Monetary Policy and Supply Disturbances: A Heterogeneous Firm Approach*

Klaus Adam - University College London & CEPR

Henning Weber - Deutsche Bundesbank

July 9, 2025

Abstract

We study optimal monetary policy design in a sticky price setup with heterogeneous firms and firm turnover. The quadratic welfare objective and the linearized Phillips curve both depend on the gap between inflation and optimal inflation, with optimal inflation fluctuating in response to productivity disturbances around a non-zero average value. Monetary policy should "look through" certain productivity disturbances and leave the nominal interest rate unchanged. Some of these productivity disturbances move output and inflation in the same direction, creating challenges for the correct identification of demand and monetary policy shocks in empirical work.

JEL-Class.No.: E31, E32, E52, E61

1 Introduction

Monetary policy models traditionally rely on representative firm frameworks that abstract from heterogeneity between firms other than price heterogeneity. In micro data, firms display many additional dimensions of heterogeneity; in particular, young firms tend to be small initially, tend to grow over time

*Klaus Adam: k.adam@ucl.ac.uk; Henning Weber: henning.weber@bundesbank.de. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Deutsche Bundesbank, the Eurosystem, or their staff.

as they accumulate experience, and tend to exit when faced with sufficiently negative shocks.

This paper develops a linear-quadratic approximation to the optimal monetary policy problem in a tractable sticky-price economy with heterogeneous firms. Specifically, we consider firm turnover and cohort- and experience-specific productivity growth. Firm heterogeneity fundamentally alters the optimal monetary policy design relative to the homogeneous firm setup pioneered by Woodford (2003, 2011), especially with respect to supply shocks.

Our setup features three supply shocks:

(i) *common shocks to productivity growth* that uniformly affect productivity of all firms,

(ii) *cohort-productivity growth shocks* that affect only the initial productivity of newly entering firm cohorts, and

(iii) *experience-accumulation growth shocks* that alter only the productivity of incumbent firms.

In addition to these shocks, every firm faces an idiosyncratic risk of becoming unproductive, inducing it to exit from the economy and being replaced by a new firm.

Considering this setup in the presence of sticky prices, we derive the linearized New Keynesian Phillips curve, the quadratically approximated welfare objective, and the dynamics of the natural rate of interest and the efficient level of output.

The main new features arising from firm heterogeneity are: (1) the welfare objective and the New Keynesian Phillips curve feature the *gap* between *inflation* and *optimal inflation*, instead of inflation; (2) the optimal average value of the inflation rate is positive for plausible model parameterizations; (3) optimal inflation fluctuates in response to supply shocks; and (4) the slope of the Phillips curve decreases with the optimal steady-state inflation rate.

Although 'divine coincidence' continues to hold for all productivity disturbances, i.e., the output gap and the inflation gap can be closed simultaneously, the implications of supply shocks differ fundamentally from those in the standard model with homogeneous firms. Specifically, shocks to the growth rate of cohort productivity (ii) and experience productivity (iii) affect the dynamics of optimal inflation. In contrast, shocks to the growth rate of common productivity (i), which are the only productivity disturbances present in homogeneous firm models, fail to do so.

This finding emerges because shocks (ii) and (iii) affect the *relative productivity* of new versus incumbent firms, unlike common productivity shocks (i). Since the prices of incumbent firms are sticky, it is efficient that the necessary relative price adjustments following changes in relative productivity are brought about by new firms, whose prices are flexible. Therefore, shocks that cause the productivity of new firms to fall (rise) relative to incumbent firms make it optimal that new firms set higher (lower) prices, leading to inflation that is optimally higher (lower) than otherwise.

Despite certain supply shocks persistently moving inflation, monetary policy should "look through" these shocks and keep the nominal rate constant. This is the case following transitory or persistent shocks to cohort productivity growth (ii) and transitory shocks to experience productivity growth (iii). Following these shocks, movements in expected optimal inflation cause the real interest rate to track the natural real rate exactly when nominal rates are kept constant.

Supply shocks not only move optimal inflation but naturally also affect the dynamics of optimal output. This creates challenges for the correct identification of demand shocks in a setting with heterogeneous firms. Specifically, a purely temporary positive shock to the growth rate of experience productivity (iii) causes both output and inflation to persistently increase under optimal monetary policy. This makes it challenging to distinguish the effects of this supply shock from the effects of a standard demand shock. Both shocks generate a persistent increase in output and inflation and a persistent fall in the natural rate of interest, without causing long-run effects on any variable. Traditional sign-based identification approaches of demand shocks are thus prone to misidentify supply shocks. The effects of this supply shock can also cause difficulties for high-frequency identification of monetary policy shocks, as it becomes more challenging to filter out central bank information effects (Jarocinsky and Karadi (2020)).

More generally, the heterogeneous firm model gives rise to an alternative theory of demand shocks that is grounded in productivity shocks, akin to the imperfect information model presented in Lorenzoni (2009), where demand-like shocks arise from information noise about aggregate productivity, or the work of Bai, Ríos-Rull and Storeletten (2025), where supply shocks generate demand-like responses in a setup with search frictions.

Our finding that supply shocks mimic demand shocks in a setup with heterogeneous firms may also explain why measured productivity responds positively to an identified monetary policy loosening in the data, as documented already in Christiano, Eichenbaum and Evans (2005): empirically identified demand shocks may be partially contaminated by supply shocks. This explanation contrasts with complementary explanations put forward in Meier and Reinelt (2024) and Baqaee, Farhi and Sangani (2024), who also consider heterogeneous firm setups with sticky prices. They emphasize that demand expansions lead to a compression of the cross-sectional mark-up distribution and increase aggregate productivity through this compression. Unlike the present paper, they do not study optimal monetary policy design.

The present analysis goes beyond Adam and Weber (2019), which focuses exclusively on the implications of firm heterogeneity for the optimal inflation rate and does not provide a linear-quadratic formulation of the optimal monetary stabilization problem. Doing so allows investigating a number of new issues, such as the identification of demand versus supply shocks and the optimality of "look-through" policy. The present analysis also connects more directly to a large body of applied work that relies on linearized structural equations.

The remainder of the paper is structured as follows. Section 2 presents the sticky price model with firm heterogeneity, and section 3 explains how firm-level productivity dynamics affect the optimal inflation rate. Section 4 presents the linear-quadratic optimal policy problem for the case with firm heterogeneity and the dynamics for optimal inflation and output. Section 5 shows that it is optimal for monetary policy to look through certain supply disturbances. Section 6 explains how supply disturbances generate very similar outcomes to demand disturbances and why this can lead to misidentification of demand shocks. Section 7 presents a generalized setup in which non-reoptimizing firms index prices to lagged inflation.

2 Economic Model

We consider a cashless sticky price economy with Calvo price rigidities, heterogeneous firms, and firm turnover, as in Adam and Weber (2019). Firms enter the economy with a cohort-specific initial productivity level, accumulate technological experience during their lifetime, and randomly exit the economy. Unlike the standard New Keynesian model, the present setup gives rise to a setting in which the average productivity of price-adjusting firms is not necessarily equal to that of non-adjusting firms. As a result, the optimal steady-state inflation rate generically differs from zero and responds optimal inflation responds to supply side disturbances. The monetary policy implications of these supply side disturbances become particularly transparent with the linear Phillips curve and the quadratic welfare approximation that we derive here.

2.1 Supply Disturbances and Nominal Rigidities

Each period $t = 0, 1, \ldots$ there is a unit mass of monopolistically competitive firms indexed by $j \in [0, 1]$. Each firm j produces output Y_{jt}, which enters as an input into the production of an aggregate consumption good Y_t according to

$$Y_t = \left(\int_0^1 Y_{jt}^{\frac{\theta-1}{\theta}} \mathrm{dj} \right)^{\frac{\theta}{\theta-1}}, \tag{1}$$

where $1 < \theta < \infty$ denotes the price elasticity of product demand. Let P_{jt} denote the price charged by firm j in period t. Firms can adjust prices with probability $1 - \alpha$ each period ($0 \leq \alpha < 1$). The arrival of such a Calvo price adjustment opportunity is thereby idiosyncratic and independent of all other exogenous random variables in the economy.

We augment this standard setting by a second price adjustment opportunity that arrives with idiosyncratic probability $\delta > 0$ each period and that is independent of the Calvo adjustment opportunities. This second adjustment opportunity, to which we refer as "δ-shock", arrives in conjunction with a firm-level productivity change, as described below. Let $\delta_{jt} \in \{0, 1\}$ denote the idiosyncratic i.i.d. random variable governing this second price and productivity adjustment and let $\delta_{jt} = 1$ indicate the arrival of a δ-shock. We will interpret $\delta_{jt} = 1$ as capturing a firm turnover event in which firm j becomes permanently unproductive, thus exits the economy, and is replaced by a newly entering firm. For simplicity, we assign the same firm index j to the new firm, which operates, however, using different technology.

One can alternatively interpret $\delta_{jt} = 1$ as a product turnover event in which the demand for firm j's product disappears and where the same firm introduces a new product that is then produced with a different technology.

Letting L_{jt} denote the amount of labor used by firm j, firm output Y_{jt} is given by

$$Y_{jt} = A_t Z_{jt} L_{jt}, \tag{2}$$

where A_t captures the productivity common across all firms and Z_{jt} firm-specific productivity. Atkeson and Kehoe (2005) refer to Z_{jt} as the 'organization capital' of firm j. Common productivity evolves according to

$$A_t = a_t A_{t-1},$$

where a_t is the (gross) growth rate of common productivity in t. Firm-specific productivity evolves according to

$$Z_{jt} = \begin{cases} g_t Z_{jt-1} & \text{if } \delta_{jt} = 0 \\ Q_t & \text{if } \delta_{jt} = 1, \end{cases} \tag{3}$$

where g_t denotes the (gross) growth rate at which firms accumulate experience as they age, e.g., from learning-by-doing effects. The variable Q_t denotes the initial productivity level of the cohort of new firms that enter in period t.[1] We assume that

$$Q_t = q_t Q_{t-1}, \tag{4}$$

where q_t is the (gross) growth rate of initial cohort productivity level.

Our setup features three different stochastic productivity trends that will all have different implication for optimal inflation and the behavior of the natural rate of interest: (1) the common growth trend a_t, (2) the experience growth trend g_t, and (3) the cohort growth trend q_t.

We can decompose each of these trends into a mean component and into fluctuations around the mean:

$$\begin{aligned} a_t &= a \cdot e^{\widehat{a}_t} \\ q_t &= q \cdot e^{\widehat{q}_t} \\ g_t &= g \cdot e^{\widehat{g}_t}, \end{aligned}$$

where a, q, g denote mean components and $\widehat{a}_t, \widehat{q}_t, \widehat{g}_t$ are arbitrary serially correlated shocks whose unconditional mean satisfies $E[e^{\widehat{a}_t}] = E[e^{\widehat{q}_t}] = E[e^{\widehat{g}_t}] = 1$. To obtain a well-defined steady state, we assume

$$(1-\delta)\left(g/q\right)^{\theta-1} < 1. \tag{5}$$

[1] Atkeson and Kehoe (2005) refer to the initial productivity endowment Q_t, which the firm is assigned when $\delta_{jt} = 1$, as the 'frontier of knowledge'.

Let s_{jt} denote the number of periods that have elapsed since firm j last experienced a δ-shock.[2] Firm-specific productivity Z_{jt} in equation (3) is then given by

$$Z_{jt} = G_{jt} Q_{t-s_{jt}},$$

where

$$G_{jt} = \begin{cases} 1 & \text{for } s_{jt} = 0 \\ g_t G_{jt-1} & \text{otherwise,} \end{cases}$$

and where Q_t follows equation (4). This shows that firm-specific productivity reflects the cohort productivity level at the time of entry times the productivity gains from accumulated experience.

As usual, we define the aggregate price level as

$$P_t \equiv \left(\int_0^1 P_{jt}^{1-\theta} \mathrm{dj} \right)^{\frac{1}{1-\theta}}, \tag{6}$$

so that the gross inflation rate is given by

$$\Pi_t \equiv P_t / P_{t-1}.$$

Cost minimization in the production of final output Y_t implies

$$Y_{jt} = \left(P_{jt} / P_t \right)^{-\theta} Y_t. \tag{7}$$

Aggregation across firms yields the aggregate technology

$$Y_t = \frac{A_t Q_t}{\Delta_t} L_t, \tag{8}$$

where L_t denotes aggregate hours worked and

$$\Delta_t = \int_0^1 \left(\frac{Q_t}{G_{jt} Q_{t-s_{jt}}} \right) \left(\frac{P_{jt}}{P_t} \right)^{-\theta} \mathrm{dj} \tag{9}$$

the endogenous component to aggregate productivity.

[2]This means that $\delta_{j,t-s_{jt}} = 1$ and that $\delta_{j,\widetilde{t}} = 0$ for $\widetilde{t} = t - s_{jt} + 1, ..., t$.

2.2 Optimal Price Setting

Firms choose prices and hours worked to maximize expected discounted profits. While price adjustment is subject to adjustment frictions, factor inputs can be chosen flexibly. Firms' sales may be subject to a linear sales subsidy τ_t (sales tax if $\tau_t < 0$).

Each period t, firms receive a δ-shock with probability δ and - conditional on not receiving a δ-shock - a Calvo shock with probability $1-\alpha$. If firm j receives either of the two shocks, it can freely choose a new optimizing price $P^{\star}_{jt}$, otherwise the firm's price is given by its previous price ($P_{jt} = P_{jt-1}$).[3]

Appendix A.1 shows that the optimal reset price $P^{\star}_{jt}$ is given by

$$\frac{P^{\star}_{jt}}{P_t}\left(\frac{Q_{t-s_{jt}}G_{jt}}{Q_t}\right) = \left(\frac{\theta}{\theta-1}\frac{1}{1+\tau}\right)\frac{N_t}{D_t}, \tag{10}$$

where the aggregate variables N_t and D_t are independent of the firm index j and defined in the appendix and τ denotes the steady-state value of the sales subsidy.[4] The previous equation shows that the optimal relative reset price of a firm depends on aggregate variables and on the ratio between its own productivity ($A_t Q_{t-s_{jt}} G_{jt}$) and the productivity of a new firm in period t ($A_t Q_t$). In the standard New Keynesian setting with homogeneous firm, this productivity ratio is always equal to one, while this will generally not be the case in the present setting.

As is standard in the New Keynesian literature, we assume that the steady-state sales subsidy τ eliminates the distortions arising from monopolistic competition:

$$\frac{\theta}{\theta-1}\frac{1}{1+\tau} = 1. \tag{11}$$

The sales subsidy τ_t is allowed to fluctuate around this steady-state level and will give rise to mark-up disturbances in the Phillips curve.

2.3 Household Problem

Following Christiano, Trabandt and Walentin (2010), we consider a representative household with balanced-growth consistent preferences that are sepa-

[3] The case where non-reoptimized prices are indexed to lagged inflation will be considered in section 7.

[4] Fluctuations of τ_t enter the variable D_t, see appendix A.1.

rable between consumption C_t and hours worked L_t:

$$E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln C_t - \frac{L_t^{1+\psi}}{1+\psi} \right), \tag{12}$$

where $\beta \in (0,1)$ denotes the discount factor and ξ_t a preference shock, which evolves according to

$$\xi_t / \xi_{t-1} = e^{\widehat{\xi}_t},$$

with $\widehat{\xi}_t$ denoting a serially correlated shock with $E[e^{\widehat{\xi}_t}] = 1$ that implies $E[\xi_t] = 1$. The household faces the flow budget constraint

$$C_t + \frac{B_t}{P_t} = \frac{W_t}{P_t} L_t + \int_0^1 \frac{\Theta_{jt}}{P_t} \text{ dj} + \frac{B_{t-1}}{P_t}(1 + i_{t-1}) - T_t,$$

where B_t denotes nominal government bond holdings, i_{t-1} the nominal interest rate, W_t the nominal wage rate, Θ_{jt} nominal profits from ownership of firm j, and T_t lump sum taxes.

