Strike while the Iron is Hot: Optimal Monetary Policy with a Nonlinear Phillips Curve*

Peter Karadi[1,2], Anton Nakov[1,2], Galo Nuño[3,2], Ernesto Pastén[4], and Dominik Thaler[1]

[1]European Central Bank
[2]CEPR
[3]Bank for International Settlements and Bank of Spain
[4]Central Bank of Chile

August 11, 2024

Abstract

We study the Ramsey optimal monetary policy within the Golosov and Lucas (2007) state-dependent pricing framework. The model provides microfoundations for a nonlinear Phillips curve: the sensitivity of inflation to activity increases after large shocks due to an endogenous rise in the frequency of price changes, as observed during the recent inflation surge. In response to large cost-push shocks, optimal policy leverages the lower sacrifice ratio to reduce inflation and stabilize the frequency of price adjustments. At the same time, when facing total factor productivity shocks, an efficient disturbance, the optimal policy commits to strict price stability, similar to the prescription in the standard Calvo (1983) model.

JEL codes: E31, E32, E52

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

*We are grateful to Vladimir Asryan, Andres Blanco, Davide Debortoli, Eduardo Engel, Aurélien Eyquem, Jordi Galí, Mark Gertler, Mishel Ghassibe, Francesco Lippi, Albert Marcet, Virgiliu Midrigan, Giorgio Primiceri, Xavier Ragot, Morten Ravn, Tom Sargent, Edouard Schaal, and Jaume Ventura, as well as to participants at various conferences and seminars for their comments and suggestions. This manuscript was previously circulated as ´Strike the Iron while it's Hot: Optimal Monetary Policy with (S,s) Pricing´.The views expressed here are those of the authors only, and do not necessarily represent those of the BIS, Bank of Spain, the Central Bank of Chile, ECB, or the Eurosystem.

1 Introduction

The recent inflation surge has been accompanied by a significant increase in the frequency of price changes (Montag and Villar, 2023; Cavallo et al., 2023; Blanco et al., 2024a). Concurrently, empirical evidence reveals marked nonlinearities in the estimated Phillips curve, which characterizes the relationship between economic activity and inflation (Benigno and Eggertsson, 2023; Cerrato and Gitti, 2023). Traditional models of price setting (Calvo, 1983), which form the basis for optimal monetary policy analysis (Woodford, 2003; Galí, 2008), cannot explain the above observations. In contrast, state-dependent pricing models are well suited to capturing both of them: firms' price-adjustment decisions lead to endogenous variation in the repricing frequency and, thus, in the slope of the Phillips curve. Among the state-dependent models, the Golosov and Lucas (2007) menu cost model has emerged as a benchmark for positive analysis. However, normative aspects of the model have received scant attention, including how the central bank should respond to large inflation surges. It is precisely this crucial gap that our paper aims to bridge.

Our analysis arrives at a novel insight: in response to large cost-push shocks, the Ramsey optimal monetary policy in the Golosov and Lucas (2007) model commits to quashing inflation and leaning against changes in the repricing frequency – a "strike while the iron is hot" policy. The reason is that the cost of the anti-inflationary policy in terms of output is smaller when the frequency of price changes increases in response to the shocks – as the Phillips curve becomes steeper, the sacrifice ratio falls. Furthermore, as we show algebraically, optimal policy requires full inflation stabilization after total factor productivity shocks – a divine coincidence result after efficient shocks as in the Calvo model.

In our state-dependent price-setting model, a representative household consumes a continuum of differentiated goods and provides labor in a centralized, frictionless market. Each good is produced by a single firm with only labor subject to aggregate productivity shocks, aggregate cost-push shocks, and idiosyncratic quality shocks.[1] Firms must incur a small, fixed, "menu cost" to adjust their prices. Thus, firms' pricing decisions are characterized by

[1] We depart from the Golosov and Lucas (2007) model in this regard, which instead assumes idiosyncratic productivity shocks. We do so for ease of computation while its implications on our results are innocuous (see also Midrigan, 2011; Alvarez et al., 2021).

an (S, s) rule: when prices are within an endogenous band around the optimal reset price, firms keep them constant; otherwise, they pay the menu cost and update their price. The central bank sets the nominal interest rate. The model is calibrated to match the frequency and magnitude of price changes in the U.S. (Nakamura and Steinsson, 2008). We contrast the implications of our model to those of a Calvo model recalibrated to generate the same Phillips curve slope for small shocks (Auclert et al., 2024). This recalibration compensates for a special feature of the baseline model: the endogenous "selection" of large price changes, which raises the flexibility of the price level.

We start by exploring the implications of our menu cost economy assuming the central bank follows a Taylor rule. We show that the similarity of our baseline menu cost model and the (suitably recalibrated) Calvo model for small shocks fails to generalize in the presence of large shocks (Blanco et al., 2024a; Cavallo et al., 2023). The reason is the repricing frequency: it increases endogenously in our baseline model as shocks become large, while it stays constant in the Calvo model. The repricing frequency raises price flexibility and generates a nonlinear relationship between inflation and the output gap: a nonlinear Phillips curve. When shocks are small, the repricing frequency remains close to its steady-state value and the slope of the Phillips curve stays equal to that of the Calvo model. As shocks become larger, the frequency increases, prices become more flexible and the slope of the Phillips curve steepens.

Next, we move to the core of the paper: the optimal design of monetary policy in a menu cost model with a nonlinear Phillips curve. To this end, we solve the fully nonlinear Ramsey problem under commitment. We first analyze the Ramsey steady state. The model features a slightly positive steady-state inflation rate, at around 0.3%. This contrasts with the Calvo model, where optimal inflation is exactly zero. In our menu cost model, positive inflation reduces the frequency and thus helps firms to economize on costly price adjustments. In particular, it counterbalances the impact of too frequent price increases relative to price decreases, a consequence of the asymmetry of the profit function: firms dislike more negative price misalignments when the demand for their product is high, relative to positive misalignments when the demand is low.

We turn next to the optimal systematic response to shocks under a "timeless perspective"

(Woodford, 2003). We first analyse cost-push shocks. The optimal response to such shocks in the Calvo model is to "lean against the wind": the central bank temporarily drives output below its efficient level to contain the inflationary impact of a positive cost-push shock. The relationship between the inflation level and the change in the output gap is characterized by a near-linear target rule. In the menu cost model, the central bank also "leans against the wind" and, for small shocks, follows a linear target rule with a similar slope. The reason, however, is different. In the Calvo model, the central bank trades off the distortion associated with *average markup* volatility, linked to the variance in the output gap, to that due to higher *price dispersion*, proportional to the variance of inflation. In the menu cost model, price dispersion actually can decrease with inflation as new adjusters are endogenously selected from those with the most misaligned prices. There is however a new distortion, namely the *losses due to menu costs*, which increase with inflation and tend to dominate the welfare effect of inflation on price dispersion.

While the balance between these distortions in the Calvo and our menu cost model is similar for small shocks, the differences become relevant as the shock size increases. For large shocks, we find that monetary policy should be tighter in the menu cost model: the optimal prescription is to react more aggressively against inflation than in the case of small shocks or in a counterfactual fixed-frequency Calvo setting. The outcome is a nonlinear target rule that significantly dampens the inflation surge for a unit decline in the output gap as the shocks become large. Finally, we show that the source of the nonlinearity in the target rule is almost exclusively due to the nonlinearity in the Phillips curve. We show that the planner's preferences in our baseline model can be approximated quite well by the planner's preferences in the Calvo model, which are near quadratic over the output gap and inflation. The Phillips curve, however, is different in the two frameworks: it is nonlinear in our baseline, while near-linear in Calvo. When we counterfactually insert the quadratic preferences of the Calvo model into our baseline, the target rule still remains nonlinear, confirming that it is the nonlinearity of the Phillips curve and, thus, the nonlinearity of the sacrifice ratio that drives the nonlinearity of the target rule. As inflation diverges more and more from its steady state value, the output cost of containing inflation diminishes, so the central bank contains it more strongly.

Next we show analytically that the optimal response to TFP shocks is characterized by the "divine coincidence" (Blanchard and Galí, 2007). In other words, optimal policy stabilizes both inflation and the output gap relative to the efficient level.

The well-known time inconsistency problem of monetary policy is also present in our menu cost model, although it is muted relative to Calvo. In both models, when the steady state is inefficient, monetary policy has the incentive to stimulate output via an unexpectedly easy policy (Galí, 2008). However, in the menu cost model, such a policy is less effective on output and more inflationary because the ensuing increase in the repricing rate raises the flexibility of the aggregate price level. The time-inconsistent motive to ease is thus considerably weaker.

We assess the optimal monetary policy response to the 2022-2023 inflation surge through the lens of our model. We construct a scenario that captures key features of the inflation surge in the US. We argue that a combination of both aggregate and relative-price shocks is needed to explain the evidence documented in the micro price data by Montag and Villar (2023), who found large increases both in the frequency and the dispersion of price changes. Relative-price shocks are especially relevant, as they generate a reason for an "efficient" increase in frequency. The scenario generates realistic inflation dynamics when the monetary policy follows an inertial Taylor rule: inflation surges to around 9 percent temporarily and stays persistently above the central bank's 2 percent inflation target for a considerable amount of time. In the counterfactual optimal policy scenario, we find that monetary policy tightens aggressively and, at some output cost, keeps the inflation surge temporary with a peak of 4 percent above trend inflation.

Overall, our findings highlight the significance of an aggressive anti-inflationary stance by the central bank in the face of large shocks. By committing to policies that lean against inflation and stabilize the repricing frequency, the central bank can foster a more favorable macroeconomic outcome.

Related literature. Our paper builds on the seminal article by Golosov and Lucas (2007). They propose a menu cost model (Barro, 1972; Sheshinski and Weiss, 1977; Caballero and Engel, 1993) that has become the backbone of a positive literature studying the relationship between monetary non-neutrality and the distribution of price changes at the micro level

(Midrigan, 2011; Costain and Nakov, 2011; Alvarez et al., 2016), as well as the impact of large aggregate shocks on inflation and activity (Karadi and Reiff, 2019; Alexandrov, 2020; Auer et al., 2021). The model describes well firms' price-setting behavior in diverse environments with both low and high inflation (Nakamura and Steinsson, 2008; Gagnon, 2009; Alvarez et al., 2019). This price-setting framework provides a microfounded state-dependent alternative to the canonical time-dependent Calvo (1983) model, with widely different implications in terms of both the extent of monetary non-neutrality and price-flexibility as a response to large shocks. Indeed, most familiar price-setting models, such as the random-menu cost model of Dotsey et al. (1999); Alvarez et al. (2021), the Calvo-plus model of Nakamura and Steinsson (2010), the rational inattention model by Woodford (2009), or the control cost model by Costain and Nakov (2019), lie on a spectrum bracketed by these two polar cases. Normative from the Golosov and Lucas (2007) model can provide qualitative insights that generalize to a wide class of price-setting frameworks.

To the best of our knowledge, our paper is the first to solve for optimal monetary policy in this canonical menu cost model. Its main distinctive feature, relative to the textbook analysis based on Calvo (1983), such as in Woodford (2003) and Galí (2008), is a state-dependent relationship between inflation and the output gap – a "nonlinear Phillips curve" – which has received new empirical support following the recent inflation surge (Benigno and Eggertsson, 2023; Cerrato and Gitti, 2023; Blanco et al., 2024a). Our conclusion prescribing aggressive anti-inflationary policy after large shocks is a direct consequence of this nonlinearity: a higher Phillips curve slope implies a favourable inflation-output trade-off that optimal policy should exploit. Crucially, the framework includes firms facing idiosyncratic shocks, which are essential to explain the large size of observed price changes and raises relevant normative questions by providing an underlying cause for efficient relative price adjustments.

Solving dynamic optimal policy in response to aggregate shocks in this framework complements previous research on optimal monetary policy, which has restricted attention to menu cost settings with a representative firm and small aggregate shocks (Nakov and Thomas, 2014), sector-specific productivity shocks (Caratelli and Halperin, 2023)[2] or to optimal

[2]The former finds no significant difference between Calvo and an stylized menu cost model. The latter shows that in the face of sector-specific shocks optimal policy can be characterized as *nominal wage targeting*.

steady-state inflation rate (Adam and Weber, 2019; Blanco, 2021; Nakov and Thomas, 2014).

This paper proposes a new algorithm to solve Ramsey optimal policy in heterogeneous-agent models, building on González et al. (2024). The algorithm (i) makes the infinite-dimensional planner's problem finite-dimensional by approximating the infinite-dimensional value and distribution functions by piece-wise linear functions; (ii) accounts for the discrete price-adjustment choice using an endogenous grid; (iii) derives the FOCs of the planner's problem by symbolic differentiation; and (iv) solves the resulting set of equilibrium conditions nonlinearly under perfect foresight over the sequence space. Our approach complements other methods to solve for Ramsey policy in heterogeneous-agent models (Bhandari et al., 2021; Le Grand et al., 2022; Dávila and Schaab, 2022; Nuño and Thomas, 2022; Smirnov, 2022).

2 Model

The economy consists of a representative household, monopolistic producers facing fixed menu costs to update their prices, and a central bank that sets the nominal interest rate.

2.1 Households

A representative household consumes C_t, supplies working hours N_t and saves in one-period, zero-net supply nominal bonds B_t. The household maximizes

$$\max_{C_t, N_t, B_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(C_t, N_t), \tag{1}$$

subject to

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

where T_t are lump-sum taxes, W_t is the nominal wage, D_t are lump-sum dividends from firms, and Q_t is the price of the nominal bond. Aggregate consumption C_t is

$$C_t = \left\{ \int \left[A_t(j) C_t(j) \right]^{\frac{\epsilon-1}{\epsilon}} dj \right\}^{\frac{\epsilon}{\epsilon-1}}, \tag{3}$$

where $C_t(j)$ is the quantity purchased of product $j \in [0,1]$ and $A_t(j)$ is the quality of product j, following a random walk with stochastic volatility in logs:

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

and ε_t is an i.i.d Gaussian innovation. The demand for product j is,

$$C_t(j) = A_t(j)^{\epsilon-1} \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} C_t, \tag{4}$$

and the aggregate price index is

$$P_t = \left[\int_0^1 \left(\frac{P_t(j)}{A_t(j)} \right)^{1-\epsilon} dj \right]^{\frac{1}{1-\epsilon}}. \tag{5}$$

We assume separable utility of the CRRA class, $u(C_t, N_t) = \frac{C_t^{1-\gamma}}{1-\gamma} - \upsilon N_t$. Solving for the FOCs, we obtain the labor supply condition,

$$w_t = \upsilon C_t^{\gamma}, \tag{6}$$

where $w_t = W_t/P_t$ is the real wage. The Euler equation is

$$1 = \mathbb{E}_t \left[\Lambda_{t,t+1} e^{i_t - \pi_{t+1}} \right], \tag{7}$$

where $i_t \equiv \log(-Q_t)$ is the nominal interest rate, and

$$\Lambda_{t,t+1} \equiv \beta \frac{u'(C_{t+1})}{u'(C_t)}. \tag{8}$$

2.2 Monopolistic producers

Production of good j is

$$Y_t(j) = A_t \frac{N_t(j)}{A_t(j)}, \tag{9}$$

where $N_t(j)$ is the labor input, and A_t is aggregate productivity. [3]

[3] These shocks allow for idiosyncratic variation in reset prices by introducing a single state at the firm level and thus economizing on the dimensionality of the optimal monetary policy problem.

The nominal profit function is

$$\begin{aligned} D_t(j) =& P_t(j)Y_t(j) - (1-\tau_t)W_t N_t(j) \\ =& P_t(j)^{1-\epsilon} A_t(j)^{\epsilon-1} \left(\frac{1}{P_t}\right)^{-\epsilon} C_t - (1-\tau_t)\frac{W_t}{A_t} A_t(j)^{\epsilon} \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} C_t \end{aligned} \tag{10}$$

where τ_t is an employment subsidy financed by lump-sum taxes. Notice that we have used the equilibrium condition $Y_t(j) = C_t(j)$. The real profit function thus is

$$\Pi_t(j) \equiv \frac{D_t(j)}{P_t} = C_t \left(\exp\left(p_t(j)\right)\right)^{1-\epsilon} - C_t(1-\tau_t)\frac{w_t}{A_t}\left(\exp\left(p_t(j)\right)\right)^{-\epsilon} = \Pi(p_t(j), w_t, A_t), \tag{11}$$

where w_t is the real wage and

$$p_t(j) \equiv \log\left(\frac{P_t(j)}{A_t(j)P_t}\right)$$

is the quality-adjusted (log) *relative price*. When prices do not change in nominal terms, $p_t(j)$ evolves according to

$$p_t(j) = p_{t-1}(j) + \log\left(\frac{P_{t-1}(j)}{A_t(j)P_t}\right) - \log\left(\frac{P_{t-1}(j)}{A_{t-1}(j)P_{t-1}}\right) = p_{t-1}(j) - \sigma_t \varepsilon_t(j) - \pi_t.$$

From now on, we drop the index j for ease of notation. Without loss of generality, a firm resets its price with probability $\lambda_t(p)$. Price resetting involves the firm paying a fixed menu cost η (in labor units). The optimal reset price p_t^* maximizes the firm's value, $p_t^* = \arg\max V_t(p)$, taking into account that this new price may not change for a random period of time. The firm's value is given by the equation

$$\begin{aligned} V_t(p) &= \Pi(p, w_t, A_t) + \mathbb{E}_t\left[\left(1 - \lambda_{t+1}\left(p - \sigma_{t+1}\varepsilon_{t+1} - \pi_{t+1}\right)\right)\Lambda_{t,t+1} V_{t+1}(p - \sigma_{t+1}\varepsilon_{t+1} - \pi_{t+1})\right] \\ &+ \mathbb{E}_t\left[\lambda_{t+1}\left(p - \sigma_{t+1}\varepsilon_{t+1} - \pi_{t+1}\right)\Lambda_{t,t+1}\left(\max_{p'} V_{t+1}(p') - \eta w_{t+1}\right)\right]. \end{aligned}$$

which comprises the current period profits $\Pi(\cdot)$ and the discounted continuation value $V_{t+1}(\cdot)$ evaluated when the price does not change at $t+1$ with probability $1 - \lambda_{t+1}(\cdot)$, and when the firm sets a new price after paying the menu cost at $t+1$, with probability $\lambda_{t+1}(\cdot)$. As in Golosov and Lucas (2007) fixed menu cost model, the adjustment probability is given by

$$\lambda_t(p) = \mathbb{1}[L_t(p) > 0]$$

where $\mathbb{1}[\cdot]$ is the indicator function, and

$$L_t(p) \equiv \max_{p'} V_t(p') - \eta w_t - V_t(p)$$

is the *gain from adjustment* (or loss from inaction), net of the menu cost.

2.3 Monetary policy rule

The central bank controls the short-term nominal interest rate i_t. In Section 4, we assume that the central bank follows a simple Taylor (1993) rule:

$$i_t = \rho_i i_{t-1} + (1-\rho_i)(-\log\beta + \phi_\pi \pi_t + \phi_y(y_t - y_t^e)), \tag{12}$$

where ρ_i is the smoothing in the Taylor rule, y_t^e is the efficient-level of output, and $\phi_\pi > 1$ and ϕ_y are parameters. In Section 5, we assume instead that the central bank follows the optimal policy with commitment.

2.4 Aggregation

Firms' individual price-setting decisions give rise to a distribution of prices. Let the density of quality-adjusted log relative prices at the end of period t be $g_t(p)$. The definition of the aggregate price index can then be written as:

$$1 = \int e^{p(1-\epsilon)} g_t(p)\, dp. \tag{13}$$

Individual firms' labor demand aggregates up to

$$N_t = \frac{C_t}{A_t}\int e^{p(-\epsilon)} g_t(p)\, dp + \eta \int \lambda_t(p - \sigma_t\varepsilon_t - \pi_t) g_{t-1}(p) dp, \tag{14}$$

such that the total number of hours worked equals the total use of labor for production (the first term on the right-hand side) and the aggregation of labor allocated to price adjustment (the second term) – note that $\int \lambda_t(p-\sigma_t\varepsilon-\pi_t)g_{t-1}(p)dp$ is the *frequency* of price adjustments.

Consider next the law of motion of the price density function:

$$g_t(p) = (1 - \lambda_t(p)) \int g_{t-1}(p + \sigma_t \varepsilon + \pi_t) d\xi(\varepsilon) + \delta(p - p_t^*) \int \lambda_t(\tilde{p}) \left(\int g_{t-1}(\tilde{p} + \sigma_t \varepsilon + \pi_t) d\xi(\varepsilon) \right) d\tilde{p},$$

where $\delta(\cdot)$ is the Dirac delta function. The first line describes the evolution of the mass of firms that do not change their nominal prices: their real quality-adjusted price is affected by idiosyncratic quality shocks and aggregate inflation. The second line captures the effect of price updating: the mass of all updating firms is relocated to the optimal reset price p_t^*.

2.5 Aggregate Shocks

The logarithm of aggregate productivity follows a first-order autoregressive process

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

where $\rho_A \in [0, 1]$ is the persistence and $\varepsilon_{A,t}$ the innovation. Likewise, we assume that the (lump-sum tax-financed) employment subsidy τ_t follows an autoregressive process which is interpretable as temporary *cost-push* shocks:

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

where τ is the steady-state employment subsidy, and $\rho_\tau \in [0, 1]$. We also consider autoregressive *dispersion* shocks to the idiosyncratic volatility of quality shocks:

$$\sigma_t - \sigma = \rho_\sigma (\sigma_{t-1} - \sigma) + \varepsilon_{\sigma,t},$$

where $\sigma > 0$ is the steady-state volatility, and $\rho_\sigma \in [0, 1]$. We use these shocks to gauge the implications of shocks to price dispersion. Finally, Section 4 also assumes i.i.d. shocks to the Taylor rule (12), $\varepsilon_{r,t}$.

2.6 Equilibrium

In order to achieve high accuracy in the computation, we find it convenient to recast the problem in terms of the *distance* x between actual (log-) prices p and the optimal (log-) reset

price p^*. Therefore, we define a new state variable $x_t \equiv p_t - p_t^*$. The dynamics of x_t are given by

$$x_t \equiv p_t - p_t^* = x_{t-1} + p_t - p_{t-1} - p_t^* + p_{t-1}^* = x_{t-1} - \sigma_t \varepsilon_t - \pi_t^*,$$

where $\pi_t^* \equiv p_t^* - p_{t-1}^* + \pi_t$ is the inflation rate of the (quality-adjusted) reset price. The advantage of this reformulation is that after a price reset, x_t always jumps back to zero.

Profits can then be expressed as

$$\Pi(x_t, p_t^*, w_t, A_t) = C_t \left(\exp\left(x_t + p_t^*\right)\right)^{1-\epsilon} - C_t(1-\tau_t)\frac{w_t}{A_t}\left(\exp\left(x_t + p_t^*\right)\right)^{-\epsilon},$$

and the Bellman equation can thus be re-written as

$$\begin{aligned} V_t(x) &= \Pi(x, p_t^*, w_t, A_t) \\ &+ \mathbb{E}_t\left[\left(1 - \lambda_{t+1}\left(x - \sigma_{t+1}\varepsilon_{t+1} - \pi_{t+1}^*\right)\right)\Lambda_{t,t+1} V_{t+1}(x - \sigma_{t+1}\varepsilon_{t+1} - \pi_{t+1}^*)\right] \\ &+ \mathbb{E}_t\left[\lambda_{t+1}\left(x - \sigma_{t+1}\varepsilon_{t+1} - \pi_{t+1}^*\right)\Lambda_{t,t+1}\left(V_{t+1}(0) - \eta w_{t+1}\right)\right]. \end{aligned} \tag{15}$$

The optimality of price updating implies the following conditions for the lower and upper bounds of the inaction region s_t and S_t, which form the so-called (S, s) band

$$V_t(0) - \eta w_t = V_t(s_t), \tag{16}$$

$$V_t(0) - \eta w_t = V_t(S_t). \tag{17}$$

The optimality of the reset price requires $V_t'(0) = 0$. $V_t'(0)$ can be expressed as the sum of the marginal effect of x on current profits and on the continuation value conditional on not-updating and updating the price (where $\phi(\cdot)$ is the standard normal pdf):[4]

$$\begin{aligned} 0 = V_t'(0) = \Pi_t'(0) &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{s_{t+1}}^{S_{t+1}} V_{t+1}(x') \left.\frac{\partial \phi\left(\frac{x - x' - \pi_{t+1}^*}{\sigma_{t+1}}\right)}{\partial x}\right|_{x=0} dx' \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}}\left(\phi\left(\frac{-S_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right)\left(V_{t+1}(0) - \eta w_{t+1}\right). \end{aligned} \tag{18}$$

[4] These three conditions are sufficient only if the firm's value function is convex in x_t. We check convexity ex-post.

