behind this non-monotonicity is that increasing b changes both the volatility of the rate surprise and the informativeness of a given surprise. When b is large, the former

[14] This inference is generically imperfect. When $b = \varphi_n/(\gamma\varphi_\rho)$, for $\varphi_\rho > 0$, the rate reveals exactly the combination of assessments that determines next period's policy (even though neither assessment is separately revealed). Since $\varphi_n \geq \varphi_\rho$, however, that b is never attained in a discretionary equilibrium as Proposition 2 shows.

[15] Naturally, this comparison takes the continuation policy characterized in Lemma 1 as fixed.

effect dominates the latter, and thus the overall update is dampened. In turn, when b is small, the additional co-movement between rates and future output is stronger than the added volatility.

Since this belief sensitivity χ depends on the composition b, *the central bank's own systematic reaction shapes the strength of monetary transmission.* This observation drives everything that follows. Figure 2 summarizes the channel.

Note that fully transitory demand assessments isolate the intuition but are not necessary for the mechanism. In fact, even if $\varphi_n = \varphi_\rho$, generically the rate remains an imperfect signal of next period's policy (see footnote 14 for the knife-edge exception).

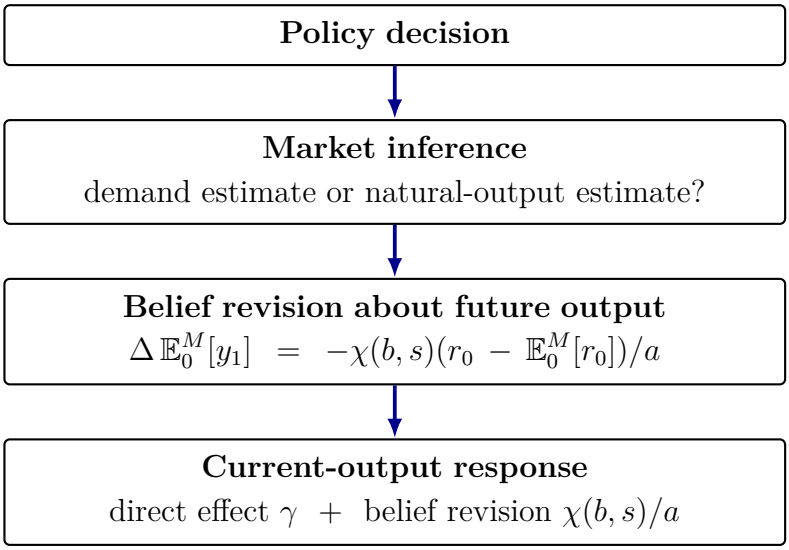

Figure 2: The belief-revision channel.

On the interpretation of the model. Like any model, ours abstracts from many details to isolate our key mechanism of interest. Two of these simplifications warrant further discussion to clarify the interpretation of the results.

1. On the private-sector side, the model summarizes the forward-looking elements of the economy in households' expectations of future output. Through a standard consumption–savings decision, these expectations shape current consumption demand and, since prices are fully rigid, current output (Appendix A). Neither the belief-revision mechanism nor its endogeneity to systematic policy hinges on this particular representation. In richer and more realistic models, the relevant expectations may be formed by financial-market participants and embedded in longer-term rates, mortgage and credit conditions, and asset valuations. Revisions in these expectations would then shape the spending and investment decisions of households and firms and,

through them, macroeconomic outcomes. The same belief-revision mechanism—and the dependence of its strength on systematic policy—would thus carry over to such environments.

2. On the policy side, the model collapses the rich, multidimensional set of decisions involved in conducting monetary policy into a single scalar choice: the current policy rate. The mechanism of interest, however, is not exclusive to such literal policy-rate choices. Statements, press conferences, minutes, and speeches are read by market participants as signals of the central bank's assessment and, consequently, shift the expected policy path, current financial conditions, and ultimately macroeconomic activity. In fact, a direct implication of the core feature of our framework—the fact that policy is a signal—is precisely that rate-setting and communication are not so different after all. This becomes especially apparent in the scale-invariance of $\chi(b, s)$ above, which is essentially the formal counterpart of the phenomenon described by Stein (2014) in the context of central-bank communication: *the softer* [the central bank] *talks, the more the market leans in to hear better and, thus, the more the whisper gets amplified.* This, then, is a particular implication of the broader mechanism we study, namely, the endogeneity of the sensitivity of market beliefs—and hence, of policy transmission—to the systematic conduct of monetary policy.

3.3 Discretionary Equilibrium

We first focus on *discretionary* equilibria at $t = 0$, namely, equilibria in which the central bank sets $t = 0$ policy taking the market's belief map (Proposition 1) as given. This is the outcome that follows if the central bank cannot commit to a rule before observing its $t = 0$ assessments.[16]

Formally, fix the *market's conjectured* linear rule and the market belief map it induces. After observing its assessments, the central bank solves

$$\arg\min_{r \in \mathbb{R}} \mathbb{E}_0^C \left[(y_0 - y_0^n)^2 \right] \tag{17}$$

subject to (16). Crucially, the function $r \mapsto \mathbb{E}_0^M[y_1 \mid r_0 = r]$ in (16) (or, equivalently, the parameters (a, b)) is taken as given: the central bank does not internalize the fact that its policy reaction changes the coefficients (a, b) and thus policy transmission. Loss from date 1 onward is omitted because, by Lemma 1, it is independent of the current rate.

The equilibrium concept is the natural one, a fixed point in which the central bank's

[16] The reason behind this time inconsistency will become clear in Section 3.4.

optimal reaction to its assessments coincides with the market's conjectured rule, and thus the market's beliefs are consistent with the central bank's actual rule. The following definition formalizes this notion.

Definition 1 (Linear discretionary equilibrium)**.** *A policy rule $r_0(\widehat{y}_0^n, \widehat{\rho}_0)$ of the form (12), together with its associated state-contingent output allocation y_0, is a linear discretionary equilibrium if: (i) the market beliefs induced by the rule satisfy (14); (ii) for every realization of the central bank's assessments, the rate prescribed by the rule solves (17), taking the induced belief map as given; and (iii) output satisfies the IS curve (1).*

For the interested reader, Appendix C recasts the economy as a game and defines the discretion and commitment solution concepts in more transparent game-theoretic terms.

Unless stated otherwise, *discretionary equilibrium* refers to an equilibrium in which the central bank's $t = 0$ rate-choice problem has a unique solution. This rules out a knife-edge case, namely, one in which the direct and the belief-revision effects of a rate adjustment cancel out exactly and thus output is independent of the policy rate and the central bank is indifferent among all rate choices (even though only one validates this as an equilibrium). We exclude these indifference cases from the analysis since they are economically uninteresting and empirically irrelevant.

Proposition 2 (The unique linear discretionary equilibrium)**.** *There exists a unique linear discretionary equilibrium. It is such that $b_d \in (0, (1 - \lambda_n\varphi_n)\gamma^{-1})$, and satisfies*

$$a_d = 1, \qquad b_d(\gamma + (1 - \lambda_n)\chi(b_d, s)) = 1 - \lambda_n\varphi_n. \tag{18}$$

Intuitively, due to the linear-quadratic nature of the central bank's problem, certainty equivalence holds: taking as given the sensitivity of expectations to rate movements, the central bank adjusts its policy reaction to minimize average inefficient output fluctuations, considering both the direct effect of the current rate (the change in the price of consumption today relative to consumption tomorrow) and the belief-revision effect on expectations of next-period output. This (conditional) expected output gap is given by

$$\mathbb{E}_0^C[y_0 - y_0^n] = [b(\gamma a + (1 - \lambda_n)\chi(b, s)) - (1 - \lambda_n\varphi_n)]\,\widehat{y}_0^n + \gamma(1 - a)(\widehat{\rho}_0 - \bar{\rho}). \tag{19}$$

Since the belief revision effect on the demand side is mean zero, the central bank disregards that component of transmission and $a_d = 1$ follows. On the other hand, the central bank's reactions to its output target estimate do shift expectations on average, and thus it picks its policy to achieve its desired average output movements considering both components of transmission. This explains the upper bound $(1-\lambda_n\varphi_n)\gamma^{-1}$, since that would be the

desired reaction if only the standard transmission were present. Since overall transmission is stronger with the belief-revision component, a less aggressive reaction is optimal (note that the target total transmission is not one-to-one since the market's prior beliefs "do part of the work", by anticipating a future rate response of $-\lambda_n\varphi_n/\gamma$ times the natural output realization on average).

Importantly, as (19) shows, the belief-revision impact is not simply χ, since that element of transmission is discounted by what the markets already expect the policy rate to be. The natural-output component of that expected rate is $-ab\lambda_n y_0^n$, which contributes $-b\chi\lambda_n y_0^n$ to the revision in expected output. In addition, the target pass-through of natural output assessments is less than one, since a fraction $\lambda_n\varphi_n$ of it is already incorporated by the market via its ex ante beliefs about future policy and its knowledge of the current natural output realization.

Finally, equilibrium uniqueness follows from the fact that increasing the central bank's reaction coefficient b strictly increases the output response to a given natural-output assessment: the larger direct response more than offsets any weakening of the belief revision channel. Consequently, the LHS of the second equality in (18) is strictly increasing for $b > 0$, so exactly one composition coefficient delivers the desired stabilization. Note that the strength of transmission per unit of policy-rate change *is not* in general increasing in b, since χ is hump-shaped as discussed above.

3.4 The Optimal Commitment Policy

Now consider the problem where the central bank commits to a $t = 0$ linear policy rule *before observing its* $t = 0$ *assessments*. Crucially, this entails internalizing that the policy rule shapes belief sensitivity χ.

The central bank's optimal commitment policy solves

$$\arg\min_{(a,K)\in\mathbb{R}^2} \mathbb{E}^C\left[(y_0 - y_0^n)^2\right] \tag{20}$$

subject to (12) and (16), with the market belief map induced by the chosen coefficients (a, K). See Appendix C for a game-theoretic interpretation of the commitment problem and a discussion of the differences between commitment and the discretionary equilibrium above.

The source of time inconsistency can be understood by analyzing the central bank's incentives. Before observing its assessments, the central bank knows that some of them will be wrong. Consequently, it would like to commit to a systematic mapping from assessments into rates that limits how strongly those errors are propagated through market beliefs and

into output. Under the conditions developed below, the optimal rule places a lower relative weight on the natural-output assessment.

Once the assessments have been realized, however, the market's response to any rate decision is already pinned down by the systematic rule it expects. A one-time deviation cannot change that "belief map". Taking it as given, then, the central bank chooses the rate to set its conditional expected output gap to zero, accounting for both the direct and belief-revision channels of monetary transmission. Consequently, generically it does not follow the ex ante rule that was chosen to shape the market interpretation.

Remark 1 (The scope of commitment)**.** *The commitment problem above restricts policy relative to the Ramsey benchmark in two ways: (i) commitment applies only to $t = 0$ policy, not the full path (i.e., it entails no forward guidance; policy reverts to its discretionary equilibrium at $t = 1$); and (ii) the policy rule is linear in the central bank's assessments. The former is a simplification that allows us to focus on the main mechanism of interest and highlights the novel nature of the time inconsistency problem that is the object of this paper (in contrast to the standard time inconsistency problem studied, for example, in Woodford (2003)). The latter, on the other hand, is a meaningful restriction, unlike in equivalent economies without private information, where the optimal unconstrained Ramsey policy is linear. This is because of the nonlinear nature of the market's belief updating as a function of the policy rule. However, in the spirit of the literature on simple rules, this is a natural family of policies to study. Furthermore, this allows the policy to be compared with the ones in related environments without private information. The quantitative analysis in Section 5 relaxes the former restriction while keeping the latter.*

Because the rule is chosen before the central bank observes its assessments, the commitment objective integrates over its ex ante beliefs specified in Section 2.2. The following lemma provides a convenient representation of the objective function of the commitment problem.[17]

[17] Lemma 2 uses the scale–composition parameterization $b = K/a$ and therefore applies to $a \neq 0$. Policy rules with $a = 0, K \neq 0$ and the constant rule $(a, K) = (0, 0)$ are evaluated separately in raw (a, K) coordinates in Lemma 3.

Lemma 2 (Loss function)**.** *For every rule with $a \neq 0$, the central bank's ex ante loss is given by*

$$L(a,b;s) = \left(b(\gamma a + \chi(b,s)) - (1-\lambda_n\varphi_n) - \lambda_n b\chi(b,s)\right)^2 \sigma^2_{n,C} + (a-1)^2\gamma^2\sigma^2_{\rho,C} \\ + \underbrace{\left(\gamma + \chi(b,s)\right)^2 \sigma^2_{\varepsilon,C} + (1-\lambda_n\varphi_n + \lambda_n b\chi(b,s))^2\sigma^2_{\eta,C}}_{\equiv \mathcal{P}(b,s)}. \tag{21}$$

The loss has a natural conditional mean-variance interpretation. Its first two terms capture systematic stabilization errors—that is, variation in the output gap's conditional mean across realizations of the central bank's assessments. The first measures the failure of output to track the natural-output assessment one for one. The second measures the failure to shut down the effect of demand shocks on output.

The third and fourth terms, collectively denoted $\mathcal{P}(b,s)$, capture output-gap volatility caused by assessment errors. Demand-assessment errors affect current output directly, with coefficient γ, and also generate a policy surprise that changes expected future output, with coefficient $\chi(b,s)$. The resulting total pass-through is then $\gamma + \chi(b,s)$. On the other hand, natural output assessment errors entail a one-to-one impact on the output gap that is then attenuated by the ex ante market expectations of future policy, based on the actual current natural output. Furthermore, the market's expected rate is $\mathbb{E}_0^M[r_0] = \bar{\rho} + a(\rho_0 - \bar{\rho}) - ab\lambda_n y_0^n$. Seeing a high realization of natural output y_0^n that the central bank does not observe, it expects the latter to cut rates by $ab\lambda_n\eta_0$ more than it actually does. So the realized surprise rises by $ab\lambda_n\eta_0$: the central bank looks hawkish. The market reads that as news that the assessment is low, and revises expected future output down by $\chi b\lambda_n\eta_0$, which entails an output gap movement. Thus, the total pass-through of natural-output assessment errors is $1 - \lambda_n\varphi_n + \lambda_n b\chi(b,s)$.

When demand assessments are transitory, the belief revision component of the pass-through arises entirely because the market partly attributes the policy surprise to persistent natural-output news, regardless of its actual source. With persistent demand assessments, part of it also comes from the share of the surprise that the market attributes to demand-assessment news.

The following proposition presents the paper's main normative result: commitment reduces the pass-through of assessment errors into output and, under the conditions stated below, generically lowers the sensitivity of market beliefs.[18]

Proposition 3 (The Optimal Commitment Policy)**.** *Let (a_c, K_c) be a solution of the commitment problem (20) and suppose $a_c \neq 0$. Let $b_c \equiv K_c/a_c$ and set*

$$m \equiv 1 - \frac{\varphi_n(1+\lambda_n)}{2}.$$

Then,

- $\mathcal{P}(b_c, s) \leq \mathcal{P}(b_d, s)$*;*
- *if either $\lambda_n = 0$ or*

$$\gamma^2 s\left(\varphi_n - 2\varphi_\rho m\right) \;\geq\; \varphi_n m^2, \tag{22}$$

then

$$\chi(b_c, s) \leq \chi(b_d, s) \tag{23}$$

and the inequality is strict whenever $\lambda_n > 0$. Further, whenever $\lambda_n = 0$, the inequality is also strict except possibly when (22) holds with equality and $\varphi_n > (2-\varphi_n)\varphi_\rho$.

The sensitivity comparison in Proposition 3 concerns $\chi(b, s)$, which, as discussed above, corresponds to the belief sensitivity to a policy surprise normalized by the scale of the current rule. It does not establish any comparison regarding the size of the response per basis point, $|\chi(b, s)/a|$.[19]

When $\lambda_n = 0$, commitment reduces the sensitivity of beliefs and consequently weakens the amplification of demand-assessment errors. The proposition extends this result to $\lambda_n > 0$ under an explicit sufficient condition that aligns the incentives to limit the propagation of demand and natural-output assessment errors.

[18]Note that the proposition does not assert the existence of a commitment optimum. Indeed, this does not hold in general due to the discontinuity at the constant policy rule. Intuitively, at every rule with $a \neq 0$, the market learns about the central bank's assessments from the policy decision (and this learning is in fact independent of the value of a as illustrated in Proposition 1). However, at $(a, K) = (0, 0)$, the market learns nothing. This discontinuity can create a situation in which the central bank can always reduce its loss by choosing a nonzero a closer to zero, without attaining the infimum. Lemma 3 provides a sufficient condition for existence of a commitment optimum. Furthermore, an optimum with $a_c = 0$ is possible and has a different economic interpretation: if $K_c \neq 0$, the central bank responds only to its natural-output assessment, while if $K_c = 0$, the rule is constant and conveys no information. These cases are not empirically relevant and thus are not considered in the analysis from here onward.

[19]In the small-error approximation of Corollary 1, Proposition 4 additionally establishes $a_c > a_d = 1$. Thus, under those conditions, the size of the response per basis point also falls.

At the discretionary rule, the first two terms of L are exactly zero: conditional on its assessments, the central bank sets its expected output gap to zero as detailed above. A small departure from that rule therefore creates expected stabilization losses only to second order. By contrast, changing the rule can generate a first-order reduction in output volatility caused by assessment errors $\mathcal{P}$. Commitment internalizes this margin, whereas a discretionary bank takes assessment-error volatility as independent of its current rate choice.

Intuitively, policy affects the propagation of assessment errors by shaping both the sensitivity of market beliefs and the policy surprises generated by these errors. Demand-assessment errors propagate to output with coefficient $\gamma + \chi(b, s)$, so reducing their amplification calls for lower belief sensitivity. When $\lambda_n > 0$, natural-output assessment errors propagate to the output gap with coefficient $-[1 - \lambda_n\varphi_n + \lambda_n b\chi(b, s)]$. Concretely, this propagation depends not only on the sensitivity of market beliefs χ but also directly on the rule's composition b: a lower b makes expected policy less dependent on realized natural output, reducing the surprise generated when the central bank's assessment differs from that realization. When these two channels conflict, commitment may reduce overall assessment-error propagation while increasing belief sensitivity.

Condition (22) aligns these two mechanisms. To understand it, recall that χ is (minus) the covariance-to-variance ratio relating future output to the scale-free policy surprise. Condition (22) first requires $\varphi_n > 2\varphi_\rho m$. This makes the natural-output component strong enough relative to the persistent demand component that increasing b can raise belief sensitivity. The accompanying lower bound on $\gamma^2 s$ then ensures that the increase in covariance dominates the increase in signal variance, making χ locally increasing in b. The condition therefore places the discretionary composition on the increasing portion of $\chi(\cdot, s)$, so lowering b reduces both χ and $b\chi$. Naturally, this condition is sufficient, not necessary: the motive to lower the sensitivity of beliefs can dominate even when the two channels are not aligned. Indeed, our numerical exercise shows that commitment lowers belief sensitivity well beyond the parameterizations where the sufficient condition holds.[20]

[20] A sweep over 32,400 parameter configurations shows that $\chi(b_c, s) > \chi(b_d, s)$ in only 3.3% of cases, even though the condition above holds in only 57%. Further, for the most part whenever the ranking reverses it does so by very little: $\chi(b_c, s)/\chi(b_d, s)$ exceeds 1.01 in 1.47% of all configurations and 1.05 in only 0.66%; plus, the median ratio among the cases where $\chi(b_c, s)/\chi(b_d, s) > 1$ is 1.0068. The parameters considered are $\gamma \in \{0.3, 0.5, 1\}$, $\varphi_n \in \{0.9, 0.98, 1\}$, $\varphi_\rho \in \{0, 0.3, 0.6, 0.9\}$, $s \in \{0.02, 0.1, 0.3, 1, 3\}$, $\lambda_n \in \{0.1, 0.3, 0.6, 0.9, 0.99\}$, $\sigma^2_{\rho,C} \in \{0.25, 1, 4\}$, $\sigma^2_{\varepsilon,C} \in \{0.05, 0.2, 1\}$ and $\sigma^2_{\eta,C} \in \{0.05, 0.2, 1, 5\}$ ($\sigma^2_{n,C}$ is normalized to 1 throughout).

How does the central bank reduce the sensitivity of market beliefs? The following corollary shows that, when assessment errors are small relative to variation in the assessments, optimal commitment implements this reduction with a lower composition parameter.[21]

Corollary 1 (Local commitment composition)**.** *Suppose* (22) *holds strictly. For all sufficiently small* $\sigma^2_{\varepsilon,C}/\sigma^2_{n,C}$ *and* $\sigma^2_{\eta,C}/\sigma^2_{n,C}$*, any optimal commitment rule has composition*

$$b_c < b_d.$$

A lower value of b weakens belief sensitivity and reduces the pass-through of demand-assessment errors. When $\lambda_n > 0$, it also reduces the pass-through of natural-output-assessment errors through the factor $b\chi(b,s)$. Note that the result concerns the *relative weight* on natural output, not the raw natural-output coefficient ab.

When $\lambda_n > 0$, condition (22) is sufficient but not necessary for a lower composition parameter in the small-error regime. The natural-output-error channel depends on $b\chi(b,s)$, which is locally increasing at the discretionary composition and therefore favors a reduction in b. This force may dominate even when the demand-error channel, operating via χ, favors an increase in b. When $\lambda_n = 0$, on the other hand, only the demand-error component depends on policy and hence, away from the knife-edge case where $\chi_b(b_d, s) = 0$, commitment tweaks policy in the direction that reduces χ.

Thus, the reduction in assessment-error propagation need not always be implemented through a lower composition parameter. Corollary 1 provides conditions under which it is. As we show in Section 5, this scenario is the empirically relevant one.

4 Empirical Evidence

This section explores the empirical relevance of the model's belief-revision channel using the reaction of market expectations to monetary policy surprises. First, we document substantial revisions in inflation expectations following policy-path surprises. Then, we show that variation across periods is consistent with the model's prediction discussed in Section 3.2: weaker policy responses generate smaller surprises but larger belief revisions per basis point.

[21] Proposition 4 in Appendix E also shows that $a_c > 1$. This "compensating" scale increase is a consequence of the fact that commitment shapes only $t = 0$ policy: assessment-error pass-through $\mathcal{P}(b,s)$ is independent of a. In Section 5, commitment also changes future policy, and scale affects how inferred assessments propagate into future allocations. A simple example illustrates this: if $a \to 0$, the market does not update much about the future, not because the information content of the current rate is low (it is independent of scale as discussed above), but because in this scenario future outcomes are independent of the central bank's assessments due to policy being unresponsive.

4.1 Empirical Specification

Let mps_t denote a revision in the expected one-year interest-rate path at monetary policy event t. We focus on inflation expectations due to data availability.[22]

Let $H \in \{0\text{–}2, 2\text{–}5\}$ index the two inflation horizons, measured in years, and let $\pi_{t,H}$ denote the corresponding annualized forward breakeven inflation rate. We use its change around the event, $\Delta\pi_{t,H}$, to proxy the revision in market inflation expectations.

For each horizon H, the coefficient of interest is β_H in the regression

$$\Delta\pi_{t,H} = \alpha_H + \beta_H mps_t + u_{t,H}, \tag{24}$$

where the measurement windows are described below. We estimate this regression separately for each of the four periods described below.

The natural theoretical counterpart of β_H is the model's belief response per unit of policy news, $-\chi(b,s)/a$. The correspondence is not exact: the empirical coefficient concerns inflation expectations and news about the expected one-year policy path, while our baseline model concerns expected output and the current policy rate. Nevertheless, both measure how much market expectations change per unit of policy news and, as argued below, the model's prediction about the cross-period variation in $-\chi(b,s)/a$ applies to β_H as well.

4.2 Measurement and Implementation Details

This subsection discusses the source of the data and the selection of events in the specification (24) together with the sample selection.

Monetary policy surprises. We obtain both the event calendar and the policy-news measure from the U.S. Monetary Policy Event-Study Database (USMPD) introduced by Acosta et al. (2025).[23] We use two types of USMPD events: *Monetary Events* and minutes releases. *Monetary Events* combine the policy statement and the subsequent press conference (if there is one): the event window runs from 10 minutes before the statement release to 60 minutes after the press conference begins. When there is no press conference, USMPD uses the 30-minute statement window running from 10 minutes before to 20 minutes after the release. For minutes releases, USMPD uses a 30-minute window from 10 minutes before to 20

[22] While the baseline model does not have inflation, it is conceptually straightforward to incorporate it, as shown in the quantitative model in Section 5. Furthermore, in that model, as in many others of the New Keynesian variety, market expectations of inflation and output are closely related.

[23] Available at `https://www.frbsf.org/research-and-insights/data-and-indicators/us-monetary-policy-event-study-database/`

minutes after the release. For each event, we take mps_t directly from the database's OIS1Y series. It is given by the intraday change in the fixed rate on a one-year U.S. dollar overnight-indexed swap over the relevant USMPD event window. Thus, mps_t captures unexpected news about the policy path conveyed by the entire FOMC event—through the policy statement and associated press conference, or through the release of minutes, *not simply* an unexpected change in the current federal funds rate. Under the expectations hypothesis, this change approximately measures the market's revision in the expected overnight policy-rate path over the following year, up to movements in the risk premium. The coefficient β_H in (24), then, measures how expected inflation is revised per unit of policy-path news at an FOMC event. In particular, it is not intended to identify the causal effect of an exogenous Taylor-rule intercept shock.[24] This is the appropriate object to study in our context, since we are interested in how the market updates its beliefs about the Fed's assessment of the economy and thus about future macroeconomic aggregates and not in the causal effect of a one-off Fed funds shock.

Inflation swaps. We measure breakeven inflation using daily Bloomberg data on two- and five-year U.S. inflation swaps indexed to the non-seasonally-adjusted headline Consumer Price Index for All Urban Consumers (CPI-U). Letting $\pi_{t,h}$ be the swap rate at time t for maturity h, we construct the forwards as $\pi_{t,h_1 \to h_2} = (h_2 \pi_{t,h_2} - h_1 \pi_{t,h_1})/(h_2 - h_1)$. In order to compute the change in expectations around the monetary event, we subtract the value on the last trading day before the event from the value on the first trading day after it.

Sample. Our sample is limited by historical inflation swap data, which starts on July 21, 2004. Thus, the first event is an FOMC meeting on August 10, 2004. Our last monetary event is an FOMC meeting on April 29, 2026. We exclude observations associated with unscheduled FOMC meetings. We also exclude events in calendar years 2020–2021 because of the Covid-19 crisis during those years. The exceptional conditions may have made market reactions to monetary policy surprises unrepresentative of other periods. The final sample contains 318 scheduled FOMC events: 63 in the pre-GFC period (August 10, 2004 to May 21, 2008), 25 in the GFC period (June 25, 2008 to December 16, 2009), 160 in the post-GFC/pre-Covid period (January 6, 2010 to December 11, 2019), and 70 in the post-2021 period (January 5, 2022 to April 29, 2026). It comprises 159 policy announcements (including the associated press conference when applicable) and 159 releases of FOMC minutes.

[24] As discussed in Bauer and Swanson (2023), to quantify the dynamic causal effects of an intercept shock to a Taylor rule, we would need to take additional steps. See also Jarociński and Karadi (2026).

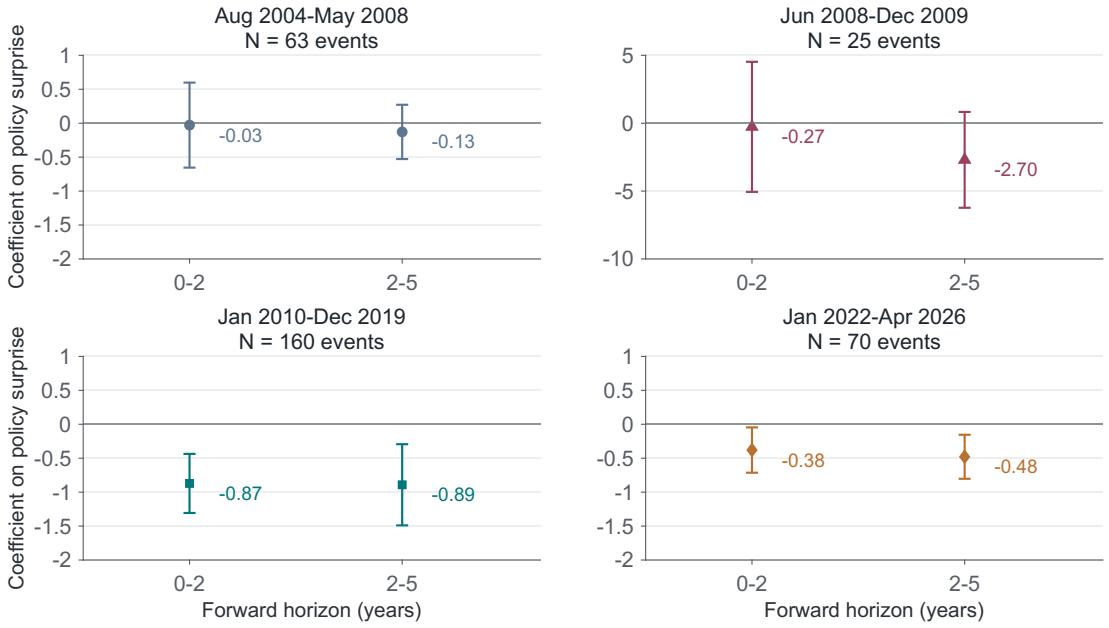

Figure 3: Estimates of (24) using the USMPD OIS1Y surprise. Markers and adjacent numbers report coefficients; whiskers show 95% HC1 confidence intervals. Horizontal axes show forward horizons in years. All panels report percentage points per percentage-point surprise. Note the GFC panel (June 2008–December 2009) has a different y-axis.

4.3 Results

Figure 3 shows the estimated coefficient in (24) for different maturities and subsamples. Point estimates are negative in each case, but their precision differs substantially across regimes. The response is statistically different from zero at the 5% level at both horizons during 2010–2019, at both horizons during 2022–2026, and at neither horizon during the GFC.[25] None of the pre-GFC responses is significant at that level.

Furthermore, the variation across periods is consistent with our model's prediction discussed in Section 3.2. Recall that holding policy composition, the information structure, and continuation policy fixed, a lower policy scale a generates smaller policy surprises and larger belief revisions per basis point, $|\chi(b,s)/a|$. Intuitively, when policy responds less to the central bank's assessments, a given rate surprise induces a larger revision in market beliefs about those assessments.

Recall that the estimated coefficient measures the response of inflation expectations to news about the expected one-year policy path, while $-\chi(b,s)/a$ measures the response of expected output to news about the current policy rate. However, these differences do not invalidate the prediction above. First, the prediction follows from Bayesian updating in a

[25]Since the sample is quite limited (25 events), we interpret these magnitudes cautiously.

linear forecasting problem. Scaling down the policy signal leaves the information it conveys about the central bank's assessments unchanged, so an optimal forecaster increases the response per unit of signal proportionally. Thus, it applies to inflation just as it does to output. Moreover, the prediction still holds (though *directionally*, not with the exact proportionality) if the surprise measure combines current and expected future policy rates as our OIS1Y measure does.[26] Consequently, the qualitative prediction remains relevant when policy news is measured using OIS1Y.

Table 1 summarizes expectations responses and policy-path surprise volatility across the three comparison periods. The estimated sensitivity of market beliefs per basis point is substantially larger during 2010–2019: a 25-basis-point surprise implies a cumulative expected-price response over five years of -1.10 percentage points, compared with -0.11 before the GFC and -0.55 during 2022–2026. A Wald test rejects equality of the cumulative response coefficients across these three periods ($p = .006$).

At the same time, policy-path surprises were substantially smaller during 2010–2019. Their standard deviation was approximately 60% lower than in either comparison period (see Table 1 for the magnitudes). This sharp reduction is suggestive of weaker near-term policy responses, corresponding to a lower scale a in the model.[27] Naturally, lower surprise volatility could also be a consequence of less assessment variation or greater policy predictability. However, the accompanying sharp increase in belief sensitivity per basis point provides additional support for our interpretation.

A natural explanation of the smaller near-term policy surprises during that period is the fact that the zero lower bound was either binding or close to binding for a significant share of it.[28] Importantly, the argument behind the model's prediction does not require the weaker

[26] Let $z = (r_0 - \mathbb{E}_0^M[r_0])/a$. In the baseline model, $\Delta\,\mathbb{E}_0^M[y_1] = -\chi z = -\gamma\Delta\,\mathbb{E}_0^M[r_1]$. Consider a *policy-path surprise*, i.e., a weighted average of the surprise in the current rate and the surprise in the expected future rate

$$\Delta R = w(r_0 - \mathbb{E}_0^M[r_0]) + (1-w)\Delta\,\mathbb{E}_0^M[r_1].$$

Given that surprise, the updating coefficient can be written as $-\chi/[wa + (1-w)\chi/\gamma]$. Thus, for $a, \chi > 0$ and $0 < w \leq 1$, decreasing a also generates stronger belief revisions per unit of policy (path) news, holding the other determinants fixed.

