NBER WORKING PAPER SERIES

SUPPLY SHOCKS IN A HETEROGENEOUS-FIRM NEW KEYNESIAN MODEL: THE ENTRY MULTIPLIER

Florin O. Bilbiie
Marc J. Melitz

Working Paper 28258
http://www.nber.org/papers/w28258

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020, Revised July 2026

This is a major revision of and supersedes our previous paper titled "Aggregate-Demand Amplification of Supply Disruptions: The Entry-Exit Multiplier". We thank the Editor Nir Jaimovich and three anonymous referees, as well as Andrea Ajello, Susanto Basu, Cristiano Cantore, Vasco Carvalho, Larry Christiano, Giancarlo Corsetti, Joel David, Martin Eichenbaum, Andrea Ferrero, Francesco Furlanetto, Jordi Galí, Pierre-Olivier Gourinchas, Basile Grassi, Masahige Hamano, Antoine Lepetit, Virgiliu Midrigan, Sara Moreira, Anthony Savagar, Petr Sedlacek, Riccardo Silvestrini, Vincent Sterk, Paolo Surico, Mathias Thoenig, Michael Woodford, Francesco Zanetti, and participants in numerous seminars and conferences for useful comments and discussions. Bilbiie thanks the Janeway Institute and the Keynes Fund at Cambridge for financial support, as well as Bocconi University, IGIER, the Baffi Center and EIEF Rome for support and hospitality during part of writing this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Supply Shocks in a Heterogeneous-Firm New Keynesian Model: The Entry Multiplier
Florin O. Bilbiie and Marc J. Melitz
NBER Working Paper No. 28258
December 2020, Revised July 2026
JEL No. E31, E32, E40, E60

ABSTRACT

We study productivity shocks in a New Keynesian model with endogenous entry, selection, and nominal rigidities. Adjustment along the extensive margin fundamentally alters the transmission of TFP shocks. Under sticky prices, productivity disturbances generate a large “entry multiplier”: firm-entry-exit responds much more strongly than under flexible prices, even when output is allocatively efficient to first order. Introducing wage stickiness breaks this neutrality. Adverse TFP shocks reduce profits, trigger exit, and generate a negative output gap while remaining inflationary. Productivity shocks therefore behave as true supply shocks, without resorting to ad hoc cost-push or markup disturbances. Unlike standard markup shocks, TFP shocks in our framework imply procyclical profits and entry, consistent with the data. When wages are sufficiently sticky, expansionary monetary policy raises entry, closes the negative output gap, and improves welfare. The model remains analytically tractable and isomorphic to the standard New Keynesian framework, with entry and selection appearing as simple wedges.

Florin O. Bilbiie
University of Cambridge
florin.bilbiie@gmail.com

Marc J. Melitz
Harvard University
Department of Economics
and NBER
mmelitz@harvard.edu

An online appendix is available at http://www.nber.org/data-appendix/w28258

1 Introduction

How do productivity shocks affect inflation, economic slack, and the conduct of monetary policy? In the standard New Keynesian framework, transitory productivity shocks are not "true" supply shocks: Although *adverse* productivity shocks induce higher inflation, they also generate over-production in the sense of a *positive* output gap relative to the "flexible-price" model counterpart without nominal rigidities. Thus, those adverse productivity shocks call for a *contractionary* monetary policy response. In order to explain stagflationary episodes, the New Keynesian literature typically introduces cost-push or markup shocks, which shift marginal costs independently of productivity. These shocks are ad hoc, in the sense that they lack microfoundations. And while useful for matching inflation and output dynamics, they generate counterfactual implications for profits, entry, and firm dynamics.

This paper offers an alternative mechanism. We show that once we allow for firm entry, productivity shocks on their own (without invoking any additional markup shocks) take back their role as "true" supply shocks: adverse shocks are stagflationary; and call for an expansionary monetary policy response. Endogenous entry introduces a key new channel: an *entry multiplier* that amplifies the response of entry to productivity shocks under nominal rigidities. This entry response generates first-order effects for aggregate demand, inflation, and welfare.

