# Strike the Iron while it's Hot: Optimal Monetary Policy with (S,s) Pricing*


Peter Karadi[1], Anton Nakov[2], Galo Nuño[3], Ernesto Pastén[4], and Dominik Thaler[5]

[1,2,5]European Central Bank
[3]Bank for International Settlements and Bank of Spain
[4]Central Bank of Chile and Toulouse School of Economics
Emails: [1]peter.karadi@ecb.int, [2]anton.nakov@ecb.int, [3]galo.nuno@bde.es,
[4]ernesto.pasten@gmail.com, [5]dominik.thaler@ecb.int


April 25, 2024

PRELIMINARY AND INCOMPLETE – PLEASE DO NOT CIRCULATE


**Abstract**

We study the Ramsey optimal monetary policy in a variant of the Golosov and Lucas (2007) state-dependent pricing framework. The model captures the endogenous increase in the repricing rate after large shocks, as experienced during the recent inflation surge. A novel insight is that in response to large cost-push shocks optimal policy under commitment should strike down inflation and stabilize the repricing rate, countering the inflationary aspirations of firms and dampening state dependence. Along the optimal commitment path, the central bank benefits from a lower sacrifice ratio and achieves lower inflation. At the same time, when facing demand or efficiency shocks (e.g. TFP shocks), the optimal policy is to commit to full price stability, just like in the standard Calvo model with exogenous timing of price changes.


*The views expressed here are those of the authors, 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., 2023b). Concurrently, empirical evidence reveals marked time variations in the slope of the estimated Phillips curve, which characterizes the relationship between economic activity and inflation (Cerrato and Gitti, 2023). Traditional models of price setting (Calvo, 1983), which form the basis for optimal monetary policy analysis (Galí, 2008; Woodford, 2003), are inadequate to explain these observations. In contrast, state-dependent pricing models are well suited to capturing them. Among these models, the Golosov and Lucas (2007) menu-cost model has emerged as a benchmark for positive analysis. However, the normative aspects of this model have not received much attention. It is precisely this crucial gap that our paper aims to bridge.

Our analysis arrives at a novel insight: Ramsey optimal monetary policy in the Golosov and Lucas (2007) model should commit to quashing inflation and lean against changes in the repricing rate. This conclusion extends to scenarios with large cost-push shocks, which broadly capture the recent inflation experience. The optimal inflation path is one which balances price dispersion costs (which are mitigated by a more flexible price level) with reduced menu costs (which are lower with less flexible prices). Our findings thus extend prevailing views in New Keynesian economics and propose a more nuanced approach for central banks facing large-scale economic perturbations.

We show that the Golosov and Lucas (2007) menu-cost model can reproduce salient features of the recent inflation dynamics as a plausible response to a large cost-push shock. Most notably, the framework captures the sizable increase in the frequency of price changes, which endogenously increases the flexibility of the aggregate price level. The variation in frequency is important: a linearly-scaled small shock, which would imply a broadly unchanged frequency response, would underestimate the large shock's impact on inflation by over 20 percent. Additionally, it would also underestimate the magnitude of the interest-rate policy response and the strength of the accompanying downturn.

We also show that the variation in frequency causes state dependence in monetary policy. As the frequency increases, a policy change has an amplified impact on inflation and a mitigated impact on activity, reducing the sacrifice ratio of a policy tightening. In other words, the Phillips curve becomes non-linear, and its slope increases substantially with frequency, for a realistic range of frequency. In the case of large technology, preference, and cost-push shocks, the effectiveness of policy in stabilizing inflation increases exactly when the shock exerts an outsized relative influence on inflation. For realistic parameter values, however, an unchanged policy rule achieves worse stabilization properties after large shocks than after small ones. In other words, it pays off to be more aggressive in the fight against inflation after large shocks.

The normative part of the paper focuses on the Ramsey problem within (a variant of) the

Golosov and Lucas (2007) menu-cost model. It establishes five novel results. First, it shows that the welfare-maximizing long-run inflation rate is close to zero, but slightly positive, at around 0.3%. This is different from the canonical Calvo (1983) model, where the optimal inflation is exactly zero. In the state-dependent model, positive inflation reduces 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, which is the consequence of the asymmetry of the profit function: firms dislike negative price misalignments when the demand for their product is large, relative to positive misalignments when the demand is low.

Second, the optimal policy is time inconsistent and the magnitude of the time inconsistency is amplified by the state dependence of price setting. Similarly to the Calvo (1983) model, when the steady state is inefficient, monetary policy can generate a temporary boom by an unexpectedly easy policy (Galí, 2008). The magnitude of time inconsistency is amplified in the state-dependent model by (i) higher steady-state inefficiency due to idiosyncratic price dispersion and (ii) the impact of the unexpected policy easing on the flexibility of price level through affecting the repricing rate.

Concerning the optimal dynamic response to shocks, our analysis finds that optimal monetary policy should commit to stabilizing the repricing rate. This commitment serves the purpose of counteracting the inflationary aspirations of firms and mitigates the state-dependence of firms' decisions. This is true for various types of shocks, which we overview now as results three to five.

Third, the optimal response to efficient (technology, preference) shocks is characterized by "divine coincidence" (Blanchard and Galí, 2007). In other words, it is optimal to lean against frequency, fully stabilize inflation, and close the output gap as a response to these shocks.

Fourth, the optimal response to small cost-push shocks is also similar to, but not exactly the same as, the Calvo (1983) model. In particular, the relationship between the price level and the output gap is characterized by a linear target rule, with a slope that is close to, but slightly steeper than in the linearized Calvo (1983) model. The result implies that even though the sacrifice ratio in the Golosov and Lucas (2007) model is low due to excess price flexibility caused by the endogenous selection of large price changes, it is almost fully offset by the lower welfare costs of inflation due to the more flexible prices, which limits the extent of relative-price distortions.

Fifth, it is optimal for the central bank to lean against frequency increases and stabilize inflation after large cost-push shocks. Along the path of optimal commitment, we observe tangible benefits for the central bank. There is a discernible reduction in the sacrifice ratio, contributing to lower inflation. Even though higher frequency also reduces the price dispersion costs of inflation, the reduced relative importance of inflation stabilization is insufficient to offset the gains from disinflation in the presence of a reduced sacrifice ratio. This nuanced understanding of the optimal policy stands in stark contrast to the outcomes predicted by the standard New Keynesian model with exogenous timing of price adjustment, which fails to capture the non-linear dynamics revealed by our analysis.

Overall, our findings highlight the significance of a proactive approach by the central bank in the face of large shocks. By committing to policies that suppress inflation and stabilize the repricing rate, the central bank can foster a more favorable macroeconomic outcome.

**Literature Review** The macroeconomic framework of the paper generates a non-linear Phillips curve, which has received renewed empirical support following the recent inflation surge (Phillips, 1958; Benigno and Eggertsson, 2023; Cerrato and Gitti, 2023). The non-linear Phillips curve in our paper is microfounded by price-adjustment frictions in a heterogeneous-firm dynamic stochastic general equilibrium framework. In particular, the paper assumes the presence of small, fixed (menu) costs of price adjustments (Barro, 1972; Sheshinski and Weiss, 1977; Caplin and Spulber, 1987) following closely the quantitative model of Golosov and Lucas (2007). The framework provided foundations for subsequent work aimed at establishing a connection between stylized facts of price-setting in microdata and monetary non-neutrality (Dotsey et al., 1999; Gertler and Leahy, 2008; Nakamura and Steinsson, 2008; Midrigan, 2011; Alvarez et al., 2016; Auclert et al., 2022). The framework naturally gives rise to a non-linear Phillips curve after large shocks and high inflation rates through the endogenous adjustment in the frequency of price changes. Higher frequency makes the price level more flexible leading to a variable inflation-output tradeoff (Karadi and Reiff, 2019; Alvarez and Neumeyer, 2020; Costain et al., 2022; Alexandrov, 2020; Blanco et al., 2022). The question of optimal policy in the framework, however, has received scant attention in the literature. This is where the main contribution of the paper lies.

To the best of our knowledge, our paper is the first to solve optimal policy in the Golosov and Lucas (2007) model. Previous research has addressed optimal policy in the time-dependent Calvo (1983) model (Woodford, 2003; Galí, 2008) and restricted attention to the question of optimal long-run inflation rate in menu cost economies (Adam and Weber, 2019; Blanco, 2021). One reason for the lack of direct predecessors is the computational difficulty due to the "curse of dimensionality". This problem arises because a menu cost model needs to keep track of the whole distribution of the gap between actual and desired prices to predict the way that prices respond to shocks and monetary policy. This distribution is not degenerate because of idiosyncratic shocks, which help the model to reproduce key features in microdata. We tackle this difficulty by solving the problem under perfect foresight over the impulse-response space. Differently from related recent work solving heterogeneous agent models over the impulse-response space (Boppart et al., 2018; Auclert et al., 2022), our methodology does not require linearization. Therefore it suitable for analysing non-linear responses to large shocks. To solve for optimal policy, we need a high level of accuracy that requires a careful optimization of the solution algorithm. The framework allows us to numerically assess optimal responses to efficient (demand and productivity) and inefficient (cost-push and idiosyncratic volatility) shocks and characterize the non-linear target rule, which generalizes its counterpart in the linear-quadratic Calvo (1983) framework.