2.4 Government

To close the model, we consider a government which faces the budget constraint

$$\frac{B_t}{P_t} = \frac{B_{t-1}}{P_t}(1 + i_{t-1}) + \tau_t \int_0^1 \left(\frac{P_{jt}}{P_t} \right) Y_{jt} \text{ dj} - T_t.$$

The government levies lump sum taxes T_t, so as to implement a bounded state-contingent path for government debt B_t/P_t.[5] Since we consider a cashless limit economy, there are no seigniorage revenues, even though the central bank controls the nominal interest rate. We furthermore assume that monetary policy is not constrained by a lower bound on nominal interest rates.

3 Firm-Level Productivity Dynamics and Optimal Inflation

This section illustrates the firm-level productivity dynamics implied by our setup and the crucial role played by inflation in achieving efficiency of relative

[5] The household's transversality condition will then automatically be satisfied in equilibrium.

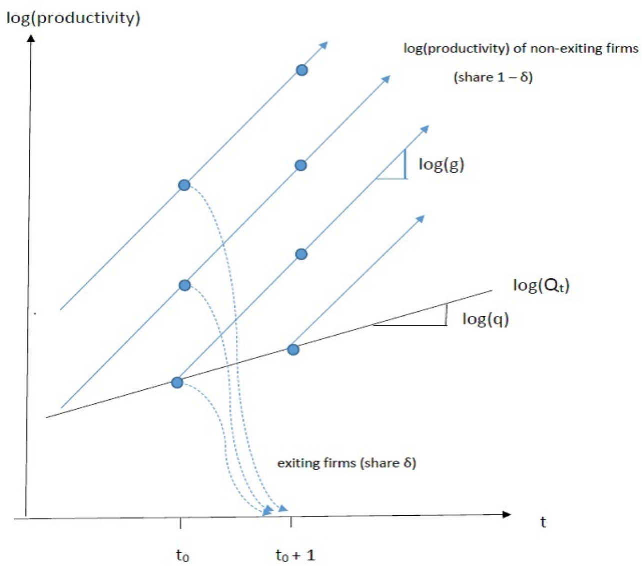

Figure 1: Productivity dynamics in a setting with firm entry and exit

prices.[6]

Figure 1 illustrates the firm-level productivity dynamics for the empirically plausible case in which the cohort trend exceeds unity ($q > 1$), but is less strong than the experience trend ($g > q$). To simplify the exposition, the figure depicts the deterministic productivity dynamics and also abstracts from the common productivity trend a, which does not affect the relative productivity between firms.

The line labeled $\log(Q_t)$ in figure 1 indicates the cohort trend and captures the productivity of newly entering firms at each point in time. Lines departing from the trend line of the cohort capture the productivity dynamics of the entering cohorts over time. Since $g > q$, the productivity of existing firms grows faster than the productivity of new entrants, so that incumbent firms are more productive and thus larger than newly entering firms. In experience-adjusted terms, however, newly entering firms are the most productive firms in the economy. The dashed arrows indicate the productivity

[6]To insure that the paper is self-contained, this section reviews some of the key insights from Adam and Weber (2019). Readers familiar with this paper can directly proceed with the next section.

losses of exiting firms that have been hit by a δ-shock. For simplicity, the figure assumes that their productivity permanently drops to zero. As should be clear from the figure, the entry and exit dynamics imply an exponential distribution for firm age.[7]

Importantly, once firms have entered the economy, their productivity relative to other incumbent firms remains unchanged: all incumbent firms sustain the same experience trend (g) and the same common productivity trend (a). (The latter is not shown in the figure.) This holds true even for the case with shocks. Therefore, if firms' initial relative prices (inversely) reflect firms' relative productivities, there is no need for prices of incumbent firms to be adjusted in response to productivity disturbances.[8] To obtain efficiency of *all* relative prices, it is thus sufficient that newly entering firms set a relative price that (inversely) reflects their productivity relative to that of the average incumbent firm. Since newly entering firms in figure 1 have lower productivity than the average incumbent firm, the prices of newly entering firms in figure 1 should optimally be higher than the price of the average incumbent firm, causing inflation to be optimal.[9] In this way, the optimal inflation rate is related to the relative productivity between new and incumbent firms.

4 The Optimal Monetary Policy Problem

Despite the presence of heterogeneous firms with different accumulated experience and different initial cohort productivity, it is possible to aggregate output across firms in closed form and to derive a linear-quadratic approximation of the policymaker's Ramsey problem. The approximate problem features a generalized quadratic objective function and a generalized linear New Keynesian Phillips Curve. The approximate policy problem transparently reveals how the stochastic productivity trends (a_t, g_t, q_t), the tax/subsidy shocks (τ_t) and the preference shocks (ξ_t) shape the monetary policy tradoffs in the pres-

[7]Coad (2010) shows that such an age distribution is empirically plausible and how it generates, together with (productivity) growth shocks, a Pareto distribution for firm size, in line with the observed firm size distribution.

[8]This is not true following mark-up disturbances as we discuss later on.

[9]Note that the average price of exiting firms in the period before being hit by the exit shock is equal to the average price of incumbent firms, so that the prices of exiting firms gets replaced by the higher prices of newly entering firms. This is consistent with empirical evidence in Bils (2009) who shows that a large part of inflation is due to the higher prices charged for new products.

ence of firm heterogeneity. Since productivity trends also affect the steady-state value of optimal inflation and the natural rate of interest, we start by considering the effects of these trends on the efficient steady-state outcomes. Subsequently, we present the linear-quadratic stabilization problem around this steady state.

4.1 Optimal Steady-State Inflation and Real Rate

The optimal (gross) steady-state inflation rate is

$$\Pi^\star = \frac{g}{q},$$

where g denotes the (gross) experience productivity trend and q the (gross) cohort productivity trend. This is a special case of the multi-sector steady-state result derived in Adam and Weber (2023), see also Adam and Weber (2025) for a literature overview. The optimal gross inflation rate is larger than one whenever $g > q$. Incumbent firms then accumulate experience productivity at a rate faster than the rate at which successively entering cohorts become more productive. As a result, young firms are less productive than incumbent firms, in line with empirical evidence. As explained in section 3, this causes positive inflation to be optimal in steady state.

Interestingly, the strength of this effect is independent of the product turnover rate δ. Although a higher δ implies that more firms set a high relative price, it also implies that incumbent firms had less time to accumulate experience, so that the relative price does not need to be as high as with lower δ. As it turns out, the two effects exactly offset each other, so that the optimal steady-state inflation rate is independent of δ.

The (gross) real steady-state interest rate in the efficient equilibrium, R^e, is equal to

$$R^e = \frac{aq}{\beta}.$$

It decreases with the discount factor β and increases with the steady-state ouput and productivitiy growth rate aq.[10] With a positive steady-state growth rate of the economy, $aq > 1$, the real (gross) interest rate satisfies $R^e > 1$. Moreover, the nominal (gross) steady-state interest rate is larger

[10]The experience productivity trend g generates only level effects for aggregate productivity, because accumulated experience is ultimately lost when firms exit the economy.

than one if

$$\frac{ag}{\beta} > 1. \tag{13}$$

We can then abstract from the lower bound constraint on nominal interest rates in our linear-quadratic approximation for sufficiently small economic disturbances.

4.2 The Linear-Quadratic Policy Problem

To approximate the policymaker's Ramsey problem, we impose the following assumptions:

Assumption 1 1. Initial prices in $t = -1$ inversely reflect firms' relative productivities, to a first-order approximation:

$$P_{j,-1} = \Phi \frac{1}{Q_{-1-s_{j,-1}} G_{j,-1}} + O(2) \qquad \text{for all } j \in [0,1] \text{ and some } \Phi > 0,$$

where $O(2)$ is a second-order approximation error.

2. The steady-state sales subsidy is efficient, i.e., equation (11) holds.
3. The (gross) steady-state nominal interest rate is larger than one, i.e., inequality (13) holds.

These points are all standard in the sticky-price literature: the first insures that inefficient price dispersion is of second order in the initial period; the second guarantees that the sticky price steady state is efficient, so that it is sufficient to approximate the nonlinear Phillips curve and other equilibrium conditions to first order; the third point insures that the zero lower bound constraint on nominal rates does not bind in the steady state.[11]

We then approximate the optimal policy problem around the efficient steady state in which subsidy shocks are absent ($\tau_t = \tau$). We define the log-deviations of inflation and optimal inflation from (nonzero) optimal steady-state inflation as

$$\pi_t \equiv \ln \Pi_t - \ln \Pi^\star, \tag{14}$$

$$\pi_t^\star \equiv \ln \Pi_t^\star - \ln \Pi^\star. \tag{15}$$

[11] Unlike in the canonical model without economic growth, this does not follow from time discounting alone.

The output gap is defined as

$$x_t \equiv \ln Y_t - \ln Y_t^e, \tag{16}$$

where Y_t^e denotes the (stochastic) efficient level of output, as defined in Appendix A.3. We define the log-deviation of the natural rate of interest from its steady-state value as

$$r_t^n \equiv \ln R_t^e - \ln R^e \tag{17}$$

and the log-deviation of the nominal interest rate from its steady-state value as

$$\hat{\imath}_t \equiv \ln(1+i_t) - \ln(ag/\beta). \tag{18}$$

The mark-up shock, which captures the effect of time variation in the sales subsidy, is defined as

$$u_t \equiv \frac{\kappa}{1+\psi}\left(\ln\frac{1}{1+\tau_t} - \ln\frac{1}{1+\tau}\right), \tag{19}$$

where

$$\kappa \equiv \frac{(1-\alpha\rho)(1-\alpha\beta\rho)}{\alpha\rho}(1+\psi). \tag{20}$$

Using these definitions, we can state our main result:

Proposition 1 *Suppose Assumption 1 holds. A linear-quadratic approximation to the optimal monetary policy problem is then given by*

$$\max_{\{x_t,\pi_t,\hat{\imath}_t\}_{t=0}^{\infty}} - E_0 \sum_{t=0}^{\infty} \beta^t \left((\pi_t - \pi_t^\star)^2 + \lambda x_t^2 \right) \tag{21}$$

$$s.t.:$$

$$\pi_t - \pi_t^\star = \beta E_t\left[\pi_{t+1} - \pi_{t+1}^\star\right] + \kappa x_t + u_t \tag{22}$$

$$x_t = E_t x_{t+1} - (\hat{\imath}_t - E_t \pi_{t+1}) + r_t^n, \tag{23}$$

with $\lambda \equiv \kappa/\theta$ *and*

$$\pi_t^\star = \rho\pi_{t-1}^\star + (1-\rho)\cdot(\widehat{g}_t - \widehat{q}_t), \tag{24}$$

where $\pi_{-1}^\star$ *given and*

$$\rho \equiv (1-\delta)(\Pi^\star)^{(\theta-1)} < 1.$$

The natural rate evolves according to

$$r_t^n = E_t\left[\widehat{a}_{t+1} + \widehat{g}_{t+1} - \widehat{\xi}_{t+1} - \pi_{t+1}^\star\right]. \tag{25}$$

The proof of the proposition can be found in appendices B and C.[12] The policy problem in proposition 1 differs from the one that arises in a canonical homogeneous firm setup along four important dimensions.

First, it follows from equation (24) that the New Keynesian Phillips curve (22) with firm heterogeneity can alternatively be expressed as

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + u_t + z_t, \tag{26}$$

where $z_t \equiv \pi_t^\star - \beta E_t \pi_{t+1}^\star$. Viewed through the lens of the canonical homogeneous firm model, the Phillips Curve (26) thus features two kinds of mark-up shocks, namely shocks that produce a trade-off between output gap (x_t) and inflation gap ($\pi_t - \pi_t^\star$) stabilization, namely the mark-up shocks u_t, and shocks that do not produce such a trade-off, namely the mark-up shocks z_t). The latter shocks nevertheless cause the optimal inflation rate (actual and expected) to move over time.

Interestingly, equation (24) shows that purely *temporary* shocks to experience or cohort productivity growth ($\widehat{g}_t$, $\widehat{q}_t$) give rise to *persistent* mark-up disturbances z_t in the Phillips Curve. The heterogeneous firm model thus delivers an alternative theory of persistent mark-up disturbances that is grounded in productivity movements. This is of interest because persistent mark-up shocks have been found to be quantitatively important when bringing the New Keynesian Phillips Curve to the data (Smets and Wouters (2003, 2007)). At the same time, the traditional micro-foundations for these shocks, i.e., variations in taxes/subsidies or firm monopoly power, have been called into question, e.g., Chari, Kehoe and McGratten (2009).

Second, the policy problem (21)-(23) is (generically) approximated around a steady state with nonzero optimal inflation ($\ln \Pi^\star = \ln g/q \lesseqgtr 0$), while the homogeneous firm model is approximated around a zero optimal steady-state inflation rate. For the empirically plausible case in which young firms are initially less productive and thus smaller than older firms, we have $g/q > 1$ and thus $\ln \Pi^\star > 0$. Similarly, when the model is interpreted as one featuring product turnover instead of firm turnover, the empirical evidence in Adam and Weber (2023) for the U.K. and Adam, Gautier, Santoro, and Weber (2022) for Germany, France and Italy reveals that the relative price of products falls over the product life cycle, which also requires $g/q > 1$.

[12]The appendices derive results directly for the generalized case with price indexation considered in section 7. For the special case without price indexation, $\varphi = 0$, this delivers the result in proposition 1.

Third, the optimal inflation rate is time-varying and depends on the shocks to experience productivity growth $\widehat{g}_t$ and cohort productivity growth $\widehat{q}_t$, see equation (24). In the canonical homogeneous firm model, only mark-up shocks move the optimal inflation rate. In the presence of firm heterogeneity, this is also true for productivity disturbances.

Lastly, the Phillips curve slope κ in equation (20) generally differs from that in the standard model and depends on the steady-state inflation rate via the parameter ρ. For the special case in which the optimal steady-state inflation rate is given by $\ln \Pi^\star = \ln g/q = 0$, the slope of the Phillips curve is

$$\kappa|_{\Pi^\star=1} = \frac{(1-\alpha(1-\delta))\,(1-\alpha(1-\delta)\beta)}{\alpha(1-\delta)}(1+\psi), \tag{27}$$

where $\alpha(1-\delta)$ is the overall degree of price rigidity taking into account that newly entering firms can also freely set their price. The slope of the Phillips curve is then identical to that in the canonical homogeneous firm model with overall price rigidity $\alpha(1-\delta)$. However, in the present setup, the Phillips curve slope in equation (20) falls with the optimal steady-state inflation rate $\Pi^\star$ because

$$\frac{\partial \kappa}{\partial \Pi^\star} < 0$$

for all $\Pi^\star$ satisfying the existence condition (5). The Phillips curve is therefore flatter than in the canonical New Keynesian model when the optimal steady-state inflation rate is larger than zero ($\ln \Pi^\star > 0$).

Despite these fundamental differences, the monetary policy problem (21)-(23) is isomorphic to the problem in the canonical New Keynesian model with homogeneous firms (Woodford (2011)). Specifically, in the absence of a binding lower-bound constraint, the Euler equation (23) does not constrain the problem and can be eliminated. The resulting simpler optimal policy problem, which involves only the output gap (x_t) and the inflation gap ($\pi_t - \pi_t^\star$), is then the same as in the canonical model if one replaces the inflation gap ($\pi_t - \pi_t^\star$) by inflation (π_t). Therefore, all lessons from the canonical homogeneous firm model about the optimal behavior of the output gap and inflation carry over to the output gap and the inflation *gap* in the model with firm heterogeneity.

The next sections explore in greater detail the dynamics of optimal inflation and efficient output growth.

4.3 The Dynamics of Optimal Inflation

Equation (24) reveals that a one-time positive shock to the growth rate of experience productivity ($\widehat{g}_t$) persistently increases optimal inflation. In contrast, a positive one-time positive shock to cohort productivity growth ($\widehat{q}_t$) persistently decreases optimal inflation. The optimal inflation rate remains unaffected following shocks to the growth rate of common productivity ($\widehat{a}_t$) and the growth rate of the time discount factor ($\widehat{\xi}_t$).[13]

To understand the underlying economic forces, note that shocks to common productivity ($\widehat{a}_t$) and shocks to the growth rate of the discount factor ($\widehat{\xi}_t$) both do not affect the *relative* productivity between new and incumbent firms. As a result, the efficient relative price between new and incumbent firms remains unchanged following these disturbances.