Appendix A derives this expression. The law of motion of the density is

$$g_t(x) = (1-\lambda_t(x))\int g_{t-1}(x+\sigma_t\varepsilon+\pi_t^*)d\xi(\varepsilon) + \delta(x)\int \lambda_t(\tilde{x})\left(\int g_{t-1}(\tilde{x}+\sigma_t\varepsilon+\pi_t^*)d\xi(\varepsilon)\right)d\tilde{x}, \tag{19}$$

Finally, the aggregate price index and labor market clearing can be expressed as

$$1 = \int e^{(x+p_t^*)(1-\epsilon)} g_t(x)\,dx, \tag{20}$$

$$N_t = \frac{C_t}{A_t}\int e^{(x+p_t^*)(-\epsilon)} g_t(x)\,dx + \eta\int \lambda_t(x+p_t^*-\sigma_t\varepsilon_t-\pi_t^*)g_{t-1}(x)dx. \tag{21}$$

Equations (6), (7), (15)-(21), together with a policy such as (12), define an *equilibrium* in $g(\cdot)$, $V_t(\cdot)$, C_t, N_t, w_t, i_t, p_t^*, s_t, S_t, π_t^*.

3 Calibration

Our baseline calibration is presented in Table 1. There are four blocks: household's preferences, firms' price setting behavior, the Taylor rule used in Section 4, and shocks processes.

This calibration relies as much as possible on the existing literature. Regarding preferences, the monthly discount factor is $0.96^{1/12}$. The elasticity of substitution across products is $\epsilon = 7$, so the frictionless net markup is $1/6$. We follow Midrigan (2011) in assuming log utility in consumption, so the relative risk aversion coefficient is $\gamma = 1$; and the weight on leisure in utility is $\nu = 1$. These assumptions yield that the real wage equals aggregate consumption, $w_t = C_t$ (and nominal wages equal nominal aggregate demand).

Regarding price setting, we set the menu cost $\eta = 3.59\%$, and the standard deviation of idiosyncratic shocks in the absence of dispersion shocks to be $\sigma = 2.36\%$. These choices allow us to match in the Ramsey steady state the 8.7% of average frequency of price changes as well as the 8.5% of average absolute size of non-zero price changes documented by Nakamura and Steinsson (2008) for the U.S.[5] For the Taylor rule governing monetary policy in Section 4, we assume $\phi_\pi = 1.5$ and $\phi_y = 0.5/12$ as in Taylor (1993) (adjusted for our monthly calibration). A common assumption is a smoothing component in the Taylor rule, in our

[5]The calibration is conducted at the optimal steady state inflation rate of 0.25%.

Table 1: Parameter values

		Households	
β	$0.96^{1/12}$	Discount rate	Golosov and Lucas (2007)
ϵ	7	Elasticity of substitution	Golosov and Lucas (2007)
γ	1	Risk aversion parameter	Midrigan (2011)
υ	1	Utility weight on labor	Set to yield $w = C$
		Price setting	
η	0.036	Menu cost	Set to match 8.7% of frequency of price changes and 8.5% of absolute size of price changes documented in Nakamura and Steinsson (2008)
σ	0.024	Std dev of quality shocks	
		Monetary policy	
ϕ_π	1.5	Inflation coefficient in Taylor rule	Taylor (1993)
ϕ_y	0.5/12	Output gap coefficient in Taylor rule	Taylor (1993)
ρ_i	$0.75^{1/3}$	Smoothing coefficient	
		Shocks	
ρ_A	$0.95^{1/3}$	Persistence of the TFP shock	Smets and Wouters (2007)
ρ_τ	$0.9^{1/3}$	Persistence of the cost-push shock	Smets and Wouters (2007)
ρ_σ	$0.75^{1/3}$	Persistence of the dispersion shock	

case, $\rho_i = 0.75^{1/3}$. Finally, the persistence of shocks is taken from Smets and Wouters (2007), once transformed from quarterly to monthly frequency: $0.95^{1/3}$ for aggregate productivity shocks, $0.9^{1/3}$ for aggregate subsidy shocks (interpreted as cost-push shocks). The standard deviation of idiosyncratic quality shocks at the product level is assumed to be $0.75^{1/3}$.

4 The nonlinear Phillips curve

As a prerequisite for our normative analysis, we now explore the key positive properties of our model. For this, we characterize its behavior through three equations: an IS curve, a Taylor rule, and the price-setting block distilled into a nonlinear Phillips curve. We contrast the framework to the three-equation new Keynesian model, which forms the basis for textbook optimal monetary policy analysis (Woodford, 2003; Galí, 2008). The main difference regards the Phillips curve.

The key distinctive feature of our model is the endogenous response of the frequency of price changes. This leads to a nonlinear Phillips curve and, consequently, a state-dependent

inflation-output trade-off involved in monetary policy decisions. In the standard linearized Calvo framework, the frequency of price changes is constant, the Phillips curve is log-linear and the inflation-output trade-off implied by its slope is constant. In contrast, in our setup they depend on the size of aggregate shocks and the state of the economy. This nonlinearity is quantitatively important when large shocks trigger a response in the frequency of price changes – for instance, as observed during the post-COVID inflationary episode.

4.1 A three-equation framework

To start, we briefly sketch a three-equation framework and discuss its similarities and departures from the new Keynesian framework.

The first equation is the IS curve which can be derived from the Euler equation (7):

$$\tilde{y}_t^e = -\frac{1}{\gamma}\left(i_t - \mathbb{E}_t[\pi_{t+1}] - r_t^e\right) + \mathbb{E}_t[\tilde{y}_{t+1}^e] = -\frac{1}{\gamma}\sum_{j=0}^{\infty}\mathbb{E}_t[r_{t+j} - r_{t+j}^e] \tag{22}$$

where $\tilde{y}_t^e \equiv \log(Y_t/Y_t^e) - \log(Y/Y^e)$ denotes the gap between output and its efficient level (adjusted for the steady-state gap), i_t is the monetary policy instrument – the nominal, one-period-ahead, interest rate, $\mathbb{E}_t\pi_{t+1}$ is expected inflation, and r_t^e is the efficient real interest rate. The efficient output level and interest rates are those prevailing in a counterfactual efficient allocation absent price rigidities (see Appendix B.1). This IS curve implies that the output gap is determined by the cumulative sum of future real interest rate gaps ($r_t = i_t - \mathbb{E}_t[\pi_{t+1}]$). The IS curve is globally log-linear.

The second equation, which we use for positive analysis only, is the Taylor rule in equation (12) describing the endogenous response of monetary policy, which we repeat here for convenience:

$$i_t = \rho_i i_{t-1} + (1-\rho_i)\left(\phi_\pi \pi_t + \phi_y \tilde{y}_{t+1}^e\right).$$

This expression is globally log-linear by assumption. Therefore, the first two equations cannot generate nonlinear dynamics. Our baseline model shares these equations with the canonical new Keynesian model.

This leads us to the third equation, the new Keynesian Phillips curve. In the Calvo (1983) model, which assumes a constant Poisson probability that individual firms' prices are

updated, it can be approximated up to a first order by a single equation:

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa \tilde{y}_t^e + u_t \tag{23}$$

where u_t is the cost-push shock. The Phillips curve *slope* is captured by parameter κ. That is, the impact of a unit change in the output gap on inflation does not depend on the size of the shock or the state of the economy.

Our nonlinear menu cost model the Phillips curve is different. Nevertheless, as shown by Auclert et al. (2024), aggregate dynamics in menu cost models are well-approximated by equation (23) *when shocks are small*, i.e., when there is no significant response of the frequency of price changes, with one caveat: κ in equation (23) should not be calibrated using the frequency of price changes from micro data, but instead a fictitious higher frequency to mimic the "selection effect" inherent in the menu cost model. As shown by Caplin and Spulber (1987), Caballero and Engel (2007), Golosov and Lucas (2007), Alvarez et al. (2021), the "selection effect" refers to the positive correlation existing in menu cost models between the expected magnitude of a price change and the probability of the price change, when an aggregate shock hits. This feature is absent in Calvo (1983), so even if the frequency of price changes is the same in both models and remains constant to shocks, aggregate prices exhibit higher flexibility in menu cost models.

We deviate from this equivalence result by exploring the monetary policy implications of our menu cost model when *shocks are large* in the sense that the model predicts a significant response of the frequency of price changes. This is what we do next.

4.2 Nonlinear New Keynesian Phillips Curve

This subsection stresses that the relationship between inflation and the output gap is strongly nonlinear for large shocks in the menu cost model. Thus, equation (23) is not an accurate description of aggregate inflation dynamics even if adjusted via recalibration of κ.

To make this point, we simulate generalized impulse responses to i.i.d monetary policy shocks of varying sizes, starting from the deterministic steady state. We store the impulse responses of inflation and the output gap.[6] Solving the new Keynesian Phillips curve (23)

[6]These impulse responses are computed nonlinearly under perfect foresight. For small shocks, this is equivalent

forward with $u_t = 0$ yields,

$$\pi_t = \kappa \sum_{j=0}^{\infty} \beta^i \mathbb{E}_t \tilde{y}^e_{t+j}. \tag{24}$$

With this equation in mind, Figure 1 shows, in panel (a), the relationship between inflation π_t and the cumulative discounted sum of output gaps, $\sum_{j=0}^{\infty} \beta^i \mathbb{E}_t \tilde{y}^e_{t+j}$, in our menu cost model (solid blue line); and for the standard Phillips curve (Calvo recalibrated as in Auclert et al., 2024, the dashed red line). For small shocks both models exhibit the same slope by construction. For large shocks, the Phillips curve becomes steeper and eventually even bends backwards in the menu cost model, while it stays close to linear in the Calvo model.

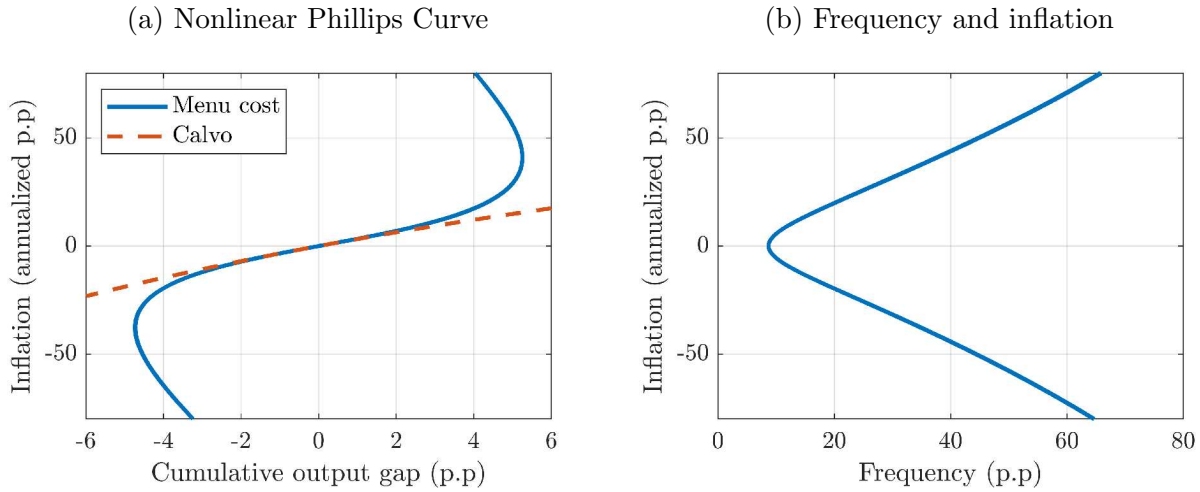


Figure 1: Inflation-output trade-offs in the menu cost and Calvo models.

Note: The figure is produced by computing the impulse responses to monetary policy shocks of different magnitudes assuming no persistence of the Taylor rule ($\rho_i = 0$). The solid blue line is the baseline menu cost model and the dashed red line is the recalibrated Calvo model.

To provide intuition, panel (b) in Figure 1 depicts the relationship between frequency of price changes and inflation. When shocks are small the frequency of price changes remains close to its steady-state level, 8.7% in our calibration. However, as shocks get larger, the response of both inflation and frequency rise.[7] This makes aggregate prices more flexible, which implies that the responsiveness of the cumulative output gap decreases in the shock

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. (2023): an unexpected once-and-for-all large shock that hits the economy in the deterministic steady state.

[7]The u-shape of the frequency response is a robust feature of menu cost models: the new price increases that are triggered by the small shock are almost fully offset by the new decreases that are canceled. For large shocks, however, price decreases disappear and price increases generate a large frequency response (Gagnon, 2009; Karadi and Reiff, 2019; Alvarez and Neumeyer, 2020; Alexandrov, 2020).

size. At frequencies over 40% the Phillips curve becomes backward bending. At this level, monetary policy impulses reach their maximum effectiveness in stimulating activity, and any larger policy easing raises inflation with smaller output effects. However, such shock sizes are fairly extreme. For the rest of the analysis we restrict our attention to shocks that can be large, but not as large as to go beyond this point.

While the location of the inflation-output gap pairs on the Phillips curve in Figure 1 is informative about the *size dependence* of the impact of shocks, the slope of the Phillips curve at each point is informative about the *state dependence* of marginal shocks. The slope reflects the state-dependence of the inflation-output trade-off involved in monetary policy decisions: how much cumulative output gap must decline to reduce inflation by a unit, also known as the *sacrifice ratio* of monetary policy. This slope almost doubles when frequency reaches 20%, a magnitude documented during the post-COVID inflation surge (Montag and Villar, 2023).[8] While in a low-frequency and low-inflation environment, the sacrifice ratio is high, it becomes much lower once frequency and inflation increases.[9]

Finally, the Phillips curve relationship features a certain degree of *sign-dependence or asymmetry*, even if it is small. The intuition relies on the asymmetry of firms' profit function: low relative prices lead to high demand, while high relative prices lead to low demand. Therefore, a negative deviation of the relative price from the optimum causes higher losses than a positive deviation. Hence, an inflationary aggregate shock (which reduces the relative prices of non-adjusters), calls for more aggressive price increases and makes aggregate prices more flexible than a disinflationary aggregate shock of the same absolute size. Higher shock persistence increases the expected cumulative impact on the relative price if firms do not adjust, amplifying this asymmetry.

At this point it is important to stress, that the true relationship between inflation and output is not determined by one stable *structural* equation alone, but by a set of equations and variables that jointly determine a relationship inflation and output. This relationship is dynamic and depends on potentially all state variables of the model, both past and future,

[8] As a complement to Figure 1, Figure 12 in Appendix D.1 depicts the "slope of the Phillips curve" for positive shocks of different sizes, which give rise to a range of realistic frequency values (the relationship is analogous for negative shocks).

[9] Blanco et al. (2024b) also discuss how the sacrifice ratio changes with the level of inflation

including the price distribution and aggregate shocks. Despite this, we find numerically that this relationship is relatively stable, so that we can characterize some of its relevant qualitative features as we did above (see Section 7 below).

4.3 Size-dependence in response to a cost-push shock

To provide a benchmark for uncoming optimal policy analysis, Figure 2 shows the responses of inflation (panel a), output gap (panel b), frequency of price changes (panel c) and interest rates (panel d) to cost-push shocks of different magnitudes under a Taylor rule. The shock is a persistent decline in employment subsidy τ_t. Inflation and frequency are evaluated on impact, output gap is evaluated at its peak response (5-6 months after impact) and the real interest rate gap is the cumulated. The figure displays both the menu cost model (solid blue line) and the (recalibrated) Calvo model (dashed red line).

In Figure 2, the nonlinearity is quantitatively small when the shock is small enough to yield a response of frequency below 10 p.p.. In this region, the Calvo model behaves similarly.[10] For larger cost-push shocks, the menu cost model yields a larger response of inflation (panel a). As the frequency approaches 20% (up 11 p.p. from the steady state of 8.7%), the response of inflation in the menu cost model is significantly larger than in the Calvo model. The intuition again hinges on the increase in the frequency of price changes as a result of the shock, which makes the price level endogenously more flexible. As the Taylor rule is the same for both models, the interest rate responds more in the menu cost model (panel d). This leads to a marginally amplified response of the output gap in the menu cost model (panel b).

5 Optimal monetary policy problem and computational approach

We turn next to the analysis of optimal monetary policy. We consider optimal monetary policy under commitment. In this section we introduce the central bank's problem and present a new computational method to deal with the complexities associated with the high-dimensionality of this problem.

[10] Figure 15 in Appendix D.3 shows how the full impulse responses to a small cost-push shock are similar to those of the recalibrated Calvo model.

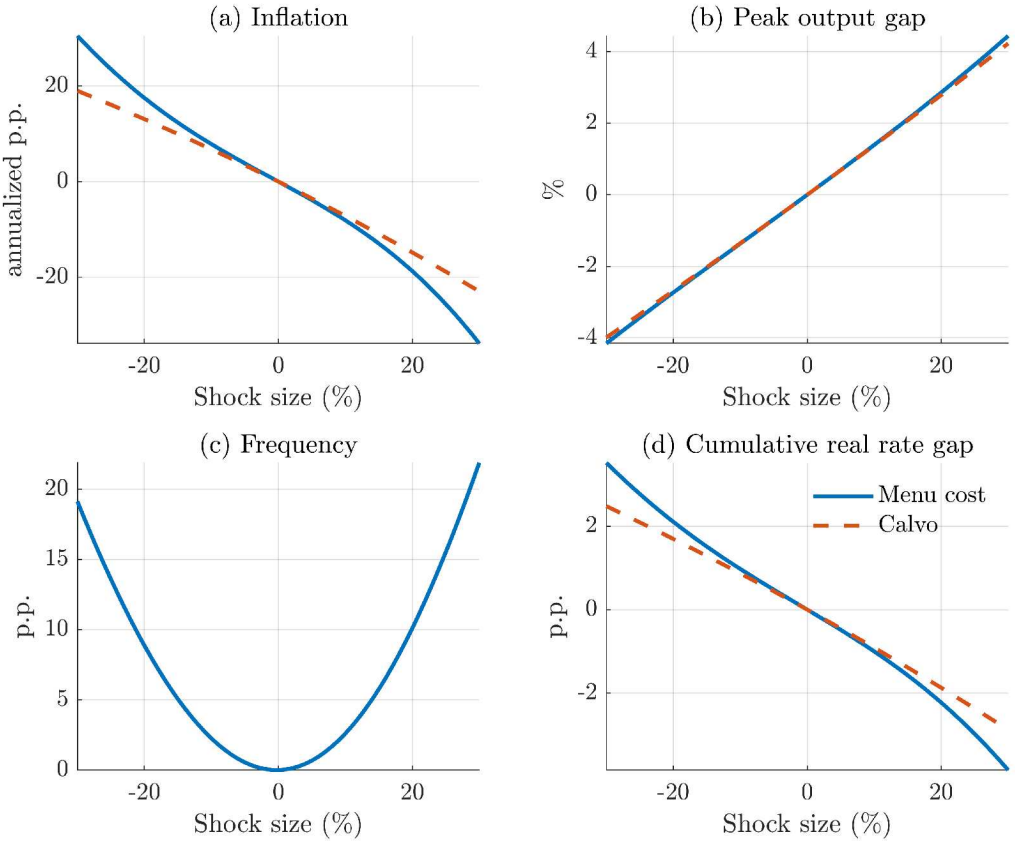


Figure 2: Response to a cost-push shock for different shock magnitudes.

Note: The figure displays the difference in the value of inflation and the frequency of price changes between the period of the shock arrival and the value in the deterministic steady state. The output gap is evaluated at the peak of the impact relative to the deterministic steady state. Panel (d) shows the difference of the real interest rate from steady state, cumulated from $t = 1$ to $t = \infty$.

5.1 Ramsey problem

The central bank maximizes households' welfare under commitment (Galí, 2008). The problem implies that the central bank selects the path for all equilibrium variables subject to all the competitive equilibrium conditions:

$$\max_{\substack{\{g_t^c(\cdot), g_t^0, V_t(\cdot), C_t, \\ w_t, p_t^*, s_t, S_t, \pi_t^*\}_{t=0}^{\infty}}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\gamma}}{1-\gamma} - \upsilon \frac{C_t}{A_t} \left(\int e^{(x+p_t^*)(-\epsilon_t)} g_t^c(p)\, dx + g_t^0 e^{(p_t^*)(-\epsilon)} \right) - \upsilon \eta g_t^0 \right)$$

subject to

$$
\begin{aligned}
w_t &= \upsilon C_t^{\gamma}, \\
V_t(x) &= \Pi(x, p_t^*, w_t, A_t) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{s_{t+1}}^{S_{t+1}} \left[V_{t+1}(x') \phi \left(\frac{(x - x') - \pi_{t+1}^*}{\sigma_{t+1}} \right) \right] dx' + \\
&\quad \Lambda_{t,t+1} \left(1 - \frac{1}{\sigma_{t+1}} \int_{s_{t+1}}^{S_{t+1}} \left[\phi \left(\frac{(x - x') - \pi_{t+1}^*}{\sigma_{t+1}} \right) \right] dx' \right) \left[\left(V_{t+1}(0) - \eta w_{t+1} \right) \right], \\
V_t(s_t) &= V_t(0) - \eta w_t, \\
V_t(S_t) &= V_t(0) - \eta w_t, \\
0 &= \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{s_{t+1}}^{S_{t+1}} V_{t+1}(x') \left. \frac{\partial \phi \left(\frac{x - x' - \pi_{t+1}^*}{\sigma_{t+1}} \right)}{\partial x} \right|_{x=0} dx' \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi \left(\frac{-S_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}} \right) - \phi \left(\frac{-s_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}} \right) \right) \left(V_{t+1}(0) - \eta w_{t+1} \right). \\
g_t^c(x) &= \frac{1}{\sigma_t} \int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1}) \phi \left(\frac{x_{-1} - x - \pi_t^*}{\sigma_t} \right) dx_{-1} + g_{t-1}^0 \phi \left(\frac{-x - \pi_t^*}{\sigma_t} \right), \\
g_t^0 &= 1 - \int_{s_t}^{S_t} g_t^c(x) dx, \\
1 &= \int e^{(x + p_t^*)(1-\epsilon)} g_t^c(x) \, dx + g_t^0 e^{(p_t^*)(1-\epsilon)}.
\end{aligned}
$$

where $\phi(\cdot)$ is the probability density function of a normal random variable, g_t^c is the continuous part of the price-gap density, g_t^0 is the frequency of repricing, that is, the probability mass of updated prices, and s_t and S_t are the endogenous boundaries of the inaction region for x. The nominal interest rates consistent with this path of nominal and real variables can then be recovered from the household's Euler equation (7). Importantly, the constraint set of the planner's problem is continuous and differentiable despite the fact that the individual firm's price policy function is not. This is so because each firm has zero mass, and thus the discontinuity in a single firm's behavior does not lead to a discontinuity in aggregates. Furthermore, note that both $V_t x)$ and $g_t^c(x)$ are continuously differentiable over the relevant range (s_t, S_t).

5.2 Computational method

We propose a new algorithm similar to that in González et al. (2024) but applied to discrete time. The idea is to represent the problem as a high-dimensional dynamic programming

problem in which the Bellman equation and the law of motion of the price-gap distribution are constraints.

The problem of the Ramsey planner is complicated by the fact that the value function $V_t(\cdot)$ and the distribution $g_t(\cdot)$ are infinite-dimensional variables. This poses a challenge when solving the optimal monetary policy problem, as we need to compute the first-order conditions (FOCs) with respect to these infinite-dimensional variables.[11] Our algorithm first discretizes the planner's objective and constraints (the private equilibrium conditions) and then determines the FOCs, instead of first determining the FOCs for the planner's continuous space problem, and then discretizing them. Thus we transform the original infinite-dimensional problem into a high-dimensional problem, in which the value function and the state density are replaced by large vectors with a dimensionality equal to the number of grid points used to approximate the individual state space. This approximation needs to be smooth and accurate enough to capture the higher-order effects of policy.

