Strike while the Iron is Hot: Optimal Monetary Policy under State-Dependent Pricing

Peter Karadi
European Central Bank and CEPR

Anton Nakov
European Central Bank and CEPR

Galo Nuño
Bank of Spain and CEPR

Ernesto Pastén
Central Bank of Chile

Dominik Thaler
European Central Bank

This version: April 9, 2026
First version: July 30, 2024
Latest version: Please click here

Abstract

We characterize optimal monetary policy under state-dependent pricing. The framework gives rise to nonlinear inflation dynamics: large inflationary shocks trigger more frequent price adjustments, making the price level more flexible. Following large cost-push shocks, the central bank exploits this increased flexibility and curbs inflation more aggressively and at lower output cost than in standard time-dependent models. The optimal policy is well approximated by a nonlinear Taylor rule, in which the anti-inflationary stance increases more than proportionally with inflation. When faced with a total factor productivity shock — a shock to efficient output — the optimal policy maintains strict price stability.

JEL codes: E31, E32, E52

Keywords: State-dependent pricing, large shocks, nonlinear Phillips curve, optimal monetary policy

We are grateful to Guido Ascari, Andres Blanco, Luigi Bocola, Davide Debortoli, Eduardo Engel, Aurélien Eyquem, Jordi Galí, Erwan Gautier, Mark Gertler, Mishel Ghassibe, Basil Halperin, Francesco Lippi, Albert Marcet, Alberto Martin, Virgiliu Midrigan, Giorgio Primiceri, Xavier Ragot, Morten Ravn, Tom Sargent, Edouard Schaal, Raphael Schoenle, Mathias Trabandt and Jaume Ventura, as well as to participants at various conferences and seminars for their comments and suggestions. Sergi Barcons and Filippo Boggetti provided excellent research assistance. We acknowledge the use of AI tools (Gemini and Claude) for language polishing and stylistic improvements, as well as Refine for verifying algebraic computations. All AI-assisted suggestions were carefully reviewed, and the authors take full responsibility for the final content and accuracy of this paper. All errors are our own. The views expressed here are those of the authors only and do not necessarily represent those of the Bank of Spain, the Central Bank of Chile, the ECB, or the Eurosystem.

1. Introduction

A central question in macroeconomics is the optimal design of monetary policy. The canonical literature relies on "time-dependent" pricing frameworks (Calvo 1983), which assume a fixed frequency of price adjustment (Woodford 2003; Galí 2008). In contrast, a rapidly growing literature on state-dependent pricing — also known as menu-cost models — features firms that endogenously choose the timing of price changes. This allows the adjustment frequency and the degree of price stickiness to respond to economic conditions. The recent global inflation surge provides compelling evidence that this frequency responds strongly to the macroeconomic environment: in the U.S., it more than doubled at the inflation peak in 2022 (Montag and Villar 2023). Despite this evidence, the normative implications of state-dependent pricing remain underexplored. This paper addresses that gap by characterizing optimal monetary policy in a quantitatively disciplined state-dependent pricing model.

Our analysis yields a novel insight: optimal monetary policy under state-dependent pricing is inherently nonlinear. The central bank should respond to cost-push-driven inflation with increasing aggressiveness — the larger the inflationary shock, the more resolutely policy should tighten. The key driving force behind this result is that large inflationary shocks increase the repricing frequency and price flexibility, which reduces the output costs of disinflation. We characterize this policy as "striking while the iron is hot": because price flexibility is higher when inflation is elevated (the "iron is hot"), the central bank should tighten forcefully ("strike") to reduce inflation with minimal disruption to real economic activity. We show that this nonlinearity is quantitatively relevant for inflation surges of the magnitude experienced in 2021–2023.

Model Structure. We construct a menu-cost model that captures salient features of U.S. price setting, both in normal times and during the recent inflation surge. The model generalizes the seminal framework of Golosov and Lucas (2007). A representative household consumes the final good and supplies labor in a frictionless labor market. The final good aggregates a variety of differentiated goods. Each good is produced by a single firm using labor and intermediate inputs. Aggregate productivity, cost-push shocks, and firm-level shocks shift firms' desired prices. Firms pay a fixed menu cost to adjust nominal prices, resulting in an endogenous (S, s) decision rule: prices are kept unchanged within an inaction range and are only adjusted when they differ sufficiently from the optimal price. A central bank sets the nominal interest rate.

We depart from the original Golosov and Lucas (2007) framework in two empirically motivated dimensions. First, we assume that firm-level shocks follow a fat-tailed distribution rather than a Gaussian one (Midrigan 2011). This provides flexibility to match the high kurtosis and large average size of absolute price changes observed in microdata, as well as the strong comovement between inflation and repricing frequency that Gaussian models struggle to replicate (Blanco et al. 2024a). Second, we embed strategic complementarity via a roundabout production structure (Basu 1995)

to obtain a degree of monetary non-neutrality consistent with time-series evidence (Nakamura and Steinsson 2010). At the same time, by generating "pricing cascades", strategic complementarity delivers substantial inflationary effects even in response to moderate cost shocks (Ghassibe and Nakov 2025).

Dissecting Optimal Policy. To understand optimal monetary policy in this model, we begin with a tractable, simplified model version. It abstracts from roundabout production and fat-tailed shocks and introduces a "night" sub-period during which only firms are active and prices are fully flexible. The latter assumption transforms the firms' dynamic problem into a sequence of static problems while preserving the key underlying mechanism: the endogenous response of the repricing rate to aggregate shocks. In the simple model, both welfare and the central bank's choice set can be mapped into the space of inflation and the output gap (the deviation of output from its efficient level).

The model features three distortions. The first two arise from deviations of markups from their efficient levels: one from the *average* markup and the other from the *dispersion* of markups across goods. The third distortion reflects the resource costs of price adjustment. Welfare losses from these distortions can be expressed as functions of the output gap and inflation: output gap deviations affect welfare through the average markup, while inflation affects welfare through price adjustment costs and markup dispersion. As a result, welfare is approximately quadratic in inflation and in the output gap.

Due to price rigidities, monetary policy has real effects. The central bank's choice set is given by an upward-sloping, nonlinear Phillips curve: higher inflation is associated with a larger output gap, but also with a higher frequency of price adjustment, which diminishes the effectiveness of monetary policy and steepens the curve.[1] Consequently, the output cost of disinflation falls with inflation.

Cost-push shocks shift the Phillips curve and generate a trade-off between inflation and output-gap stabilization for the central bank. The optimal policy response to cost-push shocks follows from the interaction between the Phillips curve and the welfare function. Large inflationary shocks call for a more-than-proportional tightening of policy relative to small shocks. The intuition is that the steepening Phillips curve makes disinflation progressively less costly in output terms, or, put differently, reduces the effectiveness of monetary policy in stimulating output. At the same time, the relative welfare weight the policymaker places on inflation versus output stabilization remains remarkably stable, even at high levels of inflation. As a result, policy becomes increasingly anti-inflationary. This implies that the nonlinearity in optimal policy arises almost entirely from the changing costs of disinflation, rather than from shifts in the policymaker's preferences.

[1] Other reasons why the Phillips curve can be nonlinear are explored in Benigno and Eggertsson (2023) and Erceg et al. (2024).

We also demonstrate analytically that the optimal response to aggregate TFP shocks and time-preference shocks is characterized by the so-called "divine coincidence" (Blanchard and Galí 2007b), whereby monetary policy fully stabilizes both inflation and the output gap simultaneously.

Comparison with Calvo Pricing. The contrast with the canonical Calvo model is stark. Under Calvo pricing, the frequency of price adjustment is fixed by assumption, so the Phillips curve is near-linear and the output costs of disinflation are constant, independent of the inflation level. This yields an optimal policy response that is proportional to the inflationary shock: doubling the cost-push shock doubles the optimal policy response, unlike in our model.

However, the nonlinearity of optimal policy under menu costs is not a trivial consequence of the endogenous steepening of the Phillips curve alone; it also depends on the welfare function. Indeed, in the linearized Calvo framework, exogenously raising the frequency of adjustment and thus the slope of the Phillips curve simultaneously lowers the relative weight of inflation in the welfare function, so that the optimal policy remains unchanged. In our model, this offset does not occur: the welfare weights remain stable even as the Phillips curve steepens, and the increased price flexibility translates into more aggressive optimal policy. The stability of the welfare function is thus as crucial for our result as the nonlinear Phillips curve.

Notably, the welfare cost of inflation is somewhat higher in a Calvo model calibrated to match the Phillips curve slope at zero inflation of the baseline model (Auclert et al. 2024). This is partly because inflation *reduces* markup dispersion in the menu cost model, where firms with large price misalignments choose to adjust (selection effect), while inflation *increases* dispersion in the Calvo framework, where adjusting firms are chosen at random. Therefore, on the basis of its welfare function alone, the Calvo model would call for *less* inflation tolerance than our menu cost framework.[2] Despite this lower tolerance, however, the Calvo model prescribes *weaker* policy responses to large shocks. The reason is that the output cost of disinflation *does not* decrease in inflation in the Calvo model with its linear Phillips curve.

Quantitative Results and Policy Prescriptions. These insights extend to our fully dynamic baseline model. We calibrate the model parameters to match the monthly frequency and kurtosis of price changes, alongside the mean absolute size of price changes in the U.S. prior to the recent inflation surge (see, e.g., Nakamura and Steinsson 2008; Alvarez et al. 2016). We find that the model with fat-tailed idiosyncratic shocks captures these moments while delivering a realistic frequency response during an inflation surge. Furthermore, an empirically motivated share of intermediate inputs in production delivers a flatter Phillips curve than the standard menu-cost model, aligning its slope with the range of empirical estimates.

[2] This explains why previous research found higher optimal trend inflation rates under state-dependent frameworks than under Calvo (Burstein and Hellwig 2008; Blanco 2021).

Close to the steady state, the implications of our state-dependent model and the Calvo benchmark are similar. However, the nonlinearity rapidly becomes quantitatively significant as inflation rises. At 10% inflation, the slope of the Phillips relationship is about 4.6 times higher than under Calvo pricing. The optimal policy response to a cost-push shock is correspondingly more restrictive: the 10-year real interest rate increases by 75 basis points more than the standard Calvo response. Thus, while at low inflation levels the Calvo model provides a good approximation to the menu-cost model, this equivalence breaks down at levels seen in 2022, when U.S. inflation reached 10% and the repricing frequency more than doubled.[3]

We confirm the robustness of our main results in alternative state-dependent price setting models and for alternative parameterizations. First, we contrast optimal policy in the CalvoPlus model (Nakamura and Steinsson 2008) and the random menu cost model (Dotsey et al. 1999; Blanco et al. 2024a; Gagliardone et al. 2025) to our baseline model. Second, we assess the quantitative importance of our two new features: the fat-tailed idiosyncratic shocks and the roundabout economy. Third, we show that alternative models yield similar quantitative conclusions about the nonlinearity of optimal policy when calibrated to match the increase in price adjustment frequency observed in 2022.

We translate our theoretical findings into practical policy prescriptions by casting them in terms of a Taylor rule. We find that a linear Taylor rule, calibrated to mimic Ramsey optimal policy for small shocks, performs poorly for large shocks. Instead, the policymaker can significantly improve welfare by adopting a nonlinear Taylor rule in which the response coefficient of the interest rate to inflation increases smoothly with the level of inflation. This policy exploits the increased price flexibility and the attendant decline in the output costs of disinflation during high-inflation episodes. In a counterfactual scenario capturing features of the 2021–2023 inflation surge, such a policy would have led to at least 4 percentage points lower peak inflation with only mild output costs.

Related literature. Our paper contributes to the literature on state-dependent pricing by extending the seminal framework of Golosov and Lucas (2007) to better capture price-setting dynamics during both tranquil periods and inflation surges.[4] We add fat-tailed idiosyncratic shocks and strategic complementarity via a roundabout production structure to replicate the high kurtosis of price changes and the strong inflation-

[3] In Appendix D, we establish two further results in the baseline model. First, we show that the classic time-inconsistency problem remains present but is attenuated relative to the Calvo model. In both frameworks, an inefficient steady state creates an incentive to stimulate output via unexpected easing. However, in our model, such policy is less effective because the resulting spike in the repricing rate increases price flexibility, thereby weakening the motive to inflate the economy. Second, the model features an essentially zero Ramsey steady-state inflation rate, approximately the same as in the standard Calvo model, which prescribes exactly zero inflation.

[4] For foundational work, see Barro (1972); Sheshinski and Weiss (1977); Caballero and Engel (1993). For recent quantitative studies focusing on monetary non-neutrality, see Gertler and Leahy (2008); Midrigan (2011); Costain and Nakov (2011); Alvarez et al. (2016); Auclert et al. (2024).

frequency comovement observed in recent data (Blanco et al. 2024a; Montag and Villar 2023).[5] Relatedly, Blanco et al. (2024c) present a tractable model with a similarly strong feedback loop between inflation and frequency, which generates a substantial variation in the inflation-output tradeoff in the U.S. over the last decades.

To our knowledge, this paper is the first to solve for the Ramsey optimal monetary policy response to aggregate shocks in a heterogeneous-firm menu-cost model. This framework departs from the canonical textbook analysis (Woodford 2003; Galí 2008) based on Calvo (1983), where the frequency of price adjustment is an exogenous constant. We show that optimal policy is more aggressively anti-inflationary during an inflation surge caused by large fluctuations in costs because the variation in frequency generates a favorable inflation-output gap tradeoff after large shocks, a feature absent from the linear-quadratic Calvo framework, where the tradeoff remains invariant to shock size.

Our work complements previous research on optimal monetary policy in state-dependent price-setting models, focusing on a representative firm (Nakov and Thomas 2014), sector-specific productivity shocks (Caratelli and Halperin 2023) or optimal steady-state inflation (Adam and Weber 2019; Blanco 2021; Nakov and Thomas 2014).[6]

Finally, we contribute to the computational literature by proposing a new algorithm to solve Ramsey problems in heterogeneous-agent models. We extend the approach of González et al. (2024) to discrete-time problems with inaction regions, solving nonlinearly for optimal policy under perfect foresight over the sequence space (Auclert et al. 2021). Our approach complements other methods to solve Ramsey policy in heterogeneous-agent models (Bhandari et al. 2021; Le Grand et al. 2022; Dávila and Schaab 2022; Nuño and Thomas 2022).

2. Baseline Model

The economy consists of a representative household, a representative final-good producer, a continuum of monopolistically competitive firms, and a central bank. The household consumes the final good, supplies labor, and saves in nominal bonds. The final-good producer aggregates differentiated goods into a single consumption good using CES technology. Each atomistic firm produces a differentiated good by combining labor and intermediate inputs using a constant-returns-to-scale technology subject to idiosyncratic shocks. Each firm also faces menu costs when adjusting prices. Monetary policy is conducted by a central bank that sets the nominal interest rate.

[5] Changes in frequency have been documented both in response to large shocks to costs (Karadi and Reiff 2019; Alvarez and Neumeyer 2020; Auer et al. 2021; Gagliardone et al. 2025; Gautier et al. 2025), and in response to changes in trend inflation (Gagnon 2009; Alvarez et al. 2019; Nakamura and Steinsson 2018); they have received new empirical support following the recent U.S. inflation surge (Montag and Villar 2023; Cavallo et al. 2024).

[6] Nakov and Thomas (2014) find no difference between Calvo and a random menu cost model under the optimal policy. Caratelli and Halperin (2023) show that, in the face of sector-specific shocks, optimal policy can be characterized as nominal wage targeting.

Time is discrete, and we solve the model under perfect foresight, ignoring aggregate uncertainty. Hence, we suppress the expectations operator. Aggregate shocks are one-off events.

2.1. Household

A representative household consumes C_t, supplies labor N_t, and trades one-period nominal bonds B_t, which are in zero net supply. The household maximizes present discounted utility,

$$\max_{\{C_t,N_t,B_t\}} \sum_{t=0}^{\infty} \beta^t \left[\log(C_t) - \chi N_t\right] \tag{1}$$

subject to the budget constraint

$$P_t C_t + Q_t B_t + T_t = B_{t-1} + W_t N_t + D_t, \tag{2}$$

where P_t is the nominal price of the final good, T_t denotes lump-sum taxes, W_t the nominal wage, D_t firm dividends, and $Q_t \equiv 1/R_t$ the price of a nominal bond paying one unit of currency next period. The gross nominal interest rate is R_t.

Intratemporal maximization yields the household's labor-supply condition,

$$w_t = \chi\, C_t, \tag{3}$$

where $w_t \equiv W_t/P_t$ is the real wage.

Let $\pi_t \equiv \log(P_t/P_{t-1})$ denote the inflation rate. The consumption Euler equation is

$$1 = \beta \frac{C_t}{C_{t+1}} \frac{R_t}{\exp(\pi_{t+1})}. \tag{4}$$

Defining the real rate as $R_t^r = \frac{R_t}{\exp(\pi_{t+1})}$, iterating the Euler equation forward to infinity, using the fact that consumption and the real rate converge to their respective steady-state values $\bar{C}$ and $\bar{R}^r = \frac{1}{\beta}$, and taking logs, we can derive the following relationship:

$$c_t - \bar{c} = -\sum_{i=0}^{\infty} (r_{t+i}^r - \bar{r}^r), \tag{5}$$

where lower-case letters refer to logarithms. The log deviation of consumption from steady state in period t is equal to the negative of cumulative deviations of the net real interest rate from its steady state.

2.2. Final-good producer

A representative final-good producer aggregates differentiated goods $Y_t(j)$ into a single final good Y_t. It chooses the bundle $\{Y_t(j)\}_{j\in[0,1]}$ to minimize period-t costs,

$\int_0^1 P_t(j)\, Y_t(j)\, dj$, subject to the Dixit and Stiglitz (1977) aggregator

$$Y_t = \left\{ \int_0^1 \left[A_t(j) Y_t(j) \right]^{\frac{\epsilon-1}{\epsilon}} dj \right\}^{\frac{\epsilon}{\epsilon-1}}, \tag{6}$$

where $P_t(j)$ denotes the nominal price of good $j \in [0,1]$ and the term $A_t(j)$ is an idiosyncratic and stochastic quality shifter. The latter drives both relative demand and relative productivity in a computationally convenient way, as we show later (Midrigan 2011; Alvarez et al. 2016). The stochastic process for this quality shifter is

$$\log A_t(j) = \log A_{t-1}(j) + \varepsilon_t(j), \tag{7}$$

where the innovation ε_t is drawn from a mixture of two normal distributions, $N(0, \sigma_1)$ and $N(0, \sigma_2)$, with $\sigma_1 \leq \sigma_2$, using probability weights $\varpi \in (0,1)$ and $1 - \varpi$, respectively. This yields a leptokurtic distribution denoted by $\hat{\Phi}(\cdot)$.

Demand for each good j is given by

$$Y_t(j) = A_t(j)^{\epsilon-1} \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} Y_t \tag{8}$$

and the nominal price of the final good is

$$P_t = \left[\int_0^1 \left(\frac{P_t(j)}{A_t(j)} \right)^{1-\epsilon} dj \right]^{\frac{1}{1-\epsilon}}. \tag{9}$$

2.3. Differentiated good producers

A monopolistic firm j produces the corresponding differentiated good j with a constant-returns-to-scale technology

$$Y_t(j) = A_t \frac{N_t(j)^{\alpha} M_t(j)^{1-\alpha}}{A_t(j)}, \tag{10}$$

where $N_t(j)$ denotes labor input, $M_t(j)$ is the final good used as intermediate input, and A_t is a potentially time-varying aggregate productivity component. The idiosyncratic quality shifter $A_t(j)$ enters as an inverse productivity component, reflecting that higher-quality goods are more costly to produce.

The static cost minimization problem. In each period, firm j chooses labor and intermediate inputs to minimize costs

$$\min_{N_t(j), M_t(j)} \; W_t N_t(j) + P_t M_t(j)$$

subject to (10) for a given $Y_t(j)$, where P_t denotes the final (and intermediate inputs) good price. The first-order condition for intratemporal cost-minimization is:

$$\alpha P_t M_t(j) = (1-\alpha) W_t N_t(j). \tag{11}$$

This implies the following expression for nominal marginal costs:

$$MC_t(j) = \frac{A_t(j)}{A_t \alpha^{\alpha}(1-\alpha)^{1-\alpha}} w_t^{\alpha} P_t. \tag{12}$$

The dynamic pricing problem. Firm j chooses its nominal price $P_t(j)$ to maximize the expected discounted sum of profits subject to its downward-sloping demand given by (8). The firm incurs a fixed "menu cost" η measured in labor units each time it updates its price. In particular, firm j chooses in each period whether to update its price to the optimal nominal reset level $P_t^*(j)$ or to keep it unchanged as in the previous period, $P_{t-1}(j)$. Per-period nominal profits are

$$D_t(j) = \left[P_t(j) - (1-\tau_t) MC_t(j)\right] Y_t(j), \tag{13}$$

where τ_t is a proportional subsidy financed by lump-sum taxes, which follows an AR(1) process. The subsidy τ_t is the same for all differentiated good producers, and its steady-state level is such that it offsets the effect of market power in the differentiated goods market on average across firms. Innovations in τ_t are interpreted as "cost-push shocks". Taken literally, they capture variation in subsidies, taxes, or market power, that drives a wedge between marginal costs and prices. More generally, they also serve as a reduced-form representation of other unmodeled shocks and frictions, such as oil price shocks, international supply chain disruptions, or sectoral productivity shocks in the presence of wage rigidities (e.g. Blanchard and Galí 2007a, Rubbo 2023). For the central bank, these shocks generate a trade-off between stabilizing inflation and closing the output gap, defined as the gap between actual output and its efficient level.[7]

We express the firm's pricing problem in terms of the *price gap*,

$$x_t(j) \equiv \log P_t(j) - \log P_t^*(j) = p_t(j) - p_t^*(j)$$

where

$$p_t(j) \equiv \log\left(\frac{P_t(j)}{A_t(j) P_t}\right), \qquad p_t^*(j) \equiv \log\left(\frac{P_t^*(j)}{A_t(j) P_t}\right).$$

Thus, the price gap $x_t(j)$ is the log difference between firm j's actual price $P_t(j)$ and its optimal reset price $P_t^*(j)$.[8] Equivalently, we can express this gap as the difference

[7] The efficient level of output is the one that would prevail in an economy absent nominal frictions and monopolistic markup distortion.

[8] In general, $P_t^*(j) \neq P_t^f = \frac{\epsilon}{\epsilon-1}(1-\tau_t) MC_t(j)$, where P_t^f is the optimal flexible price. One way to approximate the pricing problem is to cast it as minimizing the losses from quadratic deviations of prices from P_t^f, which is valid under special assumptions on the persistence of shocks affecting $(1-\tau_t) MC_t(j)$. These assumptions are not met in our case due to the mean-reversion of our aggregate shock processes.

between the current and the optimal quality-adjusted relative prices, denoted by $p_t(j)$ and $p_t^*(j)$, respectively.

The random walk process (7) for firm-level quality shifters implies that the price gap is the firm's only idiosyncratic state variable (Midrigan 2011). In the aggregate steady state, with only idiosyncratic shocks, there is a non-degenerate distribution of optimal nominal reset prices $P_t^*(j)$, but the optimal *quality-adjusted* reset price $p_t^*(j)$ is constant and identical across firms. Along the perfect foresight transition paths, $p_t^*(j)$ varies but is still common to all firms. Therefore, we suppress the index j and interpret the index t as capturing the effects of all *aggregate* variables and disturbances.

The optimal pricing of a firm with price gap x_t follows a standard (S, s) rule: prices remain unchanged if $x_t \in [s_t, S_t]$, and are reset to p_t^* (so, $x_t = 0$) otherwise. If the firm keeps its nominal price unchanged, its price gap evolves according to

$$x_t = x_{t-1} - \varepsilon_t - \pi_t^*, \tag{14}$$

where

$$\pi_t^* \equiv p_t^* - p_{t-1}^* + \pi_t \tag{15}$$

is inflation in the quality-adjusted reset price. Let $V_t(x)$ denote the firm's end-of-period value function expressed in terms of the price gap, i.e., after the firm has adjusted its price or not. Any dependence of $V_t(x)$ on endogenous or exogenous aggregate variables is subsumed in the time index. The optimality conditions of the firm's pricing problem are

$$V_t'(0) = 0, \tag{16}$$

$$V_t(0) - \eta w_t = V_t(s_t), \tag{17}$$

$$V_t(0) - \eta w_t = V_t(S_t). \tag{18}$$

Eq. (16) ensures that $x = 0$ maximizes the firm's value, while equations (17) and (18) ensure indifference between adjusting and not adjusting at the trigger points s_t and S_t.

The end-of-period value function is

$$\begin{aligned} V_t(x) = \Pi_t(x) &+ \Lambda_{t,t+1} \int_{s_{t+1}}^{S_{t+1}} V_{t+1}(x')\, \hat{\phi}\left(x - x' - \pi_{t+1}^*\right) dx' \\ &+ \Lambda_{t,t+1} \left[1 - \int_{s_{t+1}}^{S_{t+1}} \hat{\phi}\left(x - x' - \pi_{t+1}^*\right) dx'\right] \left[V_{t+1}(0) - \eta w_{t+1}\right], \end{aligned} \tag{19}$$

which comprises the sum of current real profits $\Pi_t(x)$ and the discounted expected continuation value. This value accounts for two distinct scenarios: if the firm maintains its price at $t+1$, the resulting price gap x' is determined by a quality shock with density $\hat{\phi}(\cdot)$. Non-adjustment occurs when x' remains within the interval $[s_t, S_t]$. Conversely, if the firm opts to adjust, the gap is reset to $x' = 0$ at the expense of a menu cost ηw_{t+1}.

The real profit function $\Pi_t(x)$ for a firm with price gap x follows from the definition

of x, p_t^*, nominal profits in (13) and marginal costs in (12):

$$\Pi_t(x) = \left[e^{x+p_t^*} - \frac{(1-\tau_t) w_t^{\alpha}}{A_t \alpha^{\alpha} (1-\alpha)^{1-\alpha}} \right] \left(e^{x+p_t^*} \right)^{-\epsilon} Y_t. \tag{20}$$

Appendix B shows that the derivative of the value function at the optimum satisfies

$$\begin{aligned} V_t'(0) = \Pi_t'(0) + \Lambda_{t,t+1} \int_{s_{t+1}}^{S_{t+1}} V_{t+1}(x') \left. \frac{\partial}{\partial x} \hat{\phi}\left(x - x' - \pi_{t+1}^*\right) \right|_{x=0} dx' \\ + \Lambda_{t,t+1} \left[\hat{\phi}\left(-S_{t+1} - \pi_{t+1}^*\right) - \hat{\phi}\left(-s_{t+1} - \pi_{t+1}^*\right) \right] \left[V_{t+1}(0) - \eta w_{t+1} \right]. \end{aligned}$$

2.4. Distributional dynamics, aggregation, and general equilibrium

Price gap distribution. We now obtain expressions for the aggregate price level, price dispersion, the goods and labor markets clearing conditions. For this, it is useful to characterize the distribution of end-of-period price gaps, denoted $g_t(x)$, consisting of a continuous density $g_t^c(x)$ and a mass point (Dirac delta) at $x = 0$, g_t^0:

$$g_t(x) = g_t^c(x) + g_t^0 \delta(x), \tag{21}$$

where the continuous component evolves according to

$$g_t^c(x) = \int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1}) \hat{\phi}(x_{-1} - x - \pi_t^*) \, dx_{-1} + g_{t-1}^0 \, \hat{\phi}(-x - \pi_t^*) \tag{22}$$

for $x \in [s_t, S_t]$ and zero otherwise, and

$$g_t^0 = 1 - \int_{s_t}^{S_t} g_t^c(x) \, dx. \tag{23}$$

The first term on the right-hand side of (22) describes the evolution of the density of price gaps for firms that kept their nominal prices unchanged at $t-1$, while the second term captures the distribution of current price gaps for firms that did adjust. Outside the "inaction region" $[s_t, S_t]$ there is no mass because firms whose price gaps would fall outside this interval optimally reset in the current period, thereby creating the mass point at zero, given by (23).

Aggregate price level. Following (9), the *real* aggregate price level must integrate to one,

$$1 = \int_{s_t}^{S_t} e^{\left(x+p_t^*\right)(1-\epsilon)} g_t(x) \, dx. \tag{24}$$

Price dispersion. Aggregate price dispersion is

$$\Delta_t = \int_{s_t}^{S_t} \left(e^{-\epsilon(x+p_t^*)} \right) g_t(x) \, dx, \tag{25}$$

such that $\int_0^1 A_t(j)Y_t(j)\,dj = Y_t\Delta_t$.

Goods-market clearing. Goods-market clearing requires

$$Y_t = C_t + \int_0^1 M_t(j)\,dj,$$

so that total output equals consumption plus the use of intermediates in production.