The adverse productivity shock that we consider first is a classic negative total factor productivity (TFP) shock: a downward shift in the production function, such as the one associated with severe restrictions on the availability of inputs. This can be regarded as a metaphor for the aftermath of the COVID recession, supply chain and wartime disruptions, the Ukraine war.[1]We first isolate the entry multiplier in its most distilled form. To this end, we derive a model with endogenous entry but flexible wages, and show that TFP shocks amplify the response of firm entry under sticky prices – relative to flexible prices. Nominal rigidities induce changes in profits that trigger entry dynamics, setting off a feedback loop to aggregate productivity – which is endogenous due to the impact of input product variety on TFP. The intuition for a negative shock is as follows: Firms wish to increase their price to reflect their increased marginal cost. With sticky prices they cannot, so they are "stuck" with their suboptimal price. This induces further losses and triggers further exit, engendering an additional (endogenous) aggregate productivity decrease that amplifies the initial impulse.[2]

[1]The recent COVID-19 crisis highlighted the importance of the extensive margin: a very sharp response in entry and exit of businesses and product varieties; an older working paper version of this paper reviews empirical evidence for these entry dynamics in the COVID crisis.

[2]This mechanism captures an intuition that is more general than the inability to reset prices. It applies more gener-

Yet in this simplest modeling form, output coincides with its flexible-price counterpart to first order: despite large movements in entry, there is no output gap. This "envelope" or neutrality result—which we derive analytically—provides a clean benchmark: nominal rigidities magnify entry responses yet leave allocative efficiency unaffected absent wage stickiness. Introducing sticky wages fundamentally changes this conclusion. We show that adverse TFP shocks then generate a negative output gap; that is, a recession relative to the flexible-price-and-wage allocation. Lower productivity reduces profits and expected profitability. Firms cannot adjust their prices, amplifying losses and exit leading to a persistent contraction in the mass of producers. Sticky wages prevent hours from adjusting efficiently, so actual output falls by more than potential. In this environment, TFP shocks behave as true supply shocks: they are inflationary and simultaneously create slack.

This mechanism overturns a counterfactual— yet central—prediction of the standard New Keynesian model. There, negative TFP shocks raise inflation but imply a positive output gap, observationally indistinguishable from demand expansions. In our framework, the comovement is reversed. Inflation and the output gap move in opposite directions, i.e. there is more slack as in stagflationary episodes, but without relying on exogenous markup shocks. Instead, entry and exit act as an endogenous cost-push wedge in the Phillips curve, which we characterize analytically. Importantly, unlike standard cost-push shocks, TFP shocks in our model generate procyclical profits and entry, consistent with the data.

Our next main result concerns monetary policy. When wages are sufficiently sticky (relative to prices), a monetary expansion raises profits, stimulates entry, and closes the negative output gap induced by the adverse TFP shocks. In this region, monetary policy can fully restore efficient output and improve welfare. This stands in sharp contrast to the standard New Keynesian prescription, where contractionary policy is required in response to either adverse TFP shocks or stagflationary markup shocks. In our framework with endogenous entry, expansionary policy is appropriate precisely because productivity disturbances reduce demand through endogenous capacity destruction. We characterize analytically the degree of wage stickiness required for entry to respond positively to monetary expansions. When this condition holds, monetary policy operates through the extensive margin: higher demand raises profitability, induces entry, and expands productive capacity. This channel is absent in representative-firm models and reverses the standard prediction that supply-driven inflation necessarily calls for tightening.

ally to profitability shocks induced by nominal rigidities. Thus, this is a reduced form for frictions that impinge upon intensive-margin adjustments, with negative consequences for profitability.