An additional challenge in our particular problem is that a simple discrete-state approximation may fail, as the private equilibrium conditions include discrete choices. Therefore, we approximate the distribution and value function not by discrete functions on a predetermined grid, but by *piece-wise linear* functions over an *endogenous* grid. The endogenous grid is selected to always include the two boundaries of the inaction region (points $x = s$ and $x = S$) and the optimal price gap ($x = 0$). Furthermore, we explicitly take the mass point at 0 into account in the distribution. Integrals to compute expectations are evaluated algebraically, conditional on those piecewise linear functions.

As we show in Appendix E, the Bellman equation can thus be approximated over a grid of price gaps x as

$$\mathbf{V}_t = \mathbf{\Pi}_t + [\mathbf{A}_t \mathbf{V}_{t+1} - \mathbf{b_{t+1}} \eta w_{t+1}]$$

[11] There are a number of proposals in the literature to deal with this problem. Nuño and Thomas (2022), Smirnov (2022), and Dávila and Schaab (2022) deal with the full infinite-dimensional planner's problem in continuous time. This implies that the Kolmogorov forward (KF) and the Hamilton-Jacobi-Bellman (HJB) equations are constraints faced by the central bank. They derive the planner's FOCs using calculus of variations, thus expanding the original problem to also include the Lagrange multipliers, which in this case are also infinite-dimensional. These papers solve the resulting differential equation system using the upwind finite-difference method of Achdou et al. (2021). Bhandari et al. (2021) make the continuous cross-sectional distribution finite-dimensional by assuming that there are N agents instead of a continuum. They then derive standard FOCs for the planner. In order to cope with the large dimensionality of their problem, they employ a perturbation technique. Le Grand et al. (2022) employ the finite-memory algorithm proposed by Ragot (2019). It requires changing the original problem such that, after K periods, the state of each agent is reset. This way the cross-sectional distribution becomes finite-dimensional.

where $\mathbf{V}_t$ and $\mathbf{b}_t$ are vectors with the value function and the expected adjustment probability evaluated at different grid points, respectively, and $\mathbf{A}_t$ is a matrix that captures the idiosyncratic transitions due to firm-level quality shocks and aggregate inflation. Similarly, the law of motion of the density for $x \neq 0$ is

$$\mathbf{g}_t^c = \mathbf{F}_t \mathbf{g}_{t-1}^c + \mathbf{f}_t g_{t-1}^0.$$

where $\mathbf{g}_t^c$ and $\mathbf{f}_t$ are vectors representing the probability distribution function and the scaled and shifted normal distribution, respectively, $\mathbf{F}_t$ is a matrix that captures the evolution of the price distribution due to firm-level quality shocks and aggregate inflation, and

$$g_t^0 = 1 - \mathbf{e}_t^\intercal \mathbf{g}_t^c.$$

is the mass point at $x = 0$. Here $\mathbf{e}_t$ is a vector of weights corresponding to the trapezoid rule. The labor market clearing condition and the definition of the price index can be written in a similar form. The computational Appendix E provides futher details.

Once we have the discretized version of the problem, we find the planner's FOCs by symbolic differentiation. This delivers a large-dimensional system of difference equations, as we have Lagrange multipliers associated with each gridpoint of the value function or the probability function.

Next, we find the Ramsey steady state. To do so, we construct a nonlinear multidimensional function mapping one variable, in our case inflation, to the rest of the steady-state equilibrium variables. We then combine this function with the planner's FOCs. As the system is linear in the Lagrange multipliers, the solution boils down to finding the zero of a nonlinear function of the initial variable (i.e., inflation), for instance, using the Newton method. Finally, to compute the dynamics of the Ramsey problem, we solve the system of difference equations nonlinearly in the sequence space also using the Newton method.

The symbolic differentiation and the two applications of the Newton algorithm can be conveniently automated using several available software packages. In our case, we employ Dynare (Adjemian et al., 2023), but the approach is also compatible with the nonlinear sequence-space Jacobian toolboxes. This algorithm can be employed to compute optimal

policies in a large class of heterogeneous-agent models. Compared to other techniques, it stands out for being easy to implement. González et al. (2024) show that this algorithm delivers the same results as computing the FOCs by hand using calculus of variations and then discretizing the model. Our algorithm runs in a few minutes on a normal laptop.

6 Optimal monetary policy: results

We now proceed to investigate the model's normative prescriptions.

6.1 The steady state under the optimal policy

The solution of the Ramsey planner's problem has a steady state featuring a slightly positive inflation of 0.25%.[12] This is different from the standard New Keynesian model with Calvo pricing (Galí, 2008), where the optimal inflation in the Ramsey steady state is zero. The value of inflation in the Ramsey steady state in the menu cost model is very close to the value of steady-state inflation that maximizes steady-state welfare, which in turn is also very close to the value of inflation that minimizes the frequency of price adjustments.

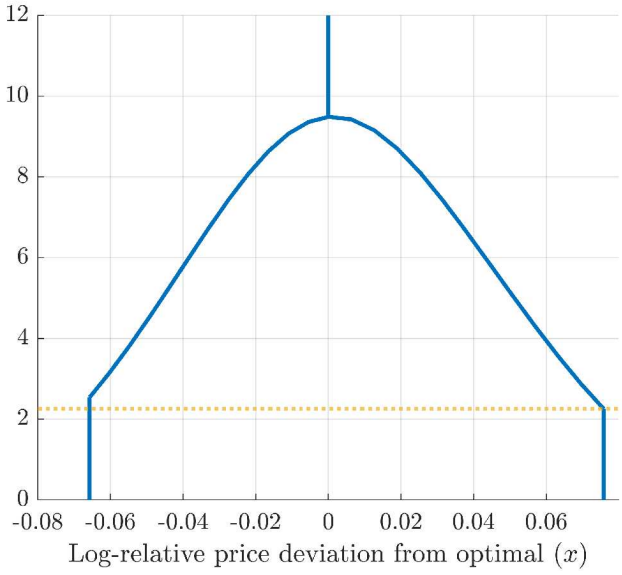


Figure 3: Steady-state price-gap density.

Note: 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.

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

What explains the positive optimal inflation? The key is the asymmetry of the profit function (11). For a firm, a negative price gap is more undesirable than a positive price gap of the same size because a negative price gap $-x$ leads to much larger sales at a markup loss of $-x$, while the positive price gap x leads to only somewhat smaller sales at a markup gain of x. This implies that the (S, s) band is asymmetric: the lower threshold s_t is closer to the optimal price than the upper one S_t (see Figure 3). 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 this effect is small, but not negligible.

6.2 Timeless optimal monetary response to cost-push shocks

While the Ramsey steady state is a prerequisite to any dynamic analysis, our main focus is on the systematic monetary policy response to shocks. We therefore analyze *timeless* optimal monetary policy (Woodford, 2003; Galí, 2008). This corresponds to the optimal monetary policy starting from the Ramsey steady state when a shock hits the economy and all of the Lagrange multipliers are initialized at their steady-state values.

We start with the analysis of cost-push shocks. Such shocks have been proposed as relevant drivers of the recent inflation surge (see also Section 8). More importantly, in the standard New Keynesian model they are well known to break the divine coincidence, and call for a policy of "leaning against the wind". That is, the central bank tolerates a temporary increase in inflation to cushion the decline in output below its efficient level. Such a policy exploits the Phillips curve relationship and is thus the relevant case to explore the implications of a nonlinear Phillips curve.

Impulse response to a cost-push shock. The optimal monetary policy response to a cost-push shock in our model is also characterized by leaning against the wind. The solid blue line in Figure 4 shows the response of the economy to an inflationary large cost-push

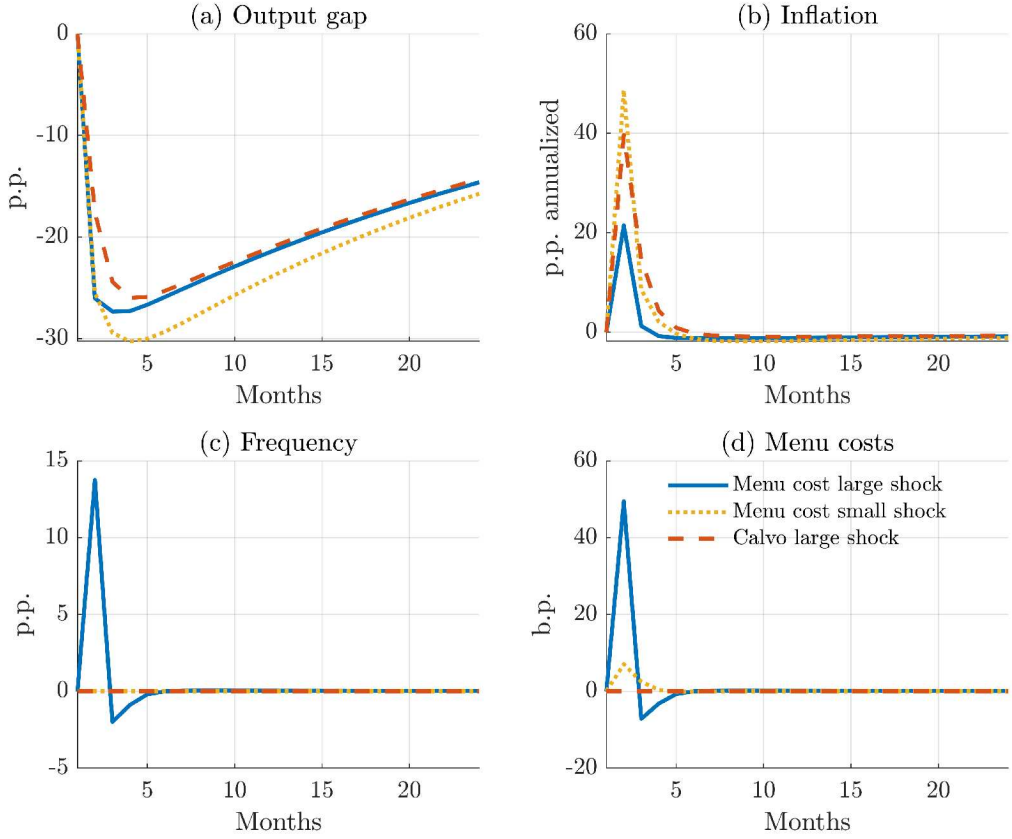


Figure 4: Impulse responses to a large cost-push shock in the Calvo model, and to a large and a small cost-push shock in the menu cost model under the optimal monetary policy.

Note: all displayed variables except frequency are linearly scaled in the small shock according to the ratio between the large and small shocks.

shock under the optimal policy, which pushes the repricing frequency above 20%. Inflation goes up (panel b) and the efficient output gap drops (panel a).

However, due to the nonlinearity of the Phillips curve, the policy is now size dependent. The figure compares the response to a large shock (solid blue line) with that of a small shock that is scaled linearly by the relative shock size (yellow dotted line). The optimal monetary policy is tighter for the large shock: it prevents inflation from increasing as much as in the small-shock case (panel b).

The figure also displays the optimal policy response to a large shock in the case of a Calvo model calibrated to replicate the slope of the Phillips curve (red dashed line). The figure shows that the tighter policy in the menu cost model leads to a somewhat more adverse

output-gap response relative to the Calvo model (panel a). For small shocks, the policy in the two models approximately coincides (not shown).

Impact of shock size. Figure 5 illustrates the optimal response to a cost-push shock for a whole range of shock sizes. The response of 4 variables are plotted as a function of the shock size in both the menu cost model (solid blue line) and the Calvo model (dashed red line). The Calvo model is almost linear in the shock size and behaves almost identical to the menu cost model for small shocks. The deviation between the two lines illustrates the nonlinearity of the economy under optimal policy for large shocks. The larger the shock the more inflation and frequency are contained as the central bank tightens monetary policy more aggressively. The central bank thus ´strikes while the iron is hot'.

Second, there is a certain sign-dependence, as the optimal policy response to positive and negative shocks differ. This is related to the asymmetries discussed in Section 4.

Nonlinear targeting rule. Optimality in the linearized Calvo model, requires a particularly simple relationship, known as the targeting rule: inflation and the change in the output gap are proportional to each other at each point in time. The proportionality factor is given by the elasticity of substitution $-1/\epsilon$.

Motivated by this result, Panel a of Figure 6 shows the relationship between annualized inflation and the change in output gap on impact in response to cost-push shocks of different sizes. The slope of the "target rule" relationship at zero output gap in the menu cost model (blue line) is very close to $-1/\epsilon$ (dotted yellow line), the slope of the target rule in the linear Calvo model. In the case of small shocks, the change in frequency is negligible, and thus the logic of the Calvo framework still applies. In the menu cost model the relationship is nonlinear (blue solid line). At a frequency of around 20% (11.3% in the figure), the slope of the optimal target rule becomes significantly lower than its slope at the steady-state frequency of 8.7%. The relationship means that after a large shock that increases frequency substantially, the central bank is stabilizing inflation more relative to the output gap on the margin than after small shocks. The central bank thus "leans against frequency", tightening policy more aggressively in the case of a large shock that increases frequency. In the nonlinear Calvo model (red dashed line) the nonlinearity is negligible.

Welfare decomposition. To understand the rationale for the optimal policy, we de-

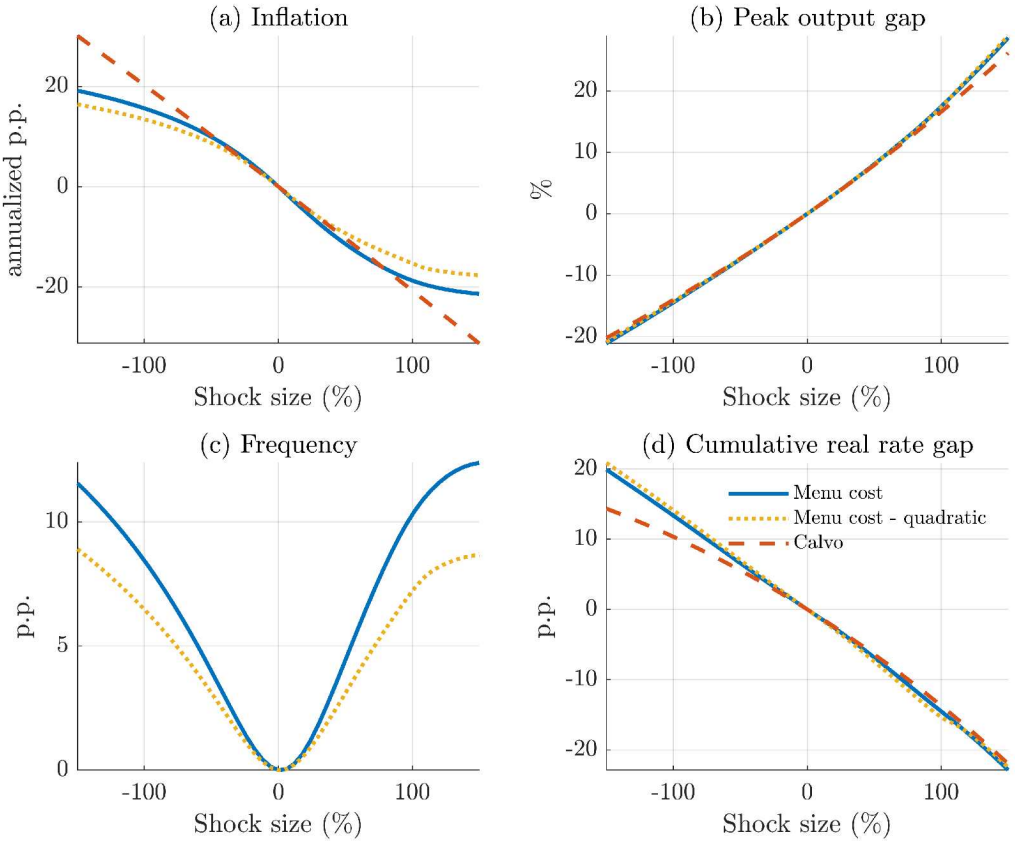
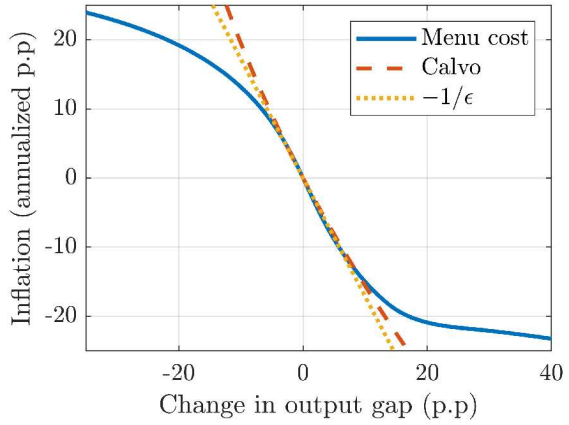


Figure 5: Optimal response to a cost-push shock for different shock magnitudes.

Note: The figure displays the difference in the value of inflation, output gap, and repricing frequency between the period after the shock arrival and the value in the deterministic steady state. The real interest rate is evaluated over the entire life of the shock.

compose the welfare gap relative to the first-best efficient equilibrium ($W_0 - W_0^e$), into three components. The *misallocation* caused by non-zero markups ($\mu_t(i) = p_t(i) - mc_t(i)$) reduces welfare relative to the first-best, where all markups are zero. Misallocation, in turn, can be decomposed into two components. First, the *average markup gap*, that is the degree of misallocation caused by the average markup $\overline{\mu}_t$. Second, *price dispersion*, which affects the dispersion of the demeaned markups $\zeta_t^{\mu-\overline{\mu}}$. The average markup gap thus describes *aggregate* over- or under-consumption, while the price dispersion gap refers to the inefficient *relative* consumption of different goods. Finally, the third component weighing down welfare is the utility loss from the labor allocated to price adjustment (*menu costs*). In Appendix B.2, we derive that the welfare gap equals

(a) Nonlinear targeting rule



(b) Frequency and inflation

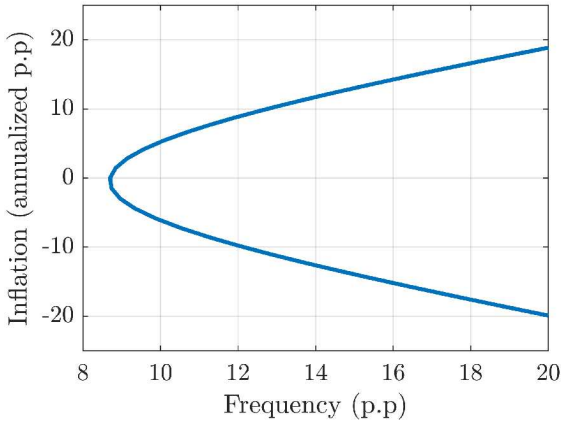


Figure 6: Optimal target rule under optimal policy.

Note: Panel (a) shows the relationship between annualized inflation and the change in output gap on impact as a function of a series of cost-push shocks under the optimal policy (blue solid line); and contrasts it to the optimal target rule in the Calvo model (dashed red line) and in the linearized Calvo model (dotted yellow line). Panel (b) shows the frequency at each inflation level.

$$W_0 - W_0^e = \sum_{t=0}^{\infty} \beta^t \left(\underbrace{-\log(1-\tau_t) - \overline{\mu}_t - \left(\frac{1}{e^{\overline{\mu}_t}(1-\tau_t)} - 1 \right)}_{\text{Average markup gap}} \right.$$

$$\left. \underbrace{- \left(\frac{1}{e^{\overline{\mu}_t}(1-\tau_t)} (\zeta_t^{\mu-\overline{\mu}} - 1) \right)}_{\text{Price dispersion}} \underbrace{-\eta g_t^0}_{\text{Menu costs}} \right). \quad (25)$$

Figure 7 displays the decomposition of the welfare gap conditional on cost push shocks of different sizes for the menu cost and the Calvo models.[13] Note that the above decomposition applies to both models, even though the last term is trivially zero in Calvo. The two components in the Calvo model both increase with the absolute shock size. In the menu cost model, however, price dispersion does not increase, but *decreases*, in absolute shock size.

[13] We have set the steady-state labor subsidy τ to offset the average markup distortions in steady state, i.e. $\overline{\mu} = -\log(1-\tau)$. The steady state tax ($\tau = 14.6\%$) is close, but somewhat higher than the inverse of the elasticity of substitution $1/7 = 14.3\%$. Thus, the steady-state welfare gap comes exclusively from the markup dispersion ($\zeta^{\mu-\overline{\mu}}$) and the menu costs. In our baseline calibration, these terms are similar in magnitude and equal to 0.3 percent of steady state output. Menu costs of such magnitude are reasonable: Zbaracki et al. (2004) estimates menu costs including managerial and marketing costs around 1.2 percent of firm revenues, which is higher, but of a similar magnitude.

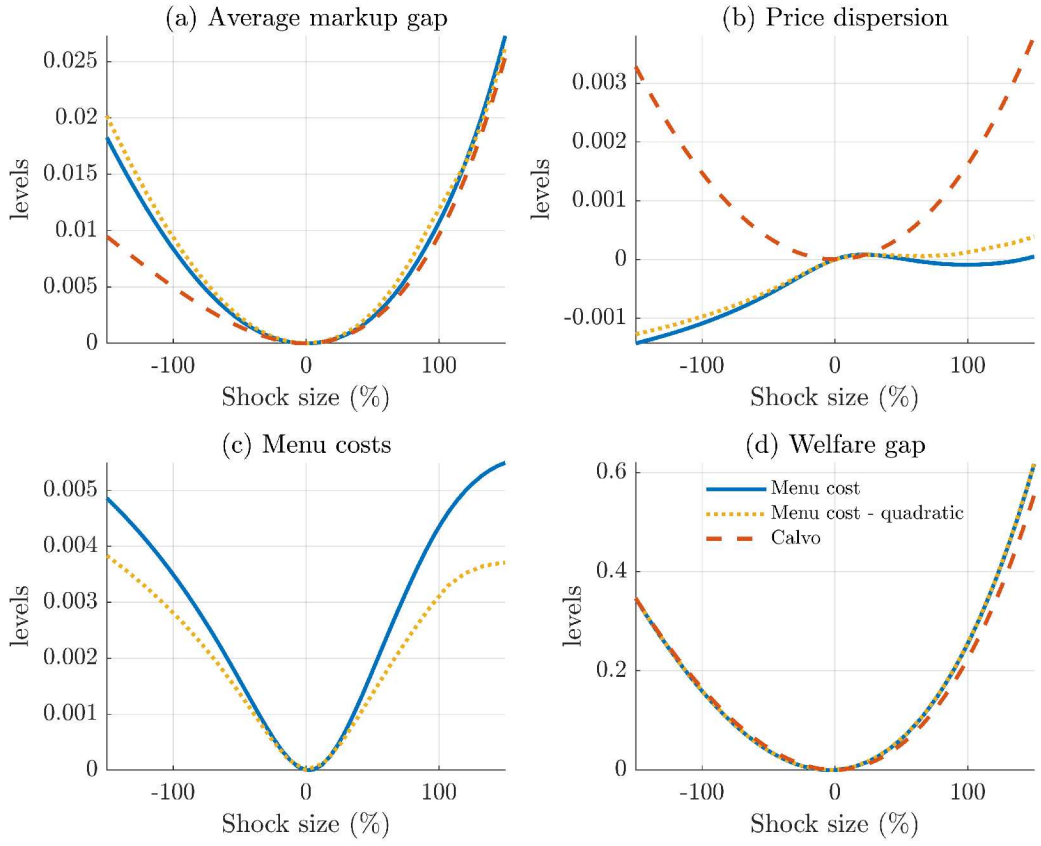


Figure 7: Welfare decomposition of a cost-push shock for different shock magnitudes.
Note: The figure displays the difference in the different components of the welfare gap in equation (25). Panels a, b, c report impact effects, panel d reports the cumulative discounted welfare gap.

This is due to the selection effect. Quantitatively, price dispersion plays a minor role in the welfare gap. Instead, in the menu cost model the price adjustment costs are important. Like price dispersion in Calvo, they increase with shock size.