Combining goods market clearing with the production technology (10) and cost minimization (11) and then using (25) gives

$$\frac{C_t}{Y_t} = 1 - \left(\frac{(1-\alpha)w_t}{\alpha}\right)^{\alpha}\frac{\Delta_t}{A_t}. \tag{26}$$

Labor-market clearing. Equilibrium in the labor market requires

$$N_t = \left[\frac{\alpha}{(1-\alpha)w_t}\right]^{1-\alpha}\frac{\Delta_t}{A_t}Y_t + \eta g_t^0, \tag{27}$$

so that total labor supply N_t equals the sum of labor used in production $\int_0^1 N_t(j)\,dj = \left[\frac{\alpha}{(1-\alpha)w_t}\right]^{1-\alpha}\frac{\Delta_t}{A_t}Y_t$, and labor employed in price adjustment, ηg_t^0. The latter term is the menu cost η multiplied by the fraction of firms that actually incur the cost, g_t^0.

Private equilibrium. We are now ready to define the private equilibrium:

DEFINITION 1 (Private Equilibrium). *Given an initial distribution of price gaps $g_{-1}^c(x)$, g_{-1}^0, a sequence of nominal interest rates $\{R_t\}_{t\geq 0}$, and aggregate shocks $\{A_t, \tau_t\}_{t\geq 0}$, a* private equilibrium *is a sequence $\{w_t, N_t, Y_t, C_t, V_t(x), S_t, s_t, p_t^*, \pi_t, \pi_t^*, g_t(x), g_t^c(x), g_t^0, \Delta_t\}_{t\geq 0}$ that satisfies the private equilibrium conditions: labor supply (3), the Euler equation (4), the definition of reset inflation (15), the conditions for the optimal pricing problem p_t^*, s_t, S_t in (16), (17), and (18), the definition of the value function $V_t(\cdot)$ in (19), and the distribution of price gaps $g_t(\cdot)$ in equations (21), (22) and (23), the aggregate price index P_t in (24), aggregate price dispersion Δ_t in (25), and the goods and labor market clearing conditions (26) and (27).*

Thus, the central bank has a single degree of freedom to choose the nominal interest rate R_t.

2.5. Aggregate exogenous disturbances

Aggregate productivity (in logs) and the subsidy (in levels) follow first-order autoregressive processes:

$$\log(A_t) = \rho_A \log(A_{t-1}) + \varepsilon_{A,t} \tag{28}$$

$$\tau_t - \tau = \rho_\tau(\tau_{t-1} - \tau) + \tau\varepsilon_{\tau,t} \tag{29}$$

where ρ_A and $\rho_\tau \in [0, 1)$ are persistence parameters, τ is the steady state subsidy, and $\varepsilon_{A,t}$ and $\varepsilon_{\tau,t}$ are one-off shocks.

2.6. Ramsey problem

The nominal interest rate is set by a benevolent central bank under full commitment. The problem of the central bank is to select paths for all equilibrium variables so as to maximize household welfare

$$\max_{\substack{\{w_t, N_t, Y_t, C_t, V_t(x), S_t, s_t, p_t^*, \\ \pi_t, \pi_t^*, g_t(x), g_t^c(x), g_t^0, \Delta_t, R_t\}_{t=0}^{\infty}}} \sum_{t=0}^{\infty} \beta^t \left(\log C_t - \chi N_t\right)$$

subject to the private equilibrium conditions defined above, initial conditions for $g_{-1}^c(x)$ and g_{-1}^0 and the transversality conditions, given some paths for the exogenous variables.

3. Inspecting optimal monetary policy in a simple model

To build intuition for how optimal monetary policy should be conducted during episodes of elevated inflation, this section derives our main results in a simplified version of the model. We concentrate on the difference between small and large shocks and contrast the conclusions with an analogously simplified Calvo framework.

Our analysis in this section is analogous to the familiar linear–quadratic textbook treatment of optimal monetary policy (e.g. Galí 2008). However, we avoid local approximations in order to preserve the model's nonlinearity. In Section 4, we calibrate our full model and solve it numerically, and show that the results presented here generalize and are quantitatively relevant for episodes of elevated inflation similar to the 2021–2023 period in the U.S.

3.1. A simple model

The simplified version deviates from the baseline in three ways: First, labor is the sole input in production, which eliminates roundabout production ($\alpha = 1$). Second, firm-specific shocks are normally distributed ($\varpi = 0$). Third, each period is divided into two sub-periods: *night* and *day*. The day is as in the baseline model. During the night, only firms are active and may freely reset their prices. Since firms reset prices freely at night, the distribution of price gaps collapses to a mass point at $x = 0$. Furthermore, firms choose daytime prices solely to maximize current-period profits, since any price gap remaining from the day will be costlessly eliminated overnight. The state-dependent nature of price adjustment is preserved: the number and identity of firms that change their prices continue to respond endogenously to aggregate and idiosyncratic shocks as well as to monetary policy. However, since firms solve a

sequence of static maximization problems, future expectations drop out entirely from the pricing decision.

This tractability allows us to cast the central bank's problem as a two-dimensional optimization problem with (i) welfare as a function of inflation and output gap only and (ii) a structural relationship between them – the Phillips curve. The equilibrium conditions of the simple model are simplified versions of those in the baseline model. We report them in Appendix A. For illustrative purposes, we calibrate the simple model using an analogous strategy to the baseline model to be discussed in Section 4.1. Importantly, the model generates a realistic increase in frequency under an inflation surge.[9]

3.2. Phillips curve

The central bank's choice set is given by the possible allocations consistent with a private equilibrium. We show first that these equilibria determine a relationship between inflation and the output gap: the Phillips curve. We characterize this relationship in terms of the output gap (output relative to the efficient allocation $\hat{Y} \equiv Y/A$) in the following proposition, dropping time subscripts for brevity in this section. Absent TFP shocks, output and output gap coincide.

PROPOSITION 1. *Private equilibria in the simple model can be characterized by a single static equation in inflation and output gap as follows*

$$1 = \int_s^S e^{(p)(1-\epsilon)} \frac{1}{\sigma} \phi\left(\frac{p+\pi}{\sigma}\right) dp + \left(\frac{\epsilon\,(1-\tau)}{\epsilon-1}\hat{Y}\right)^{1-\epsilon} \left[1 - \int_s^S \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right) dp\right], \qquad (30)$$

where s and S are implicit functions: $S(\hat{Y},\tau)$ and $s(\hat{Y},\tau)$ are the two roots of the equation $\left[\frac{\epsilon}{\epsilon-1}(1-\tau)\hat{Y}\right]^{1-\epsilon} - \left(\frac{\epsilon}{\epsilon-1}\right)^{-\epsilon}\left[(1-\tau)\hat{Y}\right]^{1-\epsilon} - \eta = e^{z(\hat{Y},\tau)(1-\epsilon)} - (1-\tau)\hat{Y}e^{z(\hat{Y},\tau)(-\epsilon)}$ *for* $z = s, S$.

PROOF. See Appendix A. □

Panel (a) of Figure 1 shows the Phillips curve in the calibrated simple model and compares it to the case of Calvo pricing. In both cases, the curves are increasing: under sticky prices, a policy easing that raises inflation also raises output. The key difference between the two models is that the Phillips curve is nonlinear in the menu cost model, while it is near-linear in the Calvo model. In particular, the menu cost Phillips curve is convex: higher inflation is associated with smaller increments in the output gap.

[9] The calibration of the simple model (see Table A3) shares the parameters of the baseline model described in Section 4.1, with three exceptions: First, we abstract from roundabout production and leptokurtic shocks ($\alpha = 1$ and $\varpi = 0$). Second, we assume that cost-push shocks have no persistence ($\rho_\tau = 0$). Third, the menu cost η and the standard deviation of idiosyncratic shocks σ are calibrated so that the model matches the frequency of price changes of 8.7% at zero inflation while producing an increase in frequency by 14 percentage points at 10% inflation. These two data moments are essential drivers of our key results. Matching them in a standard model with Gaussian idiosyncratic shocks implies an absolute size of price changes that falls well below that observed in the data, as emphasized by Blanco et al. (2024b). We also consider a Calvo version of the model with a Calvo parameter calibrated so that the Phillips curve has the same slope at zero output gap. This implies a probability of adjustment $(1-\theta)$ of 35.9%.

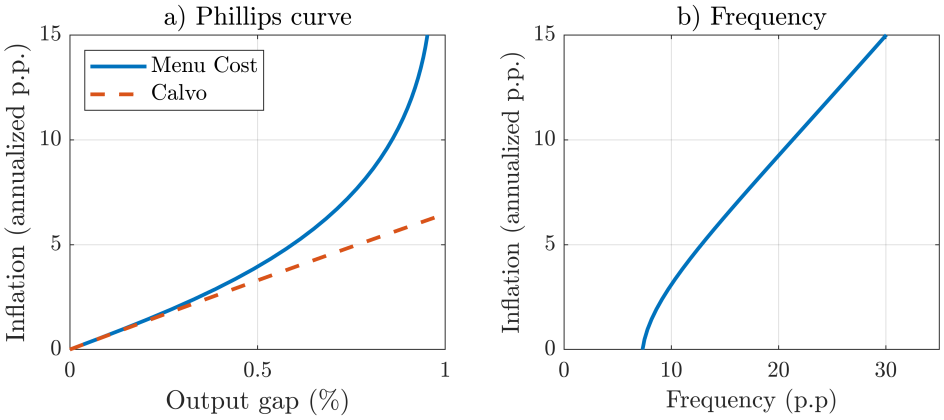

FIGURE 1. Phillips curve

Panel (a) plots the Phillips curve implicitly defined by equation (30), as well as the corresponding curve in the case of Calvo pricing. The output gap is $\log(\hat{Y})$ and annualized inflation is 12π. Panel (b) displays the mapping between frequency and inflation in the menu cost model.

Panel (b) of Figure 1 depicts the relationship between the frequency of price changes and inflation in the menu cost model. The frequency is constant by construction in the Calvo model. In the menu cost model, when inflation is low, the frequency of price changes remains close to its steady-state level. The economy thus behaves locally like a Calvo economy, as emphasized by Auclert et al. (2024).[10] However, as inflation increases, frequency rises as more firms decide to update their prices. This makes average prices more flexible, reducing the responsiveness of output to changes in inflation, thus steepening the Phillips curve.[11]

The nonlinear response of the frequency of price changes to inflation is an inherent feature of the menu cost model (See Appendix G for an illustration). To see why, first abstract from aggregate shocks. The firms follow an (S, s) pricing rule, and adjust their prices only if idiosyncratic shocks push their price gaps — the distance between their current and desired price — outside of the inaction thresholds. Now consider a small inflationary shock (See Figure A5, panel (a)). As inflation reduces the relative price of each firm, the shock shifts the entire price-gap distribution slightly. While this induces some firms to increase their prices, it also causes other firms to forgo price cuts. Because the density of firms near both inaction thresholds is similar, these two effects offset each other, leaving the total frequency of adjustments nearly constant. Now consider a large inflationary shock (See Figure A5, panel (b)). A significant shift in the distribution pushes a massive number of firms beyond the price-increase threshold. Simultaneously, the pool of firms considering price cuts is exhausted, meaning there

[10] For the similarity, the Calvo economy needs to be suitably calibrated, as we do in our exercises. In particular, the Calvo parameter — which determines the extent of price rigidity — needs to be adjusted to reflect the selection of large changes in the menu cost framework, which raises price flexibility.

[11] At very high frequencies the Phillips curve becomes backward-bending (not shown). At this point, monetary policy reaches its maximum effectiveness in stimulating activity, and any further inflation reduces output. However, such levels of inflation and frequency are fairly extreme. For the rest of the analysis, we restrict our attention to inflation levels that can be large, but not as large as to go beyond this point.

are no further "forgone cuts" to offset the surge in hikes. Consequently, total adjustment frequency increases. This transition from stable to accelerating frequency creates the characteristic nonlinear relationship between inflation and price flexibility observed in the data (Gagnon 2009; Karadi and Reiff 2019; Alvarez and Neumeyer 2020; Alexandrov 2020; Cavallo et al. 2024).

The Phillips curve describes the choice set of the policymaker when choosing inflation. Its *slope* reflects the state dependence of the inflation-output trade-off involved in monetary policy decisions: The inverse of the slope measures by how much the output gap must decline in order to reduce inflation by one percentage point. While in a low-frequency and low-inflation environment the slope of the Phillips curve coincides in the Calvo and menu cost models, it becomes steeper in the menu cost model once frequency and inflation increase,[12] leading to a decline in the output cost of disinflation. Its *location* depends on the cost-push shock τ: it shifts the Phillips curve $\pi(\log \hat{Y})$ leftwards or rightwards in parallel, as Appendix A shows.

3.3. Welfare

The central bank maximizes the household's welfare subject to the private equilibrium conditions. Since these constraints are static in the simple model, the central bank maximizes welfare by maximizing period utility. By combining the equilibrium conditions with the utility function, we first describe how the underlying distortions, namely misallocation and price-adjustment costs, affect utility. Second, we link these welfare distortions to the output gap and inflation in the simple model.

PROPOSITION 2. *Let $U - U^e$ be the central bank's utility gap relative to the utility under the efficient allocation, expressed in efficient-consumption-equivalent units. Let the welfare-relevant markup be the relative price of firm j divided by the welfare-relevant marginal cost: $\mu(j) = \frac{P(j)/P}{WRMC(j)}$, where $WRMC(j) \equiv wA(j)/A$. Then the utility gap can be expressed as a function of the average welfare-relevant markup ($\overline{\mu}$), the markup dispersion (ζ^{μ}), and price adjustment costs as*

$$U - U^e = \underbrace{\underbrace{-\log\overline{\mu} - \left(\frac{1}{\overline{\mu}} - 1\right)}_{\text{Average markup}} - \underbrace{\frac{1}{\overline{\mu}}(\zeta^{\mu} - 1)}_{\text{Markup dispersion}}}_{\text{Misallocation}} - \underbrace{\eta g^0}_{\text{Adjustment costs}}, \tag{31}$$

where the average markup is $\overline{\mu} \equiv \left(\int \mu(j)^{1-\epsilon} dj\right)^{\frac{1}{1-\epsilon}}$, the markup dispersion is $\zeta^{\mu} \equiv \int (\mu(j)/\overline{\mu})^{-\epsilon} dj$, and ηg^0 are the price adjustment costs in labor units.

PROOF. See Appendix A. □

[12] Blanco et al. (2024c) and Costain et al. (2022) also discuss how the slope of the Phillips curve steepens as inflation increases.

Proposition 2 shows that welfare costs are driven by two components: First, *misallocation* caused by the deviation of firms' relative prices from the welfare-relevant marginal costs. Second, labor is wasted to conduct price adjustments.[13] Misallocation can be further decomposed into terms driven by the average markup ($\overline{\mu}$) and the dispersion of markups (ζ^{μ}). The average welfare-relevant markup describes the *average* overproduction or underproduction, while the markup dispersion refers to the inefficient allocation of production across goods.

The utility gap and its components can be expressed as functions of inflation and output gap. This is analogous to the Calvo case, where welfare can also be expressed as a function of the output gap and inflation.[14] This is formulated in Proposition 3.

PROPOSITION 3. *In the simple model, the utility gap can be expressed as*

$$\begin{aligned} U - U^e \;=\; & \underbrace{\log(\hat{Y}) - (\hat{Y} - 1)}_{\textit{Average markup}} \\ & \underbrace{-\hat{Y}\left(\int_s^S e^{p(-\epsilon)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp + \left(1 - \int_s^S \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right)e^{p^*(-\epsilon)} - 1\right)}_{\textit{Markup dispersion}} \\ & \underbrace{-\eta\left[1 - \int_s^S \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right]}_{\textit{Adjustment costs}} \end{aligned} \tag{32}$$

where s, S and p^* are implicit functions of inflation.[15] *The utility function depends only on inflation and output gap: the average markup term in equation* (32) *depends only on the output gap, whereas the markup dispersion and adjustment costs terms depend only on inflation. TFP and cost-push shocks do not affect the utility gap.*

PROOF. See Appendix A. □

Figure 2 illustrates this decomposition for the menu cost model and contrasts it to the analogous decomposition in the Calvo model. Three results emerge. First, the relationship between average markups and the output gap is identical in both frameworks (panel b). Second, markup dispersion decreases with absolute inflation in the menu cost model while it increases in the Calvo model (panel c). Given the one-to-one mapping between markup and price dispersion, equations (14) and (15) imply that higher absolute inflation pushes all firms, and in particular the large mass of firms centered at zero price gap, away from their optimal price. In the menu cost model, a "selection effect" counters this inflationary shift: as inflation shifts firms away from the zero gap, it simultaneously triggers those with the largest misalignments to

[13] The welfare decomposition of Proposition 2 straightforwardly generalizes to the full model. It also applies to the Calvo model, in which case the last term (adjustment costs) is zero, by construction.

[14] A second-order approximation of welfare in the Calvo model is quadratic, $-\frac{1}{2}\left[\hat{y}^2 + \epsilon\left(\frac{\theta}{1-\theta}\right)\pi^2\right]$, where $\hat{y}$ denotes log deviations of $\hat{Y}$ from the zero-inflation steady state (Woodford 2003).

[15] $s(\pi)$, $S(\pi)$ and $p^*(\pi)$ solve the (S, s) band conditions, and the definition of the price level.

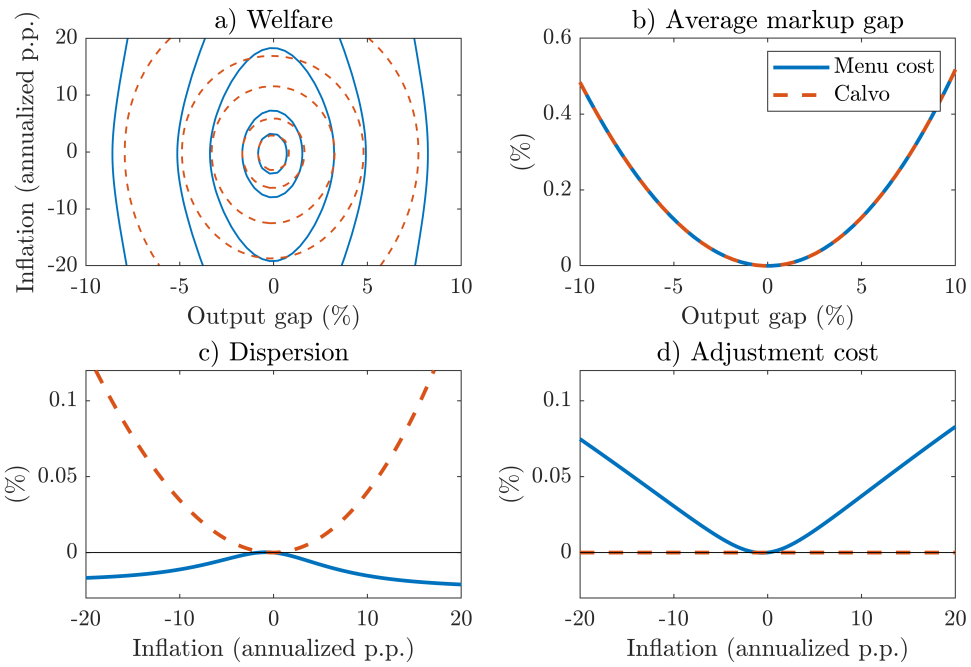

FIGURE 2. Welfare decomposition

Note: Decomposition of welfare differences according to equation (31). Welfare gaps are expressed in % of efficient consumption.

reset (see Figure A5 in the Appendix). Because firms jump back to the zero price gap upon adjustment, the concentration of firms at this optimal reset point is sufficient to compress the overall distribution. In contrast, the Calvo model lacks this selection mechanism: the probability of adjustment remains independent of a firm's position, meaning the mass of firms simply shifts further away from the zero gap as inflation rises, resulting in an unconditional increase in price dispersion. Third, as inflation rises, so does the fraction of adjusting firms in the menu cost model (see panel b in Figure 1). Since price changes are costly, the welfare cost due to price adjustment rises as inflation deviates from zero (panel d). In contrast, price changes are costless in the Calvo model.

When the welfare cost of markup dispersion and price adjustment are added up, a U-shaped relationship with inflation emerges (not shown). The relationship between welfare and inflation then resembles that in the Calvo model, which is also U-shaped. However, in the Calvo model, this relationship stems exclusively from the allocative inefficiency generated by price dispersion while in the menu cost model it stems from the use of resources for price adjustments.

Quantitatively, the welfare losses from nominal rigidities are somewhat smaller in the menu cost model (Burstein and Hellwig 2008; Blanco 2021; Blanco et al. 2024c). This is reflected in the different degrees of ellipticity of the iso-welfare curves in panel a of Figure 2.

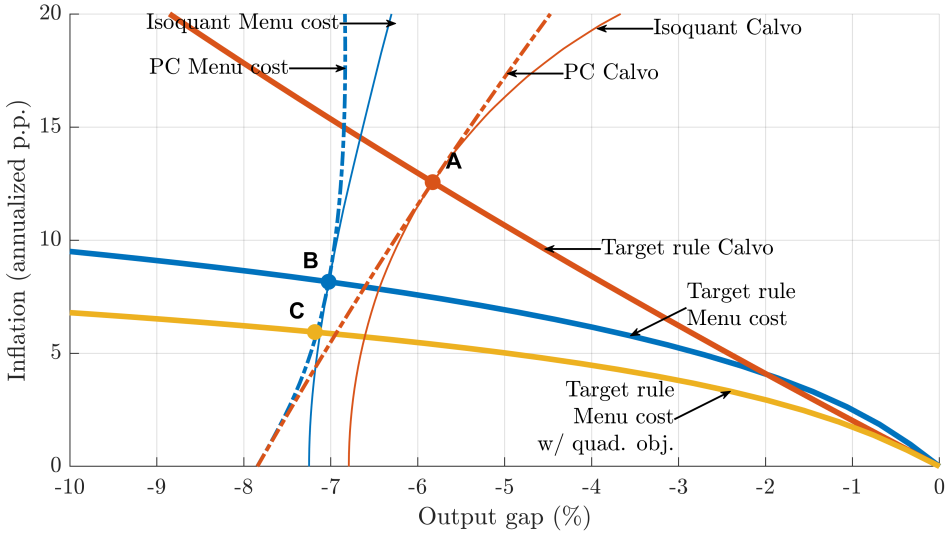

FIGURE 3. Optimal policy and the target rule

This figure combines the welfare functions from panel a in Figure 2 with the Phillips curves from panel a in Figure 1 to derive the target rule.

3.4. Optimal response to cost-push shocks: Strike while the iron is hot

We can now set up the central bank's problem as a static maximization problem in the two-dimensional output-gap-inflation space. The central bank chooses inflation π and output gap $\hat{Y}$ so as to maximize the objective (32) subject to the Phillips curve (30). We focus on the cost-push shock τ in this section.

Figure 3 represents the central bank's problem and its solution graphically. It shows the Phillips curve (PC, dashed lines) for a particular value of the exogenous cost-push shock τ, and the utility isoquant (thin solid lines) that is tangent to that PC. The optimal policy is defined by their tangency point, points A and B for Calvo and menu cost pricing, respectively. The "target rule" traces these points for different levels of the cost-push shock (thick solid line). It summarizes the values of inflation and output gap for different cost-push shock magnitudes under the optimal policy.

The target rule is near-linear under Calvo pricing (solid red line): inflation increases proportionally to the decline in the output gap. This is the consequence of the near-linear Phillips curve and a welfare function that is approximately quadratic. This result is well known in the literature (see, for example, Galí 2008): the central bank tolerates an increase in inflation, despite the misallocation it brings about via price dispersion, to partially cushion the adverse effect of a cost-push shock on the average markup, captured by a fall in the output gap.

Under menu costs, however, the target rule is nonlinear (blue). It is concave, which implies that, as the cost-push shock becomes larger, its impact on inflation is progressively attenuated under the optimal policy. The central bank's tolerance for inflation decreases in the magnitude of the shock; its stance is increasingly anti-inflationary – it *strikes while the iron is hot*.

This raises the question of why. First, as established, the relative welfare losses from inflation are lower in the menu cost model than in Calvo. This induces the central bank to tolerate *more* inflation in the menu cost model for small shocks, when, by construction, the slope of the Phillips curve is similar to the Calvo model. This is evident in Figure 3, where the target rule of the menu cost model lies above the Calvo counterpart for annualized inflation rates up to 4%. Second, the inherent nonlinearity of the Phillips curve in the menu-cost framework significantly reduces the output cost of disinflation following large shocks. This change in the trade-off drives the increasingly aggressive anti-inflationary stance of optimal policy as shock magnitudes rise. The latter flexibility effect quickly dominates the initial motive to tolerate higher inflation; consequently, the menu-cost target rule drops below the Calvo rule for inflation rates exceeding 4%, diverging sharply by the time inflation reaches 10%.

To confirm the primary role of the nonlinearity of the Phillips curve in driving the nonlinearity of the target rule, we compute the optimal policy assuming that the central bank faces the Phillips curve of the menu cost framework, but it has the welfare objective of the Calvo model. That is, we look for the tangency point of the dashed blue Phillips curve with the corresponding solid red iso-welfare curve (not shown) in Figure 3. This point is marked by "C". The yellow line represents the locus of these tangency points across all magnitudes of cost-push shocks. It defines the target rule for a hypothetical central bank that maintains a Calvo-based welfare objective while operating within a menu-cost environment. Although this "Calvo-objective" rule is flatter — reflecting the relatively higher inflation aversion inherent in that framework — its concavity remains similar to the baseline menu-cost rule. This leads us to the conclusion that the nonlinearity of the target rule is driven by the convexity of the Phillips curve, rather than by the shape of the objective function.

Let us restate this result relating it to the underlying mechanism. For small shocks, the change in the frequency of price adjustment is negligible, so the logic of the Calvo framework applies: the central bank tolerates some inflation to partially cushion the decline in the output gap. As inflation rises, however, the frequency of price adjustment increases and prices become more flexible. Consequently, the Phillips curve steepens and the output cost of disinflation falls. As avoiding inflation becomes cheaper, the central bank becomes increasingly reluctant to tolerate it — even though it is generally somewhat less costly than in the Calvo model. Hence, following a large cost-push shock that raises the frequency of price adjustment, the central bank stabilizes inflation more aggressively relative to the output gap than after small shocks.

We can equivalently interpret the result in terms of the effectiveness of monetary policy. As the Phillips curve steepens, the *transmission of monetary policy to real activity* weakens: achieving a given positive effect on the output gap would require a substantially larger increase in inflation. In light of this decline in effectiveness, the central bank becomes increasingly reluctant to use inflation to stimulate output.

To understand why this outcome is significant, consider the standard linear-

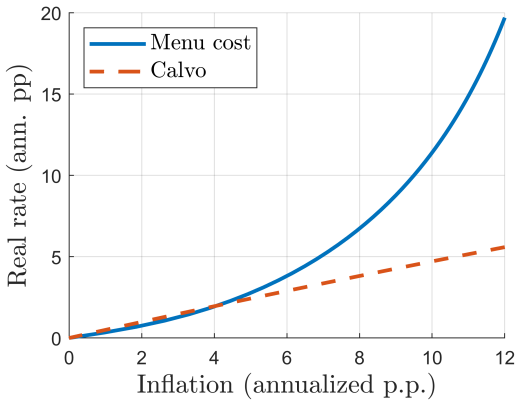

FIGURE 4. Strike while the iron is hot: real rate tightening

The figure displays the relationship between inflation and the real interest rate under optimal policy and different cost-push shock sizes.

quadratic Calvo framework. In that model, changes in repricing frequency do not alter the optimal target rule because the objective function and the Phillips curve change in perfectly offsetting ways. Specifically, while a higher frequency of price adjustments steepens the Phillips curve, it simultaneously reduces the welfare cost of inflation. This is because greater price flexibility limits the impact of inflation on price dispersion, the primary source of allocative inefficiency in that model (Galí 2008). These two effects – a steeper trade-off and a lower relative welfare weight on inflation – neutralize each other, such that the slope of the target rule depends solely on the elasticity of substitution ($-1/\epsilon$) and is invariant to the adjustment frequency.

In our menu-cost environment, however, the relative stability of the objective function prevents this neutralization. Because the welfare weights do not decline to offset the surge in price flexibility during large shocks, the nonlinearity of the Phillips curve is allowed to drive the optimal policy response. We thus conclude that the stability of the welfare objective is just as critical to the nonlinearity of optimal policy as the nonlinearity of the Phillips curve.

The central bank implements the optimal policy in terms of inflation and output gap by setting the nominal interest rate appropriately. Because the Euler equation (5) links the output gap to the real interest rate, the target rule can also be expressed as an interest rate rule. In the simple model with an i.i.d. cost-push shock, the Euler equation and goods market clearing imply the static relation $c - \bar{c} = \hat{y} = -(r^r - \bar{r}^r)$, so that the real rate moves inversely with the output gap. The interest rate rule is therefore the mirror image of the target rule in Figure 3. Figure 4 illustrates this relationship by plotting inflation against the real interest rate under optimal policy for different sizes of a cost-push shock. When inflation is high, the central bank sets the nominal rate so as to achieve a more-than-proportionate increase in the real rate. In other words, when firms adjust prices more frequently, the policy stance responds more aggressively to inflation. This provides a complementary way to express the "strike while the iron

is hot" prescription in terms of interest rates. We will return to this figure when we discuss the implementation of optimal policy in Section 4.5.