This is different for shocks to the growth rate of experience productivity $\widehat{g}_t$. A temporary positive growth rate shock, for instance, permanently increases the productivity of incumbent firms, while leaving the productivity of newly entering firms unaffected (also in future periods). Efficiency then requires that the price charged by new firms increases relative to the price charged by incumbent firms. Since incumbent firms' prices are partially sticky, it is optimal to implement the required relative price change via an upward adjustment in the nominal price charged by new firms, i.e., by higher inflation.[14] Since the productivity increase for incumbent firms only slowly fades via the gradual exit of incumbent firms, the increase in optimal inflation following a purely temporary shocks to $\widehat{g}_t$ is rather persistent.

Correspondingly, a temporary positive shock to the growth rate of cohort productivity $\widehat{q}_t$ causes new firms to become permanently more productive, while the productivity of incumbent firms is now unaffected. The price charged by new firms must therefore fall in relation to that charged by incumbent firms. Again, it is efficient to implement this relative price adjustment via new firms. However, this time, new firms must charge lower prices, leading to lower optimal inflation. As before, this effect is rather persistent and only fades as new (and more productive) cohorts of young firms gradually

[13]Optimal inflation does not respond to the mark-up shocks (u_t) because these shocks are absent in the efficient equilibrium around which we approximate the policy problem. We discuss the optimal response to mark-up shocks in section 4.5.

[14]Having incumbent firms reduce their nominal prices is inefficient as price stickiness prevents all incumbent firms from doing so at the same time, causing inefficient price dispersion.

replace older firm cohorts.

Interestingly, if $\widehat{g}_t$ or $\widehat{q}_t$ were to move permanently, say due to structural changes in the economy, then the optimal inflation rate would also move permanently. In fact, the coefficients in equation (24) imply that the long-term effect of permanent changes in $\widehat{g}_t$ or $\widehat{q}_t$ transmits one for one to optimal inflation, albeit with different signs, independently of the firm turnover rate (δ).

However, the transitional dynamics following such permanent changes depend on the firm turnover rate. To understand why, consider the special case with $\Pi^\star = 1$, so that $\rho = 1 - \delta$. Equation (24) then simplifies to

$$\pi_t^\star = (1 - \delta)\, \pi_{t-1}^\star + \delta \cdot (\widehat{g}_t - \widehat{q}_t)\,, \tag{28}$$

and shows that the impact effect of the shocks $\widehat{g}_t$ and $\widehat{q}_t$ on the optimal inflation rate is lower for lower turnover rates δ: fewer firms will then have to adjust relative prices following these shocks. At the same time, the autoregressive coefficient $(1-\delta)$ in equation (28) increases with lower turnover, making the transition longer lasting. Taken together, the long-run effect of permanent shocks is *independent* of the turnover rate.

In the more general case with $\Pi^\star \neq 1$, impact and autoregressive coefficients are adjusted to take account of the fact that the steady-state expenditure share of new firms is not only a function of the relative number of new firms, but also of their relative productivity. For example, if new firms are less productive than incumbent firms in the steady state ($g > q$, which implies $\Pi^\star > 1$), then their expenditure share is lower than δ, so that the impact coefficient in equation (24) is smaller than δ.

4.4 The Dynamics of Efficient Output Growth

The efficient growth rate of real output is given by

$$\widehat{\gamma}_t^e = \rho\widehat{\gamma}_{t-1}^e + (\widehat{a}_t - \rho\widehat{a}_{t-1}) + \rho(\widehat{g}_t - \widehat{g}_{t-1}) + (1 - \rho)\widehat{q}_t\ , \tag{29}$$

and depends on all three supply disturbances, see appendix B.5.

A temporary increase in common productivity growth $\widehat{a}_t$ increases output growth once but has no additional effects on output growth in subsequent periods, so the economy is catapulted to a permanently higher output level.

A temporary increase in experience growth $\widehat{g}_t$ also increases output in the current period, with an effect that is close to one reflecting the large share of incumbent firms in the economy. However, experience growth shocks result in negative output growth in subsequent periods because the long-term output effect of these shocks is zero, since no firm remains in the economy forever. Therefore, a temporary shock to experience productivity growth results in a persistent increase in the level of output that fades out slowly over time.

In contrast, a temporary increase in the growth rate of cohort productivity $\widehat{q}_t$ results in a permanent increase in long-run output. Yet, unlike with shocks to common productivity growth ($\widehat{a}_t$), the effect is initially small and gradually builds up over time, as only few new firms enter the economy in each period.

Equation (29) also shows that permanent changes in common productivity growth $\widehat{a}_t$ increase output growth permanently and one for one. Permanent changes in cohort productivity growth $\widehat{q}_t$ also permanently increase output growth, but the effect again builds up only gradually over time. In contrast, permanent changes to experience productivity growth $\widehat{g}_t$ only lead to a temporary increase in output growth and ultimately only cause a level shift in output.

4.5 Optimal Stabilization Policy

This section discusses optimal monetary policy and its implementation via targeting and instrument rules, highlighting similarities and differences with the canonical New Keyensian model.

Under commitment, the first-order conditions to problem (21)-(23) imply the targeting criterion

$$\pi_t - \pi_t^\star = -\frac{\lambda}{\kappa}(x_t - x_{t-1}). \tag{30}$$

The targeting criterion features the inflation gap on the left-hand side and the (negative) change in the output gap on the right-hand side. Relative to the canonical New Keynesian model, the targeting criterion additionally features the time-varying optimal inflation rate $\pi_t^\star$.

As in the canonical homogeneous firm model, commitment to the targeting criterion (30) implements the Ramsey outcome as the locally unique rational expectations equilibrium. In the Ramsey outcome, the output gap and the inflation gap can both be closed in response to the disturbances $(\widehat{a}_t, \widehat{q}_t, \widehat{g}_t, \widehat{\xi}_t)$, while mark-up shocks u_t generate a trade-off between output gap and inflation *gap* stabilization.

Defining the natural nominal interest rate as

$$\hat{\imath}_t^n \equiv r_t^n + E_t \pi_{t+1}^\star, \tag{31}$$

a Taylor rule of the form

$$\hat{\imath}_t = \hat{\imath}_t^n + \alpha_\pi \left(\pi_t - \pi_t^\star\right) \tag{32}$$

implements the Ramsey optimal monetary policy response to the disturbances $(\widehat{a}_t, \widehat{q}_t, \widehat{g}_t, \widehat{\xi}_t)$ as a locally unique equilibrium outcome, provided $\alpha_\pi > 1$. However, this Taylor rule does not implement the optimal policy response to mark-up disturbances u_t.

An interesting difference between the models with homogeneous versus heterogeneous firms concerns the determinacy property of Taylor rules of the form (32), including Taylor rule variants without time-varying intercepts: ensuring local determinacy only requires $\alpha_\pi > 1$, as in the canonical New Keynesian model. This differs from findings in Gorodnichenko and Coibion (2011) and Ascari and Sbordone (2014) who consider determinacy properties in the canonical homogeneous firm model, linearized around a non-optimal non-zero steady-state inflation rate. They find that conditions for determinacy become tighter for higher steady-state inflation rates, unlike in the heterogeneous firm model considered here.

5 "Looking Through" Supply Disturbances

This section shows that it is optimal for monetary policy to "look through" temporary disturbances to experience productivity growth $\widehat{g}_t$ and temporary or persistent disturbances to cohort productivity growth $\widehat{q}_t$. In particular, it is optimal that nominal interest rates stay constant following these shocks, while output and inflation move.

Monetary policymakers often express their desire to "look through" supply disturbances.[15] However, this desire is hard to justify through the lens of a canonical New Keynesian model with homogeneous firms. In that framework, common productivity disturbances move the natural real rate but not inflation and thus require corresponding adjustments in the nominal rate. In fact, since the optimal policy response implies no movement in inflation, the canonical New Keynesian model does *not* justify a look-through approach

[15] See for instance Powell (2023) or Lagarde (2025).

in any way. The same holds for mark-up disturbances. These generate a policy trade-off in the canonical model between output gap and inflation stabilization and therefore require an appropriate interest rate response.

Interestingly, these results continue to be true in the heterogeneous firm setup: neither disturbances to common productivity growth nor mark-up disturbances rationalize a look-through approach.

However, the heterogeneous firm setup features two new supply disturbances $\widehat{g}_t$ and $\widehat{q}_t$ that can justify a look-through approach as being optimal. In particular, purely temporary movements in $\widehat{g}_t$ and both temporary and persistent movement in $\widehat{q}_t$ do not move the natural nominal rate, which can be expressed according to:[16]

$$\hat{\imath}_t^n = E_t \left[\widehat{a}_{t+1} + \widehat{g}_{t+1} - \widehat{\xi}_{t+1} \right]. \tag{33}$$

Following these disturbances, it is optimal to *not* adjust the policy rate, despite the fact that these disturbances generate persistent movements in output growth and inflation, see the discussions in sections 4.3 and 4.4.

Figure 2 illustrates this outcome by depicting the Ramsey optimal response of inflation, detrended output and nominal rates.[17] The top row considers a positive and temporary experience productivity growth shock $\widehat{g}_0 > 0$ and the bottom row a positive, temporary cohort productivity growth shock $\widehat{q}_0 > 0$.[18] The figure shows that output and inflation respond persistently and in the same direction for experience productivity shocks, but in opposite directions for cohort productivity shocks. The optimal response of policy rates is to "look through" these responses by keeping nominal rates constant.

The look-through approach continues to be optimal when shocks to cohort productivity growth $\widehat{q}_t$ are persistent because expectations of future cohort productivity growth do not affect the natural nominal rate in equation (33). The reason for this outcome is that the effects of these shocks on the natural real rate and on the expected optimal inflation rate offset each other exactly. [19]

[16] This follows from equations (25) and (31).

[17] Detrended output is defined as $\widetilde{y}_t = \ln Y_t - \ln \overline{\Gamma}^e_t$, where $\overline{\Gamma}^e_t$ is the deterministic output level emerging in the absence of the shock, as derived in appendix B.5.

[18] The figure assumes $\rho = 0.95$ and considers shocks of unit size. No other parameter matters for impulse response dynamics in this figure.

[19] This is not true for persistent experience productivity growth shocks $\widehat{g}_t$. Following

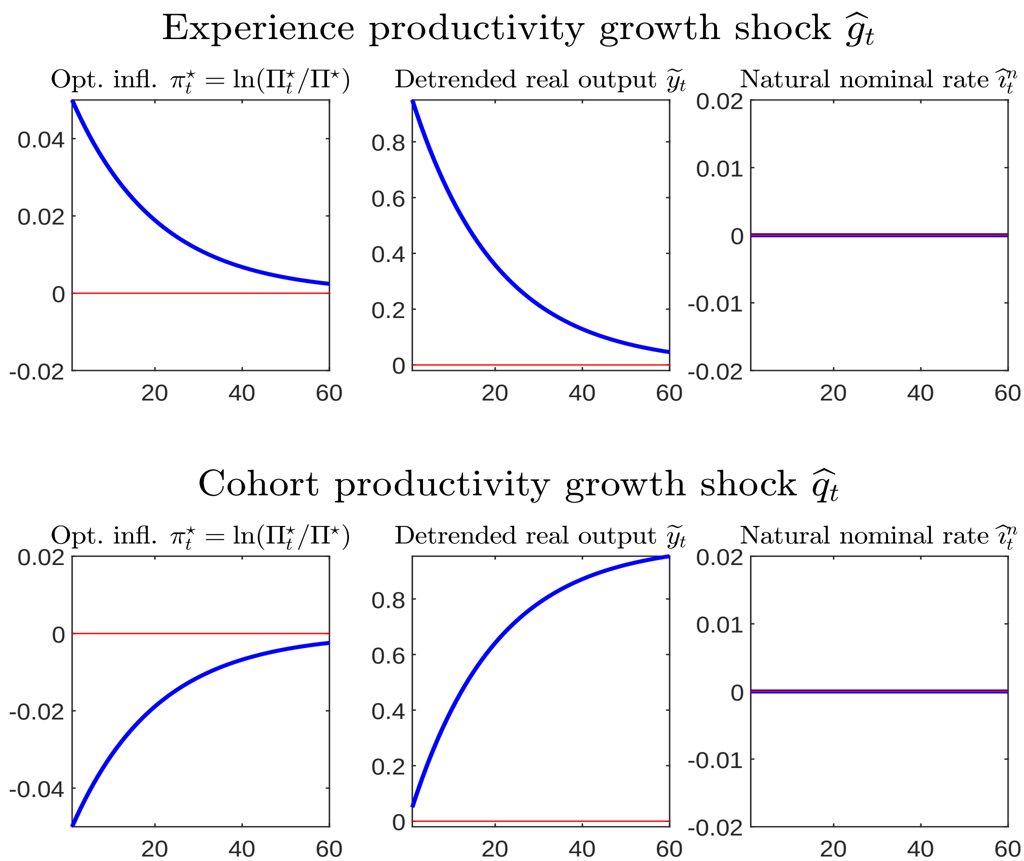

Figure 2: Looking through supply shocks: Impulse responses of temporary shocks to $\widehat{g}_t$ and $\widehat{q}_t$.

This implies that a period in which the initial productivity of young firms is persistently accelerating, as is generally considered to have happened during the New Economy boom in the mid-to-late 1990's in the United States, is a period in which aggregate growth accelerates, inflationary pressures become weaker, all the while nominal interest rates optimally stay constant. This finding can broadly rationalize Alan Greenspan's response to the New Economy boom during this period.

---

such shocks, monetary policy must optimally increase nominal interest rates, see equation (33).

6 Supply Shocks Mimic Demand Shocks

This section shows that temporary shocks to experience productivity growth $\widehat{g}_t$ are empirically very hard to distinguish from demand shocks such as monetary policy shocks or discount factor shocks. In particular, we illustrate how sign-based identification approaches of demand shocks may actually identify supply shocks. Similar challenges can arise with high-frequency identification approaches of monetary policy shocks.

These findings are of interest because they provide an alternative explanation for the somewhat puzzling empirical observation that measures of aggregate productivity increase following identified monetary policy loosenings. The literature attributes these productivity responses to the presence of steady-state misallocations that are temporarily reduced (aggravated) after positive (negative) demand shocks; see Meier and Reinelt (2024) and Baqaee et al. (2024). Our findings suggest that they could alternatively emerge because shocks to experience productivity growth are mistakenly interpreted as monetary policy or demand shocks.

6.1 Pitfalls of Sign-Based Demand Shock Identification

This section shows that traditional sign-based identification approaches of demand shocks may actually also identify supply shocks within our heterogeneous firm setup.

To illustrate this point, suppose the economy is at its deterministic steady state in $t = -1$. In $t = 0$, a purely temporary, positive shock to experience productivity growth $\widehat{g}_0 > 0$ hits the economy. This leads to a persistent fall in the natural real rate:[20]

$$r_t^n = -\rho^{t+1}(1-\rho) \cdot \widehat{g}_0 \tag{34}$$

Under optimal monetary policy, the nominal rate follows its natural level and simply remains at its steady-state value:

$$\hat{\imath}_t = \hat{\imath}_t^n = 0,$$

where the second equality follows from equation (33). Letting $\overline{\Gamma}_t^e$ denote the deterministic output level in the absence of the shock and $\widetilde{y}_t = \ln Y_t - \ln \overline{\Gamma}_t^e$

[20]This follows from equations (24) and (25).

the log deviation from this deterministic level, we get[21]

$$\widetilde{y}_t = \rho^{t+1}\widehat{g}_0 = -\frac{1}{1-\rho} r_t^n,$$

which shows that output persistently increases. The inflation response is given by[22]

$$\pi_t = \rho^t(1-\rho) \cdot \widehat{g}_0 = -\frac{1}{\rho} r_t^n$$

and also displays a persistent increase.