Alternative central bank objective. While the previous explanation addresses the motivation of the central banker, it does not answer the question of why optimal policy in the menu cost model displays such a strong degree of nonlinearity. To answer that question, and following the previous discussion, we consider a purely quadratic objective

$$-\sum_{t=0}^{\infty}\beta^t\frac{1}{2}\left[\mathbb{E}\left(\pi_t-\pi_{ss}\right)^2+\frac{\kappa}{\epsilon}\mathbb{E}(\tilde{y}_t^e)^2\right]. \tag{26}$$

In the Calvo model, this expression is exact up to a second-order approximation: The average markup gap equals the square of the output gap up to a second-order approximation.

The price dispersion gap depends on the square of inflation, up to second order. In the menu cost model, we impose that quadratic objective. In this case, the first term is still an exact second-order approximation of the welfare cost of the average markup. The second term serves as an approximation of the relationship between the inflation and the sum of price dispersion and menu costs. This makes sense for the following reason. In the menu cost model, the menu costs depend linearly on frequency, which we found to depend on inflation in a u-shaped fashion, as discussed in Section 4. Price dispersion is less important. Thus we can approximate the sum of those two gaps by a quadratic function of inflation.

The dotted yellow line in Figure 5 above displays the optimal response in the menu cost model as a function of shock size in the case of a quadratic objective. The result is strikingly similar to the result in the baseline case. The same holds for the welfare components and cumulative welfare shown by Figure 7, the nonlinear target rule or the impulse responses (not shown). Thus, the quadratic objective turns out to be a very reasonable approximation to the true nonlinear objective function even for large shocks. This is so even though the π_t^2 term is not a complete second-order approximation to the two latter terms.[14] As far as the objective function is concerned, the central bank thus faces a very similar trade-off in the Calvo and in the menu cost model.

This result furthermore confirms that the main source of the nonlinearity in the menu cost model is not to be found in the exact nonlinear shape of the welfare function, but in the nonlinearity of the Phillips curve relationship.

There is a different way to interpret this exercise. Consider the Calvo model, and approximate the nonlinear objective by the quadratic objective above. This approximation has basically no effects in Calvo (not shown), so that the red dashed line in Figure 5 remains a valid description of optimal policy. Now replace the near-linear Phillips Curve in the Calvo model by the nonlinear Phillips curve relationship from the menu cost model. This yields the yellow dotted line. Thus, inserting a nonlinear Phillip Curve explains the nonlinearity of optimal policy.

Intuition based on a simple model with a nonlinear Phillips curve. In the linearized Calvo model, with a quadratic objective and a linear Phillips curve, the target

[14] This approximation neglects state dependencies. However they seem to play a quantitatively minor role.

rule takes the shape (Galí, 2008)

$$\pi_t = -\frac{1}{\epsilon}\left(\tilde{y}_t^e - \tilde{y}_{t-1}^e\right) \tag{27}$$

Now consider a small instructive variation of the linear-quadratic textbook model, meant to resemble key features of the state-dependent model documented before. Assume that the Phillips curve has an inverse S-shape, in the spirit of what we have shown in Section 4, and is given by:[15]

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa \tilde{y}_t^e + \kappa_3 (\tilde{y}_t^e)^3$$

with constants $\kappa, \kappa_3 > 0$. Furthermore assume that the objective (26) remains the same. We have shown in the previous paragraphs that this objective approximates the objective in the menu cost model well. We can again derive a simple target rule:

$$\pi_t + \frac{\frac{\kappa}{\epsilon}\tilde{y}_t^e}{\kappa + 3\kappa_3(\tilde{y}_t^e)^2} - \frac{\frac{\kappa}{\epsilon}\tilde{y}_{t-1}^e}{\kappa + 3\kappa_3(\tilde{y}_{t-1}^e)^2} = 0$$

Which, for $t = 1$, reads

$$\pi_t = -\frac{\frac{\kappa}{\epsilon}\tilde{y}_t^e}{\kappa + 3\kappa_3(\tilde{y}_t^e)^2}$$

This equation has the same qualitative features as the target rule in Figure 6. It has a slope of $-\frac{\kappa/\epsilon}{\kappa} = -\frac{1}{\epsilon}$ at 0 and flattens out as the output gap $\tilde{y}_1^e$ deviates from 0 either direction. Policy thus also strikes while the iron is hot in this simple model. It does so because the sacrifice ratio drops as larger shocks push the economy along the nonlinear Phillips curve.

Let's get back to the standard Calvo model again for one final piece of intuition. Surprisingly, perhaps, the target rule (27) is independent of the frequency of price changes, given by the exogenous Calvo parameter θ. The reason behind this result is that two opposing effects perfectly offset each other. On the one hand, frequency raises the slope of the Phillips curve $\kappa = (1-\theta)(1-\beta\theta)/\theta$, and therefore reduces the sacrifice ratio making the cost of reducing inflation lower. On the other hand, higher frequency reduces the relative weight of inflation in the welfare function, given by ϵ/κ, because it reduces the relative price distortions caused by inflation. The latter does evidently not happen in our ad-hoc model because we keep the

[15] This simplified Phillips curve features no inversion. This contrasts with the menu cost model. Nonetheless, the analysis in this part is based on shocks small enough not to make the inverted part of the Phillips curve relevant.

objective constant by assumption. And it does not happen in the menu cost model either, because the predominant cost of inflation is due to the adjustment costs, which, unlike price dispersion, don't get less bad (get worse) when the frequency increases.

Taken together these considerations are consistent the numerical finding that the nonlinearity of the Phillips Curve explains the nonlinear strike-while-the-iron-is-hot policy.

6.3 Timeless optimal monetary response to TFP shocks

Next, we consider TFP shocks, which, unlike before, affect the efficient allocation. In the standard New Keynesian model with Calvo prices, the response to such shocks is characterized by strict price stability: the central bank steers real interest rates to replicate the path of natural interest rates, which leads to inflation and the output gap remaining at zero. This is commonly known as the "*divine coincidence*" (Blanchard and Galí, 2007).

A version of the divine coincidence also holds in our economy. As we have shown in Section 6.1, the Ramsey plan features a positive level of trend inflation in the long run. In response to an TFP shock, optimal timeless commitment policy keeps inflation at its steady-state level. We prove this formally in Appendix C. As inflation remains constant, the frequency of repricing also stays constant. In other words, the optimal policy offsets the dynamic impact of the efficient shocks in a form of "dynamic divine coincidence".

The conclusion is that strict targeting of the optimal steady-state inflation rate simultaneously minimizes inefficient output fluctuations and the costs of nominal rigidities. Notice that the shape of the Phillips curve plays no role in this result and thus the prescription is the same for small and large shocks.

6.4 Time-0 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 faces no previous pre-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-0 problem" (Woodford, 2003).

The solid blue lines in Figure 8 show the time path under the optimal policy. The labor

subsidy is set to zero in this exercise, which, therefore, ceases to offset any markup distortions caused by the firms' market power. The optimal policy is time-inconsistent: without pre-commitment, the central bank engineers a temporary expansion. Thereby, 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 8 shows the equivalent time-0 response in the Calvo model. The figure shows that the incentive to surprise is substantially weaker 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 hike, which causes an increase in the frequency of price changes. As a result, the output gap increases by less than it would have with exogenous frequency. That is, the output boost from a given amount of inflation is lower than under Calvo. Since the planners objective function isn't significantly different than under Calvo (as we saw before), the social planner thus eases less aggressively.

There is a countervailing force that raises the time inconsistency in our baseline model relative to the Calvo model. Namely, the markup distortions are higher, as discussed below, and a labor subsidy of $\tau = 1/\epsilon$ is insufficient to bring the distortions caused by the average markup to zero, as it is the case in the Calvo model. A time-0 optimal policy, therefore, stays time inconsistent even with a $\tau = 1/\epsilon$ labor subsidy (not shown). The optimal policy easing in this scenario, however, is small, two orders of magnitude smaller than those under no labor subsidy. Therefore, this channel is too weak to counteract the opposite effect caused by the more flexible price level detailed in the previous paragraphs.

A corollary to the negligibility of the time inconsistency with an appropriate labor subsidy is that the analysis in the previous sections, where we adopted a timeless perspective, would go through without any quantitatively relevant changes also if we adopted a time-0 perspective.

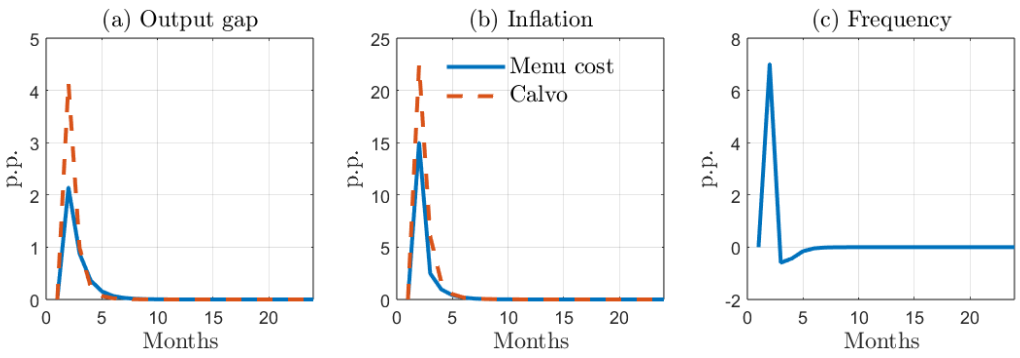


Figure 8: Time inconsistency of the optimal policy.

Note: The figure compares the time-0 optimal policies in the menu cost model and in the Calvo model.

7 Robustness

In this section, we show the robustness of the nonlinear Phillips curve and the nonlinear target rule to some parameter choices.

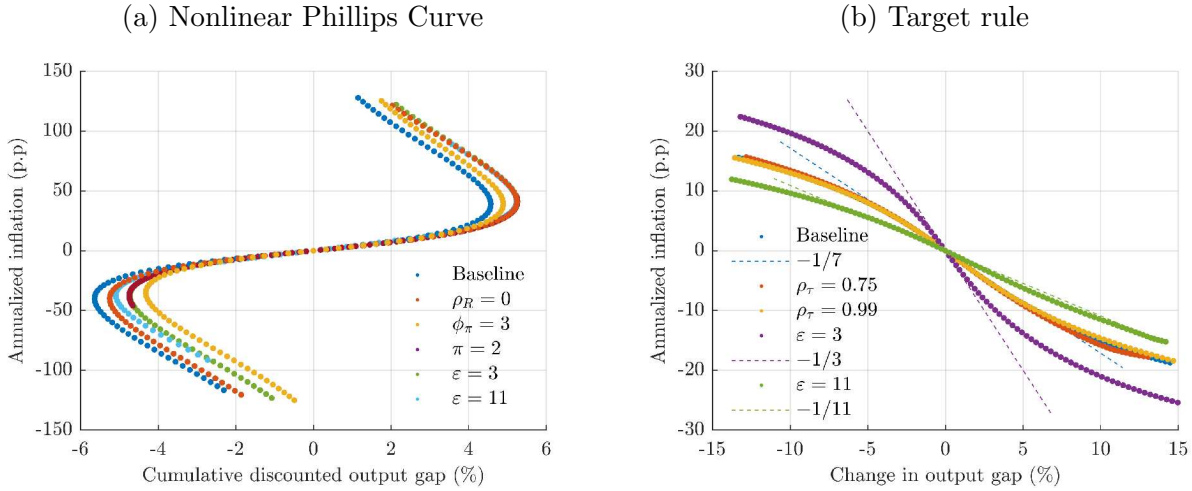


Figure 9: Robustness of the nonlinear Phillips curve and the target rule

Note: The figures recreate the nonlinear Phillips curve (a) and the target rule (b) for alternative parameter values. The Phillips curve shows robustness to alternative Taylor rule coefficients: no smoothing ($\rho_R = 0$), more aggressive anti-inflationary stance ($\phi_\pi = 3$), 2 percent inflation target ($\pi^* = 2\%$). And both show robustness to alternative elasticity of substitution parameters $\epsilon = 3, 11$.

Panel (a) of Figure 7 shows the robustness of the nonlinear Phillips curve. The figure reports the relationship between the impact effect of annualized inflation and the cumulative discounted output gap for different i.i.d. monetary policy shocks of varying sizes as Figure 1. It reports how the relationship changes for varying Taylor rule parameters: no smoothing ($\rho_R = 0$), more aggressive anti-inflationary stance ($\phi_\pi = 3$), 2 percent inflation target

($\pi^* = 2\%$). Additionally, it shows the relationship for varying the elasticity of substitution parameter $\epsilon = 3, 11$[16]. The figure shows that the relationship is not structural in the sense that it is not independent of the policy rule, but it is quantitatively robust in the sense that modifying the Taylor rule parameters and the elasticity of substitution parameters modify the figure only slightly. It's qualitative features concerning size-dependence and asymmetry described in Section D.1 stay unchanged.

Panel (b) of Figure 9 contrasts the relationship between annualized impact inflation and the change in the output gap on impact under Ramsey optimal policy in the baseline with alternatives with varying degree of persistence of the cost-push shock ($\rho_\tau = 0.75, 0.99$); with alternative values of the elasticity of substitution parameter ($\epsilon = 3, 11$). It also shows linear lines with slope $-1/\epsilon$ for $\epsilon = 3, 7, 11$. The figure shows that (i) the target rule is influenced by the persistence of the underlying shock, but the variation is quantitatively small. Furthermore, (ii) the elasticity of substitution plays a key role in determining the slope of the target rule just as in the Calvo model. Lastly, (iii) the qualitative features of the nonlinearity are robust: it is optimal to strike-it-while-it's-hot for a wide range of parameter values.

8 Implications for the 2022-2023 US inflation surge

Scenario description. We use our model to assess its optimal policy prescriptions in a stylized scenario akin to the 2022-2023 post-Covid US inflation surge. Besides the large increase in repricing frequency, a further notable feature of the inflationary episode was the significant *increase* in the dispersion of price changes (Montag and Villar, 2023). This is relevant because large aggregate shocks with uniform impact on firms' optimal reset prices, including demand, supply, and cost-push shocks in our framework, would *decrease* the price-change dispersion. The reason is that the new price changes, resulting from price gaps pushed over one of the inaction thresholds by the shock, are necessarily closer to the mean of price changes than the canceled price adjustments, which are around the other inaction threshold. The contradictory evidence, therefore, suggests that concurrent forces raised the

[16] We recalibrate the menu cost and the idiosyncratic quality shock volatility such that the frequency and the size of price changes stay constant across calibrations

dispersion of reset prices, like an increase in the volatility of idiosyncratic demand or supply shocks.[17] In this section, we construct a scenario as a combination of a cost-push shock, a dispersion shock, and a monetary policy that follows an inertial Taylor rule. The scenario broadly captures key features of the 2022-2023 US inflation surge. Then we use our model to characterize the counterfactual optimal Ramsey commitment policy.

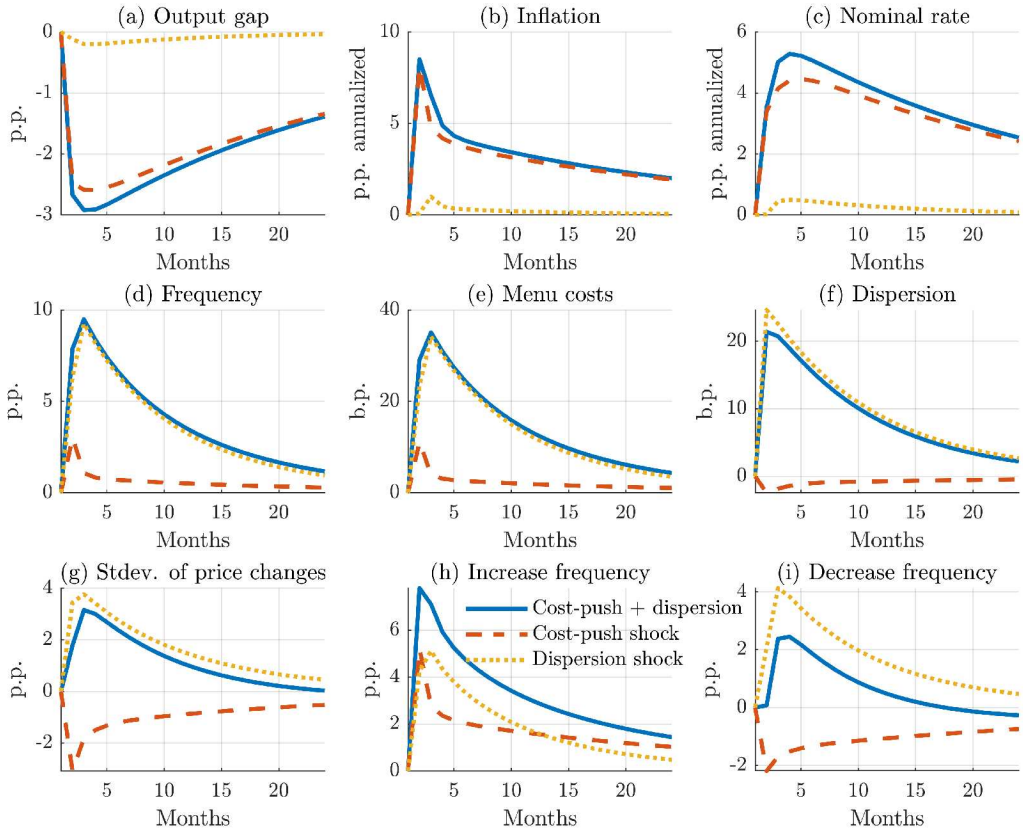


Figure 10: Inflation surge caused by a cost-push and dispersion shocks.
Note: The figure shows a scenario (blue solid line) as a response to a combination of a large cost push τ_t and a dispersion σ_t shock, which captures some key features of the 2022-2023 inflation surge. The figure also shows the individual responses to the cost-push shock (red dashed line) and the dispersion shock (yellow dotted line) components.

Figure 10 shows the impulse responses to a combination of a cost-push and a dispersion shock, alongside the impulse responses to the two shock components. The two shock broadly capture both the large increase in the frequency of price changes (which increases by 9.5 p.p. in the model versus 11 p.p. in the data), and the increase in the standard deviation of the

[17] Similar idiosyncratic volatility shock and its interaction with aggregate shocks is analyzed in Vavra (2014). Berger and Vavra (2019) explore the possibility of increased dispersion caused by heterogeneous response to aggregate shocks.

price changes (which increases by around 3 p.p. in the model versus 1.5 percentage points during the 2022-2023 inflation surge (Montag and Villar, 2023)).

The scenario captures broad features of the 2022-2023 US inflation surge. The shocks raise the inflation rate from the steady-state 2 percent to around 10.5 percent.[18] Around half of the inflation surge is temporary and recedes quickly, while the remaining half is persistent and declines slowly. The nominal interest rate increases quickly by slightly over 5 percent, broadly in line with the magnitude of the Fed's 2022 policy tightening. The output gap declines by 3 percentage points.[19] The figure also shows that the scenario captures the sizable and persistent increase in the frequency of price changes.

The figure also confirms that the standard deviation of price changes prove the necessity of a dispersion shock accompanying the aggregate shock and helps identify its magnitude. The standard deviation of price changes declines as a response to the cost-push shock, while it increases significantly after the dispersion shock. Their combination in the baseline leads to an overall increase as in the data. Notably, the exercise also captures the around 2 percentage-point *increase* in the frequency of price decreases, even though it is not a moment we explicitly target. An aggregate shock alone would have reduced the frequency of price decreases, while a sole dispersion shock would have increased it by more than 4 percentage points. All in all, the large increase in the overall frequency is driven mostly by the increase in price increases reaching around 14% from around 6%, similar to the data (Montag and Villar, 2023).

How would optimal Ramsey policy have reacted to such a combination of shocks? The answer is not obvious, as the dispersion shock generates variation in optimal relative prices and, consequently, a welfare-enhancing increase in the frequency of price changes that monetary policy might not want to counteract. Figure 11 shows the counterfactual optimal policy response. It shows that the optimal policy would prescribe a substantially tighter policy in

[18] We have reduced the interest-rate smoothing parameter of the inertial Taylor rule from 0.75 quarterly rate to 0.33 quarterly rate to capture the unusually fast pace of interest-rate increase observed in 2022.

[19] The decline in output is contrary to the evidence. A more realistic scenario should also include a positive demand shock raising output, caused, for example, by heightened consumption demand from excess savings built up during the Covid-19 lock-downs and from fiscal transfers. Excluding such a demand shock is a conservative choice: a scenario with demand shock would prescribe an even more aggressive optimal policy tightening than our baseline. The reason is that monetary policy would optimally fully offset an efficient demand shock, while it leans only partially against the inefficient cost-push shock in our scenario.

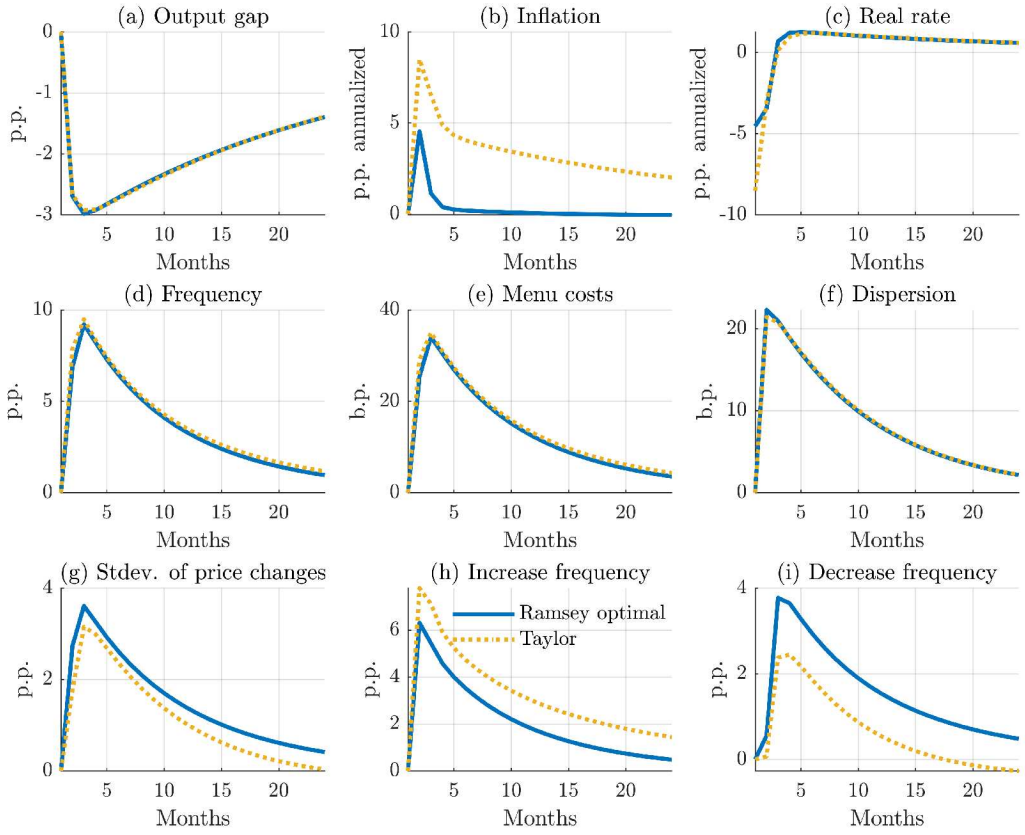


Figure 11: Optimal monetary policy response to a joint cost-push and dispersion shock.
Note: The figure shows the optimal Ramsey commitment response (blue solid line) to a combination of a large cost push τ_t and a dispersion σ_t shocks. It contrasts it to the policy following an inertial Taylor rule (yellow dotted line).

the initial periods than the Taylor rule. The optimal policy would have led to an increase in inflation by only around 4 percentage points at the cost of a somewhat lower output gap. This is substantially lower than the 8.5 percentage points increase in inflation under the Taylor rule. The policy would have reduced the frequency of price changes only marginally: it would have stayed persistently high. The low optimal inflation, furthermore, would be accompanied by a substantial shift from price increases towards price decreases, which are not costlier than price increases in our framework.[20]

[20]There is no reason for downward nominal rigidity in our framework, up to first order. Price increases are more frequent than price decreases even under zero inflation only due to the asymmetry of the profit function, a higher-order feature. It raises the incentives of firms to adjust their relative prices if it is below the optimal reset price than when it is above it.

