Forecasts, Forward Guidance, and Credibility*

Lukasz A. Drozd

Federal Reserve Bank of Philadelphia

lukasz.drozd@phil.frb.org

September 2026

Preliminary. Do not cite.

Abstract

The two largest U.S. inflation outbreaks began with supply shocks but featured different paths of expectations. We model a public that second-guesses a delegated central bank on three things—whether it will be held to its word, whether its forecast is right, whether its objective is what it says—each with its own channel into expectations. Enforcement skepticism is bounded, by exposure alone, at any stake; forecast skepticism is not, and its cost is convex; the mandate fixes the third. In the post-COVID look-through, inflation peaked at 5.5 percent with expectations below 3 because the public accepted the forecast. The risk was that acceptance failing, not enforcement.

Keywords: central-bank communication, forward guidance, inflation targeting, credibility, asymmetric information, optimal monetary policy.

JEL codes: E31, E52, E58, E61, E65, D82, D83.

*I thank João Gomes, Urban Jermann, and Alessandro Dovis for insightful comments at various stages of this project. Alessandro's comments helped clarify the institutional interpretation developed here and in a companion paper. The views expressed herein are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Correspondence: Lukasz A. Drozd, Federal Reserve Bank of Philadelphia, Research Department, 10 N. Independence Mall West, Philadelphia, PA 19106.

1 Introduction

The Great Inflation of the 1970s and the post-COVID inflation surge present a striking contrast. Supply disruptions and monetary accommodation featured in both episodes, but their resolution differed sharply. The Great Inflation became persistent and ended with costly disinflation. After COVID, inflation fell substantially without the deep contraction many had anticipated, an outcome often described as *immaculate disinflation*. The unusually low output cost of the recent disinflation makes it an important departure from earlier experience (Nelson, 2022; Forbes et al., 2025).

A prominent explanation emphasizes central bank credibility. A record of low inflation, reinforced by a transparent targeting framework, allowed central banks to accommodate supply pressures without generating a comparable increase in expected inflation. Recent work argues that this credibility gave central banks greater scope to look through the post-COVID shocks (Nakamura et al., 2025; English et al., 2024). But what, precisely, must the public believe? Confidence that a central bank will honor its policy commitments is different from confidence in its assessment that inflationary pressures will fade. An assurance that inflation will decline can rest on either belief, with different implications when the assurance proves wrong.

This paper develops a theory of flexible inflation targeting with imperfect accountability to distinguish these channels. Society delegates monetary policy to a central bank operating under a publicly understood standing rule. The rule remains fixed in the short run, but the central bank can announce future departures in response to its assessment of economic conditions. The rule and the previously announced departure together specify the current policy obligation,

$$a_t = \bar{\theta}(\widetilde{S}_t) + \Delta_t^p.$$

Institutional sanctions discipline deviations from this obligation. The same accountability applies whether the central bank has issued additional guidance or allowed the standing rule to speak for it.

The arrangement provides two forms of flexibility. Before policy comes due, the central bank can adjust its obligation through an announced departure. At delivery, finite sanctions permit it to deviate when honoring that obligation becomes sufficiently costly. Policy therefore remains a choice, but the standing rule and prior announcements affect the consequences of that choice. The strength of enforcement is persistent and privately observed. The public learns about it from realized policy, and its beliefs influence how much of an announced departure it expects the central bank to deliver.

We use this framework to separate what the public can second-guess about a delegated central bank into three objects—whether it will be held to its word, whether its forecast is right, whether its objective is what it says—each with its own channel into expectations. The mandate fixes the third. The theory bounds the first. The application prices the second, by asking whether the same supply disturbance can generate a modest response of inflation expectations in one case and a large, persistent response in another. It considers a central bank that repeatedly expects supply pressure to fade more quickly than it does. When the public accepts that assessment, the standing rule can accommodate inflation surprises while expectations remain comparatively low: current inflation is high, but the conditions inducing it are expected to recede. If the public instead expects the pressure to persist, it anticipates continued accommodation.

The theory places this institution in a standard linear-quadratic New Keynesian economy. The assessment of future conditions is public; enforcement is the central bank's only private information. The framework builds on Drozd (2026), which studies communication when the assessment is private and enforcement is known. The audit here has two components: one rises with the delivery miss and is borne in proportion to the hidden exposure, the other is levied on the output gap and is borne in full by both modes, offsetting a systematic inflationary incentive without altering announcement incentives. After the gap component, both modes share stabilization objectives and choose policy sequentially; they differ only in the sanction for a delivery miss. We use *enforcement credibility* for the public's belief about the mode, distinguishing it from acceptance of the forecast.

Two properties of the framework capture the key trade-offs and make the application possible. First, bounded accountability limits how far any policy obligation can move delivery from the response the central bank prefers when policy comes due. This *communication frontier* applies to the rule as well as to additional guidance. An obligation near the frontier is sensitive to changes in conditions or enforcement. With the direct sanction alone, a more ambitious promise eventually induces less delivery. The analysis also bounds the additional discipline supplied by future reputational consequences.

Second, we establish conditions under which both modes choose the announcement preferred by the high-enforcement mode. The low-enforcement mode can imitate while delivering less of the adjustment the promise requires. Announcements then influence expected policy without revealing enforcement strength. The public learns from realized inflation, after accounting for the rule's response to observed conditions.[1] Learning is

[1]Under the conditions in Corollary 1, the pool is undefeated in the sense of Mailath et al. (1993). The comparison concerns pure-strategy announcement equilibria and does not establish uniqueness.

limited when the modes deliver similar policies, so stable inflation need not by itself establish strong enforcement.

The application is calibrated to the United States. The Federal Reserve entered the post-COVID episode with a two-percent target and guidance under its flexible average inflation targeting framework that envisioned inflation moderately above target to offset past shortfalls (Federal Open Market Committee, 2020; Ihrig and Waller, 2024). We represent that guidance as an inherited departure from the standing rule. Each year, the central bank expects rapid normalization, while realized supply pressure declines more slowly. The baseline public accepts the forecast. The rule accommodates much of the resulting surprise, so realized inflation rises much more than short-term expectations. As pressure subsides, inflation declines and the output gap closes.

Counterfactuals retain the realized disturbance and the announcements and change what the public believes about one primitive at a time: the probability of the low-enforcement mode, or the persistence of the disturbance. The central bank, misperceiving how it is read, announces as in the baseline and delivers optimally against the expectations it faces. Skepticism about enforcement has modest effects, because it reaches only the term of expected inflation that the communication frontier bounds, and the gap between the two modes' responses is capped by exposure alone, at any stake. A public that assigns the disturbance more persistence than the central bank publishes reaches the terms the frontier does not bound, and the response is convex: modest departures are absorbed while the delivery miss stays inside the sanction's responsive range, and large ones are not, so that at the realized persistence the same shock produces a surge, with enforcement intact. Letting the central bank learn how it is read and re-optimize its announcements changes nothing, because beyond that range no announcement buys different delivery. The risk of looking through a supply shock therefore lies less in whether promises will be kept than in whether the public accepts the central bank's assessment of the economy.

The distinction between the standing rule and additional guidance also matters for interpreting the favorable outcome. In the calibration, much of the inflation–expectations gap remains when the central bank issues no additional guidance. The accepted assessment and the policy response already communicated by the rule do most of the work. The application therefore examines the risks of flexible targeting as a whole, including the role of guidance, rather than attributing the outcome to new announcements alone.

The remainder of the paper proceeds as follows. Section 2 relates the analysis to earlier work. Section 3 presents the institution and the recursive policy problem. Section 3.2 characterizes delivery, the communication frontier, and the pooling of announcements. Section 4 compares responses to the same supply disturbance under alternative public

readings and enforcement conditions. The appendix contains supplementary results, proofs, the numerical algorithm, data construction, and equilibrium checks.

2 Related Literature

The institution relates to the monetary delegation literature on incentive contracts (Walsh, 1995) and costly political override (Lohmann, 1992). The audit's gap component follows Drozd (2026): it removes a preference bias from the entire objective, including the announcement decision. A separate bounded sanction disciplines delivery while leaving policy a choice in both enforcement modes. This instrument draws on Halac and Yared (2022), who study compliance with fiscal rules under bounded, jointly costly penalties.[2]

The closest antecedent for the announcement game is King et al. (2008). They distinguish beliefs about the policymaker's type from the credibility of a particular inflation plan, obtain pooling at the strong type's preferred announcement using the refinement of Mailath et al. (1993), and show that more demanding promises can reduce expected compliance. Their static illustration also allows bounded random costs of deviation. We share their logic of imitation and equilibrium selection. Here announcements specify departures from a standing rule in response to a public forecast. Both enforcement modes choose delivery sequentially and differ only in sanction exposure. We establish conditions under which the high-enforcement mode's preferred pool is undefeated and preferred by both modes to separating equilibria.

Learning from noisy inflation connects the paper to Lu et al. (2016) and King and Lu (2026), who study committed policy when the public infers commitment capacity from observed outcomes. The alternative type follows a fixed rule in the former and optimizes myopically in the latter. Our framework makes delivery an endogenous choice in both modes under common stabilization objectives. Credibility measures the public's belief about the enforcement supporting those choices.[3]

Limits to imperfectly credible guidance are established in earlier work. Schaumburg and Tambalotti (2007) study commitment plans subject to exogenous opportunities for reoptimization. Haberis et al. (2019) study departures from an existing rule when the probability of early reversion increases with the temptation to renege, so extending lower-for-longer guidance need not provide additional stimulus. We characterize how much policy adjustment an announcement induces under a bounded sanction while the standing

[2] A separate approach restricts feasible policy through inflation caps (Athey et al., 2005) or history-dependent inflation bounds (Waki et al., 2018).

[3] King and Lu (2022) also study the distinction between explicit inflation targeting and imperfectly communicated policy intentions.

rule remains active. The direct enforcement limit exists even when the mode is known; learning affects expected delivery and the additional discipline supplied by credibility.[4]

Related papers design rules and announced paths under imperfect credibility. Kostadinov and Roldán (2025) show how gradual disinflation can strengthen reputational incentives to follow an announced path. Dovis and Kirpalani (2021) show that a more lenient rule can preserve uncertainty about commitment when initial reputation is low; their pooling concerns delivered policy, whereas ours concerns announcements whose delivery can differ across modes. Elkina (2026) studies the tradeoff between discretion to respond to private current conditions and reputation building. We take the standing rule and enforcement technology as given and study guidance chosen around them. This also distinguishes our question from Davig and Foerster (2023), who study communication of the systematic reaction function itself through targets, tolerance bands, and projections.

The information underlying guidance connects the paper to Bassetto (2019), who shows how private information about preferences or beliefs can give forward guidance value in a cheap-talk setting (see also Crawford and Sobel, 1982). Fujiwara and Waki (2022) show that disclosure of information about future fundamentals can reduce welfare. Gáti and Handlan (2025) study forecast communication when the public is uncertain about the precision of the central bank's information. We study delivery conditional on a public assessment. An announcement remains consequential even after the forecast is public because it changes the benchmark against which delivery is sanctioned. Learning about enforcement is endogenous; alternative public readings of forecasts and policy objectives are maintained counterfactual assumptions in the application.

The application relates to work on credibility and supply shocks. Nakamura et al. (2025) document that stronger inflation-fighting credibility helped central banks look through post-COVID inflation. Christoffel and Farkas (2025) quantify the de-anchoring risk of look-through policies, while Bocola et al. (2025) show how reputation building can induce a stronger response to supply shocks when reputation is low. Our application distinguishes skepticism about the forecast motivating guidance from skepticism about its enforcement and the central bank's objective.

[4] Batista et al. (2023) use sustainable plans to discipline deviations through the loss of future policy effectiveness. Our direct institutional sanction operates alongside the continuation consequences of Bayesian learning.

3 Theory

The central bank inherits a standing rule, can announce departures in advance, and chooses delivered policy subject to an audit. The rule and the announcement jointly define its obligation. Accountability is imperfect, so the obligation affects policy without mechanically determining it. We first specify the economy and the institution, then characterize delivery, announcements, and learning about enforcement. Each of the three policy risks considered in the paper gets its channel here: the audit's gap component fixes the objective the public may question; the delivery audit sets how far a promise can move policy, and the announcement game what an announcement reveals about whether it will be enforced; and the public assessment is the forecast the rule responds to, which the public may or may not accept. The framework builds on Drozd (2026); the assessment is public here, while enforcement is persistent and privately observed.

3.1 Environment

Time is discrete, $t = 0, 1, \ldots$. The economy consists of a private sector and a delegated monetary authority, which we call the central bank. The publicly observed exogenous state s_t contains the cost-push disturbance $u(s_t)$, the natural real rate $r^n(s_t)$, and other public conditions relevant for stabilization. Inflation π_t and the output gap x_t satisfy

$$\pi_t = \beta \widetilde{\mathbb{E}}_t \pi_{t+1} + \kappa x_t + u(s_t) + \varepsilon_t^{\pi}, \tag{1}$$

$$x_t = \widetilde{\mathbb{E}}_t x_{t+1} - \frac{1}{\sigma}[i_t - \widetilde{\mathbb{E}}_t \pi_{t+1} - r^n(s_t)], \tag{2}$$

where $\beta \in (0, 1)$ is the discount factor, $\kappa > 0$ the slope of the Phillips curve, $\sigma > 0$ the inverse intertemporal elasticity of substitution, and i_t the nominal interest rate. A tilde denotes conditioning on public information: $\widetilde{\mathbb{E}}_t$ includes the date-t announcement but not the policy delivered at t. The cost-push term is the reduced form of the inflation–output tradeoff that arises in multisector economies when the targeted price index differs from the divine-coincidence index (Woodford, 2003; Galí, 2015; Rubbo, 2024).

The central bank chooses the conditional inflation center θ_t; realized inflation adds control noise,

$$\pi_t = \theta_t + \varepsilon_t^{\pi}, \qquad \varepsilon_t^{\pi} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\pi^2), \qquad \sigma_\pi > 0. \tag{3}$$

Here σ_π is the standard deviation of control noise, which is independent of the economic state and the enforcement mode. Subtracting the noise from (1) gives the conditional-mean

Phillips curve

$$\theta_t = \beta \widetilde{\mathbb{E}}_t \pi_{t+1} + \kappa x_t + u(s_t), \tag{4}$$

so that, at a given center and cost-push disturbance, higher expected inflation requires a lower output gap. Society evaluates policy by the discounted sum of the standard quadratic stabilization loss. The inflation weight is normalized to one, and $\eta > 0$ is the weight on output-gap stabilization:

$$\ell_t = \frac{1}{2}(\theta_t^2 + \eta x_t^2). \tag{5}$$

Writing the loss in the center rather than in realized inflation drops the policy-independent constant $\sigma_\pi^2/2$. When the nominal rate is unrestricted, the rate supporting the chosen allocation is

$$i_t = r^n(s_t) + \widetilde{\mathbb{E}}_t \pi_{t+1} + \sigma(\widetilde{\mathbb{E}}_t x_{t+1} - x_t); \tag{6}$$

determinacy requires an appropriate off-path response, and at the effective lower bound the restriction on i_t adds an implementability constraint. We assume an unrestricted rate throughout.

In addition to s_t, a scalar assessment v_t of next period's conditions—information about the persistence or propagation of a disturbance—is public knowledge at date t: the central bank and the public see the same forecast. In the present specification v_t forecasts future cost-push pressure; its formation is not a choice in the model, and it may prove wrong.[5] All agents know the transition kernels

$$v_t \sim Q_v(\cdot \mid s_t, v_{t-1}), \qquad s_{t+1} \sim Q_s(\cdot \mid s_t, v_t), \tag{7}$$

The first kernel governs assessments; the second governs future public conditions conditional on the current assessment. These transitions are independent of the enforcement mode and policy choices. Observing the assessment therefore provides no additional information about enforcement. The risk that the public rejects the forecast is therefore separate from the risk that it second-guesses enforcement; the model treats them as separate.

[5]Central banks publish their projections; the model takes the assessment as public and does not study how it is formed or released. If it were the central bank's private information, the announcement would have to convey it as well; Appendix B.5 discusses that case, which the companion paper studies.

3.1.1 Preferences and the Audit's Two Components

The central bank may prefer a positive output gap $x^* > 0$, giving it the period loss $\ell_t^{CB} = \frac{1}{2}[\theta_t^2 + \eta(x_t - x^*)^2]$. Without a corrective institution, this preference creates a systematic incentive to inflate even when cost-push pressure is absent, and the incentive survives both the standing rule and an announcement the central bank keeps: a biased central bank announces the departure it wants and delivers it exactly, so a sanction on the delivery miss alone never fires. The mandate therefore gives the audit a second component, levied on the observed output gap,

$$T_t = \eta x^* x_t, \qquad \ell_t^{CB} + T_t = \ell_t + \frac{\eta}{2}(x^*)^2. \tag{8}$$

Because the central bank is risk neutral, this is the expected cost of a sanction whose probability is affine in the gap, scaled by the same stake Ψ as the miss component below: the two are components of one audit, not two instruments. A probability lies in $[0, 1]$, so the component is affine with a constant and holds over a bounded range of gaps; at the calibration of Section 4 the bound binds only beyond twenty points of output gap, far outside any path considered here, and the constant shifts the objective without touching a first-order condition.

The gap is the base that works. It nets out expectations, $x_t = [\pi_t - \beta e_t - u(s_t)]/\kappa$, so the component cancels $\eta x^*/\kappa$ in the conditions for the announcement and for delivery alike, and it needs no center to be measured from; a component on the inflation level would tax βe_t and distort the announcement. As in Drozd (2026), the correction therefore reaches the whole objective rather than the delivery first-order condition alone. Lemma 2 in Appendix G extends the argument to the recursive problem and shows why a component on inflation alone would leave announcement incentives distorted.

The component requires knowledge of x^*. It is borne in full by both modes: the exposure χ_j below scales the miss component alone, so the hidden mode changes what the central bank delivers and never what it wants. For implementation, the auditor can use

$$\widetilde{T}_t = \frac{\eta x^*}{\kappa}\big[\pi_t - \beta e_t - u(s_t)\big], \qquad e_t \equiv \widetilde{\mathbb{E}}_t \pi_{t+1}.$$

We assume the public expectation e_t is available to the auditor along with realized inflation and the public state. The mean Phillips curve and zero-mean control noise give $\mathbb{E}[\widetilde{T}_t \mid \theta_t, e_t, s_t] = \eta x^* x_t$. Risk neutrality makes the two implementations equivalent. The component supplies no additional information about the inflation center because it uses only the same noisy inflation observation and public variables.

We can consequently write the central bank's stabilization objective as (5), omitting the constant. This removes the risk about the objective: with the bias gone, the objective is the target, and the public has nothing there to second-guess. The correction does not commit future policy. Even with aligned stabilization objectives, a policy that improves expectations today may be costly to deliver tomorrow. The rule, guidance, and delivery audit address that remaining problem.

3.1.2 The Standing Rule

The standing regime combines a numerical long-run inflation target, normalized to zero, with a stationary rule for the inflation center written on *public conditions*,

$$\widetilde{S}_t \equiv (s_t,\, v_{t-1}), \tag{9}$$

the realized state and the assessment of it that was public one period earlier.[6] The rule is $\bar{\theta}(\widetilde{S}_t)$: strict inflation targeting is $\bar{\theta} \equiv 0$, and a flexible, publicly understood regime is any other function of public conditions. The rule can depend only on $\widetilde{S}_t$: it does not respond to the current assessment v_t, which concerns next period's conditions, to credibility, or to the central bank's own announcements, defined next. The inherited assessment lets a standing rule distinguish anticipated pressure from a forecast error. This distinction matters because expected future accommodation affects current inflation before the pressure arrives. For given expected inflation, the static response to an additional unit of current cost-push pressure is $D \equiv \eta/(\eta + \kappa^2)$, derived in Section 3.1.5. The leading example below assigns this response to forecast errors and a smaller response to anticipated pressure. Forward guidance allows the central bank to adjust the resulting policy benchmark. The rule is imposed; its coefficients are not the solution to an optimal rule-design problem.

Leading example. In the application of Section 4 the assessment is a forecast of next period's disturbance, $v_t = \widehat{u}_{t+1|t}$, and the rule is linear in the inherited forecast and the surprise,

$$\bar{\theta}(\widetilde{S}_t) = \phi_{IT}\, \widehat{u}_{t|t-1} + D\,(u_t - \widehat{u}_{t|t-1}), \qquad 0 \le \phi_{IT} < D, \tag{10}$$

For a correctly forecast AR(1) disturbance, $u_t = \rho u_{t-1} + \varepsilon_t$, the rule responds to the innovation at D and to the anticipated component at ϕ_{IT}.

Figure 1 anchors this rule to commitment policy. For a unit innovation with $\rho = 0.70$, correctly forecast, it plots the inflation center under the rule and under the exact commitment solution of Appendix B.1, $\theta_t = \psi u_t - (1-\delta)p_{t-1}$, where p_{t-1} is the accumulated

[6] Davig and Foerster (2023) give an institutional account of announced policy rules of this kind.

price-level gap and $0 < \delta < 1$ (Clarida et al., 1999; Woodford, 2003). Both accommodate the innovation on impact and then lean against its foreseen continuation; that shared shape is what the forecast loading $\phi_{IT} < D$ captures. They part afterwards. Commitment's response depends on past policy, turns negative as the pressure fades, and restores the price-level path; the rule holds a constant fraction ϕ_{IT} of foreseen pressure and never reverses. The standing rule is thus a coarse, history-free version of commitment. It is imposed rather than designed, and its forecast loading is an institutional restriction, not a representation of Ramsey policy. Forward guidance is the discretionary departure from it.

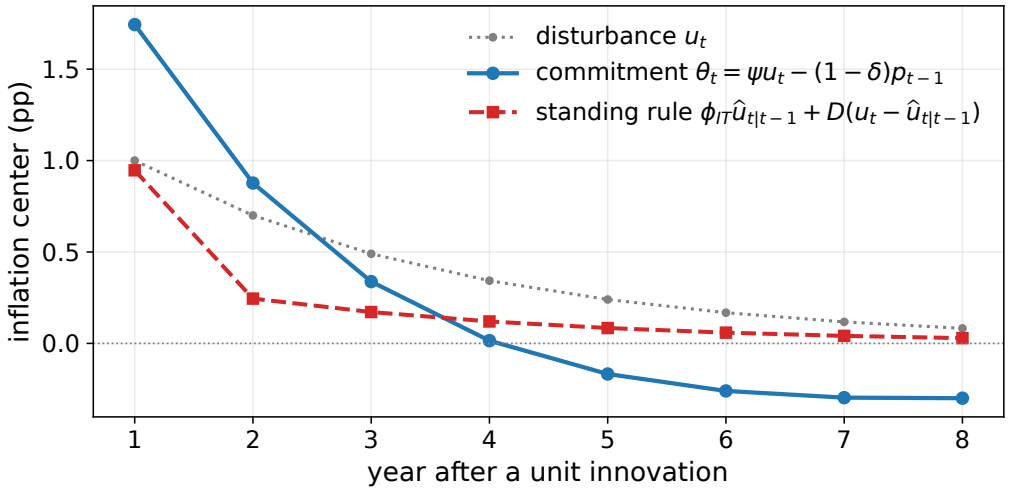

Figure 1: Standing rule and commitment after a unit innovation

Notes: Inflation center after a unit innovation to $u_t = \rho u_{t-1} + \varepsilon_t$ with $\rho = 0.70$, at the calibration of Section 4, with a correct forecast and no enforcement or control noise. Commitment is the solution of Appendix B.1 with $\delta = 0.803$ and $\psi = 1.744$; the standing rule is (10) with $\phi_{IT} = 0.35$. Table 7 reports the same paths per unit of current pressure, with discretion and static-share delivery.