3.5. Optimal response to TFP and demand shocks: A divine coincidence result

Next, we turn to the optimal response to other shocks. First, we consider total factor productivity shocks, which affect the efficient allocation. In the standard New Keynesian model with Calvo prices — also in its static version — the optimal response to such shocks features strict price stability with both inflation and the output gap stabilized at zero. This is commonly known as the "divine coincidence" (Blanchard and Galí 2007b). A version of the divine coincidence also holds in our economy.

PROPOSITION 4. *Under optimal policy, inflation and the output gap do not respond to TFP shocks.*

PROOF. See Appendix A. □

As in the standard Calvo model, TFP shocks affect neither the Phillips curve (30) nor the objective function (32). Hence the optimal inflation and output gap levels are not affected either. Indeed, the central bank ensures that both inflation and the output gap remain constant, and consequently the frequency of repricing as well as the price gap distribution stay constant. The rate of inflation that is optimal absent shocks thus simultaneously minimizes inefficient output fluctuations (the average markup gap) and the costs of nominal rigidities (markup dispersion and adjustment costs) also under TFP shocks.[16] Notice that — unlike for cost-push shocks — the shape of the Phillips curve plays no role in this result, which therefore holds equally for small and for large shocks.

The divine coincidence also extends to demand shocks in the form of time-preference shocks – which could be introduced into our model by replacing parameter β with an exogenous AR(1) shock process. This modification only affects the Euler equation (4), which is not a binding constraint for the planner. Therefore, time-preference shocks do not affect the optimal-inflation schedule in Figure 3.

4. Optimal monetary policy in the full model

We now turn to the full model. The key difference relative to the simple model is that dynamics now play a role, as agents are forward-looking, and the distribution of prices is a state variable. Furthermore, roundabout production and fat-tailed shocks allow the model to be brought closer to the data quantitatively. We analyze *timeless* Ramsey

[16] This optimal inflation rate is not exactly zero due to the model's nonlinearities, even if it is very close to zero in our calibration. We discuss optimal steady state inflation for the baseline model in Appendix D.1.

TABLE 1. Calibration

Param	Value	Description	Reference
		Households and final producer	
β	$0.96^{1/12}$	Discount rate	Golosov and Lucas (2007)
χ	1	Utility weight on labor	Set so that $w = C$
ϵ	6	Elasticity of substitution	Gagliardone et al. (2025)
		Differentiated goods producers	
α	0.2	Labor share	Nakamura and Steinsson (2010)
η	0.064	Menu cost	Calibrated to match
σ	0.0306	Std dev. of quality shocks	moments of steady state
σ_1/σ_2	0.065	Ratio of low vs. high stdev	price change distribution
ϖ	0.934	Share of low volat. shock	at 2% inflation
τ	0.1672	Steady-state subsidy	Set so average markup is zero
		Aggregate shocks	
ρ_A	$0.95^{1/3}$	TFP shock persistence	Smets and Wouters (2007)
ρ_τ	$0.9^{1/3}$	Cost-push persistence	Smets and Wouters (2007)

policy (Woodford 2003). [17] In doing so, we extend the static analysis of Section 3 to a dynamic environment.

4.1. Calibration and computation

We calibrate the model at a monthly frequency. Table 1 summarizes the parameter values employed. Table 2 contrasts some relevant moments with analogous moments in the data.

Household Preferences and Market Structure. We set the household discount factor, $\beta = 0.96^{1/12}$, corresponding to an annual real interest rate of 4%, as in Golosov and Lucas (2007). We assume the relative welfare weight of labor $\chi = 1$, so the real wage equals consumption $w_t = C_t$ in equilibrium. We target a pre-subsidy steady-state gross markup of 1.2 (e.g., Gagliardone et al. 2025), implying an elasticity of substitution among goods of $\epsilon = 6$. We then set the subsidy τ to offset the average markup in the

[17] The timeless policy corresponds to the optimal monetary policy response to shocks once the potential time-inconsistency issues have disappeared. It is computed assuming that all states and all backward-looking Lagrange multipliers start at their respective values in the Ramsey steady state (RSS) when a shock arrives. The RSS corresponds to the steady state with only idiosyncratic shocks and with trend inflation set at the optimal long run level determined by the Ramsey planner. We defer the analysis of the steady state of the Ramsey problem and of the time inconsistency of optimal policy to Appendix D. There, we show that the optimal policy is virtually time consistent, provided that the subsidy τ offsets the average markup distortion in the steady state, as in our baseline calibration.

Ramsey steady state, which implies that output is at its efficient level.[18]

Technology and Firm-Level Shocks. We introduce fat-tailed idiosyncratic shocks. Such a distribution has been shown to help the model capture both the fat tails of the steady state price change distribution (Midrigan 2011), and the magnitude of frequency increase after large shocks (Karadi and Reiff 2019), which is impossible in simple models with Gaussian distribution (Blanco et al. 2024a). We choose a flexible parametric form, a mixture of two Gaussian distributions. The distribution is determined by three parameters: the standard deviation of the low and high volatility distributions (σ_1 and σ_2, respectively) and the probability of drawing from the low-volatility distribution (ϖ). We denote by σ the standard deviation of the overall distribution. Along with the menu cost η, these parameters are jointly calibrated to match four key micro-moments of U.S. price adjustment. The first three moments are the monthly frequency of price changes of 8.7% (Nakamura and Steinsson 2008), average absolute size of price changes of 8.5% (Nakamura and Steinsson 2008) and kurtosis of price changes of 4 (Alvarez et al. 2016), all three at the steady state under a 2% trend inflation. The fourth data moment is a 14 percentage point increase in the frequency of price changes, which is the difference between the peak frequency in 2022 (around 22.7%) (Montag and Villar 2025) and the steady state frequency (8.7%). To capture this moment in our model, we measure the frequency increase associated with a monetary policy shock that drives inflation up to 10% from an (approximately optimal) 0% inflation under a conventional Taylor rule ($\log(R_t/R) = 1.5\pi_t + 0.5\log(Y_t/Y)$). Jointly matching the four moments is an indication of the validity of the underlying framework, because our numerical explorations reveal that the three steady state moments (frequency, size, kurtosis) determine a narrow range of possible values for the fourth target statistic (frequency increase). Our empirical target falls within this range. The calibration yields a menu cost $\eta = 0.064$ and a shock process defined by $\sigma = 0.0306$, $\varpi = 0.934$, and $\sigma_1 = 0.065\sigma_2$.

Firms' production technology exhibits a roundabout structure as in Basu (1995). We calibrate the labor input share to $\alpha = 0.2$ in the baseline model, which is within the range 0.1-0.5 found reasonable by Nakamura and Steinsson (2010). As emphasized by that paper, the parameter flattens the Phillips curve through its impact on the strength of strategic complementarities among firms. Under our calibrated values, the slope of the Phillips relationship[19] (0.08) is within the range of estimates (0.006-0.106) reported by Hazell et al. (2022). We calibrate the corresponding Calvo model to match the same slope. Under the same intermediate-input share ($\alpha = 0.2$), this requires a monthly adjustment probability ($1 - \theta$) of 30% – equivalent to a textbook Calvo model without intermediate inputs ($\alpha = 1$) that has a monthly adjustment probability of 15%.

[18] In the standard Calvo model, this means $\tau = 1/\epsilon$. In our model the steady state features price dispersion due to the idiosyncratic shocks, which implies a slightly different τ.

[19] We define the Phillips relationship as the relationship between impact inflation and impact output gap under i.i.d. monetary policy shocks and a conventional Taylor rule. Around the steady state its slope closely approximates that of the linearized New Keynesian Phillips curve.

TABLE 2. Moments

	Steady state at 2% inflation			10% inflation	Phillips curve
	Frequency	Size	Kurtosis	Freq. incr.	slope
Data	8.7%	8.5%	4	14 pp	0.006-0.106
Baseline	8.7%	8.5%	4	14 pp	0.08

The calibration targets are from Nakamura and Steinsson (2008) (frequency and size), Alvarez et al. (2016) (kurtosis), Montag and Villar (2025) (frequency increase), and Hazell et al. (2022) (Phillips curve slope).

Aggregate Shocks. Finally, we calibrate the persistence of aggregate shocks to match quarterly estimates from Smets and Wouters (2007), adjusted to a monthly frequency. The persistence of the TFP shock is set to $\rho_A = 0.95^{1/3}$. Similarly, the persistence of the cost-push shock is set to $\rho_\tau = 0.9^{1/3}$.

Computation. The main challenge in solving the optimal monetary policy problem presented in Section 2.6 is that the central bank's problem contains infinite-dimensional choice variables $(g(x), V(x))$. To tackle this difficulty, we solve this problem with a novel nonlinear algorithm. The core idea is to represent it as a high- but finite-dimensional optimization problem in which the Ramsey planner chooses sequences of all equilibrium variables under perfect foresight. The algorithm has five ingredients: (i) approximation of the value and distribution functions by piecewise linear functions on a grid for x; (ii) an endogenous time-varying grid for x to precisely cover the inaction region; (iii) analytical integral evaluation over x; (iv) symbolic differentiation for the derivation of first order conditions; and (v) a Newton solver for the resulting high-dimensional set of non-linear equilibrium conditions. This algorithm builds on González et al. (2024), whose general approach we extend to discrete time and tailor to the context of models with inaction regions (steps i-iii).[20] Appendix H presents the details.

The Ramsey problem as defined in Section 2.6 is continuous and differentiable – even though the individual firm's pricing rules are not. This is so because each firm is atomistic with zero mass, and thus the discontinuity in a single firm's behavior does not lead to a discontinuity in aggregates. Our algorithm maintains this differentiability (steps i and ii), which sets it apart from other common approaches to solve for aggregate dynamics in menu cost models.[21] Therefore we can characterize optimal policy by first order conditions (step iv).

[20] The algorithm achieves accuracy of order 10^{-8} and the results are robust to further refinement beyond the baseline 300 periods and 35 gridpoints. The random menu cost model in the robustness section requires 225 gridpoints for stability.

[21] Note that $V_t(x)$ and $g_t^c(x)$ are continuously differentiable with respect to x over the relevant range (s_t, S_t). In the numerical implementation, we approximate these functions using a linear spline (step i), which is continuous but not differentiable. However, the approximated problem remains continuously differentiable, because the spline approximations enter the model only through integrals.

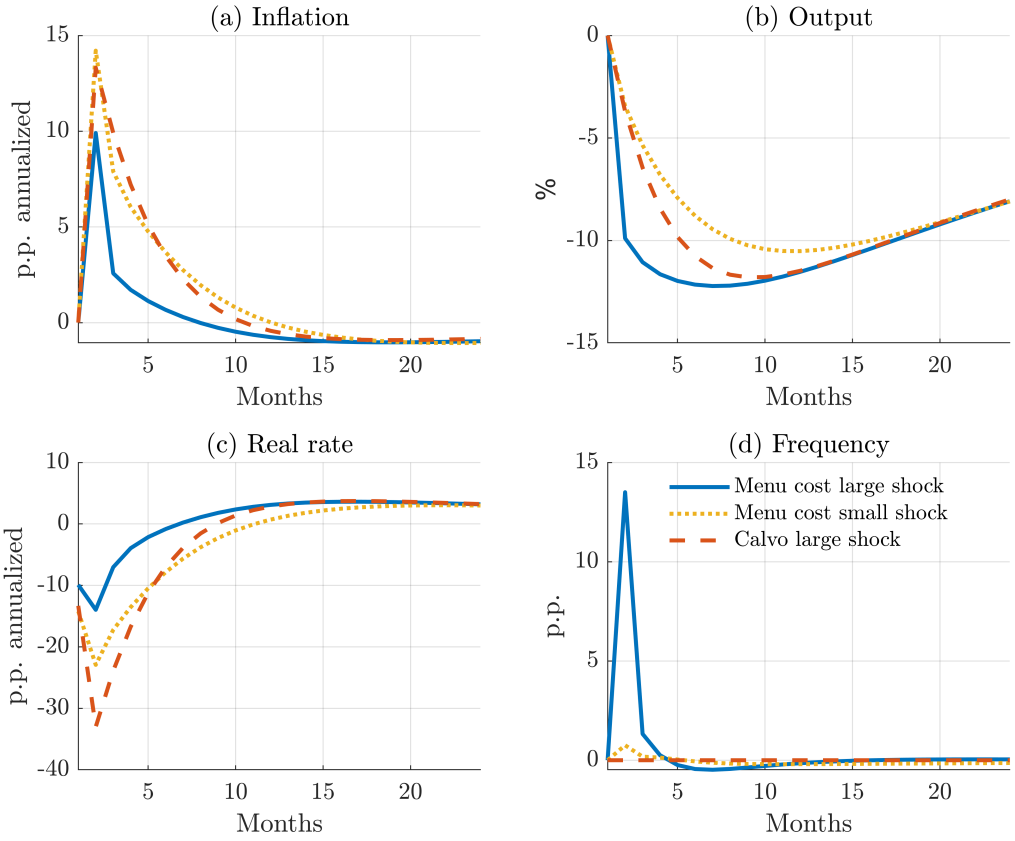

FIGURE 5. Impulse responses to a cost-push shock under the optimal monetary policy

The figure shows impulse responses in deviations from steady state to a large cost-push shock in the baseline menu cost model (blue solid line); and it contrasts the responses with those of a linearly rescaled small cost-push shock in the baseline model (yellow dotted line) and a large cost-push shock in the Calvo model (red dashed line).

4.2. Nonlinear optimal monetary policy response to cost-push shocks

We begin with cost-push shocks, presenting impulse responses that track their dynamic effects. Figure 5 shows the impulse responses to a large cost-push shock τ_t (blue solid line) in the baseline model under optimal policy starting from the Ramsey steady state, and contrasts them with impulse responses to a small cost-push shock (yellow dotted line); and to the same large cost-push shock in the Calvo model (red dashed line).[22] The small shock is linearly scaled to make it comparable with the large shock. The size of the large shock is calibrated to generate a 10% inflation surge, which is broadly in line with the inflation rate observed during the 2021–2023 inflation surge in the U.S. (Montag and Villar 2023). The large shock implies an increase in the effective marginal

[22]The impulse responses are computed nonlinearly under perfect foresight. For small shocks, this is equivalent to the first-order approximation to the stochastic problem, as discussed by Boppart et al. (2018). For large shocks, its interpretation is similar to that in Cavallo et al. (2024): an unexpected once-and-for-all large shock that hits the economy in the deterministic steady state.

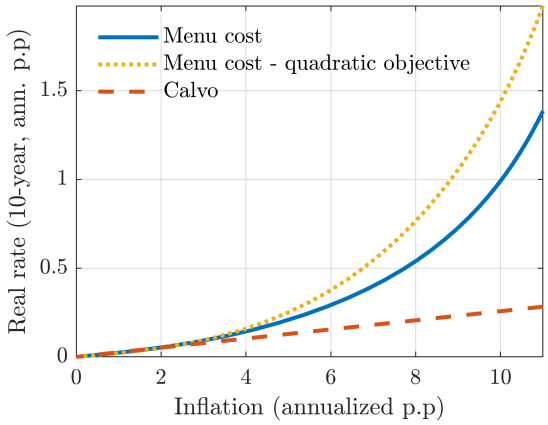

FIGURE 6. Strike while the iron is hot

The figure displays the relationship between impact inflation and annual 10-year real interest rate (in deviations from steady state) under optimal policy and cost-push shocks of different sizes.

cost of 3.5% on impact.[23] In general equilibrium, wage adjustments dampen this effect such that marginal costs increase by only 0.9%.

The central bank optimally tolerates a temporary inflation increase (panel a) to partially cushion the decline in output (panel b), as it did in the simple model.[24] Eventually, inflation turns negative such that the price level returns to its initial level in the long run. The optimal policy implies a temporary decline in the real interest rate alongside the surging inflation, but entails a commitment to a persistently tight policy in the future.

Optimal monetary policy in the menu cost model is nonlinear, as the difference between the impulse responses to the large shock and the linearly scaled small shock shows. After the large cost-push shock, which raises the frequency of price changes (panel d), the optimal policy prescribes a more contractionary real interest rate path (panel c) than after a linearly scaled small shock, when frequency is basically unchanged. This tighter policy leads to a more muted increase in inflation (panel a) and a stronger drop in output (panel b). In sum, the central bank is more anti-inflationary for the large shock than for the small shock, similarly to what we saw in the simple model in Section 3. That is, the central bank strikes while the iron is hot.

This pattern is not specific to the two shock sizes illustrated above. Figure 6 illustrates the relationship between inflation and the 10-year real interest rate deviations under optimal policy for cost-push shocks of different sizes. The 10-year rate is defined here as the average real interest rate over a 10-year horizon. It summarizes the policy commitment in a single statistic.[25] It can thus be interpreted as a summary measure of

[23] The small shock is calibrated to be in the linear region of the model (a 0.03% increase in marginal cost on impact), and the impulse responses are rescaled by the ratio of the two shocks.

[24] For cost-push shocks, output equals the output gap, as this type of shock yields no variation of efficient output.

[25] The real rate response almost fully dissipates after ten years, so the 10-year real rate closely approximates the cumulative real rate deviation over the life of the shock. Unlike in the simple model, the

the "policy stance". When inflation is high, the central bank adjusts the policy stance more than proportionally to inflation, as it did in the simple model (cf. Figure 4).

Intuition. In the baseline model, there is no longer a structural relationship between contemporaneous inflation and the output gap.[26] Nor can the welfare function be expressed solely as a function of these two variables. Nevertheless, it is useful to frame the central bank's problem in terms of a constraint – like the Phillips curve – and an objective – welfare – and to apply the intuition developed in the simple model. The separation is not only heuristic; we can numerically decompose the contributions of (i) the objective and (ii) the constraints to our results in the full model, just as we have done in the simple one.

We begin by examining the central bank's constraint. To illustrate it, we plot the relationship between output and inflation on impact following transitory monetary policy shocks of different sizes under a Taylor rule without interest-rate smoothing, starting from the Ramsey steady state (panel a of Figure 7). We refer to this as the *Phillips relationship*. As in the simple model (cf. Figure 1), this relationship is upward-sloping and convex: the Phillips relationship steepens, that is, the output costs of disinflation decline as inflation rises. As in the simple model, this is because the frequency of price adjustments eventually increases as inflation accelerates (see panel b).

This suggests that the nonlinearity of the Phillips relationship, as opposed to any change in the welfare function, may be the reason behind the nonlinearity of the optimal policy also in the full model. To confirm this hypothesis, we need to abstract from the role of the welfare objective. Therefore, we proceed as in Section 3.4 and compute the Ramsey policy combining the dynamic menu cost framework with the quadratic approximation of the welfare objective in the Calvo model. That is, we replace the true period welfare function by the quadratic loss function of a Calvo economy with roundabout production, which takes the familiar form $\hat{y}_t^2 + \lambda\pi_t^2$ where $\lambda = \frac{\theta}{1-\theta}\frac{\epsilon}{\alpha(2-\alpha)}$ (this is a special case of the loss function in Rubbo 2023). The results are shown in Figure 6 (yellow dotted line). As in the simple model, the policy response is similar to that under the true objective, and even more nonlinear than in the baseline. This confirms that the key reason behind the nonlinearity of the stance-inflation relationship is the nonlinearity of the Phillips relationship.

Quantitative importance. How big are the differences between the near-linear Calvo model and our menu cost model? Are the nonlinearities of the menu cost model sufficiently strong to matter quantitatively? The slope of the relationship between policy stance and inflation under optimal policy in Figure 6 is almost indistinguishable

relationship between stance and inflation depicted in Figure 6 is no longer structural and should rather be interpreted as descriptive.

[26] Rather, the relationship between inflation and the output gap is a multidimensional one, described by a block of dynamic equations that depend on state variables and expectations about future variables. The same is true in the nonlinear Calvo model: the familiar Phillips curve emerges only after linearization. A similar argument applies to the welfare function.

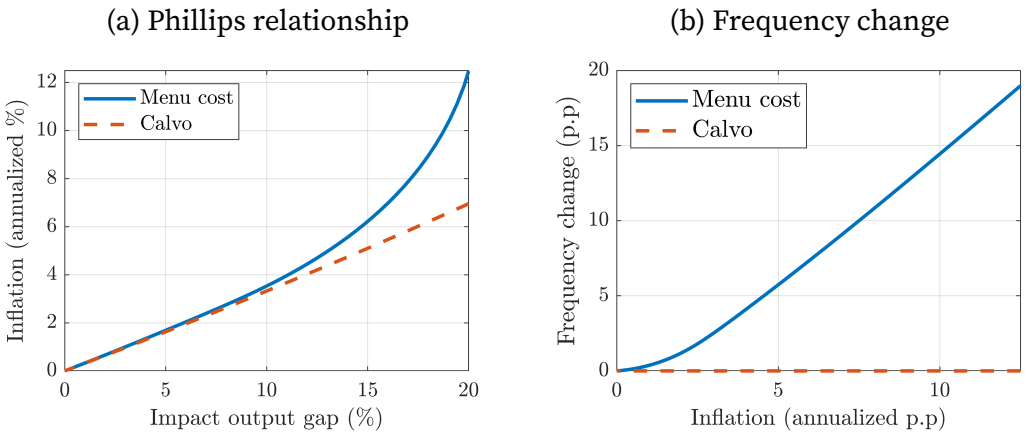

FIGURE 7. The Phillips relationship and frequency

Panel (a) of the figure displays the relationship between inflation and the output gap on impact under a Taylor rule $\log(R_t/R) = 1.5\pi_t + 0.5\log(Y_t/Y)$ and monetary policy shocks of different sizes, starting from the steady state. Panel (b) shows the relationship between inflation and frequency change, on impact.

at zero inflation from that under Calvo, which is given by $\frac{1}{120}\frac{\epsilon}{2-\alpha}$ up to first order.[27] The Calvo model thus delivers a good approximation of optimal policy for low levels of inflation.

However, the nonlinearity of the menu cost model becomes quantitatively significant quite quickly as inflation rises. At 10% inflation, for example, the slope of the Phillips relationship is about 4.6 times higher than under Calvo pricing (Figure 7). At the same inflation level, the optimal policy response to a cost-push shock is much more restrictive, almost 75 basis points higher than the 25 basis points response under Calvo pricing (Figure 6). Thus, while at low inflation levels the Calvo model is a sufficiently good approximation of the menu cost model, this equivalence breaks down at inflation levels such as those seen in 2022, when inflation reached approximately 10% and frequency more than doubled.

4.3. Optimal monetary policy response to TFP shocks

Next, we consider TFP shocks. As in the simple model, a version of the divine coincidence holds in our baseline economy as well.[28]

PROPOSITION 5. *The timeless Ramsey policy keeps inflation constant at the steady state level in response to aggregate TFP shocks.*

PROOF. See Appendix C. □

[27] This coefficient can be derived using the second-order approximation of welfare and the first-order approximation of the Phillips curve (taken from Rubbo (2023) for the case of a single-sector roundabout economy with log-linear utility), and using equation (5) to convert output into the real rate. The factor 1/120 accounts for the annualization of inflation and for the averaging of the real rate.

[28] Nakov and Thomas (2014) show that the divine coincidence holds in a representative-firm state-dependent pricing model for TFP shocks. We extend this result to a setting with firm heterogeneity, in which the adjustment frequency responds smoothly to the shock.

The level of inflation in the Ramsey steady state is not exactly zero — unlike under Calvo pricing — but is very close to it.[29]

Unlike in the simple model, the divine coincidence does not hold exactly for demand shocks like time-preference shocks. However, deviations from divine coincidence are quantitatively small, as Figure A3 in Appendix E.2 shows. Thus, the divine coincidence holds approximately.

4.4. Robustness and sensitivity analysis

We now assess the robustness of the nonlinear optimal monetary policy in alternative state-dependent price setting models and parameterizations. First, we contrast optimal policy in the CalvoPlus model (Nakamura and Steinsson 2008) and the random menu cost model (Dotsey et al. 1999; Gagliardone et al. 2025) with our baseline model. Second, we assess the quantitative importance of two assumptions, the fat-tailed idiosyncratic shocks and the roundabout economy, for our results. Third, we show that results in the simple model remain robust in an analogously calibrated full model, including quantitatively.

Alternative state-dependent models. We start by examining the robustness of our results in two alternative popular state-dependent price-setting models — the CalvoPlus model and the random menu cost (MC) model.

The CalvoPlus and the random menu cost models are variations of the baseline menu cost model, where the menu cost η is stochastic. In the CalvoPlus model, price adjustment is free with probability $1-\theta$, as in the Calvo (1983) model, and with probability θ the menu cost takes a fixed positive value (η). In the case of the random menu cost model, price adjustment is free with probability $1-\theta$; otherwise the menu cost is drawn from the uniform distribution $U[0,\eta]$. To facilitate comparison with our baseline model, we introduce roundabout production and fat-tailed idiosyncratic shocks to these models.

The extensions give us a single extra parameter: the probability of free price changes, $1-\theta$. We set this parameter to one-half of the steady state frequency, which is a similar share as in Gagliardone et al. (2025). We then calibrate four parameters: (i) the menu cost in the CalvoPlus and the maximum menu cost in the random menu cost model, (ii) the standard deviation of idiosyncratic shocks, (iii) the share of small idiosyncratic shocks, and (iv) the relative standard deviation of small and large shocks to match four moments: frequency and the kurtosis of price change distribution and the size and interquartile range of the absolute price change distribution. Notice that the fourth target differs from that used for calibrating our baseline model (frequency at 10% inflation). We adopt this alternative target because, given the value of θ, it is impossible to simultaneously match the three steady-state moments and the frequency

[29] This is due to the asymmetry of the profit function (20). See Appendix D.

TABLE 3. Moments of alternative state-dependent models

	Steady state at 2% inflation				10% inflation		Phillips curve
	Frequency	Size	Kurtosis	IQR	Freq. incr.	Shock size	slope
Data	8.7%	8.5%	4	11%	14 pp		0.006-0.106
Baseline	8.7%	8.5%	4	7%	14 pp	3.5%	0.08
CalvoPlus	8.7%	8.5%	4	11%	4 pp	4.0%	0.036
Random MC	8.7%	8.5%	4	11%	4 pp	5.3%	0.04

The table shows the ability of alternative state-dependent models to capture moments of the price change distribution. The table shows that the CalvoPlus and the random menu cost models can match a fourth moment, the interquartile range (IQR) of the steady state absolute price change distribution, but significantly underestimate the extent of frequency increase at 10% inflation (4 pp in both cases instead of 14 pp in the data and the baseline). The calibration targets are from Nakamura and Steinsson (2008) (frequency and size), Alvarez et al. (2016) (kurtosis), Montag and Villar (2025) (IQR and frequency increase), and Hazell et al. (2022) (PC slope).

at 10%.[30] Table 3 contrasts the ability of these models to match data moments. The calibrated parameters are listed in Table A1 in the Appendix.

$$\text{CalvoPlus: } \tilde{\eta} = \begin{cases} \eta & \text{w/ prob } \theta \\ 0 & \text{w/ prob } 1-\theta \end{cases} \qquad \text{Random MC: } \tilde{\eta} = \begin{cases} U[0,\eta] & \text{w/ prob } \theta \\ 0 & \text{w/ prob } 1-\theta \end{cases}$$

These frameworks are tractable extensions of the fixed menu cost model. They are proposed, first, because they raise the real effects of monetary policy by reducing the endogenous selection of large price changes. Indeed, they generate a lower Phillips curve slope than our baseline. However, the Phillips curve slope is affected by both the extent of nominal stickiness and the degree of strategic complementarities between prices. Due to high strategic complementarities caused by a realistically high intermediate inputs share ($1-\alpha = 0.8$), the slope is already quite low and an additional change in nominal stickiness relative to our baseline has only a muted additional impact (Gertler and Leahy 2008). Second, these models can capture a wider range of moments of the steady-state price-change distribution. As Table 3 confirms, the models can indeed better capture the interquartile range (IQR) of the price change distribution documented in the data (Montag and Villar 2025).[31] However, at the same time, the models significantly underestimate the extent of frequency-inflation comovement. Instead of a 14 pp increase observed in 2022, they only generate 4 pp in both the CalvoPlus and random menu cost models.[32] This has implications for the models' normative policy prescriptions.

All three state-dependent price-setting models prescribe a monetary policy in

[30] In particular, when we attempt to capture also the frequency-inflation co-movement, this pushes the calibrated parameters back towards our baseline model (with $1-\theta = 0$), in which case we cannot match the steady state interquartile range of price changes.

[31] Notice that these models have been calibrated to match the IQR, while the baseline was not.

[32] Gagliardone et al. (2025) show that the random menu cost framework captures well the approximate doubling of frequency during the surge (from around 2% to around 30%) of producer-price inflation in Belgium.