9 Conclusion

This paper analyzes the Ramsey problem within a menu cost model à la Golosov and Lucas (2007). A key contribution is the identification of a new incentive for the central bank: In the presence of large *cost-push* shocks, optimal monetary policy should commit to quelling inflation more aggressively, relative to small shocks, and to stabilizing the repricing frequency. Along the trajectory of optimal commitment, the central bank utilizes the reduction in the sacrifice ratio, leading to lower inflation at a cost of a slightly more marked decline in output. Optimal policy thus *strikes while the iron is hot.* This policy prescription diverges markedly from that of the standard New Keynesian model with exogenous timing of price adjustment, which fails to capture such nonlinear dynamics. When confronted with aggregate demand or TFP shocks, our findings indicate that the optimal policy in the menu cost model involves a commitment to full price stability, akin to the standard Calvo model with exogenous timing of price changes.

In sum, our research underscores the importance of an aggressive anti-inflationary policy by the central bank in the face of large shocks. By committing to policies that curb inflation and stabilize the repricing frequency, the central bank can deliver a more favorable macroeconomic outcome. Our analysis is confined to the case of nominal price rigidities in the seminal fixed menu cost price setting model; we leave for future research the interaction with wage rigidities and assessment of optimal policy in more complex price-setting frameworks, which can match features of micro price setting better.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2021) "Income and wealth distribution in macroeconomics: A continuous-time approach," *The Review of Economic Studies*, Vol. 89, pp. 45–86.

Adam, Klaus and Henning Weber (2019) "Optimal Trend Inflation," *American Economic Review*, Vol. 109, pp. 702–737.

Adjemian, Stéphane, Houtan Bastani, Michel Juillard, Frédéric Karamé, Ferhat Mihoubi, Willi Mutschler, Johannes Pfeifer, Marco Ratto, Sébastien Villemot, and Normann Rion (2023) "Dynare: Reference Manual Version 5," working papers, HAL.

Alexandrov, Andrey (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, Fernando, Martin Beraja, Martín Gonzalez-Rozada, and Pablo Andrés Neumeyer (2019) "From Hyperinflation to Stable Prices: Argentina's Evidence on Menu Cost Models," *The Quarterly Journal of Economics*, Vol. 134, pp. 451–505.

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi (2016) "The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach," *American Economic Review*, Vol. 106, pp. 2817–51.

Alvarez, Fernando, Francesco Lippi, and Aleksei Oskolkov (2021) "The Macroeconomics of Sticky Prices with Generalized Hazard Functions*," *The Quarterly Journal of Economics*, Vol. 137, pp. 989–1038.

Alvarez, Fernando and Pablo Andres Neumeyer (2020) "The Passthrough of Large Cost Shocks in an Inflationary Economy," in Gonzalo Castex, Jordi Galí, and Diego Saravia eds. *Changing Inflation Dynamics,Evolving Monetary Policy*, Vol. 27 of Central Banking, Analysis, and Economic Policies Book Series: Central Bank of Chile, Chap. 2, pp. 007–048.

Auclert, Adrien, Rodolfo Rigato, Matthew Rognlie, and Ludwig Straub (2024) "New Pricing Models, Same Old Phillips Curves?" *The Quarterly Journal of Economics*, Vol. 139, pp. 121–186.

Auer, Raphael, Ariel Burstein, and Sarah M Lein (2021) "Exchange Rates and Prices: Evidence from the 2015 Swiss Franc Appreciation," *American Economic Review*, Vol. 111, pp. 652–686.

Barro, Robert J. (1972) "A Theory of Monopolistic Price Adjustment," *Review of Economic Studies*, Vol. 39, pp. 17–26.

Benigno, Pierpaolo and Gauti 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.

Berger, David and Joseph Vavra (2019) "Shocks versus Responsiveness: What Drives Time-Varying Dispersion?" *Journal of Political Economy*, Vol. 127, pp. 2104–2142.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J Sargent (2021) "Inequality, Business Cycles, and Monetary-Fiscal Policy," *Econometrica*, Vol. 89, pp. 2559–2599.

Blanchard, Olivier and Jordi Galí (2007) "The Macroeconomic Effects of Oil Price Shocks: Why Are the 2000s so Different from the 1970s?" in *International Dimensions of Monetary Policy*: National Bureau of Economic Research, Inc, pp. 373–421.

Blanco, Andres, Corina Boar, Callum J Jones, and Virgiliu Midrigan (2024a) "Nonlinear Inflation Dynamics in Menu Cost Economies," NBER Working Papers 32094, National Bureau of Economic Research.

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

Blanco, Andrés, Corina Boar, Callum J. Jones, and Virgiliu Midrigan (2024b) "The Inflation Accelerator," NBER Working Papers 32531, National Bureau of Economic Research, Inc.

Boppart, Timo, Per Krusell, and Kurt Mitman (2018) "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative," *Journal of Economic Dynamics and Control*, Vol. 89, pp. 68–92.

Caballero, Ricardo and Eduardo Engel (1993) "Heterogeneity and Output Fluctuations in a Dynamic Menu-Cost Economy," *The Review of Economic Studies*, Vol. 60, pp. 95–119.

Caballero, Ricardo J and Eduardo MRA Engel (2007) "Price Stickiness in Ss models: New Interpretations of Old Results," *Journal of Monetary Economics*, Vol. 54, pp. 100–121.

Calvo, Guillermo A. (1983) "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, Vol. 12, pp. 383 – 398.

Caplin, Andrew S. and Daniel F. Spulber (1987) "Menu Costs and the Neutrality of Money," *The Quarterly Journal of Economics*, Vol. 102, pp. 703–726.

Caratelli, Daniele and Basil Halperin (2023) "Optimal monetary policy under menu costs," unpublished manuscript.

Cavallo, Alberto, Francesco Lippi, and Ken Miyahara (2023) "Large Shocks Travel Fast," NBER Working Papers 31659, National Bureau of Economic Research, Inc.

Cerrato, Andrea and Giulia Gitti (2023) "Inflation Since COVID: Demand or Supply," unpublished manuscript.

Costain, James and Anton Nakov (2011) "Distributional Dynamics under Smoothly State-Dependent Pricing," *Journal of Monetary Economics*, Vol. 58, pp. 646 – 665.

——— (2019) "Logit Price Dynamics," *Journal of Money, Credit and Banking*, Vol. 51, pp. 43–78.

Dávila, Eduardo and Andreas Schaab (2022) "Optimal Monetary Policy with Heterogeneous Agents: A Timeless Ramsey Approach," *Working Paper*.

Dotsey, Michael, Robert G. King, and Alexander L. Wolman (1999) "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *The Quarterly Journal of Economics*, Vol. 114, pp. 655–690.

Gagnon, Etienne (2009) "Price Setting during Low and High Inflation: Evidence from Mexico," *The Quarterly Journal of Economics*, Vol. 124, pp. 1221–1263.

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

Golosov, Mikhail and Robert E. Lucas (2007) "Menu Costs and Phillips Curves," *Journal of Political Economy*, Vol. 115, pp. 171–199.

González, Beatriz, Galo Nuño, Dominik Thaler, and Silvia Albrizio (2024) "Firm Heterogeneity, Capital Misallocation and Optimal Monetary Policy," Working Paper Series 2890, European Central Bank.

Karadi, Peter and Adam Reiff (2019) "Menu Costs, Aggregate Fluctuations, and Large Shocks," *American Economic Journal: Macroeconomics*, Vol. 11, pp. 111–146.

Le Grand, Francois, Alais Martin-Baillon, and Xavier Ragot (2022) "Should monetary policy care about redistribution? Optimal monetary and fiscal policy with heterogeneous agents,"Technical report, Paris School of Economics.

Midrigan, Virgiliu (2011) "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *Econometrica*, Vol. 79, pp. 1139–1180.

Montag, Hugh and Daniel Villar (2023) "Price-Setting During the Covid Era," *FEDS Notes*.

Nakamura, Emi and Jón Steinsson (2008) "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, Vol. 123, pp. 1415–1464.

——— (2010) "Monetary Non-neutrality in a Multisector Menu Cost Model," *The Quarterly Journal of Economics*, Vol. 125, pp. 961–1013.

Nakov, Anton and Carlos Thomas (2014) "Optimal Monetary Policy with State-Dependent Pricing," *International Journal of Central Banking*, Vol. 36.

Nuño, Galo and Carlos Thomas (2022) "Optimal Redistributive Inflation," *Annals of Economics and Statistics*, pp. 3–63.

Ragot, Xavier (2019) "Managing Inequality over the Business Cycles: Optimal Policies with Heterogeneous Agents and Aggregate Shocks," 2019 Meeting Papers 1090, Society for Economic Dynamics.

Sheshinski, Eytan and Yoram Weiss (1977) "Inflation and Costs of Price Adjustment," *Review of Economic Studies*, Vol. 44, pp. 287–303.

Smets, Frank and Rafael Wouters (2007) "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *The American Economic Review*, Vol. 97, pp. 586–606.

Smirnov, Danila (2022) "Optimal Monetary Policy in HANK," Unpublished manuscript, Universitat Pompeu Fabra.

Taylor, John B (1993) "Discretion versus Policy Rules in Practice," in *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 195–214, Elsevier.

Vavra, Joseph (2014) "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation," *The Quarterly Journal of Economics*, Vol. 129, pp. 215–258.

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

——— (2009) "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics*, Vol. 56, pp. S100–S124.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen (2004) "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *The Review of Economics and Statistics*, Vol. 86, pp. 514–533.

ONLINE APPENDIX

A Optimality condition of the reset price

If the value function is convex, the optimal reset price is fully characterized by the system of first order conditions in Section 2.6.[21] This Appendix presents the derivation of $V_t'(0)$. To start, we reproduce here the value function presented in equation (15):

$$\begin{aligned} V_t(x) &= \Pi(x, p_t^*, w_t, A_t) \\ &\quad +\mathbb{E}_t\left[\left(1-\lambda_{t+1}\left(x-\sigma_{t+1}\varepsilon_{t+1}-\pi_{t+1}^*\right)\right)\Lambda_{t,t+1}V_{t+1}(x-\sigma_{t+1}\varepsilon_{t+1}-\pi_{t+1}^*)\right] \\ &\quad +\mathbb{E}_t\left[\lambda_{t+1}\left(x-\sigma_{t+1}\varepsilon_{t+1}-\pi_{t+1}^*\right)\Lambda_{t,t+1}\left(V_{t+1}\left(0\right)-\eta w_{t+1}\right)\right]. \end{aligned}$$

Since the only source of uncertainty is the idiosyncratic shocks, $V_t(x)$ becomes

$$\begin{aligned} V_t(x) &= \Pi(x, p_t^*, w_t, A_t) \\ &\quad +\Lambda_{t,t+1}\int\left[\left(1-\lambda_{t+1}\left(x-\sigma_{t+1}\varepsilon-\pi_{t+1}^*\right)\right)V_{t+1}(x-\sigma_{t+1}\varepsilon-\pi_{t+1}^*)\phi(\varepsilon)\right]d\varepsilon \\ &\quad +\Lambda_{t,t+1}\left(V_{t+1}\left(0\right)-\eta w_{t+1}\right)\int\left[\lambda_{t+1}\left(x-\sigma_{t+1}\varepsilon-\pi_{t+1}^*\right)\phi(\varepsilon)\right]d\varepsilon, \end{aligned}$$

where $\phi\left(\cdot\right)$ denotes the standard normal p.d.f. The term $x-\sigma_{t+1}\varepsilon-\pi_{t+1}^*$ is the state of the firm at $t+1$, conditional on the state x at t and the realization $\varepsilon_{t+1}=\varepsilon$ of the shock at $t+1$. Denoting the state at $t+1$ as $x'\equiv x-\sigma_{t+1}\varepsilon-\pi_{t+1}^*$, such that $\varepsilon\equiv\frac{x-x'-\pi_{t+1}^*}{\sigma_{t+1}}$, and applying the correspoding change of variable to the integral yields

$$\begin{aligned} V_t(x) &= \Pi(x, p_t^*, w_t, A_t)+\frac{\Lambda_{t,t+1}}{\sigma_{t+1}}\int\left[\left(1-\lambda_{t+1}\left(x'\right)\right)V_{t+1}(x')\phi\left(\frac{x-x'-\pi_{t+1}^*}{\sigma_{t+1}}\right)\right]dx' \\ &\quad +\left(V_{t+1}\left(0\right)-\eta w_{t+1}\right)\frac{\Lambda_{t,t+1}}{\sigma_{t+1}}\int\left[\lambda_{t+1}\left(x'\right)\phi\left(\frac{x-x'-\pi_{t+1}^*}{\sigma_{t+1}}\right)\right]dx'. \end{aligned}$$

The probability of updating a price λ_{t+1} given any state of the nature is either 0 in the "inaction region" or 1 otherwise. Defining the inaction region as the (s_t, S_t) band, we restrict the first integral in the latter expression. We also replace the second integral, which is the probability mass of updating the price, by 1 minus the probability mass of not updating the

[21] We verify convexity ex post.

price. Thus:

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

where $\Phi(\cdot)$ denotes the standard normal c.d.f. and, to simplify notation, we suppress the dependence of Π on aggregate variables.

Finally, taking the derivative of $V_t(x)$ with respect to x and reformulating, we get $V_t'(x)$:

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

which must be evaluated at $x = 0$.

B Efficiency and welfare analysis

B.1 Efficient and natural level of output

This Appendix derives efficient output, efficient real interest rate and natural output.

Efficient output. We obtain it as the solution to a social planning problem. The problem maximizes household welfare in equation (1) subject to (i) the aggregate consumption equation (3), (ii) aggregate labor supply in ($N_t = \int_i N_t(j)$) and (iii) product-level production functions in (9) 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

$$\max_{N_t(j)} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\left[A_t \left(\int N_t(j)^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \right]^{1-\gamma}}{1-\gamma} - \upsilon \int N_t(j) di,$$

subject to $\int N_t(j)di = N_t$. The solution implies that the efficient output fluctuates with aggregate productivity, but is independent of demand shocks as well as of cost-push shocks. In particular, the efficient level of output is

$$Y_t^e = C_t^e = A_t N_t^e = \upsilon^{-1/\gamma} A_t^{1/\gamma}. \tag{28}$$

For our parametrization, $\upsilon = 1$ and $\gamma = 1$, we thus have that

$$\begin{aligned} N_t^e =& 1, \\ C_t^e =& A_t. \end{aligned} \tag{29}$$

The efficient labor supply is equal across products and the efficient product-level consumption varies across products j inversely proportional to the product-level quality, in particular

$$\begin{aligned} N_t^e(j) =& N_t^e \\ C_t^e(j) =& \frac{A_t N_t^e}{A_t(j)}. \end{aligned}$$

Efficient real interest rate. It is implicitly defined by the Euler equation after substituting in efficient consumption:

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

Natural output. It is defined as the counterfactual output with flexible prices. Under flexible prices, firms maximize their real profit function (11) in each period t by choosing

$$\frac{P_t^n(j)}{A_t(j)P_t^n} = \frac{\epsilon}{\epsilon-1}(1-\tau_t)\frac{w_t^n}{A_t}.$$

The expression implies that the quality-adjusted relative price is homogeneous across products j. Together with the definition of the price-level in equation (5) in the main text, this implies that the natural level of the quality-adjusted log relative price is zero ($p_t(j) = 0$). Or equivalently, the natural level of relative price is equal to the quality: $P_t^n(j)/P_t^n = A_t(j)$. The product-level demand function and the unit quality-adjusted relative price implies that product-level natural consumption is inversely proportional to the quality of product j:

$$C_t^n(j) = \frac{1}{A_t(j)} C_t^n.$$

Furthermore, the natural real wage, output and labor are given by the following closed-form

expressions:

$$\begin{aligned} w_t^n =& A_t \frac{\epsilon-1}{\epsilon}\frac{1}{1-\tau_t}, \\ Y_t^n =& C_t^n = \left(\frac{w_t^n}{\upsilon}\right)^{1/\gamma}, \\ N_t^n =& \frac{Y_t^n}{A_t}. \end{aligned}$$

Notably, the productivity shock affects the natural and the efficient output similarly, but the cost-push (labor-tax) shocks only affect the natural level of output.

B.2 Welfare decomposition

This Appendix derives the welfare decomposition presented in equation (25) in the main text. We start by obtaining expressions that are used at the end for the welfare decomposition.

Markups. The real marginal cost of firm j is

$$MC_t(j) = \frac{\partial\left((1-\tau_t)w_t N_t(j)\right)}{\partial Y_t(j)} = \frac{(1-\tau_t)w_t A_t(j)}{A_t},$$

where we have used that $N_t(j) = A_t(j)Y_t(j)/A_t$.

The (log-) markup $\mu_t(j)$ is the (log-) difference between the relative price and the real marginal cost:

$$\mu_t(j) = \log\frac{P_t(j)}{P_t} - \log\frac{(1-\tau_t)w_t A_t(j)}{A_t} = \log\frac{P_t(j)}{A_t(j)P_t} - \log\frac{(1-\tau_t)w_t}{A_t} = p_t(j) - mc_t, \quad (30)$$

where $p_t(j)$ is the quality-adjusted relative price and

$$mc_t \equiv \log\left(MC_t(j)/A_t(j)\right) = \log\left((1-\tau_t)w_t/A_t\right) = \log(1-\tau_t) - \log A_t + \log\upsilon + \gamma\log Y_t,$$

is the 'aggregate component' of the marginal cost. Notice that we have employed eq. (6).

Aggregate markup and output. The aggregate (log-) markup $\overline{\mu}_t$ is

$$\overline{\mu}_t = \log\left(\int e^{\mu_t(j)(1-\epsilon)}dj\right)^{\frac{1}{1-\epsilon}} = \log\left(\int \frac{e^{p_t(j)(1-\epsilon)}}{e^{mc_t(1-\epsilon)}}dj\right)^{\frac{1}{1-\epsilon}} = \log\frac{1}{e^{mc_t}}\left(\int e^{p_t(j)(1-\epsilon)}dj\right)^{\frac{1}{1-\epsilon}} = -mc_t,$$

where we used the observation that the average quality-adjusted relative price is one (eq. 13). Therefore

$$\overline{\mu}_t = -\log(1-\tau_t) + \log A_t - \log\upsilon - \gamma\log Y_t \quad (31)$$

or equivalently

$$e^{\overline{\mu}_t} = \frac{A_t}{\upsilon(1-\tau_t)Y_t^{\gamma}} \quad (32)$$

expressing the tight relationship between average markup and the output.

Taking into account eq. (28), the efficient output gap can be expressed

$$\log Y_t - \log Y_t^e = \frac{1}{\gamma}\left(-\log(1-\tau_t) - \overline{\mu}_t\right), \tag{33}$$

which is proportional to the negative average markup.

Markup dispersion. The dispersion of the quality-adjusted relative prices (ζ_t^p) is

$$\zeta_t^p = \int e^{p_t(j)(-\epsilon)} g(p_t(j))dj = \int e^{(\mu_t(j)+mc_t)(-\epsilon)} g(\mu_t(j)+mc_t)dj = \int e^{(\mu_t(j)-\overline{\mu}_t)(-\epsilon)} g(\mu_t(j)-\overline{\mu}_t)dj \equiv \zeta_t^{\mu-\overline{\mu}}$$

where $\zeta_t^{\mu-\overline{\mu}}$ is the dispersion of the demeaned markups which equals price dispersion.

Welfare. Finally, consider the case with $\gamma = 1$. We can express the difference between welfare (W_t) from the welfare in the efficient equilibrium (W_t^e) subject to ashock in period 0 as

$$\begin{aligned}
W_0 - W_0^e = \sum_{t=0}^{\infty} \beta^t \left(U_t - U_t^e\right) &= \sum_{t=0}^{\infty} \beta^t \left((\log C_t - N_t) - (\log C_t^e - N_t^e)\right) \\
&= \sum_{t=0}^{\infty} \beta^t \left((\log Y_t - N_t) - (\log Y_t^e - 1)\right) \\
&= \sum_{t=0}^{\infty} \beta^t \left(-\log(1-\tau_t) - \overline{\mu}_t - \frac{C_t}{A_t}\int e^{p(-\epsilon)} g_t(p)\, dp - \eta g_t^0 + 1\right) \\
&= \sum_{t=0}^{\infty} \beta^t \left(-\log(1-\tau_t) - \overline{\mu}_t - \left(\frac{1}{e^{\overline{\mu}_t}(1-\tau_t)}\zeta_t^{\mu-\overline{\mu}} - 1\right) - \eta g_t^0\right) \\
= \sum_{t=0}^{\infty} \beta^t \left(-\log(1-\tau_t) - \overline{\mu}_t - \left(\frac{1}{e^{\overline{\mu}_t}(1-\tau_t)} - 1\right) - \left(\frac{1}{e^{\overline{\mu}_t}(1-\tau_t)}(\zeta_t^{\mu-\overline{\mu}} - 1)\right) - \eta g_t^0\right)
\end{aligned}$$

where U_t^e is the utility in the efficient equilibrium and where we have used (29) and $C_t = Y_t$ in line 1 and (33) and (14) in line 2 and (32) and $C_t = Y_t$ and $\gamma = 1$ in line 3. The final expression decomposes welfare into terms related to average markup, markup dispersion, and adjustment costs.

C Response to TFP shocks

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

The planner's problem is:

$$\max_{\substack{\{g_t^c(\cdot), g_t^0, V_t(\cdot), C_t, \\ w_t, p_t^*, s_t, S_t, \pi_t^*\}_{t=0}^{\infty}}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log C_t - \upsilon L_t\right)$$

subject to

$$\begin{aligned}
w_t &= \upsilon C_t, \\
L_t &= \frac{C_t}{A_t}\left(\int e^{(x+p_t^*)(-\epsilon_t)} g_t^c(p)\, dx + g_t^0 e^{(p_t^*)(-\epsilon)}\right) - \upsilon\eta g_t^0 \\
V_t(x) &= \Pi(x, p_t^*, w_t, A_t) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{s_t}^{S_t} \left[V_{t+1}(x')\phi\left(\frac{(x-x') - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right] dx' + \\
&\quad \Lambda_{t,t+1}\left(1 - \frac{1}{\sigma_{t+1}}\int_{s_t}^{S_t}\left[\phi\left(\frac{(x-x') - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right] dx'\right)\left[\left(V_{t+1}(0) - \eta w_{t+1}\right)\right], \\
V_t(s_t) &= V_t(0) - \eta w_t, \\
V_t(S_t) &= V_t(0) - \eta w_t, \\
0 &= \Pi_t'(0) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{s_{t+1}}^{S_{t+1}} V_{t+1}(x') \left.\frac{\partial\phi\left(\frac{x-x'-\pi_{t+1}^*}{\sigma_{t+1}}\right)}{\partial x}\right|_{x=0} dx' \\
&\quad + \Lambda_{t,t+1}\left(\phi\left(\frac{-S_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right)\left(V_{t+1}(0) - \eta w_{t+1}\right). \\
g_t^c(x) &= \frac{1}{\sigma_t}\int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1})\phi\left(\frac{x_{-1} - x - \pi_t^*}{\sigma_t}\right) dx_{-1} + g_{t-1}^0\phi\left(\frac{-x - \pi_t^*}{\sigma_t}\right), \\
g_t^0 &= 1 - \int_{s_t}^{S_t} g_t^c(x) dx, \\
1 &= \int e^{(x+p_t^*)(1-\epsilon)} g_t^c(x)\, dx + g_t^0 e^{(p_t^*)(1-\epsilon)}.
\end{aligned}$$

We now transform it in a convenient fashion. First, normalize the constraints involving $V_t(x)$ by A_t and substitute for the wage $w_t = \upsilon C_t$ and the discount factor $\Lambda_{t,t+1} = \beta\frac{C_t}{C_{t+1}}$.