Detrended aggregate productivity also increases in line with detrended output, but ultimately all variables return to their pre-shock value in detrended terms. Since the supply shock produces no long-run response for any variable, it satisfies the sign-based identifying restrictions typically used in empirical demand shock identification, which require output and inflation to move in the same direction without displaying permanent effects; see, for instance, Giannone and Primiceri (2024) and Bergholt et al. (2025).

Monetary policy surprises also produce a persistent increase in output and inflation within our setup. Specifically, we consider a setting in which monetary policy announces in $t = 0$ to implement a path for nominal rates that generates the same lower real interest rate path (34), as induced for the natural rate by the supply shock. We then have the following result:

Lemma 2 *Suppose the economy is in steady state in $t = -1$ and monetary policy unexpectedly announces in $t = 0$ to implement a path for nominal rates that gives rise to the path of real interest rates (34). We then have*

$$\begin{pmatrix} \pi_t \\ \widetilde{y}_t \end{pmatrix} = \begin{pmatrix} -\frac{\kappa}{(1-\rho)(1-\beta\rho)} \\ -\frac{1}{(1-\rho)} \end{pmatrix} r_t^n. \tag{35}$$

The proof of lemma 2 can be found in appendix D. In line with conventional wisdom, temporarily lower real rates persistently increase output and inflation, without generating long-run effects. The difficulty for sign-based

[21]We focus on deviations of output from trend rather than on the output gap, because the latter is not directly observed in empirical work. Appendix B.5 shows how the path of detrended output can be computed. For the case here, where the output gap is closed, output simply evolves according to $\widetilde{y}_t = \widehat{\gamma}_t^e + \widetilde{y}_{t-1}$, where $\widehat{\gamma}_t^e$ is determined in equation (29).

[22]This follows from equation (24).

demand shock identification is that temporary experience growth shocks can do the same as illustrated above.

One difference between the experience growth shock considered above and the monetary policy surprise is that the nominal rate falls in the latter case but stays constant in the former (under optimal policy). The following lemma shows that this difference is not key for our findings:

Lemma 3 *Suppose all agents learn in $t = 0$ about a sequence of preference shocks $\{\widehat{\xi}_t\}_{t=0}^{\infty}$ that induce the natural rate path*

$$r_t^n = \rho^{t+1}(1-\rho)\widehat{g}_0 - \frac{\kappa}{1-\beta\rho}\rho^{t+2}\widehat{g}_0, \tag{36}$$

but that monetary policy keeps nominal rates constant at the steady state value. The response of the economy is then again given by equation (35).

The proof of lemma 3 can be found in appendix E. Lemma 3 shows that preference shocks give rise to a persistent move of output and inflation in the same direction, when nominal rates do not move, as is the case with the supply shock $\widehat{g}_0$. This shows that there exists considerable potential to misidentify supply shocks as demand shocks when relying on commonly used sign-based identification restrictions.

6.2 Pitfalls of High-Frequency Monetary Policy Shock Identification

The potential to misidentify supply shocks as demand shocks exists even when using sophisticated high-frequency identification approaches, as pioneered by Kuttner (2001). These approaches impose that interest rate surprises occur within a narrow time window, typically set around central bank press conferences.

To illustrate this point, consider again an economy that in $t = -1$ is at its deterministic steady state. In $t = 0$, the central bank reveals during its press conference that the economy has been hit by a positive shock to experience productivity growth that is going to be partly reversed in the future ($\widehat{g}_0 > 0$ and $\widehat{g}_t < 0$ for $t > 0$ with $-\sum_{t>0}\widehat{g}_t < \widehat{g}_0$).

Figure 3 illustrates the response of output, inflation and nominal interest rates to such a supply shock surprise.[23] Under optimal policy, nominal rates

[23] The figure assumes $\rho = 0.95$ and considers the supply shock $\widehat{g}_0 = 1$, $\widehat{g}_t = -0.1(0.75)^t$ for $1 \leq t \leq 24$ and $\widehat{g}_t = 0$ for $t > 24$.

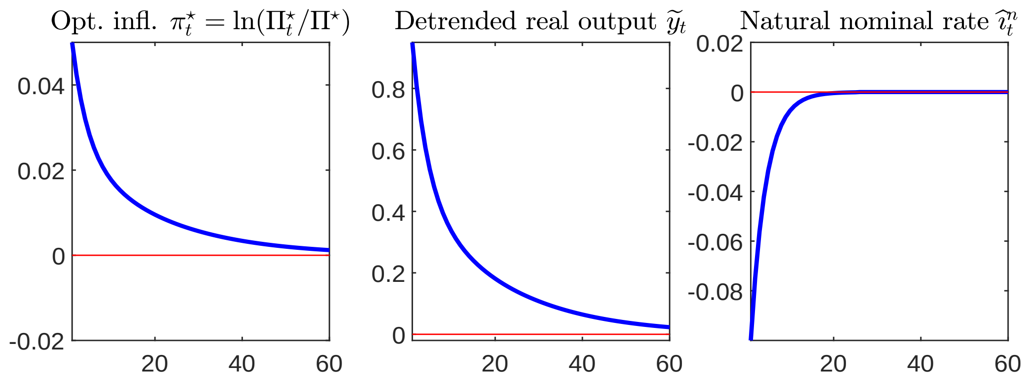

Figure 3: Impulse responses of a temporary and partly reversed shock to $\widehat{g}_t$.

decrease, which follows directly from equation (33), while output persistently increases due to increased experience productivity. The inflation response follows directly from equation (24). The impulse response looks very much like that associated with a monetary policy surprise loosening.

It is well-known, however, that central bank press conferences also reveal information about the central bank's view about the state of the economy. Therefore, it has become standard in high-frequency identification to attempt to filter out so-called central bank 'information effects'. Following Jarocinski and Karadi (2020), much of the empirical literature requires that a surprise loosening of monetary policy is associated with a positive stock market response. The key objective is to rule out that the surprise loosening of monetary policy is due to the central bank announcing unexpectedly bad news about the economy, which then results in a surprise loosening that is associated with output and inflation potentially falling, while output and inflation would rise following a monetary policy shock, i.e., a loosening that is not due to the state of the economy.

In the present setup, the surprise loosening in response to the experience growth shocks considered above is associated with *positive* news about the economy. Figure 3 shows that real output temporarily increases, which is due to a temporary increase in aggregate productivity. This causes real firms profits to increase in proportion with real output.[24] Moreover, the real

[24] Appendix B.6 shows that real profits move proportionately with real output when the output gap is closed and there are no mark-up shocks, as is the case here.

discount rate falls, as real interest rates fall with lower nominal rates and higher inflation, see figure 3. Since the real present value of discounted profits unambigously increases, the stock market can be expected to appreciate.

The supply shock announced by monetary policy during the presse conference thus passes the standard high-frequency identification restriction for a monetary policy loosening, even when seeking to filter out central bank information effects. Interestingly, the considered supply shock gives rise to a first-order increase in detrended aggregate productivity. This may offer a complementary explanation for the empirical finding that measured productivity rises following identified monetary policy shocks.[25]

6.3 Potential Remedies to Mis-Identification

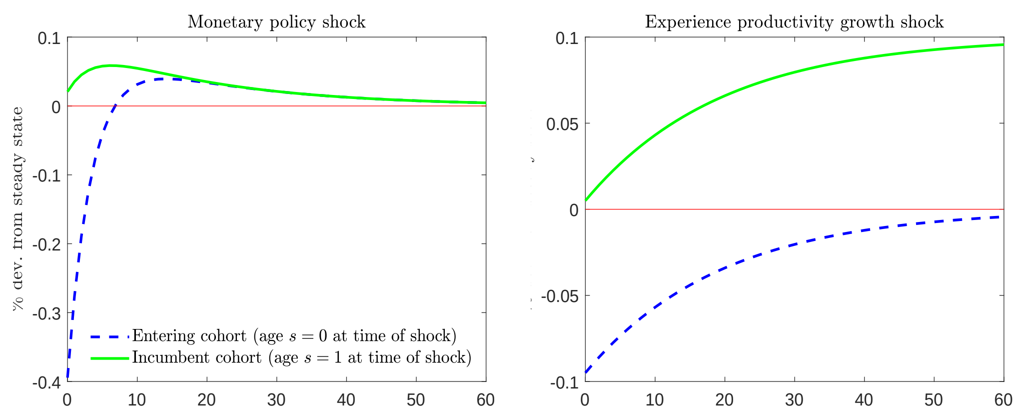

Figure 4: Cohort-level expenditure shares: Monetary policy easing shock (left) and positive experience productivity growth shock (right)

The previous section illustrated the potential to misidentify supply shocks as demand shocks when relying on aggregate variables. This section shows that micro-level information can be useful in overcoming the misidentification problem.

Figure 4 depicts impulse responses of the expenditure shares for different firm cohorts.[26] In particular, it contrasts the expenditure share response for

[25]Productivity does not respond to first order following a true monetary policy shock, e.g., the one considered in lemma 2.

[26]Appendix F describes the computation of cohort-level impulse responses underlying

firms that newly enter the economy in the period of the shock to that of firms that entered in the previous period (all in deviation from steady-state shares). The left-hand side panel shows the impulse response following the monetary policy shock considered in lemma 2, while the right-hand side panel shows the impulse response following the temporary experience productivity growth shock $\widehat{g}_0 > 0$ considered at the beginning of section 6.1.

While incumbent firms gain market share and new firms lose market share in both shock scenarios, the divergence is only temporary following the monetary policy shock. In contrast, an experience productivity growth shock leads to a permanent divergence of expenditure shares between new and incumbent firms. This shows how micro-level information may be useful in overcoming the identification challenges. However, for the case where experience shocks are partially reversed over time, as considered in section 6.2, the gap between expenditure shares also partly closes over time. Thus, empirically distinguishing such shocks from true monetary policy shocks remains challenging, despite the use of micro-level information.

7 Generalized Setup With Price Indexation

In applied work, researchers often prefer to include price indexation of non-optimizing firms into the model to get an inflation process that is inherently more persistent. This section shows that the linear-quadratic approximation of the monetary policy problem with firm heterogeneity naturally generalizes to the case with price indexation. To obtain a formulation of the problem that nests the one in the canonical model without firm heterogeneity, we consider price indexation of the form

$$P_{jt+1} = P_{jt} \left(\frac{\Pi_t}{\Pi^\star} \right)^{\varphi},$$

for all firms that do not reoptimize prices, where $\varphi \in [0, 1]$ is the indexation parameter. Price indexation only reacts to deviations of actual inflation from its optimal steady-state value $\Pi^\star$. For the special case where optimal steady-state inflation $\Pi^\star = 1$ and $\varphi = 1$, the scheme reduces to the one considered in Christiano, Eichenbaum and Evans (2005). We then have the following result:

figure 4. The figure assumes $\theta = 3, \kappa = 0.1, \alpha = 0.75, \rho = 0.95$ and $\beta = 0.99$.

Proposition 4 *Suppose Assumption 1 holds. A linear-quadratic approximation to the optimal monetary policy problem is then given by*

$$\max_{\{x_t,\pi_t,\hat{\imath}_t\}_{t=0}^{\infty}} -E_0 \sum_{t=0}^{\infty} \beta^t \left[\left(\pi_t - \pi_t^\star - \varphi \left(\pi_{t-1} - \pi_{t-1}^\star \right) \right)^2 + \lambda x_t^2 \right] \tag{37}$$

$$s.t.:$$

$$\pi_t - \pi_t^\star - \varphi \left(\pi_{t-1} - \pi_{t-1}^\star \right) = \beta E_t \left[\pi_{t+1} - \pi_{t+1}^\star - \varphi \left(\pi_t - \pi_t^\star \right) \right] + \kappa x_t + u_t$$

$$x_t = E_t x_{t+1} - \left(\hat{\imath}_t - E_t \pi_{t+1} \right) + r_t^n,$$

with π_{-1} *given,* $\lambda \equiv \kappa/\theta$*, and*

$$\pi_t^\star - \varphi \pi_{t-1}^\star = \rho \left(\pi_{t-1}^\star - \varphi \pi_{t-2}^\star \right) + (1-\rho) \cdot \left(\widehat{g}_t - \widehat{q}_t \right),$$

with $\pi_{-1}^\star$ *and* $\pi_{-2}^\star$ *given, and*

$$\rho \equiv (1-\delta) \left(\Pi^\star \right)^{(\theta-1)} < 1.$$

The natural rate evolves according to

$$r_t^n = E_t \left[\widehat{a}_{t+1} + \widehat{g}_{t+1} - \widehat{\xi}_{t+1} - \left(\pi_{t+1}^\star - \varphi \pi_t^\star \right) \right]. \tag{38}$$

The proof can be found in appendices B and C. Under commitment, the first-order conditions to this policy problem imply the generalized targeting criterion

$$\pi_t - \pi_t^\star - \varphi(\pi_{t-1} - \pi_{t-1}^\star) = -\frac{\lambda}{\kappa}(x_t - x_{t-1}), \tag{39}$$

which now features the change of the inflation *gap* on the left-hand side instead of simply the change of inflation, as is the case in the homogeneous firm model.

8 Conclusions and Outlook

Introducing a firm life cycle into New Keynesian models generates several appealing policy implications: the optimal inflation target is non-zero in steady-state and varies over time following supply disturbances; optimal monetary policy "looks through" certain supply disturbances, unlike in the homogeneous firm setup; and the identification of demand shocks is more challenging than implied by the homogeneous firm model because some supply disturbances move output and inflation in the same direction under optimal policy.

These implications of the heterogeneous firm framework also raise new challenges for monetary policy. For instance, to be able to implement optimal policy, it is key that policymakers distinguish between traditional mark-up disturbances in the Phillips curve, which generate a trade-off between output gap and inflation gap stabilization, and a new Phillips curve disturbance, which fails to generate such a trade-off because it reflects movements in the optimal inflation rate.

Also, the framework suggests that the optimal inflation rate is affected by structural change in the economy. Declining business dynamism in the form of lower firm turnover, for instance, implies that movements in optimal inflation become more persistent over time. The emergence of innovative productivity-enhancing technologies, such as artificial intelligence, affects the optimal inflation rate in ways that depend on whether these technologies enter the economy predominantly via new or existing firms.

The potentially changing nature of optimal inflation implies that monetary policy institutions are well-advised to conduct regular framework reviews reassessing whether their inflation targets are still appropriate. In fact, the present analysis suggests that even the optimal steady-state inflation target is beyond the control of monetary policy and depends on structural parameters of the economy that may gradually evolve over time.

References

Adam, K., E. Gautier, S. Santoro, and H. Weber (2022): "The Case for a Positive Euro Area Inflation Target: Evidence from France, Germany and Italy," *Journal of Monetary Economics*, 132, 140–153.

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

——— (2023): "Estimating the Optimal Inflation Target from Trends in Relative Prices," *American Economic Journal: Macroeconomics*, 15(3), 1–42.

——— (2025): "The optimal inflation target: bridging the gap between theory and policy," in *Research Handbook on Inflation*, ed. by G. Ascari, and R. Trezzi, chap. Chapter 5. Edward Elgar.

Ascari, G., and A. M. Sbordone (2014): "The Macroeconomics of Trend Inflation," *Journal of Economic Literature*, 52(3), 679–739.

Atkeson, A., and P. J. Kehoe (2005): "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 113(5), 1026–1053.

Bai, Y., V. Ríos-Rull, and K. Storesletten (2025): "Demand Shocks as Productivity Shocks," *Review of Economic Studies (forthcoming).*

Baqaee, D., E. Farhi, and K. Sangani (2024): "Supply Side Effects of Monetary Policy," *Journal of Political Economy*, 132(4), 1065–1112, Journals.