[27] In addition, Pyle and Williams (2016) find that one-year Treasury yields responded substantially less to macroeconomic surprises in 2010–2014 than in 1990–2000. This is consistent with markets expecting (near-term) policy to respond less to economic news at least during the first years of our post-GFC sample.

[28] Note, however, that our empirical measure of surprises does not correspond to plain Fed funds rate surprises. Concretely, if the Fed had been able to fully substitute its main policy tool with the "unconventional" ones (chiefly interventions in asset markets or forward guidance), the inability to move the short rate would not have necessarily translated into smaller OIS1Y surprises. It is quite plausible, though, that either the tools were less effective or the Fed was more hesitant in using them aggressively, amounting to a more muted—but, importantly, nonzero—responsiveness of the near-term policy path.

Table 1: Policy-path surprises and expectations responses across periods

Period	Expected-price response (pp)	Surprise SD (bp)
2004–2008	−0.11	4.05
2010–2019	−1.10	1.70
2022–2026	−0.55	4.62

Notes: The expected-price response is the revision in expected cumulative price growth over five years following a +25 bp OIS1Y surprise, in percentage points: $0.25(2\beta_{0\text{–}2} + 3\beta_{2\text{–}5})$. Surprise SD is the standard deviation of event-window OIS1Y changes, in basis points. The samples cover August 2004–May 2008, January 2010–December 2019, and January 2022–April 2026, with 63, 160, and 70 events, respectively.

policy responsiveness to be a policy choice, so it applies regardless of whether the cause was a "technological" constraint such as the ZLB.

Together, the smaller policy surprises and larger belief revisions per basis point are consistent with the model's prediction for weaker policy responsiveness during 2010–2019.

4.4 Measurement Robustness and Interpretation

Alternative measure of policy news. Our baseline specification measures policy news using the change in the one-year OIS rate over the USMPD event window. Appendix F.1 instead uses the fourth quarterly Eurodollar/SOFR futures contract, ED4. Although the coefficient levels differ across the two measures, the key finding is unchanged: at every horizon, the point estimate is more negative during 2010–2019 than in the pre-GFC and post-2021 periods. Hence, the cross-period ranking of sensitivity of market expectations is not specific to OIS1Y as the measure of policy-path news.

Alternative measure of inflation expectations. While our baseline uses inflation swaps to measure changes in the market's inflation expectation, a natural alternative is to use Treasury Inflation-Protected Securities (TIPS) breakevens. However, these can be affected by time-varying liquidity and other market-specific premia (Fleckenstein et al., 2014; Abrahams et al., 2016). Appendix F.2 reports the corresponding regressions using breakevens constructed from nominal Treasury and TIPS yields. The key finding persists: at both horizons, the point estimate is more negative during 2010–2019 than in the pre-GFC and post-2021 periods.

Quote construction and liquidity. The baseline inflation-swap series uses Bloomberg's LAST field. In a thin market with wide bid–ask spreads, this could partly reflect stale quotes or changes in the recorded price within the spread range instead of changes in the

underlying valuation of inflation compensation. Appendix F.3 rebuilds the swap-forward revisions and runs the specification (24) twice, once using bid-side levels at every maturity and once using ask-side levels. If the baseline result were driven mostly by movements of LAST with an otherwise unchanged bid–ask spread, the estimated coefficients would change materially when the same regressions were estimated at the two sides of the market. However, reassuringly, both the bid- and ask-based estimates preserve the main results.

Inflation-risk premia. The main limitation of the exercise is that inflation swaps measure inflation compensation under the risk-neutral measure. Their response may therefore reflect changes in inflation-risk premia as well as revisions in "physical" inflation expectations. There is no straightforward way to separate these components at the frequency of monetary-policy events: TIPS breakevens contain a corresponding risk-premium component, while survey expectations are not observed in narrow event windows or at comparable horizons (besides having other limitations). Moreover, existing evidence indicates that inflation-risk premia are generally small at shorter horizons (Haubrich et al., 2012; Ajello et al., 2020). Consequently, we interpret the estimates as informative about revisions in market inflation expectations, but acknowledge the possibility of a potentially important contribution from risk premia, particularly during the GFC and at longer horizons.

Event-window timing. The policy surprise is measured over a narrow intraday USMPD window, while the inflation-swap revision runs from the last trading day before the event to the first trading day afterward. The latter may therefore contain other news arriving outside the policy window. This may add noise or contamination from other news, but as far as we know there is no natural reason why such news should systematically generate the "ordering" documented above in the strength of the response across policy subsamples.

Taken together, the evidence documents substantial market responses to policy-path news and a joint pattern consistent with the model: during 2010–2019, larger belief revisions per basis point accompanied substantially smaller policy surprises. The ordering of the estimated expectations responses survives alternative measures of policy news and inflation compensation, as well as separate bid- and ask-based swap quotes. These findings support the empirical relevance of the inference mechanism, subject to the identification and measurement limitations discussed above.

The quantitative model below uses this evidence together with macroeconomic data and real-time Federal Reserve forecast revisions to discipline the strength of the belief-updating mechanism and assess its relevance.

5 Quantitative Analysis

This section quantifies the importance of our mechanism in a New Keynesian model estimated using U.S. data. The analysis delivers two lessons. First, how markets interpret policy decisions substantially affects the macroeconomic consequences of errors in the central bank's real-time assessments. Second, at our estimates, these costs can be substantially reduced through a much weaker response to natural output estimates. We first introduce the model and estimate it using U.S. macroeconomic data, Federal Reserve real-time forecasts, and inflation-swap responses to policy surprises. We then examine the gains from commitment and explain how changes in systematic policy achieve them. Finally, we assess which features of the data drive the results.

5.1 Environment

The private sector's behavior is described by a forward-looking IS curve and New Keynesian Phillips curve,

$$y_t = \mathbb{E}_t^M y_{t+1} - \gamma \left(i_t - \mathbb{E}_t^M \pi_{t+1} - \rho_t \right), \tag{25}$$

$$\pi_t = \beta \, \mathbb{E}_t^M \pi_{t+1} + \kappa (y_t - y_t^n) + u_t, \tag{26}$$

where y_t is output, π_t is inflation, y_t^n is natural output[29], i_t is the nominal short rate, ρ_t is the demand or discount-rate shock, and u_t is a cost-push shock. As in the baseline model, γ is the elasticity of intertemporal substitution in the IS curve, and κ is the NKPC slope. We let $\bar{\rho}$ denote the unconditional mean of the natural rate and hence the steady-state nominal rate under zero trend inflation.

The central bank minimizes a standard dual-mandate loss function with an additional rate-level penalty,

$$\frac{1}{2} \mathbb{E}_0^C \sum_{t=0}^{\infty} \beta^t \left[x_t^2 + \vartheta_\pi \pi_t^2 + \vartheta_i (i_t - \bar{\rho})^2 \right], \tag{27}$$

[29] We use "natural output" to denote the flexible-price output *absent markup shocks*. Since we assume the steady-state monopoly distortion is offset by a subsidy (see Appendix A), this coincides with the efficient level of output. However, it differs from the flexible-price output in the presence of cost-push shocks (indeed, under the standard microfoundation, the cost-push shock is proportional to the gap between the efficient level of output and its flexible-price counterpart). Hence, our terminology is different from that of Blanchard and Galí (2007) (among others), which uses "natural output" to refer to the flexible-price one.

where $x_t \equiv y_t - y_t^n$ is the output gap. The fundamental state is $X_t = (y_t^n, \rho_t - \bar{\rho}, u_t)^T$. Each component follows an independent AR(1),

$$X_t = A_X X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma_\varepsilon),$$

where $A_X \equiv \text{diag}(\varphi_n, \varphi_\rho, \varphi_u)$ with $|\varphi_j| < 1$ and $\Sigma_\varepsilon \equiv \text{diag}(\sigma_n^2, \sigma_\rho^2, \sigma_u^2)$ is a positive-definite matrix.

5.2 Information and Equilibrium

The information structure is similar to that of the baseline model, except that the central bank's estimates are explicitly derived from noisy signals. Furthermore, the market's conditional mean and variance for the central bank's assessments are implied by the signal structure via a standard steady-state Kalman filter. Finally, all parameters of the signal structure are common knowledge and the information structure admits a standard common-prior foundation.

We describe the core of the information structure below. Details can be found in Appendix G.

Beliefs. The central bank does not observe the fundamental state directly. Instead, each period it observes private correctly centered noisy signals

$$s_t^C = X_t + e_t^C, \qquad e_t^C \sim N(0, \Sigma_e),$$

where e_t^C is i.i.d. over time and independent of the fundamental innovations. Further, $\Sigma_e = \text{diag}((\sigma_n^e)^2, (\sigma_\rho^e)^2, (\sigma_u^e)^2)$, so each fundamental's signal noise is independent from the others (as are fundamentals themselves). As in the baseline model, the central bank's information consists only of these signals. In particular, it does not observe endogenous outcomes.[30]

The market, on the other hand, observes each period's current fundamentals X_t and the current policy rate i_t, but not the central bank's private signal s_t^C.

Timing. Within a period, fundamentals X_t and central-bank signals s_t^C are realized first. The market observes X_t, while the central bank observes s_t^C and revises its beliefs. Then, the central bank sets the policy rate i_t. Finally, the market observes i_t, revises its beliefs, and current output and inflation are determined by (25)–(26). In other words, the timing within

[30] If the central bank could observe past endogenous outcomes, it would change its policy to alter the informativeness of signals. Since this is not the focus of the current paper and its empirical relevance is doubtful, we shut down this channel (see Appendix B for an extended discussion).

a period is analogous to the one described in Figure 1, save for the inclusion of inflation and the replacement of the central bank's assessments with an explicit signal structure.

Finite-history approximation. Since the central bank's private assessments are persistent, agents optimally use the entire history to form their beliefs: the market uses fundamentals and policy rates to infer the central bank's current assessment, while the central bank uses signals and rates to forecast market beliefs. Thus, our economy has no finite-dimensional sufficient state. For computation, we approximate the market's conditional expectations in (25)–(26) by projections on truncated histories and, correspondingly, shrink the central bank's posterior state. Besides being a useful approximation, this has a natural counterpart in the U.S., since Greenbook/Tealbook estimates—the proxies for the central bank's private information we use—are publicly released with a lag.[31]

Our baseline estimation sets the finite history length to $H = 8$ quarters. Increasing the finite history length to $H = 12$ and re-estimating the model leaves the headline results unchanged, although some individual coefficients and intermediate magnitudes differ.[32]

Discretion and commitment. Discretionary equilibria and commitment are defined analogously to their counterparts in the baseline model. For a given approximation horizon H, market forecasts are given by the (population) linear projections implied by the policy rule and the retained histories of fundamentals and rates. A discretionary equilibrium is a fixed point between this forecast rule and the central bank's best response, which takes the forecasting rule as given. Under (linear) commitment, the central bank instead chooses ex ante a linear rule on the same finite-H information state and internalizes the forecast rule it induces.[33] Appendix G.2 provides the formal definitions while details on the solution method can be found in Appendix G.3.

5.3 Measurement and Estimation

We fix standard New Keynesian parameters to their values in Giannoni and Woodford (2005), $(\beta, \gamma, \kappa) = (0.99, 6.25, 0.024)$. We also fix $\vartheta_\pi = \vartheta_i = 1$, where inflation and rates are in quarterly units as is standard. Equal inflation and output-gap weights are a common

[31] See Appendix G.1 for a detailed discussion on the finite-history approximation.

[32] See Appendix H.1 for details.

[33] Throughout, we study stationary projection equilibria as defined in Appendix G.2. Our focus on equilibria driven by the model's fundamental and signal innovations is in the spirit of the minimum state variable approach of McCallum (1983). Establishing the unique implementation of the resulting allocation is beyond the scope of our analysis. In particular, our assumption that the central bank does not observe endogenous outcomes rules out standard direct feedback rules. More generally, as Lubik et al. (2023) show, imperfect information introduces additional difficulties in ruling out indeterminacy.

dual-mandate benchmark. There is less agreement on the rate-level weight, so we use one as a transparent reference value and report the results' sensitivity to deviations from it in Appendix H.2.

The remaining nine parameters—the innovation standard deviations, persistence parameters, and signal-noise ratios of the three fundamentals—are estimated using ten empirical moments taken from recent U.S. history: six macro moments (persistence and residual variance of inflation, the output gap, and the federal funds rate), one natural-output uncertainty moment $\mathcal{H}$, two Tealbook/Greenbook forecast-revision moments, and one inflation-swaps moment measuring the reaction of expected inflation to monetary-policy surprises.[34]

Macro moments. The first set of targets contains six quarterly macro moments: AR(1) coefficients and AR(1) residual variances for inflation, the output gap, and the policy rate. The underlying series are real GDP, the PCE price index, and the effective federal funds rate from FRED (series GDPC1, PCEPI, and FEDFUNDS, respectively). We transform real GDP to log points and extract its cyclical component using the Hamilton (2018) filter with parameters $h = 8$, $p = 4$. We interpret this component as an empirical proxy for the output gap. Inflation is the quarter-over-quarter log change in the PCE price index, and the policy rate is measured in quarterly percentage points. Thus, the output gap is in percentage log-output points, while inflation and the policy rate are in quarterly percentage points. The sample contains 120 quarters from 1990Q1 through 2019Q4 (we omit the Covid period).

Central bank information moments. The second set of targets disciplines the precision of the central bank's private assessments. First, to pin down the central bank's uncertainty about natural output, we target the average one-sided filtered posterior standard deviation of potential output from Holston et al. (2023)'s model. We begin with the data vintage distributed with the official 2026Q1 HLW estimates, discard all observations after 2019Q4, and re-estimate the model's three stages over 1961Q1–2019Q4. Holding the resulting parameter estimates fixed, we use the *one-sided* Stage-3 Kalman filter: the posterior covariance for each quarter t conditions only on observations up to t. We then average the posterior standard deviation of potential output over the 120 quarters from 1990Q1 through 2019Q4.[35]

[34] Our use of expectations data to discipline a model of monetary policy with asymmetric information is related to the quantitative analyses of Caravello et al. (2026) and Bocola et al. (2026). The latter also target the response of inflation expectations to monetary policy surprises. We complement this with Federal Reserve forecast revisions and evidence on natural output uncertainty to discipline both how markets interpret policy decisions and the precision of the central bank's underlying assessments.

[35] Although the state estimates are one-sided, this is not a fully real-time exercise for two reasons. First, the historical observations come from the revised 2026Q1 dataset. Second, the parameters are estimated once through 2019Q4 instead of recursively at each date. Unfortunately, the data in the official HLW estimates do not allow us to build forecasts that are based on real-time data alone.

This gives $\mathcal{H} = 1.7578$, which we use as a proxy for central-bank uncertainty about natural output and impose exactly in the estimation.[36] Second, we use revisions to Tealbook (formerly Greenbook) estimates to discipline the central bank's real-time uncertainty about headline PCE inflation and real GDP growth (and thus, indirectly, about the underlying shock processes). These are prepared by the Federal Reserve Board Staff before each FOMC meeting. We retrieve these from the Federal Reserve Bank of Philadelphia's Greenbook Data Sets.[37] The effective sample consists of 55 matched vintage-quarter pairs representing 46 distinct target quarters during 2000Q1–2019Q2 for each variable. In particular, for each target quarter τ, let z_τ^0 denote the current-quarter nowcast and z_τ^L the first estimate of the same quarter reported in a vintage *at least twelve calendar months later*. We target the (centered) revision-variance share

$$R_z^{\text{data}} = \frac{\text{Var}(z_\tau^L - z_\tau^0)}{\text{Var}(z_\tau^L)}.$$

This measures the size of subsequent revisions relative to the variation in the later-vintage estimates. In the model, its counterpart is the posterior-variance share

$$R_z^{\text{model}} = \frac{\text{Var}(z_\tau - \mathbb{E}_\tau^C z_\tau)}{\text{Var}(z_\tau)}.$$

Consequently, we treat the later-vintage estimate as a proxy for the ex-post realization and thus interpret the empirical revision-variance shares as proxies for the fraction of its variation that the central bank could not predict.

Inflation swaps moment. Finally, we target a moment measuring the response of market expectations based on inflation swaps and monetary policy surprises as detailed in Section 4.[38] This can be thought of as the key moment disciplining the strength of our belief-revision mechanism. Indeed, an analogous economy with imperfect but public central bank estimates would be unable to match this moment since there would be no policy surprises in the first place.

As described there, the event information and the measured OIS1Y changes come from

[36] We do so since this procedure does not yield a properly defined "sampling uncertainty" and thus it cannot be weighted in the same way as the other moments. Further, it is not really a "data moment" since its construction uses a model. Section 5.6 below discusses the sensitivity of our results to changes in the degree of natural output uncertainty.

[37] Available at `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/greenbook` (series `gPPCE` for inflation and `gRGDP` for real GDP growth).

[38] As acknowledged there, this treats changes in inflation-swap breakevens as a proxy for revisions in expected inflation, even though in principle they may also reflect changes in inflation risk premia.

USMPD while the daily inflation-swap rates come from Bloomberg. The event definitions and intraday windows are the same as in the empirical analysis. Note that the swaps reference CPI inflation, while the macro and forecast-revision moments use PCE inflation. We abstract from this distinction and map the empirical moments to the model's single inflation variable.

We use the same event sample as the empirical analysis in the previous section, pooling all 318 events to estimate the inflation-swaps target. The sample contains scheduled announcements and minutes from August 10, 2004 through April 29, 2026 and excludes the entire calendar years 2020–2021. This leaves 318 events: 159 monetary-policy announcements and 159 minutes releases.

Concretely, the target is the response of the expected cumulative price change over the first five years following a monetary-policy surprise. We use the (expected) cumulative price response instead of the response of medium-term forward inflation because our model has no endogenous source of inflation persistence. Empirically, it is built from the reactions of the 0–2 year and 2–5 year inflation-swap forwards as follows

$$\beta^{P}_{1:20} = 2\beta^{\pi}_{1:8} + 3\beta^{\pi}_{9:20},$$

where $\beta^{\pi}_{1:8}$ and $\beta^{\pi}_{9:20}$ denote the pooled full-sample 0–2 year and 2–5 year coefficients, using the specification in Section 4 with a common intercept and slope across all events. The units are percentage points of cumulative expected price growth per percentage point of annualized policy-path news. In the model, $\beta^{P}_{1:20}$ is the coefficient in the analogous (population) regression of the market's revision in expected cumulative price growth over the first five years on its revision in the expected one-year nominal policy path.[39] Appendix G.3.4 gives its finite-H counterpart.

Estimation. We estimate the parameters by minimum distance. Since the targets combine different datasets, sampling units, and frequencies, an estimate of their joint sampling covariance matrix is not directly available. Thus, we use diagonal inverse-sampling-variance weights. Appendix G.4 contains the estimation details. All model target moments, losses, and decompositions reported here and in the appendix are computed analytically.

[39] In the model, the central bank's assessments are revealed through its rate decision alone. Empirically, policy-path news also reflects explicit communication. The target moment remains similar when minutes releases are excluded: using only the 159 policy announcements, including any associated press conference, gives $\beta^{P}_{1:20} = -2.092$ (paired-bootstrap standard error 0.674), while the full sample gives -2.449.

5.4 Estimates and Model Fit

The estimated parameter values are shown in Table 2.

Table 2: Baseline Estimated Parameters

Parameter	Value	Description
σ_u	0.3627	Cost-push innovation sd
σ_n	1.0960	Natural-output innovation sd
σ_ρ	0.0971	Demand innovation sd
φ_u	0.0540	Cost-push persistence
φ_n	0.9568	Natural-output persistence
φ_ρ	0.9666	Demand persistence
q_u	0.1813	Cost-push signal-noise ratio
q_n	0.9322	Natural-output signal-noise ratio
q_ρ	0.0017	Demand signal-noise ratio

Here, q_j, for $j \in \{u, n, \rho\}$, is the ratio of the signal-noise variance to the unconditional variance of the corresponding fundamental. The fitted signal structure implies steady-state posterior-variance shares of 0.0017 for the demand state and 0.2175 for natural output. Thus, the central bank is estimated to have substantially more precise real-time information about demand than about natural output, consistent with the findings about the difficulty of measuring economic slack in real time (Orphanides, 2001; Orphanides and van Norden, 2002).

The high estimated persistence of natural output is consistent with the persistent movements in productive capacity (see, e.g., Holston et al. (2023), who allow permanent innovations to natural output and its trend growth). On the other hand, the very low estimated persistence of cost-push shocks has precedents in the literature (see, e.g., Sbordone (2007) and the estimates assuming rational expectations in Milani (2007); also used as a benchmark in Galí (2015)). The high demand persistence is less standard (albeit not unprecedented, see, e.g., Miao et al. (2020)). It may partly reflect the absence of endogenous propagation mechanisms such as consumption habits and investment adjustment costs.[40]

Taken together, persistent fundamentals and noisy signals make assessment errors persistent. Following a one-time signal error, subsequent observations gradually correct the central bank's assessment. At the estimated parameters, the effect of a natural-output signal error has a half-life of approximately 2.24 quarters, while the effect of a demand-signal error has

[40]This was already acknowledged by Smets and Wouters (2007, p. 597). They find that weakening these endogenous persistence mechanisms substantially increases the estimated persistence of demand disturbances using aggregate U.S. data.

Table 3: Targeted Moments Fit

Moment	Data	Model
AR(1): inflation	0.1321	0.0951
AR(1): output gap	0.8943	0.6326
AR(1): policy rate	0.9643	0.9642
Residual variance: inflation	0.1729	0.1582
Residual variance: output gap	1.1785	1.1972
Residual variance: policy rate	0.0103	0.0103
Natural-output uncertainty ($\mathcal{H}$)	1.7578	1.7578
Forecast-revision variance share: inflation	0.2414	0.2299
Forecast-revision variance share: real GDP growth	0.5876	0.4342
Market-expectations sensitivity $\beta^P_{1:20}$	-2.4486	-3.2574

a half-life of only 0.19 quarters (see Appendix G.3.1 for details). Since the two fundamentals have a very similar estimated persistence, this difference reflects the lower precision of natural-output information.

Table 3 presents the fit of the model in terms of the targeted moments. The largest miss is the persistence of the output gap. The two Tealbook moments and the inflation-swaps moment remain fairly close to their empirical counterparts (less than one standard error away).

5.5 The Role of Commitment

Commitment gains. We consider three policy regimes: discretion, commitment and a restricted-commitment regime where the central bank chooses an ex ante linear rule that can only condition on its current assessments of the three fundamentals. Studying this simple rule serves two purposes. First, as we show below, it captures almost all of the commitment gain, making it useful for identifying which features of policy matter most for stabilization. Second, the comparison allows us to assess how much can be achieved without direct history dependence, a central feature of standard New Keynesian commitment policies. The results are summarized in Table 4 below.

Table 4: Commitment Gains

Policy	Loss	Gain (% of D loss)	Share of full commitment gain
Discretion	1.1510	–	–
Commitment: simple rule	0.1386	87.96%	98.75%
Full commitment	0.1257	89.08%	100.00%

Full commitment lowers the stationary per-period loss by 89.08% and reduces the standard deviation of the output gap by 89.9% (see Appendix G.5 for the full loss decomposition under each policy regime). The simple rule captures 98.75% of the full commitment gain. Thus, almost all of the loss reduction can be achieved by changing how strongly policy responds to the central bank's current assessments. By construction, conditional on the central bank's current assessments, the simple rule implements the same rate regardless of past assessments and rates (and, furthermore, does not incorporate a Ramsey "promise state"). Consequently, the excellent performance of this rule documented above shows that the source of the large commitment gains is not the ability to honor past promises or to compensate for past policy mistakes caused by poor information (see, e.g., Woodford (2003) and Aoki (2006) respectively). But how can changing only the systematic response to current assessments deliver such a substantial improvement? We turn to this question next.

Changes in systematic policy. Table 5 shows how the on-impact interest-rate response to signal innovations varies across regimes.[41] Under the simple rule the absolute response to a natural-output signal innovation falls by 97.47%, compared with 18.10% for a demand-signal innovation. Thus, commitment entails a particularly sharp reduction in the response to natural output relative to demand.

Table 5: Impact Policy Responses to Signal Innovations

Signal	Discretion	Simple rule	Full commitment
Natural output	−0.0720	−0.0018	−0.0158
Demand	0.0601	0.0493	0.0484
Cost-push	0.0269	0.0316	0.0208

Notes: Interest-rate responses are in annual percentage points. Each row uses the same positive signal innovation across policy regimes, with its size set to one standard deviation of the corresponding signal noise.

As in the baseline model, a discretionary central bank takes the market's forecasting structure as given and uses the reaction of market expectations to stabilize the economy as it perceives it (recall that its problem is linear-quadratic and certainty equivalence applies). Commitment instead internalizes how its systematic policy shapes that forecasting structure and, consequently, the transmission of assessment errors into allocations.

In this richer model, policy affects the market's forecasting structure through two separate channels. It determines what the market infers about the central bank's assessments from

[41] Under discretion and full commitment, policy depends on the central bank's current estimates of the past values of fundamentals as well as their current values. A new signal about a persistent fundamental prompts an update of these past-value estimates too, so the on-impact policy response does not in general equal the corresponding estimate's coefficient times the revision in that estimate.

its decisions, and how those inferred assessments map into (expected) future policy rates. The first channel is the focus of the analysis in the baseline model. The second is absent there, since in that model commitment applied only to $t = 0$ policy, while policy from $t = 1$ onward was determined by the discretionary equilibrium. Here, commitment binds policy in all periods, so it can change how future assessments map into future policy rates. While naturally both elements shape commitment gains, the following exercises show the key role that the first channel plays.

The policy reaction to perceived demand pressure: an illustration. To make the mechanism above concrete, consider a central bank whose information suggests positive demand pressure due to demand-signal noise. A discretionary central bank naturally raises its interest rate to counteract the perceived expansionary pressure. When doing this, it anticipates that it may surprise the market and thus part of its intended stabilization will be achieved via the market revising its expectations of output and inflation downward. In doing this, it naturally takes the continuation equilibria as given (i.e., its deviation is "one off" in game-theoretic terms).

Commitment changes both the size of the policy response and the market forecasting structure through which it operates. Under the simple rule, policy responds a bit less to demand assessments, but its response to natural output estimates is reduced much more sharply. An unexpected tightening, thus, mostly conveys news about the central bank's demand assessment.

Together, these changes shape how markets interpret policy and thus the latter's effects on macroeconomic variables. Concretely, following a one-standard-deviation demand-signal error, the increase in the nominal policy rate falls from 0.0601 under discretion to 0.0493 under the simple rule, while the resulting output-gap contraction shrinks from 0.5049 to 0.0803. Thus, commitment prevents a large share of the assessment error's pass-through into allocations. Figure 4 shows the responses over time, together with the corresponding responses to a natural-output signal error.

Note that commitment also improves stabilization following a genuine demand innovation, which the central bank initially detects only partially (and hence it also induces an assessment error). Following a one-standard-deviation demand innovation, the positive output-gap response on impact changes from 0.0482 under discretion to -0.0222 under the simple rule, thus becoming much smaller and even switching sign.[42]

We next show that changing the market's interpretation can improve stabilization without

[42]The contribution of fundamental demand innovations to the loss falls by 20.69%. Appendix G.5.3 contrasts the responses to fundamental innovations and signal errors.

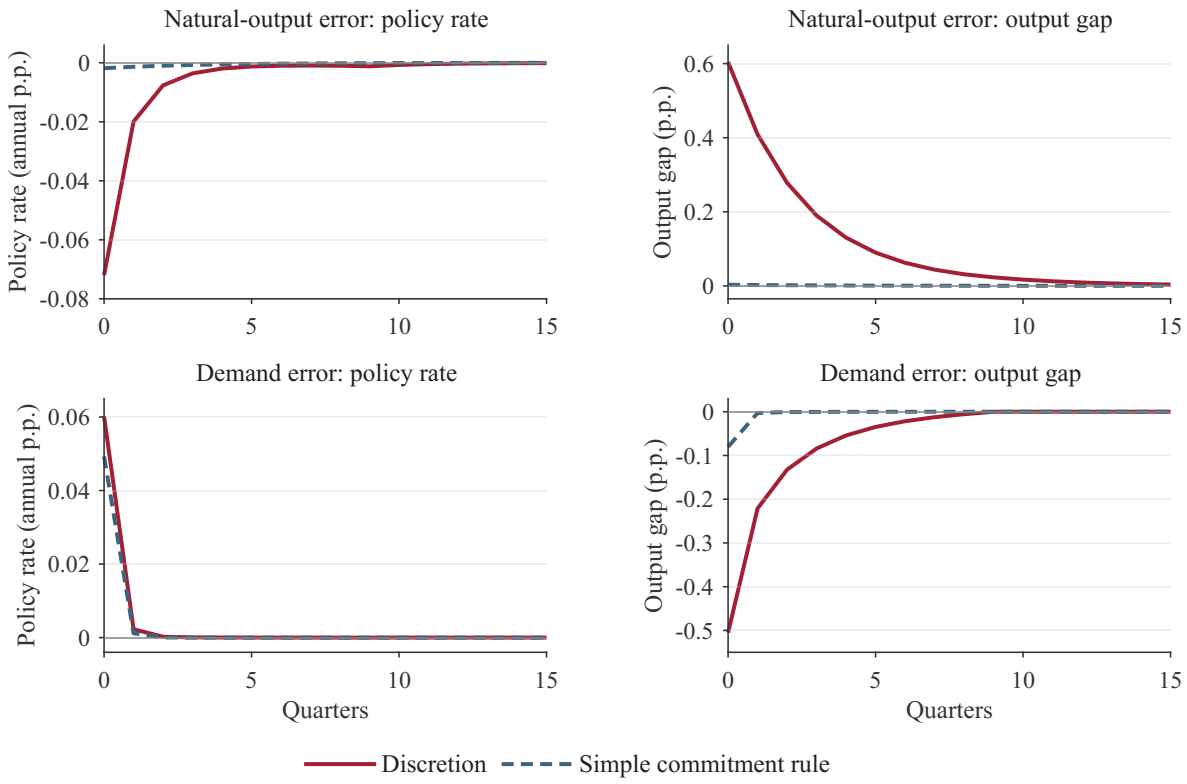

Figure 4: Responses to positive one-standard-deviation signal errors in natural output (top row) and demand (bottom row).

changing the realized interest rate path. Starting from the discretionary policy, we replace only the coefficients on current and past natural-output assessments with their simple-rule counterparts. Following a one-standard-deviation demand-signal error, the central bank's natural-output assessments remain unchanged. Since all other policy coefficients are unchanged, the two rules generate exactly the same realized nominal-rate path, as shown in Panel A of Figure 5.

While the rate decision is the same, markets interpret it differently, revising their estimate of the central bank's natural-output assessment downward much less under the counterfactual rule. On impact, this inferred decline is 0.4190 under discretion and only 0.0298 under the counterfactual, while the central bank's actual natural-output assessment is unchanged (Panel B). Because natural-output assessment errors are corrected slowly, attributing a tightening to such an error has persistent implications for expected future policy. Thus, expected next-period output falls much less and the decline in expected inflation is essentially eliminated. Furthermore, the output-gap contraction on impact is 81.18% smaller (Panel C). Note that if the central bank's assessments were public, changing the natural-output response

would not affect the macroeconomic consequences of the demand error. Appendix G.5.6 provides the details of this exercise.

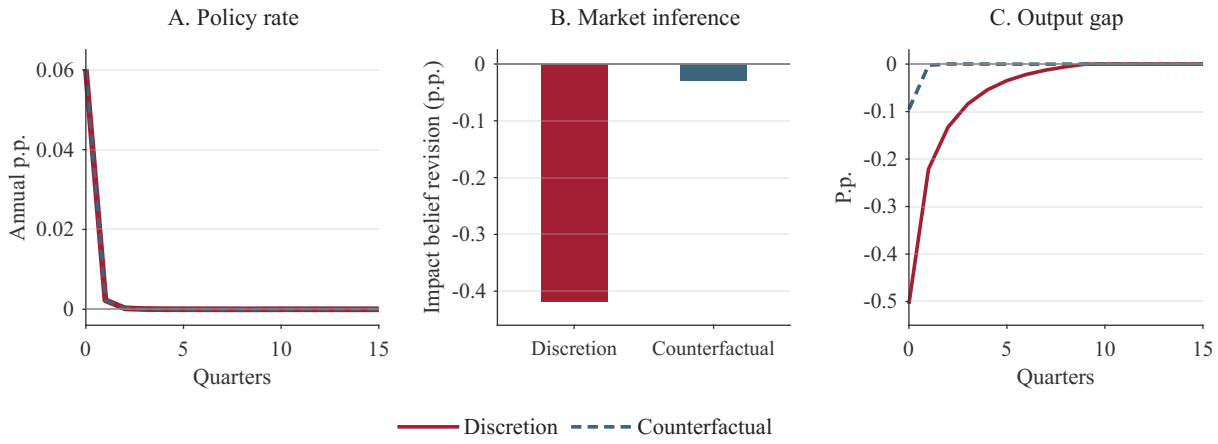

Figure 5: Responses to a positive one-standard-deviation demand-signal error. The counterfactual changes only the response to natural-output assessments. Panel B shows the market's impact revision to its estimate of the central bank's natural-output assessment.