Finally, we show that introducing firm heterogeneity and endogenous selection affects magnitudes but not mechanisms. The model retains a very similar mathematical structure to the standard New Keynesian framework. The main difference is that entry and selection appear as wedges in the Phillips curve and resource constraints. This structure allows us to derive closed-form solutions for the entry multiplier, the output gap, and welfare effects; while preserving the familiar logic of New Keynesian policy analysis.

Taken together, our results reframe the role of productivity shocks in monetary economies. Endogenous entry transforms TFP shocks into genuine supply disturbances—inflationary, contractionary, and welfare-relevant—eliminating the need for ad hoc markup shocks and restoring coherence between inflation dynamics, slack, profits, and policy prescriptions.

Related literature

A large and growing literature emphasizes the role of endogenous entry and variety with *flexible prices* for business cycles, studying macro fluctuations and normative properties i.a. Devereux, Head, and Lapham (1996), Jaimovich (2007), Jaimovich and Floetotto (2008), Colciago and Etro (2010), and Bilbiie et al (2012, 2019). Important recent extensions study the role of firm heterogeneity and selection, see i.a. Clementi and Palazzo (2015), Lee and Mukoyama (2018), Hamano and Zanetti (2017), and Edmond, Midrigan, and Xu (2020). Our framework is also related to Ghironi and Melitz (2005), who study endogenous entry and heterogeneity in productivity in an open-economy setting with flexible prices.[3] Clementi and Pallazzo (2016) study the amplifying role of endogenous entry and exit in a model with heterogeneous firms in the Hopenhayn (1992) tradition (with perfect competition and decreasing returns), and with flexible prices.

Several papers have analyzed these models with nominal rigidities, focusing on the effects and design of monetary policy, i.a. Bilbiie et al (2007, 2014); Bergin and Corsetti (2008); Lewis and Poilly (2014); Bilbiie (2021), and Gutierrez et al (2021). More recently, several subsequent contributions extended this framework to study the interaction of nominal rigidities, endogenous entry, and selection with firm heterogeneity: see Cooke and Damianovic (2021), Colciago and Silvestrini (2022), and Hamano and Zanetti (2022); The last paper uses a *sticky-wage* model to study the interaction of reallocation (cleansing of unproductive firms) and variety-reducing contractionary monetary policy, leading to lower aggregate productivity. [4] We draw on these contributions but

[3]Other recent developments include Cacciatore and Fiori (2016), Dixon and Savagar (2020), and Michelacci, Paciello, and Pozzi (2019) on unemloyment, endogenous productivity, and demand, respectively. Other earlier pioneering contributions on RBC-like models with entry include Campbell (1998), Chatterjee and Cooper (2014), and Cook (2001).

[4]The "entry multiplier" is present and operates in most such models entry and sticky *prices*, but it has not been

introduce heterogeneity in *quality*, which allows us to model *price* stickiness in a tractable way. But we also show that our model is isomorphic to one with heterogeneity in productivity, so long as ensuing TFP shocks are aggregate and idiosyncratic to firms. We then describe how selection and nominal rigidities both shape the dynamic macroeconomic responses to TFP shocks. And we show that those forces (selection and nominal rigidities) interact with one-another only to the extent that the distribution of firm size deviates from the Pareto shape benchmark (we do not explicitly model those deviations).

The standard NK model's failure to produce demand-induced recessions in response to negative supply shocks is the starting point of another recent paper by Guerrieri et al (2020).[5] The authors analyze the impact of the exogenous elimination (exit) of a sector in a 2-sector model. When the goods produced by the two sectors are complements and the intertemporal elasticity of substitution is elastic, this supply disruption induces a demand-induced recession, which the authors call "Keynesian supply shocks". In contrast, we model economy-wide supply disruptions when product-level entry is endogenous and those goods are substitutes (for our mechanism, the intertemporal elasticity of substitution needs to be *lower* than the elasticity of substitution between goods). Thus, our focus is on a different channel that is complementary to the sector-level disruptions examined by Guerrieri et al (2020). Importantly, our mechanism generates a negative comovement of inflation and the output gap, whereas in Guerrieri et al "Keynesian" supply shocks are recessionary but also deflationary. (and thus have the flavor of adverse demand shocks). Our mechanism makes adverse TFP shocks stagflationary, and thus literally *supply* shocks.[6]