With this, the constrains involving $V_t(x)$ become:

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

Second, define $\frac{V_t(x)}{A_t}\equiv\hat{V}_t(x)$, so that these constrains become

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

Finally, define $\hat{C}_t=\frac{C_t}{A_t}$. The planner's problem becomes

$$\max_{\left\{g_t^c(\cdot), g_t^0, \hat{V}_t(\cdot), \hat{C}_t, w_t, p_t^*, s_t, S_t, \pi_t^*, L_t\right\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log\left(\hat{C}\right) + \log\left(A_t\right) - \upsilon L_t\right)$$

$$\begin{aligned}
\hat{V}_t(x) &= \hat{C}_t \left(\exp\left(x_t + p_t^*\right)\right)^{1-\epsilon} - \hat{C}_t(1-\tau_t)\upsilon\hat{C}_t \left(\exp\left(x_t + p_t^*\right)\right)^{-\epsilon} \\
&\quad + \beta\hat{C}_t\hat{C}_{t+1}^{-1}\frac{1}{\sigma_{t+1}}\int_{s_{t+1}}^{S_{t+1}} \left[\hat{V}_{t+1}(x')\phi\left(\frac{(x-x') - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right] dx' \\
&\quad + \beta\hat{C}_t\hat{C}_{t+1}^{-1}\left(1 - \frac{1}{\sigma_{t+1}}\int_{s_{t+1}}^{S_{t+1}} \left[\phi\left(\frac{(x-x') - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right] dx'\right) \left[\left(\hat{V}_{t+1}(0) - \eta\upsilon\hat{C}_{t+1}\right)\right], \\
\hat{V}_t\left(s_t\right) &= \hat{V}_t\left(0\right) - \eta\upsilon\hat{C}_t, \\
\hat{V}_t\left(S_t\right) &= \hat{V}_t\left(0\right) - \eta\upsilon\hat{C}_t, \\
0 = \hat{V}_t'\left(0\right) &= (1-\epsilon)\hat{C}_t \left(\exp\left(x_t + p_t^*\right)\right)^{1-\epsilon} + \epsilon\hat{C}_t(1-\tau_t)\upsilon\hat{C}_t \left(\exp\left(x_t + p_t^*\right)\right)^{-\epsilon} \\
&\quad + \frac{1}{\sigma_{t+1}}\beta\hat{C}_t\hat{C}_{t+1}^{-1}\int_{s_{t+1}}^{S_{t+1}} \hat{V}_{t+1}(x') \left.\frac{\partial\phi\left(\frac{x-x'-\pi_{t+1}^*}{\sigma_{t+1}}\right)}{\partial x}\right|_{x=0} dx' \\
&\quad + \beta\hat{C}_t\hat{C}_{t+1}^{-1}\left(\phi\left(\frac{-S_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi_{t+1}^*}{\sigma_{t+1}}\right)\right)\left(\hat{V}(0) - \eta\upsilon\hat{C}_{t+1}\right). \\
L_t &= \hat{C}_t\left(\int e^{(x+p_t^*)(-\epsilon_t)} g_t^c\left(p\right) dx + g_t^0 e^{(p_t^*)(-\epsilon)}\right) \\
g_t^c(x) &= \frac{1}{\sigma_t}\int_{s_{t-1}}^{S_{t-1}} g_{t-1}^c(x_{-1})\phi\left(\frac{x_{-1} - x - \pi_t^*}{\sigma_t}\right) dx_{-1} + g_{t-1}^0\phi\left(\frac{-x - \pi_t^*}{\sigma_t}\right), \\
g_t^0 &= 1 - \int_{s_t}^{S_t} g_t^c(x)dx, \\
1 &= \int e^{(x+p_t^*)(1-\epsilon)} g_t^c\left(x\right) dx + g_t^0 e^{(p_t^*)(1-\epsilon)}.
\end{aligned}$$

which do not depend on A_t. Therefore, the redefined Ramsey policy does not respond to TFP shocks. Going back to the original variables definition, this implies that under optimal policy $C_t \propto A_t$ and $V_t(x) \propto A_t$ while all other variables remain constant at their steady-state values. Thus, inflation π_t also remains constant at its steady-state value.

D Additional figures

D.1 Slope of the nonlinear New Keynesian Phillips curve

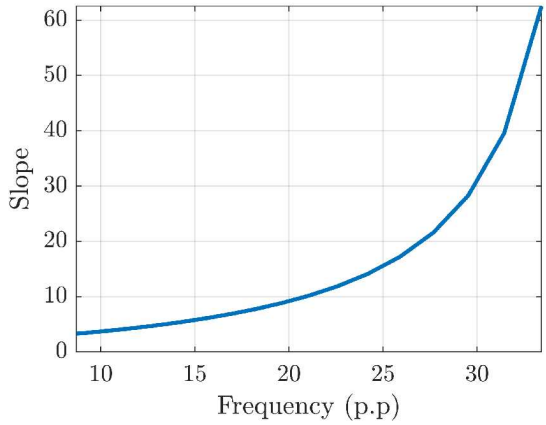


Figure 12: Slope of the nonlinear New Keynesian Phillips curve in the Golosov-Lucas model. *Note*: The figure is produced by computing the impulse responses to monetary policy shocks of different magnitudes and signs and then computing the slope.

D.2 Nonlinear New Keynesian Phillips curve

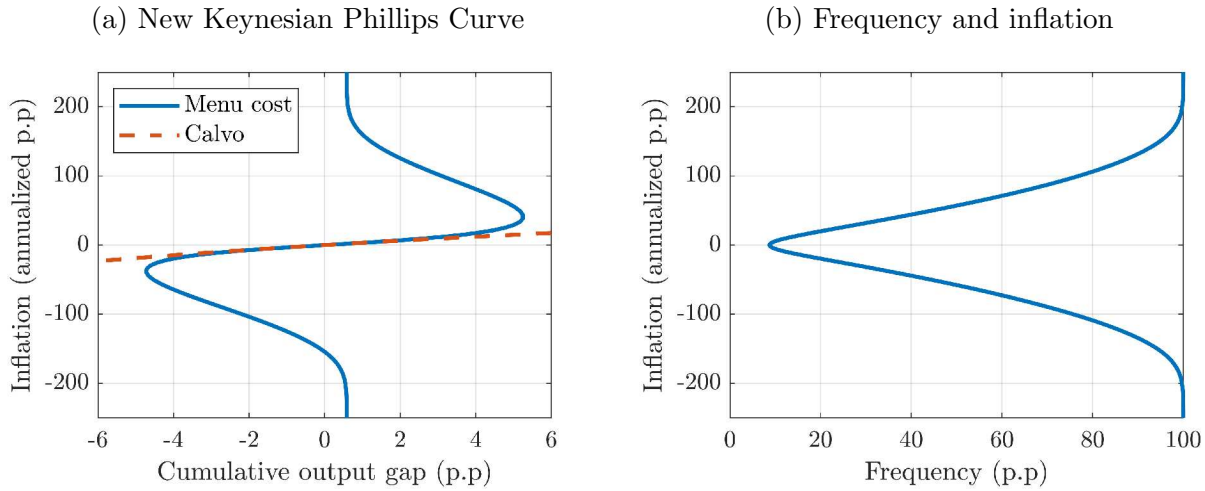


Figure 13: Inflation-output trade-offs in the menu cost model.

Note: The figure is produced by computing the impulse responses to monetary policy shocks of different magnitudes and signs.

D.3 Impulse responses under a Taylor rule

Cost-push shocks. The solid green lines on the first row of Figure 14 show the responses to a large cost-push shock. The shock is implemented as a persistent decline in the firms' employment subsidy τ_t. The shock size is calibrated to generate a 20% frequency on impact (a frequency increase of $20\% - 8.7\% = 11.3\%$, see panel d), a magnitude that has been documented during the 2022-2023 inflation surge (Montag and Villar, 2023). The exercise assumes that monetary policy follows an inertial Taylor rule (eq. 12). The shock captures some realistic features of the recent inflation surge. First, it generates a large increase in inflation (panel b), characterized by an initial temporary spike followed by a period of persistent inflation. Second, the shock causes a mild downturn as the output gap decreases (panel a).

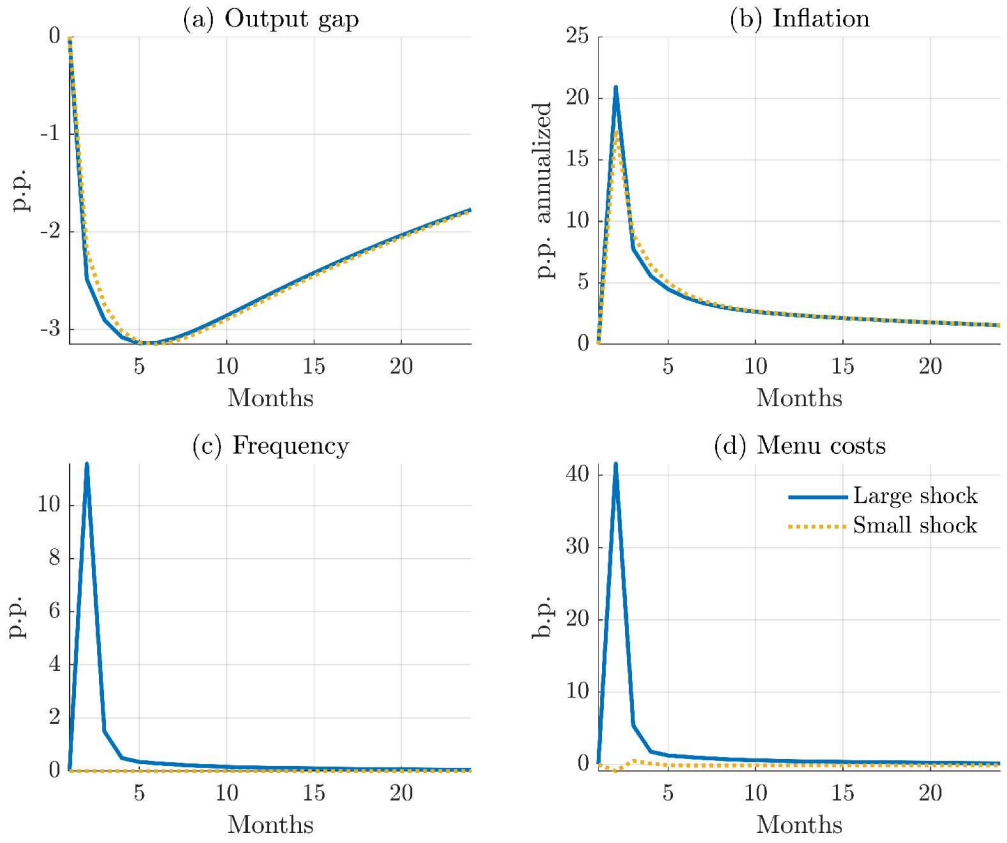


Figure 14: Impulse responses to a large and a small cost-push shock in the menu cost model. *Note*: all displayed variables except frequency are linearly scaled in the small shock according to the ratio between the large and small shocks.

In contrast, the black dashed lines show the responses to a *small* cost-push shock, scaled up linearly to the size of the large shock. The small shock is a 25 basis point decrease in the annualized employment subsidy. The difference between the figures illustrates the nonlinearity of the model. The inflation response is roughly 25% larger after the large shock than after the linearly scaled small shock. The key reason behind this is the sizable difference

between the frequency responses: while the frequency jumps after the large shock, it stays almost unchanged in response to the small shock.[22]

Comparison with the Calvo model. We compare the results above, for the small shock re-scaled, with the standard Calvo model in Figure 15. We consider two calibrations for the Calvo parameter. The dotted-dashed green line corresponds to a calibration in which the frequency of price adjustment, which in Calvo is constant, matches the frequency of price adjustments of 8.7% in the Golosov-Lucas model. In this case, the response of both inflation and output gap is much attenuated compared to the menu cost model (solid gray line).

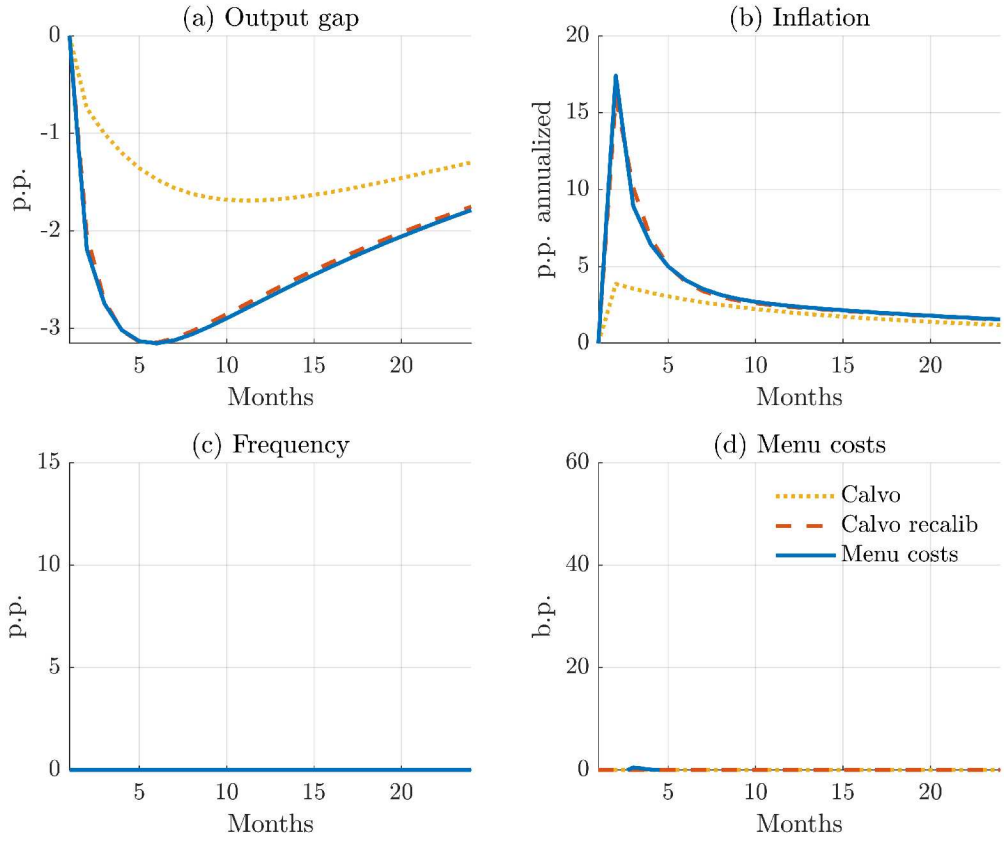


Figure 15: Impulse responses to a small cost-push shock in the menu cost and in the Calvo model. *Note*: all displayed variables except frequency are linearly-scaled in the small shock according to the ratio between the large and small shocks. The Calvo recalibrated model replicates the slope of the Phillips curve

The dashed blue line in Figure 15 shows the recalibrated Calvo model. The Calvo parameter in that case is set to to match the slope of the Phillips curve in the Golosov-Lucas model (discussed below). The recalibrated Calvo model approximately replicates the dynamics of the Golosov-Lucas model in response to small shocks. The fact that the Golosov-Lucas model can be locally approximated for small inflation levels by a re-calibrated Calvo model with a large frequency was studied by Auclert et al. (2024).

[22]The larger initial inflation bout in the case of large shocks is then compensated by a lower inflation path, relative to the small shock counterfactual, from month 3 onwards, and both paths coincide after 9 months.

Price dispersion shocks. Figure 16 shows the impulse responses to a shock to the volatility of the idiosyncratic quality shocks (σ_t). The shock raises the dispersion of optimal reset prices and generates a persistent increase in the frequency of price changes (Vavra, 2014). The observed increase in the dispersion of price changes during the recent inflation surge (Montag and Villar, 2023) indicates the presence of a similar idiosyncratic dispersion shock either parallel or as a result of the aggregate shocks (Berger and Vavra, 2019). The size of the large shock is calibrated to generate a 11.3% increase in frequency, which coincides with the peak of the cost-push shock. Notably, the shock is inflationary. This is primarily due to the asymmetry of the value function: firms with too low quality-adjusted prices face high demand, so they are relatively more motivated to *increase* their prices than firms with too high quality-adjusted prices are motivated to *decrease* theirs due to their low demand. The output gap declines primarily as a result of the high price distortions. The differences relative to the linearly scaled small volatility shock (black dashed lines) reveal a sizable nonlinearity of the model with respect to this shock. Nonetheless, notice how the impact of the shock on inflation and output is one order of magnitude lower than that of the cost-push shock presented in Figure 14.

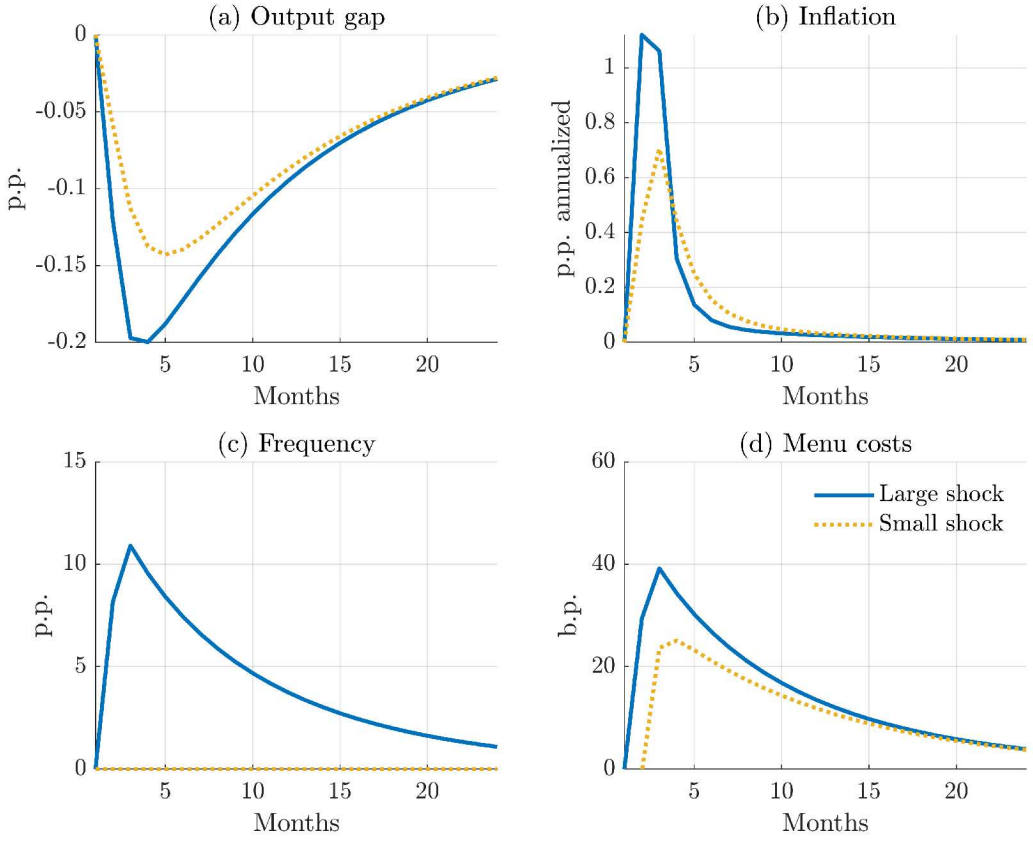


Figure 16: Impulse responses to a large and a small dispersion shock in the menu cost model. *Note*: all displayed variables except frequency are linearly-scaled in the small shock according to the ratio between the large and small shocks.

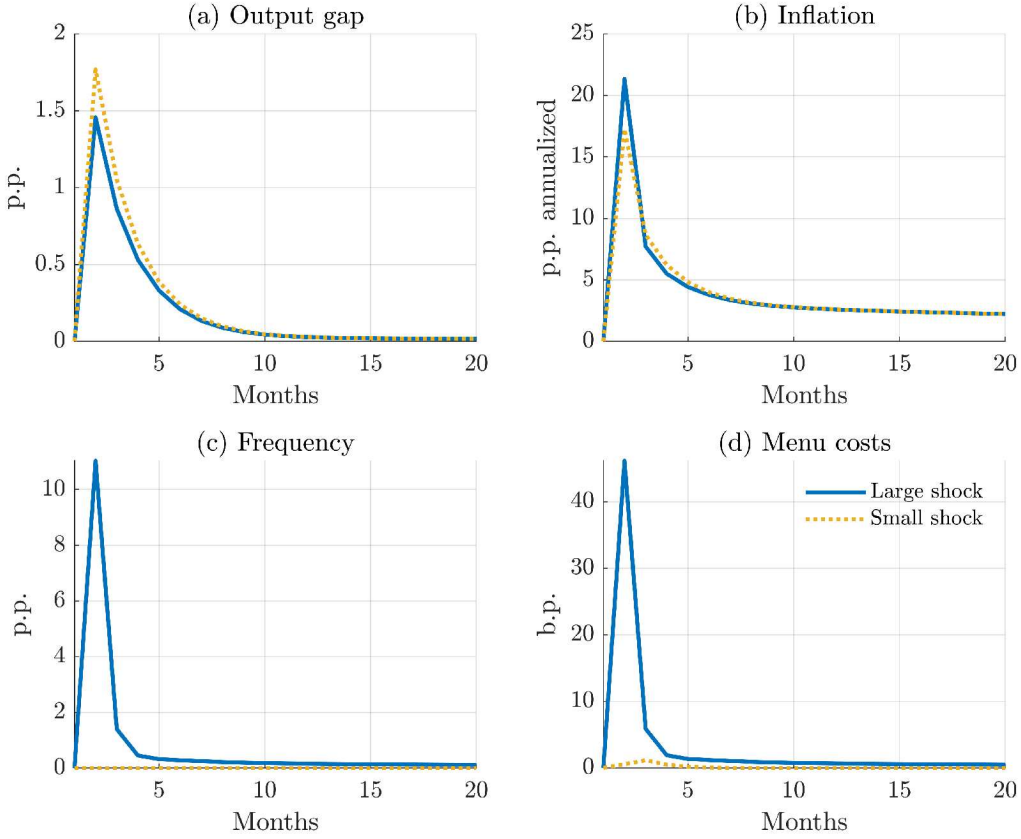


Figure 17: Impulse responses to a large and a small TFP shock in the menu cost model. *Note*: all displayed variables except frequency are linearly-scaled in the small shock according to the ratio between the large and small shocks.

TFP and monetary policy shocks. Figures 17 and 18 show analogous impulse responses to a TFP and monetary policy shocks, respectively. These shocks are calibrated to replicate the 11.3% increase in frequency. The model again displays significant nonlinearities in both cases, as the inflation surge is larger, and the output gap surge is smaller, in the case of a large shock.

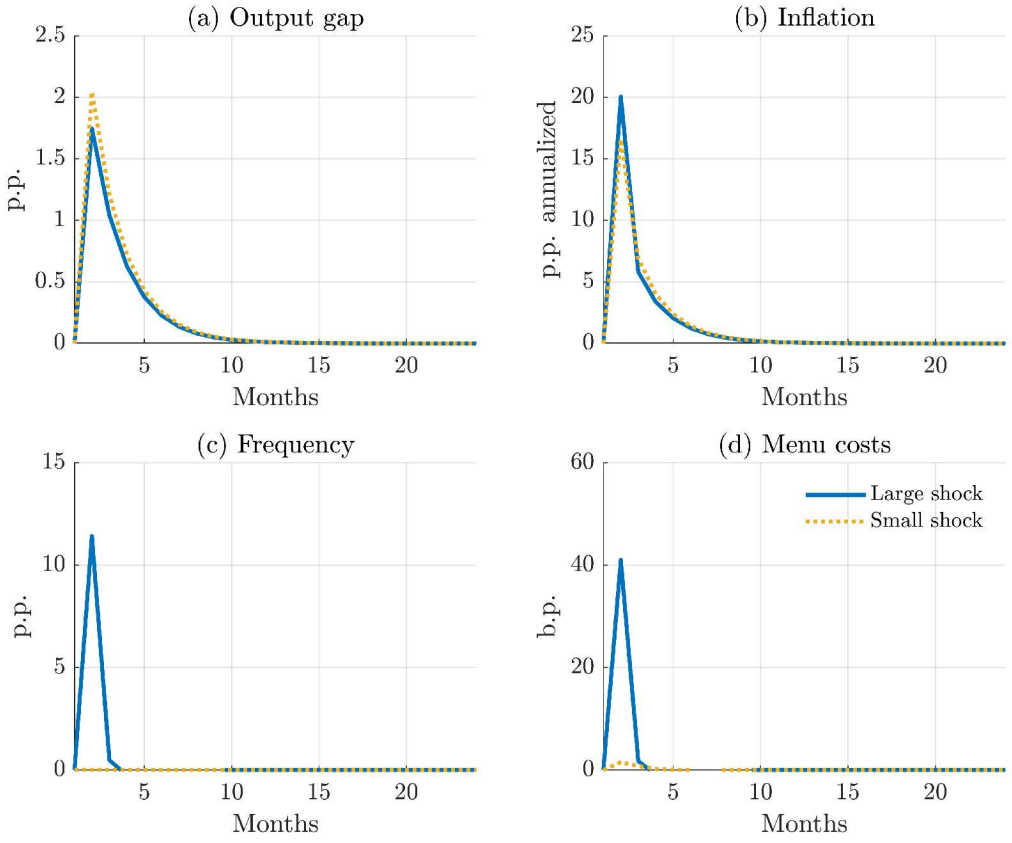


Figure 18: Impulse responses to a large and a small monetary policy shock in the menu cost model.
Note: all displayed variables except frequency are linearly-scaled in the small shock according to the ratio between the large and small shocks.