# 2 Model

Our model is a variant of the Golosov and Lucas (2007) state-dependent pricing framework. The economy consists of a representative household, a unit mass of monopolistic producers facing fixed menu costs to update their prices, and a central bank that sets the nominal interest rate.

## 2.1 Households

There is a representative household that saves in one-period bonds whose nominal value is denoted by $B_t$. Bonds are in zero net supply. Workers supply labor hours $N_t$. 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

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

where $T_t$ are lump-sum taxes, $D_t$ are the dividends received lump sum from monopolistic producers, $i_t$ is the net nominal interest rate, $W_t$ is the nominal wage. Aggregate consumption $C_t$ is

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

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

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

where $\varepsilon_t$ is an $i.i.d$ Gaussian innovation. The demand for product $i$ is,

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

and the aggregate price index is

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

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, and the Euler equation,

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

where $\pi_t \equiv \log(P_t/P_{t-1})$ is the inflation rate and $\Lambda_{t,t+1}$ is the stochastic discount factor defined as

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

## 2.2 Monopolistic producers

Production of good $i$ is

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

where $N_t(i)$ is the labor used, $A_t(i)$ is the quality, and $A_t$ is aggregate productivity.

The nominal profit function is

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

where $\tau_t$ is an employment subsidy financed by lump-sum taxes and we have used the equilibrium condition $Y_t(i) = C_t(i)$. The real profit function thus is

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

where $w_t$ is the real wage and

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

is the quality-adjusted log relative price. We assume $p_t(i)$ lies within a compact set $\Gamma = [\underline{p}, \overline{p}]$.

When prices do not change in nominal terms, $p_t(i)$ evolves according to

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

From now on, we drop the index $i$ 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 $\eta$ (in labor units). The optimal reset price 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 Bellman equation

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

which is given by the momentary 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

$$\lambda_t(p) = I[L_t(p) > 0]$$

where $I[\cdot]$ is the indicator function, and

$$L_t(p) \equiv \max_{p'} V_t\left(p'\right) - \eta w_t/A_t - V_t(p)$$

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 of the following form:

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

where $y_t^e$ is the efficient-level of output, defined in Section 4, and $\phi_\pi > 1$ and $\phi_y$ 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:

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

Individual firms' labor demand aggregates up to

$$N_t = \frac{C_t}{A_t} \int e^{p(-\epsilon)} g_t(p)\, dp + \frac{\eta}{A_t} \int \lambda_t(p - \sigma_t \varepsilon - \pi_t) g_{t-1}(p) dp, \tag{13}$$

such that the total number of hours worked equals the total use of labor for production (the first term on the right-hand side of the equation above) 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.

$$g_t(p) = \begin{cases} (1-\lambda_t(p)) \int g_{t-1}(p + \sigma_t \varepsilon + \pi_t) d\xi(\varepsilon), & \text{if } p \neq p_t^*, \\ (1-\lambda_t(p_t^*)) \int g_{t-1}(p_t^* + \sigma_t \varepsilon + \pi_t) d\xi(\varepsilon) + \int_{\underline{p}}^{\overline{p}} \lambda_t(\tilde{p}) \left( \int g_{t-1}(\tilde{p} + \sigma_t \varepsilon + \pi_t) d\xi(\varepsilon) \right) d\tilde{p}, & \text{if } p = p_t^*. \end{cases}$$

This expression for $g_t(p)$ captures that, with probability $1 - \lambda_t$, the end of the period's density is just taken from the density of prices in period $t-1$, once adjusted for quality and inflation (where $\xi(\varepsilon)$ is the Gaussian distribution of the innovation $\varepsilon$). In turn, if the actual price equals the optimal reset price, $g_t(p)$ integrates the mass of all adjusting prices and, with probability $\lambda_t$, allocates it to the optimal reset price $p_t^*$.

## 2.5 Normalization

In order to achieve high accuracy in the computations, we find it convenient to recast the problem in terms of the distance 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 $x_t\,(i)$ 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 change in prices, $x_t$ always jumps back to zero, equivalent to prices jumping to $p^*$.

Momentary profits can then be expressed as

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

and the Bellman equation can thus be re-written as

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

Notice that we have the additional condition $V'_{t+1}(0) = 0$.

The law of motion of the density then is

$$g_t(x) = \begin{cases} (1-\lambda_t(x))\int g_{t-1}(x + \sigma_t\varepsilon + \pi_t^*)d\xi(\varepsilon), & \text{if } x \neq 0, \\ (1-\lambda_t(0))\int g_{t-1}(\sigma_t\varepsilon + \pi_t^*)d\xi(\varepsilon) + \int \lambda_t(\tilde{x})\left(\int g_{t-1}(\tilde{x} + \sigma_t\varepsilon + \pi_t^*)d\xi(\varepsilon)\right)d\tilde{x}, & \text{if } x = 0. \end{cases}$$

Finally, the aggregate price index can be expressed as

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

### 2.6 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 $\varepsilon_{A,t} \sim N(0, \sigma_A^2)$, and $\rho_A \in [0,1]$ and $\sigma_A$ are parameters. Likewise, we assume that the (lump-sum tax-financed) employment subsidy $\tau_t$ follows an autoregressive process. We use this to capture temporary *cost-push* shocks.

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

where $\tau$ is the steady-state employment subsidy, $\varepsilon_{\tau,t} \sim N(0, \sigma_\tau^2)$, and $\rho_\tau \in [0,1]$ and $\sigma_\tau$ are parameters. We also consider autoregressive shocks to the idiosyncratic volatility of quality shocks:

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

where $\sigma$ is the steady-state volatility, $\varepsilon_{\sigma,t} \sim N(0, \sigma_\sigma^2)$, and $\rho_\sigma \in [0,1]$ and $\sigma_\sigma$ are also parameters. Finally, in Section 4 we assume i.i.d. shocks to the Taylor rule (4.2), $\varepsilon_{r,t} \sim N(0, \sigma_r^2)$, where $\sigma_r$ is a parameter.

## 2.7 Ramsey problem

The central bank maximizes households' welfare under commitment (see, Galí, 2008, for instance). The problem of the central bank is

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

subject to

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

where $\phi(\cdot)$ is the probability density function of a normal random variable and $s$ and $S$ are the endogenous boundaries of the inaction region for $x$. The problem implies that the central bank selects the path for all equilibrium variables subject to all the competitive equilibrium conditions. 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 behaviour does not lead to a discontinuity in aggregates. Furthermore, note that both $V(x)$ and that $g^c(x)$ are continuously differentiable over the relevant range $(s, S)$.

For small shocks, we consider a first-order approximation around the Ramsey optimal steady state. As discussed by Boppart et al. (2018), a first-order approximation to the stochastic problem can be computed by analyzing the impulse response to a one-off shock under perfect foresight (an MIT shock). This is what we are doing for small shocks. For large shocks, the interpretation is similar to that in Cavallo et al. (2023a), namely that the economy in the deterministic steady state (under the Ramsey policy) is hit by an unexpected once-and-for-all large shock.

# 3 Computational method and calibration

In this section, we describe the computational method and our baseline calibration.

## 3.1 Computational method

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. There are a number of proposals in the literature to deal with this problem. 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. 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).

Here we propose a new algorithm similar to that in González et al. (2024) but applied to discrete time. The general idea is similar to that of Nuño and Thomas (2022), Smirnov (2022), and Dávila and Schaab (2022), but instead of determining the FOCs for the planner's continuous space problem, we first discretize the planner's objective and constraints (the private equilibrium conditions) using finite differences. This transforms 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 good enough to capture the higher-order effects of policy. One challenge in our particular problem where the private equilibrium conditions include discrete choices, is that a simple discrete-state approximations typically may fail. Therefore, we approximate the distribution and value function not by a vector on a predetermined grid, but by a piecewise linear function over an endogenous grid that includes the the two boundaries of the inaction region (points $x = s$ and $x = S$) and the optimal reset price ($x = 0$). 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 A the Bellman equation can thus be approximated over a grid of price gaps $x$ as

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

where $\mathbf{V}_t$ and $\mathbf{b}_t$ are vectors with the value function and the 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 price updating. Similarly, the law of motion of the density for $x \neq 0$ is

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

where $\mathbf{g}_t$ and $\mathbf{f}_t$ are vectors with the density and the scale and shifted normal distribution, respectively, $\mathbf{F}_t$ is a matrix that captures the idiosyncratic transitions due to firm-level quality shocks, and

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

is the mass point at $x = 0$. Here $\mathbf{e}_t$ is a vector of ones in the range $(s, S)$

Second, we find the planner's FOCs by symbolic differentiation. This delivers a large-dimensional system of difference equations.

Third, 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 steady-state equilibrium variables. We then combine this function with the central bank FOCs. As the system is linear in the Lagrange multipliers, the solution just amounts to finding the zero of a nonlinear function of the initial variable (i.e., inflation), which can be easily implemented using the Newton method.

Fourth, we solve the system of difference equations non-linearly in the sequence space 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.