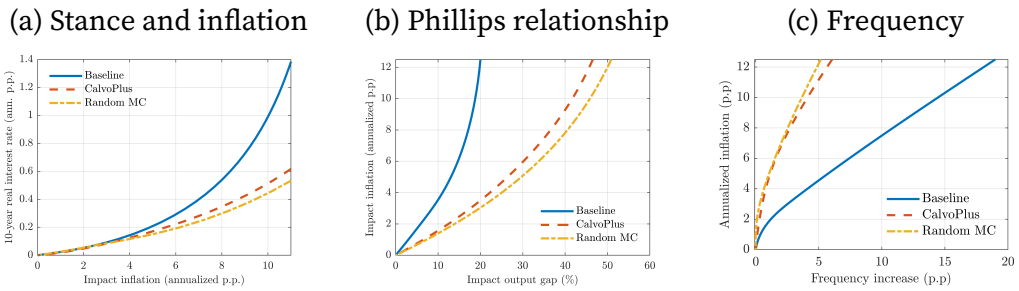

FIGURE 8. Robustness: alternative state-dependent models

Figure 8 contrasts (a) the relationship between optimal policy stance and inflation, (b) the Phillips relationship and (c) the frequency-inflation relationship in the CalvoPlus and a random menu cost model and our baseline. It shows that the optimal stance is robustly nonlinear in all three models, but its strength is underestimated by the CalvoPlus and random menu cost models, which also underestimate the true frequency inflation comovement and imply a less nonlinear Phillips relationship.

which the optimal stance tightens nonlinearly with the inflation rate (see Panel a of Figure 8). In all three models our key channel is active: the frequency of price changes increases with inflation (panel c) such that the Phillips relationship is nonlinear (see panel b). However, the nonlinearity is weaker in the CalvoPlus and the random menu cost models than in the baseline, as these models underestimate the strength of the frequency-inflation relationship, which our model matches well. Therefore they overestimate the output costs of disinflation under an inflation surge.

Relevance of fat-tailed shocks and roundabout input-output structure. Next, we assess the relevance of two key assumptions in driving our results: fat-tailed idiosyncratic shocks and strategic complementarities introduced through the roundabout production technology. For this, we recalibrate our baseline model without intermediate inputs ($\alpha = 1$) and with Gaussian idiosyncratic shocks ($\sigma_1 = \sigma_2$). Parameters are reported in Table A2.

Table 4 shows how these alternative parameterizations capture key moments. First, removing intermediate production inputs does not materially affect the model's ability to match key price setting moments both at the steady state and under 10% inflation. That said, the roundabout technology helps the model to match a more realistic Phillips curve slope and can achieve a 10% inflation under Ramsey policy with a substantially smaller cost-push shock (3.5%, versus 17.6 %). Second, without the introduction of fat-tailed idiosyncratic shocks the model would be unable to match either the kurtosis of the absolute price-change distribution or the magnitude of the frequency increase at 10% inflation. Fat-tailed shocks also contribute to the decline in the slope of the Phillips curve, through their impact on the selection of large price changes (0.42 versus 1.13, see also Midrigan 2011).

The nonlinearity of the optimal policy rule is preserved under the three parameterizations (see Panel a of Figure 9). The extent of nonlinearity is similar in the model without the roundabout economy and in the baseline. Likewise, the inflation-frequency

TABLE 4. Moments of the models without roundabout economy ($\alpha = 1$) and Gaussian idiosyncratic shocks ($\sigma_1 = \sigma_2$).

	Steady state at 2% inflation			10% inflation		Phillips curve
	Frequency	Size	Kurtosis	Freq. incr.	Shock size	slope
Data	8.7%	8.5%	4	14pp		0.006-0.106
Baseline	8.7%	8.5%	4	14pp	3.5%	0.08
No roundabout	8.7%	8.5%	4	13.7pp	17.6%	0.42
Gaussian, no roundabout	8.7%	8.5%	1.75	3.3pp	7.4%	1.13

The table shows that (i) roundabout economy ($\alpha = 0.2$) helps the model match the slope of the Phillips curve, and achieve 10% inflation without an extreme shock size. The assumption, however, does not impact the model's ability to match the magnitude of the frequency increase at 10% inflation. (ii) The fat-tailed idiosyncratic shocks contribute to matching the slope of the Phillips curve and without it the model would have trouble matching both the kurtosis of the price-change distribution and the magnitude of frequency increase at 10% inflation. The calibration targets are from Nakamura and Steinsson (2008) (frequency and size), Alvarez et al. (2016) (kurtosis), Montag and Villar (2025) (frequency increase), and Hazell et al. (2022) (PC slope).

relationship is also quite similar (see Panel c). These two models are also alike in terms of the nonlinearity of the Phillips relationship. This contrasts with the model without intermediate inputs and Gaussian shocks, where the nonlinearity of both the policy rule and the Phillips relationship is smaller. Moreover, this model significantly underestimates the frequency response at 10% inflation.

The importance of the frequency response. We have argued that fat-tailed shocks are key for our model to match the observed response of frequency at 10% inflation, while simultaneously capturing steady-state pricing moments. However, one can match the response of frequency in simpler models without fat-tailed shocks, at the cost of no longer matching some steady-state moments. Figure 10 shows the stance-inflation relationships, the Phillips relationships, and the inflation-frequency relationship for two models which are calibrated this way, and compare them to the baseline. These are the simple model from Section 3.1 and the canonical Golosov-Lucas model with Gaussian shocks and no intermediate goods. In both cases the models are calibrated to yield 8.7% frequency in the 2% inflation steady state, and to deliver an increase by 14 pp when inflation rises to 10% from its optimal level (close to 0%). Table A4 reports the implied moments.[33]

This exercise delivers two distinct results. First, the resulting stance-inflation relationship is very similar across the models (panel a). This indicates that there is a tight link between the frequency-inflation relationship (panel c) and the nonlinearity of optimal policy. Models that predict a realistically large elasticity of frequency to inflation predict a similar degree of nonlinearity. Even though the slope of the Phillips relationship differs across the three models,[34] the extent of their *nonlinearity*

[33]The models underestimate both the size of absolute price changes (3.8% instead of 8.5%) and the kurtosis of price changes.

[34]First, the slope is much higher in the simple model and in the conventional Golosov-Lucas model due to the absence of any strategic complementarities. Second, the slope is higher in the static model

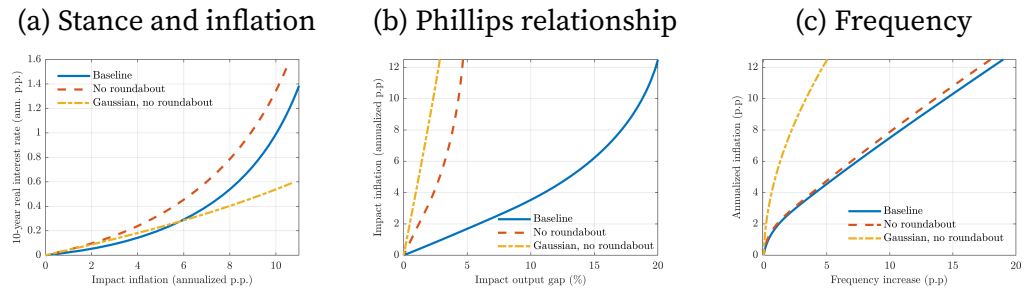

FIGURE 9. Robustness: Relevance of roundabout economy and leptokurtic shocks

This figure contrasts (a) the relationship between optimal policy stance and inflation, (b) the Phillips relationship and the (c) frequency-inflation relationship in the baseline with simplified versions of it that lack the roundabout structure and that have Gaussian instead of fat-tailed shocks.

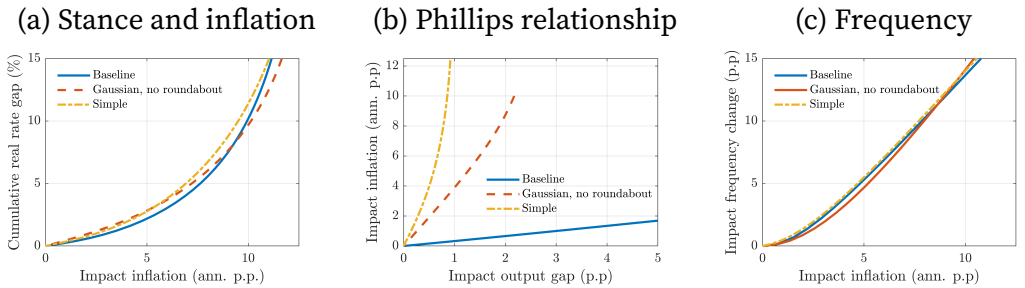

FIGURE 10. Robustness: relevance of the inflation-frequency relationship

Figure 10 contrasts (a) the relationship between optimal policy stance and inflation, (b) the Phillips relationship and the (c) frequency-inflation relationship in the simple model, the Golosov-Lucas model calibrated to capture the inflation-frequency relationship, and our baseline. It shows that the stance-inflation relationship is robustly nonlinear and quantitatively similar in the three models calibrated such that the inflation-frequency relationship is similar.

is comparable.

Second, the simple model delivers very similar stance-inflation and frequency-inflation relationships to both the baseline model and the canonical Golosov-Lucas model. This validates our use of the static model as a tractable approximation of the baseline model.

4.5. Policy implications

The previous sections established that under state-dependent price setting, optimal policy should become more anti-inflationary as inflation rises. How should a simple monetary policy rule be modified to reflect this feature, and how would such a rule have fared during the 2021–2023 inflation surge? We address these questions by extending the popular Taylor rule to approximate the Ramsey policy, and by assessing the rule's performance in a simulation calibrated to capture key features of the 2021–2023 U.S. inflation episode.

We focus on cost-push shocks throughout this section. As argued in Section 4.3, op-

than in the dynamic model, because in the former firms are myopic ($\beta = 0$).

timal policy completely offsets efficient disturbances — such as TFP or time-preference shocks — so their policy implications are straightforward in our framework: any inflation driven by efficient disturbances should be neutralized by keeping output at its efficient level.

Nonlinear Taylor rule. As already discussed, under cost-push shocks, Ramsey policy prescribes a stance that becomes increasingly aggressive as inflation rises (see Figures 4 and 6). To capture this feature, we modify the standard Taylor rule by allowing its inflation coefficient to vary smoothly with the level of inflation. Specifically, we assume the coefficient is quadratic in inflation, which, as we show shortly, approximates Ramsey policy well:

$$\log(R_t/R) = \rho_i \log(R_{t-1}/R) + (1-\rho_i)\big\{\underbrace{\phi_\pi(\pi_t)}_{\phi_{\pi,1}+\phi_{\pi,2}\pi_t^2}\pi_t + \phi_y(y_t - y_t^*)\big\}, \tag{33}$$

where ρ_i is the interest-rate smoothing coefficient, and $\phi_{\pi,1}$, $\phi_{\pi,2}$, and ϕ_y are parameters. The key departure from the standard Taylor rule is the term $\phi_{\pi,2}\pi_t^2$ within the inflation coefficient, which amplifies the responsiveness of policy to inflation whenever inflation is significantly above its steady-state level.

We calibrate some parameters of the rule directly. The monthly interest-rate smoothing coefficient is set to $\rho_i = 0.9^{1/3}$, where its quarterly equivalent is between the approximately 0.8 estimate from empirical studies (Clarida et al. 2000; Smets and Wouters 2007) and the near-unity values implied by optimal policy exercises (Levin et al. 1999; Woodford 2003). The output gap coefficient is set to $\phi_y = 0.5/12$, the monthly equivalent of the value in Taylor (1993).

The two parameters of the inflation coefficient, $\phi_{\pi,1}$ and $\phi_{\pi,2}$, are set as follows. The stance of current and future monetary policy is summarized well by the tightening in the 10-year real interest rate, which approximates the cumulated real interest rate gap (5). The blue line in Figure 11 reproduces the relationship between optimal stance and inflation under Ramsey policy. We choose $\phi_{\pi,1}$ and $\phi_{\pi,2}$ sequentially to match this relationship with the nonlinear Taylor rule. First, we set the linear coefficient ($\phi_{\pi,1}$) (imposing $\phi_{\pi,2} = 0$) to match the inflation-stance relationship under Ramsey policy and small shocks. The implied linear rule is of interest in its own right: it is a policy rule that is approximately optimal under small cost-push shocks, in the sense that the implied policy stance matches the Ramsey prescription on impact. The resulting linear coefficient is $\phi_{\pi,1} = 2.22$, not far from the original coefficient suggested by Taylor (1993) and consistent with some empirical estimates in the literature (see, for example, Clarida et al. 2000). Second, we set the nonlinear coefficient by keeping the linear coefficient fixed at $\phi_{\pi,1} = 2.22$ and matching the inflation-stance relationship at 10% inflation. The resulting nonlinear coefficient is $\phi_{\pi,2} = 135{,}000$, implying an effective inflation coefficient that rises from 2.22 under normal conditions to 11.6 following a 10-percentage-point increase in annualized inflation.

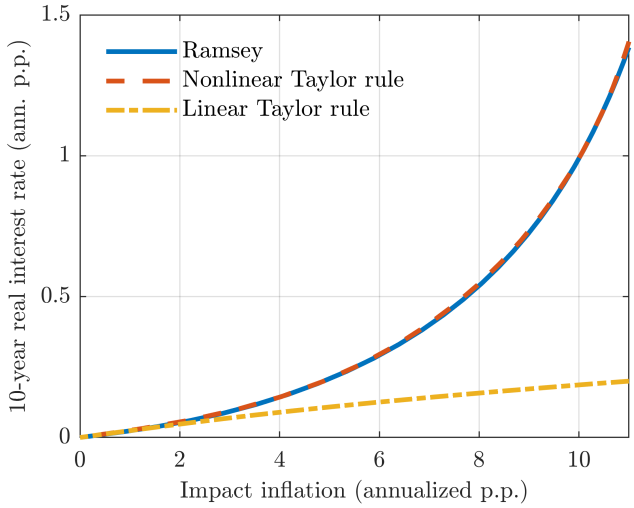

FIGURE 11. Nonlinear and linear Taylor rule and real-rate tightening.

The figure shows that the nonlinear Taylor rule with coefficients $\phi_{\pi,1} = 2.22$ and $\phi_{\pi,2} = 135{,}000$ matches the relationship between the cumulative real interest rate gap and impact inflation under Ramsey policy. A linear rule ($\phi_{\pi,2} = 0$) would imply a suboptimally accommodative policy stance at high inflation levels.

Figure 11 shows that the nonlinear rule successfully reproduces the nonlinear relationship between inflation and the optimal policy stance prescribed by the Ramsey policy. The standard linear rule ($\phi_{\pi,2} = 0$) — one that ignores the state-dependence of optimal policy — implies a policy stance that is suboptimally accommodative at high inflation levels, a property we examine in detail in the simulation below.

A simulation of the 2021–2023 inflation surge. To assess the quantitative importance of the nonlinearity of optimal monetary policy, we conduct a simulation exercise motivated by the U.S. 2021–2023 inflation surge. We assume that the central bank follows the linear Taylor rule — that is, we impose $\phi_{\pi,2} = 0$ while keeping all other parameters as described above. This is an empirically realistic benchmark, which is also an approximately optimal policy under small shocks.

We simulate a cost-push shock calibrated to drive inflation to 10%, close to the peak observed in the United States.[35] Because optimal policy fully offsets efficient and demand shocks in our framework — which, in practice, also contributed to the inflation surge — restricting attention to cost-push shocks is a conservative choice and provides a lower bound on the inflation-reducing impact of optimal policy.

The solid blue line in Figure 12 shows the response under the linear rule. By construction, inflation peaks at 10%, and the price adjustment frequency rises more than 10%, from below 9% to nearly 20%, broadly in line with U.S. experience, as discussed in Section 4.1. Output falls by around 3%, which would be a reasonable counterfactual effect under a scenario with only cost-push shocks. Papers decomposing the U.S. inflation surge also find that cost-push shocks — like oil price shocks — contributed

[35]The cost-push shock τ_t increases the marginal cost by 1% ceteris paribus.

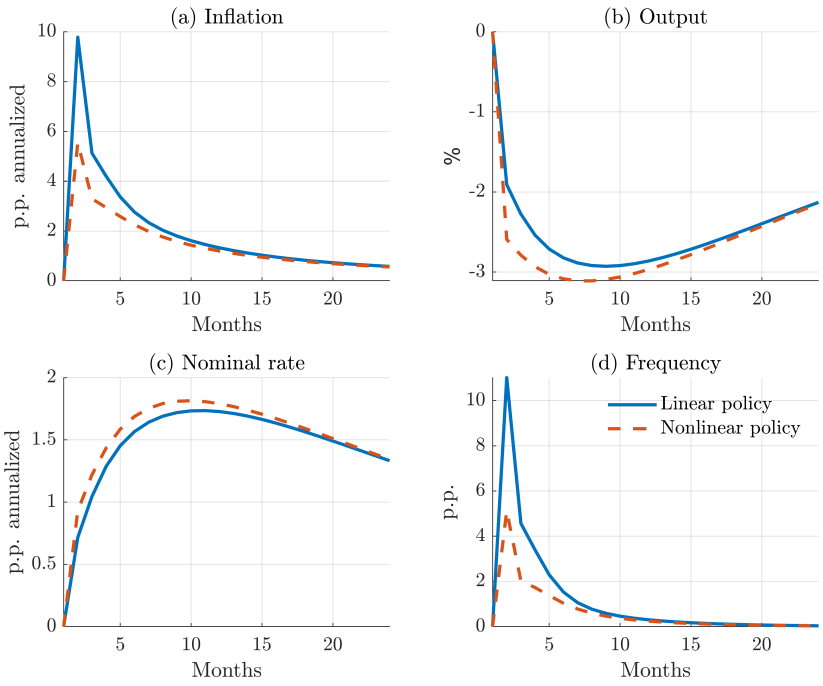

FIGURE 12. Impulse responses under linear and nonlinear Taylor rules.

The figure compares impulse responses under the linear Taylor rule ($\phi_{\pi,2} = 0$, solid blue) and the nonlinear rule ($\phi_{\pi,2} = 135{,}000$, dashed red) to a cost-push shock that drives annualized inflation to 10% under the linear rule. The nonlinear rule reduces peak inflation by approximately 4.5 percentage points at the cost of a persistently lower output by around 50 basis points.

substantially to it and would have caused a significant downturn if not counteracted by contemporaneous demand shocks (see, for example, Gagliardone and Gertler 2023). The nominal interest rate rises persistently by around 175 basis points — less than the roughly 500-basis-point tightening observed in 2022, which, however, also reflected a normalization from prior policy accommodation and an increase in the long-run real rate.

The dashed red line shows the counterfactual under the nonlinear rule. The inflation-sensitive coefficient keeps peak inflation approximately 4.5 percentage points lower, at around 5.5%, and substantially dampens the increase in the price adjustment frequency. The combination of lower inflation and a higher nominal rate (around 25 basis points) implies significantly tighter real interest rate conditions. The output cost is modest: the tighter policy stance implies an additional decline by roughly 50 basis points lasting just over a year. In sum, a strike-while-the-iron-is-hot policy allows the central bank to contain inflation following large adverse cost-push shocks at a relatively modest output cost.

To quantify the welfare implications, we compare the distance from the Ramsey optimum under each rule in consumption-equivalent terms. The nonlinear rule closes half of the welfare gap to the Ramsey policy relative to the linear rule, a substantial gain.

In the above analysis, we focused on the Taylor rule parameters that best approximate Ramsey policy, selecting the stance-inflation relationship as our target statistic. In Appendix F we show that the results are robust to finding the Taylor rule parameters that maximize conditional welfare under cost-push shocks.

5. Conclusion

This paper characterizes Ramsey optimal monetary policy in a menu cost model with roundabout production and fat-tailed idiosyncratic shocks. Our central finding is that, in response to a large aggregate cost-push shock, optimal policy commits to generating less inflation than prescribed by the standard New Keynesian model. The mechanism is intuitive: the central bank exploits the steeper Phillips curve that emerges endogenously at higher inflation rates — it strikes while the iron is hot. Quantitatively, this nonlinearity is already material at inflation rates comparable to those observed during the post-pandemic surge, suggesting first-order policy relevance. This prescription stands in sharp contrast to the standard New Keynesian model with time-dependent pricing, which, by construction, cannot capture such nonlinear dynamics. For TFP and discount-factor shocks, by contrast, optimal policy in the menu cost model closely resembles its time-dependent counterpart, prescribing a commitment to price stability.

Our analysis has focused on nominal price rigidities under commitment. Extending the framework to discretionary policy and to richer settings that include nominal wage rigidities is an important direction for future work.[36]

References

Adam, K. and H. Weber (2019, 2). Optimal Trend Inflation. *American Economic Review 109*(2), 702–737.

Alexandrov, A. (2020). The effects of trend inflation on aggregate dynamics and monetary stabilization. CRC TR 224 Discussion Paper Series, University of Bonn and University of Mannheim, Germany.

Alvarez, F., M. Beraja, M. Gonzalez-Rozada, and P. A. Neumeyer (2019). From Hyperinflation to Stable Prices: Argentina's Evidence on Menu Cost Models. *The Quarterly Journal of Economics 134*(1), 451–505.

Alvarez, F., H. Le Bihan, and F. Lippi (2016). The real effects of monetary shocks in sticky price models: A sufficient statistic approach. *American Economic Review 106*(10), 2817–51.

Alvarez, F. and P. A. Neumeyer (2020, March). The passthrough of large cost shocks in an inflationary economy. In G. Castex, J. Galí, and D. Saravia (Eds.), *Changing Inflation Dynamics,Evolving Monetary Policy*, Volume 27 of *Central Banking, Analysis, and Economic Policies Book Series*, Chapter 2, pp. 007–048. Central Bank of Chile.

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica 89*(5), 2375–2408.

[36] A step in this direction has been taken by Nuño et al. (2025), who analyze optimal monetary policy under discretion in the framework by Blanco et al. (2024c) with persistent shocks.

Auclert, A., R. Rigato, M. Rognlie, and L. Straub (2024). New pricing models, same old phillips curves? *The Quarterly Journal of Economics 139*(1), 121–186.

Auer, R., A. Burstein, and S. M. Lein (2021). Exchange rates and prices: Evidence from the 2015 swiss franc appreciation. *American Economic Review 111*(2), 652–686.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Barro, R. J. (1972). A Theory of Monopolistic Price Adjustment. *Review of Economic Studies 39*(1), 17–26.

Basu, S. (1995). Intermediate goods and business cycles: Implications for productivity and welfare. *American Economic Review 85*(3), 512–31.

Benigno, P. and G. Eggertsson (2023). It's baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve. NBER Working Papers 31197, National Bureau of Economic Research, Inc.

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2021). Inequality, business cycles, and monetary-fiscal policy. *Econometrica 89*(6), 2559–2599.

Blanchard, O. and J. Galí (2007a). The macroeconomic effects of oil price shocks: Why are the 2000s so different from the 1970s? In *International Dimensions of Monetary Policy*, pp. 373–421. National Bureau of Economic Research, Inc.

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

Blanco, A. (2021, July). Optimal Inflation Target in an Economy with Menu Costs and a Zero Lower Bound. *American Economic Journal: Macroeconomics 13*(3), 108–141.

Blanco, A., C. Boar, C. Jones, and V. Midrigan (2024a). Nonlinear inflation dynamics in menu cost economies? evidence from U.S. data. unpublished manuscript.

Blanco, A., C. Boar, C. J. Jones, and V. Midrigan (2024b). Nonlinear inflation dynamics in menu cost economies. NBER Working Papers 32094, National Bureau of Economic Research.

Blanco, A., C. Boar, C. J. Jones, and V. Midrigan (2024c, May). The Inflation Accelerator. NBER Working Papers 32531, National Bureau of Economic Research, Inc.

Boppart, T., P. Krusell, and K. Mitman (2018). Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative. *Journal of Economic Dynamics and Control 89*(C), 68–92.

Burstein, A. and C. Hellwig (2008). Welfare costs of inflation in a menu cost model. *American Economic Review 98*(2), 438–43.

Caballero, R. and E. Engel (1993). Heterogeneity and Output Fluctuations in a Dynamic Menu-Cost Economy. *The Review of Economic Studies 60*(1), 95–119.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics 12*(3), 383 – 398.

Caratelli, D. and B. Halperin (2023). Optimal monetary policy under menu costs. unpublished manuscript.

Cavallo, A., F. Lippi, and K. Miyahara (2024). Large shocks travel fast. *American Economic Review: Insights 6*(4), 558–574.

Clarida, R., J. Galí, and M. Gertler (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics 115*(1), 147–180.

Costain, J. and A. Nakov (2011). Distributional dynamics under smoothly state-dependent pricing. *Journal of Monetary Economics 58*(6), 646 – 665.

Costain, J., A. Nakov, and B. Petit (2022). Flattening of the phillips curve with state-dependent prices and wages. *The Economic Journal 132*(642), 546–581.

Dávila, E. and A. Schaab (2022). Optimal monetary policy with heterogeneous agents: A timeless ramsey approach. Working Paper.

Dixit, A. K. and J. E. Stiglitz (1977, June). Monopolistic Competition and Optimum Product Diversity. *American Economic Review 67*(3), 297–308.

Dotsey, M., R. G. King, and A. L. Wolman (1999). State-dependent pricing and the general equilibrium dynamics of money and output. *The Quarterly Journal of Economics 114*(2), 655–690.

Erceg, C. J., J. Lindé, and M. Trabandt (2024, December). Monetary Policy and Inflation Scares. IMF Working Papers 2024/260, International Monetary Fund.

Gagliardone, L. and M. Gertler (2023). Oil prices, monetary policy and inflation surges. NBER Working Papers 31263, National Bureau of Economic Research, Inc.

Gagliardone, L., M. Gertler, S. Lenzu, and J. Tielens (2025, February). Micro and macro cost-price dynamics in normal times and during inflation surges. NBER Working Papers 33478, National Bureau of Economic Research, Inc.

Gagnon, E. (2009). Price setting during low and high inflation: Evidence from Mexico. *The Quarterly Journal of Economics 124*(3), 1221–1263.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press.

Gautier, E., C. Conflitti, D. Enderle, L. Fadejeva, A. Grimaud, E. Gutiérrez, V. Jouvanceau, J.-O. Menz, A. Paulus, P. Petroulas, P. Roldan-Blanco, and E. Wieland (2025, February). Price Stickiness in the Euro Area in Times of High Inflation. unpublished manuscript.

Gertler, M. and J. Leahy (2008). A Phillips curve with an Ss foundation. *Journal of Political Economy 116*(3), 533–572.

Ghassibe, M. and A. Nakov (2025, Sep). Business cycles with pricing cascades. Working Paper Series 3123, European Central Bank.

Golosov, M. and R. E. Lucas (2007). Menu costs and Phillips curves. *Journal of Political Economy 115*(2), 171–199.

González, B., G. Nuño, D. Thaler, and S. Albrizio (2024). Firm heterogeneity, capital misallocation and optimal monetary policy. Working Paper Series 2890, European Central Bank.

Hazell, J., J. Herreno, E. Nakamura, and J. Steinsson (2022). The slope of the phillips curve: Evidence from U.S. states. *The Quarterly Journal of Economics 137*(3), 1299–1344.

Karadi, P. and A. Reiff (2019). Menu costs, aggregate fluctuations, and large shocks. *American Economic Journal: Macroeconomics 11*(3), 111–146.

Le Grand, F., A. Martin-Baillon, and X. Ragot (2022). Should monetary policy care about redistribution? Optimal monetary and fiscal policy with heterogeneous agents. unpublished manuscript.

Levin, A., V. Wieland, and J. Williams (1999). Robustness of simple monetary policy rules under model uncertainty. In *Monetary Policy Rules*, pp. 263–318. National Bureau of Economic Research, Inc.

Midrigan, V. (2011). Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica 79*(4), 1139–1180.

Montag, H. and D. Villar (2023). Price-setting during the covid era. FEDS Notes.

Montag, H. and D. V. Villar (2025). Post-pandemic price flexibility in the u.s.: Evidence and implications for price setting models. Finance and Economics Discussion Series 2025-024, Board of Governors of the Federal Reserve System (U.S.).

Nakamura, E. and J. Steinsson (2008). Five facts about prices: A reevaluation of menu cost

models. *The Quarterly Journal of Economics 123*(4), 1415–1464.

Nakamura, E. and J. Steinsson (2010). Monetary non-neutrality in a multisector menu cost model. *The Quarterly Journal of Economics 125*(3), 961–1013.

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: The information effect. *The Quarterly Journal of Economics 133*(3), 1283–1330.

Nakov, A. and C. Thomas (2014, September). Optimal Monetary Policy with State-Dependent Pricing. *International Journal of Central Banking 36*, 49–93.

Nuño, G., P. Renner, and S. Scheidegger (2025). Monetary policy with supply regimes. Mimeo.