Bergholt, D., F. Canova, F. Furlanetto, N. Maffei-Faccioli, and P. Ulvedal (2025): "What Drives the Recent Surge in Inflation? The Historical Decomposition Roller Coaster," *American Economic Journal: Macroeconomics (forthcoming).*

Bils, M. (2009): "Do Higher Prices for New Goods Reflect Quality Growth or Inflation?," *Quarterly Journal of Economics*, 124, 637–675.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2009): "New Keynesian Models: Not Yet Useful for Policy Analysis," *American Economic Journal: Macroeconomics*, 1(1), 242–66.

Christiano, L. J., M. Eichenbaum, and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Christiano, L. J., M. Trabandt, and K. Walentin (2010): "Chapter 7 - DSGE Models for Monetary Policy Analysis," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3 of *Handbook of Monetary Economics*, pp. 285–367. Elsevier.

Coad, A. (2010): "The Exponential Age Distribution and the Pareto Firm Size Distribution," *Journal of Industry, Competition and Trade*, 10(3), 389–395.

Coibion, O., and Y. Gorodnichenko (2011): "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, 101(1), 341–70.

Giannone, D., and G. Primiceri (2024): "The Drivers of Post-Covid Inflation," NBER Working Paper 32859.

Jarocinski, M., and P. Karadi (2020): "Deconstructing Monetary Policy Surprises - The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12(2), 1–43.

Kuttner, K. N. (2001): "Monetary Policy Surprises and Interest Rates: Evidencefrom the Fed Funds Futures Market," *Journal of Monetary Economics*, 47, 523–544.

Lagarde, C. (2025): "Speech at the 25th "ECB and Its Watchers Conference" in Frankfurt, Germany," https://www.ecb.europa.eu/press/key/date/2025/html/ecb.sp250312 915537d675.en.html.

Lorenzoni, G. (2009): "A Theory of Demand Shocks," *American Economic Review*, 99(5), 2050–84.

Meier, M., and T. Reinelt (2024): "Monetary Policy, Markup Dispersion, and Aggregate TFP," *Review of Economics and Statistics*, 106(4), 1012–1027.

Powell, J. (2023): "Opening Remarks to the Policy Panel at the 24th Jacques Polak Annual Research Conference," https://www.federalreserve.gov/newsevents/speech/powell20231109a.htm?

Smets, F., and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123–1175.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Woodford, M. (2003): *Interest and Prices*. Princeton University Press, Princeton.

——— (2011): "Optimal Monetary Stabilization Policy," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3B. Elsevier, Amsterdam.

A Economic Model Setup

We consider the model in Adam and Weber (2019) and augment it with a time-varying sales subsidy.

A.1 Price Setting with Time-Varying Sales Subsidy

The firm that can optimize its price in period t solves the following problem:

$$\max_{P_{jt}} E_t \sum_{i=0}^{\infty} (\alpha(1-\delta))^i \Omega_{t,t+i} \left((1+\tau_{t+i}) \frac{P_{jt+i}}{P_{t+i}} Y_{jt+i} - \frac{W_{t+i}}{P_{t+i}} \frac{Y_{jt+i}}{A_{t+i} Q_{t-s_{jt}} G_{jt+i}} \right) \tag{40}$$

$$\begin{aligned} s.t. \quad Y_{jt+i} &= \left(P_{jt+i}/P_{t+i}\right)^{-\theta} Y_{t+i}, \\ P_{jt+i+1} &= \Xi_{t+i,t+i+1} P_{jt+i}, \end{aligned}$$

where we substituted firm technology (2) and firm productivity $Z_{jt} = G_{jt} Q_{t-s_{jt}}$ into the firm's period profits. In equation (40), α denotes the Calvo probability that the firm has to keep its previous price ($0 \leq \alpha < 1$), τ_{t+i} denotes the sales subsidy and $\Omega_{t,t+i} = \beta \frac{\xi_{t+i}}{\xi_t} \frac{U_{Ct+i}}{U_{Ct}}$ denotes the household's discount factor between time t and $t+i$. The first constraint in the price-setting problem captures the firm's demand function, see equation (7), and the second constraint captures how the firm's price is indexed over time in periods in which prices are not reset optimally. The price-setting problem in equation (40) implies that the optimal product price is given by

$$P_{jt}^{\star} = \left(\frac{\theta}{\theta - 1} \frac{1}{1+\tau} \right) \frac{E_t \sum_{i=0}^{\infty} (\alpha(1-\delta))^i \Omega_{t,t+i} Y_{t+i} \left(\Xi_{t,t+i}/P_{t+i}\right)^{-\theta} \frac{W_{t+i}/P_{t+i}}{A_{t+i} Q_{t-s_{jt}} G_{jt+i}}}{E_t \sum_{i=0}^{\infty} (\alpha(1-\delta))^i \Omega_{t,t+i} Y_{t+i} \left(\frac{1+\tau_{t+i}}{1+\tau}\right) \left(\Xi_{t,t+i}/P_{t+i}\right)^{1-\theta}}.$$

Rewriting the previous equation using analogous steps as in Appendix A.2 (Price-Setting Problem of Firms) in Adam and Weber (2019) yields

$$\frac{P_{jt}^{\star}}{P_t} \left(\frac{Q_{t-s_{jt}} G_{jt}}{Q_t} \right) = \left(\frac{\theta}{\theta - 1} \frac{1}{1+\tau} \right) \frac{N_t}{D_t},$$

where numerator N_t and denominator D_t are given by

$$N_t = E_t \sum_{i=0}^{\infty} (\alpha(1-\delta))^i \Omega_{t,t+i} \frac{Y_{t+i}}{Y_t} \left(\frac{\Xi_{t,t+i} P_t}{P_{t+i}} \right)^{-\theta} \frac{W_{t+i}}{P_{t+i} A_{t+i} Q_{t+i}} \left(\frac{q_{t+i} \times \cdots \times q_{t+1}}{g_{t+i} \times \cdots \times g_{t+1}} \right)$$
$$D_t = E_t \sum_{i=0}^{\infty} (\alpha(1-\delta))^i \Omega_{t,t+i} \frac{Y_{t+i}}{Y_t} \left(\frac{\Xi_{t,t+i} P_t}{P_{t+i}} \right)^{1-\theta} \left(\frac{1+\tau_{t+i}}{1+\tau} \right).$$

Both variables can be expressed recursively according to

$$N_t = \frac{W_t}{P_t A_t Q_t} + \alpha(1-\delta) E_t \left[\Omega_{t,t+1} \frac{Y_{t+1}}{Y_t} \left(\frac{\Pi_{t+1}}{\Xi_{t,t+1}} \right)^{\theta} \left(\frac{q_{t+1}}{g_{t+1}} \right) N_{t+1} \right] \quad (41)$$
$$D_t = \frac{1+\tau_t}{1+\tau} + \alpha(1-\delta) E_t \left[\Omega_{t,t+1} \frac{Y_{t+1}}{Y_t} \left(\frac{\Pi_{t+1}}{\Xi_{t,t+1}} \right)^{\theta-1} D_{t+1} \right]. \quad (42)$$

A.2 Equilibrium Conditions

In the model of Adam and Weber (2019), we assume that utility is given by equation (12) following Christiano, Trabandt and Walentin (2010), there is no capital accumulation ($\phi = 1$) and fixed costs of production are equal to zero ($f = 0$). Under these assumptions, and when accounting for the time-varying sales subsidy, the equilibrium conditions in Appendix A.7 (Transformed Sticky Price Economy) in Adam and Weber (2019) are given by

$$1 = \left[\alpha\delta + (1-\alpha)(\Delta_t^e)^{1-\theta}\right](p_t^\star)^{1-\theta} + \alpha(1-\delta)\left(\frac{\Pi_t}{\Xi_{t-1,t}}\right)^{\theta-1} \tag{43}$$

$$\Delta_t = \left[\alpha\delta + (1-\alpha)(\Delta_t^e)^{1-\theta}\right](p_t^\star)^{-\theta} + \alpha(1-\delta)\left(\frac{q_t}{g_t}\right)\left(\frac{\Pi_t}{\Xi_{t-1,t}}\right)^{\theta}\Delta_{t-1} \tag{44}$$

$$p_t^\star = \left(\frac{\theta}{\theta-1}\frac{1}{1+\tau}\right)\frac{N_t}{D_t} \tag{45}$$

$$N_t = \frac{w_t}{\Delta_t^e} + \alpha(1-\delta)\beta E_t\left[\left(\frac{\xi_{t+1}}{\xi_t}\right)\left(\frac{\Pi_{t+1}}{\Xi_{t,t+1}}\right)^{\theta}\left(\frac{q_{t+1}}{g_{t+1}}\right)N_{t+1}\right] \tag{46}$$

$$D_t = \frac{1+\tau_t}{1+\tau} + \alpha(1-\delta)\beta E_t\left[\left(\frac{\xi_{t+1}}{\xi_t}\right)\left(\frac{\Pi_{t+1}}{\Xi_{t,t+1}}\right)^{\theta-1}D_{t+1}\right] \tag{47}$$

$$y_t = \frac{\Delta_t^e}{\Delta_t}L_t \tag{48}$$

$$\gamma_t^e = a_t q_t \Delta_{t-1}^e / \Delta_t^e \tag{49}$$

$$(\Delta_t^e)^{1-\theta} = \delta + (1-\delta)\left(\Delta_{t-1}^e q_t / g_t\right)^{1-\theta} \tag{50}$$

$$w_t = y_t L_t^{\psi} \tag{51}$$

$$1 = \beta E_t\left[\left(\frac{\xi_{t+1}}{\xi_t}\right)\left(\frac{\gamma_{t+1}^e y_{t+1}}{y_t}\right)^{-1}\left(\frac{1+i_t}{\Pi_{t+1}}\right)\right] \tag{52}$$

$$\Xi_{t-1,t} = (\Pi_{t-1}/\Pi^\star)^{\varphi} \tag{53}$$

Recall definitions $p_t^\star = P_{t,t}^\star/P_t$, $\Gamma_t^e = A_t Q_t/\Delta_t^e$, $w_t = W_t/(P_t\Gamma_t^e)$ and $y_t = Y_t/\Gamma_t^e$. The variable $\Xi_{t-1,t}$ captures either price indexation to lagged inflation, with $\varphi > 0$, or no price indexation, with $\varphi = 0$.

A.3 Efficient Level of Output, Hours and Real Rate

Equations (43) to (52) nest the efficient equilibrium in the special case with flexible prices, $\alpha = 0$, and a sales subsidy that fulfills equation (11). We denote variables in the efficient equilibrium with superscript e. Under these assumptions, equation (43) yields

$$p_t^e = 1/\Delta_t^e, \tag{54}$$

and equations (45), (46) and (47) imply $N_t^e = w_t^e/\Delta_t^e$, $D_t^e = 1$ and hence

$$w_t^e = 1, \tag{55}$$

using that $p_t^\star = p_t^e$. Equations (48) and (51) then imply that the efficient level of detrended output and the efficient level of hours worked are given by

$$y_t^e = L_t^e = 1. \tag{56}$$

The efficient (gross) real rate follows from equation (52) and is given by

$$\frac{1}{R_t^e} = \beta E_t\left(\frac{\xi_{t+1}}{\xi_t}\frac{1}{\gamma_{t+1}^e}\right). \tag{57}$$

A.4 Sticky Price Economy Expressed in Gaps

Substituting equation (54) into equation (44) and rearranging yields

$$\begin{aligned}\frac{\Delta_t}{\Delta_t^e} &= \left[\alpha\delta(\Delta_t^e)^{\theta-1} + (1-\alpha)\right]\left(\frac{p_t^\star}{p_t^e}\right)^{-\theta} \\ &\quad + \alpha(1-\delta)\left(\frac{q_t}{g_t}\right)\left(\frac{\Pi_t}{\Xi_{t-1,t}}\right)^{\theta}\left(\frac{\Delta_{t-1}}{\Delta_{t-1}^e}\right)\frac{\Delta_{t-1}^e}{\Delta_t^e}\ .\end{aligned} \tag{58}$$

Proposition 1 in Adam and Weber (2019) shows that

$$\frac{\Pi_t^\star}{\Xi_{t-1,t}^\star} = \left(\frac{1-\delta\left(\Delta_t^e\right)^{\theta-1}}{1-\delta}\right)^{\frac{1}{\theta-1}}\ . \tag{59}$$

We use this equation to rearrange the term in square brackets in equation (58), which yields

$$\frac{\Delta_t}{\Delta_t^e} = \left[1 - \alpha(1-\delta)\left(\frac{\Pi_t^\star}{\Xi_{t-1,t}^\star}\right)^{\theta-1}\right]\left(\frac{p_t^\star}{p_t^e}\right)^{-\theta}$$

$$+ \alpha(1-\delta)\left(\frac{q_t}{g_t}\right)\left(\frac{\Pi_t}{\Xi_{t-1,t}}\right)^{\theta}\left(\frac{\Delta_{t-1}}{\Delta_{t-1}^e}\right)\frac{\Delta_{t-1}^e}{\Delta_t^e}. \tag{60}$$

From equations (50) and (59), we also obtain that

$$\frac{\Delta_{t-1}^e}{\Delta_t^e} = \frac{g_t}{q_t}\left(\frac{1}{\Pi_t^\star/\Xi_{t-1,t}^\star}\right), \tag{61}$$

and we use this equation to rearrange equation (60), which yields

$$\frac{\Delta_t}{\Delta_t^e} = \left[1 - \alpha(1-\delta)\left(\frac{\Pi_t^\star}{\Xi_{t-1,t}^\star}\right)^{\theta-1}\right]\left(\frac{p_t^\star}{p_t^e}\right)^{-\theta}$$

$$+ \alpha(1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1}\left(\frac{\Pi_t/\Xi_{t-1,t}}{\Pi_t^\star/\Xi_{t-1,t}^\star}\right)^{\theta}\left(\frac{\Delta_{t-1}}{\Delta_{t-1}^e}\right) . \tag{62}$$

Furthermore, we use equation (59) to rearrange the term in square brackets in equation (43), and use equation (54) to substitute p_t^e into equation (43). This yields

$$\left(\frac{p_t^\star}{p_t^e}\right)^{1-\theta} = \frac{1 - \alpha(1-\delta)(\Pi_t/\Xi_{t-1,t})^{\theta-1}}{1 - \alpha(1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1}} . \tag{63}$$

We also divide equation (45) by p_t^e, impose $1 = \frac{\theta}{\theta-1}\frac{1}{1+\tau}$, multiply equation (46) by Δ_t^e, and denote $\widetilde{N}_t = N_t\Delta_t^e$ to obtain

$$\frac{p_t^\star}{p_t^e} = \frac{\widetilde{N}_t}{D_t}, \tag{64}$$

$$\widetilde{N}_t = w_t + \alpha(1-\delta)\beta E_t\left[\left(\frac{\Pi_{t+1}^\star}{\Xi_{t,t+1}^\star}\right)^{\theta-1}\left(\frac{\Pi_{t+1}/\Xi_{t,t+1}}{\Pi_{t+1}^\star/\Xi_{t,t+1}^\star}\right)^{\theta}\left(\frac{\xi_{t+1}}{\xi_t}\right)\widetilde{N}_{t+1}\right], \tag{65}$$

$$D_t = \frac{1+\tau_t}{1+\tau} + \alpha(1-\delta)\beta E_t\left[\left(\frac{\Pi_{t+1}^\star}{\Xi_{t,t+1}^\star}\right)^{\theta-1}\left(\frac{\Pi_{t+1}/\Xi_{t,t+1}}{\Pi_{t+1}^\star/\Xi_{t,t+1}^\star}\right)^{\theta-1}\left(\frac{\xi_{t+1}}{\xi_t}\right)D_{t+1}\right], \tag{66}$$

where we also used equation (61) to rewrite $\widetilde{N}_t$ and rewrite the equation for D_t correspondingly.

B Approximate Equilibrium Conditions

We linearize the sticky-price equilibrium conditions at the efficient steady state, in which $\Pi = \Pi^\star = g/q$, $\Xi = \Xi^\star = 1$ and the sales subsidy fulfills equation (11).