### 3.2 Calibration

Table 1 shows our baseline calibration. TBC

Table 1: Parameter values

| | | | |
|---|---|---|---|
| | | Households | |
| $\beta$ | $0.96^{1/12}$ | Discount rate | Golosov and Lucas (2007) |
| $\epsilon$ | 7 | Elasticity of substitution | Golosov and Lucas (2007) |
| $\gamma$ | 1 | Risk aversion parameter | Midrigan (2011) |
| $\upsilon$ | 1 | Utility weight on labor | Set so that $w = C$ |
| | | Price setting targets | |
| Freq. | 8.7% | Frequency of price changes | Nakamura and Steinsson (2008) |
| Size | 8.5% | Absolute size of price changes | Nakamura and Steinsson (2008) |
| | | Monetary policy | |
| $\phi_\pi$ | 1.5 | Inflation coefficient in Taylor rule | Taylor (1993) |
| $\phi_y$ | 0.125 | Output gap coefficient in Taylor rule | Taylor (1993) |
| $\rho_r$ | $0.75^{1/3}$ | Smoothing coefficient | |
| | | Shocks | |
| $\rho_A$ | $0.95^{1/3}$ | Persistence of the TFP shock | Smets and Wouters (2007) |
| $\rho_\tau$ | $0.9^{1/3}$ | Persistence of the cost-push shock | Smets and Wouters (2007) |
| $\rho_\sigma$ | $0.75^{1/3}$ | Persistence of the dispersion shock | Smets and Wouters (2007) |

## 4 Implications for monetary policy design

This section presents novel considerations for monetary policy design arising from the menu cost model in the presence of large shocks. The key distinguishing feature of the framework is the endogenous evolution of the frequency of price changes, which leads to a non-linear Phillips curve and, consequently, a state-dependent variation in the inflation-output trade-off. This is different from the time-dependent Calvo (1983) model, where the frequency is exogenously given, the Phillips curve is approximately linear and the inflation-output trade-off is approximately state-invariant.

## 4.1 Non-linear responses to large shocks

First we illustrate the non-linearity of the menu cost model. We show that this non-linearity is quantitatively relevant in realistic regions of the state space, where the frequency of price changes features realistic levels, such as those observed during the 2022-2023 inflation surge. In particular, we show impulse responses to (i) large cost-push ($\tau_t$), (ii)technology ($A_t$), (iii) relative price dispersion ($\sigma_t$), and (iv) monetary policy ($i_t$) shocks and contrast them to responses to their small counterpart.

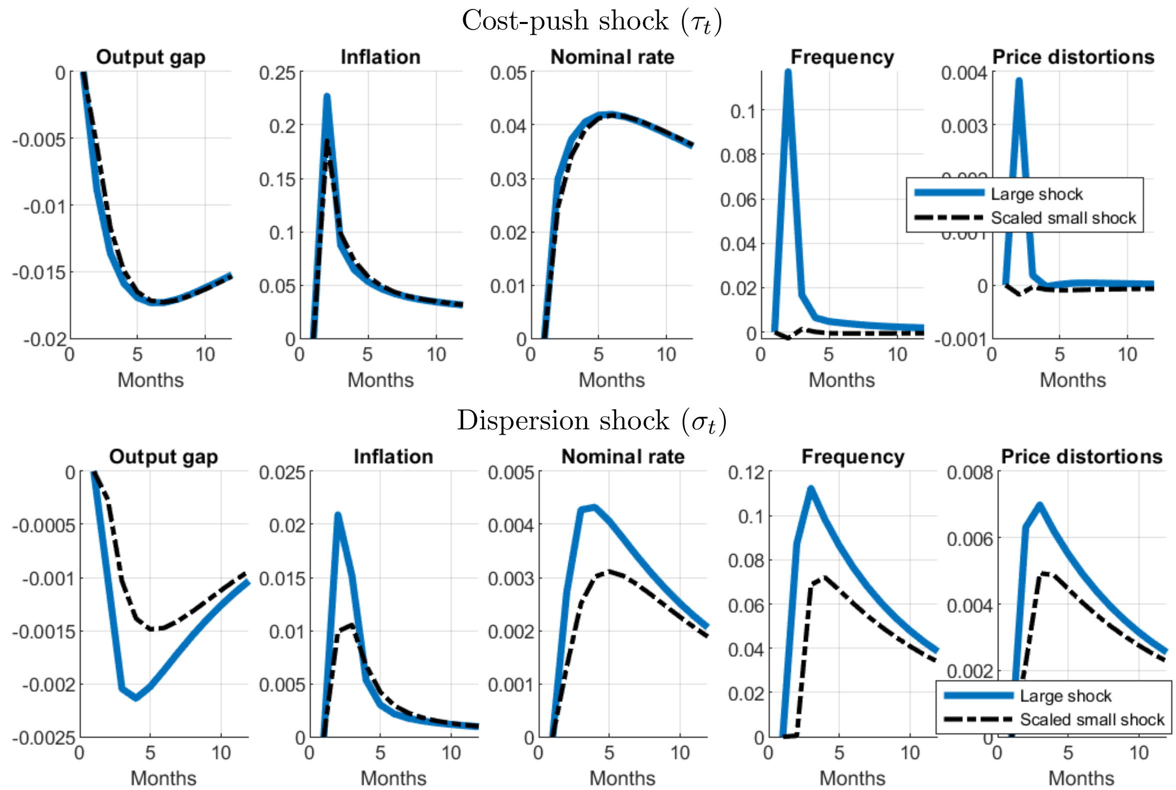



Figure 1: Impulse responses to a large and a (linearly-scaled) small cost-push shock and an idiosyncratic volatility shock

**Cost-push shocks**. The solid blue lines on the first row of Figure 1 show the responses to a large cost-push shock. The shock is implemented as a persistent decline in the firms' employment subsidy $\tau_t$. The shock size is calibrated to generate a 20% frequency on impact (a frequency increase of $20\% - 8.7\% = 11.3\%$), 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. The shock captures some realistic features of the recent inflation surge. First, it generates a large increase in inflation, characterized by an initial temporary spike followed by a period of persistent inflation. Second, the nominal interest rate follows a hump-shaped pattern. It stays substantially below the inflation rate in the first few months, causing negative real interest rates temporarily on impact. Third, the shock causes a mild downturn as the output gap decreases.

In contrast, the black dashed lines show the responses to a *small* cost-push shock, scaled linearly by the relative size of the shocks. The small shock is a 25 basis point decrease in the annualized employment subsidy. The difference between the figures illustrates the non-linearity 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. The non-linearity in 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).

In parallel with the frequency increase, the pricing distortions also increase. From equation 13 we get that

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

where the first term in the right-hand side is the the quality-adjusted relative-price dispersion and the second term are the menu costs. These two distortions drive a welfare-reducing wedge between available resources $A_t N_t$ and final consumption (expressed in labor units). Nonetheless, despite the higher pricing distortions, the output gap follows almost identical paths in response to large and small shocks. The only way to reconcile a large inflation response with a similar output gap response is that the slope of the Phillip is nonlinear and becomes higher with the shock size, a point that we develop further below.

**Price dispersion shocks**. The second row of Figure 1 shows impulse responses to a shock to the volatility of the idiosyncratic quality shocks ($\sigma_t$). 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 non-linearity of the model with respect to this shock. Nonetheless, notice how thew impact of the shock on inflation and output is one order of magnitude lower than the cost-push shock.

Notice how the impulse response to a large cost-push shock above can be roughly be described as the response to a (linearly scaled) small shock that does not change the frequency of price adjustments plus an endogenous response to the increase in frequency, which resembles a dispersion shock. The nonlinearity is thus mainly the consequence of the fact that small shocks do not affect the frequency of repricing, whereas large shocks do. That is an important conclusion that will carry over to other shocks.