These findings complement the case for caution in responding to noisy real-time information about economic slack, emphasized by Orphanides (2001, 2003). In the optimal-control benchmark of the latter paper, policy applies the full-information response coefficients to optimally filtered estimates: measurement uncertainty does not warrant further attenuation of the response to those estimates (i.e., certainty equivalence holds). In our framework, substantial stabilization gains follow from weakening the policy responses to natural output assessments that are already optimally filtered. This additional restraint limits the amplification of assessment errors through the market's interpretation of policy.

Cost-push signal errors provide a complementary illustration of the key role our central mechanism plays in the difference between the optimal commitment policy and its discretionary counterpart and between their corresponding losses. The simple commitment rule responds more strongly to these errors than discretion, yet generates much smaller movements in expected inflation and the output gap. This comparison has an intuitive interpretation. Cost-push fundamentals are almost fully transitory and their signals are not so noisy. Thus, their assessment errors are corrected very quickly. A stronger response to these assessments, together with a weaker response to natural-output assessments, makes an unexpected rate movement more indicative of a short-lived policy response and less indicative of persistent mistakes about natural output. Consequently, policy news can have smaller effects on expectations of future policy and allocations even when the contemporaneous rate response

is larger. The key consideration is thus how the entire rule shapes the market's interpretation of policy, as highlighted in the baseline model. Appendix G.5.4 contrasts fundamental innovations with signal errors and examines the interactions between policy responses.

The pass-through of assessment errors. In our model, policy decisions affect relevant outcomes both directly and by shaping market expectations of future policy. As in the theory, this creates a trade-off between stabilization conditional on the central bank's information and the propagation of assessment errors into allocations (output and inflation). To map the quantitative results to the loss decomposition in Lemma 2, note that we can write the unconditional output-gap volatility in a given period t as

$$\mathrm{Var}(x_t) = \mathrm{Var}\left(\mathbb{E}_t^C\, x_t\right) + \mathbb{E}\left[\mathrm{Var}(x_t \mid \mathcal{I}_t^C)\right].$$

The first term represents the variation in the output gap anticipated by the central bank at the time it chooses rates (i.e., after having formed its estimates for that period). The second, on the other hand, measures departures of the realized output gap from that conditional estimate, i.e., the quantitative counterpart of $\mathcal{P}$ in the baseline model.

The simple rule reduces the second component by 99.56%, while increasing the variation in the central bank's expected gap by a comparatively modest amount. It's worth stressing that the central bank's information is identical across regimes: what changes is the propagation of assessment errors to allocations (see Appendix G.5 for the full decomposition of output-gap variance under each policy regime).

Policy surprises and expectation revisions. The differences between discretion and a simple rule in the demand-signal error illustration above are reflected in how policy decisions shift inflation expectations. Since markets already observe fundamentals, these policy-induced belief revisions concern the central bank's assessments and their implications for future policy. Thus, they are an observable counterpart of the belief-revision mechanism emphasized in the theory.

We consider three statistics that are informative about this margin. First, we consider the object used in the empirical analysis and one of the targeted moments in the estimation above, $\beta^P_{1:20}$. As discussed above, it measures the revision in expected cumulative price growth over five years upon a revision of beliefs about the coming one-year policy path. Second, we consider the standard deviation of policy surprises, a measure of the "amount" of such policy news. Finally, we consider the product of the coefficient's absolute value and the

policy-surprise standard deviation, which equals the standard deviation of fitted expectation revisions, a measure of the overall pass-through of policy surprises into expectations:

$$\mathrm{SD}\left(\widehat{\Delta P^{e}}_{t,1:20}\right) = |\beta^{P}_{1:20}|\,\mathrm{SD}(mps^{M}_{t}),$$

where mps^{M}_{t} is the model's counterpart of the empirical monetary-policy surprise measure.

Table 6 reports these statistics across the three policy regimes.

Table 6: Policy News and Expectations Revisions

Policy	$\beta^{P}_{1:20}$	Policy-surprise SD	Fitted revision SD
Discretion	-3.2574	0.0308	0.1004
Commitment: simple rule	-0.0173	0.0150	0.0003
Full commitment	1.6914	0.0222	0.0375

Notes: The policy-surprise standard deviation is in annualized percentage points, while the fitted revision standard deviation is in cumulative-price percentage points.

In the estimated economy, commitment substantially reduces the absolute sensitivity of market expectations to policy news. The simple rule lowers the magnitude of the response of expected cumulative price growth over five years to a one-percentage-point policy-path surprise from about 3.26 to about 0.02, a reduction of 99.47%. Under full commitment, the coefficient becomes positive, at about 1.69, while its magnitude falls by 48.08% relative to discretion. Together with smaller policy surprises, these reductions lower the standard deviation of fitted expectation revisions by 99.74% and 62.64%, respectively.

The sign reversal under full commitment reflects a Fisherian response of the forward-looking New Keynesian block. A surprise that shifts up the market expectations of the one-year nominal rate path induces an upward revision in expected nominal rates far into the future. However, expected inflation also rises, dampening the initial increase in the real rate and leading to real rate declines later on. As a result, expected output increases. Under discretion, the surprise instead induces a persistent contraction and a fall in expected inflation. This comovement between expected nominal rates and inflation is consistent with the IS relation and the Fisherian mechanism discussed in Appendix G.5.3, particularly (113). Since markets already observe fundamentals, the news here concerns the central bank's assessments and their implications for future policy.

The simple rule brings $\beta^{P}_{1:20}$ closer to zero, while full commitment achieves slightly lower loss with a larger absolute coefficient. In general, whether the optimal commitment policy increases or decreases the sensitivity of market expectations depends on the balance between assessment-error amplification and the stabilization benefits of responsive expectations. These benefits are absent from the baseline theoretical model to isolate the mechanism

of interest, but are part of the quantitative model. First, shifting expected output and inflation allows policy to influence the economy with smaller rate deviations, decreasing the rate-level penalty given $\vartheta_i > 0$. Second, with cost-push shocks, changes in expected inflation can ease the trade-off between output-gap and inflation stabilization. Commitment balances these benefits against the propagation of assessment errors through the forecasting structure its policy rule induces.

We also compare the gains from commitment to the simple rule with the corresponding gains in two alternative economies: a full-information one and one with imperfect but public central bank information, keeping the physical environment and corresponding parameters unchanged (except, of course, the ones governing the central bank's information quality, which are moot in the full-information economy). The loss reduction relative to discretion is 18.14% under full information and 60.36% under public assessments, compared with 87.96% in our benchmark economy. The main difference is the substantially worse performance of discretion when assessments are private (commitment losses are fairly similar across these economies), which highlights the severity of the time inconsistency problem in economies where policy is transmitted to the real economy partly by changing market beliefs.[43]

Together, these results establish a quantitatively important role for the mechanism emphasized in the theory. Commitment sharply reduces reactions to mistaken natural-output assessments and, by doing that, also changes how a given policy decision is interpreted and transmitted into macroeconomic outcomes. The following subsection examines which features of the data determine both the magnitude of these gains and the associated changes in policy.

5.6 Which Data Features Drive the Results?

To assess which data features drive our key results, we examine one-at-a-time changes in the targeted moments. Changes in the HLW posterior-standard-deviation target have the largest effects on the magnitude of the commitment gains and the associated changes in policy and expectations. Thus, we focus on these perturbations, reducing the target from its baseline value, $\mathcal{H}_0 = 1.7578$, to one-half and one-quarter of it. To distinguish the direct role of natural-output uncertainty from the effects of other parameter changes, we consider two exercises. First, we change only the natural-output signal noise to match the modified moment, keeping all other structural parameters fixed. Second, we re-estimate all nine structural parameters subject to the lower uncertainty target, leaving the remaining

[43] See Appendix G.5.5 for the losses under both policy regimes and the associated gains from commitment in each economy.

targets and estimation weights as in the baseline estimation. In each case, we recompute the discretionary equilibrium and the optimal simple and full commitment policies.

We let G be the full commitment gain as a percentage of the discretionary loss and S the share of that gain captured by the simple rule. The final column of Table 7 reports R_n^C, the magnitude of the on-impact rate response to a natural-output signal error under full commitment relative to discretion. Values below one indicate a more restrained response under commitment; the closer the ratio is to one, the smaller the difference between the two policies.

Table 7: The Importance of Natural-Output Uncertainty

HLW target	G	S	R_n^C
Baseline	89.08%	98.75%	0.220
Only natural-output signal noise changes			
$\mathcal{H}_0/2$	26.31%	71.62%	0.702
$\mathcal{H}_0/4$	26.29%	72.25%	0.767
Re-estimation subject to the lower HLW target			
$\mathcal{H}_0/2$	49.26%	95.53%	0.475
$\mathcal{H}_0/4$	52.16%	96.82%	0.507

Notes: Each row compares policies within the same economy. The HLW target is a posterior standard deviation. The response ratio compares the effects of the same natural-output signal error across policy regimes.

Natural-output uncertainty is crucial for the quantitative lessons of this paper. When only the signal noise changes, halving the posterior standard deviation reduces the gain from 89.08% to 26.31%, with little further change when uncertainty is reduced to one-quarter. This reflects a sharp stabilization improvement under discretion: its loss falls from 1.151 to approximately 0.170, while the full-commitment loss remains close to 0.125. This is intuitive given our previous discussion, since more precise real-time information makes natural-output assessment errors smaller and faster to correct (concretely, their half-life falls from 2.24 quarters to 0.73 and 0.35 quarters respectively). The simple rule still captures most of the remaining gain, but direct history dependence accounts for a larger share of that gain than at the baseline.

More precise information also substantially reduces the degree of restraint prescribed by the commitment policy. At the baseline, full commitment's impact response to a natural-output signal error is only 22.0% of discretion's. This ratio rises to 70.2% and 76.7% when the uncertainty target is halved and quartered. Commitment continues to respond less strongly, but the difference between the two policies narrows considerably. Thus, smaller

and less persistent assessment errors reduce both the stabilization cost of discretion and the desirability of less aggressive reactions to natural-output readings.

Allowing the other parameters to adjust partially offsets both changes. Under re-estimation, the commitment gains are 49.26% and 52.16%, and the simple rule again captures almost all of them. The full-commitment impact response is 47.5% and 50.7% of discretion's. Hence, both the commitment gains and the degree of restraint remain substantially below their baseline levels, although the differences are smaller than when only signal noise changes.

The lower uncertainty also greatly affects the model's ability to account for the other targeted moments. Even before re-estimation, the output-gap AR(1) residual variance falls to 0.0211 and 0.0183, against a target of 1.1785 and a baseline fitted value of 1.1972. Re-estimation naturally improves the overall fit relative to changing the signal noise alone—particularly the fit of output-gap persistence—but still falls short by a significant margin: the output-gap residual variances remain only 0.0042 and 0.0049. Thus, the smaller commitment gains in these lower-uncertainty economies accompany substantially less output-gap variation than observed in the data.

By comparison, halving or quartering either the inflation or GDP-growth forecast-revision variance share target, while retaining the baseline HLW target and re-estimating the model, leaves the full commitment gains between 88.52% and 89.26%.

Overall, these exercises show that uncertainty about natural output plays a central role in the size of the commitment gains and in how much weaker commitment's response to natural output assessments is relative to discretion. Since natural output assessments are very noisy in our baseline calibration, errors are corrected slowly, and a strong policy reaction to these assessments makes market expectations particularly sensitive to policy surprises. In these exercises, more precise information reduces the impact of those errors and thus reduces the difference between discretion and commitment relative to the baseline, both in terms of losses and policy responses. Since the HLW target abstracts from important sources of uncertainty (e.g., model uncertainty in the filtering process), it may understate the broader real-time uncertainty about natural output faced by policymakers.

6 Conclusion

Monetary policy works largely by shaping expectations. Thus, its effects depend on how the private sector interprets a given decision. That interpretation depends, in turn, on how the central bank systematically conducts policy. This paper studies the implications for policy design. We show that a central bank choosing its rule ex ante internalizes this

dependence and reduces the output-gap volatility arising from the market's reaction to its imperfect real-time assessments.

Our empirical analysis documents substantial responses of market inflation expectations to policy news. In addition, during 2010–2019, larger belief revisions per basis point accompanied smaller policy-path surprises, a joint pattern consistent with the model's prediction for weaker policy responses. Together, these findings support the model's account of how market beliefs respond to monetary policy news.

In an estimated New Keynesian model incorporating this mechanism, accounting for the market interpretation of policy significantly reduces the policy loss. Full commitment entails a much less aggressive response to natural-output assessments than discretion. A rule responding only to current assessments shares this feature and captures almost all of the stabilization gain. Further counterfactuals show that changing how markets interpret policy accounts for an important part of the gain from commitment. The policy that achieves this makes market expectations less sensitive to policy surprises.

More broadly, the familiar image of monetary policy as a lever overlooks the fact that a rate decision is also an act of communication, whether intended or not: to a large extent, market participants react based on why they think the decision was made and what that implies for the future. Effective policy design must account for this interpretation. An open question is how this implicit communication through policy decisions interacts with explicit central bank communication. When can explaining a decision improve the inferences drawn from it, and how does that change the design of policy itself?

References

Abrahams, M., Adrian, T., Crump, R. K., Moench, E. and Yu, R. (2016), 'Decomposing real and nominal yield curves', *Journal of Monetary Economics* **84**, 182–200.

Acosta, M., Ajello, A., Bauer, M., Loria, F. and Miranda-Agrippino, S. (2025), Financial market effects of FOMC communication: Evidence from a new event-study database, Working Paper 30, Federal Reserve Bank of San Francisco.

Ajello, A., Benzoni, L. and Chyruk, O. (2020), 'Core and "crust": Consumer prices and the term structure of interest rates', *The Review of Financial Studies* **33**(8), 3719–3765.

Aoki, K. (2003), 'On the optimal monetary policy response to noisy indicators', *Journal of Monetary Economics* **50**(3), 501–523.

Aoki, K. (2006), 'Optimal commitment policy under noisy information', *Journal of Economic Dynamics and Control* **30**(1), 81–109.

Baeriswyl, R. and Cornand, C. (2010), 'The signaling role of policy actions', *Journal of Monetary Economics* **57**(6), 682–695.

Barro, R. J. and Gordon, D. B. (1983), 'A positive theory of monetary policy in a natural rate model', *Journal of Political Economy* **91**(4), 589–610.

Bauer, M. D., Pflueger, C. E. and Sunderam, A. (2024), 'Perceptions about monetary policy', *The Quarterly Journal of Economics* **139**(4), 2227–2278.

Bauer, M. D. and Swanson, E. T. (2023), 'An alternative explanation for the Fed Information Effect', *American Economic Review* **113**(3), 664–700.

Blanchard, O. and Galí, J. (2007), 'Real wage rigidities and the New Keynesian model', *Journal of Money, Credit and Banking* **39**(s1), 35–65.

Bocola, L., Dovis, A., Jørgensen, K. and Kirpalani, R. (2026), Monetary policy without an anchor, Working Paper Series 3218, European Central Bank.

Caballero, R. J. and Simsek, A. (2022), 'Monetary policy with opinionated markets', *American Economic Review* **112**(7), 2353–2392.

Calvo, G. A. (1978), 'On the time consistency of optimal policy in a monetary economy', *Econometrica* **46**(6), 1411–1428.

Caravello, T. E., Carrasco, A. and Martínez-Bruera, P. (2026), Reassessing central bank reputation: Beyond long-run expectations. Working paper.

Clarida, R., Galí, J. and Gertler, M. (1999), 'The science of monetary policy: A New Keynesian perspective', *Journal of Economic Literature* **37**(4), 1661–1707.

Damjanovic, T., Damjanovic, V. and Nolan, C. (2008), 'Unconditionally optimal monetary policy', *Journal of Monetary Economics* **55**(3), 491–500.

Dogra, K. (2024), Paradoxes and problems in the causal interpretation of equilibrium economics, Staff Report 1093, Federal Reserve Bank of New York.

Evans, G. W. and Honkapohja, S. (2001), *Learning and Expectations in Macroeconomics*, Princeton University Press.

Fleckenstein, M., Longstaff, F. A. and Lustig, H. (2014), 'The TIPS–Treasury bond puzzle', *Journal of Finance* **69**(5), 2151–2197.

Galí, J. (2015), *Monetary Policy, Inflation, and the Business Cycle. An Introduction to the New Keynesian Framework and Its Applications*, 2nd edn, Princeton University Press.

García-Schmidt, M. and Woodford, M. (2019), 'Are low interest rates deflationary? A paradox of perfect-foresight analysis', *American Economic Review* **109**(1), 86–120.

Giannoni, M. P. and Woodford, M. (2005), Optimal inflation-targeting rules, *in* B. S. Bernanke and M. Woodford, eds, 'The Inflation-Targeting Debate', University of Chicago Press, Chicago, pp. 93–162.

Gürkaynak, R. S., Sack, B. and Wright, J. H. (2007), 'The U.S. Treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**(8), 2291–2304.

Gürkaynak, R. S., Sack, B. and Wright, J. H. (2010), 'The TIPS yield curve and inflation compensation', *American Economic Journal: Macroeconomics* **2**(1), 70–92.

Hamilton, J. D. (2018), 'Why you should never use the Hodrick–Prescott filter', *Review of Economics and Statistics* **100**(5), 831–843.

Haubrich, J., Pennacchi, G. and Ritchken, P. (2012), 'Inflation expectations, real rates, and risk premia: Evidence from inflation swaps', *Review of Financial Studies* **25**(5), 1588–1629.

Holmstrom, B. (1982), Managerial incentive problems: A dynamic perspective, *in* 'Essays in Economics and Management in Honor of Lars Wahlbeck'.

Holston, K., Laubach, T. and Williams, J. C. (2023), Measuring the natural rate of interest after COVID-19, Staff Report 1063, Federal Reserve Bank of New York.

Jarociński, M. and Karadi, P. (2026), Monetary policy surprises with imperfect information.

Jia, C. (2023), 'The informational effect of monetary policy and the case for policy commitment', *European Economic Review* **156**, 104468.

Jimenez-Gimpel, V. (2026), Market interpretation and uncertainty in monetary transmission. Working paper, Massachusetts Institute of Technology.

Kydland, F. E. and Prescott, E. C. (1977), 'Rules rather than discretion: The inconsistency of optimal plans', *Journal of Political Economy* **85**(3), 473–491.

Lippi, F. and Neri, S. (2007), 'Information variables for monetary policy in an estimated structural model of the euro area', *Journal of Monetary Economics* **54**(4), 1256–1270.

Lubik, T. A., Matthes, C. and Mertens, E. (2023), 'Indeterminacy and imperfect information', *Review of Economic Dynamics* **49**, 37–57.

McCallum, B. T. (1983), 'On non-uniqueness in rational expectations models: An attempt at perspective', *Journal of Monetary Economics* **11**, 139–168.

McCallum, B. T. (2005), 'What is the proper perspective for monetary policy optimality?', *Monetary and Economic Studies* **23**(S1), 13–24.

McKay, A. and Wolf, C. K. (2023), 'What can time series regressions tell us about policy counterfactuals?', *Econometrica* **91**(5), 1695–1725.

Melosi, L. (2017), 'Signalling effects of monetary policy', *The Review of Economic Studies* **84**(2), 853–884.

Mertens, E. (2016), 'Managing beliefs about monetary policy under discretion', *Journal of Money, Credit and Banking* **48**(4), 661–698.

Miao, J., Wang, P. and Zha, T. (2020), 'Discount shock, price–rent dynamics, and the business cycle', *International Economic Review* **61**(3), 1229–1252.

Milani, F. (2007), 'Expectations, learning and macroeconomic persistence', *Journal of Monetary Economics* **54**(7), 2065–2082.

Morris, S. (1995), 'The common prior assumption in economic theory', *Economics and Philosophy* **11**(2), 227–253.

Orphanides, A. (2001), 'Monetary policy rules based on real-time data', *American Economic Review* **91**(4), 964–985.

Orphanides, A. (2003), 'Monetary policy evaluation with noisy information', *Journal of Monetary Economics* **50**(3), 605–631.

Orphanides, A. and van Norden, S. (2002), 'The unreliability of output-gap estimates in real time', *The Review of Economics and Statistics* **84**(4), 569–583.

Preston, B. (2005), 'Learning about monetary policy rules when long-horizon expectations matter', *International Journal of Central Banking* **1**(2), 81–126.

Pyle, B. and Williams, J. C. (2016), 'Data dependence awakens', *FRBSF Economic Letter* (2016-12).

Sbordone, A. M. (2007), 'Inflation persistence: Alternative interpretations and policy implications', *Journal of Monetary Economics* **54**(5), 1311–1339.

Smets, F. and Wouters, R. (2007), 'Shocks and frictions in US business cycles: A Bayesian DSGE approach', *American Economic Review* **97**(3), 586–606.

Stein, J. C. (1989), 'Efficient capital markets, inefficient firms: A model of myopic corporate behavior', *The Quarterly Journal of Economics* **104**(4), 655–669.

Stein, J. C. (2014), 'Challenges for monetary policy communication'. Remarks at the Money Marketeers of New York University, New York, New York.

Stein, J. C. and Sunderam, A. (2018), 'The Fed, the bond market, and gradualism in monetary policy', *Journal of Finance* **73**(3), 1015–1060.

Svensson, L. E. O. and Woodford, M. (2003), 'Indicator variables for optimal policy', *Journal of Monetary Economics* **50**(3), 691–720.

Svensson, L. E. O. and Woodford, M. (2004), 'Indicator variables for optimal policy under asymmetric information', *Journal of Economic Dynamics and Control* **28**(4), 661–690.

Townsend, R. M. (1983), 'Forecasting the forecasts of others', *Journal of Political Economy* **91**(4), 546–588.

Warsh, K. (2026), 'ranscript of chairman warsh's press conference, july 29, 2026'. Board of Governors of the Federal Reserve System.

Woodford, M. (1999), 'Optimal monetary policy inertia', *The Manchester School* **67**(S1), 1–35.

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

Appendix

A Microfoundation of the Baseline Model **53**
A.1 Environment . . . 53
A.2 Approximations . . . 55

B Information Structure **59**
B.1 General Structure . . . 59
B.2 Gaussian Beliefs and Further Assumptions in the Main Text . . . 62
B.3 Common-Prior Restrictions . . . 68

C Game Representation and Solution Concepts **71**
C.1 The Game . . . 71
C.2 Market-Consistent Profiles . . . 73
C.3 Discretion . . . 74
C.4 Commitment . . . 75

D Proofs **77**
D.1 Proof of Lemma 1 . . . 77
D.2 Proof of Proposition 1 . . . 77
D.3 Proof of Proposition 2 . . . 79
D.4 Proof of Lemma 2 . . . 82
D.5 Proof of Proposition 3 . . . 83
D.6 Proof of Corollary 1 . . . 86

E Additional Theoretical Results **86**

F Additional Empirical Results **92**
F.1 Alternative Policy-Surprise Measure . . . 92
F.2 TIPS-Based Inflation Compensation . . . 92
F.3 Bid and Ask Prices . . . 94

G Quantitative Model Details **95**
G.1 Information and Forecasting . . . 95
G.2 Definition of Equilibria . . . 98
G.3 Solving the Equilibrium and Computing Population Objects . . . 100
G.4 Estimation Details . . . 110
G.5 Additional Results . . . 111

H Robustness Exercises **118**
H.1 History-Length Sensitivity . . . 118
H.2 Rate-Level Stabilization Weight . . . 120

A Microfoundation of the Baseline Model

This appendix shows that the model in the main text can be obtained as a linear approximation of a standard economy with rigid prices.

A.1 Environment

Households. Consider a representative household with preferences over consumption C_t and labor N_t given by

$$\mathbb{E}_0^M \sum_{t=0}^{\infty} \left(\prod_{j=0}^{t-1} \delta_j \right) \left[\frac{C_t^{1-\upsilon} - 1}{1-\upsilon} - \zeta_t \frac{N_t^{1+\nu}}{1+\nu} \right], \qquad \upsilon > 0, \quad \nu > 0, \tag{28}$$

where δ_j is the one-period discount factor, the source of demand shocks. Households trade a one-period nominal bond with risk-free gross nominal return R_t set by the central bank and face the usual budget constraint

$$P_t C_t + \frac{B_{t+1}}{R_t} \leq W_t N_t + B_t + \Pi_t + T_t, \tag{29}$$

where P_t is the price of the final consumption good, W_t is the nominal wage, B_t is the household's bond holdings, Π_t denotes firm profits and T_t collects lump-sum taxes and transfers from the government. The household's Euler equation is

$$C_t^{-\upsilon} = \delta_t \, \mathbb{E}_t^M \left[R_t \frac{P_t}{P_{t+1}} C_{t+1}^{-\upsilon} \right], \tag{30}$$

where we denote by $\mathbb{E}_t^M$ the expectations operator corresponding to the representative household's beliefs (see Appendix B for details on the model's information structure). Importantly, the consumption decision is taken after the central bank sets the policy rate, so the expectation in (30) includes R_t in its conditioning information set.

Firms. As in the standard New Keynesian model, there is a continuum of intermediate goods producers $i \in [0, 1]$ with production technology

$$Y_t(i) = A_t N_t(i), \tag{31}$$

where A_t is an aggregate productivity parameter and $N_t(i)$ is the labor input of firm i. A representative final goods producer aggregates differentiated intermediate goods as

$$Y_t = \left(\int_0^1 Y_t(i)^{(\epsilon-1)/\epsilon} di \right)^{\epsilon/(\epsilon-1)}, \qquad \epsilon > 1. \tag{32}$$

As is standard, we assume an employment subsidy that exactly offsets the effect of the firms' monopolistic incentives on production and hiring is in place (paid for with lump-sum transfers).

Intermediate goods firms can reset their prices with probability $1 - \theta$ as in the standard Calvo model.

We assume the initial price distribution is symmetric and coincides with the steady-state price level (see below). Further, we consider the fully rigid-price limit $\theta \uparrow 1$. Consequently, $P_t(i) = P_t = \bar{P}$ for all i and t, i.e., there is neither inflation nor price dispersion.

Since there are neither government purchases nor investment, goods-market clearing is given by $C_t = Y_t$. Furthermore, given that prices are fully rigid, output is demand-determined.

As is standard, we define natural output Y_t^n as the equilibrium output in an equivalent economy with fully flexible prices. Because of the employment subsidy, the flexible-price allocation is efficient and thus Y_t^n is the efficient level of output. Hence, we can find it by simply equating the marginal rate of substitution between consumption and labor to the marginal product of labor. It's straightforward to verify this gives

$$(Y_t^n)^{-\upsilon} = \frac{\zeta_t (N_t^n)^{\nu}}{A_t}. \tag{33}$$

Substituting $N_t^n = Y_t^n / A_t$ and rearranging yields

$$Y_t^n = \left(\frac{A_t^{1+\nu}}{\zeta_t} \right)^{1/(\upsilon+\nu)}.$$

In the baseline model, we assume a process directly for Y_t^n. We can think of Y_t^n as a reduced-form representation of the underlying shocks, and the law of motion for log natural output y_t^n in (3) can be generated by appropriate processes for (A_t, ζ_t).

The market's observation of fundamentals is consistent with this microfoundation: households observe their preference shifters, firms observe productivity, and private-sector agents share a common information set. Demand and natural output are thus known to the private sector.

The (non-stochastic) steady state. A non-stochastic steady state is such that exogenous variables are constant, $A_t = \bar{A}$, $\zeta_t = \bar{\zeta}$, and $\delta_t = \bar{\delta} \equiv \beta \in (0,1)$. Consequently, endogenous variables are constant as well.

We approximate the model around the efficient nonstochastic steady state, in which output coincides with natural output. Since goods-market clearing also implies $\bar{C} = \bar{Y}$, the steady state is characterized by

$$\bar{C} = \bar{Y} = \bar{Y}^n = \left(\frac{\bar{A}^{1+\nu}}{\bar{\zeta}}\right)^{1/(\upsilon+\nu)}, \qquad \bar{N} = \frac{\bar{Y}}{\bar{A}}. \tag{34}$$

Furthermore, evaluating the Euler equation (30) at constant consumption and prices gives

$$\bar{R} = \beta^{-1}. \tag{35}$$

Define $\bar{\rho} \equiv -\log\beta$. The steady-state log nominal rate is then $\log\bar{R} = \bar{\rho}$. Finally, the price level $\bar{P}$ can simply be normalized to 1.

A.2 Approximations

Demand. We now log-linearize the household Euler equation. Define consumption and output as log deviations from their steady-state values,

$$c_t \equiv \log\left(\frac{C_t}{\bar{C}}\right), \qquad y_t \equiv \log\left(\frac{Y_t}{\bar{Y}}\right), \tag{36}$$

while keeping the interest rate and discount rate as simply log levels

$$r_t \equiv \log R_t, \qquad \rho_t \equiv -\log\delta_t, \tag{37}$$

with corresponding log deviations from steady state

$$\widetilde{r}_t \equiv r_t - \bar{\rho}, \qquad \widetilde{\rho}_t \equiv \rho_t - \bar{\rho}. \tag{38}$$

Since

$$\frac{C_t^{-\upsilon}}{\bar{C}^{-\upsilon}} = \exp(-\upsilon c_t), \qquad \frac{\delta_t}{\beta} = \exp(-\widetilde{\rho}_t), \qquad \frac{R_t}{\bar{R}} = \exp(\widetilde{r}_t),$$

we can divide (30) by $\bar{C}^{-\upsilon}$ and use $\beta\bar{R} = 1$ and $P_{t+1} = P_t$ to get

$$\exp(-\upsilon c_t) = \exp(\widetilde{r}_t - \widetilde{\rho}_t)\,\mathbb{E}_t^M\left[\exp(-\upsilon c_{t+1})\right]. \tag{39}$$

Let $\|\xi\|$ denote the size of the log deviations from the steady state. Doing a first-order Taylor expansion of (39) around the steady state gives

$$\begin{aligned} 1 - \upsilon c_t + \mathcal{O}(\|\xi\|^2) &= \left[1 + \widetilde{r}_t - \widetilde{\rho}_t + \mathcal{O}(\|\xi\|^2)\right] \left[1 - \upsilon\, \mathbb{E}_t^M\, c_{t+1} + \mathcal{O}(\|\xi\|^2)\right] \\ &= 1 + \widetilde{r}_t - \widetilde{\rho}_t - \upsilon\, \mathbb{E}_t^M\, c_{t+1} + \mathcal{O}(\|\xi\|^2), \end{aligned} \tag{40}$$

where the cross-terms are second order and thus included in $\mathcal{O}(\|\xi\|^2)$. Rearranging and dropping the second-order remainder yields

$$c_t = \mathbb{E}_t^M\, c_{t+1} - \frac{1}{\upsilon}\left(\widetilde{r}_t - \widetilde{\rho}_t\right). \tag{41}$$

Finally, $\widetilde{r}_t - \widetilde{\rho}_t = (r_t - \bar{\rho}) - (\rho_t - \bar{\rho}) = r_t - \rho_t$. Consequently, the log-linearized Euler equation is

$$c_t = \mathbb{E}_t^M\, c_{t+1} - \frac{1}{\upsilon}\left(r_t - \rho_t\right). \tag{42}$$

Using the equilibrium goods market clearing condition $c_t = y_t$ and defining the intertemporal elasticity of substitution $\gamma \equiv \frac{1}{\upsilon}$, equation (42) becomes

$$y_t = \mathbb{E}_t^M\, y_{t+1} - \gamma(r_t - \rho_t), \tag{43}$$

which is exactly (1).