A large empirical literature—especially based on estimated New Keynesian DSGE models—uses *cost-push* (*markup*) shocks as the primary driver of inflation fluctuations that are orthogonal to real activity, precisely because they generate the desired *negative comovement* between inflation and the output gap (or a positive comovement with unemployment or other measures of slack).

identified, isolated, and analytically characterized before. This is our paper's first contribution. Our results extend to models of entry with sunk costs where the number of firms acts as a state variable but we focus on the "static" free-entry model of entry for analytical tractability, as in Jaimovich and Floetotto (2008) for flexible prices and Bilbiie (2021) for sticky prices. For heterogeneous-firm models, see also Hamano and Pappada (2023) for an open-economy application, and Hamano (2025) for a study of aggregate and idiosyncratic volatility.

[5]In isolation, the shortcomings of the standard NK model pertaining to some of the comovements we analyze have also been studied previously. In their seminal contribution, Christiano, Eichenbaum, and Evans (2005) already demonstrated how wage stickiness is necessary in a fixed-variety model in order to replicate the positive response of profits to monetary policy found in the data (see also Christiano, Eichenbaum, and Evans, 1997).

[6]Cesa-Bianchi and Ferrero (2021) quantify empirically the contribution of sectoral shocks to aggregate fluctuations. Other contributions emphasize related supply-side mechanisms, e.g. exit following the revenue fall due to restrictions on a subset of products with rigid operating costs (Auerbach, Gorodnichenko, and Murphy (2021)); inter-sectoral linkages and complementarities (Woodford (2020)); unemployment and endogenous growth (Fornaro and Wolf (2020)), and input-output networks (Baqaee and Farhi (2020)). We abstract from such features to focus on endogenous entry and its interaction with nominal rigidities.

In canonical estimated models such as Smets and Wouters (2007), Justiniano et al (2012), and Del Negro et al (2015), markup shocks enter the Phillips curve as exogenous disturbances that move inflation independently of marginal cost. They are typically identified as the dominant source of high-frequency inflation variation. In particular, they can raise inflation while output and the output gap decline, thus replicating stagflationary episodes. Quantitatively, these models often attribute a substantial fraction of short-run inflation volatility to such shocks, even though other disturbances (e.g. demand) may matter more at lower frequencies. More broadly, both DSGE-based decompositions and VAR approaches that isolate "cost-push" innovations (e.g. oil or input cost shocks) find a similar pattern: inflation increases while activity contracts, producing a negative comovement between inflation and slack consistent with the New Keynesian notion of a Phillips-curve shift (see Bilbiie and Känzig, 2023 for recent evidence).

This empirical strategy—treating markup shocks as reduced-form wedges in the Phillips curve—has therefore become the standard way to reconcile observed inflation dynamics with relatively flat Phillips curves, although it remains somewhat controversial precisely because it relies on large exogenous residuals to account for inflation movements. But note that if those supply shocks were instead generated by negative TFP shocks, then the standard NK model would predict a *positive* output gap. The literature has thus reconciled this by assuming that those supply shocks were not TFP but cost-push shocks. But then why are profits—and also entry—not increasing during those episodes? The uncontroversial fall in entry and profits during such recessionary episodes is indirect evidence supporting the mechanism of our model. Indeed, our model also reproduces the cyclicality of profits to *both* supply and demand shocks – and it overturns the monetary policy prescription in face of stagflationary episodes.[7]

2 The Entry Multiplier and Aggregate Demand

In this section, we outline the simplest version of our model of endogenous entry with nominal rigidities, focusing on sticky prices. Households maximize the expected present value of utility defined over a consumption good C_t and hours worked L_t, where total consumption is equal to