E Computational algorithm

This appendix explains the computational method. We use a three-step approach we use to convert the original infinite-dimensional Ramsey problem into a finite-dimensional one. First, we approximate the distribution and value functions by piece-wise linear functions over a set of nodes. Second, we use endogenous nodes, such that both boundaries of the (s_t, S_t) band and the optimal reset price are "on the grid". Third, given this approximation, we evaluate integrals analytically. Step one makes the problem finite dimensional. 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 planner's infinite-dimensional problem into a finite-dimension problem in this way, we derive the planner's first order conditions. For this we use symbolic differentiation, and in particular Dynare's `Ramsey` command. The resulting set of first order conditions is then solved in the sequence space under perfect foresight. Here we employ a standard Newton method using Dynare's `perfect foresight solver` command.

To determine the appropriate initial and terminal conditions, and an initial guess for the transition paths, we need to find the non-stochastic steady state of the model. We determine the steady state of the private equilibrium conditional on a particular value of the policy instrument π using a standard Newton based solution method. We then use this function and exploit the linearity of the first order conditions wrt. the Lagrange multipliers to convert the high-dimensional problem of solving for the steady state into a one-dimensional problem, which is solved with a Newton solver. This last step is performed by Dynare's `steady` command. That is, we have to manually convert the problem into a finite-dimension problem and find the steady state conditional on a policy; the rest of the procedure uses Dynare.

The rest of the appendix explains those steps that are not straightforward applications of existing methods. It is organized as follows. First we explain how to make the planner's problem finite dimensional. For this purpose, we first define some useful auxiliary functions in Section E.1. Then we transform the equilibrium conditions to apply an endogenous grid and approximate the value and distribution functions by a piece-wise linear function in Section E.2. Finally, we evaluate the integrals analytically in Section E.3. The result is a discrete set of equations that can conveniently be represented in matrix form, which we summarized in Section E.4. Second, we explain how we determine the steady state in Section E.5.

E.1 Preliminaries

To begin with, let us normalize the variable x_t as

$$\mathsf{x_t} = \begin{cases} \frac{x_t}{s_t} & \text{if } x_t < 0 \\ \frac{x_t}{S_t} & \text{else} \end{cases} \tag{34}$$

Under this normalization, the optimal price is at $\mathsf{x_t} = 0$, the upper boundary of the (S, s) band at $\mathsf{x_t} = 1$ and the lower boundary 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 $\mathsf{x_t}$ conditional on not updating can be derived from $x_t = x_{t-1} - \sigma\varepsilon_t - \pi_t^*$:

$$\mathsf{x_t} = \begin{cases} \frac{x_t}{S_t} = \frac{x_{t-1}-\sigma_t\varepsilon_t-\pi_t^*}{S_t} = \begin{cases} \frac{x_{t-1}-\sigma_t\varepsilon_t-\pi_t^*}{S_{t-1}}\frac{S_{t-1}}{S_t} = \mathsf{x_{t-1}}\frac{S_{t-1}}{S_t} - \frac{\sigma_t\varepsilon_t+\pi_t^*}{S_t} & \text{if } \mathsf{x_t} > 0, \text{if } \mathsf{x_{t-1}} > 0 \\ \frac{x_{t-1}-\sigma_t\varepsilon_t-\pi_t^*}{s_{t-1}}\frac{s_{t-1}}{S_t} = \mathsf{x_{t-1}}\frac{s_{t-1}}{S_t} - \frac{\sigma_t\varepsilon_t+\pi_t^*}{S_t} & \text{if } \mathsf{x_t} > 0, \text{if } \mathsf{x_{t-1}} < 0 \end{cases} \\ \frac{x_t}{s_t} = \frac{x_{t-1}-\sigma_t\varepsilon_t-\pi_t^*}{s_t} = \begin{cases} \frac{x_{t-1}-\sigma_t\varepsilon_t-\pi_t^*}{S_{t-1}}\frac{S_{t-1}}{s_t} = \mathsf{x_{t-1}}\frac{S_{t-1}}{s_t} - \frac{\sigma_t\varepsilon_t+\pi_t^*}{s_t} & \text{if } \mathsf{x_t} < 0, \text{if } \mathsf{x_{t-1}} > 0 \\ \frac{x_{t-1}-\sigma_t\varepsilon_t-\pi_t^*}{s_{t-1}}\frac{s_{t-1}}{s_t} = \mathsf{x_{t-1}}\frac{s_{t-1}}{s_t} - \frac{\sigma_t\varepsilon_t+\pi_t^*}{s_t} & \text{if } \mathsf{x_t} < 0, \text{if } \mathsf{x_{t-1}} < 0 \end{cases} \end{cases} \tag{35}$$

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 $\mathsf{x_t}$ following (34). For brevity, at times we will drop the dependence on $\mathsf{x_t}$ and just write $\hat{s}_t$.

Solving (35) for $\mathsf{x_t}$, $\mathsf{x_{t-1}}$ and ε respectively, we obtain the following relations:

$$\mathsf{x_t} = \mathsf{x_{t-1}}\frac{\hat{s}_{t-1}}{\hat{s}_t} - \frac{\sigma_t\varepsilon_t + \pi_t^*}{\hat{s}_t} \tag{36}$$

$$\mathsf{x_{t-1}} = \mathsf{x_t}\frac{\hat{s}_t}{\hat{s}_{t-1}} + \frac{\sigma_t\varepsilon_t + \pi_t^*}{\hat{s}_{t-1}} \tag{37}$$

$$\varepsilon_t = \frac{\hat{s}_{t-1}\mathsf{x_{t-1}} - \hat{s}_t\mathsf{x_t} - \pi_t^*}{\sigma_t} \equiv h(\mathsf{x_{t-1}}, \mathsf{x_t}) \tag{38}$$

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

E.2 Approximating the distribution and value functions by piecewise linear functions on an endogenous grid

Now we redefine the value and distribution functions over the variable x and approximate them by piece-wise linear functions. The original infinite dimensional problem of the planner are laid out in Section 5.1. In the following, we consider each of the equations containing the distribution and value functions one by one.

E.2.1 Distribution

The distribution function is given by

$$g_t(x) = (1-\lambda_t(x))\int g_{t-1}(x+\sigma_t\varepsilon_t+\pi_t^*)d\xi(\varepsilon)+\delta(x)\int\lambda_t(\tilde{x})\left(\int g_{t-1}(\tilde{x}+\sigma_t\varepsilon+\pi_t^*)d\xi(\varepsilon)\right)d\tilde{x}$$

with

$$\int_{s_t}^{S_t} g_t(x)dx = 1 \tag{39}$$

where $\delta(x)$ is the Dirac Delta function that captures the mass point of those firms who update their prices.

We split the distribution into the continuous distribution of agents who do not update their prices plus a mass point of updaters at $x = 0$ (this is already reflected in Section 5.1):

$$g_t^c(x) = (1 - \lambda_t(x)) \int g_{t-1}^c(x + \sigma_t \varepsilon + \pi_t^*) d\xi(\varepsilon),$$
$$g_t^0 = \int \lambda_t(\tilde{x}) \int g_{t-1}^c(\tilde{x} + \sigma_t \varepsilon - \pi_t^*) d\xi(\varepsilon) d\tilde{x}.$$

Furthermore, we use equation (39) to express the latter expression as:

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

Now rewrite it using the newly defined re-normalized x where $x = \mathsf{x}\hat{s}_t$ as in equation (34): define $g_t(\mathsf{x}\hat{s}_t) \equiv \mathsf{g}_t(\mathsf{x})$ and $g_t^c(\mathsf{x}\hat{s}_t) \equiv \mathsf{g}_t^c(\mathsf{x})$ and, with a slight abuse of notation, $\lambda_t(\mathsf{x}\hat{s}_t) \equiv \lambda_t(\mathsf{x})$ and write

$$\mathsf{g}_t^c(\mathsf{x}) = (1 - \lambda_t(\mathsf{x})) \int \mathsf{g}_{t-1} \left(\frac{\mathsf{x}\hat{s}_t + \sigma_t \varepsilon + \pi_t^*}{\hat{s}_{t-1}} \right) d\xi(\varepsilon), \tag{40}$$

$$\mathsf{g}_t^0 = 1 - \int_{-1}^{1} \mathsf{g}_t^c(\mathsf{x}) \hat{s}_t(\mathsf{x}) d\mathsf{x}. \tag{41}$$

Note that for the latter expression for g_t^0 we have applied a change of variable to the integral. In particular, we have used the following substitution:

$$\begin{aligned} \int_{s_t}^{S_t} g_t^c(x) dx &= \int_{s_t}^{S_t} g_t^c(\mathsf{x}\hat{s}_t(\mathsf{x})) d\mathsf{x}\hat{s}_t(\mathsf{x}) \\ &= \int_{s_t}^{S_t} \mathsf{g}_t^c(\mathsf{x}) d\mathsf{x}\hat{s}_t(\mathsf{x}) = \int_{s_t/\hat{s}_t(\mathsf{x})}^{S_t/\hat{s}_t(\mathsf{x})} \hat{s}_t(\mathsf{x}) \mathsf{g}_t^c(\mathsf{x}) d\mathsf{x} = \int_{-1}^{1} \hat{s}_t(\mathsf{x}) \mathsf{g}_t^c(\mathsf{x}) d\mathsf{x}. \end{aligned}$$

Next we will also change the variable in the integral in the equation for $\mathsf{g}_t^c(\mathsf{x})$ (40). This change of variable is a bit more involved, so we derive it in detail here. First, we re-express (40) as

$$\mathsf{g}_t^c(\mathsf{x}) = (1 - \lambda_t(\mathsf{x})) \int \mathsf{g}_{t-1} \left(\frac{\mathsf{x}\hat{s}_t + \sigma_t \varepsilon + \pi_t^*}{\hat{s}_{t-1}} \right) \phi(\varepsilon) d\varepsilon.$$

where $\phi(\cdot)$ is the standard normal pdf.

Second, we define the value of the shock ε necessary to get from a price gap of $x_{t-1} = 0$ to a price gap of x_t as

$$\begin{aligned} \varepsilon_t^* &\equiv \varepsilon_t \in \mathbb{R} | 0 = \mathsf{x}_t \hat{s}(\mathsf{x}_t) + \sigma_t \varepsilon_t + \pi_t^* \\ &= h(0, \mathsf{x}_t) \end{aligned}$$

and then we split the integral in two parts at ε_t^*

$$\begin{aligned}\mathsf{g}_t^c(\mathsf{x}) &= (1-\lambda_t(\mathsf{x}))\int^{\varepsilon_t^*}\mathsf{g}_{t-1}\left(\frac{\mathsf{x}\hat{s}_t+\sigma_t\varepsilon+\pi_t^*}{\hat{s}_{t-1}}\right)\phi(\varepsilon)d\varepsilon \\ &+ (1-\lambda_t(\mathsf{x}))\int_{\varepsilon_t^*}\mathsf{g}_{t-1}\left(\frac{\mathsf{x}\hat{s}_t+\sigma_t\varepsilon+\pi_t^*}{\hat{s}_{t-1}}\right)\phi(\varepsilon)d\varepsilon,\end{aligned}$$

Since, for a realization ε of the shock at t,

$$\hat{s}\left(\mathsf{x}_{t-1}\right)\mathsf{x}_{t-1} = \mathsf{x}_t\hat{s}\left(\mathsf{x}_t\right)+\sigma_t\varepsilon+\pi_t^*, \tag{42}$$

we have

$$d\varepsilon = \left(\frac{\hat{s}\left(\mathsf{x}_{t-1}\right)}{\sigma_t}+\frac{\mathsf{x}_{t-1}}{\sigma_t}\frac{d\hat{s}\left(\mathsf{x}_{t-1}\right)}{d\mathsf{x}_{t-1}}\right)d\mathsf{x}_{t-1}.$$

In each of the two intervals over which the two integrals are defined, the mapping (42) is continuous and $\frac{d\hat{s}(\mathsf{x}_{t-1})}{d\mathsf{x}_{t-1}}=0$. Thus we can implement a change of variable from ε to x_{t-1} in both integrals:

$$\begin{aligned}\mathsf{g}_t^c(\mathsf{x}) &= (1-\lambda_t(\mathsf{x}))\int^{\varepsilon_t^*}\frac{s_{t-1}}{\sigma_t}\mathsf{g}_{t-1}(\mathsf{x}_{t-1})\phi\left(\frac{s_{t-1}\mathsf{x}_{t-1}-\hat{s}_t\mathsf{x}-\pi_t^*}{\sigma_t}\right)d\mathsf{x}_{t-1} \\ &+ (1-\lambda_t(\mathsf{x}))\int_{\varepsilon_t^*}\mathsf{g}_{t-1}(\mathsf{x}_{t-1})\phi\left(\frac{S_{t-1}\mathsf{x}_{t-1}-\hat{s}_t\mathsf{x}-\pi_t^*}{\sigma_t}\right)\frac{S_{t-1}}{\sigma_t}d\mathsf{x}_{t-1}.\end{aligned}$$

Finally, pasting the two integrals together again, re-denoting x_{t-1} by x' and using $h(\mathsf{x}',\mathsf{x})$

$$\mathsf{g}_t^c(\mathsf{x}) = (1-\lambda_t(\mathsf{x}))\int\frac{\hat{s}_{t-1}(\mathsf{x}')}{\sigma_t}\mathsf{g}_{t-1}(\mathsf{x}')\phi\left(h(\mathsf{x}',\mathsf{x})\right)d\mathsf{x}'$$

This concludes the change of variable.

The end of period distribution has mass on the (s,S) band, i.e. in the range $x\in[-1,1]$. We can thus restrict the boundaries of the integral accordingly:

$$\mathsf{g}_t^c(\mathsf{x}) = (1-\lambda_t(\mathsf{x}))\int_{-1}^{1}\frac{\hat{s}_{t-1}(\mathsf{x}')}{\sigma_t}\mathsf{g}_{t-1}(\mathsf{x}')\phi\left(h(\mathsf{x}',\mathsf{x})\right)\ d\mathsf{x}'.$$

So far we have rewritten the law of motion of the firm distribution g_t. We now introduce the approximation we rely on for g_t. We approximate g^c by a piece-wise linear function with equally spaced nodes $\mathsf{x}_1,\ldots,\mathsf{x}_I=-1,\ldots,0,\ldots,1$ with $\mathsf{g}_t^c(\mathsf{x}|\mathsf{x}_i<\mathsf{x}<\mathsf{x}_{i+1})\approx\mathsf{g}_t^c(\mathsf{x}_i)+\frac{\mathsf{x}-\mathsf{x}_i}{\mathsf{x}_{i+1}-\mathsf{x}_i}\frac{g_{t-1}^c(\mathsf{x}_{i+1})-g_{t-1}^c(\mathsf{x}_i)}{\mathsf{x}_{i+1}-\mathsf{x}_i}$.

Note that the auxiliary grid for x is exogenous. However, this exogenous auxiliary grid defines an *endogenous grid* for $x=\hat{s}_t\mathsf{x}$, which, at each t, exactly spans the the (s,S) band and has a node at 0. Figure 19 illustrates the use of linear interpolation with an endogenous grid as we apply it here.

From now on, g_t^c denotes the piece-wise linear approximated function and $\mathsf{g}_t^c(\mathsf{x}_i < \mathsf{x} < \mathsf{x}_{i+1})$ denotes a linear piece of it. Thus, the functions are approximated as

$$\mathsf{g}_t^c(\mathsf{x}) = (1 - \lambda_t(\mathsf{x})) \left[\sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \frac{\hat{s}_{t-1}(\mathsf{x}')}{\sigma_t} \mathsf{g}_{t-1}^c(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) \phi\left(h(\mathsf{x}', \mathsf{x})\right) \, d\mathsf{x}' + \frac{1}{\sigma_t} \mathsf{g}_{t-1}^0 \phi\left(h(0, \mathsf{x})\right) \right],$$

$$\mathsf{g}_t^0 = 1 - \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \mathsf{g}_t^c(\mathsf{x}_i < \mathsf{x} < \mathsf{x}_{i+1}) \hat{s}_t(\mathsf{x}) d\mathsf{x}.$$

Notice that, in these expressions, the integrands are continuous in the interval $\mathsf{x}_i < \mathsf{x} < \mathsf{x}_{i+1}$ 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 $\mathsf{x} \in [-1, 1]$. This is computationally efficient.

Withing this range $(1 - \lambda_t(\mathsf{x})) = 1$ so we can drop it from the expression above.

$$\mathsf{g}_t^c(\mathsf{x}) = \left[\sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \frac{\hat{s}_{t-1}(\mathsf{x}')}{\sigma_t} \mathsf{g}_{t-1}^c(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) \phi\left(h(\mathsf{x}', \mathsf{x})\right) \, d\mathsf{x}' + \frac{1}{\sigma_t} \mathsf{g}_{t-1}^0 \phi\left(h(0, \mathsf{x})\right) \right]$$

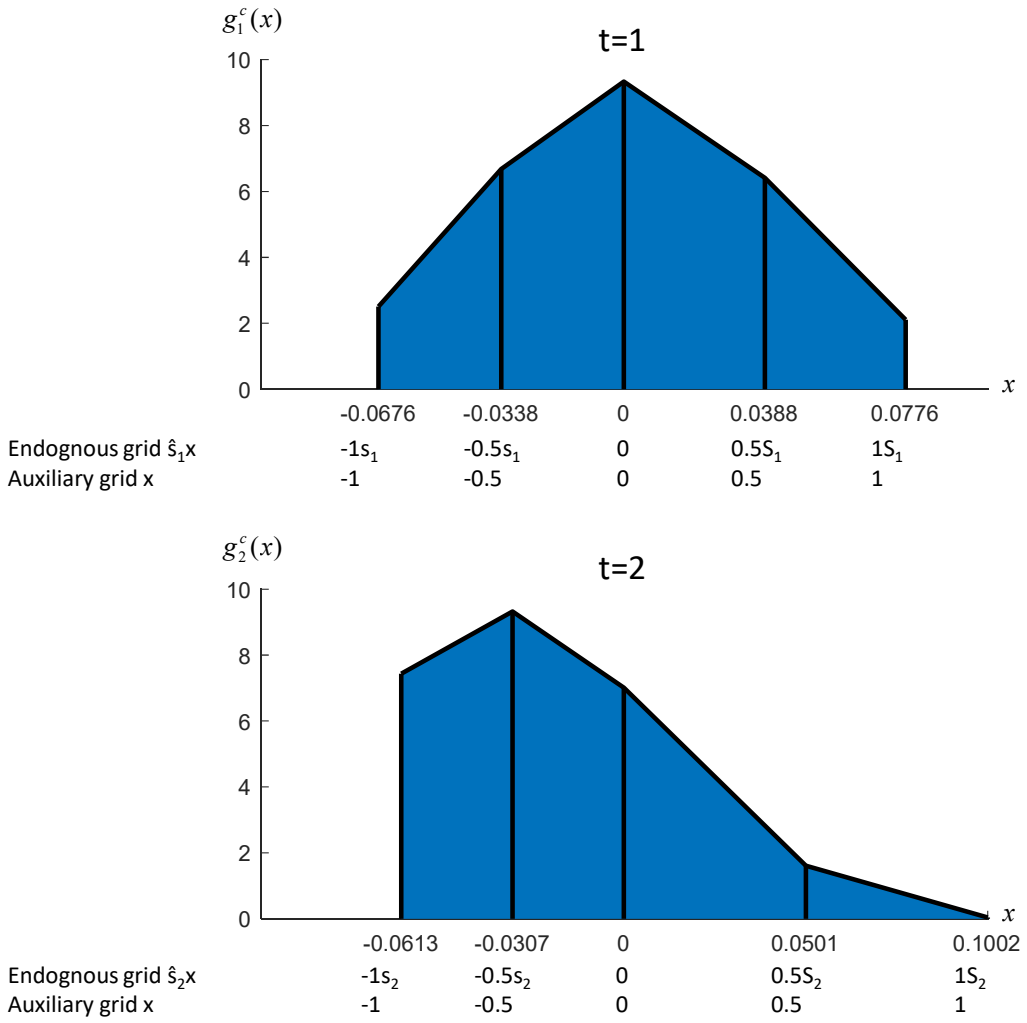


Figure 19: This figure schematically explains the linear interpolation with an endogenous grid. It shows the piece-wise linearly approximated distribution $\mathsf{g}_t^c(x)$ at two points in time, $t = 1$ and $t = 2$. The thresholds of the (S, s) band are not symmetric around 0 and differ across time. 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 $\mathsf{x} = [-1, -0.5, 0, 0.5, 1]$ by $\hat{s}_t(\mathsf{x})$: $x = \mathsf{x}\hat{s}_t$

E.2.2 Other Aggregation Equations

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

$$\begin{aligned} e^{p_t^*(\epsilon-1)} &= \int e^{x(1-\epsilon)} g_t(x)dx, \\ N_t &= \frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \int e^{x(-\epsilon)} g_t(x)d(x) + \eta \int \lambda_t(x + p_t^* - \sigma_t\varepsilon_t - \pi_t^*) g_{t-1}(x)d(x) \end{aligned}$$

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

$$e^{p_t^*(\epsilon-1)} = \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} e^{\mathsf{x}(1-\epsilon)} \mathsf{g}_t^c(\mathsf{x}_{i-1} < \mathsf{x} < \mathsf{x}_{i+1}) \hat{s}_t(\mathsf{x})d\mathsf{x} + \mathsf{g}_t^0,$$

$$N_t = \frac{C_t}{A_t} e^{p^*(-\epsilon)} \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \left(e^{\mathsf{x}(-\epsilon)} \mathsf{g}_t^c(\mathsf{x}_{i-1} < \mathsf{x} < \mathsf{x}_{i+1}) \hat{s}_t(\mathsf{x})d\mathsf{x} + \mathsf{g}_{t-1}^0 \right) + \eta \mathsf{g}_{t-1}^0.$$

E.2.3 Value Function

Recall the value function is

$$\begin{aligned} V_t(x) =& \Pi_t(x) + \Lambda_{t,t+1} \int \left(1 - \lambda_{t+1}(x - \sigma_{t+1}\varepsilon - \pi_{t+1}^*)\right) V_{t+1}(x - \sigma_{t+1}\varepsilon - \pi_{t+1}^*) d\xi(\varepsilon) \\ &+ \Lambda_{t,t+1} \left(V_{t+1}(0) - \eta w_{t+1}\right) \int \lambda_{t+1}(x - \sigma_{t+1}\varepsilon - \pi_{t+1}^*) d\xi(\varepsilon). \end{aligned}$$

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

$$\begin{aligned} \mathsf{V}_t(\mathsf{x}) = \Pi_t(\mathsf{x}) + \Lambda_{t,t+1} \int \left(1 - \lambda_{t+1}\left(\frac{\mathsf{x}\hat{s}_t + \sigma_t\varepsilon + \pi_t^*}{\hat{s}_{t-1}}\right)\right) \mathsf{V}_{t+1}\left(\frac{\mathsf{x}\hat{s}_t + \sigma_t\varepsilon + \pi_t^*}{\hat{s}_{t-1}}\right) d\xi(\varepsilon) \\ + \Lambda_{t,t+1} \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \int \lambda_{t+1}\left(\frac{\mathsf{x}\hat{s}_t + \sigma_t\varepsilon + \pi_t^*}{\hat{s}_{t-1}}\right) d\xi(\varepsilon). \end{aligned}$$