The central bank's loss function. Using the equilibrium condition $C_t = Y_t$, we can write the household's instantaneous utility as a function of output and labor

$$\frac{Y_t^{1-\upsilon} - 1}{1 - \upsilon} - \frac{\zeta_t}{1 + \nu}\left(N_t\right)^{1+\nu}. \tag{44}$$

Define the log natural output and the log output gap

$$y_t^n \equiv \log\left(\frac{Y_t^n}{\bar{Y}}\right), \qquad x_t \equiv \log\left(\frac{Y_t}{Y_t^n}\right) = y_t - y_t^n, \tag{45}$$

and the natural employment $N_t^n \equiv Y_t^n / A_t$. Since $Y_t = Y_t^n \exp(x_t)$ and $N_t = N_t^n \exp(x_t)$, (44) can be written as

$$\tilde{u}_t(x_t) = \frac{(Y_t^n)^{1-\upsilon} \exp\Big((1-\upsilon)x_t\Big) - 1}{1 - \upsilon} - \frac{\zeta_t (N_t^n)^{1+\nu} \exp\Big((1+\nu)x_t\Big)}{1 + \nu}. \tag{46}$$

Differentiating (46) with respect to the log gap therefore yields

$$\tilde{u}_t'(x_t) = (Y_t^n)^{1-\upsilon} \exp\Big((1-\upsilon)x_t\Big) - \zeta_t (N_t^n)^{1+\nu} \exp\Big((1+\nu)x_t\Big), \tag{47}$$

$$\tilde{u}_t''(x_t) = (1-\upsilon)(Y_t^n)^{1-\upsilon} \exp\Big((1-\upsilon)x_t\Big) - (1+\nu)\zeta_t (N_t^n)^{1+\nu} \exp\Big((1+\nu)x_t\Big). \tag{48}$$

Multiplying the condition defining natural output in (33) by Y_t^n and using $N_t^n = Y_t^n / A_t$ gives

$$(Y_t^n)^{1-\upsilon} = \zeta_t (N_t^n)^{1+\nu}. \tag{49}$$

Evaluating (47) and (48) at the efficient allocation $x_t = 0$ and using (49) gives

$$\tilde{u}_t'(0) = 0, \qquad \tilde{u}_t''(0) = -(\upsilon + \nu)(Y_t^n)^{1-\upsilon} < 0. \tag{50}$$

Thus, a second-order Taylor expansion around $x_t = 0$ yields

$$\tilde{u}_t(x_t) = \tilde{u}_t(0) - \frac{\upsilon + \nu}{2} (Y_t^n)^{1-\upsilon} x_t^2 + \mathcal{O}(|x_t|^3). \tag{51}$$

Note that (51) is an approximation around the time-varying efficient allocation. To express it as a constant-coefficient approximation around the steady state, use the fact that

$$(Y_t^n)^{1-\upsilon} = \bar{Y}^{1-\upsilon} \exp\Big((1-\upsilon) y_t^n\Big) = \bar{Y}^{1-\upsilon} \left[1 + (1-\upsilon) y_t^n + \mathcal{O}(\|\xi\|^2)\right]. \tag{52}$$

Since x_t^2 is already second order, the product $y_t^n x_t^2$ is third order. Hence, to second order, Y_t^n in (51) can be replaced by $\bar{Y}$ and we get

$$\tilde{u}_t(x_t) = \tilde{u}_t(0) - \frac{\upsilon + \nu}{2} (\bar{Y})^{1-\upsilon} x_t^2 + \mathcal{O}(\|\xi\|^3). \tag{53}$$

Having derived an expression for instantaneous utility, we now address the discount factors in the household's overall intertemporal utility (28). Let

$$D_t \equiv \prod_{j=0}^{t-1} \delta_j \text{ for } t \geq 1, \qquad D_0 \equiv 1. \tag{54}$$

Since $\delta_j = \exp(-\rho_j)$, $\beta = \exp(-\bar{\rho})$, and $\tilde{\rho}_j = \rho_j - \bar{\rho}$, we have

$$D_t = \beta^t \exp\left(-\sum_{j=0}^{t-1} \tilde{\rho}_j\right) = \beta^t \left[1 - \sum_{j=0}^{t-1} \tilde{\rho}_j + \mathcal{O}(\|\xi\|^2)\right]. \tag{55}$$

In the household intertemporal utility function, the first-order variation in D_t multiplies the

second-order term x_t^2 in (53) so it is third order. Hence, to second order we can simply replace D_t by β^t. Thus, combining (51)– (55), expected welfare under the central bank's beliefs can be written as

$$\mathcal{W}_0^C = \mathcal{W}_0^{n,C} - \frac{(\upsilon+\nu)\bar{Y}^{1-\upsilon}}{2}\,\mathbb{E}^C \sum_{t=0}^{\infty} \beta^t (y_t - y_t^n)^2 + \mathcal{O}(\|\xi\|^3), \tag{56}$$

where $\mathcal{W}_0^{n,C}$ is expected welfare in the equivalent flexible-price economy. Crucially, it does not depend on monetary policy.

Consequently, after dropping the policy-independent term and the positive multiplicative constant in (56), we can write the central bank's problem of maximizing its expectation of the household's utility, to second order, as

$$\min_{\{r_t\}} \quad \mathbb{E}^C \sum_{t=0}^{\infty} \beta^t (y_t - y_t^n)^2, \tag{57}$$

which is exactly (2).

Quantitative model. An analogous —yet more involved— derivation can be done for the quantitative model in Section 5, since the private-sector block differs from the baseline in two respects: (i) the Calvo parameter θ is left in its general form instead of being taken to the rigid-price limit; and (ii) the model includes a cost-push shock to the Phillips curve (which can be microfounded in the standard way, i.e., through shocks to the elasticity of substitution between varieties and thus to the optimal markup of intermediate goods firms; see, e.g., Galí (2015)). Here, y_t^n still corresponds to the efficient level of output, determined by productivity and preferences over labor. The subsidy offsets the steady-state monopoly distortion, but not its stochastic fluctuations. These fluctuations generate the cost-push term u_t, while the output gap entering the central bank's objective remains $x_t = y_t - y_t^n$.

The central bank's loss function in the quantitative model equals the standard New Keynesian stabilization objective with an additional penalty on deviations of the policy rate from its steady-state level. Such interest-rate stabilization terms are commonly used to prevent optimal policy from implying counterfactually volatile interest rates and can be microfounded using money in utility (see Woodford (2003, Chapter 6) for a detailed discussion).

An additional potential concern is that in our model agents do not hold full-information rational expectations, which may matter for the derivation of the Phillips curve (see, e.g., Preston (2005) for a detailed discussion). In our environment, however, all firms share the same information, and their beliefs are model-consistent conditional expectations with

respect to nested information sets. Consequently, the law of iterated expectations holds and the standard recursive derivation of the New Keynesian Phillips curve remains valid.[44]

B Information Structure

This appendix gives a self-contained construction of the belief structure corresponding to the first two periods of the baseline model. It has three purposes that mirror the following subsections. First, it specifies the basic primitives: the measurable space, each player's (subjective) probability measures and information at different stages. Second, it shows which of the main text's reduced-form assumptions follow from a special case of the familiar structure with a Gaussian prior and signals. Third, it states the cross-belief restrictions that a proper common-prior foundation would add and thus what elements of the baseline model are inconsistent with it while showing the information structure in the quantitative model is grounded in a common prior.

B.1 General Structure

Let

$$\Omega = \mathbb{R}^8, \qquad \mathcal{F} = \mathcal{B}(\Omega),$$

where

$$\omega = (y_0^n, \rho_0, \widehat{y}_0^n, \widehat{\rho}_0, y_1^n, \rho_1, x_{n1}, x_{\rho 1}) \in \Omega.$$

Thus $(\Omega, \mathcal{F})$ is a standard Borel space that supports all random variables relevant for the model.

There are two players, the central bank (C) and the market (M). Each player is endowed with a subjective probability measure P^i over $(\Omega, \mathcal{F})$.

[44] This applies to the exact model in which agents keep track of the full histories. However, to solve it we use a finite-dimensional state vector that truncates the histories (see Appendix G.1) and the information sets generated by the truncated state vectors are in general not nested. We analyze the sensitivity to this approximation by increasing the retained history from $H = 8$ to $H = 12$ and re-estimating the model and find that our main results still hold (see Appendix H.1 for details). By itself, this does not guarantee that the economy is close to the one with full histories, especially given that both the $H = 8$ and $H = 12$ approximations disregard the additional terms that should appear in the Phillips curve given that the information sets are not nested. We also conduct an additional exercise that holds the baseline parameters and reported policies fixed and computes the IS and Phillips-curve residuals evaluated using full-history forecasts of the baseline allocation processes instead of the truncated history ones. Reassuringly, the residuals are fairly small (see Appendix H.1 for details).

For convenience, partition the coordinates of ω into the $t = 0$ fundamental vector, the central bank's $t = 0$ assessment vector, the $t = 1$ fundamentals, and the $t = 1$ private signals:

$$X_0 \equiv (y_0^n, \rho_0), \qquad A_0 \equiv (\widehat{y}_0^n, \widehat{\rho}_0), \qquad X_1 \equiv (y_1^n, \rho_1), \qquad Z_1 \equiv (x_{n1}, x_{\rho 1}).$$

Thus, equivalently,

$$\omega = (X_0, A_0, X_1, Z_1).$$

The primitive uncertainty therefore concerns the contemporaneous fundamentals, the central bank's $t = 0$ assessments, the future fundamentals, and the private $t = 1$ signals.

The information available to the central bank and the market at the relevant points in time is represented by the following σ-fields:

$$\mathcal{I}_0^C = \sigma(\widehat{y}_0^n, \widehat{\rho}_0), \qquad \mathcal{I}_{0-}^M = \sigma(y_0^n, \rho_0),$$

and

$$\mathcal{I}_1^C = \sigma(\widehat{y}_0^n, \widehat{\rho}_0, x_{n1}, x_{\rho 1}).$$

The σ-fields and the subjective laws of each player are common knowledge. Note that $\mathcal{I}_1^C$ contains no endogenous variable: the central bank does not observe the $t = 0$ output realization before choosing r_1. This is substantive, since y_0 is informative about the central bank's $t = 0$ assessment errors. Thus, relaxing this assumption would have two implications. First, in the simplest case where $\lambda_n = 0$, y_0 fully reveals ρ_0, so even if $\varphi_\rho > 0$, the market does not use the demand-assessment $\widehat{\rho}_0$ to learn about the demand component of future policy (and thus, future output): it knows the central bank's assessment at the time r_1 is chosen will be $\bar{\rho} + \varphi_\rho(\rho_0 - \bar{\rho})$. This would imply that the entire belief-revision channel goes through natural output, reducing the richness of the analysis (since the main mechanism reduces to the case where $\varphi_\rho = 0$).

Second, and much more importantly, if $\lambda_n > 0$, y_0 becomes an imperfect signal for the central bank that contains information about both y_0^n (via the market's ex ante rate forecast and thus the size of the policy surprise) and ρ_0. Crucially, the relative informativeness about the two components *depends on the policy rule itself* (the composition b, in particular). This not only makes the problem much less tractable, but also introduces a new channel through which the central bank may want to distort its policy rule: it may want to make y_0 more informative about ρ_0 at the expense of informativeness about y_0^n (or vice versa). This mechanism is not the object of this paper. Thus, we assume that the central bank does not observe y_0 before choosing r_1, and thus cannot use it to learn about its own assessment errors. Similarly, in the quantitative model it is assumed that endogenous outcomes (π_t, y_t)

are never observed. Note that the precision of the exogenous signals remains free to be determined by the estimation, so this assumption does not restrict how informed the central bank can be about the fundamentals.

Hence, the central bank's $t = 0$ private information is summarized by its assessment vector A_0. At $t = 1$ the central bank additionally observes the private signal vector Z_1. The market observes the contemporaneous fundamentals X_0.

For any measurable $t = 0$ policy rule

$$g_0 : \mathbb{R}^2 \to \mathbb{R},$$

define the induced policy decision

$$r_0(\omega) = g_0(\widehat{y}_0^n, \widehat{\rho}_0).$$

Given such a rule, the market's post-decision σ-field is

$$\mathcal{I}_{0+}^{M,g_0} := \sigma(y_0^n, \rho_0, r_0),$$

and its $t = 1$ σ-field is

$$\mathcal{I}_1^{M,g_0} := \sigma(y_0^n, \rho_0, r_0, y_1^n, \rho_1).$$

For any measurable policy rule g_0, pre- and post-decision beliefs are represented by the regular conditional probabilities[45]

$$\pi_0^C(\omega) := P^C(\cdot \mid \mathcal{I}_0^C)(\omega), \qquad \pi_{0-}^M(\omega) := P^M(\cdot \mid \mathcal{I}_{0-}^M)(\omega),$$

$$\pi_{0+}^{M,g_0}(\omega) := P^M(\cdot \mid \mathcal{I}_{0+}^{M,g_0})(\omega),$$

and $t = 1$ beliefs are represented by

$$\pi_1^C(\omega) := P^C(\cdot \mid \mathcal{I}_1^C)(\omega), \qquad \pi_1^{M,g_0}(\omega) := P^M(\cdot \mid \mathcal{I}_1^{M,g_0})(\omega).$$

Equivalently, there exists a stochastic kernel K_{0+}^{M,g_0} on $(\Omega, \mathcal{F})$ given $\mathbb{R}^3$ such that

$$\pi_{0+}^{M,g_0}(\omega) = K_{0+}^{M,g_0}\left(y_0^n(\omega), \rho_0(\omega), r_0(\omega)\right) \qquad P^M\text{-a.s.}$$

Note that since

$$r_0 = g_0(\widehat{y}_0^n, \widehat{\rho}_0),$$

[45] These are well-defined since $(\Omega, \mathcal{F})$ is a standard Borel space.

the market's posterior after the policy decision depends on the policy rule g_0: the observed action is an endogenous signal of the central bank's private assessment.

We restrict the central bank's subjective law P^C so that its $t = 0$ assessment vector A_0 coincides with the central bank's posterior mean of the contemporaneous fundamental vector X_0 given its private information, that is,

$$A_0 = \mathbb{E}^C[X_0 \mid \mathcal{I}_0^C] \qquad P^C\text{-a.s.},$$

or component-wise,

$$\widehat{y}_0^n = \mathbb{E}^C[y_0^n \mid \mathcal{I}_0^C], \qquad \widehat{\rho}_0 = \mathbb{E}^C[\rho_0 \mid \mathcal{I}_0^C] \qquad P^C\text{-a.s.}$$

Equivalently, defining the $t = 0$ estimation-error vector

$$e_0 \equiv X_0 - A_0,$$

we have

$$\mathbb{E}^C[e_0 \mid \mathcal{I}_0^C] = 0 \qquad P^C\text{-a.s.}$$

Thus, under P^C, $(\widehat{y}_0^n, \widehat{\rho}_0)$ are posterior means conditional on the central bank's $t = 0$ information.[46]

The primitive information structure is therefore given by

$$(\Omega, \mathcal{F}), \qquad (\mathcal{I}_0^C, \mathcal{I}_{0-}^M, \mathcal{I}_1^C), \qquad (P^C, P^M).$$

For any measurable policy rule g_0, this structure induces the observable action r_0, the information fields $\mathcal{I}_{0+}^{M,g_0}$ and $\mathcal{I}_1^{M,g_0}$, and the associated conditional belief kernels, which together pin down the players' hierarchy of conditional beliefs.

B.2 Gaussian Beliefs and Further Assumptions in the Main Text

The reduced-form model in the main text specializes the structure above to a Gaussian distribution for the fundamental, assessment, and signal blocks. In this subsection we provide the necessary background and assumptions to support the Gaussian specialization.

[46] Since the market's $t = 0$ information is given by the observed fundamentals X_0, it would be of no use to define separate market $t = 0$ assessments.

Throughout, expectations taken under the central bank's and the market's subjective laws are denoted, respectively, by[47]

$$\mathbb{E}^C[\cdot] \equiv \mathbb{E}^{P^C}[\cdot], \qquad \mathbb{E}^M[\cdot] \equiv \mathbb{E}^{P^M}[\cdot],$$

and analogously for conditional variances and covariances.

Under P^C, natural output follows the following Gaussian distribution at $t = 0$:

$$y_0^n \sim \mathcal{N}(0, \sigma_{y0}^2),$$

while the discount rate follows

$$\rho_0 \sim \mathcal{N}(\bar{\rho}, \sigma_{\rho 0}^2).$$

At date 0, define the central bank's posterior errors by

$$\eta_0 \equiv y_0^n - \widehat{y}_0^n, \qquad \varepsilon_0 \equiv \rho_0 - \widehat{\rho}_0, \tag{58}$$

where the natural-output block $(\widehat{y}_0^n, \eta_0)$ and the discount-rate block $(\widehat{\rho}_0, \varepsilon_0)$ are jointly Gaussian under P^C, independent across blocks, with

$$\mathbb{E}^C[\eta_0] = \mathbb{E}^C[\varepsilon_0] = 0,$$

and

$$\mathrm{Var}^C(\widehat{y}_0^n) = \sigma_{n,C}^2, \qquad \mathrm{Var}^C(\widehat{\rho}_0) = \sigma_{\rho,C}^2, \qquad \mathrm{Var}^C(\eta_0) = \sigma_{\eta,C}^2, \qquad \mathrm{Var}^C(\varepsilon_0) = \sigma_{\varepsilon,C}^2.$$

Let

$$e_0 \equiv (\eta_0, \varepsilon_0).$$

Since $\mathcal{I}_0^C = \sigma(A_0)$ and $A_0 = \mathbb{E}^C[X_0 \mid \mathcal{I}_0^C]$,

$$\mathbb{E}^C[e_0 \mid \mathcal{I}_0^C] = 0.$$

It follows that

$$\mathrm{Cov}^C(\widehat{y}_0^n, \eta_0) = 0, \qquad \mathrm{Cov}^C(\widehat{\rho}_0, \varepsilon_0) = 0.$$

[47] Unlike in common-prior information structures, expectations of two agents may differ not only because of their different information sets, but also because of their different subjective probability measures. Thus, we retain the latter as a superscript even when specifying the conditioning information set.

Under joint Gaussianity, these orthogonality restrictions are also sufficient for the conditional mean restriction above. Together with (58), these orthogonality conditions imply the variance decompositions

$$\sigma_{y0}^2 = \sigma_{n,C}^2 + \sigma_{\eta,C}^2, \qquad \sigma_{\rho 0}^2 = \sigma_{\rho,C}^2 + \sigma_{\varepsilon,C}^2.$$

Thus, the unconditional fundamental variances are implied by the assessment and posterior-error variances instead of being additional free parameters. Equivalently, the central bank's $t = 0$ conditional law of contemporaneous fundamentals given its private assessments is

$$y_0^n \mid \mathcal{I}_0^C \sim \mathcal{N}(\widehat{y}_0^n, \sigma_{\eta,C}^2), \qquad \rho_0 \mid \mathcal{I}_0^C \sim \mathcal{N}(\widehat{\rho}_0, \sigma_{\varepsilon,C}^2),$$

which gives (5) and (6) in the main text.

The $t = 1$ variables satisfy

$$u_{n1} := y_1^n - \varphi_n y_0^n, \qquad u_{\rho 1} := \rho_1 - \bar{\rho} - \varphi_\rho(\rho_0 - \bar{\rho}),$$

so $u_{n1} = \varepsilon_n$ and $u_{\rho 1} = \varepsilon_\rho$ in (3)–(4), and

$$\nu_{n1} := x_{n1} - y_1^n, \qquad \nu_{\rho 1} := x_{\rho 1} - \rho_1,$$

where

$$u_{n1} \sim \mathcal{N}(0, \sigma_{un1}^2), \qquad u_{\rho 1} \sim \mathcal{N}(0, \sigma_{u\rho 1}^2),$$
$$\nu_{n1} \sim \mathcal{N}(0, \sigma_{\nu n1}^2), \qquad \nu_{\rho 1} \sim \mathcal{N}(0, \sigma_{\nu\rho 1}^2),$$

independently of everything else. Thus, fundamentals (y_0^n, ρ_0) evolve as Gaussian AR(1) processes with persistences φ_n and φ_ρ, respectively, and the central bank observes signals x_{n1} and $x_{\rho 1}$ at $t = 1$ that equal the fundamentals plus white Gaussian noise. Hence,

$$y_1^n \mid \mathcal{I}_0^C \sim \mathcal{N}\left(\varphi_n \widehat{y}_0^n, \varphi_n^2 \sigma_{\eta,C}^2 + \sigma_{un1}^2\right),$$

$$\rho_1 \mid \mathcal{I}_0^C \sim \mathcal{N}\left(\bar{\rho} + \varphi_\rho(\widehat{\rho}_0 - \bar{\rho}), \varphi_\rho^2 \sigma_{\varepsilon,C}^2 + \sigma_{u\rho 1}^2\right).$$

Consequently, the $t = 1$ posterior means are

$$\widehat{y}_1^n := \mathbb{E}^C[y_1^n \mid \mathcal{I}_1^C] = (1 - \kappa_{n1})\varphi_n \widehat{y}_0^n + \kappa_{n1} x_{n1},$$

$$\widehat{\rho}_1 := \mathbb{E}^C[\rho_1 \mid \mathcal{I}_1^C] = \bar{\rho} + (1 - \kappa_{\rho 1})\varphi_\rho(\widehat{\rho}_0 - \bar{\rho}) + \kappa_{\rho 1}(x_{\rho 1} - \bar{\rho}),$$

where

$$\kappa_{n1} \equiv \frac{\varphi_n^2 \sigma_{\eta,C}^2 + \sigma_{un1}^2}{\varphi_n^2 \sigma_{\eta,C}^2 + \sigma_{un1}^2 + \sigma_{\nu n1}^2},$$
$$\kappa_{\rho 1} \equiv \frac{\varphi_\rho^2 \sigma_{\varepsilon,C}^2 + \sigma_{u\rho 1}^2}{\varphi_\rho^2 \sigma_{\varepsilon,C}^2 + \sigma_{u\rho 1}^2 + \sigma_{\nu\rho 1}^2}$$

are the standard Kalman gains.

Under P^M, the conditional law of the central bank's $t = 0$ assessments given the market observables (y_0^n, ρ_0) is instead given by[48]

$$\begin{pmatrix} \widehat{y}_0^n \\ \widehat{\rho}_0 \end{pmatrix} \,\Big|\, \mathcal{I}_{0-}^M \sim \mathcal{N}\left(\begin{pmatrix} \lambda_n y_0^n \\ \bar{\rho} + \lambda_\rho(\rho_0 - \bar{\rho}) \end{pmatrix}, \begin{pmatrix} V_{n0} & 0 \\ 0 & sV_{\rho 0} \end{pmatrix} \right).$$

The parameters λ_n and λ_ρ govern how the market centers the central bank's $t = 0$ assessments on the contemporaneous fundamentals. If $\lambda_z = 1$ for some $z \in \{n, \rho\}$, the market thinks the corresponding assessment is correctly centered on average. If, on the other hand, $\lambda_z = 0$, it thinks that assessment is orthogonal to the contemporaneous fundamental. Both of these are knife-edge cases: a standard common-prior signal structure generally produces coefficients strictly between zero and one, as discussed below.

The market and the central bank may not agree on the distribution of the central bank's $t = 0$ assessments conditional on the contemporaneous fundamentals. We assume, however, that under P^M, the $t = 1$ fundamental innovations $u_{n1}, u_{\rho 1}$ and signal errors $\nu_{n1}, \nu_{\rho 1}$ have the Gaussian distributions specified above, are mutually independent, and are independent of (X_0, A_0). The market also knows that the central bank maps (A_0, Z_1) into its $t = 1$ assessments according to the filtering rules above. Hence,

$$\mathbb{E}^M[\widehat{y}_0^n \mid \mathcal{I}_{0-}^M] = \lambda_n y_0^n,$$

$$\mathbb{E}^M[\widehat{\rho}_0 \mid \mathcal{I}_{0-}^M] = \bar{\rho} + \lambda_\rho(\rho_0 - \bar{\rho}),$$
$$\mathbb{E}^M[\widehat{y}_1^n \mid \mathcal{I}_{0-}^M] = \left((1 - \kappa_{n1})\lambda_n + \kappa_{n1}\right) \varphi_n y_0^n,$$
$$\mathbb{E}^M[\widehat{\rho}_1 \mid \mathcal{I}_{0-}^M] = \bar{\rho} + \left((1 - \kappa_{\rho 1})\lambda_\rho + \kappa_{\rho 1}\right) \varphi_\rho(\rho_0 - \bar{\rho}).$$

[48] Since in the model the market observes fundamentals before any decision takes place, specifying their prior distribution under P^M is not necessary. However, for completeness, one can assume that it is the same as the central bank's.

Moreover, since the market's $t = 0$ conditional law is Gaussian and the $t = 1$ updating rule is linear, under P^M,

$$\begin{pmatrix} \widehat{y}_0^n \\ \widehat{y}_1^n \end{pmatrix} \,\Big|\, \mathcal{I}_{0-}^M \sim \mathcal{N}\left(\begin{pmatrix} \mathbb{E}^M[\widehat{y}_0^n \mid \mathcal{I}_{0-}^M] \\ \mathbb{E}^M[\widehat{y}_1^n \mid \mathcal{I}_{0-}^M] \end{pmatrix}, \begin{pmatrix} V_{n0} & c_n \\ c_n & V_{n1} \end{pmatrix} \right),$$

where

$$c_n \equiv \mathrm{Cov}^M(\widehat{y}_0^n, \widehat{y}_1^n \mid \mathcal{I}_{0-}^M) = (1 - \kappa_{n1})\varphi_n V_{n0},$$

and

$$V_{n1} \equiv \mathrm{Var}^M(\widehat{y}_1^n \mid \mathcal{I}_{0-}^M) = (1 - \kappa_{n1})^2 \varphi_n^2 V_{n0} + \kappa_{n1}^2(\sigma_{un1}^2 + \sigma_{\nu n1}^2).$$

Likewise,

$$\begin{pmatrix} \widehat{\rho}_0 \\ \widehat{\rho}_1 \end{pmatrix} \,\Big|\, \mathcal{I}_{0-}^M \sim \mathcal{N}\left(\begin{pmatrix} \mathbb{E}^M[\widehat{\rho}_0 \mid \mathcal{I}_{0-}^M] \\ \mathbb{E}^M[\widehat{\rho}_1 \mid \mathcal{I}_{0-}^M] \end{pmatrix}, \begin{pmatrix} sV_{\rho 0} & c_\rho(s) \\ c_\rho(s) & V_{\rho 1}(s) \end{pmatrix} \right),$$

where

$$c_\rho(s) \equiv \mathrm{Cov}^M(\widehat{\rho}_0, \widehat{\rho}_1 \mid \mathcal{I}_{0-}^M) = (1 - \kappa_{\rho 1})\varphi_\rho s V_{\rho 0},$$

and

$$V_{\rho 1}(s) \equiv \mathrm{Var}^M(\widehat{\rho}_1 \mid \mathcal{I}_{0-}^M) = (1 - \kappa_{\rho 1})^2 \varphi_\rho^2 s V_{\rho 0} + \kappa_{\rho 1}^2(\sigma_{u\rho 1}^2 + \sigma_{\nu\rho 1}^2).$$

The parameter s scales the market's perceived conditional variance of the central bank's $t = 0$ demand assessment. Given the normalization $V_{n0} = V_{\rho 0}$ imposed below, $s = 1$ is the symmetric-variance benchmark.

Further assumptions in the main text. In the main text, we take the above Gaussian information structure and further specialize it as follows.

- Normalize $V_{n0} = V_{\rho 0} = 1$.
- Assume the central bank receives no additional signal at $t = 1$. Equivalently, remove Z_1 from $\mathcal{I}_1^C$ and set $\kappa_{n1} = \kappa_{\rho 1} = 0$ in the updating formulas above.
- Set $\lambda_\rho = 1$.

The first assumption implies that the variance ratio of the market's beliefs about the central bank's demand and natural-output assessments reduces to s:

$$\frac{\mathrm{Var}^M(\widehat{\rho}_0 \mid \mathcal{I}_{0-}^M)}{\mathrm{Var}^M(\widehat{y}_0^n \mid \mathcal{I}_{0-}^M)} = \frac{V_{\rho 0}}{V_{n0}} s = s.$$

The second assumption entails that the central bank receives no new information between its $t = 0$ and $t = 1$ policy decisions. It therefore gives

$$\widehat{y}_1^n = \varphi_n \widehat{y}_0^n, \qquad \widehat{\rho}_1 = \bar{\rho} + \varphi_\rho(\widehat{\rho}_0 - \bar{\rho}),$$

as assumed in (7). In particular, the $t = 1$ variances and covariances are implied by each fundamental's corresponding persistence coefficient,

$$\mathrm{Var}^M(\widehat{y}_1^n \mid \mathcal{I}_{0-}^M) = \varphi_n^2, \qquad \mathrm{Cov}^M(\widehat{y}_0^n, \widehat{y}_1^n \mid \mathcal{I}_{0-}^M) = \varphi_n,$$

$$\mathrm{Var}^M(\widehat{\rho}_1 \mid \mathcal{I}_{0-}^M) = s\varphi_\rho^2, \qquad \mathrm{Cov}^M(\widehat{\rho}_0, \widehat{\rho}_1 \mid \mathcal{I}_{0-}^M) = s\varphi_\rho.$$

The restriction is imposed for expositional simplicity. More generally, new $t = 1$ information would make the persistence of the central bank's assessments weakly smaller than the persistence of the corresponding fundamentals, because the central bank would place less weight on its $t = 0$ assessment. The quantitative analysis in Section 5 relaxes this since the central bank observes new signals each period.

The third assumption is a reduced-form belief specialization for analytical transparency, and its purpose is to preserve a property of the environment that makes the analysis cleaner.

Concretely, setting $\lambda_\rho = 1$ means that the market treats the central bank's initial demand assessment as correctly centered:

$$\mathbb{E}^M[\widehat{\rho}_0 \mid \mathcal{I}_{0-}^M] = \rho_0.$$

It also implies

$$\mathbb{E}^M[\widehat{\rho}_1 \mid \mathcal{I}_{0-}^M] = \mathbb{E}^M[\rho_1 \mid \mathcal{I}_{0-}^M].$$

This is arguably a natural feature to assume, although exact centering is not generic under a standard finite-precision prior-and-signal structure. This centering assumption is chosen because it delivers the clean discretionary coefficient of one on the central bank's demand assessment. The reason is that, when $\lambda_\rho = 1$, the demand component of the policy surprise $r_0 - \mathbb{E}_0^M[r_0]$ has conditional mean zero under the central bank's beliefs. Its stochastic component is an estimation error that enters output additively, so certainty equivalence applies. The discretionary central bank therefore chooses a unit coefficient on $\widehat{\rho}_0$: since the discount rate and the policy rate enter output with the same coefficient γ, the central bank wants to offset its demand assessment one for one. A central bank with commitment, by contrast, remains free to choose $a \neq 1$ because it internalizes how the policy rule affects the pass-through of assessment errors.

B.3 Common-Prior Restrictions

The information structure above features heterogeneous priors: P^C and P^M are different subjective probability measures. While this does not compromise either player's Bayesian rationality, it does entail more flexibility than standard common-prior models. As discussed by Morris (1995), the common-prior assumption is a substantive restriction on the relation between individuals' beliefs, and there are good reasons to doubt its validity in many economically relevant real-world contexts. Nevertheless, it remains a standard assumption in economic theory, and even more so in macroeconomics. Thus, this section details the implications a common-prior foundation would have in the context of the model above.

Market's centering and uncertainty. Since the model assumes independence between blocks, the implications of a common-prior foundation are unrelated to the links between them. Thus, consider an arbitrary one in isolation. Let P be a proper shared probability measure and $x \in \mathbb{R}$ be the square-integrable and nondegenerate fundamental with prior mean zero (we assume this for simplicity; an arbitrary common knowledge prior mean does not change the argument materially). Finally, let $\mathcal{I}^C$ denote the central bank's information. By construction, the central bank's assessment of x satisfies

$$\widehat{x} = \mathbb{E}_P[x \mid \mathcal{I}^C].$$

The estimation error $x - \widehat{x}$ is orthogonal to every $\mathcal{I}^C$-measurable random variable, so in particular

$$0 = \mathbb{E}_P[(x - \widehat{x})\widehat{x}],$$

which implies $\mathbb{E}_P[x\widehat{x}] = \mathbb{E}_P[\widehat{x}^2]$. Moreover, the law of iterated expectations gives

$$\mathbb{E}_P[\widehat{x}] = \mathbb{E}_P\left[\mathbb{E}_P[x \mid \mathcal{I}^C]\right] = \mathbb{E}_P[x] = 0.$$

Putting the above equalities together yields

$$\mathrm{Cov}_P(x, \widehat{x}) = \mathrm{Var}_P(\widehat{x}). \tag{59}$$

Now, as in our model, suppose the market observes x and

$$\mathbb{E}_P[\widehat{x} \mid x] = \lambda x \tag{60}$$

for some $\lambda \in \mathbb{R}$. Note that the fact that this holds under the common-prior measure P is crucial for the argument below. In particular, it implies

$$\begin{aligned}\mathrm{Cov}_P(x, \widehat{x}) &= \mathbb{E}_P[x\widehat{x}] \\ &= \mathbb{E}_P\left[x\, \mathbb{E}_P[\widehat{x} \mid x]\right] \\ &= \lambda\, \mathbb{E}_P[x^2] = \lambda\, \mathrm{Var}_P(x). \end{aligned} \tag{61}$$

Combining this with (59) gives

$$\mathrm{Var}_P(\widehat{x}) = \lambda\, \mathrm{Var}_P(x), \tag{62}$$

which implies $\lambda \geq 0$. Consequently, the market's centering parameter λ, prior uncertainty about fundamentals and prior uncertainty about the central bank's assessment are linked. Furthermore, the law of total variance and (62) imply

$$\begin{aligned}\lambda\, \mathrm{Var}_P(x) &= \mathrm{Var}_P(\widehat{x}) \\ &= \mathrm{Var}_P\left(\mathbb{E}_P[\widehat{x} \mid x]\right) + \mathbb{E}_P\left[\mathrm{Var}_P(\widehat{x} \mid x)\right] \\ &= \lambda^2\, \mathrm{Var}_P(x) + \mathbb{E}_P\left[\mathrm{Var}_P(\widehat{x} \mid x)\right], \end{aligned}$$

and hence

$$\mathbb{E}_P\left[\mathrm{Var}_P(\widehat{x} \mid x)\right] = \lambda(1-\lambda)\, \mathrm{Var}_P(x), \tag{63}$$

which further implies $\lambda \in [0, 1]$. Whenever the conditional variance is constant in x, as in the Gaussian case, (63) pins it down directly as a function of λ and the prior variance. Crucially, without a common prior, (59) would hold under P^C, while (61) would hold under P^M so there would be extra degrees of freedom in the relation above.