B.1 Optimal Inflation Rate

Rearranging equation (59) yields $(1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1} = 1 - \delta\,(\Delta_t^e)^{\theta-1}$, and linearizing this equation in the variables $\Pi_t^\star/\Xi_{t-1,t}^\star$ and Δ_t^e yields

$$(1-\delta)(\Pi^\star/\Xi^\star)^{\theta-1}\widetilde{\pi}_t^\star = -\delta\,(\Delta^e)^{\theta-1}\,\widehat{\Delta}_t^e,$$

after defining

$$\widetilde{\pi}_t^\star = \ln(\Pi_t^\star/\Xi_{t-1,t}^\star) - \ln(\Pi^\star/\Xi^\star).$$

Using the steady-state equation $(1-(1-\delta)(\Pi^\star/\Xi^\star)^{\theta-1}) = \delta\,(\Delta^e)^{\theta-1}$ thus yields

$$\rho\widetilde{\pi}_t^\star = -(1-\rho)\widehat{\Delta}_t^e, \tag{67}$$

after using the definition of ρ in Proposition 1 given by

$$\rho \equiv (1-\delta)(\Pi^\star)^{\theta-1},$$

which also applies to the case with price indexation given that $\Xi^\star = 1$. Linearizing equation (61) yields

$$\widehat{\Delta}_{t-1}^e - \widehat{\Delta}_t^e = \widehat{g}_t - \widehat{q}_t - \widetilde{\pi}_t^\star, \tag{68}$$

and multiplying by $(1-\rho)$ and rearranging yields

$$-(1-\rho)\widehat{\Delta}_t^e = -(1-\rho)\widehat{\Delta}_{t-1}^e + (1-\rho)(\widehat{g}_t - \widehat{q}_t - \widetilde{\pi}_t^\star).$$

Substituting for $-(1-\rho)\widehat{\Delta}_t^e$ and $-(1-\rho)\widehat{\Delta}_{t-1}^e$ using equation (67) and simplifying terms yields

$$\widetilde{\pi}_t^\star = \rho\widetilde{\pi}_{t-1}^\star + (1-\rho)(\widehat{g}_t - \widehat{q}_t). \tag{69}$$

In the case without price indexation, we obtain $\widetilde{\pi}_t^\star = \pi_t^\star$ and thus

$$\pi_t^\star = \rho\pi_{t-1}^\star + (1-\rho)(\widehat{g}_t - \widehat{q}_t). \tag{70}$$

B.2 New Keynesian Phillips Curve

Taking logs of equation (64) and using that $\widetilde{N}^e = D^e$ in steady state yields

$$\widehat{p}_t^\star = \ln(\widetilde{N}_t/\widetilde{N}^e) - \ln(D_t/D^e), \tag{71}$$

denoting $\widehat{p}_t^\star = \ln(p_t^\star/p_t^e) - \ln(p^\star/p^e)$, which is equal to

$$\widehat{p}_t^\star = \ln p_t^\star - \ln p_t^e \tag{72}$$

since $p^\star/p^e = 1$. Linearizing equations (65) and (66) and subtracting the linearized equations from each other yields

$$\ln(\widetilde{N}_t/\widetilde{N}^e) - \ln(D_t/D^e) = (1-\alpha\rho\beta)\left(\ln w_t - \ln\frac{1+\tau_t}{1+\tau}\right) \\ + \alpha\rho\beta E_t[\ln(\widetilde{N}_{t+1}/\widetilde{N}^e) - \ln(D_{t+1}/D^e) + \widetilde{\pi}_{t+1} - \widetilde{\pi}_{t+1}^\star] \ ,$$

where we define

$$\widetilde{\pi}_t \equiv \ln(\Pi_t/\Xi_{t-1,t}) - \ln(\Pi^\star/\Xi^\star) \ . \tag{73}$$

Using equation (71) to substitute $\widehat{p}_t^\star$ for $\ln(\widetilde{N}_t/\widetilde{N}^e) - \ln(D_t/D^e)$ in the previous equation yields

$$\widehat{p}_t^\star = (1-\alpha\rho\beta)\left(\ln w_t - \ln\frac{1+\tau_t}{1+\tau}\right) + \alpha\rho\beta E_t[\widehat{p}_{t+1}^\star + \widetilde{\pi}_{t+1} - \widetilde{\pi}_{t+1}^\star] \ . \tag{74}$$

Linearizing equation (63) yields

$$\widehat{p}_t^\star = \frac{\alpha\rho}{1-\alpha\rho}(\widetilde{\pi}_t - \widetilde{\pi}_t^\star) \ . \tag{75}$$

Substituting this equation into equation (74) and simplifying terms yields

$$\widetilde{\pi}_t - \widetilde{\pi}_t^\star = \frac{(1-\alpha\rho)(1-\alpha\rho\beta)}{\alpha\rho}\left(\ln w_t - \ln\frac{1+\tau_t}{1+\tau}\right) + \beta E_t[\widetilde{\pi}_{t+1} - \widetilde{\pi}_{t+1}^\star] \ . \tag{76}$$

Taking logs of equation (48) and using that the log of the productivity-adjusted dispersion gap is a second-order term provided its initial value is of second order, see equation (117), yields

$$\ln y_t = \ln L_t. \tag{77}$$

Taking logs of equation (51) and substituting equation (77) yields

$$\ln w_t = (1+\psi)\ln y_t = (1+\psi)x_t, \tag{78}$$

where the second equality follows because

$$x_t = \ln y_t = \ln Y_t - \ln Y_t^e,$$

since $y_t = Y_t/\Gamma_t^e = Y_t/Y_t^e$ from equation (56). Substituting equation (78) into equation (76) yields the New Keynesian Phillips curve

$$\widetilde{\pi}_t - \widetilde{\pi}_t^\star = \kappa x_t + \beta E_t[\widetilde{\pi}_{t+1} - \widetilde{\pi}_{t+1}^\star] + u_t, \tag{79}$$

where the markup disturbance and the slope of this curve are defined in equations (19) and (20), respectively.

B.3 Price Indexation Equation

Multiplying the inverse of equation (53) by gross inflation, taking logs, subtracting $\ln(\Pi^\star/\Xi^\star)$ and using that $\Xi^\star = 1$ yields

$$\ln(\Pi_t/\Xi_{t-1,t}) - \ln(\Pi^\star/\Xi^\star) = \ln(\Pi_t/\Pi^\star) - \varphi\ln(\Pi_{t-1}/\Pi^\star). \tag{80}$$

Exploiting definitions (73) and (14) further yields

$$\widetilde{\pi}_t = \pi_t - \varphi\pi_{t-1}. \tag{81}$$

Analogously, we obtain

$$\widetilde{\pi}_t^\star = \pi_t^\star - \varphi\pi_{t-1}^\star. \tag{82}$$

B.4 Euler Equation and Natural Real Rate

Linearizing equation (52) yields

$$0 = E_t[\widehat{\xi}_{t+1} + \widehat{y}_t - \widehat{y}_{t+1} - \widehat{\gamma}_{t+1}^e + \widehat{\imath}_t - \pi_{t+1}],$$

where $\widehat{\xi}_{t+1}$ denotes the growth rate of the time preference shock in equation (12), π_{t+1} is defined in equation (14) and $\hat{\imath}_t$ is defined in equation (18) using that $1 + i = ag/\beta$. Using that $\widehat{y}_t = x_t$ since $\widehat{y}_t = \ln y_t = \ln Y_t - \ln \Gamma_t^e = \ln Y_t - \ln Y_t^e$ which is the output gap defined in equation (16), yields

$$x_t = E_t[x_{t+1} - (\hat{\imath}_t - \pi_{t+1}) + \widehat{\gamma}_{t+1}^e - \widehat{\xi}_{t+1}]. \tag{83}$$

Linearizing equation (57) yields

$$r_t^n = E_t[\widehat{\gamma}_{t+1}^e - \widehat{\xi}_{t+1}], \tag{84}$$

where r_t^n is defined in equation (17). Substituting the previous equation into equation (83) yields the Euler equation

$$x_t = E_t x_{t+1} - (\hat{\imath}_t - E_t \pi_{t+1} - r_t^n). \tag{85}$$

B.5 Efficient Output Growth and Detrended Output

Linearizing equation (49) yields

$$\widehat{\gamma}_t^e = \widehat{a}_t + \widehat{q}_t + \widehat{\Delta}_{t-1}^e - \widehat{\Delta}_t^e,$$

and substituting for $\widehat{\Delta}_{t-1}^e - \widehat{\Delta}_t^e$ using equation (68) and simplifying yields

$$\widehat{\gamma}_t^e = \widehat{a}_t + \widehat{g}_t - \widetilde{\pi}_t^\star. \tag{86}$$

To obtain a representation of $\widehat{\gamma}_t^e$ in terms of supply disturbances only, we rearrange the previous equation to obtain $\widetilde{\pi}_t^\star = -(\widehat{\gamma}_t^e - \widehat{a}_t - \widehat{g}_t)$, and substitute this expression into equation (69). After rearranging terms, this yields

$$\widehat{\gamma}_t^e = \rho \widehat{\gamma}_{t-1}^e + (\widehat{a}_t - \rho \widehat{a}_{t-1}) + \rho(\widehat{g}_t - \widehat{g}_{t-1}) + (1 - \rho)\widehat{q}_t. \tag{87}$$

To relate efficient output growth to detrended output, note that efficient output growth at the balanced growth path is given by $\gamma^e = \overline{\Gamma}_t^e / \overline{\Gamma}_{t-1}^e$, where the aggregate growth trend at this path evolves according to

$$\overline{\Gamma}_t^e = \bar{A}_t \bar{Q}_t / \Delta^e,$$

with $\bar{A}_t = a \cdot \bar{A}_{t-1}$, $\bar{Q}_t = q \cdot \bar{Q}_{t-1}$ and where $1/\Delta^e$ denotes the endogenous component of aggregate productivity in the efficient steady state. This implies

$$\widehat{\gamma}_t^e = \ln(\gamma_t^e / \gamma^e) = \ln(\Gamma_t^e / \overline{\Gamma}_t^e) - \ln(\Gamma_{t-1}^e / \overline{\Gamma}_{t-1}^e).$$

Augmenting the two RHS variables by output and rearranging yields

$$\ln(Y_t/\overline{\Gamma}^e_t) = \widehat{\gamma}^e_t + \ln(Y_{t-1}/\overline{\Gamma}^e_{t-1}) + \ln(Y_t/\Gamma^e_t) - \ln(Y_{t-1}/\Gamma^e_{t-1}).$$

Using the definition $\widetilde{y}_t = \ln(Y_t/\overline{\Gamma}^e_t)$ and that $Y^e_t = \Gamma^e_t$, which again follows from equation (56), we obtain for the previous equation that

$$\widetilde{y}_t = \widehat{\gamma}^e_t + \widetilde{y}_{t-1} + \ln(Y_t/Y^e_t) - \ln(Y_{t-1}/Y^e_{t-1}).$$

Using the definition of the output gap in equation (16) shows that log-linear detrended output evolves according to

$$\widetilde{y}_t = \widehat{\gamma}^e_t + \widetilde{y}_{t-1} + x_t - x_{t-1}. \tag{88}$$

We now consider the effect of a one-time experience growth shock, $\widehat{g}_0 > 0$, on $\widetilde{y}_t$ in a situation in which the output gap is closed, $x_t = 0$. With $x_t = 0$, the previous equation reduces to $\widetilde{y}_t = \widehat{\gamma}^e_t + \widetilde{y}_{t-1}$, and cumulating it over time yields

$$\widetilde{y}_t = \sum_{k=0}^{t} \widehat{\gamma}^e_k + \widetilde{y}_{-1}. \tag{89}$$

With only experience growth shocks, equation (87) reduces to

$$\widehat{\gamma}^e_t = \rho\widehat{\gamma}^e_{t-1} + \rho(\widehat{g}_t - \widehat{g}_{t-1}).$$

With $\widehat{g}_0 > 0$ and initial values $\widetilde{y}_{-1} = \widehat{\gamma}^e_{-1} = \widehat{g}_{-1} = 0$, this equation implies

$$\widehat{\gamma}^e_t = \begin{cases} \rho\widehat{g}_0, & t = 0, \\ -(1-\rho)\rho^t\widehat{g}_0, & t \geq 1. \end{cases}$$

Substituting this path for $\widehat{\gamma}^e_k$ into equation (89) delivers the effect on detrended output according to

$$\begin{aligned} \widetilde{y}_t &= \widehat{\gamma}^e_0 + \sum_{k=1}^{t} \widehat{\gamma}^e_k = \left(\rho - (1-\rho)\sum_{k=1}^{t} \rho^k\right)\widehat{g}_0 \\ &= \rho^{t+1}\widehat{g}_0, \end{aligned}$$

as stated in the main text.

B.6 Firm Profits

Nominal firm profits are given by

$$\Theta_t = (1+\tau_t)P_t Y_t - W_t L_t,$$

which implies that real firm profits Θ_t/P_t feature the same growth trend Γ_t^e as output Y_t and the real wage W_t/P_t. Defining detrended real profits as

$$\vartheta_t \equiv \frac{\Theta_t/P_t}{\Gamma_t^e},$$

the previous equation yields

$$\vartheta_t = (1+\tau_t)y_t - w_t L_t, \tag{90}$$

where as before, y_t and w_t denote detrended output and the detrended real wage, respectively. Detrended real profits in the efficient equilibrium, in which $\tau_t = \tau$, are given by

$$\vartheta_t^e = (1+\tau)y_t^e - w_t^e L_t^e.$$

Using equations (55) and (56) and the condition $1+\tau = \frac{\theta}{\theta-1}$, it follows that efficient detrended real profits are constant over time and equal to

$$\vartheta^e = \frac{1}{\theta-1}.$$

Linearizing equation (90) at the efficient steady state yields

$$\widehat{\vartheta}_t = (1+\tau)\frac{y^e}{\vartheta^e}\left(\widehat{y}_t - \frac{1+\psi}{\kappa}u_t\right) - \frac{w^e L^e}{\vartheta^e}(\widehat{w}_t + \widehat{L}_t), \tag{91}$$

after using the definition (19) and denoting a variable with a hat as the percentage deviation from its value in the efficient steady state. Using $1/\vartheta^e = \theta - 1$, $1+\tau = \frac{\theta}{\theta-1}$ and $y^e = w^e = L^e = 1$, equation (91) yields

$$\widehat{\vartheta}_t = \theta\left(\widehat{y}_t - \frac{1+\psi}{\kappa}u_t\right) - (\theta-1)(\widehat{w}_t + \widehat{L}_t). \tag{92}$$

Using $\widehat{y}_t = \widehat{L}_t$ from equation (77), $\widehat{w}_t = (1+\psi)\widehat{y}_t$ from equation (78) and $\widehat{y}_t = x_t$ implies that

$$\widehat{\vartheta}_t = [1-(\theta-1)(1+\psi)]x_t - \theta\left(\frac{1+\psi}{\kappa}\right)u_t, \tag{93}$$

with $\widehat{\vartheta}_t = \ln \vartheta_t - \ln \vartheta^e$. In situations without markup shocks $u_t = 0$ and a closed output gap $x_t = 0$, the previous equation implies $\widehat{\vartheta}_t = 0$, hence $\vartheta_t = 1/(\theta - 1)$ and thus

$$\Theta_t / P_t = \frac{1}{\theta - 1} \cdot \Gamma_t^e, \tag{94}$$

which shows that real profits after such shocks are proportional to the aggregate growth trend and, since $Y_t^e = \Gamma_t^e$, efficient output.