[7] See Christiano et al (2005) for the role of wage stickiness in reproducing procyclical profits after demand shocks in a DSGE model and Cantore et al (2011) and Bilbiie and Känzig (2023) for analytical results on wage stickiness and the cyclicality of real wages and profits; the latter paper also provides VAR evidence for the response of profits to identified monetary policy, TFP, and cost-push (energy price) shocks. A separate issue is the response of hours worked to supply shocks, see e.g. e.g. Gali (1999), Basu, Fernald and Kimball (2006), Christiano, Eichenbaum, and Vigfusson (2003), Chari, Kehoe, and McGrattan (2008). Cantore et al (2014) show how factor-augmenting shocks and capital-labor substitutability contributes to resolving that debate. A different literature (e.g. Kaplan and Zoch, 2020) studies the the labor *share response to demand* shocks under nominal rigidities. Our model with all three ingredients (endogenous entry, and sticky prices and wages) addresses *all* these comovements jointly.

the output of a final-good sector consisting of a CES aggregate of intermediates. We consider a standard CRRA utility:

$$U\left(C_t, L_t\right) = \frac{C_t^{1-\sigma^{-1}} - 1}{1-\sigma^{-1}} - \chi \frac{L_t^{1+\varphi}}{1+\varphi},$$

where χ is the weight on the disutility of labor. (Without loss of generality, we will choose χ so that steady-state labor is normalized to one in order to simplify the analytics.).

For now, we model nominal rigidities by assuming *fixed* prices for intermediates. In the full model, we allow for more general nominal rigidities, including both *sticky* prices and wages. At this stage, this simple modeling of *fixed* prices alone is sufficient to develop the main intuition for our aggregate-demand amplification channel.

Static Equilibrium with Endogenous Entry

At time t, the household consumes C_t, equal to final good production Y_t.[8] The latter is produced using a continuum of intermediate inputs with measure N_t:

$$Y_t = \left(\int_0^{N_t} y_t\left(\omega\right)^{\frac{\theta-1}{\theta}} d\omega\right)^{\frac{\theta}{\theta-1}}, \tag{1}$$

where $\theta > 1$ is the symmetric elasticity of substitution across intermediate goods. Let $p_t\left(\omega\right)$ denote the nominal price of good ω and $P_t = \left(\int_0^{N_t} p_t\left(\omega\right)^{1-\theta} d\omega\right)^{\frac{1}{1-\theta}}$ the price of the final good. The demand for each intermediate ω is then $y_t\left(\omega\right) = \left(p_t\left(\omega\right)/P_t\right)^{-\theta} Y_t$.[9] We assume

$$\theta > \sigma \geq 1. \tag{2}$$

The first inequality ensures that the intermediate varieties ω are Edgeworth-substitutes satisfying $\partial^2 U/\partial y(\omega_1)\partial y(\omega_2) < 0$.[10] The second inequality $\sigma \geq 1$ ensures that income effects on labor supply (modulated by the coefficient of risk aversion σ^{-1},the inverse of which also determines the elasticity of intertemporal substitution) are weaker than substitution effects, so that the equilibrium response of labor supply to real wage improvements is generically positive. In the $\sigma = 1$ limit, the two effects offset each other and hours worked are constant under flexible prices.

There is a continuum of monopolistically competitive firms, each producing a different inter-

[8]The two will be distinct after we introduce nominal rigidities.

[9]This specification follows Ethier (1982) and Romer (1987)'s extension of the Spence-Dixit-Stiglitz aggregator. Our results carry through, albeit with some differences in interpretation, to a setup where the CES aggregate is defined over individual varieties of consumption goods instead.