3.1.3 Policy Obligations and Accountability

Forward guidance is an announced departure from the standing rule. The subscript indexes the date at which the promise comes due: Δ_t^p is announced at $t-1$, while Δ_{t+1}^p is chosen at t. The policy obligation due at t is

$$a_t \equiv \bar{\theta}(\widetilde{S}_t) + \Delta_t^p. \tag{11}$$

The central bank cannot rewrite the standing rule within the period. It can adjust its future obligation through guidance and choose a current deviation subject to the same institutional sanction that enforces the rule. Setting $\Delta_{t+1}^p = 0$ leaves the standing rule as next

period's obligation; we call this choice *no additional guidance*. Policy is still communicated through the rule.

The delivered departure and the miss are

$$\Delta_t = \theta_t - \bar{\theta}(\widetilde{S}_t), \tag{12}$$

$$d_t = \Delta_t - \Delta_t^p = \theta_t - a_t. \tag{13}$$

Thus realized inflation is $\pi_t = a_t + d_t + \varepsilon_t^\pi$. An announced departure changes the obligation; a delivery miss departs from it. These are the institution's two forms of flexibility, chosen at different dates.

The public observes inflation and audits

$$\pi_t - \bar{\theta}(\widetilde{S}_t) - \Delta_t^p = d_t + \varepsilon_t^\pi, \tag{14}$$

the miss up to control noise once the rule's own response to public conditions is removed, an application of the informativeness principle of Holmström (1979).[7] A miss can have either sign: when the announcement comes due, the central bank may prefer more accommodation than it announced, or less.

A miss has two consequences: a direct institutional cost and a change in what the public believes about enforcement. The audit applies to all misses, including deviations from the standing rule when $\Delta_t^p = 0$. The central bank operates in one of two persistent hidden modes, $j_t \in \{H, L\}$. Mode H, high enforcement, is fully exposed to a common institutional sanction; mode L, low enforcement, is exposed only to a fraction of it. The exposure χ_j scales the sanction,

$$\chi_H = 1 > \chi_L > 0. \tag{15}$$

Both modes share the same target, stabilization objective after the audit's gap component, information, and instruments; they differ only in how strongly governance, legal or career accountability, and public censure translate a miss into a contemporaneous institutional cost. Let $\Psi > 0$ be the largest expected sanction at full exposure and $\tau > 0$ the audit tolerance.[8] The sanction is imposed with probability $1 - \exp\{-[\pi_t - \bar{\theta}(\widetilde{S}_t) - \Delta_t^p]^2/(2\tau^2)\}$;

[7] The public has no separate noiseless observation of θ_t through instruments or allocations: (6) is an implementation relation. If observed instruments revealed the center, that information would enter the likelihoods below and change the learning problem.

[8] A bounded payoff sanction, rather than a restriction of the action set, follows Walsh (1995) in form and the limited-enforcement setting of Halac and Yared (2022), in which the policymaker retains full discretion and penalties are joint and bounded.

integrating over control noise, the expected sanction for a miss d in mode j is

$$K_j(d) = \chi_j \Psi \left[1 - \frac{\tau}{\sqrt{\tau^2 + \sigma_\pi^2}} \exp\left\{ -\frac{d^2}{2(\tau^2 + \sigma_\pi^2)} \right\} \right]. \tag{16}$$

The sanction is symmetric in the miss, rises with its size, and flattens at $\chi_j \Psi$. Its value at $d = 0$ is a false-alarm cost that control noise imposes even on exact delivery; incentives depend on how the sanction changes with the miss, and because the sanction is bounded, its marginal value rises and then falls (Figure 10). The largest marginal sanction, $\sup_d |K_j'(d)|$, is what limits how far an announcement can move delivered policy.

The mode follows a Markov chain. With λ_t the public's probability of low enforcement at the start of period t, the transition probabilities $P_{jk} = \Pr(j_{t+1} = k \mid j_t = j)$ are

$$P_{HL} = (1 - \rho_\lambda)\bar{\lambda}, \qquad P_{LL} = (1 - \rho_\lambda)\bar{\lambda} + \rho_\lambda, \tag{17}$$

where $0 \leq \rho_\lambda < 1$ is the persistence of the mode and $0 < \bar{\lambda} < 1$ the ergodic share of low enforcement; the remaining probabilities follow from the row sums. We call the public's belief about this mode *enforcement credibility*, or simply credibility within the theoretical model. It concerns compliance with the policy obligation, not the accuracy of the public assessment or the stability of policy preferences. Beliefs move in two steps within the period. An announcement p is read with probability $\lambda_t(p)$ of low enforcement, which Bayes' rule pins down on the equilibrium path. Realized inflation then updates that reading. Conditional on the announcement and the state, mode j places the inflation center at $\bar{\theta}(\widetilde{S}_t) + \Delta^*_{t,j}$, so the two modes generate normal inflation distributions with different means and the common variance σ_π^2, and the posterior probability of low enforcement, q_t, satisfies

$$\frac{q_t}{1 - q_t} = \frac{\lambda_t(p)}{1 - \lambda_t(p)} \frac{\mathcal{L}_L^\pi(\pi_t \mid p, \Omega_t)}{\mathcal{L}_H^\pi(\pi_t \mid p, \Omega_t)}, \tag{18}$$

$$\lambda_{t+1} = (1 - \rho_\lambda)\bar{\lambda} + \rho_\lambda q_t, \tag{19}$$

where $\mathcal{L}_j^\pi$ is the inflation density generated by mode j's delivered policy and Ω_t is the state defined below. Higher realized inflation favors the mode whose center is higher: the low-enforcement mode when the announcement restrains the central bank relative to the departure it prefers on current grounds (its static preference, defined in Section 3.1.5), and the high-enforcement mode when the announcement commits it to more accommodation than it wants ex post. Equation (19) carries the posterior into the next period's prior,

allowing for a change of mode. In the equilibrium of Section 3.2 both modes issue the same announcement, so $\lambda_t(p) = \lambda_t$ on the path: the public reads the announcement at λ_t, and only realized inflation moves the belief.

3.1.4 Timing

The state at the start of the period is public, $\Omega_t = (\widetilde{S}_t, v_t, \lambda_t, \Delta_t^p)$: public conditions, the assessment of next period's conditions, credibility, and the announcement due. The central bank also knows its enforcement mode, carried as a subscript j_t; it is the central bank's only private information. Table 1 gives the order of events within the period. The new announcement moves expected inflation before the announcement due is delivered, so the announcement and the delivery are separate decisions: the new announcement changes current activity, and current delivery takes the resulting expectation as given.

Table 1: Timing within period t

Stage	Event	Information and decision
1	State	The public and the central bank observe Ω_t; the central bank also knows its enforcement mode j_t.
2	Guidance	The central bank announces Δ_{t+1}^p, the departure for the following period, or issues no additional guidance.
3	Expected inflation	The public reads the announcement with probability $\lambda_t(\Delta_{t+1}^p)$ of low enforcement and forms $e_t = \widetilde{\mathbb{E}}_t \pi_{t+1}$.
4	Delivered policy	Taking Δ_{t+1}^p and e_t as given, the central bank chooses Δ_t, delivering the announcement due, Δ_t^p.
5	Inflation and policy audit	Inflation π_t is realized; the public observes the audit variable (14), and the policy-miss sanction is assessed.
6	Transition	The public updates q_t from inflation, the mode transitions, λ_{t+1} follows from (19), Δ_{t+1}^p becomes due, and next period's conditions s_{t+1} and assessment v_{t+1} arrive.

Announcement game. The central bank's only private information is its enforcement mode j_t. It signals through Δ_{t+1}^p, and the public forms the expectation e_t that enters the current Phillips curve. A *reading* is the belief $\lambda_t(\Delta_{t+1}^p)$ attached to an announcement. In equilibrium each mode chooses an optimal announcement given these readings, and on-path readings satisfy Bayes' rule. Different announcements reveal the mode in a separating equilibrium; a common announcement leaves the belief at the prior in a pooling equilibrium.

Off-path readings must also deter deviations. Section 3.2.2 gives conditions for pooling at the high-enforcement mode's preferred announcement and assesses that equilibrium

using the undefeated criterion of Mailath et al. (1993). The criterion compares the proposed equilibrium with alternative equilibria that could explain an unexpected announcement.[9]

3.1.5 Delegated Policy Problem

The central bank's problem is recursive: each period it inherits an announcement due and a level of credibility, issues new guidance, and chooses current delivery. The application of Section 4 solves this recursion year by year. The results of Section 3.2 isolate one announcement and summarize all later decisions by a continuation value; that episode is one step of the recursion.

At the delivery stage e_t has already been formed and used in private decisions. Substituting the mean Phillips curve (4) into (5) gives the current loss as a function of the delivered departure,

$$\ell_t(\Delta_t) = \frac{1}{2}\left\{[\bar{\theta}(\widetilde{S}_t) + \Delta_t]^2 + \eta\,\frac{[\bar{\theta}(\widetilde{S}_t) + \Delta_t - \beta e_t - u(s_t)]^2}{\kappa^2}\right\}, \tag{20}$$

a quadratic in Δ_t with curvature $A \equiv 1 + \eta/\kappa^2$. The departure the central bank prefers on current grounds alone, its static preference, is the minimizer

$$\Delta_t^d = D\,[u(s_t) + \beta e_t] - \bar{\theta}(\widetilde{S}_t), \qquad D \equiv \frac{\eta}{\eta + \kappa^2}, \tag{21}$$

so the central bank accommodates the share D of cost pressure and expected inflation, a smaller share the more it cares about inflation relative to activity. Completing the square,

$$\ell_t(\Delta_t) = \ell_t(\Delta_t^d) + \frac{A}{2}(\Delta_t - \Delta_t^d)^2. \tag{22}$$

The stabilization cost of delivering a departure is therefore quadratic in its distance from the static preference.

Let $\mathcal{V}_j(\Omega_t)$ be the central bank's expected loss in mode j before its announcement, and $\mathcal{V}_j(\Omega_t \mid \Delta_{t+1}^p)$ the same loss once the new announcement has been issued, excluding the communication cost defined below. The discounted continuation loss after a miss d is

$$\mathcal{R}_j(d \mid \Omega_t, \Delta_{t+1}^p) = \beta\,\mathbb{E}_j\left[\mathcal{V}_{j_{t+1}}(\Omega_{t+1}) \mid \Omega_t, \Delta_{t+1}^p, \Delta_t = \Delta_t^p + d\right], \tag{23}$$

[9] Appendix C.5 gives its definition and implications.

which integrates over control noise, the audit, the posterior (19), the mode transition, and the next public state, with all future decisions optimal. The recursion is

$$\mathcal{V}_j(\Omega_t \mid \Delta^p_{t+1}) = \min_{\Delta_t} \left\{ \ell_t(\Delta_t) + K_j(\Delta_t - \Delta^p_t) + \mathcal{R}_j(\Delta_t - \Delta^p_t \mid \Omega_t, \Delta^p_{t+1}) \right\}, \tag{24}$$

$$\mathcal{V}_j(\Omega_t) = \min \left\{ \mathcal{V}_j(\Omega_t \mid 0),\ \gamma + \inf_{\Delta^p_{t+1} \neq 0} \mathcal{V}_j(\Omega_t \mid \Delta^p_{t+1}) \right\}. \tag{25}$$

The first line determines delivery. The central bank weighs the stabilization benefit of departing from its obligation against the sanction and the effect on future credibility. The second line determines next period's obligation. Leaving it at the standing rule incurs no issuance cost. A nonzero departure costs $\gamma > 0$, representing the institutional cost of formulating and authorizing an additional commitment. It changes current expectations and creates a future obligation. When its gross benefit is smaller than γ, the central bank relies on the standing rule. Current policy nevertheless remains a choice subject to the audit.

3.1.6 Equilibrium

The public expects from an announcement the policy it believes will be delivered. Let $\mathbb{E}_j[\Delta^*_{t+1} \mid \Delta^p_{t+1}]$ denote the expected next-period departure conditional on the current mode being j; it integrates over next period's state, the transition from j to the next mode, and optimal delivery by that mode. Expected inflation is

$$e_t = \widetilde{\mathbb{E}}_t[\bar{\theta}(\widetilde{S}_{t+1})] + \left(1 - \lambda_t(\Delta^p_{t+1})\right) \mathbb{E}_H[\Delta^*_{t+1} \mid \Delta^p_{t+1}] + \lambda_t(\Delta^p_{t+1})\, \mathbb{E}_L[\Delta^*_{t+1} \mid \Delta^p_{t+1}], \tag{26}$$

the expected rule plus the departure each mode would deliver, weighted by the public's reading of the announcement. We study a pool at the announcement preferred by the high-enforcement mode, read at the prior λ_t. Section 3.2 establishes conditions under which both modes choose this announcement.

Definition 1 (Pooling equilibrium)**.** *A pooling equilibrium consists of an announcement policy $\Delta^p_{t+1}(\Omega_t)$ common to both modes, delivery policies, expectations, beliefs, and value functions such that: (i) the announcement solves (25) for mode H when the public reads it at $\lambda_t(\Delta^p_{t+1}) = \lambda_t$ and forms e_t through (26); (ii) delivery solves (24) in each mode; (iii) the public reads any other announcement as the low-enforcement mode's, $\lambda_t(p) = 1$ for $p \neq \Delta^p_{t+1}(\Omega_t)$, and neither mode gains from such an announcement; (iv) beliefs after inflation follow (18)–(19); and (v) an interest-rate rule delivers the chosen allocation.*

Condition (iii) requires both modes to prefer the proposed announcement to every deviation. Proposition 2 states this requirement. Appendix A.3 gives the additional comparisons with separating equilibria and the equilibrium-selection conditions. The application evaluates the relevant inequalities along its reported paths.

3.2 Equilibrium Properties

The audit's gap component has removed the systematic inflation bias. The remaining problem is to sustain a policy obligation when future stabilization incentives may favor a different response. The next two subsections are the enforcement channel: how far an obligation can move delivered policy, and what announcements reveal about whether it will be enforced. The last is the forecast channel: what changes when the public reads the economy differently from the central bank, with the institution unchanged.

3.2.1 Communication Frontier—Laffer Curve of Communication

At delivery, the stabilization cost of moving away from the static preference rises quadratically with the distance. The sanction for missing the obligation is bounded, and so is the discipline that future credibility can add, because noisy inflation moves the public's belief only gradually. Against a cost that keeps rising, a bounded penalty can hold policy only within a limited distance of what the central bank would otherwise choose.

To state this formally we follow one announcement through a policy episode. It is issued in Period 1; it comes due in Period 2, when the state s_2 and with it the static preference Δ_2^d are realized; and its consequences for credibility arrive in Period 3. Later decisions are summarized by the continuation loss $\mathcal{R}_j$, whose marginal effect on the cost of a miss is bounded by a constant ρ_1. Two regularity conditions are maintained: the value of credibility is bounded, and the curvature A of the stabilization loss dominates the concave part of the sanction together with the curvature that credibility can add (Appendix A.1).

Proposition 1 (Communication frontier)**.** *At a given delivery state and expected inflation, mode j chooses a unique departure, and it satisfies*

$$|\Delta_2^* - \Delta_2^d| \leq \frac{\chi_j \Psi \tau}{A\sqrt{\mathrm{e}}\,(\tau^2 + \sigma_\pi^2)} + \frac{\rho_1}{A}.$$

The first term is the largest adjustment the direct sanction alone can induce.

The bound is the *communication frontier*. It measures distance from the policy the central bank would choose on current grounds, not from the standing rule: a large announced

departure is easy to deliver if it is close to that static preference, and a small one is costly if conditions have moved. Stronger actual enforcement—a higher stake Ψ or exposure χ_j—widens the direct limit. At the calibration it is 0.264 percentage points for mode H and 0.066 for mode L, and continuation incentives can add at most ρ_1/A.

The frontier binds the rule as much as guidance. When no additional guidance is issued, $\Delta_2^p = 0$, the obligation is the rule's own center, enforced with the same limited force. With the direct sanction alone, delivered policy then lies between the obligation a_2 and the static preference, and close to the point where the two coincide,

$$\theta_2^* - a_2 = \frac{A}{A + K_j''(0)}(\theta_2^d - a_2), \qquad \theta_2^d \equiv \bar{\theta}(\widetilde{S}_2) + \Delta_2^d,$$

up to higher-order terms. Stronger enforcement reduces a small delivery miss without removing the incentive to accommodate.

How much of an announcement is delivered depends on where the miss falls on the sanction. The sanction is convex on the *responsive range* $|d| < \sqrt{\tau^2 + \sigma_\pi^2}$, where a larger miss raises the marginal sanction; beyond it the marginal sanction falls, so a larger miss adds less and less to the cost of nonperformance. Near a zero miss the fraction of an additional announcement that is delivered is $\omega_j = K_j''(0)/[A + K_j''(0)]$, the curvature of the sanction at a zero miss relative to that curvature and the curvature of the stabilization loss together. This is the *traction* of communication; at the calibration $\omega_H = 0.22$ and $\omega_L = 0.07$. The gap between the two modes' tractions is at most $(1 - \sqrt{\chi_L})/(1 + \sqrt{\chi_L}) = 0.33$ at any stake, attained at $\Psi = 128.5$ (Corollary 3); at the calibration it is 0.15. The stake raises both modes' traction; exposure alone sets how far apart they can be. Traction is state dependent. The slope of the delivery map is the traction at each announcement: it falls from ω_j at a zero miss to zero at the frontier and turns negative beyond it, a Laffer curve of communication. Where a given announcement sits on that curve depends on the static preference, which moves with conditions and with expected inflation, so the same words buy different amounts of policy in different states. Traction also depends on credibility. The public weights the two modes' responses, $(1 - \lambda_1)\omega_H + \lambda_1\omega_L$, so skepticism about enforcement lowers the effect of the same announcement on expectations and moves the public's frontier toward mode L's. Actual and perceived enforcement therefore jointly determine what an announcement is worth.

Figure 2 traces the whole map. Inside the responsive range delivery rises with the announcement; at the point where the marginal sanction peaks it reaches the frontier; beyond it, delivery falls back toward the static preference and more ambitious guidance is abandoned. Mode H reaches its maximum at an announcement of 1.82 percentage

points, and mode L has the same shape at a quarter of the height. The public expects the credibility-weighted mixture of the two: at the reference state, at most 0.240 points when $\lambda = 0.12$ and 0.145 when $\lambda = 0.60$. The circles are the recursive solutions with continuation incentives and lie on the direct-sanction curves. The full delivery map, the bounds on continuation effects, and the sensitivity of ambitious guidance near the frontier are in Appendices A.2 and B.3.

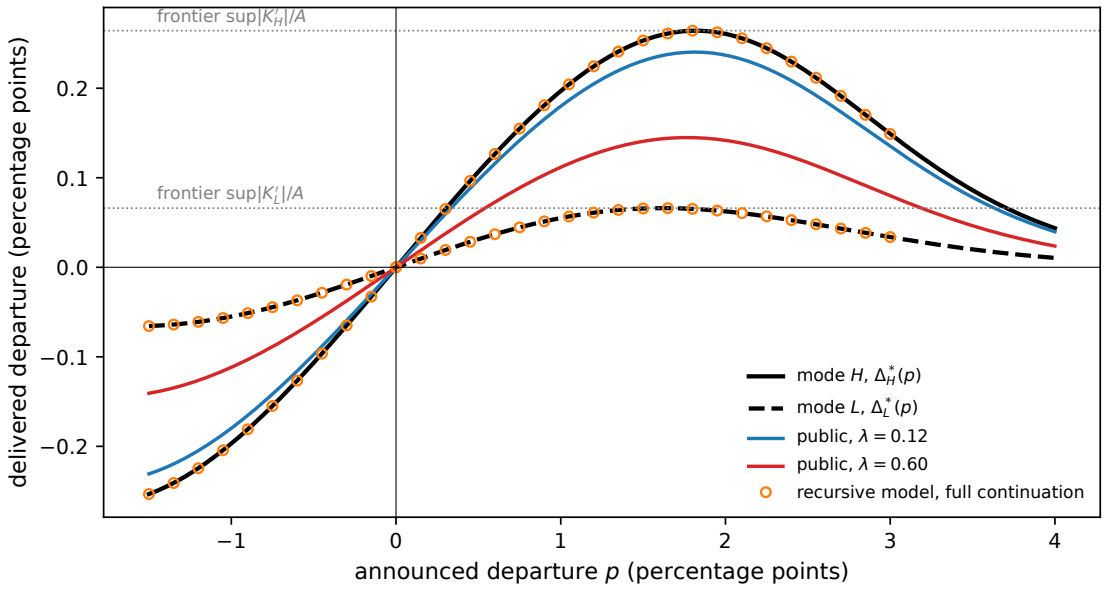

Figure 2: Announced and delivered departures

Notes: Direct-sanction delivery maps at a static preference of zero and the application calibration. The public weights the two modes by the probability λ of low enforcement. Circles are recursive solutions including continuation incentives at the reference state ($\lambda = 0.12$), with no further guidance. Dotted lines mark the direct-sanction frontiers.

3.2.2 Pooling and Credibility

The modes can announce the same policy while delivering different amounts of it. Imitation is inexpensive for mode L, because weaker enforcement lets it depart further from the promise; whether it imitates depends on what being believed is worth, in current expectations and in future credibility.

Let $U_j(p, \lambda)$ be mode j's expected episode loss from announcement p, including delivery, continuation, and issuance costs, when the public assigns probability λ to low enforcement. Let p^* minimize $U_H(\cdot, \lambda_1)$.

Proposition 2 (Pooling of announcements)**.** *Both modes announcing p^*, read at the prior λ_1, constitutes an equilibrium with all other announcements attributed to mode L if and only if*

$$U_j(p^*, \lambda_1) \leq U_j(m, 1) \quad \textit{for every } m \neq p^*, \qquad j = H, L. \tag{27}$$

In this equilibrium, announcements do not reveal enforcement. The public updates its belief from realized inflation.

The condition says that neither mode gains from an announcement the public reads as weak enforcement, and the two modes face it differently. Mode H obtains its preferred announcement under the pooled reading. Mode L accepts an announcement it would not otherwise choose in exchange for that reading, so for it the condition requires the reading to be worth the cost of imitation.