Nuño, G. and C. Thomas (2022). Optimal Redistributive Inflation. *Annals of Economics and Statistics GENES*(146), 3–63.

Rubbo, E. (2023, July). Networks, phillips curves, and monetary policy. *Econometrica 91*(4), 1417–1455.

Sheshinski, E. and Y. Weiss (1977). Inflation and Costs of Price Adjustment. *Review of Economic Studies 44*(2), 287–303.

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

Taylor, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 39, pp. 195–214. Elsevier.

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

APPENDIX

Contents

Appendix A. The simple model

This appendix lays out the simple model and provides the proofs for the related propositions.

Model description. As explained in the main text in Section 3, prices are reset overnight in the simple model. Thus, all dynamics are muted, such that the model boils down to a sequence of static models.[37] For this reason, we remove the time subscript in the simple model.

The equilibrium conditions of the simple model coincide with those of the full model in large part. The dynamic equations of the firm sector change due to night-time price adjustment, as we now explain.

Due to the free night-time adjustment, the firm's Bellman equation (19) reduces to its current profit function (20). The optimal reset price therefore corresponds to a constant markup over marginal cost, which is simply given by the wage, since there is no roundabout production. Suppressing time subscripts in this section for brevity, this yields

$$e^{p^*} = \frac{\epsilon}{\epsilon - 1}(1-\tau)\frac{w}{A}. \tag{A1}$$

The price-adjustment thresholds, previously defined by (17)–(18), are now defined by the condition that current profits under the status-quo price, $\Pi(x)$, equal the profits obtained under the optimal price net of the menu cost, $\Pi(0) - \eta w$. Thus,

$$\left(e^{p^*(1-\epsilon)} - (1-\tau)\frac{w}{A}e^{-\epsilon p^*}\right)C - \eta w = \left(e^{(p^*+s)(1-\epsilon)} - (1-\tau)\frac{w}{A}e^{-\epsilon(p^*+s)}\right)C, \tag{A2}$$

$$\left(e^{p^*(1-\epsilon)} - (1-\tau)\frac{w}{A}e^{-\epsilon p^*}\right)C - \eta w = \left(e^{(p^*+S)(1-\epsilon)} - (1-\tau)\frac{w}{A}e^{-\epsilon(p^*+S)}\right)C, \tag{A3}$$

where we have made use of the fact that absent roundabout production $Y = C$.

Furthermore, free night-time adjustment changes the law of motion of the price gap distribution. It implies that $g^c_{-1}(x) = 0$, $g^0_{-1} = 1$, and $p^*_{-1} = 0$. Given this, and that we abstract from leptokurtic shocks, the daytime price-gap distribution is thus Gaussian with mean $-(\pi + p^*)$ and variance σ^2:

$$g^c(x) = \frac{1}{\sigma}\phi\left(\frac{x + \pi + p^*}{\sigma}\right), \qquad x \in [s, S], \tag{A4}$$

where ϕ is the density of the standard normal variable.[38]

The four static equilibrium conditions remain unchanged from the baseline model: the intratemporal labor supply (3), the frequency of price changes (23), the labor-market clearing (27), and the definition of the price level (9). Hence, there are eight

[37] Alternatively, the static model version can be seen as a particular case of the complete model in which we set $\beta = 0$ and assume that the initial distribution is such that all firms had set the same price last period ($g^c_{-1}(x) = 0$, $g^0_{-1} = 1$, $p^*_{-1} = 0$).

[38] To see this, plug (15), $g^c_{-1}(x) = 0$, $g^0_{-1} = 1$ and $p^*_{-1} = 0$ into (22) and note that $\phi(x) = \phi(-x)$.

equations and nine variables $\{\pi, w, C, N, s, S, g^0, g^c(\cdot), p^*\}$, leaving the policymaker to choose the inflation rate π.[39] We summarize those equations here:

$$
\begin{aligned}
e^{p^*} &= \frac{\epsilon}{(\epsilon-1)}(1-\tau)\frac{w}{A}, && \text{(A5)}\\
(e^{p^*})^{1-\epsilon} &- (1-\tau)\frac{w}{A}(e^{p^*})^{-\epsilon} - \eta\frac{w}{C} = (e^{(p^*+s)})^{1-\epsilon} - (1-\tau)\frac{w}{A}(e^{(p^*+s)})^{-\epsilon}, && \text{(A6)}\\
(e^{p^*})^{1-\epsilon} &- (1-\tau)\frac{w}{A}(e^{p^*})^{-\epsilon} - \eta\frac{w}{C} = (e^{(p^*+S)})^{1-\epsilon} - (1-\tau)\frac{w}{A}(e^{(p^*+S)})^{-\epsilon}, && \text{(A7)}\\
g^c(x) &= \frac{1}{\sigma}\phi\left(\frac{x+\pi+p^*}{\sigma}\right), && \text{(A8)}\\
w &= C, && \text{(A9)}\\
g^0 &= 1 - \int_s^S g^c(x)dx, && \text{(A10)}\\
1 &= \int_s^S e^{(x+p^*)(1-\epsilon)} g^c(x)\,dx + g^0 e^{p^*(1-\epsilon)}, && \text{(A11)}\\
N &= \frac{C}{A}\left(\int_s^S e^{(x+p^*)(-\epsilon)} g^c(x)\,dx + g^0 e^{p^*(-\epsilon)}\right) + \eta g^0. && \text{(A12)}
\end{aligned}
$$

The only equation that remains dynamic is the Euler equation (4). However, this equation is redundant and does not constrain the planner. It only serves to back out the interest rate consistent with the optimal policy, so we can ignore it. The equilibrium in the simple model thus boils down to a sequence of static equilibria.

In the above, as in the main text, we define the distribution and value functions as a function of the price gap x, as is common in the state-dependent pricing literature. However, for the analysis of the simple model, it is convenient to define them rather as a function of the price level $p \equiv x + p^*$. The (S, s) bands will also be renormalized accordingly; by abuse of notation, we continue to denote them s, S. After this change of variable, the system reads:

$$
\begin{aligned}
e^{p^*} &= \frac{\epsilon}{(\epsilon-1)}(1-\tau)\frac{w}{A}, && \text{(A13)}\\
(e^{p^*})^{1-\epsilon} &- (1-\tau)\frac{w}{A}(e^{p^*})^{-\epsilon} - \eta\frac{w}{C} = (e^{s})^{1-\epsilon} - (1-\tau)\frac{w}{A}(e^{s})^{-\epsilon}, && \text{(A14)}\\
(e^{p^*})^{1-\epsilon} &- (1-\tau)\frac{w}{A}(e^{p^*})^{-\epsilon} - \eta\frac{w}{C} = (e^{S})^{1-\epsilon} - (1-\tau)\frac{w}{A}(e^{S})^{-\epsilon}, && \text{(A15)}\\
g^c(p) &= \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right), && \text{(A16)}\\
w &= C, && \text{(A17)}\\
g^0 &= 1 - \int_s^S g^c(p)dp, && \text{(A18)}\\
1 &= \int_s^S e^{p(1-\epsilon)} g^c(p)\,dp + g^0 e^{p^*(1-\epsilon)}, && \text{(A19)}
\end{aligned}
$$

[39] To ensure that firms have no incentive to deviate from a symmetric nighttime reset price, we assume that the anticipated value of τ for the following day is such that $\pi = 0$ is optimal.

$$N = \frac{C}{A}\left(\int_s^S e^{p(-\epsilon)} g^c(p)\, dp + g^0 e^{p^*(-\epsilon)}\right) + \eta g^0. \tag{A20}$$

Note that the firm's decisions are much simpler than in the full model, since they are static. The reset price maximizes firms' current profits by setting a constant markup (A13). Furthermore, firms keep their (logged quality-adjusted real) price p unchanged as long as current profits exceed profits under the optimal price p^* minus the menu cost, that is, when $\Delta\Pi(p) \equiv \Pi(p) - (\Pi(p^*) - \eta w) > 0$. $\Delta\Pi(p)$ has exactly two roots.[40] One root (s) is smaller than p^* and one root (S) is larger than p^*. The function $\Delta\Pi(p)$ is positive between them. Thus, these two roots define the (S, s) bands, which characterize the optimal update decision.

Phillips curve - Proposition 1. We now show how we derive the Phillips curve displayed in Proposition 1. First, we use equations (A16)-(A18) to eliminate g^0, w and $g^c(p)$ from equations (A13), (A14), (A15) and (A19). Then we use the resulting version of (A13) to eliminate p^* in the remaining 3 equations. Finally, we define the output gap relative to the efficient level of output $Y^{\text{eff}} = A$ as $\hat{Y} \equiv \frac{Y}{A} = \frac{C}{A}$. This leaves us with the following equations:

$$\left(\frac{\epsilon}{\epsilon-1}(1-\tau)\hat{Y}\right)^{1-\epsilon} - \left(\frac{\epsilon}{\epsilon-1}\right)^{-\epsilon}\left((1-\tau)\hat{Y}\right)^{1-\epsilon} - \eta = e^{S(1-\epsilon)} - (1-\tau)\hat{Y}e^{S(-\epsilon)}, \tag{A21}$$

$$\left(\frac{\epsilon}{\epsilon-1}(1-\tau)\hat{Y}\right)^{1-\epsilon} - \left(\frac{\epsilon}{\epsilon-1}\right)^{-\epsilon}\left((1-\tau)\hat{Y}\right)^{1-\epsilon} - \eta = e^{s(1-\epsilon)} - (1-\tau)\hat{Y}e^{s(-\epsilon)}. \tag{A22}$$

$$1 = \left[\int_s^S e^{(p)(1-\epsilon)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp + \left(\frac{\epsilon(1-\tau)}{\epsilon-1}\hat{Y}\right)^{1-\epsilon}\left[1 - \int_s^S \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right]\right],$$

The first two equations implicitly define the functions $\mathsf{s}(\hat{Y}, \tau)$ and $\mathsf{S}(\hat{Y}, \tau)$. A simple closed-form solution for these limits exists if $\epsilon = 2$. For $\epsilon = 3$ and $\epsilon = 4$, a more cumbersome closed-form solution exists. Beyond that we have not found any closed form solution. Plugging those two functions into the last equation, we arrive at the Phillips curve in Proposition 1. That is, we have compressed equations (A13)– (A19) into a single equation relating inflation and output – the Phillips curve:

$$1 = \int_{\mathsf{s}(\hat{Y},\tau)}^{\mathsf{S}(\hat{Y},\tau)} e^{p(1-\epsilon)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp + \left(\frac{\epsilon(1-\tau)}{\epsilon-1}\hat{Y}\right)^{1-\epsilon}\left[1 - \int_{\mathsf{s}(\hat{Y},\tau)}^{\mathsf{S}(\hat{Y},\tau)} \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right] \tag{A23}$$

[40] To see this consider the function $f(P) \equiv \Delta\Pi(\log P)P^{\epsilon}$. Note that $f(P)$ is written in terms of the price in levels, not in logs. $f(P)$ is positive at the optimal price $P = \exp(p^*)$, negative at $P \to 0$ and at $P \to \infty$, continuous and concave for positive P. Thus it has 2 roots. The log of those roots are the only roots of the function $\Delta\Pi(p)$.

The effect of the cost-push shock τ on the Phillips curve. Note that the terms $\hat{Y}$ and τ only appear as the product $\hat{Y}(1-\tau)$ in this expression. Thus, if we express this Phillips curve in terms of the log of $\hat{Y}$, $\pi(\log(\hat{Y});\tau)$—as we have done in all figures in section 3—changes in τ lead to parallel horizontal shifts. To see this, consider a particular combination of $\bar{\pi}$, $\bar{\hat{Y}}$, $\bar{\tau}$ satisfying the above Phillips curve. Now consider a different level of τ such that $(1-\tau) = (1-\bar{\tau})x$. To satisfy the above equation for $\pi = \bar{\pi}$, it must be that $\hat{Y} = \bar{\hat{Y}}/x$. Thus $\log(\hat{Y}) = \log(\bar{\hat{Y}}) - \log(x)$. This is a parallel horizontal shift of the function $\pi(\log(\hat{Y});\tau)$.

Welfare - Proposition 3. Next, we prove Proposition 3. The central bank's objective is given by the household's utility function, which we can write in terms of the output gap

$$U = \log(C) - \chi N = \log(\hat{Y}) + \log(A) - \chi N \tag{A24}$$

Using first the labor-market clearing condition (A20) to eliminate N in the utility function, and then the definition of frequency (A18) to eliminate g^0 and then the distribution (A16) to eliminate g^c, we arrive at the following:

$$\begin{aligned} U &= \log(\hat{Y}) - \hat{Y}\left(\int_s^S e^{p(-\epsilon)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp + \left(1-\int_s^S \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right)e^{p^*(-\epsilon)}\right) \\ &\quad -\eta\left(1-\int_s^S \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right) + \log(A) \end{aligned} \tag{A25}$$

Using the firms' reset price (A13) to eliminate $\frac{w}{A}$ in the (S,s) conditions (A14), (A15) we get:

$$e^{p^*(1-\epsilon)} - \frac{\epsilon-1}{\epsilon}e^{p^*(1-\epsilon)} - \eta = e^{s(1-\epsilon)} - \frac{\epsilon-1}{\epsilon}e^{s(-\epsilon)}e^{p^*}, \tag{A26}$$

$$e^{p^*(1-\epsilon)} - \frac{\epsilon-1}{\epsilon}e^{p^*(1-\epsilon)} - \eta = e^{S(1-\epsilon)} - \frac{\epsilon-1}{\epsilon}e^{S(-\epsilon)}e^{p^*}, \tag{A27}$$

Equations (A26), (A27), and the definition of the price level (A19) together implicitly define functions $\mathsf{s}(\pi)$, $\mathsf{S}(\pi)$ and $\mathsf{p}^*(\pi)$. Plugging these into the welfare function (A25) we arrive at the expression in the text:

$$\begin{aligned} U &= \log(\hat{Y}) - \hat{Y}\left(\int_{\mathsf{s}(\pi)}^{\mathsf{S}(\pi)} e^{(p)(-\epsilon)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp + \left(1-\int_{\mathsf{s}(\pi)}^{\mathsf{S}(\pi)} \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right)e^{\mathsf{p}^*(\pi)(-\epsilon)}\right) \\ &\quad -\eta\left[1-\int_{\mathsf{s}(\pi)}^{\mathsf{S}(\pi)} \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right] + \log(A) \end{aligned} \tag{A28}$$

This welfare function depends only on inflation and consumption. It also applies to Calvo, with a few modifications: $\eta = 0$, $-\mathsf{s}(\pi) = \mathsf{S}(\pi) = \infty$ and the term for the endogenous price updating probability $1 - \int_{\mathsf{s}(\pi)}^{\mathsf{S}(\pi)} \frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp$ is to be replaced by the

exogenous Calvo price updating probability. In the Calvo case without idiosyncratic shocks, this representation of the welfare function, when approximated to second order, yields the well-known loss function $-\frac{1}{2}\left[\hat{y}^2+\epsilon\left(\frac{\theta}{1-\theta}\right)\pi^2\right]$ (see Galí 2008) where the 'hat' denotes log deviations from the efficient allocation.

In the menu cost model which we study here, we can decompose the welfare gap relative to the efficient allocation ($C^{\text{eff}}=A$, $L^{\text{eff}}=1$) into 3 terms:

$$\begin{aligned} U-U^{\text{eff}} &= \underbrace{\log(\hat{Y})-(\hat{Y}-1)}_{\text{Average markup gap}} \\ &\quad \underbrace{-\hat{Y}\left(\int_{s(\pi)}^{S(\pi)} e^{(p)(-\epsilon)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp+\left(1-\int_{s(\pi)}^{S(\pi)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right)e^{p^*(\pi)(-\epsilon)}-1\right)}_{\text{Markup dispersion}} \\ &\quad \underbrace{-\eta\left(1-\int_{s(\pi)}^{S(\pi)}\frac{1}{\sigma}\phi\left(\frac{p+\pi}{\sigma}\right)dp\right)}_{\text{Adjustment costs}} \\ &= \underbrace{-\log\bar{\mu}-\left(\frac{1}{\bar{\mu}}-1\right)}_{\text{Average markup}}-\underbrace{\frac{1}{\bar{\mu}}(\zeta^{\mu}-1)}_{\text{Markup dispersion}}-\underbrace{\eta g^0}_{\text{Adjustment costs}}. \end{aligned}$$

The first equation follows directly from (A28). The second equation is explained in the proof of Proposition 2.

Welfare - Proposition 2. Finally, we prove Proposition 2. We start by proving three lemmas. The first describes the relationship between output and the average welfare-relevant markup, the second the relationship between price and markup dispersion and the third characterizes the efficient allocation.

LEMMA A1. *Let the average welfare-relevant markup* $\bar{\mu}\equiv\left(\int\mu(j)^{1-\epsilon}dj\right)^{\frac{1}{1-\epsilon}}$, *where the welfare-relevant markup is the relative price of firm* j *divided by its welfare-relevant marginal cost:* $\mu(j)=\frac{P(j)/P}{WRMC(j)}$, *and* $WRMC(j)\equiv wA(j)/A$. *Then in any market equilibrium there is a relationship between the average welfare-relevant markup and output:*

$$\log Y=\log A-\log\bar{\mu} \tag{A29}$$

or equivalently

$$Y=\frac{A}{\bar{\mu}} \tag{A30}$$

PROOF. In the proof, we first derive the real welfare-relevant marginal cost and define an aggregate component that is common to all firms. We show that this aggregate component is what affects the average welfare-relevant markup. Then, we derive an expression for the average efficient markup gap, which proves the lemma.

The real welfare-relevant marginal cost of firm j is

$$WRMC(j) = \frac{\partial\left(wN(j)\right)}{\partial Y(j)} = \frac{wA(j)}{A},$$

where we have used that $N(j) = A(j)Y(j)/A$.

Let the common real welfare-relevant marginal cost *wrmc* be defined as

$$wrmc \equiv (WRMC(j)/A(j)) = (w/A) = Y/A, \tag{A31}$$

where we used the labor supply condition ($w = C$) and the definition of output ($Y = C$) which together ensure that $w = Y$.

The welfare-relevant markup $\mu(j)$ is the relative price divided by the real welfare-relevant marginal cost:

$$\mu(j) = \frac{P(j)}{P} \Big/ \frac{wA(j)}{A} = \frac{P(j)}{A(j)P} \Big/ \frac{w}{A} = \frac{e^{p(j)}}{wrmc},$$

where $p(j)$ is the logarithm of the quality-adjusted relative price.

The average welfare-relevant markup $\overline{\mu}$ is

$$\overline{\mu} = \left(\int \mu(j)^{1-\epsilon} dj\right)^{\frac{1}{1-\epsilon}} = \left(\int \frac{e^{p(j)(1-\epsilon)}}{wrmc^{1-\epsilon}} dj\right)^{\frac{1}{1-\epsilon}} = \frac{1}{wrmc}\left(\int e^{p(j)(1-\epsilon)} dj\right)^{\frac{1}{1-\epsilon}} = \frac{1}{wrmc}, \tag{A32}$$

where we used the observation that the average quality-adjusted relative price is one in equilibrium.

The lemma follows from equations (A31) and (A32). □

We define the complete density $g(p) \equiv g^c(p) + g^0\delta(p - p^*)$, which includes both the continuous term $g^c(p)$ defined in equation (22) and the Dirac delta $\delta(p)$ times the frequency g^0 defined in equation (23). The second lemma shows the relationship between price dispersion and markup dispersion.

LEMMA A2. *Let the dispersion of the quality-adjusted relative prices be* $\zeta^p \equiv \int e^{p(-\epsilon)} g(p) dp$. *Let the markup dispersion be* $\zeta^\mu \equiv \int (\mu(p)/\overline{\mu})^{-\epsilon} g(\log \mu(p) - \log \overline{\mu}) dp$. *Then*

$$\zeta^p = \zeta^\mu. \tag{A33}$$

PROOF.

$$\zeta^p = \int e^{p(-\epsilon)} g(p) dp = \int e^{(\log \mu(p) + \log wrmc)(-\epsilon)} g(\log \mu(p) + wrmc) dp = \int e^{(\log \mu(p) - \log \overline{\mu})(-\epsilon)} g(\log \mu(p) - \log \overline{\mu}) dp = \zeta^\mu$$

□

The third lemma establishes the efficient levels of output and labor. The lemma implies that the efficient output fluctuates with aggregate productivity but is independent of both demand shocks and cost-push shocks.

LEMMA A3. *Let Y^e be the efficient output and N^e be the efficient labor, then*

$$\begin{aligned} N^e &= 1, \\ Y^e &= A. \end{aligned} \tag{A34}$$

PROOF. We obtain the efficient output as the solution to a social planning problem. The problem maximizes household welfare in equation (1) subject to (i) the aggregate consumption equation (6), (ii) the aggregate labor supply condition ($N_t = \int_i N_t(j)$) and (iii) product-level production functions in (10) with respect to product-level consumption and labor ($C_t(j), N_t(j), j \in [0,1], t = 0, 1, 2, \dots$).

After some algebra, the optimization problem simplifies to

$$\max_{N_t(j)} \sum_{t=0}^{\infty} \beta^t \log \left[A_t \left(\int N_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}} \right] - \chi \int N_t(j) dj,$$

subject to $\int N_t(j) dj = N_t$.

It is straightforward to see that the optimization problem implies that the efficient labor supply is equal across products ($N_t^e(j) = N_t^e$, for all $t = 1, 2, \dots$). Furthermore, optimality requires that $N_t^e = 1$ for all $t = 1, 2, \dots$. From this, it is clear that $Y_t^e = A_t N_t^e = A_t$. □

COROLLARY A1. *The efficient product-level consumption ($C^e(j)$) varies across products j inversely proportional to the product-level quality, in particular*

$$C^e(j) = \frac{AN^e}{A(j)}.$$

Under perfect foresight, the efficient real interest rate is implicitly defined by the Euler equation after substituting in efficient consumption:

$$r_t^e = -\log \beta - (1 - \rho_A) \log A_t$$

With Lemmas A1, A2, and A3, we are ready to prove Proposition 2. It is repeated here for convenience.

PROPOSITION A1. *Let $U - U^e$ be the central bank's utility gap relative to the utility under efficient allocation expressed in efficient-consumption-equivalent units. The utility gap can be expressed as a function of the average welfare-relevant markup ($\overline{\mu}$), the markup dispersion*

(ζ^{μ}), *and price adjustment costs as*

$$U - U^e = \underbrace{\underbrace{-\log\overline{\mu} - \left(\frac{1}{\overline{\mu}} - 1\right)}_{\text{Average markup}} - \frac{1}{\overline{\mu}}\underbrace{(\zeta^{\mu} - 1)}_{\text{Markup dispersion}}}_{\text{Misallocation}} - \underbrace{\eta g^0}_{\text{Adjustment costs}}, \tag{A35}$$

where the markup dispersion is $\zeta^{\mu} \equiv \int (\mu(j)/\overline{\mu})^{-\epsilon} g(\log\mu(j) - \log\overline{\mu}) dj$, *and* ηg^0 *are the price adjustment costs in labor units.*

PROOF. We can express the difference between utility (U) and utility in the efficient equilibrium (U^e) as

$$\begin{aligned}
U - U^e &= (\log C - \chi N) - (\log C^e - N^e) \\
&= (\log Y - \chi N) - (\log A - 1) \\
&= -\log\overline{\mu} - \frac{Y}{A}\int e^{p(-\epsilon)} g(p)\, dp - \eta g^0 + 1 \\
&= -\log\overline{\mu} - \left(\frac{1}{\overline{\mu}}\zeta^{\mu} - 1\right) - \eta g^0 \\
&= -\log\overline{\mu} - \left(\frac{1}{\overline{\mu}} - 1\right) - \frac{1}{\overline{\mu}}(\zeta^{\mu} - 1) - \eta g^0
\end{aligned}$$

where U^e is the utility in the efficient equilibrium. The first step of the derivation uses the definition of output $Y = C$ and the efficient output $Y^e = A$. The second step uses Lemma A1 and the labor market equilibrium (27). The third step uses Lemma A2. □

Divine coincidence for TFP shocks - Proposition 4. Notice that both the objective (A28) and the constraint (A23) are independent of TFP A. Thus, the solution to the central bank's optimization problem—the optimal level of inflation—is independent of TFP.

Appendix B. Optimality condition of the reset price

If the *post-decision* value function $V(\cdot)$ is convex, the optimal reset price is fully characterized by the system of first-order conditions in Section 2.3.[41] This appendix presents the derivation of $V_t'(0)$.

To start, we reproduce the value function presented in equation (19), which we then rewrite using $\hat{\Phi}(\cdot)$ to denote the c.d.f. of $\hat{\phi}(\cdot)$

$$\begin{aligned}
V_t(x) &= \Pi_t(x) + \Lambda_{t,t+1}\int_{s_{t+1}}^{S_{t+1}} \left[V_{t+1}(x')\hat{\phi}\left(x - x' - \pi^*_{t+1}\right)\right] dx' \\
&\quad + \Lambda_{t,t+1}\left(1 - \int_{s_{t+1}}^{S_{t+1}} \left[\hat{\phi}\left(x - x' - \pi^*_{t+1}\right)\right] dx'\right)\left(V_{t+1}(0) - \eta w_{t+1}\right) \\
V_t(x) &= \Pi_t(x) + \Lambda_{t,t+1}\int_{s_{t+1}}^{S_{t+1}} \left[V_{t+1}(x')\hat{\phi}\left(x - x' - \pi^*_{t+1}\right)\right] dx'
\end{aligned}$$

[41] We verify convexity ex post.

$$+\Lambda_{t,t+1}\left(1-\left[\hat{\Phi}\left(x-s_{t+1}-\pi^*_{t+1}\right)-\hat{\Phi}\left(x-S_{t+1}-\pi^*_{t+1}\right)\right]\right)\left(V_{t+1}(0)-\eta w_{t+1}\right)$$

Taking the derivative of $V_t(x)$ with respect to x and reformulating, we get $V'_t(x)$:

$$\begin{aligned}
V'_t(x) &= \Pi'_t(x)+\Lambda_{t,t+1}\frac{\partial\int_{s_{t+1}}^{S_{t+1}}V_{t+1}(x')\hat{\phi}\left(x-x'-\pi^*_{t+1}\right)dx'}{\partial x}\\
&\quad+\Lambda_{t,t+1}\left(\hat{\phi}\left(x-S_{t+1}-\pi^*_{t+1}\right)-\hat{\phi}\left(x-s_{t+1}-\pi^*_{t+1}\right)\right)\left(V_{t+1}(0)-\eta w_{t+1}\right)\\
&= \Pi'_t(x)+\Lambda_{t,t+1}\int_{s_{t+1}}^{S_{t+1}}V_{t+1}(x')\frac{\partial\hat{\phi}\left(x-x'-\pi^*_{t+1}\right)}{\partial x}dx'\\
&\quad+\Lambda_{t,t+1}\left(\hat{\phi}\left(x-S_{t+1}-\pi^*_{t+1}\right)-\hat{\phi}\left(x-s_{t+1}-\pi^*_{t+1}\right)\right)\left(V_{t+1}(0)-\eta w_{t+1}\right)
\end{aligned}$$

which must be evaluated at $x=0$.

Appendix C. Optimal response to TFP shocks: divine coincidence

This appendix proves Proposition 5. It shows that, in response to a TFP shock, optimal timeless commitment policy keeps inflation at its steady-state level $\pi_t=\pi$.