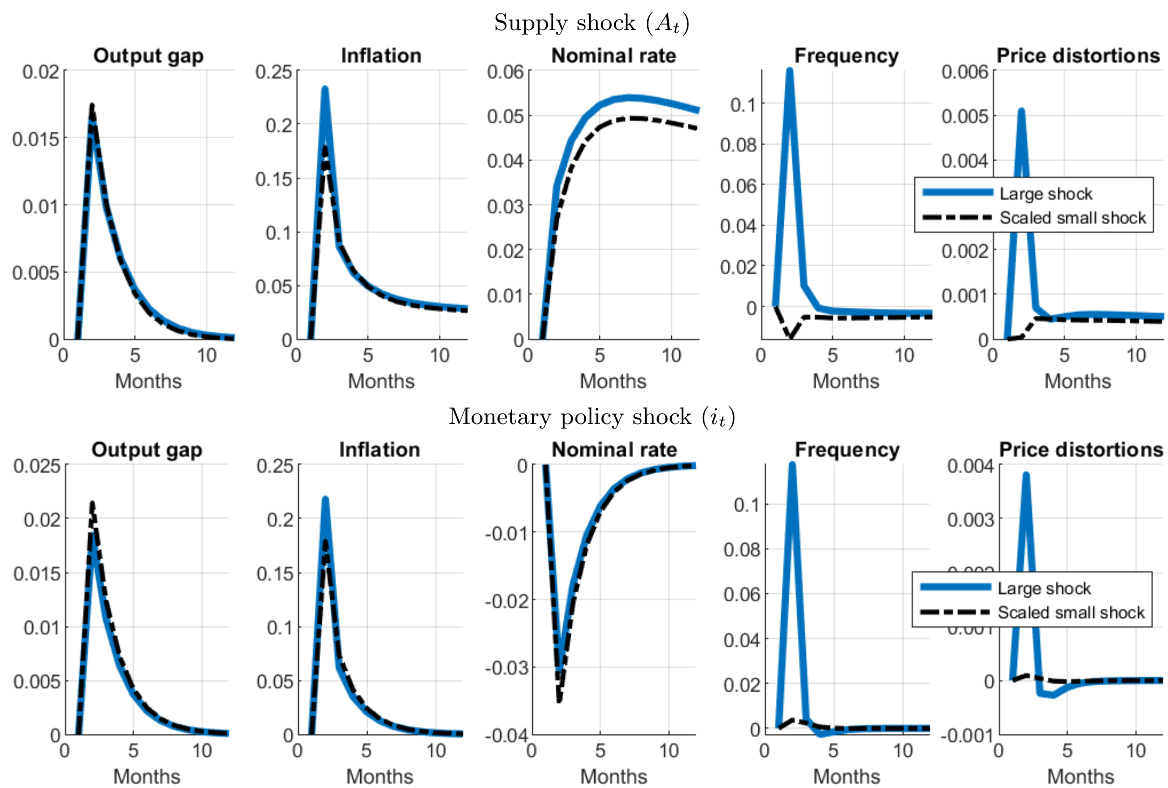



Figure 2: Responses to large and small productivity and monetary policy shocks

**TFP and monetary policy shocks**. Figure 2 shows analogous impulse responses to large aggregate shocks (blue solid line) and linearly-scaled small aggregate shocks (black dashed line). The first row shows responses to a productivity shock ($A_t$), and the second row shows responses to a monetary policy shock ($i_t$). The large shocks are calibrated to generate a 20% frequency on impact. Similar to the case of cost-push shocks, the response to the large shocks can be roughly decomposed into the response to a small shock in which the frequency of repricing remains constant plus a dispersion shock, which increases inflation and price distortions.

## 4.2 The non-linear Phillips curve

In order to provide some intuition about the working of this model, here we derive a 3-equation 'macro-block' of the model. This helps us to contrast it to the well-understood 3-equation New Keynesian model.

**New Keynesian IS curve and the Taylor rule**. The first equation of the macro block is the familiar new Keynesian IS curve. It can be derived from the Euler equation (7):

$$\tilde{y}_t^e = -\frac{1}{\gamma}\left(i_t - \pi_{t+1} - r_t^e\right) + \tilde{y}_{t+1}^e. \tag{15}$$

where the efficient output gap ($\tilde{y}_t^e \equiv \log Y_t/Y_t^e$) measures the (log) deviation of output from its efficient level, their difference expressed as its deviation from the steady state output gap, and $r_t^e$ is the efficient real interest rate.[1] The efficient output level and interest rates are those prevailing in a counterfactual efficient allocation, which is the solution to the social planning problem (See Appendix B). Under perfect foresight, IS equation (15) is globally log-linear.[2]

The second equation of the macro block, the Taylor rule (equation 4.2) is similarly globally log-linear. We repeate here for convenience, .

$$i_t = \rho_r i_{t-1} + (1-\rho_r)\left[\phi_\pi(\pi_t) + \phi_y(y_t - y_t^e)\right],$$

The global log-linearity of these equations mean that they are not responsible for the non-linear evolution of the variables of interest, like the output gap, inflation, interest rates, which are all expressed in log terms. The non-linearity, instead, comes from the third equation, the Phillips curve, which expresses the relationship between inflation and the output gap.

**Linearized New Keynesian Phillips Curve**. For small shocks, the menu-cost model implies a relationship between inflation and output gap that is characterized by a log-linear new Keynesian Phillips curve (NKPC) (Auclert et al., 2022). It takes the form

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa \tilde{y}_t^e + u_t \tag{16}$$

The expression relates inflation $\pi_t$ to expected future inflation $\mathbb{E}_t[\pi_{t+1}]$, the output gap, and a cost-push shock $u_t$ term. The parameter $\kappa$ determines the magnitude of the inflation-output trade-off.[3]

[1] Steady state output gap is negative in our model. The reason is that due to idiosyncratic shocks and price adjustment frictions, steady state output is subject to positive pricing distortions even under zero trend inflation and no aggregate shocks.

[2] We express the relationship in perfect foresight to simplify the algebra and to be consistent with our computational implementation. The algebra becomes easier because we can ignore covariance terms. These terms are quantitatively small in case market participants expect small shocks in the future.

[3] Auclert et al. (2022) show that the parameter in the Golosov and Lucas (2007) model is higher than in a Calvo (1983) model calibrated to the same steady-state frequency of price changes. The difference is the consequence of the so-called ´selection effect´. While in Calvo (1983) model it is random which firm can adjust

Here $\kappa$ captures the slope of the Phillips curve, that is, the impact on inflation of a change in the output gap. The slope is constant in the log-linearized formula, and therefore does not depend on the size of the shock or the state of the economy. The relationship between inflation and output gap, however, is valid only for small shocks in a neighborhood of the steady state. Our focus in this section is to see whether the relationship generalizes in the presence of large aggregate shocks and how the Phillips curve slope described by $\kappa(\Theta)$ varies across the state space $(\Theta)$.

**Non-linear New Keynesian Phillips Curve**. The relationship between inflation and output is necessarily non-linear in a menu-cost model. To see why, consider a series of monetary policy shocks of an ever-increasing size. If the shocks are small, the frequency of price changes is low: firms update prices occasionally, and the economy behaves similarly to a standard sticky-price economy à la Calvo (1983). When the shocks are large, however, the frequency of price changes increases: firms update prices more often thus reducing the degree of price stickiness. For a sufficiently large shock, the frequency reaches 100%, that is, prices become completely flexible, and monetary policy loses its ability to influence activity. The output gap falls then below zero due to the negative effects of pricing distortions from high inflation.

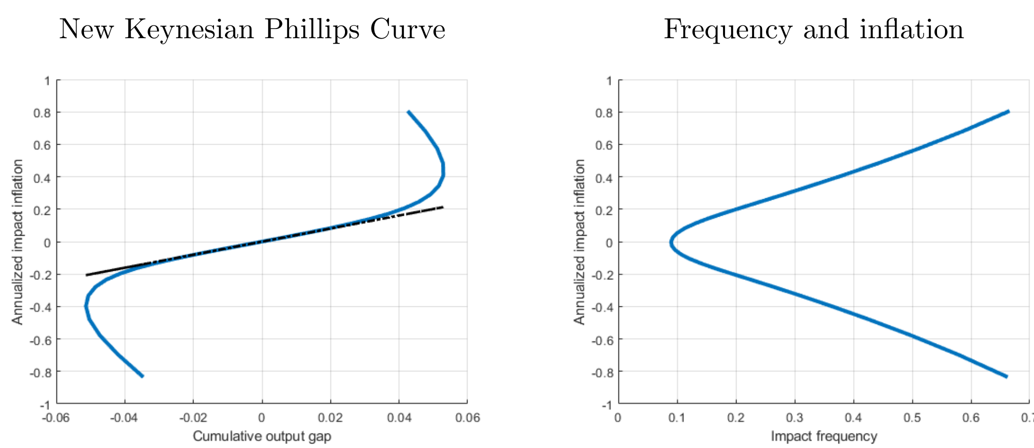



Figure 3: Inflation-output trade-off

The non-linear relationship between inflation and output gap is complex. It is state dependent, shock-size dependent and shock-sign dependent as well. It is state dependent because it depends potentially on all state variables of the model, including the initial relative price distribution, as well as the history of past aggregate shocks. It is shock-size dependent, because, as argued before, large shocks have larger relative impact on inflation relative to the output gap. And it is shock-sign dependent, because the profit function of the firm is

their prices, in Golosov and Lucas (2007) the firms that adjust are far from their optimum. Thereby, large price adjustments in Golosov and Lucas (2007) are endogenously selected, which makes the price level more flexible and inflation more responsive to changes in the output gap.

asymmetric: low relative prices lead to high demand, while high relative prices to low demand. Therefore, a negative deviation of the relative price from the optimum causes higher losses than a similar sized positive deviation. An aggregate shock, therefore, which generates a positive inflation and reduces relative prices can be expected to generate stronger endogenous responses and more flexible prices, than those causing negative inflation and higher relative prices.

Despite the complexity of the Phillips curve, we can characterize numerically some of its relevant features. To illustrate the *shock-size dependence* of $\kappa(\Theta)$, Figure 3 plots the relationship between annualized inflation $12\pi_t$ on the y-axis and the cumulative output gap $\sum \beta^i \mathbb{E}_t \tilde{y}^e_{t+i}$ on the x-axis for varying shock sizes (blue solid line). Specifically, we simulate random realizations of i.i.d monetary policy shocks of varying size initiated from the stationary distribution of quality-adjusted relative prices at the non-stochastic steady state. We store impulse responses of inflation and output gap. Each point on the curve shows a particular realization of inflation and cumulative output gap for a particular shock size.[4]

$$\pi_t = \kappa(\Theta_t) \sum_{i=0}^{\infty} \beta^i \mathbb{E}_t \tilde{y}^e_{t+i} + \sum_{i=0}^{\infty} \beta^i \mathbb{E}_t u_{t+i} \tag{17}$$

The figure confirms that the relationship between activity and prices is shock-size dependent. To help assessing the quantitative relevance of this non-linearity, the figure also depicts a linearized NKPC (black dashed line) calibrated to have identical slope to the true NKPC for small shocks. The right panel shows the the frequency of price changes at different inflation levels. The frequency is exactly 8.7% when the inflation $\pi_t$ is zero. This is the deterministic steady state, when there are no aggregate shocks but idiosyncratic shocks yield a non-degenerate distribution of price changes. For higher magnitudes of shocks, frequency of price adjustments increases gradually.