The above implies that the endpoint values for λ have stark consequences. If $\lambda = 0$, equation (62) implies that $\widehat{x}$ is constant P-almost surely. If $\lambda = 1$, then $\mathrm{Var}_P(\widehat{x}) = \mathrm{Var}_P(x)$ and thus, by orthogonality

$$\mathrm{Var}_P(x) = \mathrm{Var}_P(\widehat{x}) + \mathrm{Var}_P(x - \widehat{x}) \Longrightarrow x = \widehat{x}$$

P-almost surely. Thus, neither centering endpoint is compatible with the market having uncertainty about the central bank's assessment in a common-prior environment.

While the argument above is general, below we develop it for the special case of a Gaussian prior-plus-signal at $t = 0$. We do so because it makes the intuition for the above constraint transparent and further illustrates that *every* interior λ can be obtained that way.

Gaussian specialization. Let the common-prior block be such that

$$x \sim \mathcal{N}(0, V), \qquad z = x + \nu, \qquad \nu \sim \mathcal{N}(0, R), \qquad x \perp \nu,$$

where x is the fundamental, z the signal observed by the central bank, ν its noise, and $V, R > 0$. Standard Bayesian updating implies

$$\widehat{x} = \mathbb{E}[x \mid z] = \lambda z, \qquad \lambda \equiv \frac{V}{V + R} \in (0, 1).$$

If the market observes the fundamental x, its beliefs about the central bank's assessment satisfy

$$\mathbb{E}[\widehat{x} \mid x] = \lambda x, \qquad \mathrm{Var}(\widehat{x} \mid x) = \lambda^2 R = \lambda(1 - \lambda)V.$$

Thus, every interior value of the centering parameter, $\lambda \in (0, 1)$, is consistent with a common prior in the Gaussian prior and signal setting. Further, the intuition for why the endpoints are not is transparent: as $\lambda \to 0$ or $\lambda \to 1$, the market's uncertainty regarding the central bank's assessment approaches zero. If the central bank puts zero weight on its private information, the common prior mean is common knowledge and the market knows its assessment. Analogously, if the central bank puts full weight on its signal, the signal has no noise and equals the fundamental. Since the market knows the fundamental, it also knows the central bank's assessment.

The baseline model, on the other hand, leaves the centering and conditional variance as free parameters and, further, assumes the ρ-centering equals 1, which is not consistent with non-zero uncertainty about the central bank's assessment under a common prior. The baseline model is thus not compatible with a common prior in its current form, but it would become compatible if the ρ-centering were strictly less than 1, $\lambda_n > 0$, and the above relationships between the centering and conditional variance were imposed. In fact, the quantitative model in Section 5 is fully consistent with a common prior since the centering parameters are the ones implied by the rest of the information structure (in particular, they are consistent with the steady-state Kalman filter; see Appendix G for details). Concretely, a single Gaussian distribution generates fundamentals and the central bank's signal errors, and its parameters and the signal structure are common knowledge.

An important implication of the lack of a common-prior foundation for the model is that there is no natural ex ante welfare measure, since there is more than one possible ex ante distribution under which to compute expected values. We discuss this issue below.

Implications for "welfare" statements. In Section 3, the central bank's objective and all ex ante loss comparisons are evaluated under the central bank's ex ante measure P^C. Absent a common prior, it is not straightforward to take these as actual welfare statements, even if one accepts the stylized economy under consideration and takes at face value the central bank's loss as a reasonable measure of welfare losses in each state of the world. Thus, the claims there are stated in terms of central bank loss only. Going beyond that would require taking a stance on the ex ante distribution, which is beyond the scope of the paper.

Importantly, this is not an issue in the richer model of Section 5, which does admit a common-prior foundation.[49]

C Game Representation and Solution Concepts

This appendix represents the $t = 0$ interaction between the central bank and the market in Section 2 as an extensive-form game. It restates the discretion and commitment solution concepts of Section 3 in terms of strategies and unilateral deviations, taking the sequentially optimal continuation from $t = 1$ onward as given.

The definitions use only the primitives of Appendix B.1—the measurable space $(\Omega, \mathcal{F})$, the σ-fields representing each player's information, and the subjective laws (P^C, P^M). The discussion of linear rules and the pointwise version of the market's posterior also uses the Gaussian specialization of Appendix B.2. Since the parameter s is a known constant, it is suppressed from the notation.

C.1 The Game

There are two players, the central bank (C) and the market (M).[50] Recall the partition $\omega = (X_0, A_0, X_1, Z_1)$, where $X_0 = (y_0^n, \rho_0)$ are the $t = 0$ fundamentals, $A_0 = (\widehat{y}_0^n, \widehat{\rho}_0)$ the central bank's $t = 0$ assessments, X_1 the $t = 1$ fundamentals and Z_1 the central bank's $t = 1$ signals. The extensive form is as follows.[51]

[49] Of course, the appropriateness of the central bank's loss function as a proxy for welfare losses in the economy depends on the extent to which the linear model is a legitimate approximation of a more standard one. On that, see Appendix A, particularly the discussion of the quantitative model at the end.

[50] For the purposes of this section, we take output to be "mechanically" determined by the IS equation (1) (which, as Appendix A shows, is a linear approximation of the household's utility-maximizing behavior together with market clearing), so the market *actions* are its forecasts of next-period output. Given this and the fact that the market does not have incentives to alter its forecast even if it could affect aggregate expectations as argued in Remark 2 below, we can safely think of the market as being one player and not an aggregation of a continuum of agents.

[51] The main focus of our analysis is the $t = 0$ game since that is where the policy rate plays a signaling role that makes the game-theoretic formulation useful. Thus, we omit some notation for the $t = 1$ game and fully omit the $t \geq 2$ stages.

1. A state ω is realized. The central bank observes the realization of A_0 and the market observes that of X_0.

2. The central bank chooses $r_0 \in \mathbb{R}$.

3. The market observes the realization of r_0 and chooses a forecast $f_0 \in \mathbb{R}$ of y_1. Output at $t = 0$ is then determined by the IS curve (1), $y_0 = f_0 - \gamma(r_0 - \rho_0)$.

4. At $t = 1$, the central bank observes Z_1, with information $\mathcal{I}_1^C$, and chooses r_1. Then, the market observes X_1 and forecasts y_2, and output at $t = 1$ is determined by (1).

5. For $t \geq 2$ it is common knowledge that $(y_t^n, \rho_t) = (0, \bar{\rho})$ and thus a complete-information version of the stage game repeats. We omit the resulting continuation strategies below since the structure is fully standard.

A strategy for the central bank is a pair of measurable maps

$$g_0 : \mathbb{R}^2 \to \mathbb{R}, \qquad g_1 : \mathbb{R}^4 \to \mathbb{R},$$

where $r_0 = g_0(\widehat{y}_0^n, \widehat{\rho}_0)$ and $r_1 = g_1(\widehat{y}_0^n, \widehat{\rho}_0, x_{n1}, x_{\rho 1})$.[52] Equivalently, in the terminology of Appendix B, r_t is $\mathcal{I}_t^C$-measurable. A strategy for the market is a measurable map

$$\mu : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \qquad f_0 = \mu(X_0, r_0),$$

together with its (degenerate) $t = 1$ counterpart. Note that μ is defined at *every* $r_0 \in \mathbb{R}$. The central bank's ex ante loss and the market's $t = 0$ forecast loss are

$$\mathcal{L}^C = \mathbb{E}^C \left[\sum_{t \geq 0} \beta^t \left(y_t - y_t^n \right)^2 \right], \qquad \ell_0^M = \left(f_0 - y_1 \right)^2,$$

respectively. The market minimizes the conditional expectation of ℓ_0^M, while the central bank minimizes $\mathcal{L}^C$.

Remark 2 (The market's forecast does not affect its target)**.** *By Lemma 1, $y_1 = \widehat{y}_1^n + \gamma(\rho_1 - \widehat{\rho}_1)$ depends only on the central bank's $t = 1$ assessments and the contemporaneous demand fundamental (recall we assumed the central bank does not observe y_0). Hence, y_1 is unaffected by the $t = 0$ forecast f_0 and thus the market has no incentive to distort its forecast in order to influence the object it is forecasting. This implies that (i) the market's best response is its*

[52] In the baseline specialization, there are no $t = 1$ informative signals so g_1 depends on A_0 alone, i.e., it is a function from $\mathbb{R}^2$ to $\mathbb{R}$ (see Appendix B). We keep the general form and thus the slight abuse of notation for clarity.

conditional expectation of y_1 and (ii) the analysis that takes the market to be one agent is identical to the one that would arise from a continuum of atomistic forecasters.[53]

Write g_1^* for the $t = 1$ strategy in Lemma 1. It is the unique sequentially optimal continuation and does not depend on g_0. Since our analysis of different solution concepts pertains only to the $t = 0$ game, y_1 below is the outcome induced by g_1^* unless stated.

C.2 Market-Consistent Profiles

As argued above, the market's optimal strategy (forecast) is always its conditional expectation of y_1. Thus, our solution concepts of interest take this as given. The following definition formalizes it.

Definition 2. *Given a policy rule g_0, let $\mathcal{M}(g_0)$ denote the set of optimal market strategies, i.e., μ such that for almost every (X_0, r_0) under the joint law induced by g_0 and P^M,*

$$\mu(X_0, r_0) \in \arg\min_{f \in \mathbb{R}} \mathbb{E}^M \left[(f - y_1)^2 \,\middle|\, X_0, r_0 \right].$$

The set of market-consistent profiles is $\mathcal{E} \equiv \{(g_0, \mu) : \mu \in \mathcal{M}(g_0)\}$.

Because the loss is quadratic, $\mu \in \mathcal{M}(g_0)$ if and only if

$$\mu(x, r) = \mathbb{E}^M \left[y_1 \mid X_0 = x, r_0 = r \right] \qquad \text{for } P^M\text{-a.e. } (x, r). \tag{64}$$

For linear rules, the corresponding formula is given by Proposition 1. Note that conditional expectations are defined only almost surely. To determine the potential deviations of the central bank and thus characterize equilibria (however defined; see below), we need a pointwise version. For the constant rule, Bayes' rule determines the forecast at $r = \bar{\rho}$ but leaves forecasts following deviations unspecified. The following definition specifies the market's forecast at every fundamental realization and policy rate, using the Gaussian updating formula for nonconstant rules and the natural "no-update" convention for the constant rule.

Definition 3. *Let $g_0^{a,K}$ denote a rule of the form (12). If $(a, K) \neq (0, 0)$, define $\mu^{a,K}(x, r)$ for every $(x, r) \in \mathbb{R}^2 \times \mathbb{R}$ by the affine Gaussian updating formula in Proposition 1 evaluated at $X_0 = x$. Letting $x = (y^n, \rho)$, for the constant rule $a = K = 0$, define*

$$\mu^{0,0}(x, r) := \varphi_n \lambda_n y^n \qquad \text{for every } (x, r) \in \mathbb{R}^2 \times \mathbb{R},$$

i.e., the prior mean of y_1 conditional on $X_0 = x$.

[53] Of course, this does matter for the derivation of the IS curve so for the purposes of the full model, the standard continuum interpretation is the appropriate one.

For nonconstant rules, the selected affine formula agrees almost surely with the market's conditional expectation. Evaluating it at every (x, r) specifies a continuous market response also at histories where that conditional expectation is not pinned down.

The constant rule is different. Bayes' rule pins down the forecast at $r = \bar{\rho}$ but says nothing about deviations. The constant-rule response is a *convention*, albeit a natural one: after a deviation from the constant rule, the market does not revise its assessment of y_1. This is a very simple form of equilibrium refinement and is used for ruling out a constant-rate discretionary equilibrium (e.g., in Proposition 2). In the commitment case, the central bank is bound by its rule, so only the market's on-path forecast matters.

By Lemma 1 and the discussion surrounding (17), the central bank's policy problems are separable across periods. The optimality conditions below can therefore be written using the relevant current-period loss.

C.3 Discretion

Definition 4 (Discretionary equilibrium)**.** *A profile* (g_0, g_1, μ) *is a* discretionary equilibrium *if*

(i) $(g_0, \mu) \in \mathcal{E}$*;*

(ii) for P^C*-almost every realization of* A_0*,*

$$g_0(A_0) \in \arg\min_{r \in \mathbb{R}} \ \mathbb{E}^C \left[(y_0 - y_0^n)^2 \ \middle| \ \mathcal{I}_0^C \right]$$

when the market plays μ*; and*

(iii) for P^C*-almost every realization of* (A_0, Z_1)*,*

$$g_1(A_0, Z_1) \in \arg\min_{r_1 \in \mathbb{R}} \ \mathbb{E}^C \left[(y_1 - y_1^n)^2 \ \middle| \ \mathcal{I}_1^C \right].$$

It is a linear discretionary equilibrium if in addition g_0 *takes the form* (12) *and* $\mu = \mu^{a,K}$ *as specified in Definition 3.*

Condition (i) requires the market's forecast to be given by its conditional expectation of future output almost surely under the distribution induced by the rule and its prior beliefs P^M. For linear equilibria, we specify the market's response at every fundamental realization and rate by selecting the continuous affine Gaussian forecast for nonconstant rules and the no-update response for the constant rule. Conditions (ii) and (iii) make the central bank's action optimal.

Crucially, the phrase in the main text *taking the belief map as given* has a precise game-theoretic meaning. The coefficients (a, b) that determine $-\chi(b, s)/a$ describe the market's conjecture about the policy rule, even though in equilibrium, they coincide with the coefficients of the central bank's strategy. Condition (ii) holds the market's response function μ fixed while the central bank varies r. The central bank fully anticipates how its rate choice changes the market's forecast through $\mu(X_0, r)$, but—as is standard—a unilateral rate deviation does not change that response function. As will be clear below, with commitment the central bank can choose a binding rule before its assessments are realized, anticipating how that rule shapes the market's belief updating upon seeing a given policy decision.

Remark 3 (On linearity)**.** *Although we study* linear *fixed points of (i)–(ii), the central bank's optimization is unrestricted: it can choose any measurable policy rate at each point in time. For a linear policy conjecture with $a \neq 0$, Proposition 1 implies that the market's strategy μ is affine in the observed rate. Under the nonzero-total-transmission convention in Section 3.3, the central bank's $t = 0$ loss is a strictly convex quadratic in the rate, and its unique pointwise best response is affine in $\mathbb{E}^C[X_0 \mid \mathcal{I}_0^C] = A_0$. Hence, the class of linear policies is closed under the (unrestricted) best-response mapping.*

Linearity does*, however, restrict the class of market conjectures over which uniqueness claims hold: we do not rule out equilibria supported by nonlinear market conjectures.*

Finally, when $P^C = P^M$, the discretionary-equilibrium conditions are those of a perfect Bayesian equilibrium in the corresponding common-prior signaling game, with the central bank's assessment vector A_0 as its type. The analytical model does not impose that restriction; Appendix B.3 discusses the resulting scope and contrasts it with the common-prior quantitative model.

C.4 Commitment

Commitment adds a "$t = -1$ move" to the game. Before observing A_0, the central bank publicly chooses and commits to an interest rate rule from

$$\mathcal{G} := \left\{ g_0^{a,K} : (a, K) \in \mathbb{R}^2 \right\},$$

where $g_0^{a,K}$ corresponds to a generic rule of the form (12). The market observes the rule before forming its $t = 0$ forecast. Importantly, commitment technology binds the policy choice at $t = 0$ alone. After $t = 0$, the central bank again acts sequentially, and hence uses g_1^*.

Definition 5 (Commitment policy)**.** *For every $g_0^{a,K} \in \mathcal{G}$, let its market response be $\mu_{g_0} \equiv \mu^{a,K}$ as in Definition 3. A* commitment policy*, thus, is a rule $g_0 \in \mathcal{G}$ solving*

$$\min_{g_0 \in \mathcal{G}} \mathbb{E}^C \left[(y_0 - y_0^n)^2 \right], \qquad r_0 = g_0(A_0), \quad f_0 = \mu_{g_0}(X_0, r_0), \quad g_1 = g_1^*.$$

Note that by definition, the continuation equilibria at $t \geq 1$ are the same under the two solution concepts.

The distinction between the two solution concepts can now be stated more precisely. In discretionary equilibria, a unilateral central bank deviation changes the chosen rate while holding the market's strategy μ fixed, as explained above. Under commitment, the central bank instead acts ex ante by choosing a publicly observed rule. The market's complete strategy specifies a response μ_g following *every possible rule* g. Thus, holding the market's strategy fixed while the central bank deviates from g to $\tilde{g}$ changes the relevant continuation response from μ_g to $\mu_{\tilde{g}}$. It is in this precise sense that a central bank with commitment power *internalizes the effect of its rule on the market's interpretation.*

The time inconsistency arises after A_0 is observed. Absent commitment (i.e., enforcement of the pre-announced rule), the central bank would choose a rate while holding the selected map μ_{g^c} fixed. Unless g^c is itself a discretionary equilibrium, the announced rule fails condition (ii) and the central bank would deviate from what it previously announced.

Remark 4 (On the choice of intercept)**.** *Fixing the intercept in (12) at $\bar{\rho}$ is without loss under either solution concept. Since it is known to the market, an intercept c does not affect the policy surprise and hence does not affect belief revision. Under commitment, for fixed (a, K), replacing $\bar{\rho}$ by c adds $\gamma^2(c - \bar{\rho})^2$ to expected loss, so the optimal intercept is $\bar{\rho}$. Under discretion, the same conclusion follows from the fixed-point condition: under the maintained unique-solution convention, matching the intercept of the central bank's pointwise best response to the conjectured intercept gives $c = \bar{\rho}$.*

The above is a partial-commitment problem. If the $t = 0$ choice also included commitment to a $t = 1$ rule

$$r_1 = \bar{\rho} + a_1(\widehat{\rho}_1 - \bar{\rho}) - K_1 \widehat{y}_1^n,$$

the normalized sensitivity of the market's $t = 0$ forecast would become

$$\chi_{a_1,K_1}(b, s) = \frac{\gamma K_1 \varphi_n b + \gamma a_1 \varphi_\rho s}{b^2 + s},$$

which reduces to $\chi(b, s)$ at $(a_1, K_1) = (1, 1/\gamma)$. Full commitment would therefore give the central bank an additional margin: it could use the future rule's response coefficients to affect current inference. That problem is not studied here.

Finally, note that the restriction to $\mathcal{G}$ binds under commitment: the central bank's ex ante possibilities are *restricted* to linear rules. An unrestricted commitment problem would instead allow all measurable rules and require specifying a market-response selection for each such rule. As discussed in Section 3.4, the market's updating is a nonlinear function of the rule, so the unrestricted optimum is not necessarily linear, unlike in the corresponding economies without private information. The paper does not study this.

D Proofs

Throughout the proofs we write

$$h(b,s) := \gamma + \chi(b,s), \qquad \Pi_n(b,s) := 1 - \lambda_n\varphi_n + \lambda_n b\chi(b,s),$$

$$e(a,b,s) := b\left(\gamma a + \chi(b,s)\right) - \Pi_n(b,s),$$

so that h and Π_n are the pass-through of demand- and natural-output-assessment errors respectively, e is the systematic stabilization error, and consequently the loss in Lemma 2 can be written as

$$L(a,b;s) = e(a,b,s)^2\sigma^2_{n,C} + (a-1)^2\gamma^2\sigma^2_{\rho,C} + h(b,s)^2\sigma^2_{\varepsilon,C} + \Pi_n(b,s)^2\sigma^2_{\eta,C}.$$

D.1 Proof of Lemma 1

For $t \geq 2$ there is no residual uncertainty and $(y^n_t, \rho_t) = (0, \bar{\rho})$, so the central bank attains $y_t = y^n_t = 0$ by setting $r_t = \bar{\rho}$. At $t = 1$ the problem is just slightly less straightforward: since expectations are not affected by the policy decision due to the deterministic nature of next-period outcomes, minimizing $\mathbb{E}^C_1[(y_1 - y^n_1)^2]$ subject to the IS equation yields $r_1 = \widehat{\rho}_1 - \widehat{y}^n_1/\gamma$. Substituting into $y_1 = \mathbb{E}^M_1 y_2 - \gamma(r_1 - \rho_1) = -\gamma(r_1 - \rho_1)$ (as $\mathbb{E}^M_1 y_2 = 0$) yields $y_1 = \widehat{y}^n_1 + \gamma(\rho_1 - \widehat{\rho}_1)$, i.e., $\mathbb{E}^C_1 y_1 = \widehat{y}^n_1$.

D.2 Proof of Proposition 1

First, consider the case where the rate is not constant, i.e., $(a, K) \neq (0, 0)$. Since r_0 and y_1 are jointly Gaussian under the market beliefs, the conditional expectation of the latter is simply the linear projection on the former. Further, by Lemma 1, assessment persistence,

and the mean-zero restrictions in Section 2.2, $y_1 = \widehat{y}_1^n + \gamma(\rho_1 - \widehat{\rho}_1)$ has prior mean $\varphi_n \lambda_n y_0^n$ under the market's beliefs. Thus,

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n \lambda_n y_0^n + \frac{\mathrm{Cov}_0^M(y_1, r_0)}{\mathrm{Var}_0^M(r_0)} \left(r - \mathbb{E}_0^M[r_0]\right).$$

Given $r_0 = \bar{\rho} + a(\widehat{\rho}_0 - \bar{\rho}) - K\widehat{y}_0^n$, and the market beliefs, we have $\mathrm{Var}_0^M(r_0) = a^2 s + K^2$. Furthermore,

$$\mathrm{Cov}_0^M(\widehat{y}_1^n, \widehat{y}_0^n) = \varphi_n \,\mathrm{Var}_0^M(\widehat{y}_0^n) = \varphi_n.$$

Moreover, (4) and (7) imply

$$\rho_1 - \widehat{\rho}_1 = \varepsilon_\rho - \varphi_\rho(\widehat{\rho}_0 - \rho_0).$$

Because ρ_0 is in the market's information set and ε_ρ is independent of the central bank's $t = 0$ assessments,

$$\mathrm{Cov}_0^M(\rho_1 - \widehat{\rho}_1, \widehat{\rho}_0) = -\varphi_\rho \,\mathrm{Var}_0^M(\widehat{\rho}_0) = -\varphi_\rho s.$$

The remaining cross-block covariances vanish by independence. Therefore,

$$\begin{aligned}
\mathrm{Cov}_0^M(y_1, r_0) &= \mathrm{Cov}_0^M(\widehat{y}_1^n + \gamma(\rho_1 - \widehat{\rho}_1), a\widehat{\rho}_0 - K\widehat{y}_0^n) \\
&= a\,\mathrm{Cov}_0^M(\widehat{y}_1^n, \widehat{\rho}_0) - K\,\mathrm{Cov}_0^M(\widehat{y}_1^n, \widehat{y}_0^n) \\
&\quad + \gamma a\,\mathrm{Cov}_0^M(\rho_1 - \widehat{\rho}_1, \widehat{\rho}_0) - K\gamma\,\mathrm{Cov}_0^M(\rho_1 - \widehat{\rho}_1, \widehat{y}_0^n) \\
&= -K\varphi_n - \gamma a \varphi_\rho s.
\end{aligned}$$

Thus,

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n \lambda_n y_0^n - \left(\frac{K\varphi_n + \gamma a \varphi_\rho s}{a^2 s + K^2}\right)(r - \mathbb{E}_0^M[r_0]).$$

For $a \neq 0$, substitute $K = ab$ to obtain

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n \lambda_n y_0^n - \chi(b, s)\left(\frac{r - \mathbb{E}_0^M[r_0]}{a}\right).$$

For $a = 0$ and $K \neq 0$, note that the rate r fully reveals $\widehat{y}_0^n$ and nothing about $\widehat{\rho}_0$. Thus, the market simply scales the rate to recover $\widehat{y}_0^n$ and we get

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n \lambda_n y_0^n - \left(\frac{\varphi_n}{K}\right)(r - \mathbb{E}_0^M[r_0]).$$

Finally, if $(a, K) = (0, 0)$, the rate is constant and thus trivially the market does not learn anything from it, so the conditional expectation is simply the prior mean

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n \lambda_n y_0^n.$$

The expressions for output in each case follow directly from the IS equation and the above conditional expectations.

D.3 Proof of Proposition 2

We first rule out $a = 0$ in raw (a, K) coordinates. If $K = 0$, the conjectured rate is constant and thus the market makes no update whatsoever upon observing the central bank's choice. Then, the market's prior expectation of next period's output is $\varphi_n \lambda_n y_0^n$, given its belief that the central bank's contemporaneous assessment has a mean $\lambda_n y_0^n$. Hence, the central bank's expected output gap following an arbitrary rate r is

$$\varphi_n \lambda_n \widehat{y}_0^n - \gamma(r - \widehat{\rho}_0) - \widehat{y}_0^n.$$

More precisely, using the definitions of the central bank's assessments, its conditional expected loss is

$$\mathbb{E}_0^C[(y_0 - y_0^n)^2] = [-\gamma(r - \widehat{\rho}_0) - (1 - \varphi_n \lambda_n)\widehat{y}_0^n]^2 + \mathbb{E}_0^C\left[(\gamma(\rho_0 - \widehat{\rho}_0) - (1 - \varphi_n \lambda_n)(y_0^n - \widehat{y}_0^n))^2\right].$$

The second term is independent of r. Hence, the unique minimizer of the loss satisfies

$$r_0^* - \bar{\rho} = (\widehat{\rho}_0 - \bar{\rho}) - \left(\frac{1 - \varphi_n \lambda_n}{\gamma}\right) \widehat{y}_0^n,$$

which evidently does not coincide with the constant conjecture. Thus, $(a, K) = (0, 0)$ cannot be a discretionary equilibrium.

Now suppose $a = 0$ and $K \neq 0$. The conjectured rule is $r_0 = \bar{\rho} - K\widehat{y}_0^n$, so $\mathbb{E}_0^M[r_0] = \bar{\rho} - K\lambda_n y_0^n$. The projection in Proposition 1 gives the belief map following an arbitrary current rate r as

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n \lambda_n y_0^n - \frac{\varphi_n}{K}(r - \bar{\rho} + K\lambda_n y_0^n) = -\frac{\varphi_n}{K}(r - \bar{\rho}). \tag{65}$$

Given this, the central bank's expected output gap is

$$-\left(\gamma + \frac{\varphi_n}{K}\right)(r - \bar{\rho}) + \gamma(\widehat{\rho}_0 - \bar{\rho}) - \widehat{y}_0^n.$$

In particular, its conditional expected loss is

$$\mathbb{E}_0^C[(y_0 - y_0^n)^2] = \left[-\left(\gamma + \frac{\varphi_n}{K}\right)(r - \bar{\rho}) + \gamma(\widehat{\rho}_0 - \bar{\rho}) - \widehat{y}_0^n\right]^2 + \mathbb{E}_0^C\left[(\gamma(\rho_0 - \widehat{\rho}_0) - (y_0^n - \widehat{y}_0^n))^2\right].$$

The second term is independent of r. If $\gamma + \varphi_n/K \neq 0$, the rate-choice problem therefore has the unique solution

$$r_0^* - \bar{\rho} = \frac{\gamma}{\gamma + \varphi_n/K}(\widehat{\rho}_0 - \bar{\rho}) - \left(\frac{1}{\gamma + \varphi_n/K}\right)\widehat{y}_0^n.$$

Its coefficient on $\widehat{\rho}_0 - \bar{\rho}$ is nonzero since $\gamma > 0$, while the conjectured rule's coefficient is zero, so the best response cannot coincide with the conjecture. If instead $\gamma + \varphi_n/K = 0$, equivalently $K = -\varphi_n/\gamma$, expected loss is independent of r and the rate choice is not unique. Thus no discretionary equilibrium satisfying the unique-solution convention can have $a = 0$. Consequently, it can be written in scale–composition form (a, b) and we will do so from here onward.

Fix a conjectured rule (a, b) and write the total effect of a change in the policy rate (keeping the induced belief map fixed) as

$$H := \gamma + \frac{\chi(b, s)}{a}.$$

The market's prior expected policy rate is given by

$$\begin{aligned}\mathbb{E}_0^M[r_0] &= \bar{\rho} + a\left(\mathbb{E}_0^M[\widehat{\rho}_0] - \bar{\rho}\right) - ab\,\mathbb{E}_0^M[\widehat{y}_0^n] \\ &= \bar{\rho} + a(\rho_0 - \bar{\rho}) - ab\lambda_n y_0^n,\end{aligned}$$

where the second equality uses $\mathbb{E}_0^M[\widehat{\rho}_0] = \rho_0$ and $\mathbb{E}_0^M[\widehat{y}_0^n] = \lambda_n y_0^n$ from (9). Thus, the belief map induced by the conjecture entails the following expected output after observing an arbitrary current rate r:

$$\mathbb{E}_0^M[y_1 \mid r_0 = r] = \varphi_n\lambda_n y_0^n - \frac{\chi(b, s)}{a}\left(r - \bar{\rho} - a(\rho_0 - \bar{\rho}) + ab\lambda_n y_0^n\right).$$

Substituting this into the IS curve gives

$$\begin{aligned}y_0 - y_0^n &= -\gamma(r - \rho_0) - \frac{\chi(b, s)}{a}\left(r - \bar{\rho} - a(\rho_0 - \bar{\rho}) + ab\lambda_n y_0^n\right) - (1 - \varphi_n\lambda_n)y_0^n \\ &= -H(r - \bar{\rho}) + (\gamma + \chi(b, s))(\rho_0 - \bar{\rho}) - b\chi(b, s)\lambda_n y_0^n - (1 - \varphi_n\lambda_n)y_0^n.\end{aligned}$$

By definition of the central bank's assessments, its conditional expected loss can be written as

$$\begin{aligned}\mathbb{E}_0^C[(y_0 - y_0^n)^2] = \; & [-H(r - \bar{\rho}) + (\gamma + \chi(b,s))\,(\widehat{\rho}_0 - \bar{\rho}) - b\chi(b,s)\lambda_n \widehat{y}_0^n - (1 - \varphi_n\lambda_n)\widehat{y}_0^n]^2 \\ & + \mathbb{E}_0^C\left[((\gamma + \chi(b,s))\,(\rho_0 - \widehat{\rho}_0) - (1 + b\chi(b,s)\lambda_n - \varphi_n\lambda_n)(y_0^n - \widehat{y}_0^n))^2\right].\end{aligned}$$

The second term is independent of r. The unique-solution convention gives $H \neq 0$, so the central bank's best response is

$$r_0^* - \bar{\rho} = \frac{\gamma + \chi(b,s)}{H}(\widehat{\rho}_0 - \bar{\rho}) - \left(\frac{1 + b\chi(b,s)\lambda_n - \varphi_n\lambda_n}{H}\right)\widehat{y}_0^n.$$

Note that the policy-dependent part of the central bank's objective is the square of an affine function with slope $-H$. Thus, the best response above is the unique global minimizer (strict convexity holds given $H \neq 0$).

Equilibrium requires this response to coincide with $r_0 - \bar{\rho} = a(\widehat{\rho}_0 - \bar{\rho}) - ab\widehat{y}_0^n$, and hence

$$a = \frac{\gamma + \chi(b,s)}{H}, \qquad ab = \frac{1 + b\chi(b,s)\lambda_n - \varphi_n\lambda_n}{H}.$$

Substituting the definition of H into the first equality yields $a\gamma + \chi(b,s) = \gamma + \chi(b,s)$, and therefore $a_d = 1$. Rearranging the second equality then gives

$$\underbrace{b_d\,(\gamma + (1 - \lambda_n)\chi(b_d,s))}_{\equiv G(b_d;s)} = 1 - \varphi_n\lambda_n,$$

which is (18).

It remains to establish existence and uniqueness. Expanding G yields

$$G(b;s) = \gamma b + \frac{\varphi_n b^2 + \gamma\varphi_\rho bs}{b^2 + s}(1 - \lambda_n).$$

For $b < 0$, its second term satisfies

$$\frac{\varphi_n b^2 + \gamma\varphi_\rho bs}{b^2 + s}(1 - \lambda_n) \leq \frac{\varphi_n b^2}{b^2 + s}(1 - \lambda_n) < \varphi_n(1 - \lambda_n) \leq 1 - \lambda_n,$$

while $\gamma b < 0$. Hence, $G(b;s) < 1 - \lambda_n \leq 1 - \varphi_n\lambda_n$ for all $b < 0$ and thus there can't be an equilibrium with $b < 0$. Furthermore, G is clearly continuous in b, and

$$\frac{\partial G}{\partial b} = \frac{\gamma b^4 + \gamma\,(2 - (1 - \lambda_n)\varphi_\rho)\,b^2 s + \gamma\,(1 + (1 - \lambda_n)\varphi_\rho)\,s^2 + 2(1 - \lambda_n)\varphi_n bs}{(b^2 + s)^2} > 0$$

for $b > 0$. Thus, G is strictly increasing for $b > 0$ and it equals zero at $b = 0$. Finally, let $\bar{b} = (1 - \lambda_n \varphi_n)\gamma^{-1}$. Note that

$$G(\bar{b}; s) = \gamma \bar{b} + (1 - \lambda_n)\bar{b}\chi(\bar{b}, s) = (1 - \lambda_n \varphi_n) + (1 - \lambda_n)\bar{b}\chi(\bar{b}, s) > 1 - \lambda_n \varphi_n,$$

since $\chi(\bar{b}, s) > 0$, $\bar{b} > 0$ and $\lambda_n < 1$. Hence, $G(\cdot; s)$ crosses $1 - \varphi_n \lambda_n$ exactly once, at some $b_d \in (0, \bar{b})$. Thus, the discretionary equilibrium is unique and satisfies $b_d \in (0, \bar{b})$ as stated.