In the application, guidance restrains the central bank relative to its static preference. Being read as high enforcement therefore lowers expected inflation, while mode L can imitate without delivering all of the restraint. At sufficiently large announcements the reading stops being worth the imitation, and separation would require mode H to announce there. In the 2021 state the pooled optimum is 0.47, and the smallest announcement above it that mode L would not imitate is 0.80; mode H prefers the pool. Further comparisons show the pool preferred by both modes to separation and undefeated in the sense of Mailath et al. (1993); they are stated in Appendix A.3. The application uses the pool preferred by mode H and checks the inequalities at each date.[10]

Learning from delivery. In the pool the announcement leaves the belief at λ_1, so whatever the public learns about enforcement it learns from delivered policy. The modes deliver different centers, and after inflation π_2 the posterior odds of low enforcement are

$$\frac{q_2}{1-q_2} = \frac{\lambda_1}{1-\lambda_1} \exp\left\{ \frac{\Delta_L^* - \Delta_H^*}{\sigma_\pi^2} \left[\pi_2 - \bar{\theta}(\widetilde{S}_2) - \frac{\Delta_H^* + \Delta_L^*}{2} \right] \right\}, \tag{28}$$

before the mode transition. The audit strips out the rule's response to observed conditions first, so a rule-prescribed response to a supply surprise is not evidence of weak enforcement. The direction of the update depends on which mode delivers the higher center: at the reference state in Figure 3, mode H delivers 0.20 and mode L 0.06, so there higher inflation favors mode H. The speed of the update depends on how far apart the centers are relative to control noise. At the reference state a year of inflation moves the belief by

[10] Appendix F reports the incentive margins and the other equilibria retained by the numerical screen. The conditions do not establish a unique pooling equilibrium.

a few hundredths, and along the application paths the marginal continuation effect of a miss never exceeds 1×10^{-2}, against a direct marginal sanction of 1.62 at the 2021 delivery: learning adds little to the delivery incentive in this calibration. If the centers coincide, inflation is uninformative and the belief moves only through the exogenous transition; low inflation alone does not build credibility. Indeed, when cost-push pressure is absent, the rule prescribes zero, and no promise is inherited, both modes sustain a zero center without additional guidance (Appendix G.1).

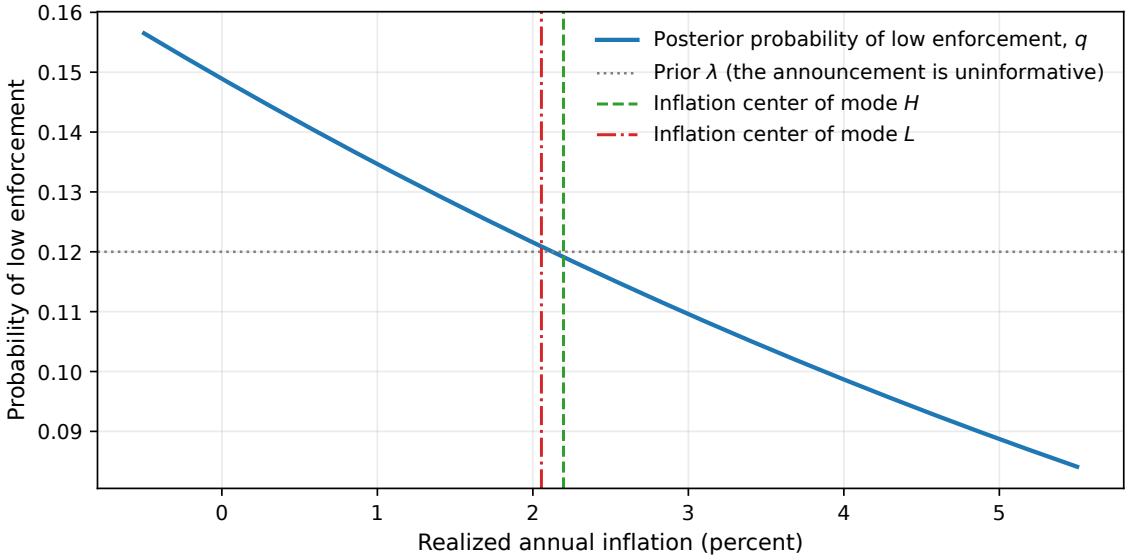

Figure 3: Posterior probability of low enforcement

Notes: The posterior probability of low enforcement as realized inflation varies, at the reference state of the theory exhibits (no cost-push pressure, $\lambda = 0.12$) with an announced departure of 1.0 due and no new announcement. The horizontal line is the belief after the announcement, which equals the prior because both modes announce the same departure; the vertical lines are the two modes' inflation centers, from the full delivery problem in the recursion.

3.2.3 Policy Readings and Fragility

A public forecast influences policy before the forecasted conditions arrive. If the public expects supply pressure to fade, it anticipates less accommodation from the rule and from subsequent delivery. If it instead expects persistence, expected inflation rises. Equation (21) shows the consequence: at delivery, an increase in e_t raises the preferred inflation center by βD per unit. Holding inflation down then requires a larger contraction. Finite accountability limits how far the central bank can resist this change in the stabilization tradeoff. The anchor is fragile in this sense: the same announcements, delivered by the same institution, support different expectations depending on how the public reads them.

The public can thus question three different things. Skepticism about *enforcement* changes the weight placed on the modes' deliveries. Skepticism about the *forecast* changes the economic conditions under which those deliveries are expected. Skepticism about *objectives* treats accommodation as a lasting policy preference rather than a response to a temporary disturbance; the mandate fixes this third one, since the target and the audit's gap component leave no lasting preference for accommodation to question. Only enforcement is privately observed and learned in the theoretical model. The application imposes alternative beliefs about enforcement and about the persistence of the disturbance to examine the risks of policy being understood differently from how the central bank intends it.

4 Application: Inflation Risk after Look-Through Policies

The post-COVID look-through is the test case. The Federal Reserve accommodated pressure it expected to fade, and the public went along: Figure 4 shows a large rise in realized inflation with a much smaller rise in one-year expectations. This section asks which of the three risks that outcome was exposed to, and finds that it was the forecast, not enforcement.

The central bank repeatedly expects supply pressure to fade quickly, but the disturbance proves persistent. Sargent and Williams (2026) rationalize the Federal Reserve's rate decisions of the 2020s with the same assessment: a central bank estimating a drifting Phillips curve whose persistence had fallen before the pandemic reads the surge as transitory. The baseline public accepts that assessment. Their account has no public deciding whether to accept it; this one does, and the alternative readings, which retain the same realized supply path but question enforcement or the persistence of the disturbance, measure the risk of acceptance failing.

The exercise has three steps. We first explain how the targeting institution accommodates repeated surprises while expectations remain comparatively low, separating the rule's contribution from additional guidance. We then hold announcements fixed and vary the public's reading, allowing delivery to adjust. Finally, we let the central bank reoptimize its announcements and vary actual sanction strength. A separate benchmark asks what would have happened had the forecast proved correct.

The computation is annual. From 2021 on, it solves the recursion in Section 3.1.5, using the pool characterized by Proposition 2. The inherited announcement and the public's belief carry across years. At each date the central bank evaluates continuation along the forecast path, integrating over control noise and enforcement transitions; the

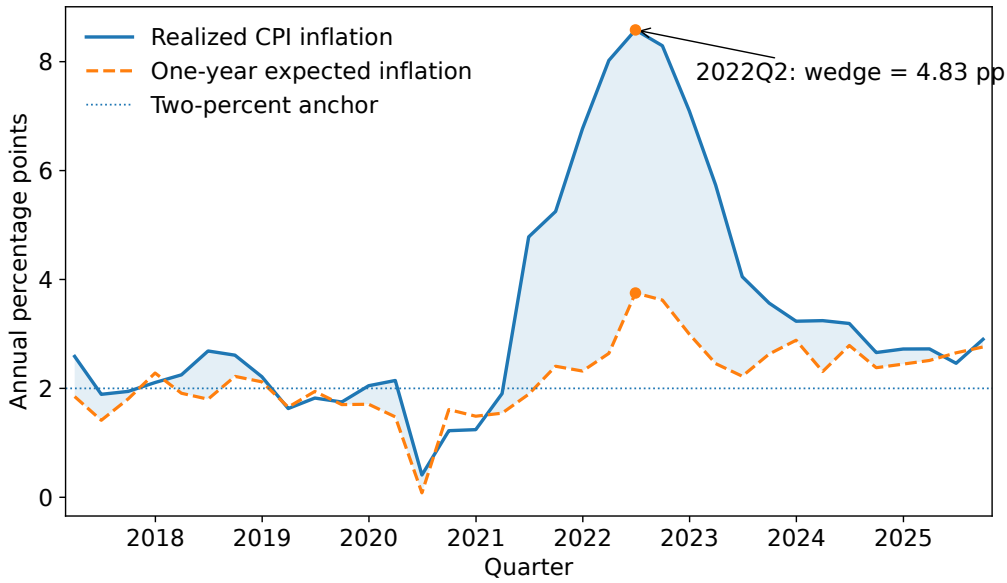

Figure 4: Inflation and one-year expectations

Notes: Quarterly averages of year-over-year CPI inflation and the Federal Reserve Bank of Cleveland's model-based one-year expected inflation. Realized inflation peaked near 9 percent while one-year expectations remained below 4 percent. The two-percent line is a target reference. Appendix E describes the series.

problem is re-solved when the next realized disturbance arrives. This deterministic-forecast approximation does not integrate over future supply innovations. Under forecast skepticism, the public instead evaluates future policy along a more persistent path. The 2020 allocation and initial promise are imposed. Appendices D and F describe the solution and its numerical checks.

4.1 Calibration

The disturbance peaks in 2021 and then decays at $\rho_u = 0.70$. Each year, the public assessment instead assigns persistence $\rho^{CB} = 0.29$, so $\widehat{u}_{t+1|t} = \rho^{CB} u_t$. The baseline public accepts this forecast. The policy obligation for the following year is

$$\bar{\theta}(\widetilde{S}_{t+1}) + \Delta^p_{t+1} = \phi_{IT} \widehat{u}_{t+1|t} + \Delta^p_{t+1} + D(u_{t+1} - \widehat{u}_{t+1|t}). \tag{29}$$

The forecast made at t enters next period's public conditions, $\widetilde{S}_{t+1} = (s_{t+1}, v_t)$. The audit subtracts the full state-contingent center in (29). It therefore treats the rule's response to a forecast error as prescribed policy, rather than as a delivery miss.

Rule parameterization. The forecast loading $\phi_{IT} = 0.35$ is imposed, not estimated. It prescribes some accommodation of anticipated pressure. Further accommodation can be announced in advance or chosen at delivery, subject to a sanction for the latter. At $\phi_{IT} = 0$, anticipated pressure receives no accommodation from the rule, but surprises still receive the response D. At $\phi_{IT} = D$, the rule is Du_t. This is the static response with zero expected future inflation; full discretion also responds to expected inflation through (21).

The difference $(D - \phi_{IT})\widehat{u}_{t+1|t}$ measures the departure needed to bring the forecast component to the static share D. It is a useful reference for the announcement, not its optimality condition: the central bank also considers expected inflation, sanctions, and continuation values. For 2022 the difference is 0.50, while the chosen announcement is 0.47. Appendix B.1 compares the imposed rule with commitment under a correctly forecast shock. The rule captures some restraint on anticipated pressure but omits commitment's history dependence and eventual reversal. The application evaluates guidance around this rule; it does not establish that the rule is optimal.

Table 2 reports the parameters. The model assumes a moderately sloped Phillips curve, $\kappa = 0.15$, and an output-gap weight $\eta = 0.40$ from the production calibration of Drozd and Tavares (2026); the enforcement block is that of the theory exhibits, with $\Psi = 18$ in the units of (5). The enforcement, belief, and interpretation parameters specify illustrative scenarios. The intertemporal-elasticity parameter σ affects the supporting interest rate but not these unrestricted-rate allocations.

The forecast persistence $\rho^{CB} = 0.29$ is motivated by the December 2021 Summary of Economic Projections. Its median core-PCE projection falls from 4.4 percent in 2021 to 2.7 percent in 2022, so the ratio of deviations from the two-percent target is $(2.7-2)/(4.4-2) \simeq 0.29$ (Federal Open Market Committee, 2021). This inflation-forecast ratio is a calibration proxy for the assessment's persistence, not an estimate of cost-push persistence; it is consistent with the real-time estimates in Sargent and Williams (2026).

The 2021 cost-push peak, $u_{2021} = 2.86$, is set so that 2021 inflation matches the December-over-December core CPI inflation rate of 5.5 percent (Bureau of Labor Statistics, 2022); later observations and the expectations series are untargeted.

As a descriptive check, the slope of four-quarter changes in one-year expected inflation on four-quarter changes in realized inflation is 0.347 for core CPI (HAC standard error 0.064) and 0.244 for headline CPI (standard error 0.031), over 2018Q1–2024Q4. These moments document the limited response of expectations. They do not identify the sanction scale, and neither does the episode: the results that carry the paper hold at every stake (Proposition 1, Proposition 2, Corollary 3), and the loss moves by less than one percent between the two stakes compared in Section 4.4. Appendix E describes their construction.

Activity is measured by the employment shortfall: the prime-age employment-to-population ratio against its 2019Q4 level, fourth quarter over fourth quarter, read one for one as the output gap. The prime-age restriction removes the demographic movements in the headline ratio, and the series is a common proxy for slack (Abraham and Kearney, 2020). Output growth on the change in this series over 1990–2019 gives 0.82 (standard error 0.21), and in the model, where output is labor input, the mapping is one by construction.

The pandemic year 2020 is an initial condition. Inflation is set at the 2020 core CPI rate of 1.6 percent and the output gap at −4.0, and the central bank carries into 2021 an announced departure of $\Delta^{p}_{2021} = 0.50$, representing the September 2020 guidance that policy would tolerate inflation moderately above target for some time (Federal Open Market Committee, 2020; Ihrig and Waller, 2024). Its magnitude is a calibration choice motivated by the half-point core-PCE shortfall of 2020; the online appendix repeats the exercise at 0.25 and 1.00 with the 2021 shock recalibrated, and the comparisons below keep their ordering. The gap is set shallower than the measured shortfall, −5.28, because the 2020 collapse was substantially temporary layoff; the value enters no result.

4.2 Immaculate Post-COVID Disinflation

The baseline illustrates accommodation within a continuing targeting institution. The public expects the disturbance to fade and the rule to prescribe less inflation as it does. Repeated surprises raise realized inflation without causing the public to abandon that outlook.

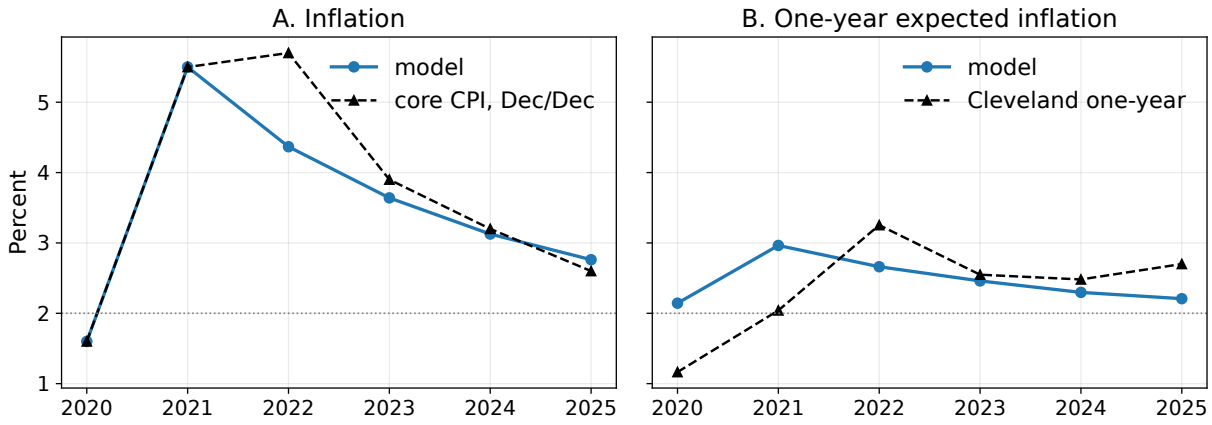

Figure 5: Baseline path and post-pandemic data

Notes: The model's inflation path is calibrated to the 2020 and 2021 core CPI observations; later observations and the Cleveland one-year expectations are untargeted. Core CPI is the December-to-December change and the Cleveland series the annual average of monthly one-year expectations, so the comparison is descriptive.

Table 2: Calibration

Parameter	Value	Target or rationale
Preferences and technology		
Discount factor β	0.96	Standard, annual
Phillips-curve slope κ	0.15	Production calibration, flat curve
Output-gap weight η	0.40	Inflation weight normalized to one
Shock process and information		
Cost-push persistence ρ_u	0.70	Decay of realized core-CPI deviation, 2021–25
Public persistence assessment ρ^{CB}	0.29	December 2021 SEP forecast ratio; proxy
Cost-push path u_{2021}	2.86	Matches 2021 core CPI, Dec/Dec
Initial condition, 2020	$\theta = -0.4$, $x = -4.0$	2020 core CPI and activity collapse; imposed, audit uninformative
Standing rule		
Rule loading ϕ_{IT}	0.35	Imposed accommodation of forecast pressure
Announced departure inherited from 2020 for 2021	0.50	Makeup guidance of September 2020; the 2020 core-inflation shortfall from target
Audit and enforcement		
Sanction scale Ψ (baseline)	18	Theory-section calibration; $\chi_j \Psi$ is the largest expected sanction
Communication cost γ	0.04	Fixed cost of a nonzero announcement; no guidance once it is worth less than this
Audit tolerance τ	1.10	Set to σ_π
Control noise σ_π	1.10	Annual inflation control error
Exposures χ_H, χ_L	1.00, 0.25	Mode H fully exposed; mode L one quarter
Beliefs and readings		
Prior λ_0; ergodic share $\bar{\lambda}$; persistence ρ_λ	0.12; 0.05; 0.87	High initial credibility, persistent modes
Enforcement skepticism: prior λ_0	0.60	Public enters the episode skeptical of enforcement
Forecast skepticism: forecast weight ω_v; feared ρ^{skept}	0.50; 0.85	Public forecast given half weight; feared persistence above the realized one
Public persistence ρ^{pub} (scenarios)	0.39–0.70	From the published ρ^{CB} to the realized ρ_u; enforcement unchanged

Notes: Enforcement and belief parameters are those of the theory exhibits of Section 3.2. The shock process and the initial conditions are specific to the episode.

The central bank enters 2021 with the announcement $\Delta^{p}_{2021} = 0.50$ made in 2020 and a probability of low enforcement of 0.111. Realized pressure is 2.86, against the inherited forecast of 0.04. The standing rule accommodates most of this surprise. The new assessment anticipates much less pressure in 2022: 0.83.

That forecast affects expectations through both the standing rule and expected delivery. If it proves correct, the rule prescribes a 2022 center of 0.29 above target (2.29 percent). The central bank announces an additional departure of 0.47. The public expects it to deliver 0.67, allowing for the incentive to exceed the announcement. Expected inflation for 2022 is therefore 2.29 + 0.67 = 2.96 percent. The public observes the forecast and anticipates policy under the rule even when no additional guidance is issued.

In 2022, pressure is 2.00, exceeding the previous forecast. The rule's response to this surprise raises its center to 1.40 above target. The central bank delivers a further departure of 0.97, missing its announcement by 0.49. Inflation is 4.37 percent, above the expectation formed a year earlier by 1.40 points. Of this difference, 1.11 comes from the rule's response to the forecast error and 0.29 from delivery exceeding what the public had expected. Table 3 reports the full sequence, and Figure 6 compares the announcement due with delivered policy and the static preference.

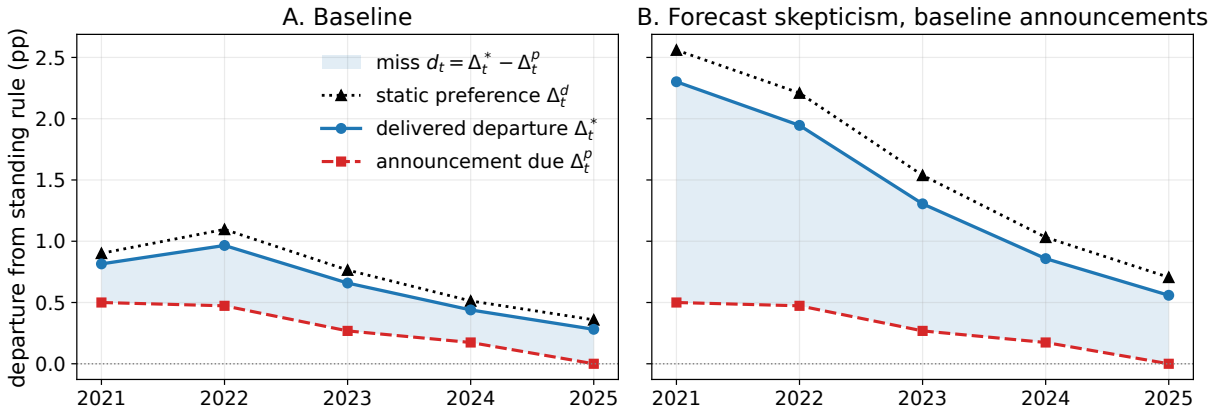

Figure 6: Look-through dilemma year by year

Notes: Departures from the standing rule, in percentage points. The announcement due was issued the previous year. The static preference minimizes current stabilization loss at given expected inflation, without sanctions or continuation incentives. The shaded band is the miss between the announcement and mode H's delivered departure. Panel A is the baseline; panel B holds announcements fixed and changes the public's forecast as in Section 4.3.

As pressure declines, inflation falls and the output gap closes. The gross value of further guidance falls below the communication cost $\gamma = 0.04$ from 2024 onward. Enforcement credibility changes little: the probability of low enforcement moves from 0.12

Table 3: Baseline path

	2020 (initial)	2021	2022	2023	2024	2025
Cost-push u_t	0.15	2.86	2.00	1.40	0.98	0.69
Standing rule $\bar{\theta}(\widetilde{S}_t)$	0.14	2.68	1.40	0.98	0.69	0.48
Announced departure due Δ_t^p	0.00	0.50	0.47	0.27	0.17	0.00
Static preference Δ_t^d	–	0.90	1.10	0.76	0.51	0.36
Delivered departure Δ_t^* (mode H, realized)	-0.54	0.82	0.97	0.66	0.44	0.28
Delivered departure of mode L, $\Delta_{L,t}^*$	–	0.88	1.06	0.73	0.49	0.34
New announcement Δ_{t+1}^p	0.50	0.47	0.27	0.17	0.00	0.00
Expected departure for $t+1$	0.13	0.67	0.46	0.32	0.20	0.14
Expected inflation	2.14	2.96	2.66	2.46	2.30	2.21
Inflation	1.60	5.50	4.37	3.64	3.13	2.76
Output gap	-4.00	-1.92	-1.81	-1.36	-0.94	-0.83
Probability of low enforcement λ_t	0.120	0.111	0.103	0.096	0.090	0.084

Notes: Inflation and expected inflation in percent; the cost-push disturbance in percentage points of annual inflation; the output gap in percent; the standing rule as a deviation from the two-percent target; the remaining policy objects as percentage-point departures from the standing rule. The announcement made in year t applies to year $t+1$; its expected departure is the departure the public expects to be delivered in $t+1$. The 2020 column is the imposed initial condition, and every entry from 2021 on is solved.

to 0.084. The two modes' inflation centers are close relative to control noise, so realized inflation provides little information about enforcement.