The central bank's problem after eliminating w_t, π_t and $g_t(\cdot)$ is:

$$\max_{\substack{\left\{g^c_t(\cdot),g^0_t,V_t(\cdot),C_t,N_t,\\ p^*_t,s_t,S_t,\pi^*_t,\Delta_t\right\}_{t=0}^{\infty}}}\ \sum_{t=0}^{\infty}\beta^t\left(\log\left(C_t\right)-N_t\right)$$

subject to

$$\begin{aligned}
N_t &= \left[\frac{\alpha}{(1-\alpha)}\right]^{1-\alpha}\Delta_t C_t^{\alpha}A_t^{-1}\left[1-\left(\frac{(1-\alpha)C_t}{\alpha}\right)^{\alpha}A_t^{-1}\Delta_t\right]^{-1}+\eta g^0_t\\
\Delta_t &= \int_{s_t}^{S_t}e^{\left(x+p^*_t\right)(-\epsilon)}g^c_t\left(x+p^*\right)dx+g^0_t e^{\left(p^*_t\right)(-\epsilon)}\\
V_t(x) &= \left(\left(e^{x+p^*_t}\right)^{1-\epsilon}-\frac{(1-\tau_t)C_t^{\alpha}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}A_t}\left(e^{x+p^*_t}\right)^{-\epsilon}\right)C_t\left[1-\left(\frac{(1-\alpha)C_t}{\alpha}\right)^{\alpha}A_t^{-1}\Delta_t\right]^{-1}\\
&\quad+\Lambda_{t,t+1}\int_{s_{t+1}}^{S_{t+1}}V_{t+1}(x')\hat{\phi}\left(x-x'-\pi^*_{t+1}\right)dx\\
&\quad+\Lambda_{t,t+1}\left[1-\int_{s_{t+1}}^{S_{t+1}}\hat{\phi}\left(x-x'-\pi^*_{t+1}\right)dx'\right]\left[V_{t+1}(0)-\eta C_{t+1}\right],\\
V_t\left(s_t\right) &= V_t\left(0\right)-\eta C_t,\\
V_t\left(S_t\right) &= V_t\left(0\right)-\eta C_t,\\
0 &= \left((1-\epsilon)\left(e^{x+p^*_t}\right)^{1-\epsilon}+\epsilon\frac{(1-\tau_t)C_t^{\alpha}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}A_t}\left(e^{x+p^*_t}\right)^{-\epsilon}\right)C_t\left[1-\left(\frac{(1-\alpha)C_t}{\alpha}\right)^{\alpha}A_t^{-1}\Delta_t\right]^{-1}\\
&\quad+\Lambda_{t,t+1}\int_{s_{t+1}}^{S_{t+1}}V_{t+1}(x')\left.\frac{\partial\hat{\phi}\left(x-x'-\pi^*_{t+1}\right)}{\partial x}\right|_{x=0}dx'\\
&\quad+\Lambda_{t,t+1}\left(\hat{\phi}\left(-S_{t+1}-\pi^*_{t+1}\right)-\hat{\phi}\left(-s_{t+1}-\pi^*_{t+1}\right)\right)\left(V_{t+1}(0)-\eta C_{t+1}\right).
\end{aligned}$$

$$
\begin{aligned}
g_t^c(x) &= \int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1})\hat{\phi}\left(x_{-1}-x-\pi_t^*\right)dx_{-1} + g_{t-1}^0\hat{\phi}\left(-x-\pi_t^*\right),\\
g_t^0 &= 1-\int_{s_t}^{S_t} g_t^c(x)dx,\\
1 &= \int_{s_t}^{S_t} e^{\left(x+p_t^*\right)(1-\epsilon)} g_t^c(x)\,dx + g_t^0 e^{\left(p_t^*\right)(1-\epsilon)},
\end{aligned}
$$

given some initial conditions for $g_{-1}^c(x)$ and g_{-1}^0.

We now transform it in a convenient fashion. First, normalize the constraints involving $V_t(x)$ by $A_t^{1/\alpha}$, divide and multiply the expectation terms by $A_{t+1}^{1/\alpha}$ and substitute for the discount factor $\Lambda_{t,t+1} = \beta\frac{C_t}{C_{t+1}}$. With this, the constraints involving $V_t(x)$ become:

$$
\begin{aligned}
\frac{V_t(x)}{A_t^{1/\alpha}} &= \left(\left(e^{x+p_t^*}\right)^{1-\epsilon} - \frac{1-\tau_t}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\frac{C_t^\alpha}{A_t}\left(e^{x+p_t^*}\right)^{-\epsilon}\right)\frac{C_t}{A_t^{1/\alpha}}\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\frac{C_t^\alpha}{A_t}\Delta_t\right]^{-1}\\
&\quad+\beta\frac{A_{t+1}^{1/\alpha}}{A_t^{1/\alpha}}\frac{C_t}{C_{t+1}}\int_{s_{t+1}}^{S_{t+1}}\left[\frac{V_{t+1}(x')}{A_{t+1}^{1/\alpha}}\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)\right]dx'\\
&\quad+\frac{A_{t+1}^{1/\alpha}}{A_t^{1/\alpha}}\beta\frac{C_t}{C_{t+1}}\left(1-\int_{s_{t+1}}^{S_{t+1}}\left[\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)\right]dx'\right)\left[\left(\frac{V_{t+1}(0)}{A_{t+1}^{1/\alpha}}-\eta\frac{C_{t+1}}{A_{t+1}^{1/\alpha}}\right)\right],\\
\frac{V_t(s_t)}{A_t^{1/\alpha}} &= \frac{V_t(0)}{A_t^{1/\alpha}}-\eta\frac{C_t}{A_t^{1/\alpha}},\\
\frac{V_t(S_t)}{A_t^{1/\alpha}} &= \frac{V_t(0)}{A_t^{1/\alpha}}-\eta\frac{C_t}{A_t^{1/\alpha}},\\
0 &= \left((1-\epsilon)\left(e^{x+p_t^*}\right)^{1-\epsilon} + \epsilon\frac{1-\tau_t}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\frac{C_t^\alpha}{A_t}\left(e^{x+p_t^*}\right)^{-\epsilon}\right)\frac{C_t}{A_t^{1/\alpha}}\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\frac{C_t^\alpha}{A_t}\Delta_t\right]^{-1}\\
&\quad+\beta\frac{A_{t+1}^{1/\alpha}}{A_t^{1/\alpha}}\frac{C_t}{C_{t+1}}\int_{s_{t+1}}^{S_{t+1}}\frac{V_{t+1}(x')}{A_{t+1}^{1/\alpha}}\left.\frac{\partial\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)}{\partial x}\right|_{x=0}dx'\\
&\quad+\beta\frac{A_{t+1}^{1/\alpha}}{A_t^{1/\alpha}}\frac{C_t}{C_{t+1}}\left(\hat{\phi}\left(-S_{t+1}-\pi_{t+1}^*\right)-\hat{\phi}\left(-s_{t+1}-\pi_{t+1}^*\right)\right)\left(\frac{V_{t+1}(0)}{A_{t+1}^{1/\alpha}}-\eta\frac{C_{t+1}}{A_{t+1}^{1/\alpha}}\right).
\end{aligned}
$$

Second, define $\frac{V_t(x)}{A_t^{1/\alpha}} \equiv \hat{V}_t(x)$, and rewrite $\frac{C_t^\alpha}{A_t} = \left(\frac{C_t}{A_t^{1/\alpha}}\right)^\alpha$ so that these constraints become

$$
\begin{aligned}
\hat{V}_t(x) &= \left(\left(e^{x+p_t^*}\right)^{1-\epsilon} - \frac{1-\tau_t}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\left(\frac{C_t}{A_t^{1/\alpha}}\right)^\alpha\left(e^{x+p_t^*}\right)^{-\epsilon}\right)\frac{C_t}{A_t^{1/\alpha}}\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\left(\frac{C_t}{A_t^{1/\alpha}}\right)^\alpha\Delta_t\right]^{-1}\\
&\quad+\beta\frac{A_{t+1}^{1/\alpha}}{C_{t+1}}\frac{C_t}{A_t^{1/\alpha}}\int_{s_{t+1}}^{S_{t+1}}\left[\hat{V}_{t+1}(x')\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)\right]dx'
\end{aligned}
$$

$$+\beta\frac{A_{t+1}^{1/\alpha}}{C_{t+1}}\frac{C_t}{A_t^{1/\alpha}}\left(1-\int_{s_{t+1}}^{S_{t+1}}\left[\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)\right]dx'\right)\left[\left(\hat{V}_{t+1}(0)-\eta\frac{C_{t+1}}{A_{t+1}^{1/\alpha}}\right)\right],$$

$$\hat{V}_t(s_t) = \frac{V_t(0)}{A_t^{1/\alpha}}-\eta\frac{C_t}{A_t^{1/\alpha}},$$

$$\hat{V}_t(S_t) = \frac{V_t(0)}{A_t^{1/\alpha}}-\eta\frac{C_t}{A_t^{1/\alpha}},$$

$$\begin{aligned}0 = &\left((1-\epsilon)\left(e^{x+p_t^*}\right)^{1-\epsilon}+\epsilon\frac{1-\tau_t}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\left(\frac{C_t}{A_t^{1/\alpha}}\right)^\alpha\left(e^{x+p_t^*}\right)^{-\epsilon}\right)\frac{C_t}{A_t^{1/\alpha}}\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\left(\frac{C_t}{A_t^{1/\alpha}}\right)^\alpha\Delta_t\right]^{-1}\\
&+\beta\frac{A_{t+1}^{1/\alpha}}{C_{t+1}}\frac{C_t}{A_t^{1/\alpha}}\int_{s_{t+1}}^{S_{t+1}}\hat{V}_{t+1}(x')\left.\frac{\partial\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)}{\partial x}\right|_{x=0}dx'\\
&+\beta\frac{A_{t+1}^{1/\alpha}}{C_{t+1}}\frac{C_t}{A_t^{1/\alpha}}\left(\hat{\phi}\left(-S_{t+1}-\pi_{t+1}^*\right)-\hat{\phi}\left(-s_{t+1}-\pi_{t+1}^*\right)\right)\left(\hat{V}_{t+1}(0)-\eta\frac{C_{t+1}}{A_{t+1}^{1/\alpha}}\right).\end{aligned}$$

Finally, define $\hat{C}_t=\frac{C_t}{A_t^{1/\alpha}}$. The central bank's complete problem becomes

$$\max_{\substack{\left\{g_t^c(\cdot),g_t^0,V_t(\cdot),C_t,N_t,\\ p_t^*,s_t,S_t,\pi_t^*,\Delta_t\right\}_{t=0}^\infty}}\ \sum_{t=0}^\infty\beta^t\left(\log\left(\hat{C}\right)+\log\left(A_t^{1/\alpha}\right)-\chi N_t\right)$$

subject to

$$N_t = \left[\frac{\alpha}{(1-\alpha)}\right]^{1-\alpha}\Delta_t\hat{C}_t^\alpha\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\hat{C}_t^\alpha\Delta_t\right]^{-1}+\eta g_t^0$$

$$\Delta_t = \int_{s_t}^{S_t}e^{\left(x+p_t^*\right)(-\epsilon)}g_t^c\left(x+p_t^*\right)dx+g_t^0e^{\left(p_t^*\right)(-\epsilon)}$$

$$\begin{aligned}\hat{V}_t(x) = &\left(\left(e^{x+p_t^*}\right)^{1-\epsilon}-\frac{1-\tau_t}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\hat{C}_t^\alpha\left(e^{x+p_t^*}\right)^{-\epsilon}\right)\hat{C}_t\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\hat{C}_t^\alpha\Delta_t\right]^{-1}\\
&+\beta\frac{\hat{C}_t}{\hat{C}_{t+1}}\int_{s_{t+1}}^{S_{t+1}}\left[\hat{V}_{t+1}(x')\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)\right]dx'\\
&+\beta\frac{\hat{C}_t}{\hat{C}_{t+1}}\left(1-\int_{s_{t+1}}^{S_{t+1}}\left[\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)\right]dx'\right)\left[\left(\hat{V}_{t+1}(0)-\eta\hat{C}_{t+1}\right)\right],\end{aligned}$$

$$\hat{V}_t(s_t) = \frac{V_t(0)}{A_t^{1/\alpha}}-\eta\hat{C}_t,$$

$$\hat{V}_t(S_t) = \frac{V_t(0)}{A_t^{1/\alpha}}-\eta\hat{C}_t,$$

$$\begin{aligned}0 = &\left((1-\epsilon)\left(e^{x+p_t^*}\right)^{1-\epsilon}+\epsilon\frac{1-\tau_t}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\hat{C}_t^\alpha\left(e^{x+p_t^*}\right)^{-\epsilon}\right)\hat{C}_t\left[1-\left(\frac{1-\alpha}{\alpha}\right)^\alpha\hat{C}_t^\alpha\Delta_t\right]^{-1}\\
&+\beta\frac{\hat{C}_t}{\hat{C}_{t+1}}\int_{s_{t+1}}^{S_{t+1}}\hat{V}_{t+1}(x')\left.\frac{\partial\hat{\phi}\left(x-x'-\pi_{t+1}^*\right)}{\partial x}\right|_{x=0}dx'\end{aligned}$$

$$+\beta\frac{\hat{C}_t}{\hat{C}_{t+1}}\left(\hat{\phi}\left(-S_{t+1}-\pi^*_{t+1}\right)-\hat{\phi}\left(-s_{t+1}-\pi^*_{t+1}\right)\right)\left(\hat{V}_{t+1}(0)-\eta\hat{C}_{t+1}\right).$$

$$g^c_t(x) = \int_{s_{t-1}}^{S_{t-1}} g^c_{t-1}(x_{-1})\hat{\phi}\left(x_{-1}-x-\pi^*_t\right)dx_{-1}+g^0_{t-1}\hat{\phi}\left(-x-\pi^*_t\right),$$

$$g^0_t = 1-\int_{s_t}^{S_t} g^c_t(x)dx,$$

$$1 = \int_{s_t}^{S_t} e^{\left(x+p^*_t\right)(1-\epsilon)}g^c_t(x)\,dx+g^0_t e^{\left(p^*_t\right)(1-\epsilon)}.$$

Notice that TFP A_t only appears in the objective in a separable way. Therefore, the redefined Ramsey policy is independent of TFP shocks. Going back to the original variables definition, this implies that under optimal policy $C_t \propto A_t^{1/\alpha}$ and $V_t(x) \propto A_t^{1/\alpha}$ while all other variables remain constant at their steady-state values. Thus, inflation π_t also remains constant at its steady-state value.

Appendix D. Optimal monetary policy: additional results

Here we investigate two additional results: optimal long-run inflation and the time-inconsistency of the Ramsey-optimal monetary policy.

D.1. Ramsey steady state inflation

The solution of the Ramsey planner's problem has a steady state featuring a slightly positive inflation of 0.3 basis points per annum.[42] This is different from the standard New Keynesian model with Calvo pricing (Galí 2008), where the optimal inflation in the Ramsey steady state is exactly zero – not only in the model's linear-quadratic approximation but also in its fully nonlinear form. The value of inflation in the Ramsey steady state in the menu cost model is very close to the value of inflation that maximizes steady-state welfare, which in turn is very close to the value of inflation that minimizes the frequency of price adjustments.

	Baseline	Nonlinear Calvo	CalvoPlus	Random menu cost
Ramsey steady state inflation	0.3bps	0	-2.6bps	-1.1bps

The table shows that Ramsey steady state inflation in state-dependent models is very close to zero.

What explains the deviation from zero optimal inflation? One relevant factor is the asymmetry of the profit function (20). For a firm, a negative price gap is more undesirable than a positive price gap of the same size because a negative price gap leads to a large sales increase at a markup loss, while a positive price gap leads to only a somewhat smaller drop in sales at a markup gain of x. This implies that the (S, s)

[42] In our numerical exploration, we have only found a single steady state.

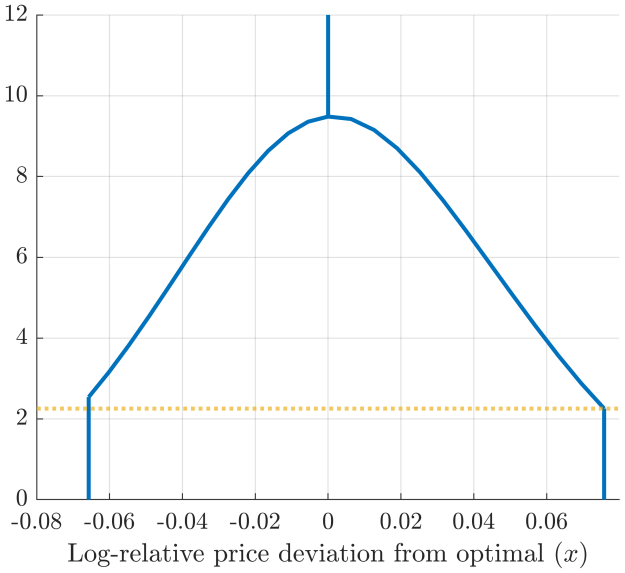

FIGURE A1. Steady-state price-gap density.

The figure displays the steady-state price-gap density $g(x)$ with zero inflation. The dashed yellow line indicates the mass of firms at the upper threshold of the (S, s) band.

band is asymmetric: the lower threshold s_t is closer to the optimal price than the upper one S_t.

Thus, in the zero inflation steady state, there is more mass of firms close to the lower threshold of the inaction band than to the upper threshold. As a result, there are more upward than downward price adjustments. Small positive inflation raises the optimal reset price p^* and shifts the (S, s) band leftwards and thus reduces the number of upward price movements by more than it increases the number of downward price movements. The frequency of price adjustments decreases and, with it, the distortions caused by menu costs. Quantitatively, however, this effect is small.

Additionally, inflation has a nonlinear and asymmetric impact on other distortions – like the price dispersion (25) – as well as equilibrium conditions – like the aggregate price level (24). They affect optimal inflation in complex ways, but their overall impact is small. Indeed, optimal inflation is close to zero also in alternative state-dependent models, but there it is slightly negative. In the CalvoPlus framework optimal steady state inflation is -2.6 basis points, while in the random menu cost model it is -1.1 basis points.

D.2. Time-zero problem

We now turn to investigating the time inconsistency of optimal policy. To assess its magnitude, we solve the optimal policy problem, starting from the price distribution in the Ramsey steady state, assuming that the central bank acts without prior commitment. In this case, the Lagrange multipliers associated with forward-looking equations are initially set to zero. This problem is often referred to as the "time-zero problem"

(Woodford 2003).

The solid blue lines in Figure A2 show the time path under the optimal policy. The subsidy is halved in this exercise, which, therefore, ceases to offset markup distortions caused by the firms' market power. The steady state of the Ramsey policy is time-inconsistent: without pre-commitment, the central bank engineers a temporary expansion. In doing so, it raises welfare by bringing output closer to its efficient level at a cost of elevated pricing distortions arising from the higher inflation.

The dashed red line on Figure A2 shows the equivalent time-zero response in the Calvo model. The figure shows that the incentive to surprise is huge in the Calvo model, also amplified by the roundabout production structure, which raises the 'effective' markup (Baqaee and Farhi 2020) and with it the desire to surprise. Notably, the effects become substantially weaker and less extreme in the menu cost model: both the inflation and output gap increases are smaller relative to the Calvo model. The reason is that the price level becomes more flexible in the state-dependent model: the unexpected easing causes a sizable inflation spike, which causes an increase in the frequency of price changes. As a result, the output gap increases by less than it would under exogenous frequency. That is, the output boost from a given amount of inflation is lower than under Calvo. Since, as we saw before, the central bank's objective function is not significantly different from that under Calvo, the central bank eases less aggressively.

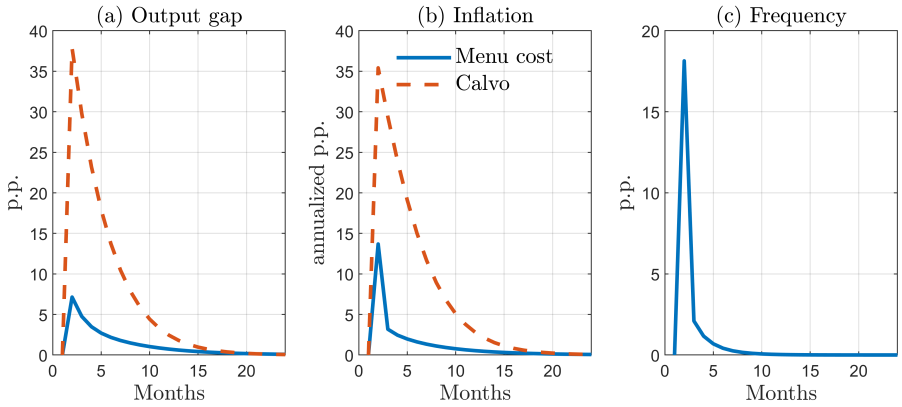

FIGURE A2. Time inconsistency of the optimal policy.

The figure compares the time-0 optimal policies in the menu cost model and in the Calvo model. Inflation is annualized as 12π.

There is a countervailing force that raises the time inconsistency in our baseline model relative to the Calvo model. Namely, due to the idiosyncratic shocks which generate markup dispersion even in steady state, the uniform subsidy τ is insufficient to fully offset the markups for all firms in the steady state. A time-0 optimal policy, therefore, stays time inconsistent for any uniform subsidy τ. This is different in the Calvo model, which features no markup dispersion in the zero inflation steady state. Here an appropriately set τ can *fully* offset the steady state markup. The time-zero problem vanishes and the stationary Ramsey plan becomes time-consistent.

However, when the subsidy τ is set so as to close the markup on average, as in our baseline calibration, the time-zero problem becomes quantitatively very small: the optimal policy easing in this scenario is two orders of magnitude smaller than those under no subsidy (not shown). Therefore, this channel is too weak to counteract the opposite effect caused by the more flexible price level detailed above.

A corollary to the negligibility of the time inconsistency with an appropriate subsidy is that the analysis in section 4, where we adopted a timeless perspective, would remain valid without any quantitatively relevant changes even if we adopted a time-zero perspective.

Appendix E. Additional tables and figures

E.1. Robustness: calibration

	Baseline	CalvoPlus	Random MC	Description	Reference
			Households and final producer		
β	$0.96^{1/12}$	$0.96^{1/12}$	$0.96^{1/12}$	Discount rate	Golosov and Lucas (2007)
χ	1	1	1	Utility weight on labor	Set so that $w = C$
ϵ	6	6	6	Elasticity of substitution	Gagliardone et al. (2025)
	Differentiated goods producers				
α	0.2	0.2	0.2	Labor share	Nakamura and Steinsson (2010)
τ	0.1672	0.1685	0.1709	Subsidy in ss	$y = y^*$ in ss
$1-\theta$	0	8.7%/2	8.7%/2	Share of 0 menu cost in ss	As in 2-product firms in Midrigan (2011)
η	0.064	0.468	2.40	Menu cost	Calibrated to match
σ	0.0306	0.0328	0.0365	Std dev. of quality shocks	moments of steady state
σ_1/σ_2	0.065	0.065	0.065	Ratio of low vs. high stdev	price change distribution
ϖ	0.934	0.926	0.9095	Share of low volat. shock	at 2% inflation
		Aggregate shocks			
ρ_A	$0.95^{1/3}$	$0.95^{1/3}$	$0.95^{1/3}$	TFP shock persistence	Smets and Wouters (2007)
ρ_τ	$0.9^{1/3}$	$0.9^{1/3}$	$0.9^{1/3}$	Cost-push persistence	Smets and Wouters (2007)

TABLE A1. Parameters of the alternative state-dependent models

	Baseline	$\alpha = 1$	$\alpha = 1, \sigma_1 = \sigma_2$	Description	Reference
			Households and final producer		
β	$0.96^{1/12}$	$0.96^{1/12}$	$0.96^{1/12}$	Discount rate	Golosov and Lucas (2007)
χ	1	1	1	Utility weight on labor	Set so that $w = C$
ϵ	6	6	6	Elasticity of substitution	Gagliardone et al. (2025)
	Differentiated goods producers				
α	0.2	1	1	Labor share	Nakamura and Steinsson (2010)
η	0.064	0.013	0.035	Menu cost	Calibrated to match
σ	0.0306	0.0306	0.0252	Std dev. of quality shocks	moments of steady state
σ_1/σ_2	0.065	0.065	1	Ratio of low vs. high stdev	price change distribution
ϖ	0.934	0.934	0.5	Share of low volat. shock	at 2% inflation
		Aggregate shocks			
ρ_A	$0.95^{1/3}$	$0.95^{1/3}$	$0.95^{1/3}$	TFP shock persistence	Smets and Wouters (2007)
ρ_τ	$0.9^{1/3}$	$0.9^{1/3}$	$0.9^{1/3}$	Cost-push persistence	Smets and Wouters (2007)

TABLE A2. Parameters of the models without roundabout economy ($\alpha = 1$) and Gaussian idiosyncratic shocks ($\sigma_1 = \sigma_2$)

	Baseline	Golosov-Lucas	Simple	Description	Reference
			Households and final producer		
β	$0.96^{1/12}$	$0.96^{1/12}$	0	Discount rate	Golosov and Lucas (2007)
χ	1	1	1	Utility weight on labor	Set so that $w = C$
ϵ	6	6	6	Elasticity of substitution	Gagliardone et al. (2025)
	Differentiated goods producers				
α	0.2	1	1	Labor share	Nakamura and Steinsson (2010)
τ	0.1672	0.1672	0.1667	Subsidy in ss	$y = y^*$ in ss
η	0.064	0.007	0.0027	Menu cost	Calibrated to match
σ	0.0306	0.011	0.0182	Std dev. of quality shocks	moments of steady state
σ_1/σ_2	0.065	1	1	Ratio of low vs. high stdev	price change distribution
ϖ	0.934	0.5	0.5	Share of low volat. shock	at 2% inflation
	Aggregate shocks				
ρ_A	$0.95^{1/3}$	$0.95^{1/3}$	0	TFP shock persistence	Smets and Wouters (2007)
ρ_τ	$0.9^{1/3}$	$0.9^{1/3}$	0	Cost-push persistence	Smets and Wouters (2007)

TABLE A3. Parameters of the canonical Golosov-Lucas model and the simple model

TABLE A4. Moments of the canonical Golosov-Lucas model and the simple model

	Steady state at 2% inflation			10% inflation		Phillips curve
	Frequency	Size	Kurtosis	Freq. incr.	Shock size	slope
Data	8.7%	8.5%	4	14pp		0.006-0.106
Baseline	8.7%	8.5%	4	14pp	3.5%	0.08
Simple	8.7%	3.77%	3.84	14pp	13.0%	0.56
Conventional GL	8.7%	3.81%	7.13	14pp	11.9%	0.95

The table shows that the simple model and the conventional Golosov and Lucas (2007) model that match the frequency increase at 10% inflation underestimate the steady state size of the price changes. The calibration targets are from Nakamura and Steinsson (2008) (frequency and size), Alvarez et al. (2016) (kurtosis), Montag and Villar (2025) (frequency increase), and Hazell et al. (2022) (PC slope).

E.2. Optimal response to demand shocks

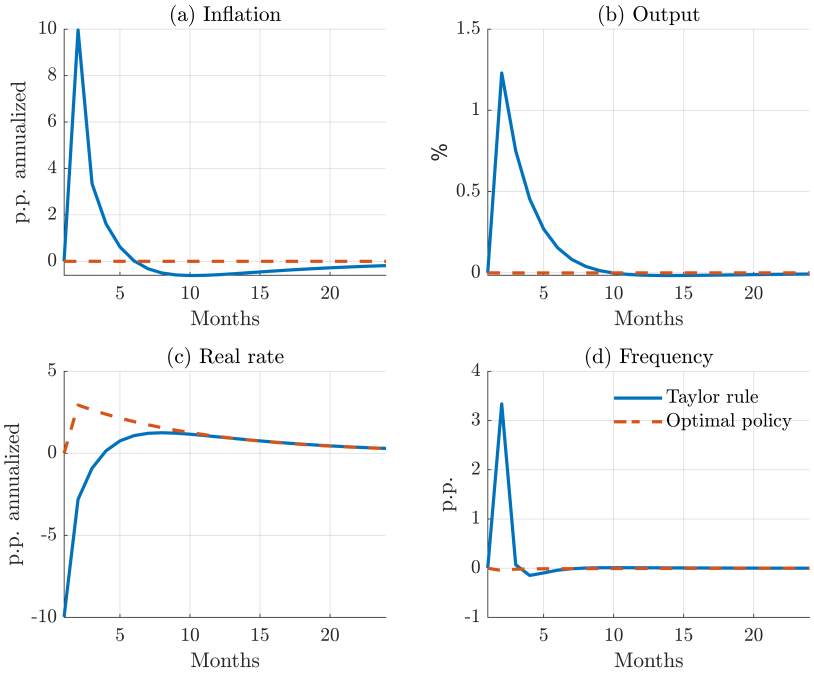

FIGURE A3. Optimal response to time-preference shocks.

The figure displays optimal response to a large shock to the rate of time-preference β_t. The latter drops by annualized 3 percentage points on impact and follows an AR(1) process in levels with persistence $\rho_\beta = \rho_A$. The figure contrasts the optimal policy to the response under a standard Taylor rule. Under a Taylor rule inflation moves considerably, highlighting that the shock is big. Under Ramsey-optimal policy, inflation is nearly flat, while the real rate closely tracks the stochastic discount factor. As a result, the response under Ramsey policy is almost indistinguishable from the response under a strict inflation targeting policy (not shown). Therefore the divine coincidence holds approximately.

Appendix F. Optimal simple rule

In this section, we report the robustness of our analysis in section 4.5 to an alternative way to approximate Ramsey policy by a simple Taylor rule. Instead of calibrating the Taylor rule coefficients to match the stance-inflation relationship, we now calibrate the coefficients to maximize household welfare under cost-push shocks. This approach is known as "optimal simple rules".

Nonlinear Taylor rule. We focus on the response to inflation ($\phi_{\pi,1}, \phi_{\pi,2}$) and persistence ($\rho_i$) and abstract from a response to output ($\phi_y = 0$). Furthermore, we remove the term $(1 - \rho_i)$ from the rule (33) to allow for super-persistence ($\rho_i > 1$), which is often found to be optimal in the optimal-simple-rules literature. This does not change the space of possible policies but facilitates the interpretation of parameters when $\rho_i > 1$. However, it implies that the parameters we find here can be compared to those in the main text only after adjusting the latter by a factor $(1 - \rho_i)$.