The figure confirms that at frequencies over 40%, the NKPC 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. At some point (not shown), as frequency reaches 100%, monetary policy becomes ineffective in influencing the output gap.[5] However, such shock sizes are fairly extreme. What does the model imply for a more realistic range of the state space?

Interestingly, for a fairly wide range of shock sizes (output gaps within the +-2% range, and frequency increases below 2 percentage points) the non-linearity is quantitatively insignificant. In this region, a linearized framework can be expected to approximate the true model quite

[4] As the solution imposes perfect foresight, expected output gaps are equal to the actual impulse-response realizations.

[5] At 100% frequency, output gap stabilizes at a level above its steady state value, as full price flexibility eliminates distortions due to idiosyncratic quality shocks.

well. For larger monetary policy impulses, however, the shocks start to exert a relatively larger impact on inflation than on output gap and the true non-linear curve deviates from its linearized counterpart. A key reason behind this development is the increase in the frequency of price changes, which makes the price level endogenously more flexible. As the frequency approaches 20%, the non-linear model is significantly different from its linearized counterpart.

Similar relationships arise for i.i.d. TFP ($A_t$) or discount rate ($\beta_t$) shocks (not shown). The results are not exactly the same, because these shocks not only influence the efficient output gap, but other features of the full model: the TFP shock affects the cost of price adjustment and the discount-rate shock the planning horizon of the price-adjusting firm. Notably, the relationship is exactly the same for i.i.d. cost-push shocks if we relate inflation to the *natural* output gap, as opposed to the efficient output gap. This is not surprising as there is a simple relationship between the natural and the efficient output gap $\tilde{y}_t^n = \tilde{y}_t^e + \log[(1-\tau_t)/(1-\tau)]$ (see Appendix B).

The results are also identical when we vary the coefficients of the Taylor rule $(\phi_\pi, \phi_y)$. Interestingly, the results are also very similar, though not exactly identical, with an inertial Taylor rule ($\rho_R > 0$).

While the location of the inflation-output gap pairs on the Phillips curve is informative about the size dependence of the shocks, the slope of the Phillips curve at each point $d\kappa(\Theta_t)/dy$, is informative about the *state dependence* of marginal shocks.It reflects the relative impact on inflation versus output gap of a small shock, provided the economy is in a particular state $\Theta_t$. Analogously, the slope reflects the state dependence of monetary policy and the inflation-output trade-off of a marginal monetary policy shock: how much cumulative output gap needs to decline to reduce inflation by a unit, also known as the *sacrifize ratio* of monetary policy. Figure 4 shows the evolution of the slope of the Phillips curve for positive shocks (the relationship is analogous for negative shocks) for a range of realistic frequency values. The slope almost doubles by the time the frequency reaches 20%. While in a low-frequency/ low-inflation environment the sacrifice ratio is high, it becomes much lower once frequency/inflation increases.

Finally, the *sign-dependence or asymmetry* of the non-linear Phillips curve is quantitatively small in case of i.i.d shocks as shown by figure 3. However, it becomes quantitatively sizable when the aggregate shocks are persistent. The reason is that persistent shocks affect the expected evolution of inflation in the near future. As explained above, due to the asymmetry of their profit and value functions, firms adjust their prices with higher probability as a response to positive inflation shocks, which reduce their relative prices than in response to negative inflation shocks, which increase them. Persistent shocks increase the expected cumulative impact on the relative price if firms do not adjust, therefore it amplifies the asymmetry.

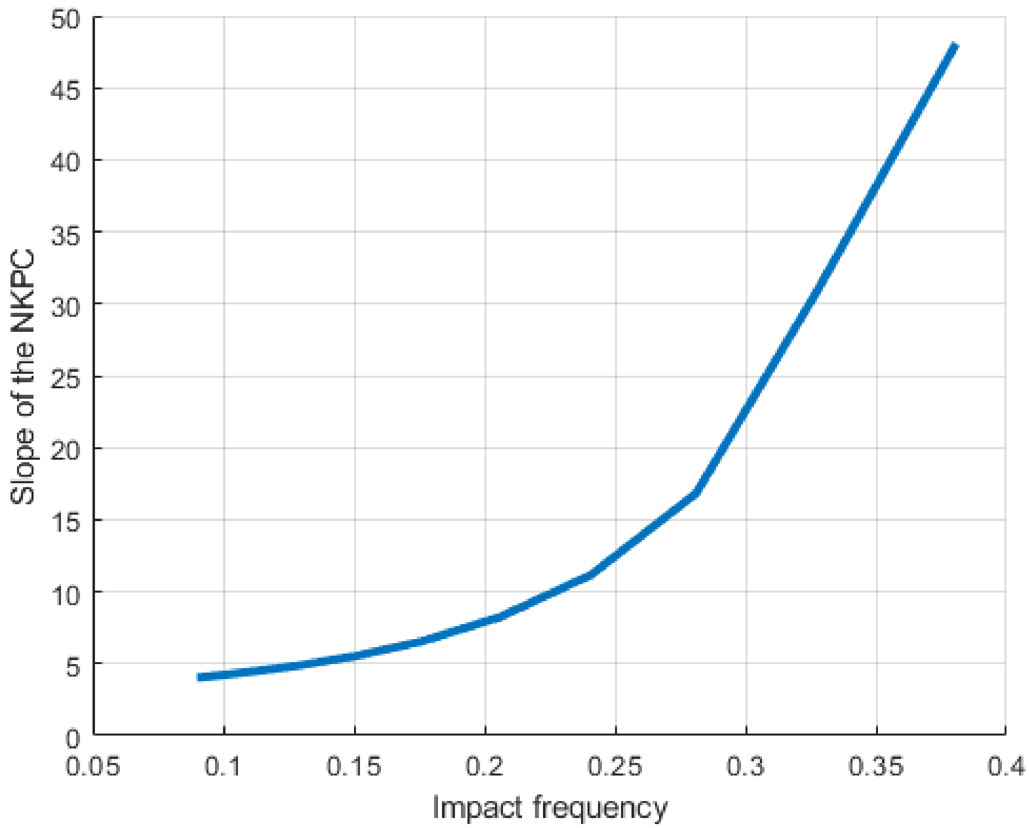



Figure 4: Frequency and the inflation-output tradeoff

# 5 Optimal monetary policy

We turn next to the analysis of optimal monetary policy. We first show the steady-state optimal inflation rate. Next we discuss the optimal response to shocks.

## 5.1 The steady state under the optimal policy

The solution of the Ramsey planners problem has a steady state state featuring slightly positive inflation 0.3%.[6] This value of inflation 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.

Why is it slightly positive and not 0? There are two asymmetries in the model that explain this. First, the profit function (11) is asymmetric: $\Pi(x) = C\left(\exp\left(x + p^*\right)\right)^{1-\epsilon} - C(1-\tau_t)\frac{w}{A}\left(\exp\left(x + p^*\right)\right)^{-\epsilon}$. Second, the consumption aggregator (3) is asymmetric. Since the asymmetry in the profit function is quantitatively much more relevant, let's have a closer

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

look at it. 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 loss of $-x$, while the positive price gap $x$ leads to smaller sales at an additional profit of $x$. This implies that the SS bands are asymmetric: the lower SS band is closer to the reset price than the upper one. Thus, in the zero-inflation steady state, there is more mass of firms at the lower SS band than at the upper, and there are more upward price adjustments than downwards. Allowing for a bit of inflation shifts the SS bands leftwards, thus reducing the number of upward price movements by more than it increases the number of downward price movements. The cost of price adjustments thus decreases. Quantitatively this effect is small and zero inflation thus is approximately optimal in the Ramsey steady state, just as in the standard New Keynesian model with Calvo prices (Galí, 2008).

### 5.2 Time-0 problem

We now 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 5 show time path under the optimal policy. The optimal policy in this case is time-inconsistent: despite starting at the optimal steady-state, the central bank engineers a short-lived tiny expansion. This happens despite the fact that there is a labor subsidy $\tau = 1/\epsilon$ in place which undoes the desired markup. In the standard Calvo model this subsidy renders the steady state efficient and thus the policy time-consistent. In our model, firm heterogeneity implies that the subsidy is not at the appropriate level to offset market power for each individual firm. The central bank has thus an incentive to temporarily expand the economy.

This new source of time inconsistency is nonetheless relatively weak when compared to traditional time inconsistency associated with removing the subsidy $\tau$ and introducing a markup distortion (red dashed lines, right axis). In this latter case the monetary expansion that the central bank optimally implements is almost two orders of magnitude larger (measured by the peak response).

### 5.3 Timeless optimal monetary response to efficient shocks

While the time-0 policy is interesting to analyze time inconsistency, our main focus is on the systematic monetary policy response to shocks. We then analyze -timeless' optimal monetary policy (Woodford, 2003; Galí, 2008). This can be obtained as the optimal monetary policy starting from the Ramsey steady state when a shock hits and all of the Lagrange multipliers are initialized at their steady-state values.