Set against the employment shortfall (Figure 8, panel C), the model reproduces the 2021 shortfall, −1.92 against −1.84, with nothing targeted; from 2022 the shortfall closes and the model's gap does not, −1.81 against −0.46 in 2022 and −0.94 against 0.25 in 2024. Over 2022–24 the model's gap sums to −4.11 point-years against 0.21. The episode is a pure cost-push disturbance here, so inflation held below full pass-through requires a negative gap; the speed of the recovery is the one dimension the application does not reproduce.

To isolate the contribution of additional guidance, we impose $\Delta_{t+1}^p = 0$ in every year, including the initial announcement, while retaining the rule, forecasts, and audit. Expected inflation then peaks at 2.88 percent, compared with 2.96 in the baseline. Guidance supports activity at the cost of higher inflation; these effects nearly offset at the baseline sanction scale. Cumulative stabilization loss is 13.28 with guidance and 13.39 without it.[11] The accepted forecast and the standing rule account for much of the inflation–expectations gap in both paths. The zero-guidance case retains a communicated policy obligation and the sanction for departing from it. Its favorable outcome is therefore an outcome of the targeting institution, even when issuing additional promises adds little.

[11]The comparison uses $\sum_{t=2021}^{2025} \frac{1}{2}(\theta_t^2 + \eta x_t^2)$ and excludes sanctions and communication costs. It therefore differs from the central bank's objective.

Had the disturbance decayed at the forecast rate, the 2022 delivery miss would have been 0.20 rather than 0.49. The vindicated path in Table 5 illustrates the cost of acting on an assessment that proves too optimistic. The next subsection changes the public's interpretation of that assessment and the institution enforcing the announcement.

4.3 Miscommunication Risks to Immaculate Disinflation

These scenarios feature a departure from the rational and belief-consistent equilibrium of the baseline model: here, the public is not surprised by the central bank, but the central bank is surprised by the public. The interpretation is miscommunication that poses policy risks. The public second-guesses the economy and the central bank in a way that is known to the central bank ex post and misperceived ex ante: the central bank assumes the baseline model's beliefs when choosing its announcement. The public fully understands the central bank's problem and forms rational expectations of its reaction function. At the end, we consider what happens when the central bank is aware of the discrepancy and best responds, that is, re-optimizes its announcements. In the baseline the two share the published persistence $\rho^{CB} = 0.29$; the realized persistence is $\rho_u = 0.70$, so the baseline public is wrong about the shock in the same way the central bank is. Everything else—the realized disturbance, the announcements, the rule, the sanction scale—is held fixed in every run.

With the leading rule (10), the inflation the public expects for next year is

$$e_t = \underbrace{\phi_{IT}\,\rho^{CB} u_t}_{\text{rule at the assessment}} + \underbrace{D\,(\rho^{pub} - \rho^{CB})\,u_t}_{\text{rule's response to the extra pressure the public expects}} + \underbrace{(1-\lambda_t)\mathbb{E}_H[\Delta^*_{t+1}] + \lambda_t \mathbb{E}_L[\Delta^*_{t+1}]}_{\text{expected delivered departure}}, \tag{30}$$

where ρ^{pub} is the persistence the public assigns the disturbance. The prior λ_0 enters the third term only. The persistence ρ^{pub} enters the second term directly and the third through the static preference (21), which rises with expected inflation at βD per unit. The frontier of Proposition 1 bounds the third term relative to the static preference, by a quarter of a point, and does not bound the static preference. Table 4 lists the primitive each scenario moves and the three terms the model implies for the 2021 announcement read for 2022.

The prior on enforcement, λ_0: 0.12 → 0.60. The public believes it probably faces the low-enforcement mode. The belief reaches the third term only, and along the baseline path there is little for enforcement to do: the obligation and the static preference are close (0.50 and 0.90 in 2021), so the two modes deliver within 0.06 of each other and even certainty of the low mode could move expected delivery by no more than that. Expectations peak

Table 4: What each scenario varies, and what the model implies for 2022

	Varied		Terms of (30), 2022				Outcome	
Scenario	λ_0	ρ^{pub}	rule	pressure	departure	e_{2021}	π_{2022}	loss
Baseline	0.12	0.29	0.29	0.00	0.67	0.96	4.37	13.3
λ_0: 0.12 → 0.60	0.60	0.29	0.29	0.00	0.70	0.99	4.38	13.5
ρ^{pub}: 0.29 → 0.39	0.12	0.39	0.29	0.28	0.91	1.47	4.62	17.3
ρ^{pub}: 0.29 → 0.50	0.12	0.50	0.29	0.56	1.29	2.13	4.96	23.2
ρ^{pub}: 0.29 → 0.60	0.12	0.60	0.29	0.83	2.00	3.13	5.54	32.4
ρ^{pub}: 0.29 → 0.70	0.12	0.70	0.29	1.11	3.40	4.80	6.60	49.4

Notes: Held fixed in every row: the realized disturbance, the central bank's published persistence $\rho^{CB} = 0.29$, the rule $\phi_{IT} = 0.35$, the sanction scale $\Psi = 18$, and, unless the row says otherwise, the baseline announcements. The left block lists the primitive varied, from the baseline value to the scenario's. The right block reports the three terms of (30) for the 2021 announcement read for 2022, their sum, realized 2022 inflation, and the 2021–25 stabilization loss.

at 2.99 percent rather than 2.96 and the loss is 13.46 rather than 13.28. The belief is about the audit's miss component, whose force the mode governs; the gap component that removes the inflation bias is borne in full by both modes. The result is not special to this calibration. The two modes' tractions differ by at most $(1 - \sqrt{\chi_L})/(1 + \sqrt{\chi_L}) = 0.33$ at any stake (Corollary 3), so a belief about enforcement moves expected delivery by at most that share of the distance between the obligation and the static preference. Here that distance is small and enforcement restrains delivery by 0.09 points in 2021; no belief about it can release more than that.

The persistence the public assigns the disturbance, ρ^{pub}: 0.29 → 0.70. The central bank publishes $\rho^{CB} = 0.29$ every year. We move the public's persistence from that value to the realized one in four steps, $\rho^{pub} = 0.39,\ 0.50,\ 0.60,\ 0.70$; at the last the public gives the published number no weight and forecasts the shock on its own. Each step adds $D(\rho^{pub} - \rho^{CB})u_t$ to the second term of (30), 1.11 points in 2022 at the endpoint, and raises the third through the static preference. Figure 7 traces the outcomes: expectations rise from 2.96 percent to 6.80, inflation from 5.50 to 9.02, the loss from 13.28 to 49.40.

The response is convex, and the reason is the frontier. Each tenth of persistence costs more than the one before: the loss rises by 4.0, 5.9, 9.3, and 17.0 across the four steps. The 2021 delivery miss grows from 0.32 at the baseline to 0.69, 1.21, 2.13, and 3.84, crossing the responsive range of 1.56 points between the second and third steps. Inside the range a larger miss raises the marginal sanction, so part of each additional point of expected pressure is absorbed by restraint; beyond it the marginal sanction falls, and the central bank slides to its static preference with the sanction fully in place. Modest departures of the

public's belief from the published assessment are absorbed by the institution. Large ones are not, and the same shock that produces the immaculate disinflation at the published persistence produces a 9.02 percent peak at the realized one. Inflation returns toward target as the pressure fades, so what is at stake is the surge, not its persistence.

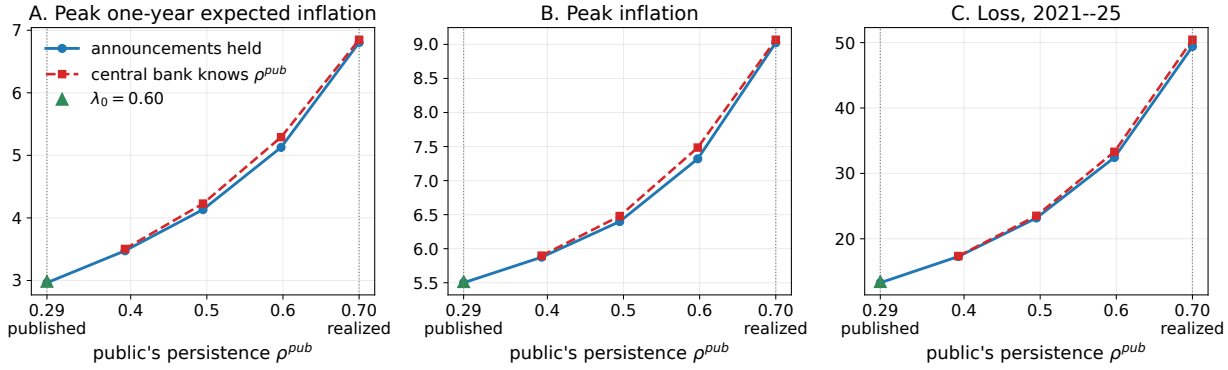

Figure 7: Expectations, inflation, and loss as the public's persistence moves from the published to the realized value

Notes: Peak one-year expected inflation, peak inflation, and the 2021–25 loss against ρ^{pub}, with the baseline announcements held (solid) and with the central bank re-optimizing its announcements knowing ρ^{pub} (dashed). The marked point is $\lambda_0 = 0.60$ at the published persistence. Realized shocks and the sanction scale are the same in every run.

Table 5: Benchmarks and public beliefs: outcomes over 2021–25

Scenario or public reading	Peak π	Peak $E\pi$	Lowest x	Total loss	Output loss
No additional guidance (standing rule)	5.34	2.88	-2.43	13.4	3.2
Vindicated, no additional guidance	5.34	2.88	-2.43	7.6	1.6
Vindicated look-through	5.50	2.96	-1.92	7.5	0.9
Baseline: shock persists	5.50	2.96	-1.92	13.3	2.1
λ_0: 0.12 $\rightarrow$ 0.60	5.52	2.99	-1.96	13.5	2.1
ρ^{pub}: 0.29 $\rightarrow$ 0.39	5.88	3.47	-2.69	17.3	3.6
ρ^{pub}: 0.29 $\rightarrow$ 0.50	6.40	4.13	-3.42	23.2	5.6
ρ^{pub}: 0.29 $\rightarrow$ 0.60	7.32	5.13	-3.64	32.4	7.3
ρ^{pub}: 0.29 $\rightarrow$ 0.70	9.02	6.80	-2.99	49.4	7.0

Notes: The benchmarks vary guidance or realized persistence. The belief rows retain the baseline announcements and change one primitive: the prior λ_0 on the low-enforcement mode, or the persistence ρ^{pub} the public assigns the disturbance, with the central bank's published persistence unchanged at $\rho^{CB} = 0.29$. Loss is the 2021–25 sum of $\frac{1}{2}(\theta_t^2 + \eta x_t^2)$, excluding sanctions and communication costs. Peaks and the lowest output gap cover the same years. Inflation centers are deviations from the two-percent target; output gaps are in percent.

When the central bank learns how it is read. The dashed lines in Figure 7 let the central bank know ρ^{pub} and re-optimize its announcements each year. It announces more, and

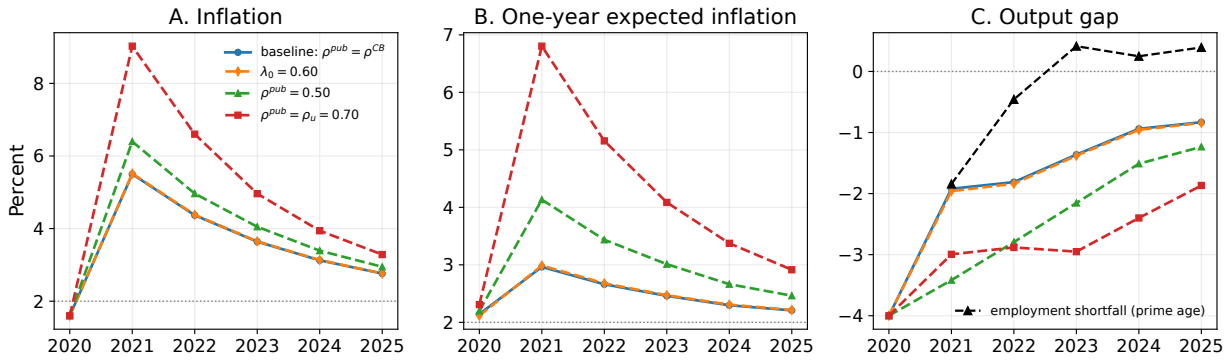

Figure 8: Same announcements, different beliefs: paths

Notes: All paths retain the baseline announcements and the realized shocks. The prior λ_0 changes the weight on the two modes' deliveries; ρ^{pub} changes the persistence the public assigns the disturbance, shown at 0.50 and at the realized 0.70. Panel C also shows the employment shortfall: the prime-age employment-to-population ratio against its 2019Q4 level, fourth quarter over fourth quarter, read one for one as the output gap (Abraham and Kearney, 2020). The 2020 initial condition is imposed and is not shown.

the outcome does not move: at the realized persistence the loss is 50.44 against 49.40. Beyond the responsive range no announcement buys different delivery, so knowing how it is read is worth nothing to the central bank once the public's belief has placed the miss there.

The conclusion we draw from these scenarios concerns nonlinearity. Enforcement skepticism changes only the term that the frontier bounds, and along the baseline path the announced departure is already close to the policy the central bank prefers, so there is little for enforcement to restrain; the gap between the two modes' responses is in any case capped by exposure alone, so its effect is small at any stake. A different public forecast of persistence changes the terms the frontier does not bound, and delivery amplifies the change, because the central bank accommodates the expectations it faces at βD per unit. Small departures from the published assessment are absorbed: the miss stays inside the responsive range, the sanction restrains delivery, and the disinflation proceeds at a modest additional cost. Large departures are not absorbed: the miss leaves the responsive range, the sanction stops restraining, the central bank delivers its static preference with enforcement in place, and the same disturbance produces a surge instead of a disinflation. The immaculate disinflation of the baseline therefore rests on the public accepting the central bank's assessment of the economy. It is robust to modest disagreement with that assessment and fragile to large disagreement. The next subsection asks what the sanction scale itself changes.

4.4 Policy Adjustment and Enforcement Strength

Can the central bank offset these risks by changing its announcements? The bottom panel of Table 6 lets it know the public's belief and re-optimize guidance at the baseline sanction scale, for the prior $\lambda_0 = 0.60$ and for the public's persistence at the realized value. A separate comparison, in the top panel and Figure 9, varies the sanction scale itself with baseline beliefs. The former changes what the public believes; the latter changes the institution's delivery incentives.

Knowing it faces $\lambda_0 = 0.60$, the central bank announces slightly more and the outcome moves by hundredths, for the reason given in Section 4.3: at the baseline scale both modes deliver close to their static preference and differ by 0.06 in 2021.

Knowing that the public uses the realized persistence, it announces more again, and the outcome does not move: the loss is 50.44 against 49.40 with the announcements held, and the peak of expected inflation 6.85 against 6.80. Beyond the responsive range no announcement buys different delivery, so re-optimization cannot recover the baseline; what the sanction scale itself changes is the subject of the top panel.

Table 6: Reoptimized guidance: loss decomposition

Scenario	Peak π	Peak $E\pi$	Total loss	Inflation loss	Output loss	Cum. gap	Largest $\lvert\Delta^p\rvert$	Largest expected $\lvert\Delta\rvert$
Sanction scale								
Baseline ($\Psi = 18$)	5.50	2.96	13.3	11.2	2.1	-6.9	0.47	0.67
Low enforcement ($\Psi = 8$)	5.59	3.02	13.4	12.0	1.3	-5.3	0.53	0.73
No additional guidance	5.34	2.88	13.4	10.2	3.2	-8.4	0.00	0.58
Public belief about the central bank, baseline scale, announcements reoptimized								
λ_0: 0.12 $\rightarrow$ 0.60, announcements re-optimized	5.52	2.99	13.5	11.5	2.0	-6.6	0.52	0.70
ρ^{pub}: 0.29 $\rightarrow$ 0.70, announcements re-optimized	9.07	6.85	50.4	46.1	4.3	-9.7	3.06	3.45

Notes: Announcements are reoptimized except in the no-additional-guidance comparison. The top panel varies the sanction scale; the bottom panel varies the public's reading at the baseline scale. Inflation and output losses sum $\frac{1}{2}\theta_t^2$ and $\frac{1}{2}\eta x_t^2$ over 2021–25, excluding sanctions and communication costs. The cumulative output gap is in point-years. The final columns give the largest absolute announced and expected departures over these years, in percentage points.

The top panel varies actual enforcement. At the lower sanction scale $\Psi = 8$, expectations peak at 3.02 percent and cumulative loss is 13.36, compared with 13.28 in the baseline. A weaker sanction permits delivery closer to the static preference. The public anticipates that accommodation, which raises expected inflation and worsens the current stabilization tradeoff. The loss moves by less than one percent between the two stakes. The frontier of Proposition 1 bounds what the stake can change in delivery, 0.264 points

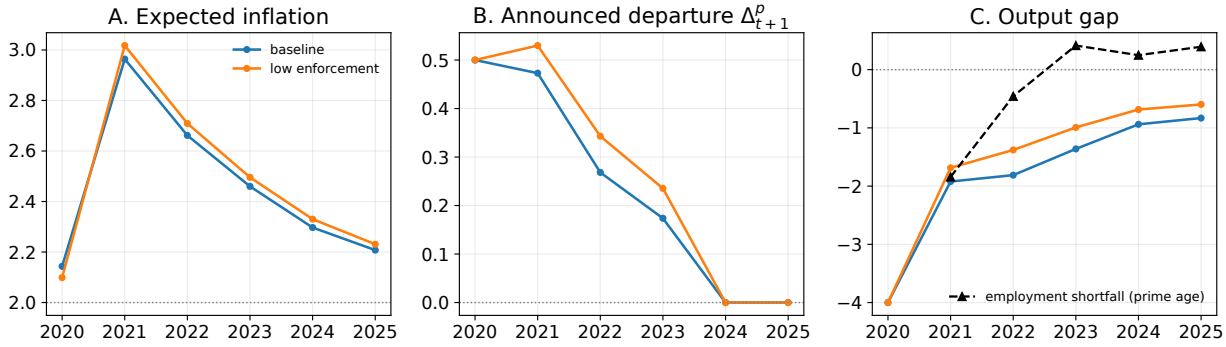

Figure 9: Paths by sanction scale

Notes: Each scenario reoptimizes the announcements and re-solves the continuation problem at its sanction scale. The exposures, the communication cost, and the variance of control noise are held fixed. Panel C also shows the employment shortfall: the prime-age employment-to-population ratio against its 2019Q4 level, fourth quarter over fourth quarter, read one for one as the output gap (Abraham and Kearney, 2020). The 2020 initial condition is imposed and is not shown.

at the baseline and 0.117 at the lower scale, and Corollary 3 bounds what beliefs about the mode can change at either.

5 Conclusion

The same supply disturbance can have very different implications for inflation expectations depending on how the public interprets the economy and the institution responding to it. This paper studies those differences in a theory of flexible inflation targeting with imperfect accountability. The standing rule and announced departures communicate a policy obligation. Bounded sanctions leave room for deviations at delivery, while observed policy can reveal the strength of enforcement.

The application distinguishes confidence in enforcement from acceptance of the central bank's assessment of the economy. When the public continues to expect rapid normalization, a rule that accommodates surprises can generate high realized inflation alongside much lower expectations. Inflation then declines as supply pressure subsides. Additional guidance has a modest effect at the baseline calibration because the rule already communicates much of the expected response.

Skepticism about enforcement is bounded by exposure alone, at any stake. A public that assigns the disturbance more persistence than the central bank publishes has much larger consequences, and the response is convex. Modest departures are absorbed while the delivery miss stays inside the sanction's responsive range; large ones are not, and at the realized persistence the same shock produces a surge with enforcement intact. The risk of looking through a supply shock therefore depends on more than confidence that promises will be kept. It depends on the public accepting the central bank's assessment of the economy, and the immaculate disinflation lasts as long as that acceptance does.

This points to a second role for the standing rule. Its first role is the one the model prices: it prescribes policy and defines the obligation the audit enforces. But by fixing the objective—a numerical target, with the audit's gap component removing the bias—the rule also settles what the central bank's words can be about. With the objective fixed, a published assessment and an announced departure can only be statements about the economy, and the public reads them as information. Without a target, the same words have no fixed referent: a judgment that pressure is transitory is as plausibly a statement of what the central bank wants as a forecast of what the economy will do, and a public that reads it the second way learns nothing from it. The model does not derive that shift in reading, since the weight the public places on the assessment is imposed. It prices the consequence, and the consequence is the difference between the two episodes with which this paper began.

References

Abraham, K. G. and Kearney, M. S. (2020). Explaining the decline in the US employment-to-population ratio: A review of the evidence. *Journal of Economic Literature*, 58(3):585–643.

Athey, S., Atkeson, A., and Kehoe, P. J. (2005). The optimal degree of discretion in monetary policy. *Econometrica*, 73(5):1431–1475.

Bassetto, M. (2019). Forward guidance: Communication, commitment, or both? *Journal of Monetary Economics*, 108:69–86.

Batista, Q., Nakata, T., and Sunakawa, T. (2023). Credible forward guidance. *Journal of Economic Dynamics and Control*, 153:104699.

Bocola, L., Dovis, A., Jørgensen, K., and Kirpalani, R. (2025). Monetary policy without an anchor. Working Paper 34436, National Bureau of Economic Research.

Bureau of Labor Statistics (2022). Consumer price index—december 2021. Release of January 12.

Christoffel, K. and Farkas, M. (2025). Managing the risks of inflation expectation de-anchoring. Working Paper Series 3082, European Central Bank.

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.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

Davig, T. and Foerster, A. (2023). Communicating monetary policy rules. *European Economic Review*, 151:104290.

Dovis, A. and Kirpalani, R. (2021). Rules without commitment: Reputation and incentives. *The Review of Economic Studies*, 88(6):2833–2856.

Drozd, L. A. (2026). Inflation targeting under delegation: Coarse rules or strict targets? Working paper, Federal Reserve Bank of Philadelphia.

Drozd, L. A. and Tavares, M. M. (2026). Capacity walls in production networks: Inflation after COVID. Working paper, Federal Reserve Bank of Philadelphia. Forthcoming.

Elkina, M. (2026). Reputation-driven design of monetary policy rules. Working paper.

English, B., Forbes, K. J., and Ubide, A., editors (2024). *Monetary Policy Responses to the Post-Pandemic Inflation*. CEPR Press, London.

Federal Open Market Committee (2020). Federal Reserve Issues FOMC Statement. Board of Governors of the Federal Reserve System. September 16.

Federal Open Market Committee (2021). Summary of economic projections. December 15.

Forbes, K., Ha, J., and Kose, M. A. (2025). Tradeoffs over rate cycles: Activity, inflation and the price level. NBER Working Paper 33825, National Bureau of Economic Research.

Fujiwara, I. and Waki, Y. (2022). The Delphic forward guidance puzzle in New Keynesian models. *Review of Economic Dynamics*, 46:280–301.

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

Gáti, L. and Handlan, A. (2025). Reputation for confidence. ECB Working Paper Series 3141, European Central Bank.