Note that by definition $\mathsf{V}_t(0) - \eta\frac{w_{t+1}}{A_{t+1}} = \mathsf{V}_t(-1) = \mathsf{V}_t(1)$ and $\mathsf{V}_t'(0) = 0$. 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 $\mathsf{V}_t(\mathsf{x})$ as

$$\begin{aligned} \mathsf{V}_t(\mathsf{x}) =& \Pi_t(\mathsf{x}) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int \hat{s}_{t+1}(\mathsf{x}') \left(1 - \lambda_{t+1}(\mathsf{x}')\right) \mathsf{V}_{t+1}(\mathsf{x}') \phi\left(h(\mathsf{x}, \mathsf{x}')\right) d\mathsf{x}' \\ &+ \Lambda_{t,t+1} \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \int \hat{s}_{t+1}(\mathsf{x}') \lambda_{t+1}(\mathsf{x}') \phi\left(h(\mathsf{x}, \mathsf{x}')\right) \frac{1}{\sigma_{t+1}} d\mathsf{x}' \end{aligned}$$

Since the price updating probability $\lambda_{t+1}(\mathsf{x}) = 1$ for any x outside the (S, s) band, we can restrict the first integral to the range $[-1, 1]$. The last term in the second line (which captures the probability of updating a price tomorrow, given the current state) can be replaced by 1 minus the probability of not updating the price tomorrow. The latter is given by an integral over the range $[-1, 1]$. So we write:

$$\begin{aligned}\mathsf{V}_t(\mathsf{x}) =&\Pi_t(\mathsf{x}) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left(1 - \lambda_{t+1}(\mathsf{x}')\right) \mathsf{V}_{t+1}(\mathsf{x}') \phi\left(h(\mathsf{x}, \mathsf{x}')\right) d\mathsf{x}' \\ &+ \Lambda_{t,t+1} \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \left(1 - \frac{1}{\sigma_{t+1}} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left(1 - \lambda_{t+1}(\mathsf{x}')\right) \phi\left(h(\mathsf{x}, \mathsf{x}')\right) d\mathsf{x}'\right).\end{aligned}$$

In the inaction region, the price updating probability $\Lambda_{t,t+1}(\mathsf{x}) = 0$, so:

$$\begin{aligned}\mathsf{V}_t(\mathsf{x}) =&\Pi_t(\mathsf{x}) + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \mathsf{V}_{t+1}(\mathsf{x}') \phi\left(h(\mathsf{x}, \mathsf{x}')\right) d\mathsf{x}' \\ &+ \Lambda_{t,t+1} \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \left(1 - \frac{1}{\sigma_{t+1}} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \phi\left(h(\mathsf{x}, \mathsf{x}')\right) d\mathsf{x}'\right).\end{aligned}$$

So far we have normalized the support of the value function. Additionally, it is convenient to normalize further the value function itself. We normalize the value function by its maximal value $\mathsf{V}_t(0)$, and denote the normalized value function by $\mathsf{v}_t(\mathsf{x})$: $\mathsf{v}_t(\mathsf{x}) \equiv \mathsf{V}_t(\mathsf{x}) - \mathsf{V}_t(0)$. The expression above can be re-written as:

$$\begin{aligned}&\mathsf{v}_t(\mathsf{x}) \equiv \mathsf{V}_t(\mathsf{x}) - \mathsf{V}_t(0) = \Pi_t(\mathsf{x}) - \Pi_t(0) \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left[\mathsf{V}_{t+1}(\mathsf{x}') \phi\left(\frac{\mathsf{x} - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \mathsf{V}_{t+1}(\mathsf{x}') \phi\left(\frac{0 - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right] d\mathsf{x}'\right) \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(-\int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left[\phi\left(\frac{\mathsf{x} - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{0 - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right] d\mathsf{x}'\right) \left(V_{t+1}(0) - \eta w_{t+1}\right) \\ &= \Pi_t(\mathsf{x}) - \Pi_t(0) \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left[\mathsf{v}_{t+1}(\mathsf{x}') \left(\phi\left(\frac{\mathsf{x} - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{0 - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right)\right] d\mathsf{x}'\right) \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(-\int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left[\phi\left(\frac{\mathsf{x} - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{0 - \mathsf{x}' - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right] d\mathsf{x}'\right) \left(-\eta w_{t+1}\right)\end{aligned}$$

Following our approach for $\mathsf{g}^c(\cdot)$, we approximate $\mathsf{v}(\cdot)$ by a piece-wise linear function with nodes $\mathsf{x}_1, \ldots, \mathsf{x}_I = -1, \ldots, 0, \ldots, 1$ with $\mathsf{v}_t(\mathsf{x} | \mathsf{x}_i < \mathsf{x} < \mathsf{x}_{i+1}) \approx \mathsf{v}_t(\mathsf{x}_i) + \frac{\mathsf{x} - \mathsf{x}_i}{\mathsf{x}_{i+1} - \mathsf{x}_i} \frac{\mathsf{v}_t(\mathsf{x}_{i+1}) - \mathsf{v}_t(\mathsf{x}_i)}{\mathsf{x}_{i+1} - \mathsf{x}_i}$.

From now on, v_t denotes the piece-wise linear approximated function and $\mathsf{v}_t(\mathsf{x}_i < \mathsf{x} < \mathsf{x}_{i+1})$

denotes a linear piece of it. Thus, this function $\mathsf{v}_t(\mathsf{x})$ is approximated as

$$\begin{aligned}\mathsf{v}_t(\mathsf{x}) =&\Pi_t(\mathsf{x}) - \Pi_t(0)\\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \hat{s}_{t+1}(\mathsf{x}')\mathsf{v}_{t+1}(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1})(\phi\left(h(\mathsf{x},\mathsf{x}')\right) - \phi\left(h(0,\mathsf{x}')\right))d\mathsf{x}'\\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(-\eta w_{t+1}\right) \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}')(\phi\left(h(\mathsf{x},\mathsf{x}')\right) - \phi\left(h(0,\mathsf{x}')\right))d\mathsf{x}'.\end{aligned}$$

E.2.4 Optimality condition for reset price

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

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

where

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

After change of variable to x, this expression becomes

$$\begin{aligned}0 = \Pi_t'(0) &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}')\mathsf{V}_{t+1}(\mathsf{x}')h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}'\\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1}-\pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1}-\pi^*_{t+1}}{\sigma_{t+1}}\right)\right)\left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right).\end{aligned}$$

Now we re-express this in terms of $\mathsf{v}(\mathsf{x})$ using $\mathsf{V}_t(\mathsf{x}) = \mathsf{v}_t(\mathsf{x}) + \mathsf{V}_t(0)$ first, and the rear-

ranging

$$
\begin{aligned}
0 &= \Pi_t'(x) + \Lambda_{t,t+1} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \left(\mathsf{v}_{t+1}(\mathsf{x}') + \mathsf{V}_{t+1}(0)\right) h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}' \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right) \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \\
&= \Pi_t'(x) + \Lambda_{t,t+1} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \mathsf{v}_{t+1}(x') h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}' \\
&\quad + \Lambda_{t,t+1} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} dx' \mathsf{V}_{t+1}(0) \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right) \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \\
&= \Pi_t'(x) + \Lambda_{t,t+1} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \mathsf{v}_{t+1}(x') h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}' \\
&\quad - \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right) \mathsf{V}_{t+1}(0) \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right) \left(\mathsf{V}_{t+1}(0) - \eta w_{t+1}\right) \\
&= \Pi_t'(0) + \Lambda_{t,t+1} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \mathsf{v}_{t+1}(\mathsf{x}') h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}' \\
&\quad + \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right) \left(-\eta w_{t+1}\right)
\end{aligned}
$$

and apply the piece-wise linear approximation of $\mathsf{v}(\mathsf{x})$:

$$
\begin{aligned}
0 = \Pi_t'(0) &+ \Lambda_{t,t+1} \sum_{i=1}^{I-1} \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') \mathsf{v}_{t+1}(\mathsf{x}_\mathsf{i} < \mathsf{x}' < \mathsf{x}_{\mathsf{i}+1}) h(0,\mathsf{x}') \frac{\phi\left(h(0,\mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}' \\
&+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}}\right)\right) \left(-\eta w_{t+1}\right).
\end{aligned}
$$

E.3 Solving for Integrals

Let us collect the approximated equations we defined so far.

$$
\begin{aligned}
\mathsf{v}_t(\mathsf{x}) =& \Pi_t(\mathsf{x}) - \Pi_t(0) \\
&+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \hat{s}_{t+1}(\mathsf{x}') \mathsf{v}_{t+1}(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) (\phi\left(h(\mathsf{x},\mathsf{x}')\right) - \phi\left(h(0,\mathsf{x}')\right)) d\mathsf{x}' \\
&+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(-\eta w_{t+1}\right) \int_{-1}^{1} \hat{s}_{t+1}(\mathsf{x}') (\phi\left(h(\mathsf{x},\mathsf{x}')\right) - \phi\left(h(0,\mathsf{x}')\right)) d\mathsf{x}',
\end{aligned} \tag{43}
$$

$$
\begin{aligned}
0 = \Pi_t'(0) + \Lambda_{t+1} \int_{-1}^{1} \hat{s}_{t+1} \mathsf{v}_{t+1}(\mathsf{x}') h(0, \mathsf{x}') \frac{\phi\left(h(0, \mathsf{x}')\right)}{\sigma_{t+1}} d\mathsf{x}' \\
+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi\left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}} \right) - \phi\left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}} \right) \right) (-\eta w_{t+1}) \,,
\end{aligned} \tag{44}
$$

$$
\mathsf{g}^c_t(\mathsf{x}) = \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \frac{\hat{s}_{t-1}(\mathsf{x}')}{\sigma_t} \mathsf{g}^c_{t-1}(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) \phi\left(h(\mathsf{x}', \mathsf{x})\right) \, d\mathsf{x}' + \frac{1}{\sigma_t} \mathsf{g}^0_{t-1} \phi\left(h(0, \mathsf{x})\right), \tag{45}
$$

$$
\mathsf{g}^0_t = 1 - \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \mathsf{g}^c_t(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) \hat{s}_t(\mathsf{x}) d\mathsf{x}, \tag{46}
$$

$$
e^{p^*_t(\epsilon-1)} = \sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} e^{(\mathsf{x})(1-\epsilon)} \mathsf{g}^c_t(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) \hat{s}_t(\mathsf{x}) d\mathsf{x} + \mathsf{g}^0_t, \tag{47}
$$

$$
N_t = \frac{C_t}{A_t} e^{p^*_t(-\epsilon)} \left(\sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} e^{\mathsf{x}(-\epsilon)} \mathsf{g}^c_t(\mathsf{x}_{i-1} < \mathsf{x} < \mathsf{x}_{i+1}) \hat{s}_t(\mathsf{x}) d\mathsf{x} + \mathsf{g}^0_{t-1} \right) + \eta \mathsf{g}^0_{t-1}. \tag{48}
$$

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 20 illustrates this graphically for the integral in the equation for $\mathsf{g}^c_t(x)$ (45).

We now determine the solution of those integrals, equation by equation. Given the coefficients of the affine functions, which depend on the values of $\mathsf{v}_{t+1}(\mathsf{g}_{t-1})$ at the grid points x_i, we can then write the solutions as a function that is linear in the elements of the vector $\mathsf{v}_{t+1}(\mathsf{x}_i)$ $(\mathsf{g}_{t-1}(\mathsf{x}_i))$. We now explain this for the simple case of the integral in equation 46. The other equations require some more tedious algebra, which we conveniently executed using symbolic math and which we omit here for brevity, but are conceptually equivalent.

E.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

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

Applying this formula to equation (46), which defines the mass point at $x = 0$, and

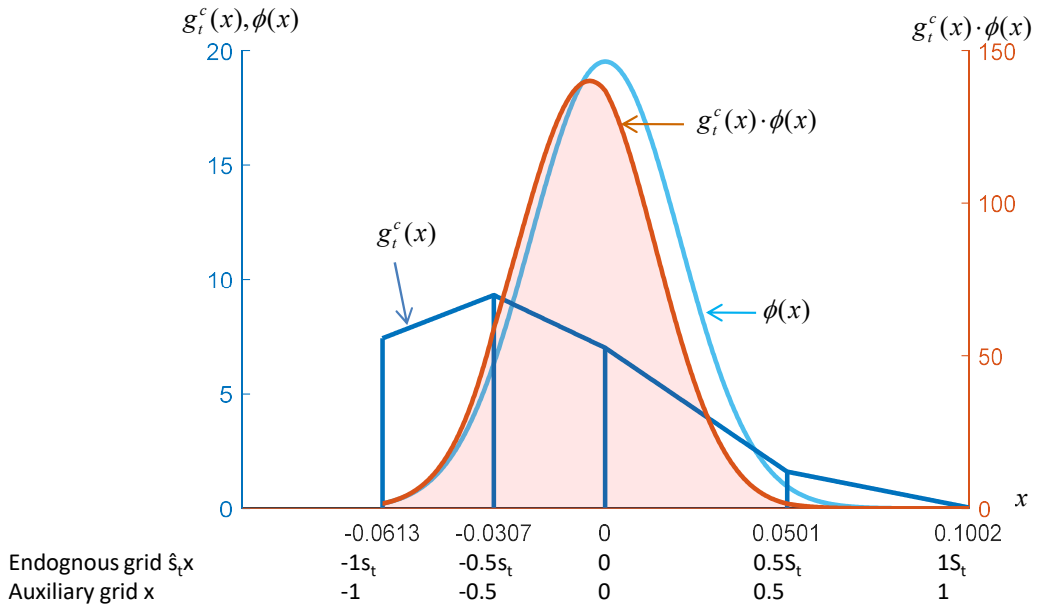


Figure 20: This figure schematically explains the analytical evaluation of integrals, given the linear approximation of the distribution and value functions. It shows the piece-wise linearly approximated distribution $\mathsf{g}_t^c(x)$ in blue, the normal pdf $\phi(x)$ in light blue and the product of the two $\mathsf{g}_t^c(x)\phi(x)$ in orange, where $x = \mathsf{x}\hat{s}_t$. The orange area thus corresponds to the term $\sum_{i=1}^{I-1} \int_{\mathsf{x}_i}^{\mathsf{x}_{i+1}} \frac{\hat{s}_{t-1}(\mathsf{x}')}{\sigma} \mathsf{g}_{t-1}^c(\mathsf{x}_i < \mathsf{x}' < \mathsf{x}_{i+1}) \phi\left(h(\mathsf{x}', \mathsf{x})\right) d\mathsf{x}'$ in equation (45).

.

re-arranging terms we get

$$\mathsf{g}_t^0 = 1 - \mathbf{e}_t^T \mathbf{g}_t^c \tag{49}$$

where $\mathbf{e}_t^T = [0.5,\ 1, \ldots, 1,\ 0.5]\Delta x$. Note that this formula corresponds to the trapezoid rule. The blue area in Figure 19 illustrates the application of the trapezoid rule.

E.3.2 Aggregate Price Index

By the same logic, the aggregate price index in (47) is computed as

$$e^{p_t^*(\epsilon-1)} = \sum_{i=1}^{I} (\mathsf{g}_t^c(\mathsf{x}_i) \mathbb{1}_{i \neq 1} d_{t,i,i-1,1-\varepsilon} + \mathsf{g}_t^c(\mathsf{x}_i) \mathbb{1}_{i \neq I} d_{t,i,i+1,1-\varepsilon}) + \mathsf{g}_t^0 \tag{50}$$

where

$$d_{t,i,j,\varepsilon} = \frac{\left(e^{(\epsilon)\mathsf{x}_i s_{t,i,j}} \left((\epsilon)\left(\mathsf{x}_i - \mathsf{x}_j\right) s_{t,i,j} - 1\right) + e^{(\epsilon)\mathsf{x}_j s_{t,i,j}}\right)}{(\epsilon)^2 \left(\mathsf{x}_i - \mathsf{x}_j\right) s_{t,i,j}}$$

and where $\mathbb{1}_{i\neq 1}, \mathbb{1}_{i\neq I}$ is an indicator function equal to 1 when i is different than 1 or I, that is whenever $\mathsf{g}_t^c(\mathsf{x}_i)$ is evaluated at the boundaries of the (S, s) band. It plays a similar role as the values 0.5 at the two ends of the vector $\mathbf{e}_t^T$ above.

Hence, we can re-write equation (50) in matrix form as

$$e^{p_t^*(\epsilon-1)} = \mathbf{d}_{t,1-\varepsilon}^T \mathbf{g}_t^c + \mathsf{g}_t^0 \tag{51}$$

where $\mathbf{g}_t^c$ is the matrix form of the distribution function g_t^c and where the vector $\mathbf{d}_{t,1-\varepsilon}$ is

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

E.3.3 Labor Market

Following the previous subsection, the labor market condition (48) is computed as

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

which we re-write in matrix form as

$$N_t = \frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \left(\mathbf{d}_{t,-\varepsilon}^T \mathbf{g}_t^c + \mathsf{g}_{t-1}^0 \right) + \eta \mathsf{g}_{t-1}^0. \tag{52}$$

E.3.4 Distribution

Once we have evaluated the integrals, the distribution function in (45) can be written as:

$$\mathsf{g}_t^c(\mathsf{x}_j) = \sum_{i=1}^{I-1} \frac{1}{2\sqrt{2\pi}} \mathsf{g}_{t-1}^c(\mathsf{x}_i) \left[\mathbb{1}_{i\neq 1} f_{t,i,i-1,j} + \mathbb{1}_{i\neq I} f_{t,i,i+1,j}\right] + \frac{1}{\sigma_t} \mathsf{g}_{t-1}^0 \phi \left(\frac{-\hat{s}_{t,j}\mathsf{x}_j - \pi_t^*}{\sigma_t} \right) \tag{53}$$

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

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

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

For compactness, define

$$\begin{aligned}
\mathbf{g}_t^c &\equiv \left[\mathsf{g}_t^c(\mathsf{x}_j)\right]_{j=1}^{I} \\
\mathbf{F}_t &\equiv \left[\sum_{i=1}^{I-1} \frac{1}{2\sqrt{2\pi}} \left[\mathbb{1}_{i\neq 1} f_{t,i,i-1,j} + \mathbb{1}_{i\neq I} f_{t,i,i+1,j}\right]\right]_{i=1,j=1}^{I,I} \\
\mathbf{f}_t &\equiv \left[\frac{1}{\sigma_t}\phi\left(\frac{-\hat{s}_t \mathsf{x}_j - \pi_t^*}{\sigma_t}\right)\right]_{j=1}^{I}
\end{aligned}$$

where $\mathbf{g}_t^c$ and $\mathbf{f}_t$ vectors with the probability mass function and the scaled and shifted normal distribution, respectively, $\mathbf{F}_t$ is a matrix that captures the idiosyncratic transitions due to firm-level quality shocks. Thus, equation 53 can be represented in matrix form as

$$\mathbf{g}_t^c = \mathbf{F}_t \mathbf{g}_{t-1}^c + \mathbf{f}_t \mathbf{g}_{t-1}^0. \tag{54}$$

E.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 $\mathsf{x}_{i_0} = 0$), the value function 43 can be written as

$$\begin{aligned}
\mathsf{v}_t(\mathsf{x}_j) = \; & \Pi_{j,t} - \Pi_{j,t}(0) \\
& + \Lambda_{t,t+1} \sum_{i=1}^{I} \frac{1}{2\sqrt{2\pi}} \mathsf{v}_{t+1}(\mathsf{x}_i) \left(\mathbb{1}_{i\neq 1}(a_{t,i,i-1,j} - a_{t,i_0,i_0-1,j}) + \mathbb{1}_{i\neq I}(a_{t,i,i+1,j} - a_{t,i_0,i_0+1,j})\right) \\
& + \Lambda_{t,t+1} \left(-\eta w_{t+1}\right) \left(\Phi\left(\frac{\mathcal{P}_{t+1,j,I}}{\sigma_{t+1}}\right) - \Phi\left(\frac{\mathcal{P}_{t+1,j,1}}{\sigma_{t+1}}\right) - \Phi\left(\frac{\mathcal{P}_{t+1,i_0,I}}{\sigma_{t+1}}\right) + \Phi\left(\frac{\mathcal{P}_{t+1,i_0,1}}{\sigma_{t+1}}\right)\right)
\end{aligned} \tag{55}$$

where

$$a_{t,i,\bar{i},j} = \frac{\sqrt{2\pi}\left(\mathcal{P}_{t+1,j,\bar{i}}\right)\left(\operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,\bar{i}}}{\sqrt{2}\sigma_{t+1}}\right) - \operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,i}}{\sqrt{2}\sigma_{t+1}}\right)\right) + 2\sigma_{t+1}\left(\exp\left(-\frac{\left(\mathcal{P}_{t+1,j,\bar{i}}\right)^2}{2\sigma_{t+1}^2}\right) - \exp\left(-\frac{\left(\mathcal{P}_{t+1,j,i}\right)^2}{2\sigma_{t+1}^2}\right)\right)}{\left|x_i \hat{s}_{t+1,i} - x_{\bar{i}} \hat{s}_{t+1,\bar{i}}\right|}. \tag{56}$$

For compactness, let us define

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

where $\mathbf{v}_t$ and $\mathbf{b}_{t+1}$ are vectors that evaluate the value function and the adjustment prob-

ability at different grid points, $\mathbf{\Pi}_t$ 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 (55) can be represented in matrix form as

$$\mathbf{v}_t = \mathbf{\Pi}_t + [\mathbf{A}_t \mathbf{v}_{t+1} - \mathbf{b}_{t+1} \eta w_{t+1}]. \tag{57}$$

E.3.6 Optimality condition for reset price

After evaluating the integral, we can write the optimality condition in (44) as

$$\begin{aligned} 0 = \Pi_t'(0) &+ \Lambda_{t,t+1} \sum_{i=1}^{I} \mathsf{v}_{t+1}(\mathsf{x}_i) \frac{1}{2} \left(\mathbb{1}_{i \neq 1} c_{t,i,i-1,i_0} + \mathbb{1}_{i \neq I} c_{t,i,i+1,i_0} \right) \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi \left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}} \right) - \phi \left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}} \right) \right) (-\eta w_{t+1}) \end{aligned} \tag{58}$$

where

$$c_{t,i,\bar{i},j} = \frac{\operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,i}}{\sqrt{2}\sigma} \right) - \operatorname{erf}\left(\frac{\mathcal{P}_{t+1,j,\bar{i}}}{\sqrt{2}\sigma} \right)}{x_i s_{t+1,i} - x_{\bar{i}} s_{t+1,\bar{i}}} - \frac{\sqrt{\frac{2}{\pi}} \exp\left(-\frac{(\mathcal{P}_{t+1,j,i})^2}{2\sigma^2} \right)}{\sigma}. \tag{59}$$

We can write this equation using matrix notation:

$$\begin{aligned} 0 = \Pi_t'(0) &+ \mathbf{c}_{t+1}^T \mathbf{v}_{t+1} \\ &+ \frac{\Lambda_{t,t+1}}{\sigma_{t+1}} \left(\phi \left(\frac{-S_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}} \right) - \phi \left(\frac{-s_{t+1} - \pi^*_{t+1}}{\sigma_{t+1}} \right) \right) (-\eta w_{t+1}) \end{aligned} \tag{60}$$

where

$$\mathbf{c}_{t+1} = \left[\Lambda_{t,t+1} \frac{1}{2} \left(\mathbb{1}_{i \neq 1} c_{t,i,i-1,i_0} + \mathbb{1}_{i \neq I} c_{t,i,i+1,i_0} \right) \right]_{i=1}^{I}. \tag{61}$$

E.4 Final equation system

Collecting the thus 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 planner's problem by the following finite dimensional planner's problem

$$\max_{\left\{ \mathbf{g}_t^c, \mathbf{g}_t^0, \mathbf{v}_t, C_t, w_t, p_t^*, s_t, S_t, \pi_t^* \right\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\gamma}}{1-\gamma} - \upsilon \left(\frac{C_t}{A_t} e^{p_t^*(-\epsilon)} \left(\mathbf{d}_{t,-\varepsilon}^T \mathbf{g}_t^c + \mathbf{g}_{t-1}^0 \right) + \eta \mathbf{g}_{t-1}^0 \right) \right)$$

subject to

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

Here the choice variables $\mathbf{v}_t$ and $\mathbf{g}_t^c$ 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.

E.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 $w = \upsilon C^{\gamma}$, and $R = (1+\pi)/\beta$ and $\pi = \pi^*$. We start with a guess for the real wage w, the optimal rest price p^*, and the boundaries of the (S, s) band s and S then:

1. Compute consumption $C = \left(\frac{w}{\upsilon}\right)^{1/\gamma}$.
2. Using $\pi = \pi^* = \overline{\pi}$, C and the 4 initial guesses, solve for that 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.

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

3. Compute the residuals of the 4 remaining equations

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

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