[10]Although $\theta > 1$ uniquely identifies the CES between varieties in consumption C_t, utility $U(C_t, L_t)$ also includes an additive leisure term. This explains the additional restriction on σ for substitutability.

mediate $\omega \in [0, N_t]$. Production requires only one factor, labor, whose productivity is scaled exogenously by a factor A_t. We will pay particular attention to the macroeconomic consequences of substantial negative TFP shocks to A_t (in our simple framework, this is identical to a downward shift in labor supply). Output supplied by firm ω is:

$$y_t(\omega) = A_t l_t(\omega),$$

where $l_t(\omega)$ is the firm's employment. The constant real marginal cost is then w_t/A_t.

We consider a symmetric equilibrium with N_t firms and drop the ω qualifier. The relative price of intermediates in units of the final good is a key object that captures the aggregate productivity benefit of input variety, also known as "increasing returns to specialization":

$$\rho_t \equiv \frac{p_t}{P_t} = N_t^{\frac{1}{\theta-1}}. \tag{3}$$

Variations in the number of goods determine endogenous aggregate productivity, an insight that is at the core of all the expanding-variety endogenous growth literature.

Let μ_t denote the firms' markup (potentially time-varying):

$$\mu_t \equiv \frac{\rho_t}{w_t/A_t}. \tag{4}$$

Firm profit (d for dividends) in period t is:

$$d_t = \rho_t y_t - w_t l_t = \left(1 - \frac{1}{\mu_t}\right) \rho_t y_t. \tag{5}$$

We assume free entry subject to an entry cost f_E in units of labor: $d_t = w_t f_E$.

The household's budget constraint is reflected in the aggregate accounting identity equating expenditures (consumption plus the fixed cost "investment" for all firms) with income (labor income and variable profits for all firms). That is:

$$Y_t + N_t w_t f_E = w_t L_t + N_t d_t \iff Y_t = w_t L_t, \tag{6}$$

using free entry. The associated scale of the firm yields the firm-level labor demand: $l_t = \frac{1}{\mu_t - 1} f_E$.We can then write the aggregate production function as:

$$Y_t = \rho_t N_t A_t l_t = N_t^{\frac{\theta}{\theta-1}} A_t \frac{1}{\mu_t - 1} f_E. \tag{7}$$

A key equation is *labor demand*, obtained by aggregating l_t across producers:

$$L_t = \frac{\mu_t}{\mu_t - 1} N_t f_E. \tag{8}$$

On the household side, the consumption-leisure choice yields *labor supply*:

$$\chi L_t^{\varphi} = (1 + s_w)\, w_t C_t^{-1/\sigma}, \tag{9}$$

with $C_t = Y_t$. s_w is the constant labor subsidy that offsets the steady state markup: $1 + s_w = \mu = \theta/(\theta - 1)$. For reasons that will become clear momentarily, this subsidy is immaterial for comparing the sticky and flexible-price equilibria under endogenous entry (without the subsidy, both equilibria would be inefficient but the comparison between them remains unaffected). We work with the subsidy from the outset since this is material in the no-entry NK model and will also become relevant in our more general model.

2.1 Aggregate Demand and Monetary Policy

An important distinction concerns input versus final-good prices and their corresponding inflation rates. We refer to the former as the producer price p_t and to the latter as the "consumer" price P_t which captures the welfare-based, idealized price index. Producer-price inflation $1 + \pi_t = p_t/p_{t-1}$ and consumer-price inflation $1 + \pi_t^c = P_t/P_{t-1}$ are related to the growth in the number of intermediates through (3):

$$\frac{1 + \pi_t}{1 + \pi_t^c} = \left(\frac{N_t}{N_{t-1}}\right)^{\frac{1}{\theta - 1}}. \tag{10}$$

This distinction is particularly important with nominal rigidities because these apply at the *individual* firm-level price p_t. The relevant inflation rate for aggregate demand is *consumer* inflation π_t^c, insofar as it determines the ex-ante real interest rate that governs inter-temporal substitution. Indeed, the solution to the household's inter-temporal problem is the standard Euler equation for consumption:[11]

$$(C_t)^{-1/\sigma} = \beta E_t \left(\frac{1 + i_t}{1 + \pi_{t+1}^c} (C_{t+1})^{-1/\sigma} \right), \tag{11}$$

where i_t is the nominal interest rate and $\beta \in (0, 1)$ is the discount rate. The model is closed by specifying the price-setting equation—delivering a Phillips curve for PPI inflation—and monetary policy. The latter is set according to a Taylor rule, responding to expected future producer price