Haberis, A., Harrison, R., and Waldron, M. (2019). Uncertain policy promises. *European Economic Review*, 111:459–474.

Halac, M. and Yared, P. (2022). Fiscal rules and discretion under limited enforcement. *Econometrica*, 90(5):2093–2127.

Holmström, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, 10(1):74–91.

Ihrig, J. and Waller, C. J. (2024). The Federal Reserve's responses to the post-COVID period of high inflation. In English, B., Forbes, K. J., and Ubide, A., editors, *Monetary Policy Responses to the Post-Pandemic Inflation*, chapter 7, pages 131–142. CEPR Press, London.

King, R. G. and Lu, Y. K. (2022). Credibility and explicit inflation targeting. Working Paper 30012, National Bureau of Economic Research.

King, R. G. and Lu, Y. K. (2026). Evolving reputation for commitment: Understanding inflation and inflation expectations. Working paper, Boston University and Hong Kong University of Science and Technology.

King, R. G., Lu, Y. K., and Pastén, E. S. (2008). Managing expectations. *Journal of Money, Credit and Banking*, 40(8):1625–1666.

Kostadinov, R. and Roldán, F. (2025). Reputation and the credibility of inflation plans. *Journal of Economic Theory*, 229:106062.

Lohmann, S. (1992). Optimal commitment in monetary policy: Credibility versus flexibility. *American Economic Review*, 82(1):273–286.

Lu, Y. K., King, R. G., and Pastén, E. S. (2016). Optimal reputation building in the New Keynesian model. *Journal of Monetary Economics*, 84:233–249.

Mailath, G. J., Okuno-Fujiwara, M., and Postlewaite, A. (1993). Belief-based refinements in signalling games. *Journal of Economic Theory*, 60(2):241–276.

Nakamura, E., Riblier, V., and Steinsson, J. (2025). Beyond the Taylor rule. NBER Working Paper 34200, National Bureau of Economic Research.

Nelson, E. (2022). How did it happen?: The Great Inflation of the 1970s and lessons for today. Finance and Economics Discussion Series 2022-037, Board of Governors of the Federal Reserve System.

Rubbo, E. (2024). Networks, Phillips curves, and monetary policy. *Econometrica*. Forthcoming.

Sargent, T. J. and Williams, N. (2026). Rationalizing Fed interest rate decisions in the 2020s. *Journal of Political Economy Macroeconomics*. Forthcoming. Working version of October 6, 2025, https://www.tomsargent.com/research/lost.pdf.

Schaumburg, E. and Tambalotti, A. (2007). An investigation of the gains from commitment in monetary policy. *Journal of Monetary Economics*, 54:302–324.

Waki, Y., Dennis, R., and Fujiwara, I. (2018). The optimal degree of monetary discretion in a New Keynesian model with private information. *Theoretical Economics*, 13(3):1319–1367.

Walsh, C. E. (1995). Optimal contracts for central bankers. *American Economic Review*, 85(1):150–167.

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

A Analytical Results

This appendix states the episode, regularity conditions, and full results summarized in Section 3.2. The recursive institution is unchanged; the episode follows one announcement through delivery and the subsequent update of credibility.

A.1 Policy Episode and Regularity Conditions

Policy episode. The results of Section 3.2 follow one announcement through three periods. In Period 1, the announcement period, the public conditions $\widetilde{S}_1$, the announcement due Δ_1^p, the assessment v, and the belief λ_1 are known; the central bank in mode j announces Δ_2^p, or no additional guidance, the public forms e_1, and the central bank then delivers the announcement due. The new announcement affects the Period-1 loss only through e_1, and $\ell_1(e_1)$ denotes that loss at the central bank's optimal Period-1 policy. In Period 2, the delivery period, the state s_2 is realized, public conditions are $\widetilde{S}_2 = (s_2, v)$, and with them the static preference Δ_2^d; the central bank delivers Δ_2, inflation $\pi_2 = \bar{\theta}(\widetilde{S}_2) + \Delta_2 + \varepsilon_2^\pi$ is observed, and the sanction is assessed. In Period 3, the credibility period, the public updates to the posterior q_2, and the central bank enters the next episode with $\lambda_3 = (1-\rho_\lambda)\bar{\lambda} + \rho_\lambda q_2$. The episode holds the mode fixed through delivery; in the application, announcement and delivery decisions overlap across years, and the continuation in (23) integrates over annual mode transitions. What makes an announcement costly is the gap between the announced departure and the departure the central bank prefers when it comes due,

$$B \equiv \Delta_2^p - \Delta_2^d. \tag{31}$$

A gap can arise even when the forecast is correct: the announcement may deliberately commit the central bank to a departure from its future static preference. Forecast errors and changes in expected inflation can alter that gap when delivery comes due. The results below condition on the Period-2 state and its static preference.

Let $\mathcal{W}_j(\lambda)$ be the central bank's expected loss in mode j at the start of the next episode when the public's belief is λ, and β_E the discount factor across episodes ($\beta_E = \beta$ in the annual application). The continuation loss after a miss $d = \Delta_2 - \Delta_2^p$ is

$$\mathcal{R}_j(d) = \beta_E \, \mathbb{E}\left[\mathcal{W}_j(\lambda_3) \mid d\right], \tag{32}$$

where λ_3 depends on the miss through the audit variable $d + \varepsilon_2^\pi$ and the posterior (18)–(19). The results rely on one property of $\mathcal{W}_j$.

Assumption 1 (Bounded value of credibility)**.** *For each mode, $\mathcal{W}_j$ is bounded on* $[0,1]$*; let* $\Delta\mathcal{W}_j$ *denote its range and* $\Delta\mathcal{W} \equiv \max_j \Delta\mathcal{W}_j$*.*

Because the miss shifts the location of a normal audit variable, the continuation loss inherits the smoothness of the normal density even when $\mathcal{W}_j$ is not smooth. Appendix C.6 shows that under Assumption 1

$$|\mathcal{R}_j'(d)| \le \rho_1 \equiv \frac{\beta_E \Delta\mathcal{W}}{\sigma_\pi\sqrt{2\pi}}, \qquad |\mathcal{R}_j''(d)| \le \rho_2 \equiv \frac{2\beta_E \Delta\mathcal{W}}{\sigma_\pi^2\sqrt{2\pi \mathrm{e}}}. \tag{33}$$

The credibility channel can add at most ρ_1 to the marginal cost of a miss and ρ_2 to its curvature. Both bounds grow with the range $\Delta\mathcal{W}$ of the continuation value and shrink with control noise. These are bounds on the discipline supplied by Bayesian learning; the corresponding limit on designed public punishments is studied in Drozd (2026).

Two constants of the sanction recur. With $V \equiv \tau^2 + \sigma_\pi^2$, the curvature of the sanction at a zero miss is $K_j''(0) = \chi_j \Psi\tau / V^{3/2}$, and the largest marginal sanction is $\sup_d |K_j'(d)| = \chi_j \Psi\tau/(\sqrt{\mathrm{e}}\, V)$, attained at $|d| = \sqrt{V}$. Since $K_j''(d) = K_j''(0)\, \mathrm{e}^{-d^2/2V}(1 - d^2/V)$, the sanction is convex exactly on the *responsive range* $|d| < \sqrt{V}$, where a larger miss raises the marginal sanction, and its curvature is never below $-2\mathrm{e}^{-3/2}K_j''(0)$. The results require the current loss to dominate the concave part of the sanction and the curvature that credibility can add.

Assumption 2 (Well-posed delivery)**.** *For each mode,* $A > 2\mathrm{e}^{-3/2}K_j''(0) + \rho_2$*. Let* $\underline{A}_j \equiv A - 2\mathrm{e}^{-3/2}K_j''(0) - \rho_2 > 0$ *denote a uniform lower bound on the curvature of the delivery objective in mode* j*.*

At the calibration of Section 4, $A = 18.8$, $K_H''(0) = 5.26$, and $\sqrt{V} = 1.56$; the margin $A - 2\mathrm{e}^{-3/2}K_H''(0) = 16.4$ is the curvature margin with the direct sanction alone.[12]

A.2 Delivery and the Communication Frontier

The results in this subsection maintain Assumptions 1 and 2. They condition on the delivery state and expected inflation. The direct-sanction problem sets $\mathcal{R}_j = 0$.

Lemma 1 (Communication frontier)**.** *(i) The delivery objective*

$$\frac{A}{2}(\Delta_2 - \Delta_2^d)^2 + K_j(\Delta_2 - \Delta_2^p) + \mathcal{R}_j(\Delta_2 - \Delta_2^p) \tag{34}$$

[12] Appendix D reports the continuation-curvature and second-order-condition checks on the computed paths; Table 8 collects the notation.

is strictly convex with curvature at least $\underline{A}_j$, so its minimizer Δ_2^* is unique and solves

$$A(\Delta_2^* - \Delta_2^d) + K_j'(d_2^*) + \mathcal{R}_j'(d_2^*) = 0, \qquad d_2^* = \Delta_2^* - \Delta_2^p. \tag{35}$$

(ii) The delivered departure lies within a fixed distance of the static preference:

$$|\Delta_2^* - \Delta_2^d| \le \frac{\sup_d |K_j'(d)| + \rho_1}{A}. \tag{36}$$

Proof. See Appendix C.1. □

The same bound applies when $\Delta_2^p = 0$: the rule is an obligation whose enforcement has limited force. With the direct sanction alone, delivered policy lies between the obligation and the static preference. Close to the point where they coincide,

$$\theta_2^* - a_2 = \frac{A}{A + K_j''(0)}(\theta_2^d - a_2) + o(|\theta_2^d - a_2|), \qquad \theta_2^d \equiv \bar{\theta}(\widetilde{S}_2) + \Delta_2^d.$$

Stronger enforcement reduces a small delivery miss without removing the incentive to accommodate. The formula follows by differentiating (35) at a zero miss with $\mathcal{R}_j = 0$. When current and future cost-push pressure are absent and the rule prescribes zero, zero guidance and a zero inflation center can be sustained in the full recursion; Appendix G.1 states this benchmark. A low-inflation outcome therefore need not require additional guidance.

Proposition 3 (Announced and delivered departures)**.** *(i) With the direct sanction alone, the delivered departure is a continuous, single-valued function of the announcement, the delivery map $\Delta_j^*(\Delta_2^p)$. For announcements at or above the static preference, it increases on $\Delta_2^d \le \Delta_2^p \le \hat{p}_j \equiv \Delta_2^d + \sup_d |K_j'(d)|/A + \sqrt{V}$, reaches the frontier $\Delta_2^d + \sup_d |K_j'(d)|/A$ at $\hat{p}_j$, and decreases beyond it, converging to Δ_2^d as the announcement grows without bound; the map is symmetric for announcements below the static preference. (ii) With the credibility channel, the delivered departure lies within $\rho_1/\underline{A}_j$ of the direct-sanction map at every announcement. (iii) Under the pooled reading, the departure the public expects from an announcement is the credibility-weighted mixture of the two modes' maps, integrated over the Period-2 state,*

$$\widetilde{\mathbb{E}}_1[\Delta_2 \mid \Delta_2^p] = \mathbb{E}_1\left[(1 - \lambda_1)\Delta_H^*(\Delta_2^p) + \lambda_1 \Delta_L^*(\Delta_2^p)\right]. \tag{37}$$

With the direct sanction alone and a common Period-2 state, the public's delivery map peaks between $\hat{p}_L$ and $\hat{p}_H$.

Proof. See Appendix C.2. □

Proposition 1 in the main text is the bound in Lemma 1(ii), using the expression for the largest marginal sanction given above.

A.3 Pooling and Equilibrium Selection

Let $U_j(p, \lambda)$ be mode j's expected episode loss when it announces p and the public assigns probability λ to low enforcement. It includes the optimized Period-1 loss, the communication cost for $p \neq 0$, and the discounted delivery and continuation losses. Define $\underline{U}_j \equiv \inf_p U_j(p, 1)$, the lowest loss the mode can achieve when read as low enforcement. The set

$$W \equiv \{p : U_L(p, 0) \geq \underline{U}_L\} \tag{38}$$

contains announcements that mode L would not prefer to imitate even if doing so identified it as mode H, relative to its best announcement read as mode L. Separation requires mode H to choose an announcement in this set. An infimum over an empty set is interpreted as $+\infty$.

Proposition 4 (Announcements do not reveal enforcement)**.** *Fix the Period-1 state and let p^* minimize $U_H(\cdot, \lambda_1)$, the announcement of mode H under the pooled reading. (i) The profile in which both modes announce p^*, the public reads p^* at λ_1 and any other announcement as the low-enforcement mode's, and credibility is updated from inflation, is a perfect Bayesian equilibrium if and only if*

$$U_j(p^*, \lambda_1) \leq U_j(m, 1) \quad \text{for every } m \neq p^*, \qquad j = H, L. \tag{39}$$

Among pooling equilibria, it is the one mode H prefers. (ii) In any equilibrium in which the two modes announce different departures, mode H's announcement lies in W and mode L's loss is at least $\underline{U}_L$. Mode L therefore prefers the pool of (i) to every such equilibrium if $U_L(p^*, \lambda_1) \leq \underline{U}_L$, and mode H does if

$$U_H(p^*, \lambda_1) \leq \inf_{p \in W} U_H(p, 0). \tag{40}$$

(iii) In the equilibrium of (i) the announcement leaves the belief at λ_1, and after inflation the posterior odds of low enforcement are

$$\frac{q_2}{1 - q_2} = \frac{\lambda_1}{1 - \lambda_1} \exp\left\{ \frac{\Delta_L^* - \Delta_H^*}{\sigma_\pi^2} \left[\pi_2 - \bar{\theta}(\widetilde{S}_2) - \frac{\Delta_H^* + \Delta_L^*}{2} \right] \right\}, \tag{41}$$

before the mode transition.

Proof. See Appendix C.4. □

Part (i) is Proposition 2 in the main text. Parts (ii) and (iii) give the comparison with separation and the updating formula.

Corollary 1 (The pool is undefeated)**.** *Let p^* be the unique minimizer of $U_H(\cdot, \lambda_1)$ and let the pool at p^* satisfy (27). (i) If $U_L(p^*, \lambda_1) \leq \underline{U}_L$ and (40) holds, the pool at p^* is undefeated. (ii) If both inequalities are strict, the pool at p^* defeats every separating equilibrium in which neither mode announces p^*. (iii) The pool at p^* defeats every pooling equilibrium at an announcement that mode L ranks strictly below p^*.*

The inequalities in part (i) rule out a preferred separating outcome for either mode, while the optimality of p^* for mode H rules out a defeating pool. Undefeatedness need not select a unique pooling equilibrium. The application uses the pool preferred by mode H. Appendix F reports the other equilibria retained by the numerical screen.

A.4 Learning Illustration

Section 3.2.2 reads the posterior (41) at the reference state and reports the speed of learning along the application paths. One case is worth stating formally. If the two modes deliver the same inflation center, the likelihood ratio in (41) equals one and the posterior is the prior, $q_t = \lambda_t$; the next prior then follows the exogenous transition

$$\lambda_{t+1} = (1 - \rho_\lambda)\bar{\lambda} + \rho_\lambda \lambda_t,$$

so repeated quiet outcomes move the belief only toward the ergodic share $\bar{\lambda}$ and never because policy was delivered. Inflation is informative about enforcement exactly when the modes would have delivered different policies.

B Supplementary Results

B.1 Commitment Benchmark for the Leading Example

We compare commitment, discretion, and prescribed policy paths for the leading example. The forecast is correct, and enforcement and control noise are absent. The disturbance follows $u_t = \rho u_{t-1} + \varepsilon_t$, with $0 \leq \rho < 1$. The planner minimizes $\mathbb{E}_1 \sum_{t \geq 1} \beta^{t-1} \ell_t$ subject to $\theta_t = \beta \mathbb{E}_t \theta_{t+1} + \kappa x_t + u_t$. Commitment begins at date 1 with no inherited multiplier, so $x_0 = 0$.

Commitment. The first-order conditions imply $\theta_t = -(\eta/\kappa)(x_t - x_{t-1})$. Define the price-level gap $p_t = \sum_{s=1}^{t} \theta_s$ and set $p_0 = 0$. Then $x_t = -(\kappa/\eta) p_t$. Substitution into the Phillips

curve gives

$$\beta \mathbb{E}_t p_{t+1} - (1 + \beta + \kappa^2/\eta) p_t + p_{t-1} = -u_t.$$

Its stable solution is

$$p_t = \delta p_{t-1} + \psi u_t, \qquad \theta_t = \psi u_t - (1 - \delta) p_{t-1}, \qquad \psi = \frac{\delta}{1 - \beta\rho\delta}, \tag{42}$$

where $\delta \in (0, 1)$ is the smaller root of $\beta\delta^2 - (1+\beta+\kappa^2/\eta)\delta + 1 = 0$. The response depends on current pressure and on the price-level gap inherited from past policy. Following a single innovation with no further shocks, p_t converges to zero: subsequent deflation reverses the initial rise in the price level.

At impact, $p_0 = 0$ implies $\theta_1 = D(u_1 + \beta \mathbb{E}_1 \theta_2)$. The commitment allocation therefore satisfies the static first-order condition at its own expected future inflation. Its advantage over discretion at that date comes from the expectation of future restraint. For a single innovation, the response per unit of current pressure is ψ at impact and $\psi[1 - (1 - \delta)/\rho]$ in the following year when $\rho > 0$. It subsequently turns negative as pressure fades.

Comparison. Discretion sets $\theta_t = u_t/(1 + \kappa^2/\eta - \beta\rho)$. The standing rule loads D on the initial innovation and ϕ_{IT} on its correctly forecast continuation. As a further reference, we impose a departure of $(D - \phi_{IT})\rho u_{t-1}$, yielding $\theta_t = D u_t$. This static-share path is not the solution of the announcement game.

Table 7 reports the response to a unit innovation at the calibration of Table 2. Here $\delta = 0.803$ and $\psi = 1.744$. The rule provides less accommodation than commitment early in the response but omits commitment's eventual reversal. Cumulative inflation is zero under commitment, 1.76 under the rule, 3.16 under static-share delivery, and 8.67 under discretion. Discounted stabilization loss is 3.58 under commitment, 6.75 under discretion, and 7.43 under the rule. The rule's early restraint is costly when the Phillips curve is flat. This comparison motivates a role for departures from the imposed rule without establishing that the rule itself is optimal.

The application differs from this benchmark in its forecast errors, enforcement, learning, and initial conditions. Along its realized geometric decay, the rule alone accommodates the fraction $\phi_{IT}\rho^{CB}/\rho_u + D(1 - \rho^{CB}/\rho_u) = 0.70$ of current pressure. Its substantial response to repeated surprises helps explain why additional guidance has a small effect at the baseline sanction scale.

Table 7: Commitment and the standing rule after a unit innovation

Year after a unit innovation	1	2	3	4	5
Disturbance u_t	1.00	0.70	0.49	0.34	0.24
Commitment, θ_t/u_t	1.74	1.25	0.69	0.04	−0.70
Discretion	2.60	2.60	2.60	2.60	2.60
Rule alone	0.95	0.35	0.35	0.35	0.35
Static-share delivery Du_t	0.95	0.95	0.95	0.95	0.95
Commitment, output gap x_t	−0.65	−0.98	−1.11	−1.12	−1.05
Commitment, price-level gap p_t	1.74	2.62	2.96	2.97	2.81

Notes: A unit innovation at date 1 followed by deterministic decay at $\rho = 0.70$, with correct forecasts and no inherited commitment multiplier. Policy paths are prescribed directly, without enforcement or control noise. Rows 2–5 report inflation per unit of current pressure; the last two rows report levels under commitment.

B.2 Local Traction

The slope of the delivery map at a zero gap gives the marginal effect of an announcement under the direct sanction alone.

Corollary 2 (Traction)**.** *At a zero gap and with the direct sanction alone,* $\partial\Delta_j^*/\partial\Delta_2^p = \omega_j = K_j''(0)/(A + K_j''(0))$*, and the marginal traction of an announcement on the departure the public expects is* $(1-\lambda_1)\omega_H + \lambda_1\omega_L$*.*

The proof and the corresponding derivative with continuation incentives are in Appendix C.2.

The two modes' tractions can never be far apart, whatever the stake.

Corollary 3 (Bound on the traction gap)**.** *Write* $k = K_H''(0) = \Psi\tau/(\tau^2 + \sigma_\pi^2)^{3/2}$*, so that* $K_L''(0) = \chi_L k$*. Then*

$$\omega_H - \omega_L = \frac{kA(1-\chi_L)}{(A+k)(A+\chi_L k)} \leq \frac{1-\sqrt{\chi_L}}{1+\sqrt{\chi_L}},$$

with equality at $k = A/\sqrt{\chi_L}$*, that is at* $\Psi = A(\tau^2 + \sigma_\pi^2)^{3/2}/(\tau\sqrt{\chi_L})$*. The bound depends on* χ_L *alone.*

Proof. Substituting $\omega_j = k_j/(A + k_j)$ with $k_H = k$ and $k_L = \chi_L k$ gives the expression. Its derivative in k has the sign of $A^2 - \chi_L k^2$, so it rises until $k = A/\sqrt{\chi_L}$ and falls after. At that point it equals $(1-\chi_L)/(1+\sqrt{\chi_L})^2 = (1-\sqrt{\chi_L})/(1+\sqrt{\chi_L})$. □

At a zero gap the modes' delivered departures differ by $(\omega_H - \omega_L)(\Delta_2^p - \Delta_2^d)$ to first order, so the corollary bounds the share of the distance between the obligation and the static preference that enforcement can make the modes deliver differently. It does not bound the level: both frontiers in Proposition 1 grow linearly in Ψ. At the calibration the

gap is 0.15 against the ceiling 0.33, attained at $\Psi = 128.5$; on the baseline path of Section 4 the realized share in 2022 is 0.15.

B.3 Fragility of Ambitious Guidance

A flat delivery map makes its inverse sensitive. Suppose the central bank wants a particular delivered departure and asks what announcement implements it. Inside the frontier the answer is to announce more than the target, by an overstatement that grows with the target. At the frontier the marginal sanction reaches its maximum. The announcement required to implement a target then becomes arbitrarily sensitive to the target and to enforcement strength.

Proposition 5 (Fragility near the frontier)**.** *Consider the direct-sanction delivery problem in mode j. (i) A target departure $\hat{\Delta}$ is implementable if and only if $A|\hat{\Delta} - \Delta_2^d| \le \sup_d |K_j'(d)|$. For $\Delta_2^d < \hat{\Delta} < \Delta_2^d + \sup_d |K_j'(d)|/A$ the implementing announcement on the responsive range is unique, $\Delta_j^p(\hat{\Delta}) = \hat{\Delta} + \delta$, where the overstatement $\delta \in (0, \sqrt{V})$ solves $K_j''(0)\,\delta\,\mathrm{e}^{-\delta^2/2V} = A(\hat{\Delta} - \Delta_2^d)$. (ii) Its sensitivities to the target and to exposure are*

$$\frac{\partial \Delta_j^p}{\partial \hat{\Delta}} = 1 + \frac{A}{K_j''(-\delta)}, \qquad \frac{\partial \Delta_j^p}{\partial \chi_j} = -\frac{\delta}{\chi_j\,(1 - \delta^2/V)}, \tag{43}$$

which tend to $+\infty$ and $-\infty$ as $\hat{\Delta}$ approaches the frontier, where $\delta \to \sqrt{V}$ and the implementing announcement converges to $\hat{p}_j$.