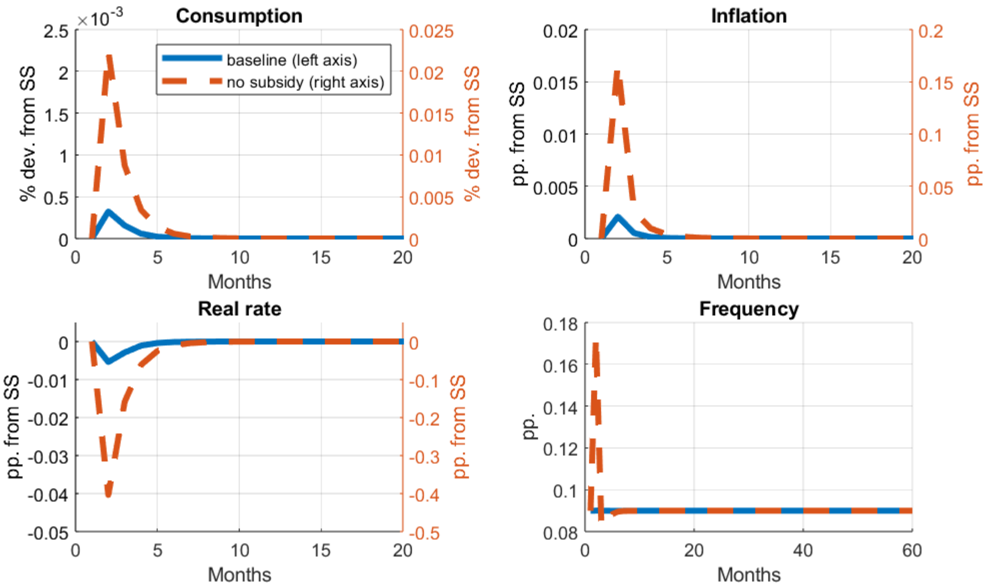



Figure 5: TFP shock

First, we consider shocks that affect the efficient allocation, such as TFP shocks. 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).

For small shocks we have already seen how the standard New Keynesian model provides a good approximation and thus the optimal monetary policy is also given by price stability. In the case of a large shock, this prescription also holds in our economy, at least approximately. As figure 6 shows for a standard negative TFP shock (dashed red line), inflation remains approximately at zero, and consumption and the real interest rate follow their natural levels (solid blue line). As inflation remains constant at zero, the frequency of repricing barely moves.[7]

The conclusion is that strict inflation targeting simultaneously minimizes inefficient output fluctuations and the costs of nominal rigidities, which in this model are the sum of adjustment costs and price dispersion. Notice that the shape of the Phillips curve plays no role for this result. In the next subsection we turn to a case when the Phillips curve matters.

[7] It changes are all due to the changes in real variable that modify the incentives of firms to update prices

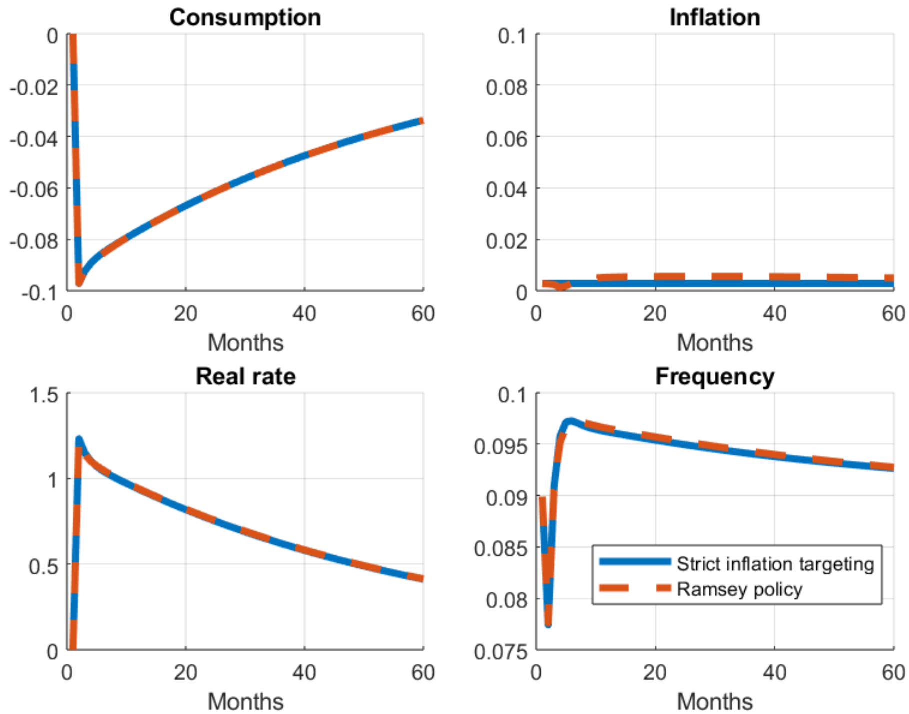



Figure 6: TFP shock

## 5.4 Timeless optimal monetary response to inefficient shocks

**Impulse responses**. Cost-push shocks are well known to break the divine coincidence in the standard New Keynesian model. For such shocks, a policy of 'leaning against the wind' is optimal, that is the central bank temporarily drives output below its efficient level to contain the inflationary impact of a positive cost-push shock. The central bank trades off this decline in the output gap with an initial bout of inflation under the commitment that inflation will be negative in the future so that the price level is eventually restored. In other words, the central bank can partially offset the inflationary impact of a cost-push shock by lowering the current output gap and by committing to lower future output gaps. Such a policy exploits the Phillips curve relationship and is thus the relevant case to explore the implications of a nonlinear Phillips curve.

The optimal monetary policy response to a cost-push shock in our model is also characterized by leaning against the wind. The first row of Figure 7 shows the response of the economy to an inflationary cost-push shock under optimal policy. Inflation goes up and the efficient output gap drops.

However, due to the nonlinearity of the Phillips curve, the policy is now size-dependent. Figure 7 compares the response to a large shock (solid blue line) with that of a small shock

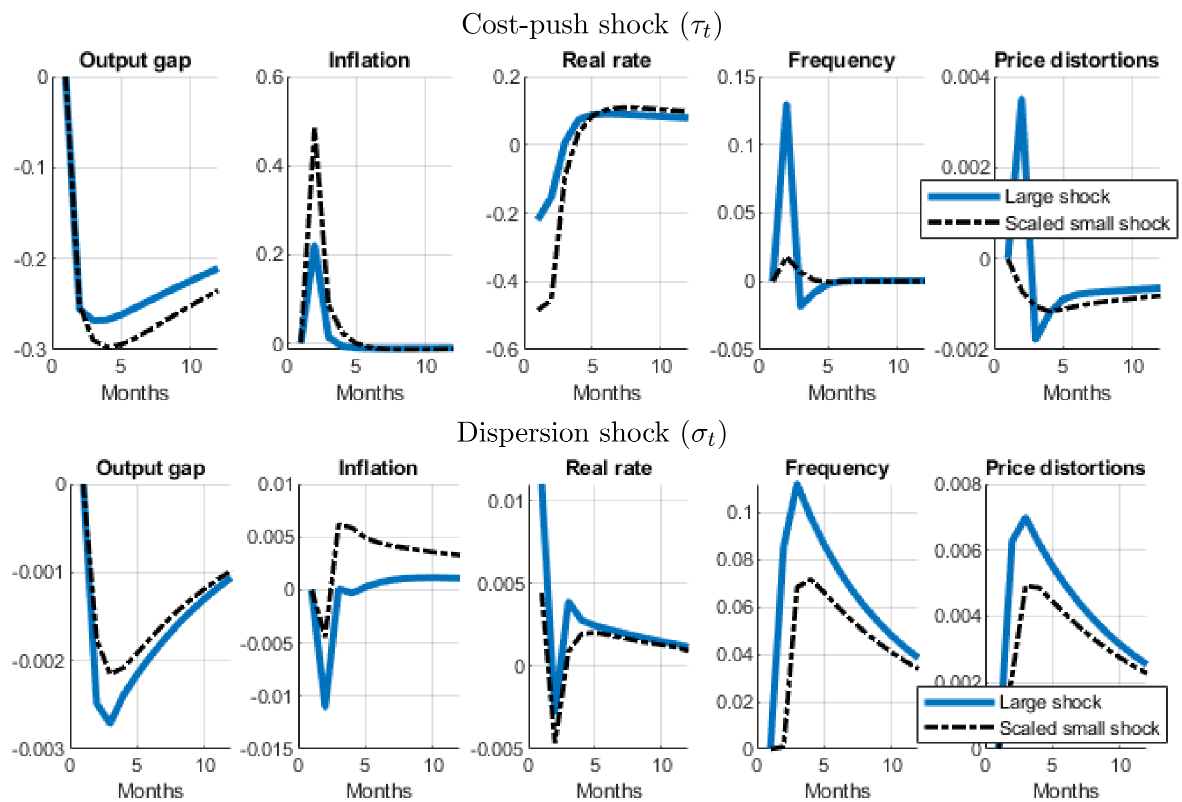



Figure 7: Impulse responses to a large and a linearly-scaled small cost-push shock and an indiosyncratic volatility shock under optimal policy

that is scaled linearly by the relative shock size (black dashed line). The optimal monetary policy, represented by the real interest rate, is tighter for the large shock: real interest rates fall by less than half on impact and then increase faster than in the case of a small shock. This prevents inflation from increasing as much as in the small-shock case. The tighter policy is also successful in preventing a decline in the output gap as large as with small shocks. Similar to the case under a suboptimal Taylor rule, frequency jumps in the case of a large shock.