[11] The full solution also implies a standard transversality condition.

inflation π_t with elasticity $\phi \geq 1$:[12]

$$1 + i_t = \beta^{-1} \left(1 + E_t \pi_{t+1}\right)^{\phi} \exp(-\varepsilon_t). \tag{12}$$

The variable ε_t has two possible interpretations: in one, it is an exogenous (demand) shock—a discretionary decision by the central bank unrelated to current economic conditions—for which we can then trace-out the general-equilibrium effects. In another interpretation, ε_t is the "intercept" of the Taylor rule, which can be actively managed by the central bank in order to cushion the effects of structural shocks. We shall consider both, distinguishing between the two when relevant.

Under a Taylor rule that fixes the ex-ante real rate $\phi = 1$, replacing it in the Euler equation (11) and using the definition of CPI inflation (10) and ρ (3) delivers:

$$\rho_t \left(C_t\right)^{-1/\sigma} = E_t \left(\rho_{t+1} \left(C_{t+1}\right)^{-1/\sigma}\right) \exp(-\varepsilon_t) \implies \rho_t \left(C_t\right)^{-1/\sigma} = \exp\left(-\epsilon_t\right), \tag{13}$$

where $\epsilon_t \equiv E_t \sum_{j=0}^{\infty} \varepsilon_{t+j}$ is the "long" interest rate – measured as the sum of the future changes in interest rates. This is a measure of the monetary stance. We leverage this policy rule for deriving some of our analytical results.

2.2 The Entry Multiplier: Closed-form Solution

In order to distill our core entry multiplier channel, we first consider an extreme form of sticky prices that are indefinitely fixed along with a restriction to the case of *log-consumption utility* ($\sigma = 1$). In the next section, we return to the more general form of preferences along with arbitrary price stickiness. Logarithmic utility in consumption implies that income and substitution effects for labor supplied cancel out. The equilibrium hours worked is then fixed at $L_t = 1$, reflecting the normalization that we used for the disutility of labor in the utility function (see 9), specifically (under $\sigma = 1$), $\chi = \theta / (\theta - 1)$.

We use hats ($\hat{}$) to denote percent deviations from steady state: $\hat{X}_t = \ln X_t - \ln X$. For any endogenous variable X_t, its steady state level (denoted without time subscript) X is determined assuming steady-state productivity $A = 1$ and no monetary shocks $\varepsilon = 0$. In the steady state, we assume that the markup is equal to its desired level determined by the elasticity of substitution between goods: $\mu = \theta / (\theta - 1)$, inflation is zero $\pi = 0$ and $1 + i = \beta^{-1}$.

[12] Results are almost identical when using a Taylor rule that responds to contemporaneous inflation π_t. The advantage of the rule with future inflation is that it nests a policy that fixes the ex ante real (PPI-deflated) rate when $\phi = 1$. Building on Bilbiie (2008), we use this in the derivation of our analytical results – including those with arbitrary price and wage stickiness.

Under *flexible prices* (denoted with a * superscript), that markup does not respond to shocks and remains constant over time: $\mu_t^* = \mu = \theta/(\theta - 1)$. The equilibrium is then fully determined by the fixed labor supply along with the aggregate production (7) and labor demand (8):

$$N_t^* = \frac{1}{\theta f_E},$$
$$Y_t^* = \frac{\theta - 1}{\theta} \left(\frac{1}{\theta f_E} \right)^{1/(\theta-1)} A_t.$$

In this simple version with representative firms and log-utility, entry does not respond to TFP A_t under flexible prices (only firm size changes): $d\hat{N}_t^*/d\hat{A}_t = 0$. In the next section, we will show that either selection or more general preferences ($\sigma > 1$) is sufficient to generate a response of N_t^* to A_t: both the extensive and intensive margins respond.