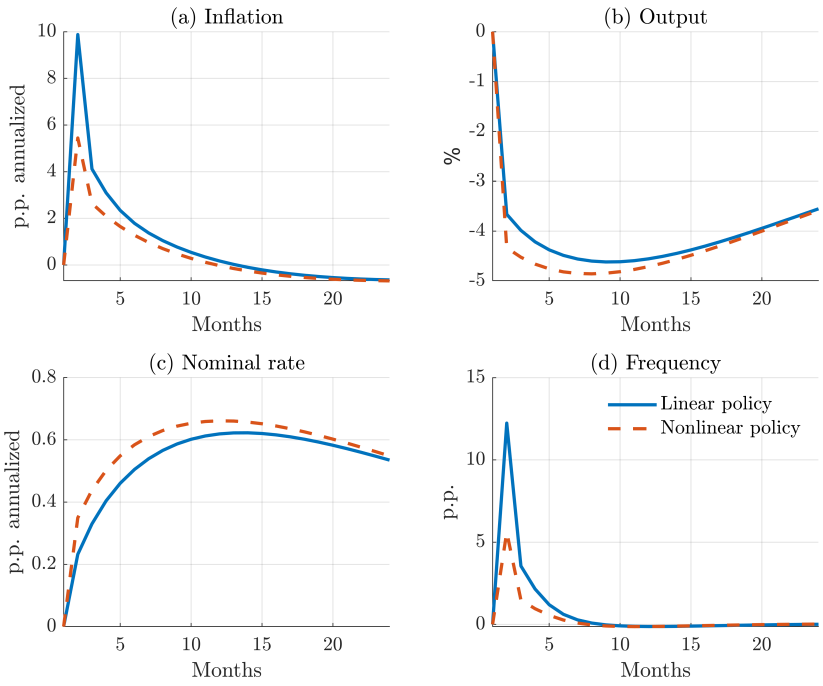

FIGURE A4. Impulse responses under linear and nonlinear optimal Taylor rules.

The figure compares impulse responses under the linear optimized Taylor rule ($\phi_{\pi,2} = 0$, solid blue) and the nonlinear optimized Taylor rule ($\phi_{\pi,2} = 2492$, dashed red) to a cost-push shock that drives annualized inflation to 10% under the linear rule. The nonlinear rule reduces peak inflation by approximately five percentage points at the cost of a lower output path.

$$\log(R_t/R) = \rho_i \log(R_{t-1}/R) + \left[\phi_{\pi,1} + \phi_{\pi,2}\pi_t^2\right]\pi_t, \tag{A36}$$

As before, we calibrate the Taylor rule in two steps. We first find the optimal linear coefficients ($\phi_{\pi,1}$, ρ_i) under a small cost-push shock. Since our Ramsey problem suffers from a mild time-zero problem,[43] we consider two small cost-push shocks of opposing signs, and find the rule that maximizes the average welfare over those two shocks. This guarantees that we find the *timelessly* optimal simple rule. Second, given those coefficients, we find the $\phi_{\pi,2}$ value that maximizes welfare under a shock that generates a 10% impact inflation under the linear policy. The shock τ_t necessary to generate 10% inflation under the linear policy implies a 1.3% increase in cost. Since this shock is big relative to the time-zero problem, we can disregard the latter in step 2. This two-step approach guarantees that the simple nonlinear rule maximizes welfare conditional on being in normal times when shocks are small and only the linear component of the rule is relevant, while at the same time maximizing welfare for large shocks as well.

The optimal coefficients are $\rho_i = 1.000584$, $\phi_{\pi,1} = 0.0278$ and $\phi_{\pi,2} = 2492$. This implies that the inflation coefficient would increase seven-fold from 0.0278 to 0.2 as inflation reaches 10%.

[43] That is, the Ramsey plan is mildly time-inconsistent in steady state. See section D

A simulation of the 2021–2023 inflation surge. As in the main text, we now simulate a cost-push shock large enough to cause 10% inflation under the linear Taylor rule, and contrast the response to that under a nonlinear rule (Figure A4). The nonlinear rule prescribes a tighter policy and implies an inflation rate that is almost 5 percentage points lower – at the peak – than under the linear rule.

The welfare differences between the two policies are also significant. The nonlinear coefficient closes the gap to the Ramsey policy by 84%. That is, the nonlinear Taylor rule closes the welfare gap to Ramsey policy six times more than the linear rule.

The exercise shows that our results are qualitatively robust to this alternative approximation of the Ramsey policy: welfare maximization requires a significantly nonlinear response to inflation, which leads to a considerably lower peak inflation rate. Even quantitatively the results of the inflation-surge simulation are fairly similar. The only caveat is that the rule features super-persistence and is thus harder to interpret and more distant from empirical Taylor rules.

Appendix G. Distribution of price gaps after a large shock

The left panel of figure A5 illustrates the steady-state distribution of price gaps. The mass of non-adjusting firms is shaded light blue. These are the firms with price gaps close to zero (that is, with prices close to the optimal price). The dashed vertical lines mark the boundaries of the inaction region s and S. Outside the inaction region are the adjusting firms: their price gaps are large enough to make it worthwhile to pay the menu cost. Upon adjustment, their price gaps are closed, creating the mass point at $x = 0$.

The right panel of figure A5 illustrates the effect of a large inflationary shock. The light blue line repeats the steady-state distribution. The dark blue line shows the distribution after the inflationary shock has shifted the entire price gap distribution to the left.[44] The inaction thresholds s and S barely move. A large mass of firms decide to increase their prices (the larger white area to the left of s) and a much smaller mass of firms decide to decrease their prices (the smaller area to the right of S). As the pool of potential price-cutters is exhausted (the thinner right tail of the distribution), fewer firms are discouraged from cutting their prices. Hence, the overall fraction of adjusters increases substantially, resulting in a larger mass point at 0. Therefore, aggregate price flexibility rises. This is why in response to large monetary shocks the Phillips curve steepens.

Appendix H. Computational algorithm

This appendix explains the computational method. We use a three-step approach to convert the original infinite-dimensional Ramsey problem into a finite-dimensional

[44] Note that the leftwards shift is not equal to 10% because (i) inflation is annualized and (ii) p^* also responds.

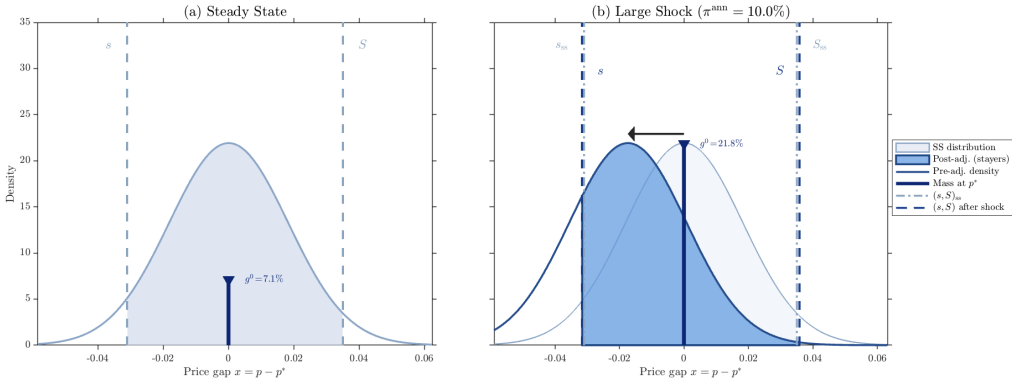

FIGURE A5. Distribution of price gaps in aggregate steady-state and after a large shock

one. First, we approximate the continuous part of the distribution $g^c(x)$ and value functions $V(x)$ by piecewise linear functions over a set of nodes. The mass point of the distribution at zero is taken into account separately, in line with the notation in the paper. Second, we use dynamic endogenous nodes, such that the grid exactly covers the inaction region $[s_t, S_t]$ and that both bounds of the inaction region ($x_t = s_t, S_t$) and the optimal reset price ($x_t = 0$) are "on the grid" for each t. Third, given this approximation, we evaluate the integrals capturing expectations over the idiosyncratic shock analytically. Step one makes the problem finite-dimensional. Having discretized the value function and the law of motion of the distribution in this way, we require those two conditions to hold at each node of the (endogenous time-varying) grid exactly at each t. This transforms the private equilibrium conditions into a large system of difference equations. Steps two and three ensure that the approximation is accurate, smooth and computationally efficient. We explain those steps in detail below.

Once we have converted the central bank's infinite-dimensional problem into a finite-dimensional problem in this way, we derive the central bank's first-order conditions. For this we use symbolic differentiation, and in particular, Dynare's `Ramsey` command. We are now left with an even larger system of difference equations, as we have FOCs for the value and distribution functions at each grid point, and the associated Lagrange multipliers. This system is made up of $(2N + n)$ equations for each t, where n is the number of scalar equilibrium conditions, N is the number of points for the x-grid (used for $g^c(x)$ and $V(x)$).

Next, we find the Ramsey steady state. To do so, we use the discretized steady-state private equilibrium conditions to construct a nonlinear multidimensional function mapping inflation to the rest of variables. We then combine this function with the steady state version of the planner's FOCs. As this system is linear in Lagrange multipliers, finding its solution boils down to finding the root of a nonlinear univariate function in inflation. To do so we use the Newton method. This step is performed by Dynare's `steady` command. The Ramsey steady state provides us with the appropriate initial and terminal conditions for the dynamic Ramsey plan, and with an initial guess

for the transition paths.

Finally, we compute the dynamics of the Ramsey plan. To do so, we solve the large system of difference equations nonlinearly in the sequence space under perfect foresight. This system is of size $(2N+n)T$, where T is the simulation horizon. Again, we employ a standard Newton method using Dynare's `perfect foresight solver` command.

The rest of the appendix explains those steps that go beyond straightforward applications of existing methods. It is organized as follows. First, we explain how to make the central bank's problem finite-dimensional. For this purpose, we first define some useful auxiliary functions in Section H.1. Then we transform the equilibrium conditions to apply an endogenous grid and approximate the value and distribution functions by a piecewise linear function in Section H.2. Finally, we evaluate the integrals analytically in Section H.3. The result is a discrete set of equations that can conveniently be represented in matrix form, which we summarize in Section H.4. Second, we explain how we determine the steady state in Section H.5.

Notice that we have turned off roundabout production and fat-tailed shocks in this appendix for ease of exposition. These two extensions are straightforward and do not affect the general idea of the method.

H.1. Preliminaries

To begin with, let us normalize the variable x_t as

$$\mathsf{x}_t = \begin{cases} \frac{x_t}{-s_t} & \text{if } x_t < 0 \\ \frac{x_t}{S_t} & \text{otherwise} \end{cases} \tag{A37}$$

Under this normalization, the optimal price is at $\mathsf{x}_t = 0$, the upper limit of the (S,s) band at $\mathsf{x}_t = 1$ and the lower limit of the (S,s) band at $\mathsf{x}_t = -1$. This will later allow us to have all critical points (s_t, S_t, p_t^*) on the grid. The law of motion of x_t conditional on not updating can be derived from $x_t = x_{t-1} - \sigma\varepsilon_t - \pi_t^*$:

$$\mathsf{x}_t = \begin{cases} \frac{x_t}{S_t} = \frac{x_{t-1}-\sigma\varepsilon_t-\pi_t^*}{S_t} = \begin{cases} \frac{x_{t-1}-\sigma\varepsilon_t-\pi_t^*}{S_{t-1}}\frac{S_{t-1}}{S_t} = \mathsf{x}_{t-1}\frac{S_{t-1}}{S_t} - \frac{\sigma\varepsilon_t+\pi_t^*}{S_t} & \text{if } \mathsf{x}_t > 0, \text{if } \mathsf{x}_{t-1} > 0 \\ \frac{x_{t-1}-\sigma\varepsilon_t-\pi_t^*}{-s_{t-1}}\frac{-s_{t-1}}{S_t} = \mathsf{x}_{t-1}\frac{-s_{t-1}}{S_t} - \frac{\sigma\varepsilon_t+\pi_t^*}{S_t} & \text{if } \mathsf{x}_t > 0, \text{if } \mathsf{x}_{t-1} < 0 \end{cases} \\ \frac{x_t}{-s_t} = \frac{x_{t-1}-\sigma\varepsilon_t-\pi_t^*}{-s_t} = \begin{cases} \frac{x_{t-1}-\sigma\varepsilon_t-\pi_t^*}{S_{t-1}}\frac{S_{t-1}}{-s_t} = \mathsf{x}_{t-1}\frac{S_{t-1}}{-s_t} - \frac{\sigma\varepsilon_t+\pi_t^*}{-s_t} & \text{if } \mathsf{x}_t < 0, \text{if } \mathsf{x}_{t-1} > 0 \\ \frac{x_{t-1}-\sigma\varepsilon_t-\pi_t^*}{-s_{t-1}}\frac{-s_{t-1}}{-s_t} = \mathsf{x}_{t-1}\frac{-s_{t-1}}{-s_t} - \frac{\sigma\varepsilon_t+\pi_t^*}{-s_t} & \text{if } \mathsf{x}_t < 0, \text{if } \mathsf{x}_{t-1} < 0 \end{cases} \end{cases} \tag{A38}$$

We now define functions to be used in the next sections to redefine the value and distribution functions. For compactness, let us adopt the notation where $\hat{s}_t(\mathsf{x}_t)$ picks the respective extremes (S,s) depending on the value of x_t following (A37). For brevity, at times we will drop the dependence on x_t and just write $\hat{s}_t \equiv \begin{cases} S_t & \text{if } \mathsf{x}_t > 0 \\ -s_t & \text{if } \mathsf{x}_t < 0 \end{cases}$.

Solving (A38) for x_t, x_{t-1} and ε respectively, we obtain the following relations:

$$x_t = x_{t-1}\frac{\hat{s}_{t-1}}{\hat{s}_t} - \frac{\sigma\varepsilon_t + \pi_t^*}{\hat{s}_t} \tag{A39}$$

$$x_{t-1} = x_t\frac{\hat{s}_t}{\hat{s}_{t-1}} + \frac{\sigma\varepsilon_t + \pi_t^*}{\hat{s}_{t-1}} \tag{A40}$$

$$\varepsilon_t = \frac{\hat{s}_{t-1}x_{t-1} - \hat{s}_t x_t - \pi_t^*}{\sigma} \equiv h_t(x_{t-1}, x_t) \tag{A41}$$

where we have defined $h_t(x_{t-1}, x_t)$ for later use.

H.2. Approximating the distribution and value functions

Now we redefine the value and distribution functions over the variable x and approximate them by piecewise linear functions. The original infinite-dimensional problem of the central bank is laid out in Section 2.6. In the following, we consider each of the equations that contain the distribution and value functions one by one.

H.2.1. Distribution

The distribution function is given by

$$g_t(x) \equiv g_t^c(x) + g_t^0\delta(x).$$

where

$$g_t^c(x) = \begin{cases} \frac{1}{\sigma}\int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1})\phi\left(\frac{x_{-1}-x-\pi_t^*}{\sigma}\right)dx_{-1} + g_{t-1}^0\phi\left(\frac{-x-\pi_t^*}{\sigma}\right), & \text{if } x \in [s_t, S_t], \\ 0, & \text{otherwise,} \end{cases} \tag{A42}$$

$$g_t^0 = 1 - \int_{s_t}^{S_t} g_t^c(x)dx.$$

where $\phi(\cdot)$ is the standard normal pdf. Now we rewrite the distribution using the newly defined re-normalized x where $x = x\hat{s}_t$ as in equation (A37): define $g_t^c(x\hat{s}_t) \equiv g_t^c(x)$ and write

$$g_t^c(x) = \begin{cases} \int_{-1}^{1} \frac{\hat{s}_{t-1}(x')}{\sigma} g_{t-1}^c(x')\phi\left(h_t(x', x)\right)dx' + g_{t-1}^0\phi\left(\frac{-x-\pi_t^*}{\sigma}\right), & \text{if } x \in [-1, 1], \\ 0, & \text{otherwise,} \end{cases} \tag{A43}$$

$$g_t^0 = 1 - \int_{-1}^{1} g_t^c(x)\,\hat{s}_t(x)dx. \tag{A44}$$

To see where this comes from, note that for the latter expression for g_t^0 we have applied a simple change of variable to the integral. In particular, we have used the

following substitution:

$$\int_{s_t}^{S_t} g_t^c(x)dx = \int_{s_t}^{S_t} g_t^c(x\hat{s}_t(x))dx\hat{s}_t(x)$$
$$= \int_{s_t}^{S_t} g_t^c(x)dx\hat{s}_t(x) = \int_{s_t/\hat{s}_t(x)}^{S_t/\hat{s}_t(x)} \hat{s}_t(x)g_t^c(x)dx = \int_{-1}^{1} \hat{s}_t(x)g_t^c(x)dx.$$

Next, we will also change the variable in the integral in the equation for $g_t^c(x)$ (A43). This change of variable is a bit more involved. First, we re-express (A42) as

$$g_t^c(x) = \begin{cases} \frac{1}{\sigma}\int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1})\phi\left(\frac{x_{-1}\hat{s}_{t-1}-x\hat{s}_t-\pi_t^*}{\sigma}\right)d(x_{-1}\hat{s}_{t-1}) + g_{t-1}^0\phi\left(\frac{-x\hat{s}_t-\pi_t^*}{\sigma}\right), & \text{if } x\in[-1,1], \\ 0, & \text{otherwise,} \end{cases}$$

Second, we split the integral in two parts at 0 (and we drop the second line of the above expression for brevity)

$$\begin{aligned} g_t^c(x) \quad = \quad & \frac{1}{\sigma}\int_{s_{t-1}}^{0} g_{t-1}^c(x_{-1})\phi\left(\frac{-x_{-1}s_{t-1}-x\hat{s}_t-\pi_t^*}{\sigma}\right)d(x_{-1}s_{t-1}) \\ & +\frac{1}{\sigma}\int_{0}^{S_{t-1}} g_{t-1}^c(x_{-1})\phi\left(\frac{x_{-1}S_{t-1}-x\hat{s}_t-\pi_t^*}{\sigma}\right)d(x_{-1}S_{t-1}) \\ & +g_{t-1}^0\phi\left(\frac{-x\hat{s}_t-\pi_t^*}{\sigma}\right) \quad \text{if } x\in[-1,1], \end{aligned}$$

Now we do a change of variable: integrate over x_{-1} instead of $x_{-1}s_{t-1}$

$$\begin{aligned} g_t^c(x) \quad = \quad & \int_{-1}^{0} \frac{-s_{t-1}}{\sigma} g_{t-1}^c(x_{-1})\phi\left(\frac{-x_{-1}s_{t-1}-x\hat{s}_t-\pi_t^*}{\sigma}\right)dx_{-1} \\ & +\int_{0}^{1} \frac{S_{t-1}}{\sigma} g_{t-1}^c(x_{-1})\phi\left(\frac{x_{-1}S_{t-1}-x\hat{s}_t-\pi_t^*}{\sigma}\right)dx_{-1} \\ & +g_{t-1}^0\phi\left(\frac{-x\hat{s}_t-\pi_t^*}{\sigma}\right) \quad \text{if } x\in[-1,1], \end{aligned}$$

Finally, pasting the two integrals together again, re-denoting x_{-1} by x' and using $h_t(x',x)$ we get expression (A43). This concludes the explanation of the change of variables.

So far we have rewritten the law of motion of the firm distribution g_t. We now introduce the approximation for g_t.

To approximate $g_t^c(x)$, we first define the variable $x \equiv x/\hat{s}_t$, that is the price gap x normalized by the corresponding threshold value $\hat{s}_t$. Second, we define a piecewise linear function over x with equally spaced nodes $x_1,\ldots,x_I = -1,\ldots,0,\ldots,1$:

$$g_t^c(x|x_i < x < x_{i+1}) = g_t^c(x_i) + \frac{x-x_i}{x_{i+1}-x_i}(g_t^c(x_{i+1}) - g_t^c(x_i)).$$

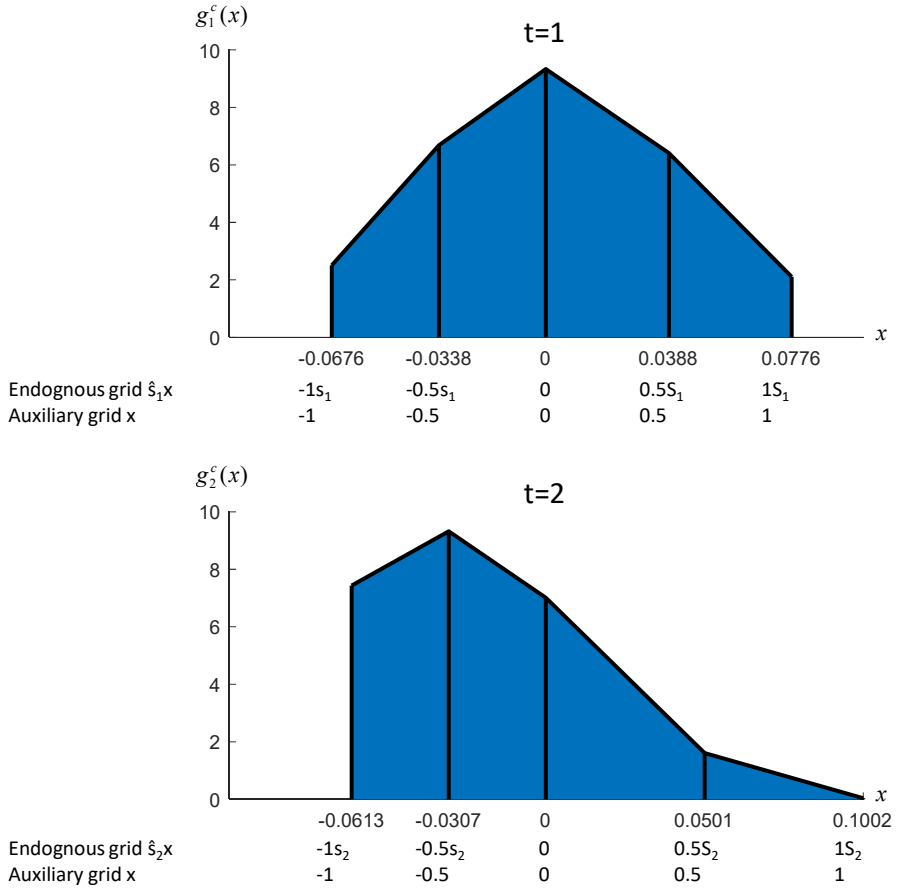

FIGURE A6. This figure schematically explains the linear interpolation with an endogenous grid. It shows the piecewise linearly approximated distribution $g_t^c(x)$ at two points in time, $t = 1$ and $t = 2$. The thresholds of the (S, s) band are not symmetric around 0 and differ across time. The endogenous grid x has I grid points, which are automatically adjusted so that half of the grid points cover the negative part of the (s, S) band and half of them cover the positive part. In this illustrative example $I = 5$ (we use a larger I when solving the model). The adjustment is obtained by multiplying the auxiliary grid $\mathrm{x} = [-1, -0.5, 0, 0.5, 1]$ by $\hat{s}_t(\mathrm{x})$: $x = \mathrm{x}\hat{s}_t$

We refer to the grid for x as the auxiliary grid. It is exogenous. Using the piecewise linear function and the definition of normalized price gap, we now approximate $g_t^c(x)$ by $g_t^c(x/\hat{s}_t) = g_t^c(\mathrm{x})$.

Why did we define the normalized price gap x? The *exogenous* grid for x over which the piecewise linear function is defined maps into an *endogenous grid* for $x = \hat{s}_t\mathrm{x}$. This endogenous grid exactly spans the (s, S) band and has a node at 0 at each t. Thus we waste no grid points outside the inaction region, which is the only interval of x that is relevant for the equilibrium conditions. Figure A6 illustrates the use of linear interpolation with an endogenous grid as we apply it here.

From now on, g_t^c denotes the piecewise linear approximated function, and $g_t^c(\mathrm{x}_i < \mathrm{x} < \mathrm{x}_{i+1})$ denotes a linear piece of it. Thus, the law of motion of the distribution is

approximated as

$$g_t^c(x) = \sum_{i=1}^{I-1} \int_{x_i}^{x_{i+1}} \frac{\hat{s}_{t-1}(x')}{\sigma} g_{t-1}^c(x_i < x' < x_{i+1}) \phi\left(h_t(x', x)\right)\, dx' + \frac{1}{\sigma} g_{t-1}^0 \phi\left(h_t(0, x)\right),$$

$$g_t^0 = 1 - \sum_{i=1}^{I-1} \int_{x_i}^{x_{i+1}} g_t^c(x_i < x < x_{i+1}) \hat{s}_t(x) dx.$$

Notice that in these expressions, the integrands are continuous in the interval $x_i < x < x_{i+1}$ since x and x' are of constant sign.

Also note that the distribution function is 0 outside the (S,s) band. Our piecewise linear g_t^c in fact is only defined over the range where the distribution has positive mass, that is, for $x \in [-1, 1]$. This is computationally efficient.

H.2.2. Other Aggregation Equations

The equilibrium conditions contain two additional aggregation equations that contain the function $g(\cdot)$, for which we use the piecewise linear approximation of $g^c(\cdot)$. Recall the aggregate price index and the labor market clearing condition

$$e^{p_t^*(\epsilon-1)} = \int e^{x(1-\epsilon)} g_t(x) dx,$$

$$N_t = \frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \int e^{x(-\epsilon)} g_t(x) d(x) + \eta g_t(x) d(x)$$

which we approximate as follows, after the change of variable to x,

$$e^{p_t^*(\epsilon-1)} = \sum_{i=1}^{I-1} \int_{x_i}^{x_{i+1}} e^{\hat{s}_t x(1-\epsilon)} g_t^c(x_i < x < x_{i+1}) \hat{s}_t(x) dx + g_t^0,$$

$$N_t = \frac{C_t}{A_t} e^{p^*(-\epsilon)} \sum_{i=1}^{I-1} \left(\int_{x_i}^{x_{i+1}} e^{\hat{s}_t x(-\epsilon)} g_t^c(x_{i-1} < x < x_{i+1}) \hat{s}_t(x) dx + g_t^0 \right) + \eta g_t^0.$$

H.2.3. Value Function

Recall the value function is

$$\begin{aligned} V_t(x) =& \Pi_t(x) + \frac{\Lambda_{t,t+1}}{\sigma} \int_{s_{t+1}}^{S_{t+1}} \left[V_{t+1}(x') \phi\left(\frac{x - x' - \pi_{t+1}^*}{\sigma} \right) \right] dx' \\ &+ \Lambda_{t,t+1} \left(1 - \frac{1}{\sigma} \int_{s_{t+1}}^{S_{t+1}} \left[\phi\left(\frac{x - x' - \pi_{t+1}^*}{\sigma} \right) \right] dx' \right) \left(V_{t+1}(0) - \eta w_{t+1} \right) \end{aligned}$$

We now express it in terms of x with $V_t(x) \equiv V_t(x\hat{s}_t)$:

$$V_t(x) = \Pi_t(x) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{s_{t+1}}^{S_{t+1}}\left[V_{t+1}(x')\phi\left(\frac{x\hat{s}_t - x'\hat{s}_{t+1} - \pi^*_{t+1}}{\sigma}\right)\right]dx'\hat{s}_{t+1}$$
$$+ \Lambda_{t,t+1}\left(1 - \frac{1}{\sigma}\int_{s_{t+1}}^{S_{t+1}}\left[\phi\left(\frac{x\hat{s}_t - x'\hat{s}_{t+1} - \pi^*_{t+1}}{\sigma}\right)\right]dx'\hat{s}_{t+1}\right)(V_{t+1}(0) - \eta w_{t+1})$$

Note that, in equilibrium it must hold that $V_t(0) - \eta\frac{w_t}{A_t} = V_t(-1) = V_t(1)$ and $V_t'(0) = 0$. The first two equalities are straightforward; the next subsection discusses the latter.