To interpret these results, we first need to understand the impact of higher frequency on the conventional New Keynesian model. In the linearized Calvo (1983) model, optimal policy is characterized by a targeting rule linking the output gap and deviations from the price level. The relationship is constant and depends only on the elasticity of substitution parameter $\varepsilon$. Strikingly, the relationship is independent of the frequency of price changes, proportional to the Calvo parameter $\theta$. 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, however, higher frequency reduces the relative weight of inflation in the welfare function, as this is given by $\varepsilon/\kappa$. The reason is that higher frequency reduces the relative-price distortions caused by inflation, so the planner is willing to tolerate more inflation.

The frequency of price changes also influences the magnitude of the Phillips-curve wedge ($u_t$) that a cost-push shock generates. The shock $u_t = \kappa(y_t^e - y_t^n)$ is equal to the product of the slope of the Phillips curve ($\kappa$), which increases with the frequency of price changes, and the difference between the efficient output ($y_t^e$) and the natural level of output ($y_t^n$) (see Galí, 2008, pp. 97.). The cost-push shock ($\tau_t$) drives the difference between the efficient and the natural output. An increase in frequency thus amplifies the impact of the shock on the Phillips curve.

In our model higher frequency increases, not decreases, distortions, as shown in Figure 7. This is because as frequency increases more firms update prices, thus paying menu costs, which more than compensates for the improvement in price distortion as prices become more flexible. That implies that the targeting rule will not be independent of frequency: as frequency increases the central bank has more incentives to quench inflation to minimize price distortions and is more effective in doing so, due to a lower sacrifice ratio. These two effects trump the amplification of the cost-push wedge and explain our results.

The second row of Figure 7 contrasts the responses to a dispersion shock. Compared to the case with a Taylor rule, the response is characterized by a more muted response of inflation: the higher frequency allows the central bank to quench inflation.

**Impact of shock size**. Figure 8 illustrates the optimal response to the cost-push shock for a whole range of shock sizes. The impact response is plotted as a function of the shock size (blue line) and is contrasted to the counterfactual linear relationship (black dashed line),

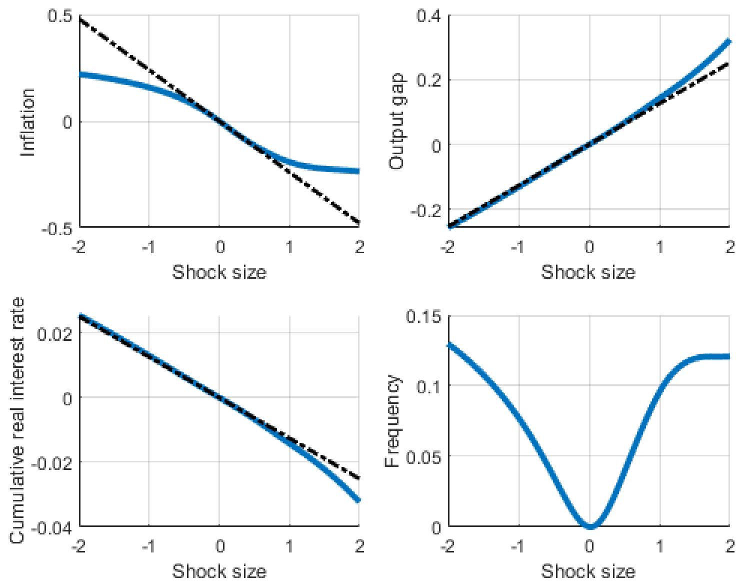


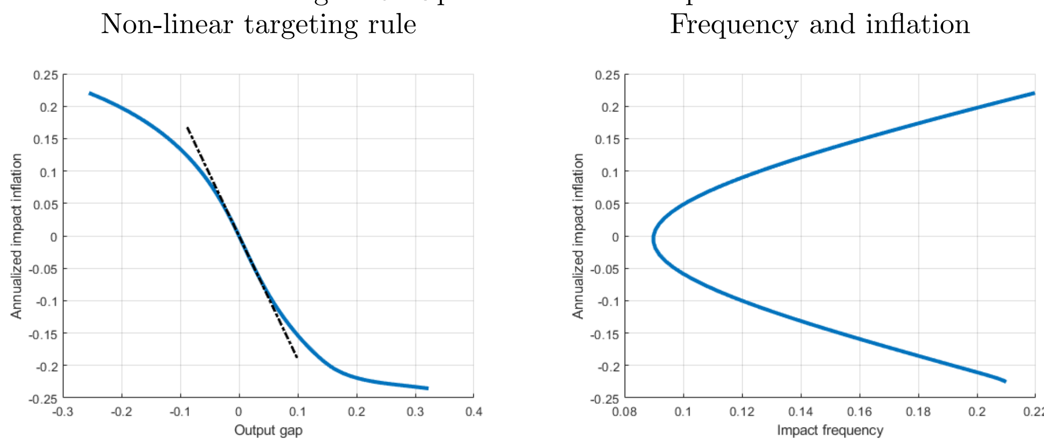

Figure 8: Optimal response to a cost push shock as a function of the shock size

which is extrapolated from the slope around small shocks (with the exception of the frequency, where this linear counterfactual is not reported). Two results should be highlighted. First, the deviation between the two lines demonstrates the nonlinearity of the economy under optimal policy for large shocks. The larger the shock the more inflation and frequency are contained. Second, there is certain sign-dependence, as the optimal policy to positive and negative shocks differ. This is related to the asymetries discussed in the previous section.

**Non-linear targeting rule**. Figure 9 shows the relationship between annualized impact inflation (which is equal to the impact on the price level) and impact output gap as a response to cost-push shocks of different sizes all with high persistence ($\rho_\tau = 0.9$, blue line). The slope of the figure at 0 is very close to $-1/\varepsilon$ (but slightly steeper), the slope of the target rule in the linearized Calvo (1983) (see Figure 11 in the appendix). Therefore in the linearized Golosov and Lucas (2007) model the higher slope of the Phillips curve due to the selection effect of large price changes significantly reduces the relative importance of inflation in the welfare function and a target rule with slope $-1/\varepsilon$ is close to optimal.

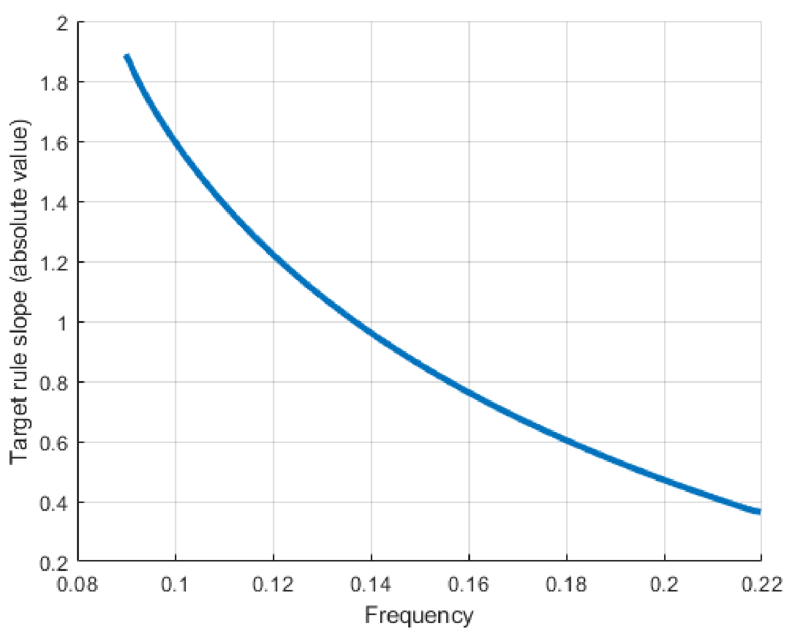
The figure also shows that the relationship is non-linear. At a frequency around 12.5%, the slope of the optimal target rule becomes significantly lower than its slope at 0. Figure 10 shows the slope explicitly as a function of the frequency. The relationship means that after a large shock that increases frequency substantially, the planner is stabilizing inflation relative to the output gap on the margin more than after small shocks. This implies that the reduction in the sacrifice ratio matters more overall than the reduction in the relative welfare weight of inflation after large shocks.

Figure 9: Optimal inflation-output trade-off




*Note: The left panel shows the relationship between annualized inflation and the output gap on impact as a function of a series of cost-push shocks under optimal policy (blue line). The figure shows that the relationship is mildly non-linear. For large shocks, its slope deviates significantly from its slope at 0, which is extended over the whole range by the black dashed line. The targeting rule shows that optimal policy leans against inflation stronger than against output gap for large shocks. The right panel shows the frequency at each inflation levels. It shows that the targeting rule is non-linear at frequency of 12.5%.*

Figure 10: Frequency and the optimal inflation-output tradeoff




*Note: The figure shows the (absolute value of the) slope of the targeting rule as a function of frequency.*

## 6 Conclusion

Our research centers on the Ramsey problem within the Golosov and Lucas (2007) menu-cost model. 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 and stabilizing the repricing rate. This commitment serves the purpose of counteracting the inflationary inclinations of firms and dampening the state-dependence inherent in firms' decisions.