Proof. See Appendix C.3. □

The implementing announcement remains finite at the frontier, but its sensitivity diverges. To raise the target by one unit requires $A/K_j''(-\delta)$ additional units of overstatement. As δ approaches $\sqrt{V}$, the marginal sanction flattens and this ratio grows without bound. An increasingly large change in the announcement is therefore needed to offset a small change in the desired policy or enforcement strength. Beyond the frontier, no announcement induces the target departure.

At the calibration, delivering half of mode H's frontier (0.132 points) takes an announcement of 0.63, an overstatement of 0.50 points; nine tenths of the frontier take 1.32, and 99 percent take 1.66, against $\hat{p}_H = 1.82$. At those three targets, the marginal announcement responses are 5.2, 9.8, and 29.6 points per point of additional delivered departure. The first-order compensation for a one-percent reduction in exposure is 0.006, 0.021, and 0.075 points of additional announcement. These are local derivatives. The credibility channel

shifts the frontier by at most ρ_1/A in either direction (Appendix C.3). The online appendix draws the implementing announcement for both modes.

B.4 When Guidance Is Issued

The communication cost gives the announcement a participation margin. Since U_j includes it, the gross value of guidance is

$$\mathcal{G}_j \equiv U_j(0, \lambda_1) - \inf_{p \neq 0} \{U_j(p, \lambda_1) - \gamma\}.$$

Mode H issues guidance when $\mathcal{G}_H > \gamma$, with no guidance selected at equality; mode L follows under the pooling conditions. For a local quadratic approximation to the objective net of the communication cost, $U_H(p, \lambda_1) - \gamma \mathbf{1}\{p \neq 0\} \approx U_H(0, \lambda_1) - bp + \frac{1}{2}cp^2$, where b is the marginal gain from an announcement and $c > 0$ its local curvature, the gross value is $b^2/(2c)$. The central bank then issues no guidance when $|b| \leq \sqrt{2c\gamma}$. In the application the value of guidance to mode H falls from 0.46 in 2021 to 0.007 in 2025, against a communication cost of 0.04, and the central bank issues no guidance from 2024 on. To mode L, guidance is worth 0.046 in 2021, less than a fifth of its value to mode H.

B.5 Private Assessments

A public assessment separates the assessment from the incentives to announce and deliver policy. If the assessment were the central bank's private information, its policy announcement would also signal its assessment. A central bank with a given assessment could change its announcement to make the public infer a different one. The public's reading would affect expected future policy through the standing rule and through the central bank's eventual delivery. The announcement optimal with the assessment already public need not be incentive compatible as a signal of that assessment. Moreover, a bounded policy-miss sanction limits how much delivery an announcement can induce, so policy guidance alone may fail to distinguish a broad range of assessments. The companion paper studies strategic forecast revelation when enforcement is known (Drozd, 2026); with hidden enforcement as well, the announcement would signal both v_t and j_t.

A strict target illustrates the importance of imperfect delivery. Consider the local linear case in which a standing rule responds to an inferred disturbance with coefficient ϕ, a central bank with no binding promise would respond with the accommodation share D of (21), and a promise moves delivered policy with local traction $\omega \in [0, 1]$. The effective

response to the inferred assessment is then

$$\phi_\omega = \omega\phi + (1-\omega)D. \tag{44}$$

With assured delivery ($\omega = 1$), a strict target ($\phi = 0$) removes this channel, which underlies the companion paper's truthful-announcement benchmark. With imperfect delivery ($\omega < 1$), even that target leaves $\phi_\omega = (1-\omega)D > 0$: the assessment predicts the central bank's ex post repair of an undelivered promise. This local comparison is an illustration, not a characterization of the joint (v, j) signaling game.

The main text takes the assessment as public, as central banks publish their projections, to isolate the risk that policy guidance is not delivered. With a private assessment, truthful communication of it requires the strategic-reporting analysis and, potentially, the model-dependent accountability studied in the companion paper.

B.6 Notation and Enforcement Illustrations

Table 8 collects the notation used in the main text and the supplementary results. Figure 10 shows the expected sanction and its derivative. Table 9 reports how the communication frontier, local traction, and the announcement at peak delivery vary with the sanction scale.

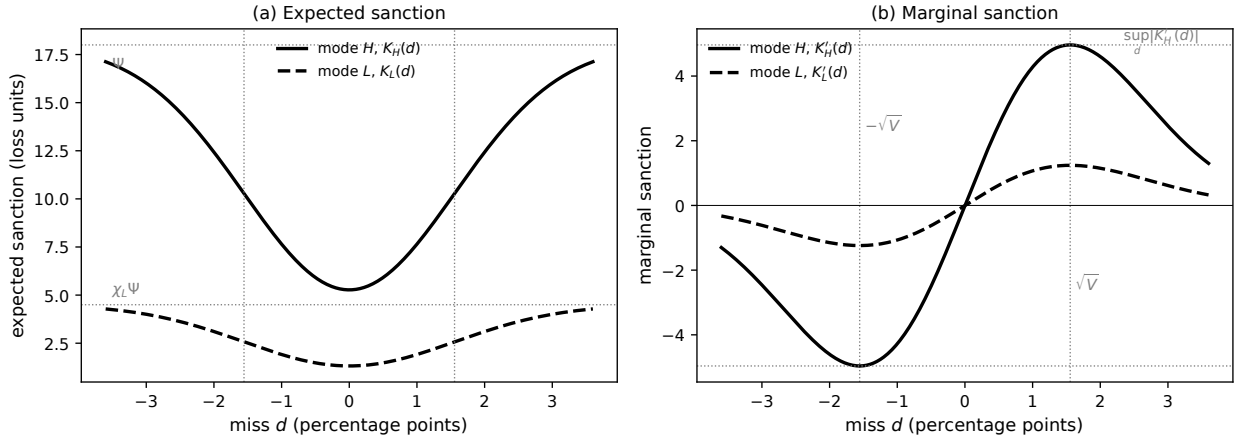

Figure 10: Expected and marginal sanctions

Notes: Panel (a): the expected sanction $K_j(d)$ of (16) at the calibration of Section 4 ($\Psi = 18$, $\chi_L = 0.25$, $\tau = \sigma_\pi = 1.1$). Panel (b): its derivative. The vertical lines mark the responsive range $|d| < \sqrt{V}$, on which the sanction is convex; the horizontal lines in panel (b) mark the largest marginal sanction of mode H, whose ratio to the curvature A of the current loss is the communication frontier of Lemma 1.

Table 8: Notation

Object	Meaning
$\widetilde{S}_t$, Ω_t	Public conditions (s_t, v_{t-1}); the state $(\widetilde{S}_t, v_t, \lambda_t, \Delta_t^p)$, public at the start of the period.
$\bar{\theta}(\widetilde{S}_t)$	Standing rule: the publicly understood inflation center, written on public conditions.
v_t	Public assessment of next period's conditions.
Δ_t^p	Announced departure due at t, issued at $t-1$; $\Delta_{t+1}^p = 0$ is no additional guidance.
a_t	Policy obligation, $\bar{\theta}(\widetilde{S}_t) + \Delta_t^p$.
T_t, x^*	The audit's gap component $\eta x^* x_t$, borne in full by both modes; the central bank's preferred output gap before it.
Δ_t, Δ_t^*	Delivered departure $\theta_t - \bar{\theta}(\widetilde{S}_t)$; its optimal value.
Δ_t^d	Static preference: the departure preferred on current grounds alone, at the expectation already formed.
d_t	Miss, $\Delta_t - \Delta_t^p$.
B	Gap in the episode, $\Delta_2^p - \Delta_2^d$.
e_t	Expected next-period inflation after the new announcement.
$j \in \{H, L\}$, χ_j	Enforcement mode, high or low; its exposure to the sanction, $\chi_H = 1 > \chi_L$.
λ_t, $\lambda_t(p)$, q_t	Probability of low enforcement at the start of the period, after announcement p, and after inflation.
$\bar{\lambda}$, ρ_λ	Ergodic share of low enforcement; persistence of the mode.
Ψ, τ, V	Sanction scale at full exposure; audit tolerance; $\tau^2 + \sigma_\pi^2$.
$K_j(d)$	Expected sanction in mode j; $K_j''(0)$ its curvature at a zero miss and $\sup_d \lvert K_j'(d)\rvert$ its largest marginal value.
A, D	Curvature of the current loss, $1 + \eta/\kappa^2$; accommodation share, $\eta/(\eta + \kappa^2)$.
γ	Communication cost of a nonzero announcement.
$\mathcal{V}_j$	Value in mode j; $\mathcal{V}_j(\cdot \mid \Delta_{t+1}^p)$ conditions on an issued announcement.
$\mathcal{R}_j(d)$	Discounted continuation loss after a miss, with future decisions optimal.
$\mathcal{W}_j(\lambda)$, $\Delta\mathcal{W}$	Value of entering the next episode with credibility λ; its range.
ρ_1, ρ_2	Bounds on the slope and curvature of $\mathcal{R}_j$ in the miss.
$\underline{A}_j$	Uniform lower bound on the curvature of the delivery objective, $A - 2e^{-3/2} K_j''(0) - \rho_2$.
$U_j(p, \lambda)$	Episode loss of mode j when it announces p and the public reads p with probability λ of low enforcement.
$\underline{U}_j$, W	Loss of mode j at its best announcement when read as low enforcement; the announcements mode L would not imitate even if believed.

Table 9: Frontier and traction across sanction scales

Scale Ψ	$K_H''(0)$	Frontier mode H	Frontier mode L	Traction ω_H	Traction ω_L	Public frontier at $\lambda = 0.12$	Peak $\hat{p}_H$
8	2.34	0.117	0.029	0.11	0.03	0.107	1.67
18	5.26	0.264	0.066	0.22	0.07	0.240	1.82

Notes: Direct-sanction results at a zero static preference, $A = 18.8$, $\sqrt{V} = 1.56$, and $\lambda = 0.12$. Mode-specific frontiers are $\sup_d |K_j'(d)|/A$; the public frontier is their belief-weighted upper bound. Traction is the derivative at a zero gap. Frontiers and peak announcements are in percentage points; curvature and traction are not. The baseline scale is $\Psi = 18$.

C Proofs

C.1 Sanction and Delivery Problem

The results condition on the public assessment and the delivery state. Write $b = \bar{\theta}(s_2, v_1)$ for the prescribed center at that state. Changing the rule's public arguments changes b but leaves $\Delta_2 = \theta_2 - b$, $d = \Delta_2 - \Delta_2^p$, and the audit variable $d + \varepsilon_2^\pi$ unchanged in form. The stabilization loss still has curvature A, with static preference $D(u_2 + \beta e_2) - b$. The delivery results therefore apply conditional on any such public benchmark. The pooling and undefeated-equilibrium arguments likewise condition on the assessment and compare only the two enforcement modes; their incentive inequalities must still be verified for the resulting payoffs. Public observability of the assessment and independence of its law from the enforcement mode are essential to this reduction.

Conditional on a miss d, the audit variable equals $d + \varepsilon_t^\pi$. Integrating the audit rule against the normal density of the control noise gives

$$\mathbb{E} \exp\left[-\frac{(d+\varepsilon_t^\pi)^2}{2\tau^2}\right] = \frac{\tau}{\sqrt{\tau^2+\sigma_\pi^2}} \exp\left[-\frac{d^2}{2(\tau^2+\sigma_\pi^2)}\right],$$

which proves (16). With $V = \tau^2 + \sigma_\pi^2$ and $K_j''(0) = \chi_j \Psi \tau / V^{3/2}$, direct differentiation gives

$$K_j'(d) = K_j''(0)\, d\, \mathrm{e}^{-d^2/2V}, \qquad K_j''(d) = K_j''(0)\left(1 - \frac{d^2}{V}\right)\mathrm{e}^{-d^2/2V}. \tag{45}$$

The absolute marginal sanction peaks at $|d| = \sqrt{V}$, with value

$$\sup_d |K_j'(d)| = \frac{K_j''(0)\sqrt{V}}{\sqrt{\mathrm{e}}} = \frac{\chi_j \Psi \tau}{\sqrt{\mathrm{e}}\, V}.$$

The curvature is smallest at $|d| = \sqrt{3V}$, where it equals $-2\mathrm{e}^{-3/2} K_j''(0)$.

Proof of Lemma 1. (i) By (33), the second derivative of the objective (34) is $A + K_j''(d) + \mathcal{R}_j''(d) \geq A - 2\mathrm{e}^{-3/2} K_j''(0) - \rho_2 = \underline{A}_j > 0$ for every d, so the objective is strictly convex with modulus $\underline{A}_j$; K_j and $\mathcal{R}_j$ are bounded, so the quadratic term makes it coercive, and it has a unique minimizer, characterized by the first-order condition (35).

(ii) Rearranging that condition and taking absolute values, $A|\Delta_2^* - \Delta_2^d| = |K_j'(d_2^*) + \mathcal{R}_j'(d_2^*)| \leq \sup_d |K_j'(d)| + \rho_1$. □

C.2 Proof of Proposition 3 and Corollary 2

We characterize the direct-sanction map from its first-order condition, locate its peak where the marginal sanction is largest, bound the effect of the credibility channel by ρ_1, and average over modes for the public's delivery map.

(i) Write $s \equiv \Delta_2^p - \Delta_2 = -d$, the negative of the miss, positive when the announcement exceeds the delivered departure, and $g(s) \equiv K_j''(0)\, s\, \mathrm{e}^{-s^2/2V}$, so that $K_j'(\Delta_2 - \Delta_2^p) = -g(s)$. The first-order condition of the direct-sanction problem is $A(\Delta_2 - \Delta_2^d) = g(s)$; since $A > 2\mathrm{e}^{-3/2}K_j''(0)$ the objective is strictly convex by the argument of Lemma 1(i) with $\rho_2 = 0$, so the condition has exactly one solution at every announcement. Adding s to both sides of $\Delta_2 = \Delta_2^d + g(s)/A$ gives the announcement as a function of the shortfall,

$$\Delta_2^p = \Delta_2^d + \frac{g(s)}{A} + s, \qquad \frac{\mathrm{d}\Delta_2^p}{\mathrm{d}s} = 1 + \frac{g'(s)}{A} \geq 1 - \frac{2\mathrm{e}^{-3/2}K_j''(0)}{A} > 0,$$

because $g'(s) = K_j''(-s) \geq -2\mathrm{e}^{-3/2}K_j''(0)$. The map $s \mapsto \Delta_2^p$ is therefore a continuous strictly increasing bijection of the real line, and the delivered departure is the continuous single-valued function $\Delta_j^*(\Delta_2^p) = \Delta_2^d + g(s(\Delta_2^p))/A$. The function g is odd, strictly increasing on $[0, \sqrt{V}]$, attains its maximum $g(\sqrt{V}) = \sup_d |K_j'(d)|$ at $s = \sqrt{V}$, and decreases to zero on $[\sqrt{V}, \infty)$. The delivered departure therefore increases with the announcement while $0 \leq s \leq \sqrt{V}$, that is, for $\Delta_2^d \leq \Delta_2^p \leq \Delta_2^d + \sup_d |K_j'(d)|/A + \sqrt{V} = \hat{p}_j$, attains $\Delta_2^d + \sup_d |K_j'(d)|/A$ there, and decreases beyond $\hat{p}_j$ toward Δ_2^d as $s \to \infty$. Announcements below the static preference are the mirror image because g is odd.

(ii) Let Δ^0 solve $A(\Delta - \Delta_2^d) + K_j'(\Delta - \Delta_2^p) = 0$ and Δ^* solve the same equation with $\mathcal{R}_j'$ added. The left-hand side of the full condition is strictly increasing in Δ with slope at least $\underline{A}_j$ and vanishes at Δ^*; at Δ^0 it equals $\mathcal{R}_j'(\Delta^0 - \Delta_2^p)$, of absolute value at most ρ_1. Hence $|\Delta^* - \Delta^0| \leq \rho_1/\underline{A}_j$.

(iii) Under the pooled reading the belief after the announcement is λ_1, so (26) at Period 1 is the mixture (37), integrated over the Period-2 state. Each mode's delivered departure lies within $(\sup_d |K_j'(d)| + \rho_1)/A$ of the static preference by Lemma 1(ii), so the public's delivery map never exceeds the expected static preference by more than $(1-\lambda_1)(\sup_d |K_H'(d)| + \rho_1)/A + \lambda_1(\sup_d |K_L'(d)| + \rho_1)/A$. With the direct sanction alone and a common Period-2 state, both maps lie below Δ_2^d when the announcement is below Δ_2^d, increase on $[\Delta_2^d, \hat{p}_L]$ and decrease on $[\hat{p}_H, \infty)$ because $\hat{p}_L < \hat{p}_H$; a mixture with positive weights therefore attains its maximum on $[\hat{p}_L, \hat{p}_H]$. □

Proof of Corollary 2. At a zero gap the direct-sanction condition $A(\Delta_2-\Delta_2^d)+K_j'(\Delta_2-\Delta_2^p)=0$ is solved by $\Delta_2=\Delta_2^p=\Delta_2^d$; differentiating it with respect to the announcement gives $\partial\Delta_j^*/\partial\Delta_2^p = K_j''(0)/(A+K_j''(0)) = \omega_j$. Differentiating the mixture gives the public's slope. With the credibility channel the same derivative is $(K_j''+\mathcal{R}_j''-\mathcal{R}_{j,dp})/(A+K_j''+\mathcal{R}_j'')$, where $\mathcal{R}_{j,dp}=\partial^2\mathcal{R}_j(d\mid p)/(\partial d\,\partial p)$ is the mixed derivative at a fixed miss; it arises because the announcement changes the equilibrium inflation centers used in the audit. The finite-difference check in Appendix D confirms that the slope is close to the direct-sanction slope at the numerical reference state. □

C.3 Proof of Proposition 5

(i) The first-order condition $A(\hat\Delta-\Delta_2^d)=-K_j'(\hat\Delta-\Delta_2^p)$ requires $A|\hat\Delta-\Delta_2^d|\le\sup_d|K_j'(d)|$, which proves necessity. For sufficiency and uniqueness on the responsive range, write $\delta=\Delta_2^p-\hat\Delta$, so that the condition reads $g(\delta)=A(\hat\Delta-\Delta_2^d)$ with g as in the proof of Proposition 3(i); by the properties of g established there it takes each value in $(0,\sup_d|K_j'(d)|)$ exactly once on $(0,\sqrt{V})$ and decreases back toward zero beyond $\sqrt{V}$, so a second, more ambitious announcement also delivers $\hat\Delta$; it lies beyond the responsive range, where by Proposition 3(i) a larger announcement lowers the delivered departure.

The endpoints follow by taking $\delta=0$ or $\delta=\sqrt{V}$, and negative target gaps follow by symmetry. The proposition concerns the responsive-range announcement.

(ii) Differentiate $A(\hat\Delta-\Delta_2^d)+K_j'(\hat\Delta-\Delta_2^p\mid\chi_j)=0$ with respect to $\hat\Delta$ and to χ_j: $A+K_j''(1-\partial\Delta_j^p/\partial\hat\Delta)=0$ and $-K_j''\,\partial\Delta_j^p/\partial\chi_j+\partial K_j'/\partial\chi_j=0$, with $\partial K_j'/\partial\chi_j=K_j'/\chi_j=-K_j''(0)\delta e^{-\delta^2/2V}/\chi_j$ and $K_j''(-\delta)=K_j''(0)e^{-\delta^2/2V}(1-\delta^2/V)$. The divergence follows from $K_j''(-\delta)\to0$ as $\delta\to\sqrt{V}$, while $\Delta_j^p(\hat\Delta)=\hat\Delta+\delta\to\Delta_2^d+\sup_d|K_j'(d)|/A+\sqrt{V}=\hat p_j$. □

Remark 1 (Fragility with the credibility channel)**.** *Under Assumptions 1 and 2, the implementable targets satisfy $A|\hat\Delta-\Delta_2^d|\le\sup_d|K_j'(d)|+\rho_1$ and, if $p\mapsto\mathcal{R}_j'(\hat\Delta-p\mid p)$ is continuous on $[\hat\Delta-\sqrt{V},\hat\Delta+\sqrt{V}]$ at the fixed announcement-stage belief, include every target with $A|\hat\Delta-\Delta_2^d|\le\sup_d|K_j'(d)|-\rho_1$. The conditioning argument makes explicit that the continuation loss can depend on the announcement as well as on the miss.*

Proof of Remark 1. Necessity of $A|\hat\Delta-\Delta_2^d|\le\sup_d|K_j'(d)|+\rho_1$ is Lemma 1(ii). For sufficiency, the map $p\mapsto-K_j'(\hat\Delta-p)-\mathcal{R}_j'(\hat\Delta-p\mid p)$ is continuous under the stated continuity condition, is at least $\sup_d|K_j'(d)|-\rho_1$ at $p=\hat\Delta+\sqrt{V}$ and at most $-\sup_d|K_j'(d)|+\rho_1$ at $p=\hat\Delta-\sqrt{V}$, so it attains every value in $[-\sup_d|K_j'(d)|+\rho_1,\sup_d|K_j'(d)|-\rho_1]$; at such

a p the target satisfies the first-order condition and, by Lemma 1(i), is the delivered departure. □

C.4 Proof of Proposition 4

We verify the equilibrium conditions of part (i) directly, show in part (ii) that any revealing equilibrium reads mode L's announcement as low enforcement, and derive the posterior odds of part (iii) from the two modes' inflation centers.

Throughout, $U_j(p, \lambda)$ is as defined in Section 3.2.2, with $e_1(p, \lambda)$ formed through (26) at $\lambda_1(p) = \lambda$ and the Period-2 delivery problem (34) evaluated at credibility λ; the equilibrium concept is defined in Appendix C.5. The proposition takes a minimizing announcement p^* as given; $\underline{U}_j$ is an infimum and need not be attained. The numerical announcement search is on the compact interval reported in Appendix D. No general existence or continuity claim is needed for the payoff comparisons below.

(i) On the proposed profile both modes announce p^*, Bayes' rule gives $\lambda_1(p^*) = \lambda_1$, and $\lambda_1(m) = 1$ for $m \neq p^*$. Mode j's loss is $U_j(p^*, \lambda_1)$ and its loss from deviating to m is $U_j(m, 1)$, so neither mode deviates if and only if (27) holds; beliefs after inflation follow Bayes' rule from the two modes' inflation distributions. Since $U_j(m, 1) \geq \underline{U}_j$, the condition holds whenever $U_j(p^*, \lambda_1) \leq \underline{U}_j$, and the two differ only through the value of being read as the low-enforcement mode at p^* itself. Any pooling equilibrium has both modes announcing some $\tilde{p}$ that Bayes' rule reads at λ_1; mode H's loss there is $U_H(\tilde{p}, \lambda_1) \geq U_H(p^*, \lambda_1)$, with equality only at p^* when the minimizer is unique.