D.4 Proof of Lemma 2

Define the estimation errors as

$$\eta \equiv y_0^n - \widehat{y}_0^n, \qquad \varepsilon \equiv \rho_0 - \widehat{\rho}_0.$$

Using the policy rule (13) and the market-centering assumptions in (9), the market's expected rate is

$$\mathbb{E}_0^M[r_0] = \bar{\rho} + a(\rho_0 - \bar{\rho}) - ab\lambda_n y_0^n.$$

The normalized policy surprise that enters the belief map is therefore

$$\begin{aligned} \frac{r_0 - \mathbb{E}_0^M[r_0]}{a} &= \widehat{\rho}_0 - \rho_0 - b(\widehat{y}_0^n - \lambda_n y_0^n) \\ &= -\varepsilon + b\eta - b(1 - \lambda_n) y_0^n. \end{aligned}$$

Likewise, substituting $\rho_0 = \widehat{\rho}_0 + \varepsilon$ into the direct policy component gives

$$r_0 - \rho_0 = (a - 1)(\widehat{\rho}_0 - \bar{\rho}) - ab\widehat{y}_0^n - \varepsilon.$$

Hence, using (16) and $y_0^n = \widehat{y}_0^n + \eta$, we get

$$\begin{aligned} y_0 - y_0^n =& -\gamma\left[(a - 1)(\widehat{\rho}_0 - \bar{\rho}) - ab\widehat{y}_0^n - \varepsilon\right] + \varphi_n \lambda_n y_0^n - \chi(b, s)(-\varepsilon + b\eta - b(1 - \lambda_n) y_0^n) - y_0^n \\ =& -\gamma\left[(a - 1)(\widehat{\rho}_0 - \bar{\rho}) - ab\widehat{y}_0^n - \varepsilon\right] + (\varphi_n \lambda_n - 1)(\widehat{y}_0^n + \eta) \\ & - \chi(b, s)(-\varepsilon + b\eta - b(1 - \lambda_n)(\widehat{y}_0^n + \eta)) \\ =& \left(b\left(\gamma a + (1 - \lambda_n)\chi(b, s)\right) - (1 - \varphi_n \lambda_n)\right) \widehat{y}_0^n - \gamma(a - 1)(\widehat{\rho}_0 - \bar{\rho}) \\ & + \left(\gamma + \chi(b, s)\right)\varepsilon + (\varphi_n \lambda_n - 1 - b\lambda_n \chi(b, s))\eta. \end{aligned}$$

Under the central-bank ex ante restrictions, the two estimates and their estimation errors are orthogonal across and within blocks, so every cross term vanishes. Consequently,

$$\begin{aligned}\mathbb{E}^C\left[(y_0 - y_0^n)^2\right] = &\left(b\left(\gamma a + (1-\lambda_n)\chi(b,s)\right) - (1-\varphi_n\lambda_n)\right)^2 \sigma_{n,C}^2 + \gamma^2(a-1)^2\sigma_{\rho,C}^2 \\ &+ \left(\gamma + \chi(b,s)\right)^2 \sigma_{\varepsilon,C}^2 + \left(\varphi_n\lambda_n - 1 - b\lambda_n\chi(b,s)\right)^2\sigma_{\eta,C}^2,\end{aligned}$$

which, after rearrangement, becomes (21).

D.5 Proof of Proposition 3

Suppose that $a_c \neq 0$ and set $b_c = K_c/a_c$. All four terms in (21) are nonnegative. The regular discretionary rule $(1, b_d)$ is feasible and makes the first two terms zero. Hence,

$$\mathcal{P}(b_c, s) \leq L(a_c, b_c; s) \leq L(1, b_d; s) = \mathcal{P}(b_d, s).$$

Now, if $\lambda_n = 0$,

$$\mathcal{P}(b, s) = (\gamma + \chi(b,s))^2\sigma_{\varepsilon,C}^2 + \sigma_{\eta,C}^2.$$

Thus, cancelling out the constants and taking square roots in the inequality above gives

$$\gamma + \chi(b_c, s) \leq |\gamma + \chi(b_c, s)| \leq \gamma + \chi(b_d, s),$$

where the final term is positive because $b_d > 0$ and thus $\chi(b_d, s) > 0$. Subtracting γ yields $\chi(b_c, s) \leq \chi(b_d, s)$.

For strictness, consider a change in b around the discretionary rule while holding $a = 1$. Let $h_d := h(b_d, s) > 0$. Differentiating (21) with respect to b and using $e(1, b_d, s) = 0$ gives

$$\begin{aligned}L_b(1, b_d; s) &= 2e(1, b_d, s)\left(h_d + b_d\chi_b(b_d, s)\right)\sigma_{n,C}^2 + 2h_d\chi_b(b_d, s)\sigma_{\varepsilon,C}^2 \\ &= 2h_d\chi_b(b_d, s)\sigma_{\varepsilon,C}^2.\end{aligned}$$

If $\chi_b(b_d, s) \neq 0$, this derivative is nonzero. Moving b a sufficiently small distance in the direction opposite its sign, while keeping $a = 1$, therefore produces a feasible rule with loss strictly below $L(1, b_d; s)$. Since (a_c, b_c) is a global minimizer,

$$\mathcal{P}(b_c, s) = \sigma_{\varepsilon,C}^2(\gamma + \chi(b_c, s))^2 + \sigma_{\eta,C}^2 \leq L(a_c, b_c; s) < L(1, b_d; s) = \mathcal{P}(b_d, s) = \sigma_{\varepsilon,C}^2 h_d^2 + \sigma_{\eta,C}^2.$$

It follows that $|\gamma + \chi(b_c, s)| < h_d$, and hence $\chi(b_c, s) < \chi(b_d, s)$.

It remains to show that this derivative cannot vanish away from the stated knife-edge. Differentiating with respect to b gives

$$\chi_b(b,s) = \frac{\varphi_n(s-b^2) - 2\gamma\varphi_\rho sb}{(b^2+s)^2}.$$

Because $(b_d^2+s)^2 > 0$, $\chi_b(b_d,s) = 0$ is equivalent to

$$\varphi_n(s-b_d^2) = 2\gamma\varphi_\rho s b_d. \tag{66}$$

Multiplying the numerator of $\chi(b_d,s)$ by $2b_d$ and using (66) gives

$$\begin{aligned} 2b_d\left(\varphi_n b_d + \gamma\varphi_\rho s\right) &= 2\varphi_n b_d^2 + 2\gamma\varphi_\rho s b_d \\ &= 2\varphi_n b_d^2 + \varphi_n(s-b_d^2) \\ &= \varphi_n(b_d^2+s). \end{aligned}$$

Dividing by $2b_d(b_d^2+s) > 0$ therefore gives

$$\chi(b_d,s) = \frac{\varphi_n}{2b_d}.$$

The discretionary condition $b_d(\gamma + \chi(b_d,s)) = 1$ then becomes

$$\gamma b_d + \frac{\varphi_n}{2} = 1,$$

so

$$b_d = \frac{1-\varphi_n/2}{\gamma}. \tag{67}$$

Next, rearrange (66) as

$$s\left(\varphi_n - 2\gamma\varphi_\rho b_d\right) = \varphi_n b_d^2.$$

Substituting (67) yields

$$\begin{aligned} s\left[\varphi_n - 2\varphi_\rho\left(1-\frac{\varphi_n}{2}\right)\right] &= \frac{\varphi_n}{\gamma^2}\left(1-\frac{\varphi_n}{2}\right)^2, \\ \gamma^2 s\left[\varphi_n - (2-\varphi_n)\varphi_\rho\right] &= \varphi_n\left(1-\frac{\varphi_n}{2}\right)^2. \end{aligned}$$

The right-hand side is strictly positive. Thus this equality can hold only if $\varphi_n > (2-\varphi_n)\varphi_\rho$, in which case it is equivalent to

$$\gamma^2 s = \frac{(1-\varphi_n/2)^2\varphi_n}{\varphi_n - (2-\varphi_n)\varphi_\rho}.$$

We have shown that $\chi_b(b_d, s) = 0$ implies the two primitive conditions in the proposition. By contraposition, thus, away from that knife-edge, $\chi_b(b_d, s) \neq 0$. Hence, the preceding argument applies and the strict inequality $\chi(b_c, s) < \chi(b_d, s)$ holds.

Suppose now that $\lambda_n > 0$ and (22) holds. From the expression for χ_b above, its sign is that of $\varphi_n(s-b^2) - 2\gamma\varphi_\rho sb$, which is strictly decreasing in b and positive at $b = 0$. Hence, there is a unique $b^* > 0$ at which $\chi_b = 0$ and $\chi(\cdot, s)$ is hump-shaped with maximum at b^* (i.e., strictly increasing on $(0, b^*)$ and strictly decreasing thereafter). Further, the computation above showing $\chi(b_d, s) = \varphi_n/(2b_d)$ used only that the numerator of χ_b vanishes at that point. Applying it at b^*, where the numerator vanishes by definition, gives $\chi(b^*, s) = \varphi_n/(2b^*)$. Evaluating that numerator at $b = m/\gamma$ and multiplying by γ^2 gives $\gamma^2 s(\varphi_n - 2\varphi_\rho m) - \varphi_n m^2$, so (22) states exactly that it is nonnegative there, which is equivalent to $\gamma b^* \geq m$. Now, note that

$$\frac{\partial}{\partial b}\left[b\chi(b, s)\right] = \frac{s\left[2\varphi_n b + \gamma\varphi_\rho(s - b^2)\right]}{(b^2 + s)^2},$$

and, by construction, for $0 < b \leq b^*$ the numerator of χ_b is nonnegative, so $\varphi_n(s - b^2) \geq 2\gamma\varphi_\rho sb \geq 0$ and hence $s \geq b^2$, making the derivative above strictly positive. The left-hand side of the discretionary condition (18), $\gamma b + (1-\lambda_n)b\chi(b, s)$, is thus strictly increasing on $(0, b^*]$; it vanishes at $b = 0$ and by the above, at b^* equals

$$\gamma b^* + (1-\lambda_n)\frac{\varphi_n}{2} \geq m + \frac{(1-\lambda_n)\varphi_n}{2} = 1 - \lambda_n\varphi_n.$$

Thus, it attains the RHS of (18), $1 - \lambda_n\varphi_n$, on $(0, b^*]$, and since the discretionary equilibrium is unique (Proposition 2), $b_d \leq b^*$.

Now, note that if $b \leq 0$,

$$\chi(b, s) - \chi(0, s) = \frac{b(\varphi_n - \gamma\varphi_\rho b)}{b^2 + s} \leq 0.$$

Since $0 < b_d \leq b^*$, the strict monotonicity of χ on $[0, b^*]$ implies $\chi(0, s) < \chi(b_d, s)$. Thus, if $b_c \leq 0$, $\chi(b_c, s) < \chi(b_d, s)$ holds, so we can focus on the case $b_c > 0$ for the rest of the proof.

Suppose now $\chi(b_c, s) > \chi(b_d, s)$. If $b_c \leq b_d$ then $0 < b_c \leq b_d \leq b^*$, where $\chi(\cdot, s)$ is increasing: a contradiction. Hence, $b_c > b_d$, and both factors of $b\chi(b, s)$ are then larger at b_c, so $b_c\chi(b_c, s) > b_d\chi(b_d, s)$. Since the two terms in $\mathcal{P}$ are positive and increasing in

$\chi(b,s)$ and in $b\chi(b,s)$ respectively, it follows that $\mathcal{P}(b_c,s) > \mathcal{P}(b_d,s)$: another contradiction. Consequently, $\chi(b_c,s) \leq \chi(b_d,s)$.

For strictness in the $\lambda_n > 0$ case, note that at $a = 1$ the first term in the loss function (21) is $e(1,b,s)^2\sigma^2_{n,C}$, which equals zero at b_d. Thus, differentiating (21) in b at $(1,b_d)$ leaves

$$L_b(1,b_d,s) = 2\left(\gamma + \chi(b_d,s)\right)\chi_b(b_d,s)\sigma^2_{\varepsilon,C} + 2\lambda_n\left(1 - \lambda_n\varphi_n + \lambda_n b_d\chi(b_d,s)\right)\left.\frac{\partial[b\chi(b,s)]}{\partial b}\right|_{b_d}\sigma^2_{\eta,C} > 0,$$

because by the arguments above, $b_d \leq b^*$ makes $\chi_b(b_d,s) \geq 0$ and the last derivative strictly positive. Thus, lowering b marginally while keeping $a = 1$ gives a rule with loss strictly below $L(1,b_d;s)$, so $\mathcal{P}(b_c,s) \leq L(a_c,b_c;s) < \mathcal{P}(b_d,s)$ and in particular $b_c \neq b_d$. Now suppose $\chi(b_c,s) = \chi(b_d,s)$. By strict monotonicity on $(0,b^*]$, $b_c < b_d \leq b^*$ can't hold, so $b_c > b_d$. In that case, $b_c\chi(b_c,s) > b_d\chi(b_d,s)$ and $\mathcal{P}(b_c,s) > \mathcal{P}(b_d,s)$: a contradiction. Hence, $\chi(b_c,s) < \chi(b_d,s)$.

D.6 Proof of Corollary 1

The result follows directly from the coefficient characterization in Proposition 4 in Appendix E.

E Additional Theoretical Results

Lemma 3 (Existence of a commitment optimum)**.** *Suppose*

$$\mathcal{P}(b_d,s) < \gamma^2\sigma^2_{\rho,C}.$$

Then, there exists a solution (a_c,K_c) to the commitment problem. Moreover, every solution satisfies $a_c \neq 0$.

Proof. For every $(a,K) \neq (0,0)$, define

$$\zeta(a,K;s) := \frac{\varphi_n K + \gamma\varphi_\rho sa}{a^2 s + K^2},$$

i.e., minus the slope of the market's output expectations with respect to the policy rate

surprise (in the (a, b) representation of policy, $\zeta = \chi/a$). Direct substitution gives the commitment loss in raw coordinates:

$$\begin{aligned}\widetilde{L}(a, K; s) = {} & \left[\gamma K + (1 - \lambda_n)K\zeta(a, K; s) - (1 - \lambda_n\varphi_n)\right]^2 \sigma_{n,C}^2 + \gamma^2(a-1)^2\sigma_{\rho,C}^2 \\ & + \left[\gamma + a\zeta(a, K; s)\right]^2 \sigma_{\varepsilon,C}^2 + \left[1 - \lambda_n\varphi_n + \lambda_n K\zeta(a, K; s)\right]^2 \sigma_{\eta,C}^2,\end{aligned}$$

which reduces to $L(a, b; s)$ in Lemma 2 when $a \neq 0$ and we set $b = K/a$. Instead, when $a = 0$ and $K \neq 0$, it gives

$$\widetilde{L}(0, K; s) = \left[\gamma K - (1 - \varphi_n)\right]^2 \sigma_{n,C}^2 + \gamma^2\sigma_{\rho,C}^2 + \gamma^2\sigma_{\varepsilon,C}^2 + \sigma_{\eta,C}^2.$$

The preceding formula is defined only for $(a, K) \neq (0, 0)$. It cannot generally be extended continuously to the origin: although $a\zeta(a, K; s)$ and $K\zeta(a, K; s)$ remain bounded as $(a, K) \to (0, 0)$, their limits depend on the direction of approach.

At the exact constant rule $(a, K) = (0, 0)$, the rate conveys no information, and hence the market retains its prior output forecast. Consequently, the associated loss is

$$L_{\text{const}} = (1 - \lambda_n\varphi_n)^2(\sigma_{n,C}^2 + \sigma_{\eta,C}^2) + \gamma^2(\sigma_{\rho,C}^2 + \sigma_{\varepsilon,C}^2).$$

Thus, the central bank's objective function with commitment is

$$\mathscr{L}(a, K; s) = \begin{cases} \widetilde{L}(a, K; s), & (a, K) \neq (0, 0), \\ L_{\text{const}}, & (a, K) = (0, 0). \end{cases} \tag{68}$$

The function $\mathscr{L}$ is continuous on $\mathbb{R}^2 \setminus \{(0, 0)\}$, but not necessarily at the origin. We prove the existence of a solution to the central bank's commitment problem by showing that the sublevel set generated by the feasible discretionary rule is compact and bounded away from this discontinuity point.

Let

$$S := \left\{(a, K) \in \mathbb{R}^2 : \mathscr{L}(a, K; s) \leq \mathcal{P}(b_d, s)\right\}$$

be the aforementioned sublevel set. The discretionary rule $(a, K) = (1, b_d)$ is feasible and satisfies

$$\mathscr{L}(1, b_d; s) = \mathcal{P}(b_d, s),$$

so S is nonempty.

Now we show that S is bounded. For every $(a, K) \neq (0, 0)$,

$$K\zeta(a, K; s) = \frac{\varphi_n K^2 + \gamma\varphi_\rho s\, aK}{a^2 s + K^2}.$$

Using $2|aK|\sqrt{s} \leq a^2 s + K^2$, it follows that

$$|K\zeta(a, K; s)| \leq \varphi_n + \frac{\gamma\varphi_\rho\sqrt{s}}{2}.$$

Consequently, for some finite constant C (independent of (a, K)),

$$|\gamma K + (1 - \lambda_n)K\zeta(a, K; s) - (1 - \lambda_n\varphi_n)| \geq \gamma|K| - C.$$

Thus, the first term of $\mathscr{L}(a, K; s)$ diverges as $|K| \to \infty$. Moreover,

$$\gamma^2(a - 1)^2\sigma^2_{\rho,C} \to \infty \qquad \text{as } |a| \to \infty.$$

Then,

$$\mathscr{L}(a, K; s) \to \infty \qquad \text{as } \|(a, K)\| \to \infty,$$

so S is bounded.

We next show that S is bounded away from $a = 0$. By assumption,

$$\mathcal{P}(b_d, s) < \gamma^2\sigma^2_{\rho,C}.$$

Thus, we can choose $\delta \in (0, 1)$ such that

$$\gamma^2(1 - \delta)^2\sigma^2_{\rho,C} > \mathcal{P}(b_d, s).$$

Then, for every *nonconstant* rule with $|a| < \delta$,

$$\mathcal{P}(b_d, s) < \gamma^2(a - 1)^2\sigma^2_{\rho,C} \leq \mathscr{L}(a, K; s).$$

Further, at the constant rule,

$$L_{\text{const}} \geq \gamma^2\sigma^2_{\rho,C} > \mathcal{P}(b_d, s).$$

Then,

$$S \subseteq \{(a, K) : |a| \geq \delta\}.$$

On the closed set $\{(a, K) : |a| \geq \delta\}$ and hence on S, the loss $\mathscr{L}$ reduces to $\widetilde{L}$. Thus, since $\widetilde{L}$ and $\mathcal{P}(b_d, s)$ are continuous, S is closed. Since we established above that it is also bounded and it is a subset of $\mathbb{R}^2$, it is compact.

By the Weierstrass theorem, $\mathscr{L}$ attains its minimum on S. By construction, any policy outside S has loss strictly greater than $\mathcal{P}(b_d, s)$, while S contains the *feasible* discretionary rule with loss $\mathcal{P}(b_d, s)$. Therefore, a minimizer of $\mathscr{L}$ on S is a global solution of the commitment problem.

Finally, every rule with $a = 0$, including the constant rule, has loss at least

$$\gamma^2 \sigma^2_{\rho,C} > \mathcal{P}(b_d, s),$$

and the global minimum is at most $\mathcal{P}(b_d, s)$. Hence, every solution satisfies $a_c \neq 0$. □

Proposition 4 (Local implementation)**.** *Suppose* (22) *holds strictly. For all sufficiently small* $\sigma^2_{\varepsilon,C}/\sigma^2_{n,C}$ *and* $\sigma^2_{\eta,C}/\sigma^2_{n,C}$*, any optimal commitment rule satisfies*

$$a_c > 1, \qquad 0 < b_c < b_d.$$

Proof. Note that L is homogeneous of degree 1 in the variance vector, so only the variance ratios matter. Thus, define the normalized error variances

$$q_\varepsilon := \frac{\sigma^2_{\varepsilon,C}}{\sigma^2_{n,C}}, \qquad q_\eta := \frac{\sigma^2_{\eta,C}}{\sigma^2_{n,C}}, \qquad q := (q_\varepsilon, q_\eta).$$

We consider sufficiently small positive pairs (q_ε, q_η), holding all other parameters fixed. Further, let $\Delta_\rho := \gamma^2 \sigma^2_{\rho,C}/\sigma^2_{n,C}$ so that

$$L(a, b; s) = \left(e(a, b, s)^2 + \Delta_\rho (a - 1)^2 + q_\varepsilon h(b, s)^2 + q_\eta \Pi_n(b, s)^2 \right) \sigma^2_{n,C}.$$

Also let

$$G(b; s) := b \left(\gamma + (1 - \lambda_n) \chi(b, s) \right).$$

The discretionary coefficient is the unique positive solution of $G(b_d; s) = 1 - \lambda_n \varphi_n$.

We first note that a global commitment optimum exists for all sufficiently small positive pairs q. Under the variance parameterization above, the sufficient condition in Lemma 3,

$$\mathcal{P}(b_d, s) < \gamma^2 \sigma^2_{\rho,C},$$

is equivalent, after dividing by $\sigma_{n,C}^2$, to

$$q_\varepsilon h(b_d, s)^2 + q_\eta \Pi_n(b_d, s)^2 < \Delta_\rho. \tag{69}$$

Since $h(b_d, s)$, $\Pi_n(b_d, s)$, and $\Delta_\rho > 0$ are fixed, this condition holds for all sufficiently small positive pairs (q_ε, q_η). Lemma 3 then implies that the commitment problem has a global minimizer and that every global minimizer has $a \neq 0$.

Now let's start at $q = (0, 0)$. Since $\Delta_\rho > 0$, the second term in the objective is zero only if $a = 1$. Conditional on $a = 1$, the first term is zero only if $G(b; s) = 1 - \lambda_n \varphi_n$, and hence only if $b = b_d$. Further, note that in raw coordinates (a, K), every policy with $a = 0$, including the constant rule, has (normalized) loss at least Δ_ρ when $q = (0, 0)$, because the central bank does not respond to its demand assessment. Thus, $(1, b_d)$ is the unique global optimum when $q = (0, 0)$.

The first-order conditions of the commitment problem can be written as

$$0 = \gamma b e(a, b, s) + \Delta_\rho (a - 1), \tag{70}$$

$$\begin{aligned} 0 = {} & (\gamma a + (1 - \lambda_n)(\chi(b, s) + b\chi_b(b, s)))\, e(a, b, s) \\ & + q_\varepsilon h(b, s)\chi_b(b, s) + q_\eta \lambda_n \Pi_n(b, s)(\chi(b, s) + b\chi_b(b, s)). \end{aligned} \tag{71}$$

We next show that $(1, b_d)$ entails uniformly lower losses than all policies outside any neighborhood of it at $q = (0, 0)$. Let

$$L_0(a, b; s) = e(a, b, s)^2 + \Delta_\rho (a - 1)^2.$$

We claim that, for every neighborhood U of $(1, b_d)$, there exists $c_U > 0$ such that

$$L_0(a, b; s) \geq c_U \qquad \text{for all } (a, b) \notin U.$$

To see this, suppose instead that there were a sequence outside U with $L_0(a_j, b_j; s) \to 0$. Policies with $a_j = 0$ are excluded because their loss is at least Δ_ρ. Otherwise, nonnegativity of the two terms in L_0 implies

$$a_j \to 1, \qquad e(a_j, b_j, s) \to 0.$$

Moreover, b_j must be bounded. If $|b_j| \to \infty$, then $\chi(b_j, s) = O(1/b_j)$, and hence

$$e(a_j, b_j, s) = \gamma a_j b_j + O(1),$$

which cannot converge to zero. Now take any convergent subsequence $b_{j_k} \to \bar{b}$. Since $a_{j_k} \to 1$ and $e(a_{j_k}, b_{j_k}, s) \to 0$, continuity of e implies

$$0 = e(1, \bar{b}, s),$$

i.e., $G(\bar{b}; s) = 1 - \lambda_n \varphi_n$. Since G attains $1 - \lambda_n \varphi_n$ only at b_d, we must have $\bar{b} = b_d$. Hence, $b_j \to b_d$. Thus, $(a_j, b_j) \to (1, b_d)$, contradicting the requirement that the sequence remain outside U.

Since (22) holds strictly, the proof of Proposition 3 gives $\gamma b^* > m$, where b^* is such that $\chi_b(b^*, s) = 0$, and also $b^* \chi(b^*, s) = \varphi_n / 2$. Hence

$$G(b^*; s) = \gamma b^* + (1 - \lambda_n)\frac{\varphi_n}{2} > m + (1 - \lambda_n)\frac{\varphi_n}{2} = 1 - \lambda_n \varphi_n = G(b_d; s),$$

so $b_d < b^*$ by strict monotonicity of $G(\cdot; s)$. Hence $\chi_b(b_d, s) > 0$ and, as also shown there, $\partial \left[b\chi(b, s) \right] / \partial b > 0$ on $(0, b^*]$ and in particular at b_d.

The above plus a standard continuity argument implies that we can choose a sufficiently small neighborhood U of $(1, b_d)$, with closure in the positive quadrant, such that

$$\chi_b(b, s) > 0, \qquad \chi(b, s) + b\chi_b(b, s) > 0$$

for all $(a, b) \in U$. Applying the argument above to this neighborhood gives $c_U > 0$ such that outside U,

$$L_q(a, b; s) \geq L_0(a, b; s) \geq c_U,$$

where $L_q = L / \sigma^2_{n,C}$ denotes the normalized loss and the leftmost inequality follows from the fact that the error terms are nonnegative. On the other hand, the feasible policy $(1, b_d)$ has loss

$$L_q(1, b_d; s) = q_\varepsilon h(b_d, s)^2 + q_\eta \Pi_n(b_d, s)^2.$$

Hence, if

$$q_\varepsilon h(b_d, s)^2 + q_\eta \Pi_n(b_d, s)^2 < c_U, \tag{72}$$

every global minimizer must lie inside U: the feasible discretionary rule has loss below c_U, while every rule outside U with $a \neq 0$ has loss at least c_U. Note that both (72) and the earlier existence condition (69) hold for all sufficiently small pairs of error-variance ratios.

Consequently, every optimum has $a_c > 0$ and $b_c > 0$. The loss is differentiable there, so each optimum satisfies the first-order conditions (70)–(71). In (71), evaluated at an optimum, the coefficient multiplying $e(a_c, b_c, s)$ is positive, and the remaining term is strictly

positive since $q_\varepsilon h(b_c, s)\chi_b(b_c, s) > 0$, while the natural-output-error term is nonnegative. Thus, $e(a_c, b_c, s) < 0$. Then, (70) implies

$$a_c - 1 = -\frac{\gamma b_c}{\Delta_\rho} e(a_c, b_c, s) > 0.$$

Finally, using the definitions of e and G,

$$G(b_c; s) - (1 - \lambda_n \varphi_n) = e(a_c, b_c, s) - \gamma b_c (a_c - 1) < 0.$$

Since G is strictly increasing for $b > 0$ and $G(b_d; s) = 1 - \lambda_n \varphi_n > 0$, it follows that $b_c < b_d$. □

F Additional Empirical Results

F.1 Alternative Policy-Surprise Measure

Figure 6 below shows the results of the specification (24) using the fourth quarterly Eurodollar/SOFR futures contract, ED4, on the sample that has both OIS1Y and ED4 data. While OIS1Y approximately measures the expected average overnight rate over the following year, ED4 measures a three-month rate roughly one year ahead. Thus, the latter places weight near the end of the one-year horizon instead of averaging expected rates over the entire path. The coefficient levels differ, but the key findings are unchanged: for every horizon and either surprise measure, the 2010–2019 estimate is more negative than the corresponding pre-GFC and post-2021 estimates.

F.2 TIPS-Based Inflation Compensation

The baseline exercise proxies the market's inflation expectations using inflation swaps. To verify that our results are not driven by specific features of that market or by the construction of the data series, we run the specification in (24) using Treasury Inflation-Protected Securities (TIPS). Concretely, we use the Federal Reserve's fitted inflation-compensation series, built from nominal Treasury and TIPS yield curves as detailed in Gürkaynak et al. (2007) and Gürkaynak et al. (2010). The fitted series account for differences in coupon and maturity structures between nominal Treasuries and TIPS, since simply subtracting their yields would not provide an apples-to-apples comparison.

Let $b_{t,h}^{\text{TIPS}}$ denote the continuously compounded zero-coupon inflation-compensation rate

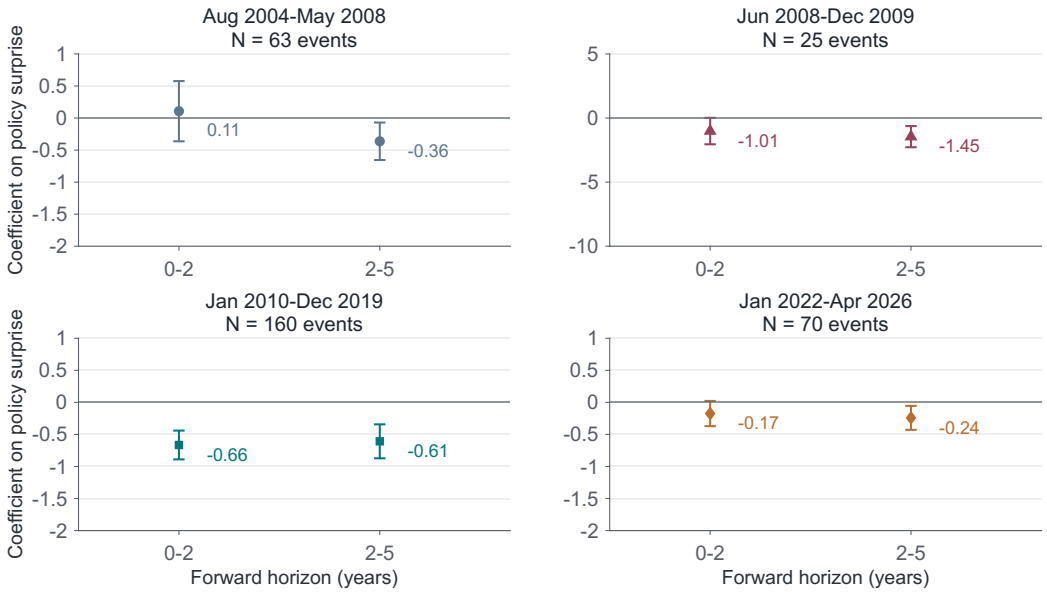

Figure 6: Estimates of (24) using the USMPD ED4 surprise. The sample and plotting conventions match Figure 3.

at maturity h. Using the two- and five-year rates, we build forward inflation-compensation rates over the same horizons used in the baseline analysis:

$$b^{\text{TIPS}}_{t,0\text{–}2} = b^{\text{TIPS}}_{t,2}, \qquad b^{\text{TIPS}}_{t,2\text{–}5} = \frac{5b^{\text{TIPS}}_{t,5} - 2b^{\text{TIPS}}_{t,2}}{3}.$$

Just as in the baseline analysis, for each policy event, the revision is the change from the last trading day before the event to the first trading day after it. Because the availability of the TIPS series differs from that of the baseline outcomes, we use the events for which the TIPS-based outcome and the policy surprise are both available. This leaves 301 events: 57 in the pre-GFC period, 25 during the GFC, 154 during 2010–2019, and 65 during 2022–2026.

Figure 7 reports the results. The central regime comparison is preserved. During 2010–2019, the point estimates are more negative at every horizon than in the pre-GFC and post-2021 subsamples. The GFC point estimates are more negative than those for 2010–2019 at both horizons, with wider confidence intervals.[54]

[54]Note that just as with the baseline measure, TIPS-based inflation compensation can also reflect inflation risk premia, on top of a liquidity premium, since regular nominal Treasuries are more liquid than TIPS. However, the fact that we recover the main results using a different set of securities makes an explanation based on inflation-swap measurement issues considerably less plausible.