B.7 Approximate Equilibrium Conditions: Summary

The equilibrium in the approximate sticky-price economy is described by the equations (69), (79), (81), (82), (84), (85), (87), (88) and (93), plus a specification of monetary policy:

$$\begin{aligned}
\widetilde{\pi}_t - \widetilde{\pi}_t^\star &= \kappa x_t + \beta E_t[\widetilde{\pi}_{t+1} - \widetilde{\pi}_{t+1}^\star] + u_t \\
x_t &= E_t x_{t+1} - (\widehat{\imath}_t - E_t \pi_{t+1} - r_t^n) \\
\widetilde{\pi}_t^\star &= \rho \widetilde{\pi}_{t-1}^\star + (1 - \rho)(\widehat{g}_t - \widehat{q}_t) \\
r_t^n &= E_t[\widehat{\gamma}_{t+1}^e - \widehat{\xi}_{t+1}] \\
\widehat{\gamma}_t^e &= \rho \widehat{\gamma}_{t-1}^e + (\widehat{a}_t - \rho \widehat{a}_{t-1}) + \rho(\widehat{g}_t - \widehat{g}_{t-1}) + (1 - \rho)\widehat{q}_t \\
\widetilde{y}_t &= \widehat{\gamma}_t^e + \widetilde{y}_{t-1} + x_t - x_{t-1} \\
\widetilde{\pi}_t &= \pi_t - \varphi \pi_{t-1} \\
\widetilde{\pi}_t^\star &= \pi_t^\star - \varphi \pi_{t-1}^\star \\
\widehat{\vartheta}_t &= [1 - (\theta - 1)(1 + \psi)] x_t - \theta \left(\frac{1 + \psi}{\kappa} \right) u_t,
\end{aligned}$$

C Approximate Welfare-Based Loss Function

We denote expected discounted lifetime utility of the representative household in equation (12) according to

$$W = E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln(C_t) - \widetilde{V}(L_t) \right), \tag{95}$$

using

$$\widetilde{V}(L_t) = \frac{L_t^{1+\psi}}{1 + \psi}. \tag{96}$$

We approximate the expected discounted utility in equation (95) around the detrended efficient equilibrium (allowing for shocks and imposing flexible prices). Using the aggregate growth trend $\Gamma_t^e = A_t Q_t / \Delta_t^e$ and the detrended version of the aggregate technology in equation (48), we rearrange equation (95) according to

$$\begin{aligned} W &= E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln(c_t \Gamma_t^e) - \widetilde{V}(L_t) \right) \\ &= E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln(c_t) - \widetilde{V}(L_t) \right) + t.i.p. \\ &= E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln(y_t) - \widetilde{V}(y_t \frac{\Delta_t}{\Delta_t^e}) \right) + t.i.p., \end{aligned} \tag{97}$$

where $t.i.p.$ denotes terms that are independent of policy and the last equation uses market clearing $c_t = y_t$. We rearrange the definition of the output gap (16) according to

$$\begin{aligned} x_t &= \ln(Y_t/\Gamma_t^e) - \ln(Y_t^e/\Gamma_t^e) \\ &= \ln y_t - \ln y_t^e, \end{aligned} \tag{98}$$

and define

$$\widehat{\Delta}_t \equiv \ln \Delta_t - \ln \Delta_t^e. \tag{99}$$

Using equations (98) and (99), we further rearrange equation (97) according to

$$\begin{aligned} W &= E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln(y_t^e \cdot e^{x_t}) - \widetilde{V}(y_t^e \cdot e^{x_t + \widehat{\Delta}_t}) \right) + t.i.p. \\ &= E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \left(\ln(e^{x_t}) - \widetilde{V}(y_t^e \cdot e^{x_t + \widehat{\Delta}_t}) \right) + t.i.p. \\ &= E_0 \sum_{t=0}^{\infty} \beta^t \xi_t \underbrace{\left(x_t - \widetilde{V}(y_t^e \cdot e^{x_t + \widehat{\Delta}_t}) \right)}_{\equiv \widetilde{U}_t} + t.i.p. \end{aligned} \tag{100}$$

We approximate the relevant part of the period utility function, denoted by $\widetilde{U}_t$, to second order around the efficient equilibrium ($\ln \Pi_t = \ln \Pi_t^\star, \ln y_t = \ln y_t^e$) in which $x_t = \widehat{\Delta}_t = 0$ and the condition $\frac{\theta}{\theta-1}\frac{1}{1+\tau} = 1$ holds.

The first derivative required for this approximation is given by

$$\left.\frac{\partial \widetilde{U}_t}{\partial x_t}\right|_{x_t=\widehat{\Delta}_t=0} = 1 - \left.\frac{\partial \widetilde{V}}{\partial L_t} y_t^e \cdot e^{x_t+\widehat{\Delta}_t}\right|_{x_t=\widehat{\Delta}_t=0} = 1 - \left.\frac{\partial \widetilde{V}}{\partial L_t}\right|_{x_t=\widehat{\Delta}_t=0} y_t^e$$
$$= 0. \tag{101}$$

The last equality uses that

$$\left.\frac{\partial \widetilde{V}}{\partial L_t}\right|_{x_t=\widehat{\Delta}_t=0} y_t^e = 1, \tag{102}$$

which follows from evaluating $\partial\widetilde{V}/\partial L_t = L_t^{\psi}$ at $L_t^e = 1$ and using $y_t^e = 1$ from equation (56). The second derivative required for approximating equation (100) is given by

$$\begin{aligned}\left.\frac{\partial^2 \widetilde{U}_t}{(\partial x_t)^2}\right|_{x_t=\widehat{\Delta}_t=0} &= -\frac{\partial^2 \widetilde{V}}{(\partial L_t)^2}\left(y_t^e \cdot e^{x_t+\widehat{\Delta}_t}\right)^2 - \left.\frac{\partial \widetilde{V}}{\partial L_t} y_t^e \cdot e^{x_t+\widehat{\Delta}_t}\right|_{x_t=\widehat{\Delta}_t=0} \\ &= -\left(\left.\frac{\partial^2 \widetilde{V}}{(\partial L_t)^2}\right|_{x_t=\widehat{\Delta}_t=0} (y_t^e)^2 + \left.\frac{\partial \widetilde{V}}{\partial L_t}\right|_{x_t=\widehat{\Delta}_t=0} y_t^e\right) \\ &= -\left(\left.\frac{\partial^2 \widetilde{V}}{(\partial L_t)^2}\right|_{x_t=\widehat{\Delta}_t=0} + 1\right),\end{aligned} \tag{103}$$

where the last equality again using equation (102) and $y_t^e = 1$. The third derivative required for approximating equation (100) is given by

$$\left.\frac{\partial \widetilde{U}_t}{\partial \widehat{\Delta}_t}\right|_{x_t=\widehat{\Delta}_t=0} = -\left.\frac{\partial \widetilde{V}}{\partial L_t} y_t^e \cdot e^{x_t+\widehat{\Delta}_t}\right|_{x_t=\widehat{\Delta}_t=0} = -\left.\frac{\partial \widetilde{V}}{\partial L_t}\right|_{x_t=\widehat{\Delta}_t=0} y_t^e$$
$$= -1, \tag{104}$$

where the last equality again uses equation (102).

Derivatives (101), (103) and (104) suffice to obtain a second-order expansion of $\widetilde{U}_t$ in equation (100) that involves all terms affected by government policy because as we show in the next section C.1, the productivity-adjusted price-dispersion gap is of second order, $\widehat{\Delta}_t \sim O(2)$, whenever its initial value is of second order, $\widehat{\Delta}_{-1} \sim O(2)$. By construction, $\widehat{\Delta}_t$ is at its maximum value at the point of approximation and hence all its first-order derivatives are zero at this point. Thus, $\widehat{\Delta}_t \sim O(2)$ necessarily holds. This result, which is well-known to arise in the standard New Keynesian model approximated around the efficient equilibrium, also emerges in the framework with heterogenous firms considered here.[27] Therefore, we neither need to consider the second derivative of $\widetilde{U}_t$ with respect to $\widehat{\Delta}_t$ nor cross derivatives of $\widetilde{U}_t$ involving $\widehat{\Delta}_t$. We now show $\widehat{\Delta}_t \sim O(2)$ formally.

C.1 Dispersion Gap is of Second Order

Substituting equation (63) into equation (62) yields

$$\frac{\Delta_t}{\Delta_t^e} = \left(1-\alpha(1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1}\right)\left(\frac{1-\alpha(1-\delta)(\Pi_t/\Xi_{t-1,t})^{\theta-1}}{1-\alpha(1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1}}\right)^{\frac{-\theta}{1-\theta}}$$
$$+\alpha(1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1}\left(\frac{\Pi_t/\Xi_{t-1,t}}{\Pi_t^\star/\Xi_{t-1,t}^\star}\right)^{\theta}\left(\frac{\Delta_{t-1}}{\Delta_{t-1}^e}\right).$$

Summarizing terms in the previous equation and using the definition

$$\rho_t \equiv (1-\delta)(\Pi_t^\star/\Xi_{t-1,t}^\star)^{\theta-1} \tag{105}$$

to rearrange it yields

$$\frac{\Delta_t}{\Delta_t^e} = (1-\alpha\rho_t)^{\frac{1}{1-\theta}}\left(1-\alpha\rho_t\left(\frac{\Pi_t/\Xi_{t-1,t}}{\Pi_t^\star/\Xi_{t-1,t}^\star}\right)^{\theta-1}\right)^{\frac{-\theta}{1-\theta}}+\alpha\rho_t\left(\frac{\Pi_t/\Xi_{t-1,t}}{\Pi_t^\star/\Xi_{t-1,t}^\star}\right)^{\theta}\left(\frac{\Delta_{t-1}}{\Delta_{t-1}^e}\right).$$

Defining

$$\widetilde{\Pi}_t \equiv \Pi_t/\Xi_{t-1,t} \tag{106}$$
$$\widetilde{\Pi}_t^\star \equiv \Pi_t^\star/\Xi_{t-1,t}^\star, \tag{107}$$

we obtain for the previous equation that

$$\frac{\Delta_t}{\Delta_t^e} = (1-\alpha\rho_t)^{\frac{1}{1-\theta}}\left(1-\alpha\rho_t\left(\frac{\widetilde{\Pi}_t}{\widetilde{\Pi}_t^\star}\right)^{\theta-1}\right)^{\frac{-\theta}{1-\theta}}+\alpha\rho_t\left(\frac{\widetilde{\Pi}_t}{\widetilde{\Pi}_t^\star}\right)^{\theta}\left(\frac{\Delta_{t-1}}{\Delta_{t-1}^e}\right).$$

[27] Derivations for the standard New Keynesian model are nested in the derivations here and are obtained when imposing the parameter values $\delta = 0$ and $\Pi_t^\star = 1$.

Taking logs of this equation, expressing inflation variables in logs and using the definition of $\widehat{\Delta}_t$ in equation (99) yields

$$\widehat{\Delta}_t = \ln\left((1-\alpha\rho_t)^{\frac{1}{1-\theta}}\left(1-\alpha\rho_t e^{(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^\star)}\right)^{\frac{-\theta}{1-\theta}} + \alpha\rho_t e^{\theta\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^\star)} e^{\widehat{\Delta}_{t-1}}\right). \tag{108}$$

Dividing definition (106) by definition (107) and taking logs, we further obtain

$$\ln\widetilde{\Pi}_t = \ln\widetilde{\Pi}_t^\star + \ln(\Pi_t/\Pi_t^\star) - \varphi\ln(\Pi_{t-1}/\Pi_{t-1}^\star) \tag{109}$$

which is linear in logs and hence implies no cross terms between $\ln\Pi_t$ and $\ln\Pi_{t-1}$ in the derivations below.

We now approximate the RHS of the equation (108) in $\ln\Pi_t$, $\ln\Pi_{t-1}$ and $\widehat{\Delta}_{t-1}$ around the point $\ln\Pi_t = \ln\Pi_t^\star$, $\ln\Pi_{t-1} = \ln\Pi_{t-1}^\star$ and $\widehat{\Delta}_{t-1} = 0$ accurate to first order and using equation (109). The first derivative required for this approximation is

$$\frac{\partial\widehat{\Delta}_t}{\partial\ln\widetilde{\Pi}_t}\cdot\frac{\partial\ln\widetilde{\Pi}_t}{\partial\ln\Pi_t} = 0, \tag{110}$$

where this derivative is equal to zero at the point of approximation because we have that

$$\begin{aligned}\frac{\partial\widehat{\Delta}_t}{\partial\ln\widetilde{\Pi}_t} &= e^{-\widehat{\Delta}_t}\cdot\Bigg((1-\alpha\rho_t)^{\frac{1}{1-\theta}}\ \theta\left(1-\alpha\rho_t e^{(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^\star)}\right)^{\frac{\theta}{\theta-1}-1}\\ &\qquad (-\alpha)\rho_t e^{(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^\star)} + \theta\alpha\rho_t e^{\theta\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^\star)} e^{\widehat{\Delta}_{t-1}}\Bigg)\\ &= 0,\end{aligned} \tag{111}$$

where the second equality follows from evaluating the derivative at the point of approximation, and equation (109) implies that

$$\frac{\partial\ln\widetilde{\Pi}_t}{\partial\ln\Pi_t} = 1. \tag{112}$$

The second derivative required for the approximation is given by

$$\frac{\partial\widehat{\Delta}_t}{\partial\ln\widetilde{\Pi}_t}\cdot\frac{\partial\ln\widetilde{\Pi}_t}{\partial\ln\Pi_{t-1}} = 0, \tag{113}$$

where this derivative again is equal to zero at the point of approximation because derivative (111) is equal to zero at this point and equation (109) implies that

$$\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_{t-1}} = -\varphi. \tag{114}$$

The third derivative required for the approximation is given by

$$\begin{aligned}\frac{\partial \widehat{\Delta}_t}{\partial \widehat{\Delta}_{t-1}} &= e^{-\widehat{\Delta}_t} \alpha \rho_t e^{\theta \ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^\star)} e^{\widehat{\Delta}_{t-1}} \\ &= \alpha \rho_t,\end{aligned}$$

where the last equality follows from evaluating this derivative at the point of approximation. We thus obtain that

$$\widehat{\Delta}_t = \alpha \rho_t \widehat{\Delta}_{t-1} + O(2). \tag{115}$$

Using the definition in Proposition 1,

$$\rho \equiv (1-\delta)(\Pi^\star)^{\theta-1},$$

we obtain $\rho_t = \rho + O(1)$ and thus can write

$$\widehat{\Delta}_t = \alpha \rho \widehat{\Delta}_{t-1} + O(2), \tag{116}$$

because the first-order terms only enter as product with other first-order terms and thus are absorbed by the residual $O(2)$. Therefore, if the initial value of the productivity-adjusted price-dispersion gap satisfies

$$\widehat{\Delta}_{-1} \sim O(2), \tag{117}$$

as we shall assume, it will be the case that $\widehat{\Delta}_t \sim O(2)$ for all $t \geq 0$.