(ii) Let an equilibrium have the two modes announcing $p_H \neq p_L$. Bayes' rule gives $\lambda_1(p_H) = 0$ and $\lambda_1(p_L) = 1$, so mode L's loss is $U_L(p_L, 1) \geq \underline{U}_L$. Optimality for mode L against the deviation to p_H, which is read as high enforcement, requires $U_L(p_L, 1) \leq U_L(p_H, 0)$; hence $U_L(p_H, 0) \geq \underline{U}_L$, which is $p_H \in W$. If $U_L(p^*, \lambda_1) \leq \underline{U}_L$, then $U_L(p^*, \lambda_1) \leq U_L(p_L, 1)$: mode L's loss in the pool is no larger than in any separating equilibrium. Mode H's loss in the separating equilibrium is $U_H(p_H, 0) \geq \inf_{p \in W} U_H(p, 0)$, so under (40) it is no smaller than $U_H(p^*, \lambda_1)$. The argument uses only the on-path beliefs, so it covers separating equilibria supported by any off-path beliefs.

(iii) On the path the announcement is common to both modes, so the belief after it is λ_1. Conditional on the mode and the announcement, inflation is normal with mean $\bar{\theta}(\widetilde{S}_2) + \Delta_j^*$ and variance σ_π^2; subtracting the two log densities gives (41), and the mode transition (19) follows. □

The information a year of inflation carries, measured by the Kullback–Leibler divergence between the two modes' inflation distributions, is $(\Delta_H^* - \Delta_L^*)^2/(2\sigma_\pi^2)$; by Lemma 1(ii) applied to each mode it is at most $[(\sup_d |K_H'(d)| + \sup_d |K_L'(d)| + 2\rho_1)/A]^2/(2\sigma_\pi^2)$.

Incentive comparison. The comparison that decides membership in W is

$$U_L(p,1) - U_L(p,0) \le U_L(p,1) - \underline{U}_L \quad \Longleftrightarrow \quad p \in W.$$

The left-hand side is the reduction in loss from being read as high rather than low enforcement at the same announcement; the right-hand side is the loss from choosing that announcement rather than the best alternative read as low enforcement.

The comparison includes current stabilization, delivery, and continuation incentives. Being read as the high-enforcement mode changes expected inflation by the difference between the two modes' expected deliveries, which is negative when the announcement restrains the central bank relative to its static preference and positive when the announcement is more ambitious than the static preference. With a negative Period-1 output gap, the envelope derivative of the optimized current stabilization loss with respect to expected inflation is $-\eta\beta x_1^*/\kappa > 0$, so being believed lowers the current loss at a restraining announcement and raises it at an ambitious one, while better credibility reduces the future loss in either case; the relative strength of the effects depends on the state and on the announcement.

Bounded sanctions alone imply neither monotonicity of this comparison nor a single threshold for W; Appendix F computes the comparison in the application. The proof of Proposition 4(ii) requires only membership in W.

C.5 Defeat and Proof of Corollary 1

An equilibrium of the announcement game at the Period-1 state is a pair of announcements (p_H, p_L) and a reading $\lambda(m) \in [0,1]$ after every announcement m, with λ given by Bayes' rule on the path (λ_1 after an announcement both modes send, 0 after one mode H alone sends, 1 after one mode L alone sends) and each mode's announcement minimizing $U_j(m, \lambda(m))$; write $U_j(\sigma)$ for mode j's loss in equilibrium σ. Following Mailath et al. (1993, Definition 2), $\sigma = (p_H, p_L, \lambda)$ *defeats* $\sigma' = (p_H', p_L', \lambda')$ if there is an announcement m such that (a) $m \notin \{p_H', p_L'\}$ and the set $K = \{j : p_j = m\}$ of modes sending m in σ is nonempty; (b) $U_j(\sigma) \le U_j(\sigma')$ for every $j \in K$, with strict inequality for some $j \in K$; and (c) $\lambda'(m)$ differs from

$$\frac{\lambda_1 \pi_L}{\lambda_1 \pi_L + (1-\lambda_1)\pi_H} \tag{46}$$

for every $(\pi_H, \pi_L) \in [0,1]^2$ with $\pi_j = 1$ for the modes in K that strictly prefer σ and $\pi_j = 0$ for the modes outside K. Expression (46) is the probability of low enforcement conditional on m coming from the modes in K, a mode indifferent between the two equilibria being allowed to send it with any probability; (c) says that σ' reads m inconsistently with the modes that would send it. An equilibrium is *undefeated* if no equilibrium defeats it. With two modes, K is $\{H, L\}$ when σ is a pool and a single mode when σ is separating, and (46) equals λ_1 when both modes strictly prefer σ, 0 when $K = \{H\}$, and 1 when $K = \{L\}$.

Proof of Corollary 1. (i) Let σ defeat the pool at p^* through m. Since p^* is the only announcement sent in the pool, $m \neq p^*$. If σ is a pool, both modes send m, so $K = \{H, L\}$ and (b) requires $U_H(m, \lambda_1) \leq U_H(p^*, \lambda_1)$, contradicting the uniqueness of the minimizer. If σ is separating, K is a single mode. For $K = \{H\}$, (b) requires $U_H(p_H, 0) < U_H(p^*, \lambda_1)$; but $p_H \in W$ by the proof of Proposition 4(ii), so $U_H(p_H, 0) \geq \inf_{p \in W} U_H(p, 0) \geq U_H(p^*, \lambda_1)$ by (40). For $K = \{L\}$, (b) requires $U_L(p_L, 1) < U_L(p^*, \lambda_1)$; but $U_L(p_L, 1) \geq \underline{U}_L \geq U_L(p^*, \lambda_1)$. No equilibrium defeats the pool.

(ii) Let σ' be separating with $p^* \notin \{p'_H, p'_L\}$, and let σ be the pool at p^* and $m = p^*$. Then (a) holds with $K = \{H, L\}$, and (b) holds strictly for both modes: $U_H(p^*, \lambda_1) < \inf_{p \in W} U_H(p, 0) \leq U_H(p'_H, 0)$ and $U_L(p^*, \lambda_1) < \underline{U}_L \leq U_L(p'_L, 1)$. With both modes strict, (c) admits only $\pi_H = \pi_L = 1$, and (46) equals λ_1. Since σ' is an equilibrium, mode H does not gain from announcing p^*: $U_H(p^*, \lambda'(p^*)) \geq U_H(p'_H, 0) > U_H(p^*, \lambda_1)$, so $\lambda'(p^*) \neq \lambda_1$ and (c) holds.

(iii) Let σ' be a pool at $p' \neq p^*$ with $U_L(p', \lambda_1) > U_L(p^*, \lambda_1)$, and take σ the pool at p^* and $m = p^*$. Then $K = \{H, L\}$, both modes strictly prefer σ, (46) is again λ_1, and $U_H(p^*, \lambda'(p^*)) \geq U_H(p', \lambda_1) > U_H(p^*, \lambda_1)$ gives $\lambda'(p^*) \neq \lambda_1$. □

Other undefeated equilibria. A pool at an announcement mode L strictly prefers to p^* cannot be defeated by the pool at p^*, because mode L does not prefer the latter. Any defeating pool must weakly improve both modes' payoffs, strictly improve at least one, and satisfy the belief condition (c); when one mode is indifferent, that condition must still be checked. A separating equilibrium can defeat a pool only through a mode that strictly prefers its separating outcome, at an announcement unsent in the pool and with a reading inconsistent with (c). The application compares these equilibria on finite grids with the numerical conventions in Appendix F.

C.6 Bounds on Continuation Incentives

Under Assumption 1, the derivative bounds used in (33) can be chosen as

$$|\mathcal{R}_j'(d)| \le \rho_1 \equiv \frac{\beta_E \Delta\mathcal{W}}{\sigma_\pi\sqrt{2\pi}}, \qquad |\mathcal{R}_j''(d)| \le \rho_2 \equiv \frac{2\beta_E \Delta\mathcal{W}}{\sigma_\pi^2\sqrt{2\pi \mathrm{e}}}. \tag{47}$$

Fix the delivery state and the public's equilibrium inflation centers. Let r denote a possible observed audit variable, an integration variable, and $h_j(r) \equiv \mathbb{E}[\mathcal{W}_j(\lambda_3) \mid r]$ the expected value of the next episode after the observation r, integrating the mode transition. Then

$$\mathcal{R}_j(d) = \beta_E \int_{\mathbb{R}} h_j(r)\, \frac{1}{\sigma_\pi}\varphi\!\left(\frac{r-d}{\sigma_\pi}\right) \mathrm{d}r,$$

where φ is the standard normal density and Φ below is its cumulative distribution function. Under Assumption 1, h_j takes values in an interval of width at most $\Delta\mathcal{W}$. Differentiating the density,

$$\mathcal{R}_j'(d) = \frac{\beta_E}{\sigma_\pi^2}\, \mathbb{E}\big[(r-d)\, h_j(r) \mid d\big], \tag{48}$$

$$\mathcal{R}_j''(d) = \frac{\beta_E}{\sigma_\pi^4}\, \mathbb{E}\big[\big((r-d)^2 - \sigma_\pi^2\big)\, h_j(r) \mid d\big]. \tag{49}$$

The weights have zero expectation, so subtracting a constant from h_j leaves both derivatives unchanged; subtract the midpoint of its range, so that $|h_j| \le \Delta\mathcal{W}/2$. For a standard normal variable Z, $\mathbb{E}|Z| = \sqrt{2/\pi}$ and $\mathbb{E}|Z^2 - 1| = 4\varphi(1)$; the triangle inequality gives (33). Differentiating the density rather than h_j means that no smoothness of $\mathcal{W}_j$ or of the posterior is needed. The bounds constrain the strength of credibility incentives, not their sign. For any two distributions of the audit variable and any continuation value in a range of width $\Delta\mathcal{W}$, the difference in expected continuation value is at most $\Delta\mathcal{W}$ times the total-variation distance between the distributions; with normal noise that distance is $2\Phi(|d-d'|/2\sigma_\pi) - 1$ for misses d and d'. This is the bound on public discipline derived for designed punishments in Drozd (2026); the Gaussian bounds above are its local form for the learning mechanism of this paper.

Conditional on the announcement and the Period-2 state, the two modes' inflation distributions are normal with means $\bar\theta(\widetilde{S}_2) + \Delta_H^*$ and $\bar\theta(\widetilde{S}_2) + \Delta_L^*$ and common variance σ_π^2. The Kullback–Leibler divergence above and the total-variation distance of the previous paragraph, evaluated at these two centers, both vanish when the two modes deliver the same departure, which, in the direct-sanction problem, happens when the announce-

ment coincides with the static preference: an announcement the central bank wants to keep teaches the public nothing about enforcement. Along the application the two delivered departures differ by a fraction of a point, and the audit moves the belief by a few hundredths per year.

D Numerical Algorithm and Checks

The paths in Section 4 and the recursive overlays in Section 3.2 are computed with one program, `code/engine_v66.py`, which solves the recursion of Section 3.1.5 by backward induction under the pooled reading. This appendix records the state, the equations solved at each stage, what the program does and assumes in each year of the episode, the construction of the counterfactuals, and the accuracy checks. Appendix F reports the verification of the equilibrium conditions.

Horizon and state. From each realized date, we solve the continuation problem along the forecast path $u_{t+h} = (\rho^{CB})^h u_t$, with a fourteen-year tail beyond 2025. The problem is solved again when the next realized shock arrives. Within a solve, future public conditions are fixed by this path. The remaining state is the pair (λ, Δ^p): the probability of low enforcement and the announcement due. The probability lies in $[(1-\rho_\lambda)\bar{\lambda}, (1-\rho_\lambda)\bar{\lambda}+\rho_\lambda]$, as implied by (19); the announcement lies in $[-1.5, 3.0]$. No announcement on a reported path reaches either bound.

Value functions and expected-departure functions are represented by tensor Chebyshev polynomials on 10×12 nodes. We fit two branches separately: no additional guidance and mode H's best nonzero announcement, including its communication cost. At each evaluation point the central bank chooses the branch with the lower loss. The continuation value and its derivatives are obtained by integrating over control noise, as described below.

The standing rule on the path is $\phi_{IT}\widehat{u}_{t|t-1} + D(u_t - \widehat{u}_{t|t-1})$, as in (29). At the last date the announcement is zero and the continuation is zero. Doubling the tail to twenty-eight years changes no reported 2021–25 figure by more than 2×10^{-7} (online appendix).

Stage problem. At a node and a candidate announcement p the public's reading is the prior at that node. The delivery stage is the fixed point of three equations in $(\Delta_H^*, \Delta_L^*, e_t)$: the first-order condition (35) of each mode, with the static preference (21) evaluated at e_t, and the expectation equation (26), in which the mode-specific expected departures are the next date's transition-weighted expected-departure functions evaluated at the

announcement p and at the posterior induced by the current audit. The continuation term is

$$\mathcal{R}_j(d) = \beta \sum_{j'} P_{jj'} \int \mathcal{V}_{j'}\big(\lambda_{t+1}(d+\varepsilon),\, p\big)\, \varphi(\varepsilon/\sigma_\pi)\, \frac{\mathrm{d}\varepsilon}{\sigma_\pi}, \qquad \lambda_{t+1}(r) = (1-\rho_\lambda)\bar{\lambda} + \rho_\lambda q_t(r),$$

where $q_t(r)$ applies (41) to the observed audit variable r using the two modes' equilibrium inflation centers at that node.

We divide the noise line at audit outcomes where the following date's choice switches between no guidance and a nonzero announcement. These outcomes are the roots of the fitted difference between the two branch values in the belief state. Each resulting segment is integrated with a Gauss–Legendre rule of 21 nodes. No segment crosses a participation boundary.

The same quadrature computes $\mathcal{R}_j'$ and $\mathcal{R}_j''$ using (48)–(49). These identities differentiate the noise density rather than the continuation value. They therefore account for jumps in next period's policy at a participation boundary. Differentiating the fitted polynomial within a branch would omit the contribution from crossing that boundary.

The fixed point is solved by Newton's method with a line search to a residual below 10^{-10}; a second start from the static preference is used if the first does not converge.

The announcement stage minimizes mode H's objective over p: a screen of 19 equally spaced candidates identifies local-minimum candidates, each of which is refined by bounded Brent minimization to a tolerance of 10^{-7}; the smallest refined value plus the communication cost γ is compared with the value of no additional guidance, and the better of the two is the announcement.

Mode L's value at the node is its objective at that announcement, including the communication cost when the announcement is nonzero (Proposition 4); its own best announcement is recorded for reporting. The next date's continuation used by mode j weights the two modes' values by the transition probabilities $P_{jj'}$. Announcements, delivered departures, and expectations are continuous; there is no announcement grid.

Exercise year by year. The application runs on annual data from 2020 to 2025 and reports 2021–25; the computational tail exists only so that the continuation values of the reported years are those of a long-horizon problem. The realized disturbance is $u_{2021} = 2.86$ in 2021 and decays at $\rho_u = 0.70$ thereafter; the assessment in every year is the forecast $\widehat{u}_{t+1|t} = \rho^{CB} u_t$ with $\rho^{CB} = 0.29$, and the public takes that forecast at face value in the baseline.

2020 is an initial condition. The allocation is imposed at the observed recession and core inflation ($\theta_{2020} = -0.4$, $x_{2020} = -4.0$) and the small disturbance $u_{2020} = 0.15$ is inherited from the predecessor calibration. No delivery decision is made, no sanction is assessed, and the audit is uninformative. The announced departure for 2021, $\Delta^p_{2021} = 0.50$, is imposed rather than chosen (Section 4.1 gives the historical basis), but expectations respond to it: the recursion is solved along the forecast path from 2020, which foresees no surge, and the expectation for 2021 is computed at that announcement exactly as at any other date. The prior for 2021 follows (19) from the initial prior $\lambda_{2020} = 0.12$.

2021 is the first solved year. The forecast made in 2020 was $\widehat{u}_{2021|2020} = 0.04$; the realized disturbance is 2.86, so the standing rule responds to the surprise with the center $\bar{\theta}(\widetilde{S}_{2021}) = 2.68$. The recursion is re-solved along the forecast path from 2021, with the announcement free at every date. At the realized state (prior 0.111, announced departure due 0.50) the central bank announces $\Delta^p_{2022} = 0.47$, and the delivered departure, mode L's hypothetical delivered departure, and expected inflation are the fixed point of the three stage equations: the static preference is 0.90 and the delivered departure 0.82, so the central bank exceeds the announcement due by 0.32 and pays the direct sanction on that miss. Inflation is 5.50 percent, the number the calibration of u_{2021} targets, and the output gap follows from (4) at the resulting expectation. The audit then updates the prior with the realized audit variable.

2022 to 2025 repeat the 2021 structure with no further assumptions: the only exogenous inputs from 2021 on are the disturbance path and the forecasts. Table 3 lists the resulting objects year by year. The gross value of guidance to mode H falls with the episode, from 0.46 in 2021 to 0.028 in 2024 and 0.007 in 2025, and the central bank issues no guidance from 2024 on, when that value falls below the communication cost $\gamma = 0.04$.

Realized paths and counterfactuals. On a realized path the actual mode is H, realized inflation equals its conditional center, and the posterior after delivery uses (41) at that outcome. The no-additional-guidance benchmark imposes a zero announcement in every year, including 2020, and switches the announcement problem off. The public-reading experiments impose, at each realized date, only the announcement the central bank made on the baseline path at that date; inside the public's expectations the central bank's future announcements remain the model's optimal policy. Their control, the baseline announcements read with the baseline beliefs, reproduces the baseline path to numerical precision, which the validation script checks.

Enforcement skepticism changes the initial prior. Forecast skepticism changes the persistence the public uses to forecast the shock to $\varpi_v \rho^{CB} + (1 - \varpi_v)\rho^{\text{skept}}$, while the

forecast, the standing rule, and the central bank's own continuation values still use ρ^{CB}; the public's expectations are solved along the feared path and supplied to the central bank's recursion by the two-recursion construction described below.

The public-persistence scenarios use the same construction as forecast skepticism with ρ^{pub} in place of the blended persistence, at $\rho^{pub} \in \{0.39, 0.50, 0.60, 0.70\}$, with the baseline announcements held and, separately, with the announcements re-optimized by a central bank that knows ρ^{pub} (`run_credence_v66.py`; `results/v66/credence_*.csv`). The former objective-skepticism scenario, a closed-form discretionary path with the sanction removed, is retired. The enforcement counterfactuals re-solve the whole recursion at each sanction scale and reoptimize every announcement.

When the public and the central bank use different forecasts, we solve two recursions at each realized date. The first follows the public's forecast path and produces its expected future rule and departure functions. The second follows the central bank's forecast path and produces its continuation values, taking the expectation functions from the first recursion as given. Its delivery and expectation fixed point therefore incorporates the public's alternative reading.

The forecast-skepticism experiment imposes the current baseline announcement in the second recursion. The forecast-skepticism regime in Section 4.4 optimizes it instead. The enforcement-skepticism regime uses the baseline forecast with a different initial belief about enforcement. The 2021 disturbance is calibrated by a one-dimensional root search to match 5.5 percent inflation at the baseline sanction scale (`code/calibrate_v66.py`).

Accuracy. The validation script records four checks in `results/v66/`. At realized decisions, fixed-point residuals are below 10^{-8}; at the underlying stage nodes they are below 10^{-10}. Along the baseline, the chosen announcement and delivery are also compared with their full objectives on grids of 91 and 601 points. No grid point improves on the reported choice by more than 10^{-6}.

The delivery second-order condition holds at each realized decision, with a minimum margin of 16.4. The largest measured continuation slope and curvature are 1×10^{-2} and 7×10^{-3}. The script also checks the stated quantitative comparisons across scenarios against the solved paths.

A coarser run uses 7×9 state nodes, a 13-point announcement screen, and 15 quadrature nodes. Across reported scenarios, its largest path difference from the standard settings is 0.021 percentage points, and its largest cumulative-loss difference is 0.04. The largest path difference occurs in the 2024 output gap under forecast skepticism with re-optimized guidance. It narrowly exceeds the validation script's 0.02-point tolerance. At

that state, holding the inherited announcement and belief fixed, doubling the standard continuation's announcement nodes changes the output gap by 0.021 points and the new announcement by 0.057 points. These discrepancies remain part of the numerical approximation; the state check does not replace the path comparison.

A separate check compares the delivery map in Proposition 3, the implementing announcement in Proposition 5, and the delivery derivative with the reference recursion. They agree to 10^{-6}, including on the declining side of the direct-sanction delivery map (`code/check_theory_formulas.py`).

Communication cost. The communication cost creates a participation boundary. At the reference state, the central bank switches from no guidance to an announcement of about 0.14 when the announcement due is about 2.9. Delivered policy changes by 0.034 because the new announcement changes expected inflation and hence the static preference.

Fitting a single polynomial across this boundary would smooth the policy jump and introduce oscillations. We instead fit the two branches separately and compare their values at each evaluation point. This preserves the discrete branch choice rather than smoothing it. Integration is likewise split at the fitted switching outcomes. The location of the boundary and the branch values remain numerical approximations.

At random states between nodes, the relative error in the fitted value functions is below 8×10^{-7} (median 8×10^{-8}). The absolute error in the expected-departure functions is below 2×10^{-5} percentage points (median 1×10^{-6}). Re-solving the baseline with twice as many announcement nodes changes the reported paths by at most 2×10^{-5}.

Accuracy of recursive overlays. The circles in Figure 2 solve delivery directly at the reference state, with a twenty-year continuation horizon and no further announcement. They do not evaluate a fitted delivery function in the announcement dimension.

Two checks assess the numerical approximation. First, a 61-point screen with local refinement examines mode H's announcement objective at every node of the reference recursion. Among 2280 objectives, it finds 0 competing candidates within 0.5 of the minimum. The same scan at the first three stages of the 2021 and 2022 application solves finds such candidates at 27 of 720 nodes. The next period's participation decision can create multiple local minima: switching between them makes the policy discontinuous and the value function kinked. The node-doubling check above measures their effect on the reported baseline path.

Second, the reference delivery map has a single change in the sign of its slope on a grid of 481 announcements. Re-solving with 14×16 nodes, a 37-point announcement screen,

and 31 quadrature nodes changes delivery by at most 1×10^{-5} and leaves the peak at 1.80 on that grid. Extending the horizon to thirty years changes it by at most 3×10^{-8}. The scripts `check_theory_convergence.py` and `check_announcement_uniqueness.py` report these comparisons and fail if they find no candidate minimum or exceed the stated accuracy tolerances. The scans are numerical diagnostics, not proofs of uniqueness between grid points.

E Data Construction

The introduction's inflation-and-expectations exhibit uses quarterly averages of monthly year-over-year CPI inflation and the Federal Reserve Bank of Cleveland's model-based one-year expected inflation. These series are distinct from long-horizon PCE forecasts. The descriptive regression reported in Section 4 regresses four-quarter changes in the quarterly-average one-year expectation on four-quarter changes in quarterly-average CPI inflation; the levels run from 2017Q1 to 2024Q4, so the four-quarter differences give $N = 28$ observations, 2018Q1–2024Q4, with Newey–West standard errors (four lags). The complement of the slope, 0.756, is a descriptive measure of how much of a change in realized inflation fails to pass into one-year expectations (`code/build_anchor_evidence.py`). The supplied quarterly file records three CPI months for 2025Q4 although BLS collected no October 2025 data; that observation is omitted from Figure 4 and enters neither the calibration nor the reported 2018Q1–2024Q4 regression sample.