After the change of variable to x', which is analogous to the change of variable applied to g_t^c previously, we can rewrite $V_t(x)$ as

$$V_t(x) = \Pi_t(x) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\left[\hat{s}_{t+1}(x')V_{t+1}(x')\phi\left(h_{t+1}(x,x')\right)\right]dx'$$
$$+ \Lambda_{t,t+1}\left(1 - \frac{1}{\sigma}\int_{-1}^{1}\left[\hat{s}_{t+1}(x')\phi\left(h_{t+1}(x,x')\right)\right]dx'\right)(V_{t+1}(0) - \eta w_{t+1})$$

So far, we have normalized the support of the value function. In addition, it is convenient to normalize the value function itself. We normalize the value function by its maximal value $V_t(0)$, and denote the normalized value function by $v_t(x)$: $v_t(x) \equiv V_t(x) - V_t(0)$. The expression above can be re-written as:

$$v_t(x) \equiv V_t(x) - V_t(0) = \Pi_t(x) - \Pi_t(0)$$
$$+ \frac{\Lambda_{t,t+1}}{\sigma}\left(\int_{-1}^{1}\hat{s}_{t+1}(x')\left[V_{t+1}(x')\phi\left(h_{t+1}(x,x')\right) - V_{t+1}(x')\phi\left(h_{t+1}(0,x')\right)\right]dx'\right)$$
$$+ \frac{\Lambda_{t,t+1}}{\sigma}\left(-\int_{-1}^{1}\hat{s}_{t+1}(x')\left[\phi\left(h_{t+1}(x,x')\right) - \phi\left(h_{t+1}(0,x')\right)\right]dx'\right)(V_{t+1}(0) - \eta w_{t+1})$$
$$= \Pi_t(x) - \Pi_t(0)$$
$$+ \frac{\Lambda_{t,t+1}}{\sigma}\left(\int_{-1}^{1}\hat{s}_{t+1}(x')\left[v_{t+1}(x')\left(\phi\left(h_{t+1}(x,x')\right) - \phi\left(h_{t+1}(0,x')\right)\right)\right]dx'\right)$$
$$+ \frac{\Lambda_{t,t+1}}{\sigma}\left(-\int_{-1}^{1}\hat{s}_{t+1}(x')\left[\phi\left(h_{t+1}(x,x')\right) - \phi\left(h_{t+1}(0,x')\right)\right]dx'\right)(-\eta w_{t+1})$$

Following our approach for $g^c(\cdot)$, we approximate $v(\cdot)$ by a piecewise linear function with nodes $x_1, \ldots, x_I = -1, \ldots, 0, \ldots, 1$ with $v_t(x|x_i < x < x_{i+1}) \approx v_t(x_i) + (x - x_i)\frac{v_t(x_{i+1}) - v_t(x_i)}{x_{i+1} - x_i}$.

From now on, v_t denotes the piecewise linear approximated function and $v_t(x_i < x < x_{i+1})$ denotes a linear piece of it. Thus, this function $v_t(x)$ is approximated as

$$v_t(x) = \Pi_t(x) - \Pi_t(0)$$
$$+ \frac{\Lambda_{t,t+1}}{\sigma}\sum_{i=1}^{I-1}\int_{x_i}^{x_{i+1}}\hat{s}_{t+1}(x')v_{t+1}(x_i < x' < x_{i+1})\left(\phi\left(h_{t+1}(x,x')\right) - \phi\left(h_{t+1}(0,x')\right)\right)dx'$$
$$+ \frac{\Lambda_{t,t+1}}{\sigma}(\eta w_{t+1})\int_{-1}^{1}\hat{s}_{t+1}(x')\left(\phi\left(h_{t+1}(x,x')\right) - \phi\left(h_{t+1}(0,x')\right)\right)dx'.$$

H.2.4. Optimality condition for the reset price

We proceed in the same way for the derivative of the value function. We start with

$$0 = V_t'(0) = \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{s_{t+1}}^{S_{t+1}} V_{t+1}(x')\left.\frac{\partial\phi\left(\frac{x-x'-\pi^*_{t+1}}{\sigma}\right)}{\partial x}\right|_{x=0} dx' + \frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(V_{t+1}(0)-\eta w_{t+1})$$

where

$$\left.\frac{\partial\phi\left(\frac{x-x'-\pi^*_{t+1}}{\sigma}\right)}{\partial x}\right|_{x=0} = \frac{1}{\sqrt{2\pi}\sigma}\frac{\pi^*_{t+1}+x'}{\sigma}e^{-\frac{1}{2}\left(\frac{-\pi^*_{t+1}-x'}{\sigma}\right)^2},$$
$$= \frac{\phi\left(\frac{-\pi^*_{t+1}-x'}{\sigma}\right)}{\sigma}\frac{\pi^*_{t+1}+x'}{\sigma}$$

After change of variable to x, this expression becomes

$$0 = \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}(x')V_{t+1}(x')h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx' + \frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(V_{t+1}(0)-\eta w_{t+1}).$$

Now we re-express this in terms of $v(x)$ using $V_t(x) = v_t(x) + V_t(0)$ first, and then rearranging

$$\begin{aligned}
0 &= \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}(x')\left(v_{t+1}(x')+V_{t+1}(0)\right)h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx' \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(V_{t+1}(0)-\eta w_{t+1}) \\
&= \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}(x')v_{t+1}(x')h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx' \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}(x')h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx'V_{t+1}(0) \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(V_{t+1}(0)-\eta w_{t+1}) \\
&= \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}(x')v_{t+1}(x')h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx' \\
&\quad - \frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)V_{t+1}(0)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(V_{t+1}(0)-\eta w_{t+1})\\
=\;&\Pi'_t(0)+\frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}(x')v_{t+1}(x')h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx'\\
&+\frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(-\eta w_{t+1})
\end{aligned}
$$

and applying the piecewise linear approximation of $v(x)$:

$$
\begin{aligned}
0=\Pi'_t(0)+\frac{\Lambda_{t,t+1}}{\sigma}\sum_{i=1}^{I-1}\int_{-1}^{1}\hat{s}_{t+1}(x')v_{t+1}(x_i<x'<x_{i+1})h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx'\\
+\frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(-\eta w_{t+1})\,.
\end{aligned}
$$

H.3. Solving for integrals

Let us collect the approximated equations defined so far.

$$
\begin{aligned}
v_t(x)=&\Pi_t(x)-\Pi_t(0)\\
&+\frac{\Lambda_{t,t+1}}{\sigma}\sum_{i=1}^{I-1}\int_{x_i}^{x_{i+1}}\hat{s}_{t+1}(x')v_{t+1}(x_i<x'<x_{i+1})(\phi\left(h_{t+1}(x,x')\right)-\phi\left(h_{t+1}(0,x')\right))dx'\\
&+\frac{\Lambda_{t,t+1}}{\sigma}(\eta w_{t+1})\int_{-1}^{1}\hat{s}_{t+1}(x')(\phi\left(h_{t+1}(x,x')\right)-\phi\left(h_{t+1}(0,x')\right))dx',
\end{aligned}
\tag{A45}
$$

$$
\begin{aligned}
0=\Pi'_t(0)+\frac{\Lambda_{t,t+1}}{\sigma}\int_{-1}^{1}\hat{s}_{t+1}v_{t+1}(x')h_{t+1}(0,x')\frac{\phi\left(h_{t+1}(0,x')\right)}{\sigma}dx'\\
+\frac{\Lambda_{t,t+1}}{\sigma}\left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma}\right)-\phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma}\right)\right)(-\eta w_{t+1})\,,
\end{aligned}
\tag{A46}
$$

$$
g^c_t(x)=\sum_{i=1}^{I-1}\int_{x_i}^{x_{i+1}}\frac{\hat{s}_{t-1}(x')}{\sigma}g^c_{t-1}(x_i<x'<x_{i+1})\phi\left(h_t(x',x)\right)\,dx'+\frac{1}{\sigma}g^0_{t-1}\phi\left(h_t(0,x)\right)\,,
\tag{A47}
$$

$$
g^0_t=1-\sum_{i=1}^{I-1}\int_{x_i}^{x_{i+1}}g^c_t(x_i<x'<x_{i+1})\hat{s}_t(x)dx,
\tag{A48}
$$

$$
e^{p^*_t(\epsilon-1)}=\sum_{i=1}^{I-1}\int_{x_i}^{x_{i+1}}e^{(x)(1-\epsilon)}g^c_t(x_i<x'<x_{i+1})\hat{s}_t(x)dx+g^0_t,
\tag{A49}
$$

$$
N_t=\frac{C_t}{A_t}e^{p^*_t(-\epsilon)}\left(\sum_{i=1}^{I-1}\int_{x_i}^{x_{i+1}}e^{x(-\epsilon)}g^c_t(x_{i-1}<x<x_{i+1})\hat{s}_t(x)dx+g^0_{t-1}\right)+\eta g^0_{t-1}.
\tag{A50}
$$

The integrals in all of these expressions can be computed analytically, since the integrands consist of affine functions multiplied by expressions that have closed-form anti-derivatives. Figure A7 illustrates this graphically for the integral in the equation for $g^c_t(x)$ (A47).

We now determine the solution of these integrals, equation by equation. Given the coefficients of the affine functions, which depend on the values of $v_{t+1}(g_{t-1})$ at the grid points x_i, we can then write the solutions as a function that is linear in the elements of the vector $v_{t+1}(x_i)$ $(g_{t-1}(x_i))$. We now explain this for the simple case of the integral in equation A48. The other equations require more tedious algebra, which we carried out using symbolic math and which we omit here for brevity, but are conceptually equivalent.

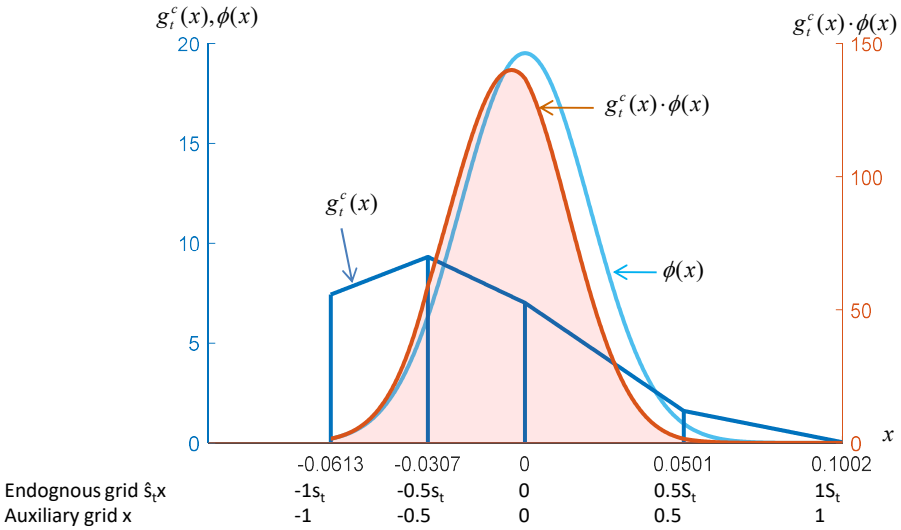

FIGURE A7. This figure schematically explains the analytical evaluation of integrals, given the linear approximation of the distribution and value functions. It shows the piecewise linearly approximated distribution $g_t^c(x)$ in blue, the normal pdf $\phi(x)$ in light blue and the product of the two $g_t^c(x)\phi(x)$ in orange, where $x = x\hat{s}_t$. The orange area thus corresponds to the term $\sum_{i=1}^{I-1} \int_{x_i}^{x_{i+1}} \frac{\hat{s}_{t-1}(x')}{\sigma} g_{t-1}^c(x_i < x' < x_{i+1}) \phi\left(h_t(x', x)\right) dx'$ in equation (A47).

H.3.1. Mass point

The integral over an affine function $f(x)$ from x_1 to x_2 is given by

$$\int_{x_1}^{x_2} f(x)\, dx = \frac{(f(x_1) + f(x_2))}{2}(x_2 - x_1)$$

thus

$$\sum_{i=1}^{I-1} \int_{x_i}^{x_{i+1}} f(x)\, dx = \sum_{i=1}^{I-1} \frac{(f(x_i) + f(x_{i+1}))}{2}(x_{i+1} - x_i).$$

Collecting the common terms on the right-hand side we get

$$\sum_{i=1}^{I-1} \int_{x_i}^{x_{i+1}} f(x)\, dx = \frac{1}{2}\left[f(x_1)(x_2 - x_1) + \sum_{i=2}^{I-1} f(x_i)(x_{i+1} - x_{i-1}) + f(x_I)(x_I - x_{I-1}) \right].$$

Applying this formula to equation (A48), which defines the mass point at $x = 0$, and re-arranging terms we get

$$\mathrm{g}_t^0 = 1 - \mathbf{e}_t^T \mathbf{g}_t^{\boldsymbol{c}} \tag{A51}$$

where $\mathbf{e}_t^T = \left[x_{t,\min(I,\,i+1)} - x_{t,\max(1,\,i-1)}\right]_{i=1}^{I} = \left[\hat{\mathrm{s}}_t \mathrm{x}_{\min(I,\,i+1)} - \hat{\mathrm{s}}_t \mathrm{x}_{\max(1,\,i-1)}\right]_{i=1}^{I}$. Note that this formula corresponds to the trapezoid rule. The blue area in Figure A6 illustrates the application of the trapezoid rule.

H.3.2. Aggregate price index

By the same logic, the aggregate price index in (A49) is computed as

$$e^{p_t^*(\epsilon-1)} = \sum_{i=1}^{I} (\mathrm{g}_t^c(\mathrm{x}_i)\mathbb{1}_{i\neq 1} d_{t,i,i-1,1-\epsilon} + \mathrm{g}_t^c(\mathrm{x}_i)\mathbb{1}_{i\neq I} d_{t,i,i+1,1-\epsilon}) + \mathrm{g}_t^0 \tag{A52}$$

where

$$d_{t,i,j,\epsilon} = \frac{\left(e^{(\epsilon)\mathrm{x}_i \hat{\mathrm{s}}_{t,i}}\left((\epsilon)\left(\mathrm{x}_i \hat{\mathrm{s}}_{t,i} - \mathrm{x}_j \hat{\mathrm{s}}_{t,j}\right) - 1\right) + e^{(\epsilon)\mathrm{x}_j \hat{\mathrm{s}}_{t,j}}\right)}{(\epsilon)^2 \left|\mathrm{x}_i \hat{\mathrm{s}}_{t,i} - \mathrm{x}_j \hat{\mathrm{s}}_{t,j}\right|}$$

and where $\hat{\mathrm{s}}_{t,i} \equiv \hat{\mathrm{s}}_t(x_i)$ and where $\mathbb{1}_{i\neq 1}$ and $\mathbb{1}_{i\neq I}$ are indicator functions equal to 1 when i is different from 1 or I, that is whenever $\mathrm{g}_t^c(\mathrm{x}_i)$ is evaluated at the bounds of the (S, s) band. It plays a similar role as the values 0.5 at the two extremes of the vector $\mathbf{e}_t^T$ above.

Hence, we can re-write equation (A52) in matrix form as

$$e^{p_t^*(\epsilon-1)} = \mathbf{d}_{t,1-\epsilon}^T \mathbf{g}_t^{\boldsymbol{c}} + \mathrm{g}_t^0 \tag{A53}$$

where $\mathbf{g}_t^{\boldsymbol{c}}$ is the vector collecting the values of the distribution function g_t^c at the grid points and where the vector $\mathbf{d}_{t,1-\epsilon}$ is

$$\mathbf{d}_{t,1-\epsilon} = \left[\mathbb{1}_{i\neq 1} d_{t,i,i-1,1-\epsilon} + \mathbb{1}_{i\neq I} d_{t,i,i+1,1-\epsilon}\right]_{i=1}^{I}.$$

Here we have adopted the notation that $[x_i]_{i=1}^{I}$ denotes an $I \times 1$ vector with elements x_i.

H.3.3. Labor market

Following the previous subsection, the labor market condition (A50) is computed as

$$N_t = \frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \left(\sum_{i=1}^{I} (\mathrm{g}_t^c(\mathrm{x}_i)\mathbb{1}_{i\neq 1} d_{t,i,i-1,-\epsilon} + \mathrm{g}_t^c(\mathrm{x}_i)\mathbb{1}_{i\neq I} d_{t,i,i+1,-\epsilon}) + \mathrm{g}_{t-1}^0\right) + \eta \mathrm{g}_{t-1}^0$$

which we re-write in matrix form as

$$N_t = \frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \left(\mathbf{d}_{t,-\epsilon}^T \mathbf{g}_t^{\boldsymbol{c}} + \mathrm{g}_{t-1}^0\right) + \eta \mathrm{g}_{t-1}^0. \tag{A54}$$

H.3.4. Distribution

Once we have evaluated the integrals, the distribution function in (A47) can be written as:

$$g_t^c(x_j) = \sum_{i=1}^{I} \frac{1}{2\sqrt{2\pi}} g_{t-1}^c(x_i) \left[\mathbb{1}_{i\neq 1} f_{t,i,i-1,j} + \mathbb{1}_{i\neq I} f_{t,i,i+1,j} \right] + \frac{1}{\sigma} g_{t-1}^0 \phi\left(\frac{-\hat{s}_{t,j} x_j - \pi_t^*}{\sigma} \right) \quad \text{(A55)}$$

where from now on, π without time subindex, denotes the scalar π, $f_{t,i,\bar{i},j}$ and $\mathcal{P}_{t,i,j}$ are defined as

$$f_{t,i,\bar{i},j} = \frac{\sqrt{2\pi}\left(\mathcal{P}_{t,\bar{i},j}\right)\left(\operatorname{erf}\left(\frac{\mathcal{P}_{t,\bar{i},j}}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{\mathcal{P}_{t,i,j}}{\sqrt{2}\sigma}\right)\right) + 2\sigma\left(\exp\left(-\frac{\mathcal{P}_{t,\bar{i},j}^{\,2}}{2\sigma^2}\right) - \exp\left(-\frac{\mathcal{P}_{t,i,j}^{\,2}}{2\sigma^2}\right)\right)}{\left|x_i \hat{s}_{t-1,i} - x_{\bar{i}} \hat{s}_{t-1,\bar{i}}\right|},$$

$$\mathcal{P}_{t,i,j} = -x_i \hat{s}_{t-1,i} + x_j \hat{s}_{t,j} + \pi_t^*.$$

For compactness, define

$$\mathbf{g_t^c} \equiv \left[g_t^c(x_j) \right]_{j=1}^{I}$$

$$\mathbf{F}_t \equiv \left[\frac{1}{2\sqrt{2\pi}} \left(\mathbb{1}_{i\neq 1} f_{t,i,i-1,j} + \mathbb{1}_{i\neq I} f_{t,i,i+1,j} \right) \right]_{j=1,i=1}^{I,I}$$

$$\mathbf{f}_t \equiv \left[\frac{1}{\sigma} \phi\left(\frac{-\hat{s}_{t,j} x_j - \pi_t^*}{\sigma} \right) \right]_{j=1}^{I}$$

where $\mathbf{g_t^c}$ and $\mathbf{f}_t$ are vectors with the probability mass function g_t^c and the scaled and shifted normal distribution at the grid points, respectively, $\mathbf{F}_t$ is a matrix that captures the idiosyncratic transitions due to firm-level quality shocks and where we have adopted the notation that $\left[x_{i,j}\right]_{j=1,i=1}^{J,I}$ denotes a $J \times I$ matrix with elements $x_{j,i}$. Thus, equation A55 can be represented in matrix form as

$$\mathbf{g_t^c} = \mathbf{F}_t \mathbf{g_{t-1}^c} + \mathbf{f}_t g_{t-1}^0. \quad \text{(A56)}$$

H.3.5. Value function

Once we have evaluated the integrals, and denoting the standard normal cdf by $\Phi(\cdot)$ and the central grid point by i_0 (i.e. for $x_{i_0} = 0$), the value function A45 can be written

as

$$\begin{aligned}\mathsf{v}_t(\mathsf{x}_j) = \; & \Pi_{j,t} - \Pi_{j,t}(0) \\ & + \Lambda_{t,t+1} \sum_{i=1}^{I} \frac{1}{2\sqrt{2\pi}} \mathsf{v}_{t+1}(\mathsf{x}_i) \left(\mathbb{1}_{i\neq 1}(a_{t,i,i-1,j} - a_{t,i_0,i_0-1,j}) + \mathbb{1}_{i\neq I}(a_{t,i,i+1,j} - a_{t,i_0,i_0+1,j}) \right) \\ & + \Lambda_{t,t+1}(\eta w_{t+1}) \left(\Phi\left(\frac{\mathcal{P}_{t+1,j,I}}{\sigma}\right) - \Phi\left(\frac{\mathcal{P}_{t+1,j,1}}{\sigma}\right) - \Phi\left(\frac{\mathcal{P}_{t+1,i_0,I}}{\sigma}\right) + \Phi\left(\frac{\mathcal{P}_{t+1,i_0,1}}{\sigma}\right) \right)\end{aligned} \tag{A57}$$

where

$$a_{t,i,\bar{i},j} = \frac{\sqrt{2\pi}\left(\mathcal{P}_{t+1,j,\bar{i}}\right)\left(\operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,\bar{i}}}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,i}}{\sqrt{2}\sigma}\right)\right) + 2\sigma\left(\exp\left(-\frac{\left(\mathcal{P}_{t+1,j,\bar{i}}\right)^2}{2\sigma^2}\right) - \exp\left(-\frac{\left(\mathcal{P}_{t+1,j,i}\right)^2}{2\sigma^2}\right)\right)}{\left|\mathsf{x}_i \hat{s}_{t+1,i} - \mathsf{x}_{\bar{i}} \hat{s}_{t+1,\bar{i}}\right|}. \tag{A58}$$

For compactness, let us define

$$\begin{aligned}\mathbf{v}_t &\equiv \left[\mathsf{v}_t(\mathsf{x}_j)\right]_{j=1}^{I}, \\ \Pi_t &\equiv \left[\Pi_{j,t} - \Pi_{j,t}(0)\right]_{j=1}^{I}, \\ \mathbf{A}_t &\equiv \left[\Lambda_{t,t+1}\frac{1}{2\sqrt{2\pi}}(\mathbb{1}_{i\neq 1}(a_{t,i,i-1,j} - a_{t,i_0,i_0-1,j}) + \mathbb{1}_{i\neq I}(a_{t,i,i+1,j} - a_{t,i_0,i_0+1,j}))\right]_{j=1,i=1}^{I,I}, \\ \mathbf{b}_{t+1} &\equiv -\left[\Lambda_{t,t+1}\left(\Phi\left(\frac{\mathcal{P}_{t+1,j,I}}{\sigma}\right) - \Phi\left(\frac{\mathcal{P}_{t+1,j,1}}{\sigma}\right) - \Phi\left(\frac{\mathcal{P}_{t+1,i_0,I}}{\sigma}\right) + \Phi\left(\frac{\mathcal{P}_{t+1,i_0,1}}{\sigma}\right)\right)\right]_{j=1}^{I}\end{aligned}$$

where $\mathbf{v}_t$ and $\mathbf{b}_{t+1}$ are vectors that evaluate the value function v_t and the adjustment probability at different grid points, Π_t is the vector of profit differences, while $\mathbf{A}_t$ is a matrix that represents the idiosyncratic transition due to firm-level quality shocks and price updating. Thus, equation (A57) can be represented in matrix form as

$$\mathbf{v}_t = \Pi_t + [\mathbf{A}_t \mathbf{v}_{t+1} - \mathbf{b}_{t+1}\eta w_{t+1}]. \tag{A59}$$

H.3.6. Optimality condition for the reset price

After evaluating the integral, we can write the optimality condition in (A46) as

$$\begin{aligned}0 = \; & \Pi_t'(0) + \Lambda_{t,t+1} \sum_{i=1}^{I} \mathsf{v}_{t+1}(\mathsf{x}_i) \frac{1}{2} \left(\mathbb{1}_{i\neq 1} c_{t,i,i-1,i_0} + \mathbb{1}_{i\neq I} c_{t,i,i+1,i_0} \right) \\ & + \frac{\Lambda_{t,t+1}}{\sigma} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma}\right) \right)(-\eta w_{t+1})\end{aligned} \tag{A60}$$

where

$$c_{t,i,\bar{i},j} = \frac{\operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,i}}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,\bar{i}}}{\sqrt{2}\sigma}\right)}{x_i \hat{s}_{t+1,i} - x_{\bar{i}} \hat{s}_{t+1,\bar{i}}} - \frac{\sqrt{\frac{2}{\pi}} \exp\left(-\frac{\left(\mathcal{P}_{t+1,j,i}\right)^2}{2\sigma^2}\right)}{\sigma}. \tag{A61}$$

We can write this equation using matrix notation:

$$\begin{aligned} 0 = {} & \Pi_t'(0) + \mathbf{c}_{t+1}^T \mathbf{v}_{t+1} \\ & + \frac{\Lambda_{t,t+1}}{\sigma} \left(\phi\left(\frac{-S_{t+1} - \pi_{t+1}^*}{\sigma}\right) - \phi\left(\frac{-s_{t+1} - \pi_{t+1}^*}{\sigma}\right) \right) (-\eta w_{t+1}) \end{aligned} \tag{A62}$$

where

$$\mathbf{c}_{t+1} = \left[\Lambda_{t,t+1} \frac{1}{2} \left(\mathbb{1}_{i \neq 1} c_{t,i,i-1,i_0} + \mathbb{1}_{i \neq I} c_{t,i,i+1,i_0} \right) \right]_{i=1}^{I}. \tag{A63}$$

H.4. Final equation system

Collecting the derived equations, and combining them with the remainder of the private equilibrium conditions (which contain no infinite dimensional objects) and the objective, we can approximate the infinite dimensional central bank problem by the following finite-dimensional problem:

$$\max_{\left\{\mathbf{g}_t^{\boldsymbol{c}}, \mathrm{g}_t^0, \mathbf{v}_t, C_t, w_t, p_t^*, s_t, S_t, \pi_t^*\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left(\log C_t - \left(\frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \left(\mathbf{d}_{t,-\epsilon}^T \mathbf{g}_t^{\boldsymbol{c}} + \mathrm{g}_t^0 \right) + \eta \mathrm{g}_t^0 \right) \right)$$

subject to

$$\begin{aligned} w_t &= C_t, \\ \mathbf{v}_t &= \Pi_t + \mathbf{A}_t \mathbf{v}_{t+1} - \mathbf{b}_{t+1} \eta w_{t+1}, \\ \mathbf{v}_{t,1} &= -\eta w_t, \\ \mathbf{v}_{t,I} &= -\eta w_t, \\ 0 &= \Pi_t'(0) + \mathbf{c}_{t+1}^T \mathbf{v}_{t+1} + \frac{\Lambda_{t,t+1}}{\sigma} \left(\phi\left(\frac{-S_{t+1} - \pi_{t+1}^*}{\sigma}\right) - \phi\left(\frac{-s_{t+1} - \pi_{t+1}^*}{\sigma}\right) \right) (-\eta w_{t+1}), \\ \mathbf{g}_t^{\boldsymbol{c}} &= \mathbf{F}_t \mathbf{g}_{t-1}^{\boldsymbol{c}} + \mathbf{f}_t \mathrm{g}_{t-1}^0, \\ \mathrm{g}_t^0 &= 1 - \mathbf{e}_t^T \mathbf{g}_t^{\boldsymbol{c}}, \\ e^{p_t^*(\epsilon-1)} &= \mathbf{d}_{t,1-\epsilon}^T \mathbf{g}_t^{\boldsymbol{c}} + \mathrm{g}_t^0. \end{aligned}$$

Here, the choice variables $\mathbf{v}_t$ and $\mathbf{g}_t^{\boldsymbol{c}}$ are vectors of length I. The rest of the choice variables are scalars. Note that the choice variables p_t^*, s_t, S_t, π_t^* implicitly appear in the problem (inside the vectors and matrices $\mathbf{A}_t$, $\mathbf{b}_t$, etc.)

As already explained at the beginning of this Appendix, we solve for the FOCs of this system by symbolic differentiation. The resulting system of FOCs is then solved in the sequence space. We next explain how we find the steady state, which serves as initial and terminal condition for dynamic simulations.

H.5. Steady state

To solve for the steady state of the private equilibrium conditions, given a policy $\bar{\pi}$, the algorithm is as follows. We rely on steady-state relationships $w = C$, $R = \exp(\pi)/\beta$ and $\pi = \pi^*$. We start with a guess for the real wage w, the optimal reset price p^*, and the bounds of the (S, s) band s and S, and then:

a. Compute consumption $C = w$.

b. Using $\pi = \pi^* = \bar{\pi}$, C and the 4 initial guesses, solve for the stationary value function using the Bellman equation and the stationary distribution using the law of motion of the distribution. Both have closed-form solutions given the guesses.

$$\begin{aligned}
\mathbf{v} &= (\mathbf{I} - \mathbf{A})^{-1} (\Pi - \mathbf{b}\eta w), \\
\mathbf{g}^c &= \left(\mathbf{I} - \mathbf{F} + \mathbf{f}\mathbf{e}^T\right)^{-1} \mathbf{f}, \\
g^0 &= 1 - \mathbf{e}^T \mathbf{g}^c
\end{aligned}$$

c. Compute the residuals of the 4 remaining equations

$$\begin{aligned}
\mathbf{v}_{t,1} &= -\eta w_t, \\
\mathbf{v}_{t,I} &= -\eta w_t, \\
0 &= \Pi'_t(0) + \mathbf{c}_{t+1}^T \mathbf{v}_{t+1} + \frac{\Lambda_{t,t+1}}{\sigma} \left(\phi \left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma} \right) - \phi \left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma} \right) \right) (-\eta w_{t+1}), \\
e^{p_t^*(\epsilon-1)} &= \mathbf{d}_{t,1-\epsilon}^T \mathbf{g}_t^{\mathbf{c}} + g_t^0.
\end{aligned}$$

d. Use a Newton method to update the 4 guesses (w, p^*, s, S) and return to step 1, until convergence of the residuals.