C.2 Welfare as Function of Output and Dispersion Gaps

Using derivatives (101), (103) and (104) and equations (116) and (117), the approximation to equation (100) is given by

$$W = E_0 \left[\sum_{t=0}^{\infty} \beta^t \xi_t \left(-\frac{1}{2} \left(\left. \frac{\partial^2 \widetilde{V}}{(\partial L_t)^2} \right|_{x_t = \widehat{\Delta}_t = 0} + 1 \right) x_t^2 - \widehat{\Delta}_t \right) \right] + t.i.p. + O(3) \ . \tag{118}$$

Since $\xi_t = 1 + O(1)$ and

$$\left.\frac{\partial^2 \widetilde{V}}{(\partial L_t)^2}\right|_{x_t=\widehat{\Delta}_t=0} = \psi\,(L_t^e)^{\psi-1} = \psi,$$

which follows from equation (96) and $L_t^e = 1$ in equation (56), we can rearrange equation (118) to obtain

$$W = -E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{1}{2}(1+\psi)x_t^2 + \widehat{\Delta}_t\right) + t.i.p. + O(3). \tag{119}$$

C.3 Dispersion Gap as Function of Inflation Gap

We now approximate equation (108) to express $\widehat{\Delta}_t$ to second-order accuracy as a function of current and lagged inflation rates, $\ln \Pi_t$ and $\ln \Pi_{t-1}$. We can rearrange the second derivative of $\widehat{\Delta}_t$ with respect to the current inflation rate according to

$$\frac{\partial}{\partial \ln \Pi_t}\left(\frac{\partial \widehat{\Delta}_t}{\partial \ln \Pi_t}\right) = \frac{\partial}{\partial \ln \widetilde{\Pi}_t}\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_t}\left(\frac{\partial \widehat{\Delta}_t}{\partial \ln \widetilde{\Pi}_t}\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_t}\right) = \frac{\partial^2 \widehat{\Delta}_t}{(\partial \ln \widetilde{\Pi}_t)^2},$$

where the second equality follows from using equation (112). Along similar lines, the second derivative of $\widehat{\Delta}_t$ with respect to the lagged inflation rate can be rearranged according to

$$\frac{\partial}{\partial \ln \Pi_{t-1}}\left(\frac{\partial \widehat{\Delta}_t}{\partial \ln \Pi_{t-1}}\right) = \frac{\partial}{\partial \ln \widetilde{\Pi}_t}\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_{t-1}}\left(\frac{\partial \widehat{\Delta}_t}{\partial \ln \widetilde{\Pi}_t}\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_{t-1}}\right) = \frac{\partial^2 \widehat{\Delta}_t}{(\partial \ln \widetilde{\Pi}_t)^2}\cdot(-\varphi)^2,$$

where the second equality follows from using equation (114). The cross derivative of $\widehat{\Delta}_t$ with respect to current and lagged inflation rates can be rearranged according to

$$\frac{\partial}{\partial \ln \Pi_{t-1}}\left(\frac{\partial \widehat{\Delta}_t}{\partial \ln \Pi_t}\right) = \frac{\partial}{\partial \ln \widetilde{\Pi}_t}\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_{t-1}}\left(\frac{\partial \widehat{\Delta}_t}{\partial \ln \widetilde{\Pi}_t}\frac{\partial \ln \widetilde{\Pi}_t}{\partial \ln \Pi_t}\right) = \frac{\partial^2 \widehat{\Delta}_t}{(\partial \ln \widetilde{\Pi}_t)^2}\cdot(-\varphi),$$

where the second equality follows from using equations (112) and (114). Taking the derivative of equation (111) with respect to the inflation variable yields

$$\frac{\partial^2 \widehat{\Delta}_t}{(\partial \ln \widetilde{\Pi}_t)^2} = e^{-\widehat{\Delta}_t}\Bigg((1-\alpha\rho_t)^{\frac{1}{1-\theta}} \theta \left(1-\alpha\rho_t e^{(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^{\star})}\right)^{\frac{\theta}{\theta-1}-2} (\alpha\rho_t)^2 e^{2(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^{\star})}$$
$$+ (1-\alpha\rho_t)^{\frac{1}{1-\theta}}\ \theta \left(1-\alpha\rho_t e^{(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^{\star})}\right)^{\frac{\theta}{\theta-1}-1} (\theta-1)(-\alpha)\rho_t e^{(\theta-1)\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^{\star})}$$
$$+ \theta^2 \alpha\rho_t e^{\theta\ln(\widetilde{\Pi}_t/\widetilde{\Pi}_t^{\star})} e^{\widehat{\Delta}_{t-1}} \Bigg),$$

and evaluating this derivative at $\ln \widetilde{\Pi}_t = \ln \widetilde{\Pi}_t^{\star}$ and $\widehat{\Delta}_t = \widehat{\Delta}_{t-1} = 0$ and simplifying terms yields

$$\begin{aligned}\frac{\partial^2 \widehat{\Delta}_t}{(\partial \ln \widetilde{\Pi}_t)^2} &= \theta \frac{(\alpha\rho_t)^2}{1-\alpha\rho_t} - \theta(\theta-1)\alpha\rho_t + \theta^2\alpha\rho_t \\ &= \theta\left(\frac{\alpha\rho_t}{1-\alpha\rho_t}\right).\end{aligned}$$

Using equation (115), the second-order approximation to equation (108) is then given by

$$\widehat{\Delta}_t = \alpha\rho_t \cdot \widehat{\Delta}_{t-1} \tag{120}$$
$$+ \frac{1}{2}\theta\left(\frac{\alpha\rho_t}{1-\alpha\rho_t}\right) \cdot \left(\left(\ln\frac{\Pi_t}{\Pi_t^{\star}}\right)^2 - 2\left(\ln\frac{\Pi_t}{\Pi_t^{\star}}\right)\left(\varphi \ln \frac{\Pi_{t-1}}{\Pi_{t-1}^{\star}}\right) + \left(\varphi\ln\frac{\Pi_{t-1}}{\Pi_{t-1}^{\star}}\right)^2\right) + O(3).$$

From definitions (14) and (15), we obtain that

$$\ln\frac{\Pi_t}{\Pi_t^{\star}} = \ln\frac{\Pi_t/\Pi^{\star}}{\Pi_t^{\star}/\Pi^{\star}} = \pi_t - \pi_t^{\star},$$

and substituting this into equation (120) and completing the squares yields that

$$\widehat{\Delta}_t = \alpha\rho_t \cdot \widehat{\Delta}_{t-1} + \frac{1}{2}\theta\left(\frac{\alpha\rho_t}{1-\alpha\rho_t}\right) \cdot \left(\pi_t - \pi_t^{\star} - \varphi(\pi_{t-1} - \pi_{t-1}^{\star})\right)^2 + O(3), \tag{121}$$

since all further terms are absorbed into the approximation residual given that $\widehat{\Delta}_t \sim O(2)$, as is shown in section C.1. Again using that $\rho_t = \rho + O(1)$ and hence

$$\frac{\alpha\rho_t}{1-\alpha\rho_t} = \frac{\alpha\rho}{1-\alpha\rho} + O(1),$$

we obtain for the second-order approximation that

$$\widehat{\Delta}_t = \alpha\rho \cdot \widehat{\Delta}_{t-1} + \frac{1}{2}\theta\left(\frac{\alpha\rho}{1-\alpha\rho}\right) \cdot \left(\pi_t - \pi_t^\star - \varphi(\pi_{t-1} - \pi_{t-1}^\star)\right)^2 + O(3), \tag{122}$$

because first-order terms only enter as products with second-order terms and thus are also absorbed into the residual $O(3)$. In the special case with $\delta = 0$, $\Pi^\star = 1$ and hence $\rho = 1$, equation (122) reduces to the equation that describes the evolution of price dispersion in the standard New Keynesian model with price indexation to lagged inflation. Summing equation (122) over time with discounting, and rearranging the result, yields

$$\sum_{t=0}^{\infty} \beta^t \widehat{\Delta}_t = \alpha\rho\beta \sum_{t=0}^{\infty} \beta^t \widehat{\Delta}_t + \frac{1}{2}\theta\left(\frac{\alpha\rho}{1-\alpha\rho}\right) \sum_{t=0}^{\infty} \beta^t \left(\pi_t - \pi_t^\star - \varphi(\pi_{t-1} - \pi_{t-1}^\star)\right)^2 + O(3) + t.i.p.$$

after absorbing the initial value $\widehat{\Delta}_{-1}$ into terms independent of policy. Rearranging the previous equation yields

$$\sum_{t=0}^{\infty} \beta^t \widehat{\Delta}_t = \frac{1}{2}\theta\left(\frac{\alpha\rho}{(1-\alpha\rho)(1-\alpha\rho\beta)}\right) \sum_{t=0}^{\infty} \beta^t \left(\pi_t - \pi_t^\star - \varphi(\pi_{t-1} - \pi_{t-1}^\star)\right)^2 + O(3) + t.i.p. \tag{123}$$

C.4 Welfare as Function of Output and Inflation Gaps

Substituting equation (123) into equation (119) yields a second-order approximation to welfare expressed in terms of the output gap and the quasi-differenced inflation gap,

$$W = -\frac{1}{2}E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{1+\psi}{\theta}\frac{(1-\alpha\rho)(1-\alpha\rho\beta)}{\alpha\rho} x_t^2 + \left(\pi_t - \pi_t^\star - \varphi(\pi_{t-1} - \pi_{t-1}^\star)\right)^2\right) + O(3) + t.i.p. \tag{124}$$

with $\rho = (1-\delta)(\Pi^\star)^{\theta-1}$. The weight on the quasi-differenced inflation gap is normalized to unity, and the weight on the output gap is equal to the slope of the Phillips curve divided by θ. The representation of approximate welfare in Proposition 1 corresponds to the case without price indexation with $\varphi = 0$.

D Proof of Lemma 2

Since the natural real rate does not move following a monetary policy shock and the real interest rate in equilibrium is given by

$$i_t - E_t \pi_{t+1} = -\rho^{t+1}(1-\rho)\widehat{g}_0,$$

the Euler equation (23) implies

$$x_t = E_t x_{t+1} + \rho^{t+1}(1-\rho)\widehat{g}_0.$$

Forward iteration yields

$$x_t = \rho^{t+1}\widehat{g}_0,$$

and using equation (34) to substitute for $\widehat{g}_0$ yields

$$x_t = -\frac{1}{1-\rho} r_t^n.$$

Since the optimal inflation rate does not respond following a monetary policy shock, the Phillips curve (22) implies

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t.$$

Forward iteration and using $x_t = \rho^{t+1}\widehat{g}_0$ yields

$$\pi_t = \frac{\kappa}{1-\beta\rho}\rho^{t+1}\widehat{g}_0,$$

an using equation (34) to substitute for $\widehat{g}_0$ yields

$$\pi_t = -\frac{\kappa}{(1-\rho)(1-\beta\rho)} r_t^n.$$

Finally, since the efficient level of output does not respond to the monetary policy shock, we have $x_t = \widetilde{y}_t$.

E Proof of Lemma 3

Assuming $\hat{\imath}_t = 0$ in equilibrium yields for the Euler equation (23) that

$$x_t = E_t x_{t+1} + E_t \pi_{t+1} + r_t^n. \tag{125}$$

We require that $x_t = E_t x_{t+1} + \rho^{t+1}(1-\rho)\widehat{g}_0$ to obtain the same solution as in lemma 2, which jointly with equation (125) yields

$$r_t^n = \rho^{t+1}(1-\rho)\widehat{g}_0 - E_t\pi_{t+1}. \tag{126}$$

The solution for inflation in lemma 2 implies that

$$\begin{aligned} E_t\pi_{t+1} &= \pi_{t+1} \\ &= \frac{\kappa}{1-\beta\rho}\rho^{t+2}\widehat{g}_0. \end{aligned}$$

Substituting this equation into equation (126) yields

$$r_t^n = \rho^{t+1}(1-\rho)\widehat{g}_0 - \frac{\kappa}{1-\beta\rho}\rho^{t+2}\widehat{g}_0, \tag{127}$$

as stated in lemma 3.

F Cohort-Level Expenditure Share

This appendix derives the evolution of cohort-level expenditure shares in response to the various (none-idiosyncratic) shocks in the model. Derivations in this section abstract from price indexation by imposing $\Xi_{t-k,t} = 1$ throughout.

Consider the expenditure share of a cohort with age s. Let $P^{\star}_{t-s,t-k}$ denote the optimal price of a firm that last experienced a δ-shock in $t-s$ and that has last reset its price in $t-k$, with $s \geq k \geq 0$. Let $m_t(s)$ denote the weighted average expenditure share in period t of the cohort of firms that last experienced a δ-shock in period $t-s$. For $s > k \geq 0$, this expenditure share is given by

$$m_t(s) = (1-\alpha)\sum_{k=0}^{s-1}\alpha^k(P^{\star}_{t-s,t-k}/P_t)^{1-\theta} + \alpha^s(P^{\star}_{t-s,t-s}/P_t)^{1-\theta}, \tag{128}$$

comprising α^s firms that have not had a chance to optimally reset prices since receiving the δ-shock and $(1-\alpha)\alpha^k$ firms that have last adjusted prices $k < s$ periods ago. For $s = 0$, the corresponding expenditure share is given by $m_t(0) = (P^{\star}_{t,t}/P_t)^{1-\theta}$. From equation (6), it follows that one can express the weighted sum of cohort expenditure shares according to

$$1 = \sum_{s=0}^{\infty}(1-\delta)^s\delta\, m_t(s), \tag{129}$$

where δ denotes the mass of firms that experience a δ-shock each period and $(1-\delta)^s$ denotes the share of those firms that have not undergone another δ-shock for s periods.

We rearrange equation (128) to obtain a tractable representation of the cohort expenditure share. Consider the optimal price $P^\star_{t-s,t}$ of a firm that sustained a δ-shock $s > 0$ periods ago, but can adjust the price in t due to the occurrence of a Calvo shock. Also, consider the price $P^\star_{t,t}$ of a firm where a δ-shock occurs in period t. The optimal price setting equation (10) then implies

$$P^\star_{t,t} = P^\star_{t-s,t}\left(\frac{g_t \times \cdots \times g_{t-s+1}}{q_t \times \cdots \times q_{t-s+1}}\right). \tag{130}$$

Using equation (130) to substitute prices $P^\star_{t-s,t-k}$ by prices $P^\star_{t-k,t-k}$ in equation (128) and using notation $p^\star_{t-k} = P^\star_{t-k,t-k}/P_{t-k}$ yields

$$\begin{aligned} m_t(s) &= (1-\alpha)\sum_{k=0}^{s-1}\alpha^k\left(p^\star_{t-k}\left(\Pi_t \times \cdots \Pi_{t-k+1}\right)^{-1}\left(\frac{g_{t-k}\times\cdots\times g_{t-s+1}}{q_{t-k}\times\cdots\times q_{t-s+1}}\right)^{-1}\right)^{1-\theta} \\ &+ \alpha^s\left(p^\star_{t-s}\left(\Pi_t\times\cdots\Pi_{t-s+1}\right)^{-1}\right)^{1-\theta}, \end{aligned} \tag{131}$$

where we obtain $m_t(0) = (p^\star_t)^{1-\theta}$ for age $s = 0$. We linearize the expenditure share at its efficient steady state in which $\Pi = \Pi^\star = g/q$. In this steady state, equation (131) implies that

$$m^e(s) = (p^e)^{1-\theta}(g/q)^{(\theta-1)\cdot s}. \tag{132}$$

Linearizing equation (131) at this steady state yields

$$\begin{aligned} \widehat{m}_t(s) = (1-\theta)\Bigg(&(1-\alpha)\sum_{k=0}^{s-1}\alpha^k\check{p}^\star_{t-k} + \alpha^s\check{p}^\star_{t-s} \\ &-\sum_{k=0}^{s-1}(1-\alpha^{k+1})\left(\widehat{g}_{t-k}-\widehat{q}_{t-k}\right) - \sum_{k=0}^{s-1}\alpha^{k+1}\pi_{t-k}\Bigg), \end{aligned} \tag{133}$$

where we denote the percentage deviation of the relative reset price from its value in the efficient steady state as $\check{p}^\star_t = \ln p^\star_t - \ln p^e$.[28] To obtain a solution

[28]This variable differs from the relative price gap used before, $\widehat{p}^\star_t = \ln p^\star_t - \ln p^e_t$, which features an additional time subscript.

for this variable, we rearrange equation (43) imposing that $\Xi_{t-1,t} = 1$ and exploiting equation (54) in steady state. This yields

$$1 = \left[\alpha(1-\rho) + (1-\alpha)(\Delta_t^e/\Delta^e)^{1-\theta}\right](p_t^\star/p^e)^{1-\theta} + \alpha\rho\left(\Pi_t/\Pi^\star\right)^{\theta-1}. \quad (134)$$

Linearizing this equation at the efficient steady state, at which $\Delta_t^e = \Delta^e$, $p_t^\star = p^e$ and $\Pi = \Pi^\star = g/q$, yields

$$(1-\alpha\rho)\check{p}_t^\star = \alpha\rho\pi_t + \frac{(1-\alpha)\rho}{1-\rho}\pi_t^\star. \quad (135)$$

For the impulse response functions (IRF), we use equations (133) and (135) jointly with the remaining approximate equilibrium conditions to compute $\widehat{m}_{t_0+s}(s)$ as a function of cohort age s assuming the shock hits at time t_0. The coefficients of this IRF are cohort-age dependent since we linearize at the efficient market share which evolves with cohort age s.