Along the trajectory of optimal commitment, tangible benefits emerge for the central bank. Notably, there is a clear reduction in the sacrifice ratio, leading to lower inflation and a more moderate decline in output. This nuanced understanding of the optimal policy diverges markedly from the predictions of the standard New Keynesian model with exogenous timing of price adjustment, which fails to capture such non-linear dynamics. Simultaneously, when confronted with demand or efficiency shocks (e.g., TFP shocks), our findings indicate that the optimal policy in the Golosov and Lucas (2007) 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 a proactive policy by the central bank in the face of substantial shocks. By committing to policies that curb inflation and stabilize the repricing rate, the central bank can deliver a more favorable macroeconomic outcome. This nuanced perspective on optimal monetary policy contributes an important layer to the ongoing discourse surrounding the Golosov and Lucas (2007) menu cost model and its implications for real-world monetary policy.

## 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, 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 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 D Rigato, Matthew Rognlie, and Ludwig Straub (2022) "New Pricing Models, Same Old Phillips Curves?"Technical report, National Bureau of Economic Research.

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, forthcoming.*

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 Jones, and Virgiliu Midrigan (2022) "Nonlinear Inflation Dynamics in Menu Cost Economies."

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.

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.

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.

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

——— (2023b) "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, Anton Nakov, and Borja Petit (2022) "Flattening of the Phillips Curve with State-Dependent Prices and Wages," *The Economic Journal*, Vol. 132, pp. 546–581.

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.

Gertler, Mark and John Leahy (2008) "A Phillips Curve with an Ss Foundation," *Journal of Political Economy*, Vol. 116, pp. 533–572.

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.

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

Phillips, A. W. (1958) “The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957,” *Economica*, Vol. 25, pp. 283–299.

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,”Technical report.

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.

# Appendix

## A Steady state computation

THIS SECTION IS TO BE UPDATED

We have $w = \upsilon C^{\gamma}$, and $R = (1+\pi)/\beta$. We solve the stationary Bellman equation on a grid of width $\Delta p$, $p_j \in \left[\underline{p}, \underline{p}+\Delta p, ..., \overline{p}\right]$, $j = 1,..,J$, so that $V_j = V(p_j)$. The expectation $\mathbb{E}\left[V(p - \sigma\varepsilon_{t+1} - \pi)|p = p_j\right]$ can be approximated as $\sum_{i=1}^{J} \mathcal{N}_{j,i} V_i$, where

$$\mathcal{N}_{j,i} = \int_{p=p_{i-1/2}}^{p_{i+1/2}} \psi\left(\frac{-p + p_j - \pi}{\sigma}\right) dp = \int_{p=p_{i-1/2}}^{p_{i+1/2}} \psi\left(\frac{p - (p_j - \pi)}{\sigma}\right) dp =$$
$$\Psi\left(\frac{p_{i+1/2} - (p_j - \pi)}{\sigma}\right) - \Psi\left(\frac{p_{i-1/2} - (p_j - \pi)}{\sigma}\right),$$

where $p_{i-1/2} \equiv (p_{i-1} + p_i)/2$, $p_{i+1/2} \equiv (p_i + p_{i+1})/2$, $\psi(\cdot)$ is the standard normal probability density function, and $\Psi(\cdot)$ is the standard normal cumulative distribution function. This is a linear approximation to the value function using a Gaussian function.

We define matrix

$$\mathbf{A} = \begin{bmatrix} \mathcal{N}_{1,1} & \mathcal{N}_{1,2} & \cdots & \mathcal{N}_{1,J} \\ \mathcal{N}_{2,1} & \mathcal{N}_{2,2} & \cdots & \mathcal{N}_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{N}_{J,1} & \mathcal{N}_{J,2} & \cdots & \mathcal{N}_{J,J} \end{bmatrix}.$$

We also define the vectors

$$\boldsymbol{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_J \end{bmatrix}, \ \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_J \end{bmatrix},$$

and

$$\mathbf{V} = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_J \end{bmatrix}, \boldsymbol{\Pi} = \begin{bmatrix} \left[e^{-\epsilon p_1}\left(e^{p_1} - w\right)\right] C \\ \left[e^{-\epsilon p_2}\left(e^{p_2} - w\right)\right] C \\ \vdots \\ \left[e^{-\epsilon p_J}\left(e^{p_J} - w\right)\right] C \end{bmatrix},$$

the Bellman equation is

$$\mathbf{V} \;=\; \boldsymbol{\Pi} + \beta\left[\mathbf{A}\left(1 - \lambda\right) \cdot \mathbf{V} + \mathbf{A}\lambda \cdot \left(\boldsymbol{\phi}'\mathbf{V} - \kappa\mathbf{w}\right)\right]$$

where $\cdot$ denotes element-by-element multiplication. To solve it, the algorithm is as follows.

Start with a guess for the real wage $w^0$, then:

1. Compute consumption $C^n = \left(\frac{w^n}{v}\right)^{1/\gamma}$.
2. Using $\pi = \overline{\pi}$, $C^n$ and $w^n$, solve the Bellman equation. To compute the distribution, solve
$$\mathbf{g} = (1-\lambda) \cdot (\mathbf{A}'\mathbf{g}) + \boldsymbol{\phi}\lambda' \left(\mathbf{A}'\mathbf{g}\right).$$
3. Compute the aggregation residual
$$resid = 1 - \sum_{j=1}^{J} e^{(1-\epsilon)p_j} g_j.$$
4. If $resid > 0$, then decrease the real wage and go back to step 1. Otherwise, increase the real wage and go back to step 1. Stop when $resid \approx 0$.

# B Efficient and natural level of output

For the welfare analysis, it is instructive to derive the efficient and natural levels of aggregate and product-level output.

**Efficient output**. The efficient level of output is the solution to a social planner problem. The problem maximizes household welfare (1) subject to (i) the aggregate consumption equation (3), (ii) aggregate labor supply ($N_t = \int_i N_t(i)$) and (iii) product-level production functions (9) with respect to product-level consumption and labor ($C_t(i), N_t(i), i \in [0,1], t = 0,1,2,\dots$).

After straightforward substitutions, the optimization problem simplifies to

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

subject to $\int N_t(i)di = N_t$. The solution implies that the efficient output fluctuates with aggregate productivity, but is independent of demand (monetary, discount-rate) shocks and cost-push (markup, labor-tax) shocks. In particular

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

The efficient real interest rate is implicitly defined by the Euler equation after substituting in efficient consumption. Under perfect foresight $R_t^e \equiv \exp r_t^e = \mathbb{E}\left[(1+i_t)/(1+\pi_{t+1})\right]$ is determined by the following expression:

$$r_t^e = log\beta_t + (1-\rho_A)\log A_t.$$

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

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

**Natural output**. The natural level of output is the counterfactual output with flexible prices. Under flexible prices, firms maximize their real profit function (11) in each period $t$ by choosing

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

The expression implies that the quality-adjusted relative price is homogeneous across products $i$. Together with the definition of the price-level (5), this implies that the natural level of the quality adjusted log relative price is zero ($p_t(i) = 0$). Or equivalently, the natural level of relative price is equal to the quality: $P_t^n(i)/P_t^n = A_t(i)$. 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 i:

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

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

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

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

The natural real interest rate is implicitly defined by the Euler equation after substituting in natural consumption. Under perfect foresight it is determined by the following expression:

$$r_t^e = log\beta_t + (1-\rho_A)\log A_t - (1-\rho_\tau)\log(1-\tau_t).$$

# C Additional figures

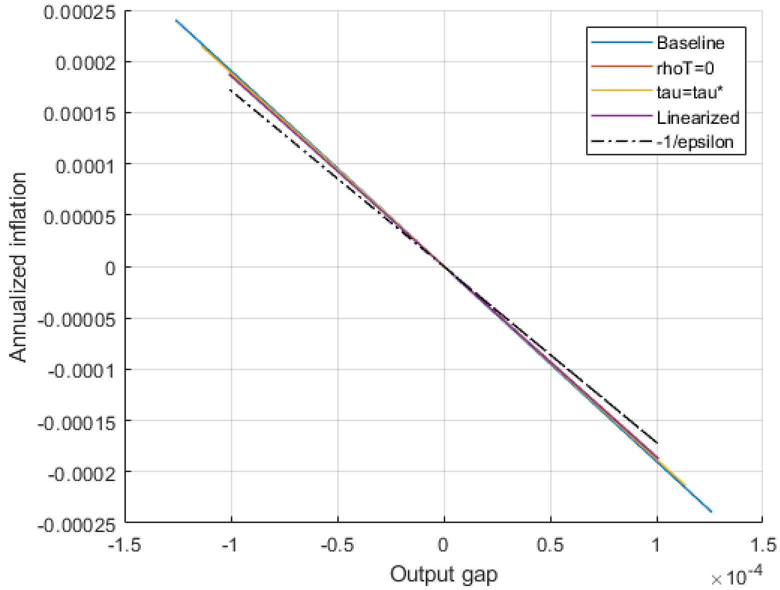



Figure 11: Optimal target rules for different parameters and small cost-push shocks