The core-CPI comparison uses FRED series `CPILFESL`. We compute monthly twelve-month inflation rates, average them by quarter, and use the same expectations series, sample, and four-lag HAC estimator as for headline CPI. The supplied monthly observations and the script `code/build_anchor_evidence.py` reproduce both regressions. The core and headline slopes are descriptive comparisons; neither is a structural estimate of credibility.

The employment shortfall is $100\log(\mathrm{EPOP}_{t,Q4}/\mathrm{EPOP}_{2019Q4})$, where EPOP is the prime-age (25–54) employment-to-population ratio, FRED series `LNS12300060`, taken at the fourth quarter of each year to match the December-over-December inflation series; annual averages fold the 2020 lockdown quarters into the level and shift the base. The scaling regression is Okun's law in growth form: annual changes in $100\log$ real output (FRED `GDPC1`, fourth quarter) on annual changes in the shortfall over 1990–2019, ordinary least squares with classical standard errors ($R^2 = 0.37$, $N = 29$); the constant absorbs trend productivity growth. No measure of potential output enters anywhere. The

script `code/build_activity_comparison.py` reproduces the series, the regression, and the comparison with the model.

F Equilibrium Checks

At each announcement state, the script `code/check_equilibrium_v72.py` rebuilds the application recursion and evaluates its objective. The states cover the baseline, the vindicated path, and the enforcement- and forecast-skepticism regimes. Mode j's objective $U_j(m, \lambda)$ is as defined in Section 3.2.2, with the following year as the next episode. It is evaluated at 91 announcements spaced 0.05 apart on the announcement interval, together with no guidance and the equilibrium announcement p^*, and at 43 readings between 0 and 1. Along the paths the announcement state is λ_t, the announced departure due, and the assessment, all taken from the reported path. In the tables an announcement of 0 is no additional guidance.

Proposition 4. Tables 10 and 11 report the objects of the proposition. The *pooling margins* are $\min_{m \neq p^*} U_j(m, 1) - U_j(p^*, \lambda_t)$ for $j = H, L$: the smallest loss a mode incurs by deviating from p^* when the public reads the deviation as the low-enforcement mode's. Both are positive at every state on the numerical screen, consistent with condition (27); the smallest values across the 20 states are 0.034 for mode H and 0.008 for mode L.

The *ranking margin* of mode L is $\underline{U}_L - U_L(p^*, \lambda_t)$; its smallest value is 0.0000, at a date without guidance. The column *unimitated above* reports the smallest announcement above p^* in the set W; the column L *reading preferred above* reports the smallest announcement above p^* at which mode H's own loss falls when the public reads it as the low-enforcement mode's, the point beyond which the central bank would rather have an announcement read that way. The *separation margin* is $\min_{m \in W} U_H(m, 0) - U_H(p^*, \lambda_t)$, the loss of mode H at its preferred announcement in W, believed, relative to the pool; the announcement attaining the minimum is reported next to it, and the following column gives the change in the departure the public expects if that announcement were believed instead of p^* read at λ_t.

The last column of Table 11 reports whether the grid contains a separating perfect Bayesian equilibrium at the state: a pair (p_H, p_L) with p_H read as high enforcement and p_L as low such that neither mode gains from the other's announcement and every other announcement can be assigned a reading on the grid at which neither mode gains from it; the least ambitious such pair is shown.

Three features of the tables are the substance of the proposition in the application. First, the numerical checks support pooling at p^* at every state examined. Margins are large where guidance is active and small at dates without guidance.

Second, separation requires an announcement that mode L does not prefer to imitate. In the 2021 baseline state, the first screened announcement above p^* in W is 0.80, compared with $p^* = 0.47$. Mode H prefers the pool to its best announcement in W by 0.330. At announcements above 0.85, it prefers to be read as low enforcement. These comparisons describe the computed state; the proposition does not require a monotone ranking of announcements by the benefit of being believed.

Third, wherever guidance is active the central bank prefers the pool to every separating equilibrium in both modes: at the 13 such states the separation margin is at least 0.013 and mode L's ranking margin at least 0.008.

At the 7 states without guidance a separating equilibrium would have mode H pay the communication cost for a small announcement that mode L does not care to imitate; the margin there lies between 0.010 and 0.046, so the pool and separation are close.

Re-solving those states and the 2021 and 2024 baseline states with twice as many announcement nodes moves every margin by at most 0.0000. Separating equilibria exist on the grid at 5 of the states (the baseline in 2024 and 2025, the vindicated path in 2022, and the enforcement-skepticism regime in 2024 and 2025); in each, mode H announces a departure mode L does not care to imitate. At the baseline sanction scale, the departure the public would expect at mode H's preferred announcement in W, were it believed, differs from the pooled optimum by at most 0.239 points.

Belief-based refinements, for comparison. Finite-grid gain-set diagnostics for the intuitive criterion and criterion D1 are reported alongside the undefeated screen. For a deviation m, let $Q_j(m) = \{\lambda : U_j(m, \lambda) \leq U_j(p^*, \lambda_1)\}$ be the readings under which mode j weakly gains from it. The D1 diagnostic attributes m to mode H when $Q_L(m)$ is a strict subset of $Q_H(m)$; the intuitive-criterion screen uses $Q_L(m) = \emptyset$. These are finite-grid comparisons with numerical tolerances, not a proof of refinement acceptance over the continuous announcement and belief spaces.

Both criteria reject the pool at p^* at many states (the D1 diagnostic at 13 of the 20 states checked): an announcement slightly above p^* pays mode H more than mode L when it is believed, because mode H delivers more of it, so the criteria attribute the deviation to mode H and conclude that mode H should make it; the pool then fails because $U_H(m, 0) < U_H(p^*, \lambda_1)$ for m slightly above p^*. The undefeated criterion instead compares deviations supported by alternative equilibria, as discussed in Section 3.2.2.

Table 10: Pooling checks by state

Scenario	Year	λ_t	p^*	Pooling margin H	Pooling margin L	Ranking margin L	Unimitated above	L reading preferred above
Baseline	2021	0.111	0.47	0.323	0.320	0.320	0.80	0.85
Baseline	2022	0.103	0.27	0.198	0.153	0.153	0.60	0.65
Baseline	2023	0.096	0.17	0.109	0.081	0.081	0.45	0.45
Baseline	2024	0.090	0	0.067	0.085	0.066	0.25	0.35
Baseline	2025	0.084	0	0.072	0.067	0.036	0.05	0.25
Vindicated	2021	0.111	0.47	0.323	0.320	0.320	0.80	0.85
Vindicated	2022	0.103	0	0.054	0.070	0.045	0.15	0.30
Vindicated	2023	0.096	0	0.048	0.045	0.005	0.05	0.10
Vindicated	2024	0.090	0	0.044	0.042	0.000	0.05	0.10
Vindicated	2025	0.085	0	0.044	0.042	0.000	0.05	0.15
Enforcement skepticism (re-opt.)	2021	0.528	0.52	0.188	0.199	0.199	0.80	0.85
Enforcement skepticism (re-opt.)	2022	0.466	0.33	0.107	0.081	0.081	0.60	0.65
Enforcement skepticism (re-opt.)	2023	0.411	0.22	0.061	0.046	0.046	0.45	0.45
Enforcement skepticism (re-opt.)	2024	0.364	0.15	0.034	0.008	0.008	0.25	0.35
Enforcement skepticism (re-opt.)	2025	0.323	0	0.038	0.053	0.024	0.05	0.25
Forecast skepticism (re-opt.)	2021	0.111	1.44	1.642	1.344	1.344	2.45	2.20
Forecast skepticism (re-opt.)	2022	0.102	0.93	0.800	0.703	0.703	1.65	1.55
Forecast skepticism (re-opt.)	2023	0.094	0.65	0.422	0.331	0.331	1.15	0.95
Forecast skepticism (re-opt.)	2024	0.088	0.43	0.368	0.295	0.295	0.85	0.45
Forecast skepticism (re-opt.)	2025	0.083	0.30	0.122	0.090	0.090	0.60	0.55

Notes: Losses in the units of the paper's loss function; announcements in percentage points of inflation. p^* is the equilibrium announcement, 0 when no guidance is issued. Margins and columns are defined in the text.

Table 11: Separation checks by state

Scenario	Year	p^*	Separation margin	Announcement of mode H in W	Δexpected if believed	Separating equilibrium on the grid (p_H, p_L)
Baseline	2021	0.47	0.330	0.80	+0.090	none
Baseline	2022	0.27	0.225	0.60	+0.067	none
Baseline	2023	0.17	0.158	0.45	+0.057	none
Baseline	2024	0	0.046	0.25	+0.051	0.25, 0
Baseline	2025	0	0.029	0.05	+0.008	-0.05, 0
Vindicated	2021	0.47	0.330	0.80	+0.090	none
Vindicated	2022	0	0.017	0.15	+0.029	0.15, 0
Vindicated	2023	0	0.044	0.05	+0.010	none
Vindicated	2024	0	0.045	0.05	+0.011	none
Vindicated	2025	0	0.045	0.05	+0.011	none
Enforcement skepticism (re-opt.)	2021	0.52	0.195	0.80	+0.059	none
Enforcement skepticism (re-opt.)	2022	0.33	0.132	0.60	+0.040	none
Enforcement skepticism (re-opt.)	2023	0.22	0.107	0.45	+0.038	none
Enforcement skepticism (re-opt.)	2024	0.15	0.013	0.25	+0.015	0.25, 0
Enforcement skepticism (re-opt.)	2025	0	0.010	0.10	+0.012	0.05, 0
Forecast skepticism (re-opt.)	2021	1.44	2.009	2.45	+0.239	none
Forecast skepticism (re-opt.)	2022	0.93	0.939	1.65	+0.179	none
Forecast skepticism (re-opt.)	2023	0.65	0.526	1.15	+0.124	none
Forecast skepticism (re-opt.)	2024	0.43	0.417	0.85	+0.122	none
Forecast skepticism (re-opt.)	2025	0.30	0.174	0.60	+0.065	none

Notes: The separation margin is the loss of mode H at its preferred announcement in W, believed, relative to the pool at p^*; a positive value means it prefers the pool. Δexpected is the change in the departure the public expects, in points of inflation, if that announcement were believed. The last column gives the two modes' announcements in the least ambitious separating equilibrium the grid contains, or "none".

The D1 diagnostic instead accepts pools at more ambitious announcements, past the point at which being believed stops helping mode H; at the baseline sanction scale those pools differ from p^* in the departure the public expects by at most 0.106 points of inflation.

Corollary 1. Table 12 reports a finite-grid comparison of pure-strategy equilibria. A *pool* at announcement m is supported on the grid when, for every other grid announcement, some grid reading deters both modes; a *separating pair* (p_H, p_L) additionally requires that neither mode gains by imitating the other. The script compares four possible forms of defeat: pool by pool, pool by separation, separation by pool, and separation by separation. The deviating announcement must be unsent in the candidate equilibrium, and every mode sending it must weakly prefer the alternative, with at least one strict preference. When both modes strictly prefer a pool, or a single mode strictly prefers its separating outcome, equilibrium deterrence implies the belief inconsistency required by Appendix C.5. With an indifferent mode, the admissible posterior range in that definition must also be considered. The reported screen treats payoff differences within 10^{-9} as indifference and uses weak-payoff dominance for these cases, without separately checking that posterior range. Its counts are therefore numerical diagnostics, rather than a complete characterization at indifference or off the grid. A grid announcement within 0.005 of p^* is counted as the pool at p^*. The sufficient conditions for the selected pool are those of Corollary 1.

Under these numerical conventions, Table 12 shows four things. First, the screen retains the pool at p^* at every one of the 20 states, consistent with the sufficient conditions in Corollary 1(i) and the numerical margins in Tables 10 and 11, and p^* is mode H's pooled optimum on the grid at every state.

Second, the grid contains 11 separating equilibria, including 1 at states with active guidance. The pool at p^* defeats every active-state separating equilibrium found on the grid. At states without guidance, mode L also announces zero in every separating equilibrium found. Because zero is already sent there, the pool at $p^* = 0$ cannot defeat those equilibria through a new announcement at zero.

Of the 10 separating equilibria at these states, 2 remain undefeated; the others are defeated by alternative pools or separating equilibria. The survivors occur at the 2025 state of the baseline, where the separation margin of mode H is 0.029 and the surviving pair is (0.05, 0), and the 2025 state of the enforcement-skepticism regime, where the separation margin of mode H is 0.010 and the surviving pair is (0.10, 0). Their survival does not mean that either mode prefers them to the selected pool: defeat also requires an announcement that is not already sent in the candidate equilibrium.

Table 12: Undefeated-equilibrium screen by state

Scenario	Year	p^*	Pool at p^* undefeated	Pooled optimum of mode L	Undefeated pools (range; number)	Δexpected span	Separating equilibria found / defeated / undefeated
Baseline	2021	0.47	yes	0.25	0.25 to 0.47; 6	0.043	0 / 0 / 0
Baseline	2022	0.27	yes	0	0 to 0.27; 6	0.052	0 / 0 / 0
Baseline	2023	0.17	yes	0	0 to 0.17; 4	0.035	0 / 0 / 0
Baseline	2024	0	yes	0	0	0.000	1 / 0 / 0
Baseline	2025	0	yes	0	0	0.000	4 / 0 / 1
Vindicated	2021	0.47	yes	0.25	0.25 to 0.47; 6	0.043	0 / 0 / 0
Vindicated	2022	0	yes	0	0	0.000	2 / 0 / 0
Vindicated	2023	0	yes	0	0	0.000	0 / 0 / 0
Vindicated	2024	0	yes	0	0	0.000	0 / 0 / 0
Vindicated	2025	0	yes	0	0	0.000	0 / 0 / 0
Enforcement skepticism (re-opt.)	2021	0.52	yes	0.35	0.35 to 0.52; 5	0.024	0 / 0 / 0
Enforcement skepticism (re-opt.)	2022	0.33	yes	0	0.20 to 0.33; 4	0.020	0 / 0 / 0
Enforcement skepticism (re-opt.)	2023	0.22	yes	0	0 to 0.22; 4	0.035	0 / 0 / 0
Enforcement skepticism (re-opt.)	2024	0.15	yes	0	0 to 0.15; 2	0.025	1 / 1 / 0
Enforcement skepticism (re-opt.)	2025	0	yes	0	0	0.000	3 / 0 / 1
Forecast skepticism (re-opt.)	2021	1.44	yes	1.10	1.10 to 1.44; 8	0.069	0 / 0 / 0
Forecast skepticism (re-opt.)	2022	0.93	yes	0.60	0.60 to 0.93; 8	0.081	0 / 0 / 0
Forecast skepticism (re-opt.)	2023	0.65	yes	0.45	0.45 to 0.65; 5	0.045	0 / 0 / 0
Forecast skepticism (re-opt.)	2024	0.43	yes	0.30	0.30 to 0.43; 4	0.030	0 / 0 / 0
Forecast skepticism (re-opt.)	2025	0.30	yes	0	0 to 0.30; 6	0.089	0 / 0 / 0

Notes: Mode L's pooled optimum minimizes $U_L(\cdot, \lambda_t)$ on the announcement grid. "Undefeated pools" are those retained by the screen described above; their range and count include p^*. Δexpected span is their largest difference from p^* in the expected departure, in percentage points. The last column counts separating equilibria found, those eliminated by p^*, and those retained, using the stated tolerance conventions.

Third, the selected pool is the only undefeated equilibrium at 5 states. Other pools can remain because mode L prefers their announcements, while mode H prefers p^*. At the 2021 baseline state, mode L's pooled optimum is 0.25, compared with $p^* = 0.47$. The screen retains 6 pools between these announcements. Relative to its own optimum, mode H loses 0.103 at mode L's optimum; mode L loses 0.046 at p^*.

Fourth, the survivors are close in what the public expects: across the undefeated pools the expected departure moves by at most 0.089 points of inflation at the baseline sanction scale and 0.043 in 2021. Re-solving nine of the states with twice as many announcement nodes leaves the undefeated verdicts, the number of undefeated pools, and the number of undefeated separating equilibria unchanged at every re-solved state. The enumeration covers equilibria in which each mode announces a single departure, as do Proposition 4 and Corollary 1.

G The Audit's Gap Component and the Low-Inflation Benchmark

The audit's gap component is the output-gap levy used in Drozd (2026). The argument below shows why the reduction to aligned stabilization objectives remains valid with repeated announcements, noisy delivery, and learning.

Lemma 2 (Objective correction under delegation)**.** *Let the central bank's period loss before the gap component be $\ell_t^{CB} = \frac{1}{2}[\theta_t^2 + \eta(x_t - x^*)^2]$, where x^* is a known constant common to both enforcement modes. Suppose the gap component $T_t = \eta x^* x_t$ is levied in every period and borne in full by both modes, the exposure χ_j applying to the miss component alone. Then:*

(i) For every feasible policy and announcement strategy, its discounted expected loss including the gap component differs from the loss under (5) by $\eta(x^*)^2/[2(1-\beta)]$. Adding the same delivery sanctions and announcement costs to the two problems preserves this equality. Announcement and delivery incentives therefore coincide at every history, including histories off the equilibrium path.

(ii) With risk neutrality and the public expectation e_t available to the auditor, the component can be implemented by $\widetilde{T}_t = (\eta x^*/\kappa)[\pi_t - \beta e_t - u(s_t)]$. This uses no separate noiseless observation of the inflation center.

(iii) A component on inflation alone, $(\eta x^*/\kappa)\pi_t$, cancels the bias in the current-loss component of the delivery first-order condition but leaves an expectation-dependent term in the objective. It does not generally preserve announcement or continuation incentives.

Proof. For part (i), expand the square:

$$\begin{aligned}\ell_t^{CB} + \eta x^* x_t &= \frac{1}{2}\theta_t^2 + \frac{\eta}{2}x_t^2 - \eta x^* x_t + \frac{\eta}{2}(x^*)^2 + \eta x^* x_t \\ &= \ell_t + \frac{\eta}{2}(x^*)^2.\end{aligned}$$

The equality holds for every feasible pair (θ_t, x_t), not just at a conditional optimum. Discounting and taking expectations therefore adds the same constant $\eta(x^*)^2/[2(1-\beta)]$ to every strategy's payoff. Delivery sanctions, announcement costs, and the transition law are otherwise identical. Conditional strategy comparisons at every history are unchanged, as are the induced policies, inflation likelihoods, and Bayesian updates. The same argument applies to the continuation problem after any candidate announcement. Value functions

are shifted by the constant, while their policy-relevant differences and derivatives are unchanged.

For part (ii), the Phillips curve gives

$$\pi_t - \beta e_t - u(s_t) = \kappa x_t + \varepsilon_t^{\pi}.$$

Hence the implemented component is $\widetilde{T}_t = \eta x^* x_t + (\eta x^*/\kappa)\varepsilon_t^{\pi}$. The second term has conditional mean zero when the central bank chooses policy. Risk neutrality and the maintained noise independence yield the same expected objective. The component is a function of public variables and the already observed inflation realization, so it does not reveal the control noise or center separately.

For part (iii), write $\mu^* = \eta x^*/\kappa$. Then

$$\begin{aligned}\ell_t^{CB} + \mu^*\theta_t - \ell_t &= -\eta x^* x_t + \frac{\eta}{2}(x^*)^2 + \mu^*\theta_t \\ &= \frac{\eta}{2}(x^*)^2 + \mu^*[\beta e_t + u(s_t)].\end{aligned}$$

At delivery, e_t and s_t are fixed, so this period's remainder does not affect the marginal current loss. An announcement changes e_t, however, and generally changes the remainder. Future expectations also enter continuation costs. Thus the current-loss first-order condition alone does not establish equivalence of the recursive problems. Subtracting $\mu^*[\beta e_t + u(s_t)]$ is what distinguishes the corrected component from an inflation-only contract. □

The component is formulaic. It requires knowledge of x^* and the observable variables specified above, and it is borne in full by both modes: the exposure χ_j scales the miss component alone, so the hidden mode concerns only the sanction for missing a policy obligation. Eliminating systematic preference bias therefore does not eliminate the problem of implementing contingent promises, and it does not establish Ramsey implementation under an arbitrary standing rule.

Discretion without the correction. The role of the gap component is visible even in the absence of shocks. Without it, delivery sanctions, or commitment, the conditional first-order condition gives

$$\theta_t = D[u_t + \beta e_t + \kappa x^*].$$

For $u_t = 0$ and a stationary rational-expectations outcome, $e_t = \theta_t$, so

$$\theta^{BG} = \frac{D\kappa x^*}{1 - \beta D} > 0 \quad \text{when } x^* > 0.$$

The gap component removes the term $D\kappa x^*$. This is an objective correction; it is not accomplished by the symmetric delivery sanction alone.

G.1 A Zero-Pressure Benchmark

The following benchmark separates the ability to sustain low inflation from the ability to learn about enforcement. It describes an economy with no further supply pressure, not a claim about the stationary distribution of the economy with recurring shocks.

Lemma 3 (Low inflation without additional guidance)**.** *Suppose the gap component is in place, current and future cost-push pressure are zero, and the standing rule prescribes zero at every state reached in this economy. Start with an inherited announcement of zero. The outcome in which both modes issue no additional guidance and choose $\theta_t = x_t = 0$ at every date attains a global lower bound on their expected discounted losses and is consistent with public expectations $e_t = 0$. Realized inflation is control noise. It conveys no information about enforcement, and beliefs evolve as*

$$q_t = \lambda_t, \qquad \lambda_{t+1} = (1 - \rho_\lambda)\bar{\lambda} + \rho_\lambda \lambda_t.$$

Proof. After the gap component and omission of its additive constant, the central bank's period cost is stabilization loss, the delivery sanction, and any issuance cost. Stabilization loss is nonnegative, $K_j(d) \geq K_j(0)$ for every miss, and the issuance cost is nonnegative. Conditional on its current mode, expected discounted cost under any feasible strategy is therefore at least

$$\mathbb{E}_j \sum_{k=0}^{\infty} \beta^k K_{j_{t+k}}(0).$$

The transition law of enforcement is independent of policy, so this bound is common to all strategies, regardless of the beliefs or expectations those strategies induce. With zero inherited and new announcements, zero centers, and $e_t = 0$, the Phillips curve gives $x_t = 0$. The outcome attains the bound in every period. No alternative announcement or delivery strategy can yield a lower expected cost.

Both modes then generate the same inflation distribution, $\mathcal{N}(0, \sigma_\pi^2)$. Expected inflation is zero, so the expectations used to construct the outcome are consistent. The likelihood ratio is one; Bayes' rule leaves the mode probability unchanged before the Markov transition, which gives the stated belief law. □

If $\lambda_0 = \bar{\lambda}$, the belief stays at $\bar{\lambda}$. Otherwise it converges to that value through the transition law, not through informative inflation observations. In a disturbed economy, reputation can be built or lost when the modes' delivered policies differ; (41) determines how informative the observations are.