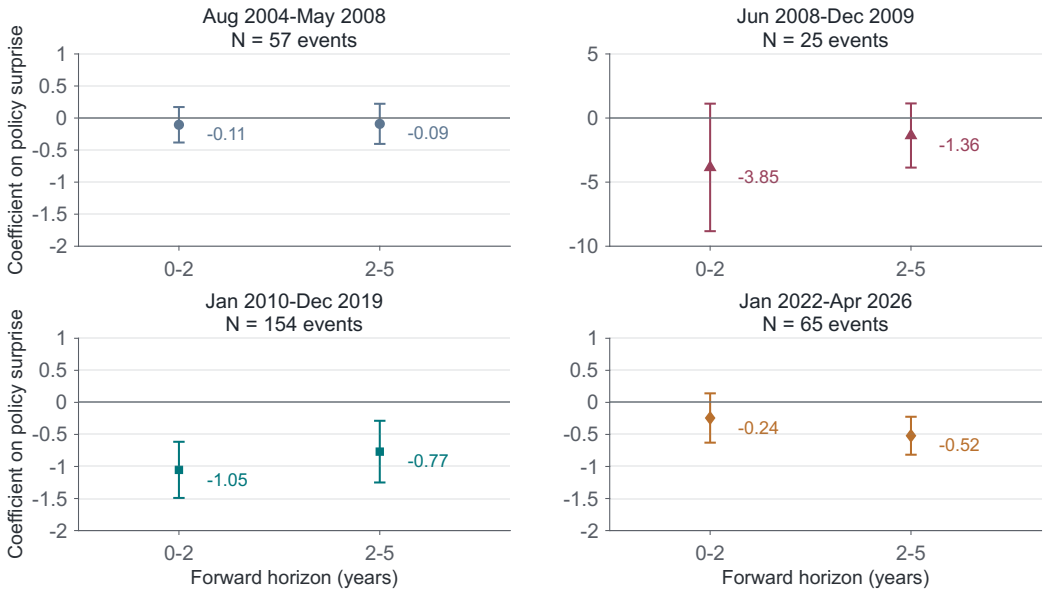

Figure 7: Estimates of (24) using OIS1Y and TIPS-based breakeven inflation. The sample includes the 301 events described in the text; plotting conventions match Figure 3.

F.3 Bid and Ask Prices

Figures 8 and 9 present results for inflation swaps using the bid and ask prices, respectively. Both use the sample and plotting conventions of Figure 3. The main regime pattern is unchanged.

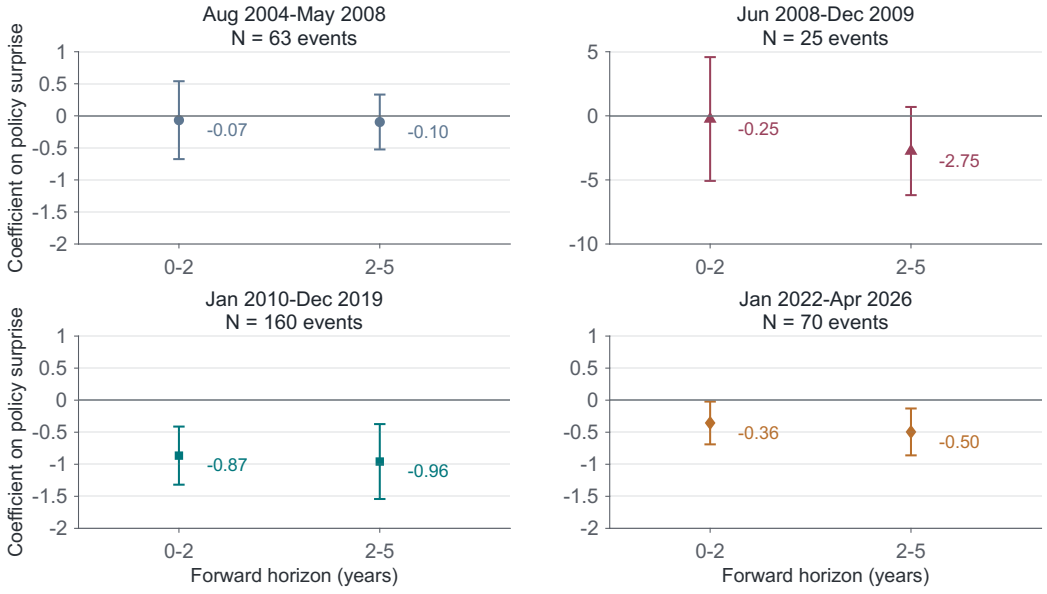

Figure 8: Estimates of (24) using OIS1Y and bid-side inflation-swap quotes.

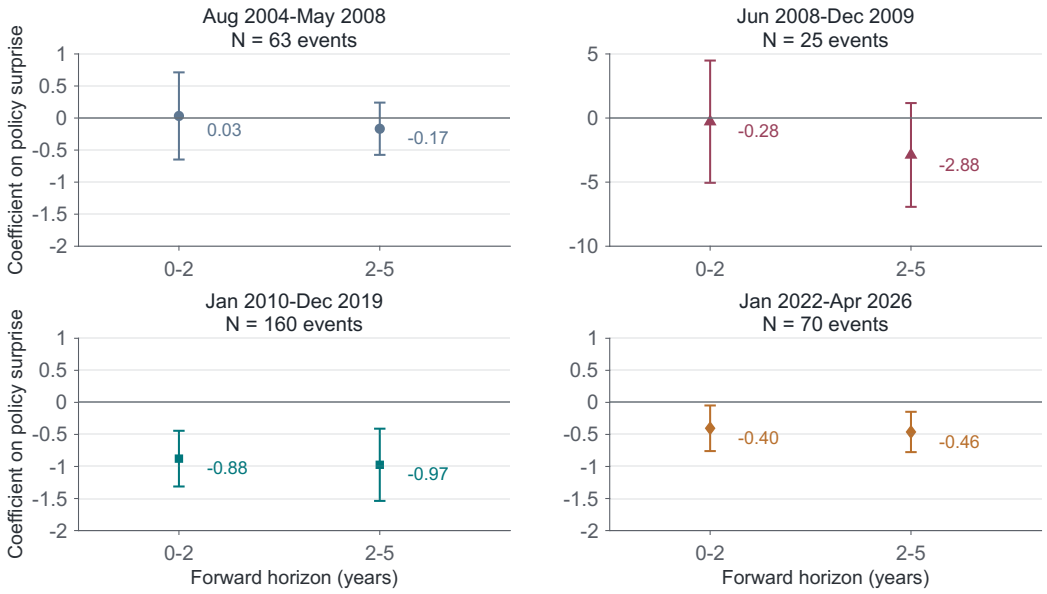

Figure 9: Estimates of (24) using OIS1Y and ask-side inflation-swap quotes.

G Quantitative Model Details

Throughout this appendix we work with the interest rate centered around its long-run mean, $\tilde{i}_t \equiv i_t - \bar{\rho}$.

G.1 Information and Forecasting

Let $W_t = (y_t, \pi_t)'$ denote the vector of endogenous variables the market forecasts one period ahead. Under the within-period timing described in Section 5, the market's information immediately before and after the current policy decision is, respectively

$$\mathcal{I}_t^{M,-} \equiv \sigma\left(\{X_\tau\}_{\tau \le t}, \{\tilde{i}_\tau\}_{\tau \le t-1}\right), \tag{73}$$

$$\mathcal{I}_t^{M} \equiv \sigma\left(\{X_\tau, \tilde{i}_\tau\}_{\tau \le t}\right). \tag{74}$$

In linear equilibria, the market forecast of next period's output and inflation is

$$\mathbb{E}[W_{t+1} \mid \mathcal{I}_t^M] = \sum_{\ell=0}^{\infty} F_\ell X_{t-\ell} + \sum_{\ell=0}^{\infty} G_\ell \tilde{i}_{t-\ell}, \tag{75}$$

where F_ℓ is 2×3 and G_ℓ is 2×1.

Note that unlike the market in Aoki (2003), the market in our model uses the entire

history of fundamentals and policy rates to predict future endogenous outcomes.[55] Consequently, in a discretionary equilibrium, policy will depend on the central bank's current expectations of the entire history of fundamentals. The reason behind this difference is that Aoki (2003) assumes that the central bank perfectly observes past realizations of fundamentals. Hence, there is no persistent information asymmetry the market has to keep track of: the only uncertainty the market faces regarding policy comes from the contemporaneous realization of the central bank's signal, whose best predictor from the market's perspective depends only on current fundamentals, since the signal error is independent across time. This amounts to a significant simplification. Note further that for our purposes that assumption would not work, since in that case the market does not extract any information from current rates that is useful to predict future rates (conditional on it already knowing contemporaneous fundamentals) and hence the mechanism we study is absent.[56]

Relation to Svensson and Woodford (2004). Our "rate-only" representation differs from the indicator-variable approach in Svensson and Woodford (2004). They also study optimal monetary policy with imperfect central-bank information in a linear forward-looking economy, but their state-space representation allows current endogenous variables to act as contemporaneous indicators of the central bank's estimate. In their notation, forward-looking variables can be written as

$$z_t = G_1 X_t + (G - G_1)\widehat{X}^C_{t|t}, \qquad z_t \equiv (y_t, \pi_t)'. \tag{76}$$

If z_t is in the private sector's date-t information set, then these variables reveal linear combinations of $\widehat{X}^C_{t|t}$ beyond what is revealed by the policy rate alone. The infinite distributed-lag representation is then bypassed: the equilibrium can be solved with finite state-space objects, Kalman gains, and covariance matrices.

Note that this is a rational-expectations fixed point, but it is not the "rate-only" information structure studied here. Concretely, in our model the only way information about the central bank's signal errors gets to the private sector is via observed policy decisions (processed using the knowledge of fundamentals that the market directly observes by assumption). Naturally, this is crucial for our mechanism. In this sense, our "rate-only"

[55] Aoki (2003)'s physical environment is the same as ours, except it does not have cost-push shocks so the divine coincidence holds. Furthermore, in their model the central bank's signals are noisy versions of endogenous objects (inflation and output) instead of fundamental shocks.

[56] Note that this would be the case even if the signal's errors weren't independent across time. At $t+1$, observing X_t allows the central bank to recover $e_t = s_t - X_t$ and subtract the component of e_{t+1} predictable from past errors from its new signal. Hence, its $t+1$ assessment depends only on the new signal-error innovation, about which i_t contains no information. Consequently, the previous policy rate does not help the market predict future policy (conditional on public fundamentals).

information structure enforces a causal information constraint that is absent from the formulation in Svensson and Woodford (2004) (this is related to the much broader critique of equilibrium models in Dogra (2024)).

Finite-history approximation. As described in the main text above, we use a finite approximation of the state in the market's forecasting operator (75). Concretely, we replace the market's projection on the full history of fundamentals and past rates with a projection on current fundamentals and policy rates and their first H lags. The central bank's state is also truncated accordingly.

Delayed publication of central-bank estimates provides a natural reason why sufficiently old policy rates may eventually contain little additional information (thus making the finite H approximation reasonable): once the underlying assessment becomes public, earlier rates and fundamentals do not contain additional information. In the U.S., Greenbook/Tealbook forecasts are released to the public with a five-year lag[57]. See Townsend (1983) for a related interpretation of finite-dimensional approximations to infinite-dimensional state spaces in rational-expectations models.

Since all relevant random variables are Gaussian, the finite-history market forecasts of future allocations are given by a linear projection of future output and inflation on the current values and first H lags of the relevant variables, i.e.,

$$M_t \equiv \mathrm{Proj}(W_{t+1} \mid Z_t^H) = \Theta Z_t^H, \tag{77}$$

where $Z_t^H \equiv (X_t', X_{t-1}', \ldots, X_{t-H}', \tilde{i}_t, \tilde{i}_{t-1}, \ldots, \tilde{i}_{t-H})'$ and Θ contains the coefficients of the finite-history population projection

$$\Theta \in \arg\min_{\widetilde{\Theta}} \mathbb{E}\left[\left\|W_{t+1} - \widetilde{\Theta} Z_t^H\right\|^2\right]. \tag{78}$$

Note that since older omitted histories are correlated with the retained regressors, Θ does not in general equal a direct truncation of the infinite-history coefficient sequences $\{F_\ell, G_\ell\}_{\ell \geq 0}$.

From here onward, M_t and Θ denote the finite-history projection and its associated coefficient matrix (since H is fixed, we drop the superscript H for notational convenience).

In the case of the central bank, its finite-history state retains its current posterior estimates of the $X_t, \ldots, X_{t-H}$, together with H lagged policy rates. Formally, denoting the finite history of fundamentals by

$$\mathcal{X}_t^H \equiv (X_t', X_{t-1}', \ldots, X_{t-H}')', \tag{79}$$

[57] `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/greenbook`

the central bank state is given by

$$\mathcal{S}_t^{C,H} \equiv \begin{pmatrix} \widehat{\mathcal{X}}_{t|t}^{C,H} \\ \tilde{i}_{t-1} \\ \vdots \\ \tilde{i}_{t-H} \end{pmatrix}, \tag{80}$$

where

$$\widehat{\mathcal{X}}_{\tau|t}^{C,H} \equiv \mathbb{E}[\mathcal{X}_\tau^H \mid \mathcal{I}_t^C], \qquad \mathcal{I}_t^C \equiv \sigma\left(\{s_\tau^C, \tilde{i}_\tau\}_{\tau \leq t}\right). \tag{81}$$

G.2 Definition of Equilibria

Discretionary equilibria. As in the simpler model, a discretionary central bank takes the market's forecasting rule Θ as given. Note that for a given market forecast $M_t = (m_{y,t}, m_{\pi,t})'$, current allocations are pinned down by the IS curve and NKPC

$$\begin{aligned} x_t &= -y_t^n + \gamma(\rho_t - \bar{\rho}) - \gamma\tilde{i}_t + m_{y,t} + \gamma m_{\pi,t}, \\ \pi_t &= u_t + \kappa x_t + \beta m_{\pi,t}. \end{aligned} \tag{82}$$

Since the market forecast is a linear function of the private sector state, so are allocations. Thus, the discretionary central bank faces a linear-quadratic problem. Defining

$$R_t(i) \equiv \begin{pmatrix} x_t(i) \\ \pi_t(i) \\ i - \bar{\rho} \end{pmatrix}, \qquad \mathcal{Q} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & \vartheta_\pi & 0 \\ 0 & 0 & \vartheta_i \end{pmatrix},$$

the discretionary problem can then be written recursively as

$$V(\mathcal{S}_t^{C,H}) = \min_i \mathbb{E}_t^C \left[\frac{1}{2} R_t(i)' \mathcal{Q} R_t(i) + \beta V(\mathcal{S}_{t+1}^{C,H}(i))\right]. \tag{83}$$

The continuation value matters because the current policy rate becomes part of future public histories. Concretely, since the market uses lagged rates to forecast future outcomes, today's policy shapes how market expectations are formed in the future. Consequently, optimal policy has an intertemporal element even under discretion, unlike in the model's full-information counterpart. Since the discretionary problem is linear quadratic, the optimal policy rule is of the form

$$\tilde{i}_t = K^D(\Theta)\mathcal{S}_t^{C,H}, \tag{84}$$

where, as the notation highlights, policy depends on the market forecasting structure Θ. This is because the central bank understands that the transmission of policy to macro outcomes depends on how the market interprets policy rates. Concretely, it is aware that expectations M_t are partly shaped by the contemporaneous policy instrument i_t.

Definition 6. *A stationary linear finite-H-history discretionary equilibrium is a set of stochastic processes $\{y_t, \pi_t, i_t, M_t\}_{t=0}^{\infty}$ and coefficient matrices (K^D, Θ) such that*

- **(Forecast Consistency)** *Given the policy rule (84), $M_t = \Theta Z_t^H$, where Θ solves the optimal forecasting problem (78) under the stationary distribution induced by the equilibrium.*
- **(Policy Optimality)** *Given the market's forecasting matrix Θ, $\tilde{i}_t = K^D \mathcal{S}_t^{C,H}$ solves (83).*
- **(Private-Sector Equilibrium)** *Given $M_t = (m_{y,t}, m_{\pi,t})'$, allocations satisfy (82), together with $y_t = x_t + y_t^n$.*

A game-theoretic definition analogous to the one in Appendix C for the baseline model applies here (and in the commitment case below).

Optimal linear commitment policy. A central bank with commitment chooses a policy rule ex ante and thus acknowledges that the market's forecasting depends on its chosen rule. We restrict the space of policies to linear rules using the same variables as the discretionary rule.[58] Then, a commitment rule is a vector K^C such that

$$\tilde{i}_t = K^C \mathcal{S}_t^{C,H}, \tag{85}$$

where each K^C implies a stationary distribution of the state with corresponding market forecasting rule $\Theta(K^C)$. Let

$$\ell_t(K^C) \equiv \frac{1}{2}\left(x_t^2 + \vartheta_\pi \pi_t^2 + \vartheta_i \tilde{i}_t^2\right)$$

denote per-period loss under that stationary distribution. We choose the commitment rule to minimize unconditional expected per-period loss under its induced stationary distribution.[59]

[58] Thus, in general this does not correspond to the unrestricted Ramsey allocation. This is done for simplicity since computing the Ramsey allocation would be much harder than in the standard full-information benchmark due to the non-linearity of the market's forecasting as a function of policy. Thus, this might be seen as providing a lower bound on the commitment gains from the unrestricted Ramsey problem (keeping the same stationary loss criterion).

[59] This is equivalent to the normalized discounted loss $(1-\beta)\,\mathbb{E}\sum_{j=0}^{\infty}\beta^j \ell_{t+j}$ if the economy starts from that stationary distribution.

This corresponds to the "fully timeless" criterion discussed by McCallum (2005), though restricted to our class of linear finite-history rules (see also Damjanovic et al. (2008) for an extended discussion). Thus, the commitment policy is given by

$$K^{C,*} \in \arg\min_{K^C} \mathbb{E}[\ell_t(K^C)], \tag{86}$$

where the latter expectation is the unconditional one corresponding to the induced stationary distribution.

Definition 7. *A stationary linear finite-H-history commitment equilibrium consists of stochastic processes $\{y_t, \pi_t, i_t, M_t\}_{t=0}^{\infty}$ and coefficient matrices (K^C, Θ) such that:*

- **(Forecast Consistency)** *Given the policy rule, $M_t = \Theta Z_t^H$, where Θ solves the market's optimal forecasting problem (78) under the stationary distribution induced by the equilibrium.*
- **(Policy Optimality)** *The policy rate satisfies $\tilde{i}_t = K^C \mathcal{S}_t^{C,H}$, where K^C solves (86).*
- **(Private-Sector Equilibrium)** *Given $M_t = (m_{y,t}, m_{\pi,t})'$, allocations satisfy (82), together with $y_t = x_t + y_t^n$.*

This formulation makes comparison between commitment and discretion clear. In the discretionary problem, the central bank takes the market's expectation formation process Θ as given when choosing i_t. Under commitment, the central bank chooses the systematic rule K^C internalizing that this shapes the informativeness of policy rates and thus the equilibrium forecast matrix $\Theta(K^C)$.

The current-assessment-only commitment equilibrium is analogous to the above except it restricts to zero all coefficients on lagged assessments and lagged policy rates.

G.3 Solving the Equilibrium and Computing Population Objects

G.3.1 State Dynamics and Forecast Projections

We detail the central bank's filtering problem, the state dynamics, and population projections for a candidate policy rule K and forecast matrix Θ. These are used to solve both discretionary and commitment equilibria.

Kalman filter. Recall the fundamental law of motion is given by

$$X_t = A_X X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim (0, \Sigma_\varepsilon),$$

and the central bank's signal equation by

$$s_t^C = X_t + e_t^C, \qquad e_t^C \sim (0, \Sigma_e),$$

where A_X, Σ_ε, and Σ_e are 3×3 primitives of the model pinning down the persistence of the fundamentals, the fundamental shock covariance, and the signal noise covariance, respectively.

The finite-history state in (79) has a law of motion given by

$$\mathcal{X}_t^H = \mathcal{A}_H \mathcal{X}_{t-1}^H + \mathcal{E}_t^H, \tag{87}$$

where

$$\mathcal{A}_H = \begin{pmatrix} A_X & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{pmatrix}, \qquad \mathcal{E}_t^H = (\varepsilon_t', 0', \ldots, 0')'.$$

Further, the central bank's signal evolution can be written as

$$s_t^C = \mathcal{A}_{s,H} \mathcal{X}_t^H + e_t^C, \qquad \mathcal{A}_{s,H} = \begin{pmatrix} I_3 & 0_{3\times 3H} \end{pmatrix}. \tag{88}$$

The central bank's filtering problem on this augmented system delivers the central bank assessments (81).

Persistence of assessment errors. The independent fundamental and signal processes allow the filter for each contemporaneous fundamental to be written separately. For component $j \in \{n, \rho, u\}$, let ℓ_j denote its steady-state Kalman gain. The evolution of the contemporaneous assessment of fundamental j is then

$$\widehat{X}_{j,t|t} = (1 - \ell_j)\varphi_j \widehat{X}_{j,t-1|t-1} + \ell_j s_{j,t}^C.$$

Consequently, starting from the steady state, a one-time signal error of size $e_{j,0}^C$ generates a sequence of assessment errors given by

$$\widehat{X}_{j,h|h} = \ell_j \left[(1 - \ell_j)\varphi_j\right]^h e_{j,0}^C, \qquad h \geq 0.$$

The decay rate of the induced assessment error is $(1 - \ell_j)\varphi_j$. At the baseline estimates, it is approximately 0.7336 for natural output and 0.0238 for demand, implying half-lives

of 2.24 and 0.19 quarters, respectively. Fundamental persistences are 0.9568 and 0.9666, respectively, so the larger half-life of natural output errors follows from the fact that its corresponding signal is particularly noisy.

State vector and stationary covariance. Rewrite the law of motion of the state vector of current and lagged fundamentals (79) as

$$\mathcal{X}_t^H = \mathcal{A}_H \mathcal{X}_{t-1}^H + \mathcal{B}_H \varepsilon_t,$$

where

$$\mathcal{B}_H = \begin{pmatrix} I_3 \\ 0_{3H\times 3} \end{pmatrix}$$

and thus $\mathcal{B}_H \varepsilon_t = \mathcal{E}_t^H$ loads the innovation into the current fundamental block. Recall, further, the law of motion of the signal (88) and let $\mathcal{K}$ denote the steady-state Kalman gain for this lag-augmented system. Then, the central bank's posterior mean defined in (81) follows the stationary updating rule

$$\widehat{\mathcal{X}}_{t|t}^{C,H} = \mathcal{K}\mathcal{A}_{s,H}\mathcal{A}_H \mathcal{X}_{t-1}^H + (I_{3(H+1)} - \mathcal{K}\mathcal{A}_{s,H})\mathcal{A}_H \widehat{\mathcal{X}}_{t-1|t-1}^{C,H} + \mathcal{K}\mathcal{A}_{s,H}\mathcal{B}_H \varepsilon_t + \mathcal{K} e_t^C.$$

Partitioning a generic fixed policy rule $K = \begin{pmatrix} K_{\mathcal{X}} & K_i \end{pmatrix}$, we can write the policy rate as

$$\tilde{i}_t = K_{\mathcal{X}} \widehat{\mathcal{X}}_{t|t}^{C,H} + K_i \begin{pmatrix} \tilde{i}_{t-1} \\ \vdots \\ \tilde{i}_{t-H} \end{pmatrix}. \tag{89}$$

To describe the joint evolution of fundamentals, central-bank assessments, and lagged policy rates, define the state vector

$$\mathcal{S}_t^H \equiv \begin{pmatrix} \mathcal{X}_t^H \\ \widehat{\mathcal{X}}_{t|t}^{C,H} \\ \tilde{i}_{t-1} \\ \vdots \\ \tilde{i}_{t-H} \end{pmatrix}. \tag{90}$$

Note that this is different from the central-bank state $\mathcal{S}_t^{C,H}$ since it also includes the realized current and lagged fundamentals $\mathcal{X}_t^H$. The state follows a linear law of motion

$$\mathcal{S}_t^H = A_{\mathcal{S}} \mathcal{S}_{t-1}^H + B_{\mathcal{S}} \eta_t, \qquad \eta_t \equiv \begin{pmatrix} \varepsilon_t \\ e_t^C \end{pmatrix}, \qquad \Sigma_\eta \equiv \begin{pmatrix} \Sigma_\varepsilon & 0 \\ 0 & \Sigma_e \end{pmatrix}, \tag{91}$$

where

$$A_{\mathcal{S}} = \begin{pmatrix} \mathcal{A}_H & 0 & 0 \\ \mathcal{K}\mathcal{A}_{s,H}\mathcal{A}_H & (I - \mathcal{K}\mathcal{A}_{s,H})\mathcal{A}_H & 0 \\ 0 & K_{\mathcal{X}} & K_i \\ 0 & 0 & \begin{pmatrix} I_{H-1} & 0 \end{pmatrix} \end{pmatrix}, \qquad B_{\mathcal{S}} = \begin{pmatrix} \mathcal{B}_H & 0 \\ \mathcal{K}\mathcal{A}_{s,H}\mathcal{B}_H & \mathcal{K} \\ 0_{H\times 3} & 0_{H\times 3} \end{pmatrix},$$

where we omit the dimensions of the zero and identity matrices to keep the notation concise. The first two block rows of $A_{\mathcal{S}}$ correspond to fundamentals and the central bank's assessments. The third includes the policy rate chosen in the previous period, and the fourth shifts the remaining rate lags. The rate-lag rows of $B_{\mathcal{S}}$ are zero because those rates were chosen before the current innovations.

Define the unconditional covariance matrix of the state as

$$\Sigma_{\mathcal{S}} \equiv \mathbb{E}[\mathcal{S}_t^H (\mathcal{S}_t^H)'], \tag{92}$$

which exists provided the matrix $A_{\mathcal{S}}$ is stable. Using the law of motion above, it is straightforward to show that the covariance matrix satisfies

$$\Sigma_{\mathcal{S}} = A_{\mathcal{S}} \Sigma_{\mathcal{S}} A_{\mathcal{S}}' + B_{\mathcal{S}} \Sigma_\eta B_{\mathcal{S}}'. \tag{93}$$

Forecast consistency. The market's optimal forecasting condition (78) can be written as the orthogonality condition

$$\mathbb{E}\left[\left(W_{t+1} - \Theta Z_t^H\right)(Z_t^H)'\right] = 0. \tag{94}$$

Since Z_t^H is a linear function of $\mathcal{S}_t^H$, we can write

$$Z_t^H = C_Z \mathcal{S}_t^H$$

for some matrix C_Z that maps the relevant components of $\mathcal{S}_t^H$ into Z_t^H. Similarly, current allocations are linear in $\mathcal{S}_t^H$ for a fixed Θ and policy rule K, so we can write

$$W_t = (y_t, \pi_t)' = C_W(\Theta, K)\mathcal{S}_t^H,$$

where the rows of C_W are defined by substituting (77), $y_t = x_t + y_t^n$ and the policy rule into (82).

Since η_{t+1} is orthogonal to $\mathcal{S}_t^H$,

$$\mathbb{E}[W_{t+1}(Z_t^H)'] = C_W A_{\mathcal{S}} \Sigma_{\mathcal{S}} C_Z', \qquad \mathbb{E}[Z_t^H (Z_t^H)'] = C_Z \Sigma_{\mathcal{S}} C_Z'.$$

Thus, for a given vector of parameters $\theta \in \mathbb{R}^9$ and policy rule K, forecast consistency is equivalent to the fixed-point condition

$$\Theta = \underbrace{C_W(\Theta, K) A_{\mathcal{S}} \Sigma_{\mathcal{S}} C_Z' \left(C_Z \Sigma_{\mathcal{S}} C_Z'\right)^{-1}}_{\equiv \mathcal{F}_\theta(\Theta, K)}. \tag{95}$$

Given the policy rule K, for any conjectured forecast matrix Θ, $\mathcal{F}_\theta(\Theta, K)$ returns the matrix of optimal projection coefficients implied by the resulting allocation process. Forecasts are consistent when those coefficients equal the conjecture.

G.3.2 Discretionary Equilibrium

Let d_X be the dimension of X_t. The law of motion for the central bank's state (80) can be written as

$$\mathcal{S}_{t+1}^{C,H} = A_C \mathcal{S}_t^{C,H} + B_C \tilde{i}_t + \xi_{t+1}^C, \qquad \mathbb{E}_t^C[\xi_{t+1}^C] = 0, \tag{96}$$

where

$$A_C = \begin{pmatrix} \mathcal{A}_H & 0_{d_X(H+1)\times H} \\ 0_{1\times d_X(H+1)} & 0_{1\times H} \\ 0_{(H-1)\times d_X(H+1)} & \begin{bmatrix} I_{H-1} & 0_{(H-1)\times 1} \end{bmatrix} \end{pmatrix}, \qquad B_C = \begin{pmatrix} 0_{d_X(H+1)\times 1} \\ 1 \\ 0_{(H-1)\times 1} \end{pmatrix}.$$

This representation uses the standard optimal forecasting implication that the expected next posterior mean is the corresponding predicted state.

Policy optimality. Consider a fixed Θ and define the central bank's forecast of the market forecast given a current demeaned rate $\tilde{i}$ as

$$\begin{pmatrix} \widehat{m}_{y,t}(\tilde{i}) \\ \widehat{m}_{\pi,t}(\tilde{i}) \end{pmatrix} \equiv \Theta \begin{pmatrix} \widehat{\mathcal{X}}_{t|t}^{C,H} \\ \tilde{i} \\ \tilde{i}_{t-1} \\ \vdots \\ \tilde{i}_{t-H} \end{pmatrix}.$$

Define the selectors

$$\iota_n = (1,0,0)', \qquad \iota_\rho = (0,1,0)', \qquad \iota_u = (0,0,1)'.$$

Using (82), the central bank's expected allocations can be written as

$$\widehat{x}_t(\tilde{i};\Theta) = -\iota_n'\widehat{X}^C_{t|t} + \gamma\iota_\rho'\widehat{X}^C_{t|t} - \gamma\tilde{i} + \widehat{m}_{y,t}(\tilde{i}) + \gamma\widehat{m}_{\pi,t}(\tilde{i}), \tag{97}$$

$$\widehat{\pi}_t(\tilde{i};\Theta) = \iota_u'\widehat{X}^C_{t|t} + \kappa\widehat{x}_t(\tilde{i};\Theta) + \beta\widehat{m}_{\pi,t}(\tilde{i}). \tag{98}$$

Splitting the forecast matrix according to the order of the regressors in Z_t^H, we get

$$\Theta =: \begin{pmatrix}\Theta_{\mathcal{X}} & \theta_0 & \Theta_-\end{pmatrix},$$

where $\Theta_{\mathcal{X}}$ multiplies $\mathcal{X}_t^H$, θ_0 multiplies the contemporaneous rate $\tilde{i}_t$, and Θ_- multiplies $(\tilde{i}_{t-1},\ldots,\tilde{i}_{t-H})'$. Let

$$\Theta_S \equiv \begin{pmatrix}\Theta_{\mathcal{X}} & \Theta_-\end{pmatrix}$$

be the part multiplying the central-bank state $\mathcal{S}_t^{C,H}$. Let d_n, d_ρ, d_u select from $\mathcal{S}_t^{C,H}$ the current posterior means of y_t^n, $\rho_t - \bar{\rho}$, and u_t, respectively. Finally define row vectors $a' = (1,\gamma)$ and $b' = (0,1)$. Then, we can write

$$\widehat{R}_t(\tilde{i};\Theta) \equiv \begin{pmatrix}\widehat{x}_t(\tilde{i};\Theta)\\ \widehat{\pi}_t(\tilde{i};\Theta)\\ \tilde{i}\end{pmatrix} = C_R(\Theta)\mathcal{S}_t^{C,H} + D_R(\Theta)\tilde{i}, \tag{99}$$

where

$$C_R(\Theta) \equiv \begin{pmatrix}C_x(\Theta)\\ C_\pi(\Theta)\\ 0_{1\times\dim(\mathcal{S}_t^{C,H})}\end{pmatrix}, \qquad D_R(\Theta) \equiv \begin{pmatrix}D_x(\Theta)\\ D_\pi(\Theta)\\ 1\end{pmatrix},$$

and

$$C_x(\Theta) \equiv -d_n' + \gamma d_\rho' + a'\Theta_S, \qquad D_x(\Theta) \equiv -\gamma + a'\theta_0, \tag{100}$$

$$C_\pi(\Theta) \equiv d_u' + \kappa C_x(\Theta) + \beta b'\Theta_S, \qquad D_\pi(\Theta) \equiv \kappa D_x(\Theta) + \beta b'\theta_0. \tag{101}$$

Thus, the central bank's contemporaneous expected loss can be written as $\frac{1}{2}\widehat{R}_t'\mathcal{Q}\widehat{R}_t$ plus a term involving posterior variances. Since the variance term is independent of the *current rate* (but not of the forecast matrix Θ), it is irrelevant for the discretionary problem.

Since the problem is linear-quadratic and we are looking for linear equilibria, we guess the continuation value is quadratic,

$$V_\Theta(\mathcal{S}) = \frac{1}{2}\mathcal{S}'P_\Theta\mathcal{S} + \text{constant}$$

for some P_Θ to be determined. Consequently, minimizing the current expected loss plus the discounted continuation value gives

$$\tilde{i}_t = K^D(\Theta)\mathcal{S}_t^{C,H}, \qquad K^D(\Theta) = -(\mathsf{R}_\Theta + \beta B_C' P_\Theta B_C)^{-1}(\mathsf{N}_\Theta' + \beta B_C' P_\Theta A_C), \tag{102}$$

where

$$\mathsf{Q}_\Theta \equiv C_R' \mathcal{Q} C_R, \qquad \mathsf{N}_\Theta \equiv C_R' \mathcal{Q} D_R, \qquad \mathsf{R}_\Theta \equiv D_R' \mathcal{Q} D_R.$$

Substituting the policy into the Bellman equation, then, pins down P_Θ as the solution to

$$P_\Theta = \mathsf{Q}_\Theta + \beta A_C' P_\Theta A_C - (\mathsf{N}_\Theta + \beta A_C' P_\Theta B_C)(\mathsf{R}_\Theta + \beta B_C' P_\Theta B_C)^{-1}(\mathsf{N}_\Theta' + \beta B_C' P_\Theta A_C). \tag{103}$$

Since we are looking for a stationary equilibrium, we further require that the optimal policy (102), after substituting the solution to (103), induces stationary state dynamics. Concretely, all the eigenvalues of the transition matrix $A_C + B_C K^D(\Theta)$ (obtained by substituting the policy rule into (96)) must have magnitudes strictly below one. This ensures that the effects of initial conditions vanish and, given finite innovation variances, a finite stationary covariance exists.

Thus, as per Definition 6 above, equations (102), (103), (94) and (82) determine the discretionary equilibria.

Numerical implementation and equilibrium selection. For an arbitrary candidate parameter vector $\theta \in \mathbb{R}^9$, the discretionary equilibrium fixed point can be written as

$$\Theta = \underbrace{\mathcal{F}_\theta\left(\Theta, K_\theta^D(\Theta)\right)}_{\equiv \mathcal{T}_\theta(\Theta)}, \tag{104}$$

where $K_\theta^D(\Theta)$ is given by (102) and $\mathcal{F}_\theta$ is defined in (95). Consequently, $\mathcal{T}_\theta$ incorporates both the forecast-consistency condition and the central bank's best response.

Thus, each evaluation of (104) entails computing the central bank's best response above, solving for the implied stationary covariance of the state (93) and using it to compute the population moments needed to update the forecast coefficients via $\mathcal{F}_\theta$.

The baseline implementation solves the fixed point (104) using Anderson-accelerated fixed-point iteration with depth five and relaxation 0.5. An additional nonlinear root finder is available as a fallback and for final polishing. Accepted solutions must satisfy the above forecast-consistency and filtering conditions at tight numerical tolerances. We also require stationary state dynamics as stated above and a (numerically) non-singular covariance matrix

of the market's forecasting regressors, ensuring finite population moments and uniquely determined forecast coefficients.

For some parameter vectors, the discretionary problem has more than one stationary fixed point. Among those, we restrict attention to locally E-stable equilibria, using the adaptive-learning criterion discussed by Evans and Honkapohja (2001).[60] Intuitively, as agents update their forecasting coefficients using observed outcomes, their revised expectations also feed back into policy and allocations. The criterion requires this feedback to correct small initial errors in the forecasting coefficients instead of reinforcing them. Formally, this entails a constraint on the eigenvalues of the Jacobian of the mapping $\mathcal{T}_\theta$ at the fixed point Θ^*,

$$\max_{\lambda \in \sigma(J_\theta)} \mathrm{Re}(\lambda) < 1, \qquad J_\theta \equiv D_\Theta \mathcal{T}_\theta(\Theta^*).$$

G.3.3 Optimal Commitment Policy

Naturally, the commitment problem uses the same central-bank filter, information state, allocation equations, and population-projection condition as the discretionary problem. The difference is that the central bank chooses the coefficients K^C directly in (86). Thus, for each candidate rule, we compute the market forecasts it induces and evaluate the resulting loss.

Forecast consistency. Fix a candidate policy K^C. Combining this rule with the fundamental processes and the Kalman filter determines the state's law of motion (91). We compute its stationary covariance as in (93) above. As in the discretionary equilibrium, we require the covariance matrix of the market's forecasting regressors to be numerically non-singular.

Note that given a policy rule, these state dynamics and their covariances do not depend on the conjectured forecast matrix Θ. Allocations, however, naturally do depend on Θ linearly via (82). Consequently, the population-projection update is affine in the forecast coefficients

$$\mathrm{vec}\, \mathcal{F}_\theta(\Theta, K^C) = a(K^C) + B(K^C)\, \mathrm{vec}\, \Theta,$$

for some matrix $B(K^C)$ and vector $a(K^C)$. Given a candidate policy rule, then, the fixed-point condition (95) can be written as a linear system in $\mathrm{vec}\, \Theta$,

$$[I - B(K^C)]\, \mathrm{vec}\, \Theta(K^C) = a(K^C). \tag{105}$$

[60] While in principle there may exist multiple locally E-stable equilibria, that did not happen in any of our numerical exercises so the "refinement" allowed us to recover a unique allocation.

We solve this directly.[61]

Policy optimization. We minimize the central bank's objective (86) over policy coefficients, recomputing the stationary covariance and consistent market forecasts (i.e., the solutions to (105)) for each candidate rule. The simple commitment rule is obtained by restricting the coefficients on lagged assessments and lagged rates to zero. Note that the central bank's objective (86) is non-convex in general. Thus, we can only guarantee that our numerical solution is a local minimum. To address this, we use multiple starting points in the numerical optimization and check local optimality of the solution.

G.3.4 Model-implied Moments

This section derives the model counterparts to the data moments used in the estimation. Once the stationary equilibrium has been found using the procedure above, the same state-space representation allows computing these moments analytically. Concretely, let v_t be any scalar, for example π_t, x_t, or $\tilde{i}_t$. Since we can write

$$v_t = c_v \mathcal{S}_t^H,$$

for some c_v, we can compute its second moments directly as

$$\mathrm{Var}(v_t) = c_v \Sigma_{\mathcal{S}} c_v', \qquad \mathrm{Cov}(v_t, v_{t-1}) = c_v A_{\mathcal{S}} \Sigma_{\mathcal{S}} c_v'.$$

Macro moments. Since all variables are measured as deviations from steady state, the AR(1) coefficient can be computed by projecting v_t on v_{t-1} alone. Thus, it can be computed as

$$\rho_v = \frac{\mathrm{Cov}(v_t, v_{t-1})}{\mathrm{Var}(v_t)} = \frac{c_v A_{\mathcal{S}} \Sigma_{\mathcal{S}} c_v'}{c_v \Sigma_{\mathcal{S}} c_v'}. \tag{106}$$

Similarly, the corresponding AR(1) residual variance can be computed as

$$\sigma^2_{v,\mathrm{AR1}} = \mathrm{Var}(v_t) - \rho_v^2 \,\mathrm{Var}(v_{t-1}) = c_v \Sigma_{\mathcal{S}} c_v' - \frac{\left(c_v A_{\mathcal{S}} \Sigma_{\mathcal{S}} c_v'\right)^2}{c_v \Sigma_{\mathcal{S}} c_v'}. \tag{107}$$

[61] The E-stability refinement used for discretion applies to $\mathcal{T}_\theta$ above, which incorporates the central bank's adjustment of its policy rule as forecasting coefficients change. The commitment optimization internalizes how each candidate policy rule shapes consistent market forecasts. We do not, however, require the resulting forecasting fixed point to be locally E-stable under a learning process in which the central bank continues to follow a given candidate rule. In fact, at the baseline estimates, this fixed-rule forecast-learning map is not locally E-stable.

Applying these formulas, we obtain the six macro moments

$$\rho_\pi, \rho_x, \rho_i, \qquad \sigma^2_{\pi,\mathrm{AR1}}, \sigma^2_{x,\mathrm{AR1}}, \sigma^2_{i,\mathrm{AR1}}.$$

Signal noise and natural-output uncertainty. For each fundamental $j \in \{n, \rho, u\}$, let

$$V_j \equiv \frac{\sigma_j^2}{1-\varphi_j^2}, \qquad q_j \equiv \frac{(\sigma_j^e)^2}{V_j}.$$

Thus, V_j is fundamental j's unconditional variance and q_j is the ratio of the signal-noise variance to the unconditional variance of the corresponding fundamental. Let $P_{t|t}^{C,H}$ denote the stationary covariance matrix from the central-bank filter (after observing the corresponding contemporaneous signal) and let $d_n^{\mathcal{X}}$ select its current natural-output component. Again letting $\theta \in \mathbb{R}^9$ denote an arbitrary vector of primitives, the model counterpart of the HLW target is

$$\mathcal{H}^{\mathrm{model}}(\theta) = \sqrt{(d_n^{\mathcal{X}})' P_{t|t}^{C,H}(\theta) d_n^{\mathcal{X}}}. \tag{108}$$

The data target is the average of the 120 one-sided, post-signal posterior standard deviations from the re-estimated HLW filter over 1990Q1–2019Q4.

Greenbook/Tealbook moments. For a given variable z_t, its model posterior-variance share is obtained by projecting it onto the central bank's state $\mathcal{S}_t^{C,H}$ defined in (80),

$$R_z^{\mathrm{model}} = \frac{\mathrm{Var}\left(z_t - \mathrm{Proj}(z_t \mid \mathcal{S}_t^{C,H})\right)}{\mathrm{Var}(z_t)} = 1 - \frac{\mathrm{Cov}(z_t, \mathcal{S}_t^{C,H})\,\mathrm{Var}(\mathcal{S}_t^{C,H})^{-1}\,\mathrm{Cov}(\mathcal{S}_t^{C,H}, z_t)}{\mathrm{Var}(z_t)}. \tag{109}$$

We compute (109) for inflation π_t and real GDP growth, $y_t - y_{t-1}$. The annualization factor applied to GDP growth cancels from the variance ratio. Because a conditional-expectation error has mean zero, its posterior variance equals its mean squared error. Thus, the posterior error-variance share coincides with the posterior MSE share. The empirical counterpart is the variance of later-vintage-minus-nowcast revisions divided by the variance of the later-vintage estimate.

Inflation-swaps moment. Let $Z_t^{H,-}$ denote the pre-policy market state, obtained by removing the contemporaneous policy rate from Z_t^H. For any variable a, define its finite-history revision upon seeing the policy rate

$$\Delta_{t,H}^M a \equiv \mathrm{Proj}(a \mid Z_t^H) - \mathrm{Proj}(a \mid Z_t^{H,-}).$$

The model revisions can then be written as

$$\Delta P^e_{t,1:20} \equiv \Delta^M_{t,H}\left(\sum_{h=1}^{20}\pi_{t+h}\right), \tag{110}$$

$$mps^M_t \equiv \Delta^M_{t,H}\left(\sum_{h=0}^{3} i_{t+h}\right), \tag{111}$$

and the relevant sensitivity coefficient is

$$\beta^P_{1:20} \equiv \frac{\mathrm{Cov}(\Delta P^e_{t,1:20}, mps^M_t)}{\mathrm{Var}(mps^M_t)}.$$

Note that both revisions are linear functions of the state $\mathcal{S}^H_t$. Consequently, the covariance and variance defining $\beta^P_{1:20}$ can be evaluated analytically using the stationary state covariance. We omit the explicit matrix expressions for the sake of brevity.

G.4 Estimation Details

Recall we denote by $\theta \in \mathbb{R}^9$ an arbitrary vector of parameters. Let j index our relevant moments and let m_j^{data} and $m_j(\theta)$ denote the corresponding data moment and its model counterpart. Further, let $g_j(\theta) \equiv m_j(\theta) - m_j^{\text{data}}$, and let v_j be the estimated sampling variance of the corresponding data moment. Our minimum-distance objective function is

$$Q(\theta) = \sum_{j\in\mathcal{M}} \frac{g_j(\theta)^2}{v_j} \tag{112}$$

where in addition we impose $\mathcal{H}^{\text{model}}(\theta) = \mathcal{H}^{\text{data}}$ and $\mathcal{M}$ contains the nine moments.

Sampling variances. The sampling variances are estimated by bootstrap. For each macro series, we use 5,000 draws from a univariate AR(1) sieve bootstrap, resampling centered AR(1) residuals and removing the first 200 quarters' simulated observations from the sample. In the case of output-gap moments, the bootstrap is conducted using the Hamilton-filtered series extracted from real GDP as introduced in the main text. For each Greenbook/Tealbook moment, we use 5,000 target/quarter resamples of the paired revision and later-vintage estimate and recompute the forecast-revision variance share. For the inflation-swaps moment, we use 10,000 bootstrap draws that resample event dates. In each draw, we re-estimate the 0–2 year and 2–5 year response coefficients using the same resampled events and recompute the combined target $\beta^P_{1:20} = 2\beta^\pi_{1:8} + 3\beta^\pi_{9:20}$.

For inflation and the policy rate, the empirical moments and their bootstrap sampling

variances are initially computed using annualized series. To express them in the quarterly units used in the model, we divide the residual-variance targets by 16 and their bootstrap sampling variances by 256 (thus, the corresponding inverse-variance weights are multiplied by 256).

Table 8: Minimum-Distance Targets and Weights

Moment	Data	Model	Sampling variance	Q contribution
AR(1) (π)	0.1321	0.0951	0.0075	0.1825
AR(1) (x)	0.8943	0.6326	0.0026	26.7151
AR(1) (i)	0.9643	0.9642	0.0016	5.7437×10^{-6}
Residual variance (π)	0.1729	0.1582	0.0051	0.0426
Residual variance (x)	1.1785	1.1972	0.0383	0.0091
Residual variance (i)	0.0103	0.0103	3.4527×10^{-6}	0.0000
Natural-output uncertainty ($\mathcal{H}$)	1.7578	1.7578	–	–
Revision-variance share (π)	0.2414	0.2299	0.0055	0.0242
Revision-variance share (GDP growth)	0.5876	0.4342	0.0446	0.5271
$\beta^{P}_{1:20}$	−2.4486	−3.2574	0.8201	0.7976
			Total	28.2983

Notes: All scored moments use inverse-sampling-variance weights. Natural-output uncertainty is imposed exactly and excluded from Q. Inflation and policy-rate residual variances are in quarterly percentage points squared. Contributions to Q are calculated from unrounded moments and sampling variances.

The estimated parameters are $(\sigma_n, \sigma_\rho, \sigma_u, \varphi_n, \varphi_\rho, \varphi_u, q_n, q_\rho, q_u)$. We impose

$$10^{-4} \le \sigma_j \le 20, \qquad 0 \le \varphi_j \le 0.995, \qquad 10^{-4} \le q_n, q_u \le 100, \qquad 10^{-4} \le q_\rho \le 1000.$$

Imposing $\mathcal{H}$ exactly leaves eight free coordinates.

For every candidate parameter vector, the estimator solves the discretionary equilibrium using the procedure described in the previous section. The reported estimate has $Q = 28.2983$, and none of the parameter constraints are binding.

G.5 Additional Results

G.5.1 Loss Decomposition

Table 9 below reports the decomposition of total loss into its three components for each of the relevant policy regimes.

Table 9: Loss decomposition in three policy regimes

Policy	Output gap	Inflation	Rate	Total
Discretion	0.9981	0.0798	0.0732	1.1510
Commitment: simple rule	0.0145	0.0769	0.0472	0.1386
Full commitment	0.0103	0.0796	0.0359	0.1257

G.5.2 Conditional Variances and Commitment Gains

Table 10 decomposes output-gap and inflation variance into variation in the central bank's conditional mean and variation around that mean. Concretely, for $z_t \in \{x_t, \pi_t\}$,

$$\mathrm{Var}(z_t) = \mathrm{Var}\Big(\mathbb{E}_t^C\, z_t\Big) + \mathbb{E}\Big[\mathrm{Var}(z_t \mid \mathcal{I}_t^C)\Big].$$

Table 10: Conditional Variances and the Sources of Commitment Gains

	Discretion	Simple rule	Full commitment
A. Output-gap variance			
Variation in the CB's expected gap	0.0005	0.0203	0.0167
Variation around the CB's expected gap	1.9957	0.0087	0.0038
Total	1.9962	0.0290	0.0205
B. Inflation variance			
Variation in the CB's expected inflation	0.1229	0.1306	0.1373
Variation around the CB's expected inflation	0.0367	0.0231	0.0218
Total	0.1596	0.1537	0.1591
C. Contributions to the gain over discretion (%)			
Combined conditional-error reduction	–	**98.79**	**97.86**
Output gap	–	98.12	97.13
Inflation	–	0.67	0.73
Conditional-mean allocation loss reduction	–	−1.36	−1.49
Rate-variance reduction	–	2.57	3.63
Total gain	–	100.00	100.00

Notes: Panels A and B report stationary variances. Panel C applies the baseline loss weights, $\vartheta_\pi = \vartheta_i = 1$, and the factor 1/2 in the objective. Negative entries indicate an increase in that loss component.

The reduction in loss attributable to variation around the central bank's expected output gap and inflation amounts to 1.0003 under the simple rule and 1.0033 under full commitment, accounting for approximately 98% of both commitment gains. Smaller rate deviations

account for the rest, with a small increase in the loss from worse conditional average stabilization. Thus, the lion's share of the gains from commitment comes from reducing deviations of realized allocations from those anticipated by the central bank.

G.5.3 Demand Innovations and Demand-Signal Errors

Table 11 documents the impact of demand-related shocks supporting the brief discussion in the main text. Each shock corresponds to a positive one-standard-deviation increase in the corresponding primitive (demand innovation and demand-signal error innovation).

Table 11: Demand Disturbances under Discretion and the Simple Rule

	Demand-signal error		Demand innovation	
	Discretion	Simple rule	Discretion	Simple rule
A. Impact responses				
Policy rate	0.0601	0.0493	0.3745	0.3067
Output gap	−0.5049	−0.0803	0.0482	−0.0222
Inflation	−0.1280	−0.0081	0.0112	−0.0759
Expected next-period inflation	−0.0804	−0.0004	0.0066	−0.0746
Ex ante real rate	0.1405	0.0497	0.3679	0.3813
B. Stationary contribution of each innovation				
Output-gap variance	0.3332	0.0065	0.0028	0.0175
Loss	0.1675	0.0033	0.0737	0.0585

Notes: Interest-rate and inflation responses are in annualized percentage points; the output gap is in percentage log-output points. The real rate is $i_t - m_{\pi,t}$, where $m_{\pi,t}$ is the market's finite-history forecast of next-period inflation. Note that each shock is a one-standard-deviation increase in the corresponding primitive, and the standard deviations are different for the two shocks. This explains the different policy rate response magnitudes (the reaction coefficient is of course the same for both since the central bank does not see them separately but only their sum).

For the demand-signal error, the smaller rate hike under the simple rule entails a smaller fall in expected inflation and a much smaller contraction. Although signal errors are independent over time, their effects on the central bank's assessments persist as subsequent signals correct the initial mistake. Concretely, after one quarter, 2.38% of a demand-signal error's initial effect remains, while the corresponding number for a natural-output signal error is 73.36% (see Appendix G.3.1 for details). Under discretion and the simple rule, the resulting transitory tightening produces the conventional contractionary and disinflationary effects. The market's belief revision under discretionary policy significantly amplifies the contractionary effect of the initial error (see Appendix G.5.6 for an exercise that shows this).

Instead, a genuine demand innovation persistently raises the natural real rate and, as the

central bank's subsequent signals arrive, its demand assessments. To understand the resulting inflation response, note that since a demand innovation leaves natural output constant, the IS equation implies

$$i_t - \mathbb{E}_t^M \pi_{t+1} = \rho_t + \frac{\mathbb{E}_t^M x_{t+1} - x_t}{\gamma}. \tag{113}$$

During the first periods following the shock, subsequent signals correct the central bank's initial underestimation of the demand innovation, while demand itself remains elevated due to its high persistence. After that "learning phase" is (essentially) over, under discretion the nominal-rate increase is approximately 1.0071 times the remaining natural-rate increase and the equilibrium features positive expected inflation. Under the simple commitment rule, that ratio is approximately 0.8098 and the equilibrium features negative expected inflation. In both cases, the resulting real rate remains close to the natural rate (given by ρ_t, since natural output growth is zero in this exercise) and thus, together with an essentially constant output gap, is consistent with (113) above and the remaining equilibrium conditions.[62]

Thus, despite its more restrained nominal tightening, the simple commitment rule induces a higher real rate and a small initial output contraction. Beyond the effects on impact, the stationary loss contribution of fundamental demand innovations falls by 20.69%.

G.5.4 Cost-Push Assessments and Policy Interactions

Cost-push assessments illustrate how the effects of policy depend on the market's interpretation of interest-rate movements. We compare genuine cost-push innovations with signal errors and then examine the interaction between the cost-push response and the other policy coefficients.

Fundamental innovations and signal errors. Table 12 reports the on-impact responses to positive one-standard-deviation innovations in the cost-push fundamental and its signal noise (starting from the steady state and setting all other innovations to zero as is standard). Since the market observes the fundamental innovation, its associated policy response is anticipated and generates no surprise. Under the simple rule, the rate rises somewhat more, while current and expected next-period inflation rise slightly less and the output gap increases substantially less.

[62] While at odds with the standard view that higher nominal rates cool down the economy, this Fisherian phenomenon is a feature of standard New Keynesian models with rational expectations (see García-Schmidt and Woodford (2019) for a discussion of this and an alternative approach). Our analysis inherits this feature from the forward-looking New Keynesian block and the rational expectations equilibrium concept.

Table 12: Cost-Push Disturbances under Discretion and the Simple Rule

	Fundamental innovation		Signal error	
	Discretion	Simple rule	Discretion	Simple rule
Policy rate	0.0632	0.0741	0.0269	0.0316
Output gap	0.0711	0.0134	-0.2262	-0.0515
Inflation	1.5458	1.5338	-0.0574	-0.0052
Expected next-period inflation	0.0893	0.0828	-0.0360	-0.0003

Notes: Impact responses to one-standard-deviation innovations ($\sigma_u = 0.3627$ and $\sigma_u^e = 0.1546$ respectively, in quarterly units). Rates and inflation are in annualized percentage points; the output gap is in percentage log-output points.

By contrast, a cost-push signal error produces an unexpected rate increase. The market does not observe the error and interprets the tightening partly as evidence of the central bank's assessments of natural output and demand. These inferences generate a substantial fall in expected inflation and a contraction under discretion. The simple rule sharply attenuates the resulting movements in expectations and allocations despite a larger rate response.

The responses are chosen jointly. As the counterfactual analysis above illustrates, each policy response affects the propagation of errors beyond its corresponding assessment. To examine these interactions, Table 13 considers changes to the cost-push coefficients starting from both discretion and the simple rule, keeping all other coefficients fixed. Market forecasts are recomputed for each rule.

Table 13: Losses under Cost-Push Coefficient Replacements

	Cost-push coefficients	
Other coefficients	Discretion	Simple rule
Discretion	1.1510	1.1180
Simple rule	0.1525	0.1386

Notes: Stationary per-period losses at the baseline parameters. The cost-push block contains the coefficients on current and past cost-push assessments. All remaining assessment coefficients and lagged-rate coefficients are held at the policy specified in the first column.

With the other coefficients fixed at discretion, replacing the cost-push block increases the loss attributable to cost-push disturbances from 0.1081 to 0.1142. However, it lowers the combined loss attributable to natural-output and demand disturbances by 0.0391, more than offsetting the cost-push loss increase. Thus, total loss falls from 1.1510 to 1.1180. As in the counterfactual exercise above, the realized rate paths following natural-output and demand disturbances are unchanged. The improvement in their effects on allocations thus

comes from the induced change in the market's interpretation of those rates. Modifying the cost-push coefficients changes how the market attributes observed rate movements to the central bank's different views and, consequently, how its expectations react to a given policy decision. Hence, the cost-push response can improve stabilization through its effects on the propagation of other disturbances, even when it worsens stabilization of cost-push disturbances themselves.

With the other coefficients fixed at the simple commitment rule, the same replacement instead lowers the cost-push contribution from 0.0867 to 0.0754. It also lowers the combined contribution of natural-output and demand disturbances by 0.0025, reducing total loss from 0.1525 to 0.1386. Thus, using only the simple rule's cost-push block improves aggregate stabilization in both comparisons, even though it worsens the stabilization of cost-push disturbances when the other coefficients remain at their discretionary values. These interactions illustrate the importance of our mechanism: the effects of a response to a particular fundamental depend on how the entire rule shapes the market's interpretation of decisions (and thus, on the other response coefficients). At the same time, at our estimates, the cost-push responses are not crucial for the large stabilization gains from the simple rule. In particular, the rule retaining the discretionary cost-push coefficients achieves 98.63% of the gain from discretion to the simple rule. Changing the cost-push block thus provides a relatively small additional improvement once the other responses have been adjusted.

G.5.5 Commitment Gains across Information Structures

We consider two environments that differ from our quantitative model only in their information structure: full information (i.e., contemporaneous fundamental shocks are common knowledge) and imperfect but public central bank information (i.e., the central bank receives imperfect signals about fundamentals, and the market observes fundamentals and those signals contemporaneously). We keep the baseline parameters at their baseline estimated values, except for the parameters governing the central bank's information quality, which in the full-information economy are such that the central bank observes the fundamentals perfectly.

In each economy, we compare discretion with the analogue of our simple commitment rule: a linear rule using only the three current fundamentals (full information) or the three contemporaneous central bank assessments (public and private assessments). Table 14 reports the losses under both policy regimes and the associated gains from commitment in each economy.

Table 14: Losses and Gains from Commitment across Information Environments

Information	Discretion loss	Simple-rule loss	Gain (%)
Full information	0.1625	0.1330	18.14
Public assessments	0.3505	0.1389	60.36
Private assessments	1.1510	0.1386	87.96

The losses under a simple commitment rule are fairly similar across information environments, while our benchmark with private central-bank assessments entails substantially higher loss under discretionary policy.

G.5.6 A Counterfactual Policy Rule

This exercise starts from the baseline discretionary rule and replaces its coefficients on natural-output assessments with the simple rule's response. Concretely, write the original discretionary rule as

$$\tilde{i}_t^D = \sum_{\ell=0}^{H} k_{n,\ell}^D \widehat{y}_{t-\ell|t}^n + R_t^D,$$

where $\widehat{y}_{t-\ell|t}^n$ is the central bank's period-t assessment of period-$t-\ell$ natural output, and R_t^D collects the demand-assessment, cost-push-assessment, and lagged-rate terms. The counterfactual rule is then

$$\tilde{i}_t^{\text{cf}} = k_{n,0}^{\text{simple}} \widehat{y}_{t|t}^n + R_t^D,$$

that is, the current natural-output coefficient is replaced and all coefficients on past natural-output assessments are set to zero. The other policy coefficients are unchanged. Crucially, the market acknowledges the new policy rule and thus adjusts its forecasting structure to be consistent with this counterfactual rule. Note that in general, this does not correspond to an equilibrium, since the rule is optimal neither under discretion nor under commitment.

We consider the effects of a one-standard-deviation demand-signal error in both the discretionary economy and this counterfactual one. Since fundamentals and signal errors are independent across blocks (i.e., natural output, demand, and cost-push), this leaves all natural-output assessments unchanged. Consequently, the implied interest rate path is the same under both rules. Nonetheless, the market's interpretation of the policy surprise triggered by the policy response to the demand assessment differs across economies, and the resulting output-gap and inflation responses differ as well. Table 15 reports the results.

Notably, the output-gap contraction falls by 81.18%. The market's estimate of the central bank's contemporaneous natural-output assessment falls by only 0.0298 instead of 0.4190.

Table 15: The consequences of a demand-signal error under a counterfactual policy rule

	Private assessments	
	Discretion	Counterfactual
A. Responses on impact		
Policy rate	0.0150	0.0150
Output gap	−0.5049	−0.0951
Inflation	−0.0320	−0.0023
Expected next-period output	−0.2854	−0.0011
Expected next-period inflation	−0.0201	+0.0000
B. Revisions to the market's beliefs about contemporaneous CB assessments		
Natural output (actual: 0)	−0.4190	−0.0298
Demand (actual: 0.0152)	0.0057	0.0126
C. Stationary contribution of demand-signal noise		
Loss	0.1675	0.0046
Policy-rate variance	0.0002	0.0002

Notes: Interest rates, inflation, and demand assessments are in quarterly percentage points; output and natural-output assessments are in percentage log-output points. Policy-rate variance is in quarterly percentage points squared. Loss is the stationary contribution of demand-signal noise under the quarterly objective with unit weights.

Beyond the effects on impact, the stationary loss contribution of demand-signal noise falls by 97.23%.

This exercise highlights the relevance of the belief-revision component of monetary transmission: changing the response to one assessment changes the consequences of an error in another, even though the policy path remains unchanged.

H Robustness Exercises

H.1 History-Length Sensitivity

To assess sensitivity of our results to the finite-history approximation, we increase the retained-history horizon from 8 to 12 lags and re-estimate the model using the same targets, weighting scheme, bounds, and fixed parameters as in the baseline. We then solve for the simple and full linear commitment rules at the new estimates. In this exercise, the discretionary and full-commitment linear rules each expand from 35 to 51 coefficients since they can condition on four additional lags of each assessment and four additional policy-rate lags. Table 16 reports the resulting values of the key policy and belief-sensitivity statistics.

Table 16: History-Length Sensitivity

Statistic	Policy	$H = 8$	$H = 12$
A. Commitment gains			
Loss	Discretion	1.1510	1.1640
	Commitment: simple rule	0.1386	0.1462
	Full commitment	0.1257	0.1331
Full-rule gain / discretionary loss	Full commitment	89.08%	88.56%
Commitment: simple rule (share of F.C. gain)	Commitment: simple rule	98.75%	98.73%
B. Expectations management			
$\beta^P_{1:20}$	Discretion	-3.2574	-3.3701
	Commitment: simple rule	-0.0173	-0.0090
	Full commitment	+1.6914	+1.4505
Policy-surprise SD	Discretion	0.0308	0.0297
	Commitment: simple rule	0.0150	0.0161
	Full commitment	0.0222	0.0211
Fitted revision SD	Discretion	0.1004	0.1001
	Commitment: simple rule	0.0003	0.0001
	Full commitment	0.0375	0.0306

Note: Each column uses the parameters estimated at the corresponding history length. The policy surprise and fitted revision are defined as in Table 6.

The full-commitment gains remain close to their baseline, amounting to 88.56% of the discretionary loss. Further, the simple rule captures 98.73% of that gain. As in our baseline, both commitment regimes reduce the size of expectations sensitivity and policy-surprise volatility relative to discretion. In the full commitment case, the reductions in the magnitude of the expectations sensitivity and fitted revision volatility are somewhat larger but still reasonably close: 56.96% and 69.45%, compared with 48.08% and 62.64% at $H = 8$. Thus, the main conclusions remain stable when the model is re-estimated with a longer retained history.

We also conduct a complementary exercise to assess the accuracy of our forecast truncation. Holding the parameters, policy rules, and implied dynamics of output and inflation fixed at their baseline values, we compute forecasts of these outcomes using the full history of fundamentals and policy rates. We replace the truncated-history forecasts in the IS and Phillips curves with these full-history forecasts and compute the resulting discrepancies between the two sides of each equation. Since the reported allocations satisfy both equations under the truncated-history forecasts, these residuals isolate the discrepancy induced by the change in forecasts. Under discretion, simple commitment, and full commitment, respectively, the IS residual standard deviations are 0.38%, 0.13%, and 2.45% of the output-gap standard deviation. The corresponding Phillips-curve residual standard deviations are

0.060%, 0.0012%, and 0.138% of the inflation standard deviation. These small residuals provide direct evidence on the accuracy of forecast truncation *at the baseline allocations*.

H.2 Rate-Level Stabilization Weight

In our baseline specification, we fix $\vartheta_i = 1$ as a transparent benchmark. To assess the sensitivity of the main results to this choice, we fix $\vartheta_i \in \{0.25, 0.5, 1.5, 2\}$ and re-estimate the remaining parameters using the same targets and weighting scheme as in the baseline. Then, we solve for the two commitment policy regimes in each estimated economy. Table 17 reports the results.

Table 17: Varying the Rate-Level Penalty ϑ_i

ϑ_i	Commitment gain	Simple rule share	R_β	R_i	Fitted revision SD ratio
0.25	91.60%	99.39%	0.565	0.826	0.467
0.50	90.55%	99.23%	0.546	0.785	0.428
1.00	89.08%	98.75%	0.519	0.720	0.374
1.50	88.13%	98.23%	0.500	0.670	0.335
2.00	87.50%	97.76%	0.484	0.630	0.305

Notes: At each value of ϑ_i, the parameters governing information and the fundamentals' processes are re-estimated using the baseline targets, weights, and bounds, and the commitment problem is solved for that economy. The row for $\vartheta_i = 1$ is the baseline estimate. The ratios compare full commitment (C) with discretion (D): $R_\beta = |\beta^P_{1:20,C}|/|\beta^P_{1:20,D}|$ and $R_i = \mathrm{SD}(mps^M_C)/\mathrm{SD}(mps^M_D)$. The fitted revision SD ratio is their product, $R_\beta R_i$.

The commitment gain remains large across these alternative estimates, ranging from 87.50% to 91.60% of discretionary loss. The simple commitment rule captures between 97.76% and 99.39% of that gain. Thus, both the substantial cost of discretion and the fact that it is largely not a consequence of history dependence are robust to changes in the rate-level weight.

Full commitment reduces both absolute expectations sensitivity and policy-surprise volatility throughout the grid. As ϑ_i rises from 0.25 to 2, R_β falls from 0.565 to 0.484 and R_i from 0.826 to 0.630, lowering the fitted-revision SD ratio from 0.467 to 0.305. These results establish that the main conclusions of the quantitative section are robust to the choice of the rate-level penalty within a reasonable range.