Under *fixed* prices, we assume that each firm's price is fixed at $p_t = p^\circ$. The markup μ_t is then endogenous and output is demand-determined. With our restriction to log-utility and the fixed-real-rate rule used above and using $C_t = Y_t$, aggregate demand (13) can be written

$$M_t \equiv P_t Y_t = \exp \epsilon_t, \text{ where } \epsilon_t \equiv E_t \sum_{j=0}^{\infty} \varepsilon_{t+j}. \tag{14}$$

This is the money demand "quantity equation": $M_t = P_t Y_t$, where M_t also has the interpretation of a money supply controlled by the central bank. Later, we solve the dynamic version of the model with a Phillips curve and Taylor rule that does not entirely neutralize PPI inflation. Writing the price index as a function of the fixed price $P_t = p^\circ N_t^{-\frac{1}{\theta-1}}$ and plugging into the money demand equation yields:$Y_t = \frac{M_t}{p^\circ} N_t^{\frac{1}{\theta-1}}$. This, along with fixed labor supply $L_t = 1$, aggregate production (7), and labor demand (8) fully determines the fixed-price equilibrium. The number of firms is then:

$$N_t = \frac{1}{f_E} \left(1 - \frac{1}{A_t} \frac{M_t}{p^\circ} \right).$$

Thus, with fixed prices, entry responds to TFP A_t. Substituting out M_t, that elasticity is $d\hat{N}_t/d\hat{A}_t = \theta - 1 > 0$. This leads to:[13]

Proposition 1. ***The Entry Multiplier.*** *The response of the number of firms N_t to the supply shock A_t is*

[13]We choose the entry cost f_E to equalize Y^* with $Y^{*\mathcal{N}}$ for $A = 1$: $f_E = (\theta - 1)^{\theta-1}/\theta^\theta$. And given $M = 1$ in steady state, we normalize the price level in the fixed-price models to equalize Y (under fixed prices and endogenous entry) with that same level of output under flexible prices ($Y^* = Y^{*\mathcal{N}}$): this requires $p^\circ = \theta/(\theta - 1)$, i.e. the same markup in the two equilibria. For the no-entry model under the same maintained optimal subsidy, we locally re-normalize the disutility parameter to $\chi^{\mathcal{N}} = 1$ so that baseline labor and output equal 1 across both models; this is a pure unit-scaling choice with no implication on any of our results.

magnified under fixed prices (relative to flexible prices):

$$\frac{d\hat{N}_t - d\hat{N}_t^*}{d\hat{A}_t} = \theta - 1 > 0.$$

This is a powerful result that operates in models with entry and nominal rigidities. In our simplest example with log utility in consumption and representative firms (no selection) – the entry multiplier takes a particularly stark form because entry does not respond to FTP with flexible prices. In our full model, we will show that this inequality still holds when entry also responds under flexible prices; and when the response with nominal rigidities is more nuanced because both prices and wages are sticky (instead of just fixed prices).

However, the intuition is simple and applies to the more general case: Sticky prices create a friction that impedes adjustment at the intensive margin and inefficiently shifts adjustment to the extensive margin. Following an exogenous productivity decline, firms would like to raise prices to reflect higher costs, but cannot. Relative to the efficient flexible-price outcome, this further reduces profits and thereby further depresses entry. Due to increasing returns to specialization, this feeds back into a further—now *endogenous*—fall in aggregate productivity. This entry-multiplier is symmetric: Following a productivity increase, the same intensive margin friction induces inefficiently high entry – beyond the positive entry response when prices are flexible. This distortion increases with with the demand elasticity θ, when a higher intensive-margin adjustment is efficient. We return to this comparative static for the impact on aggregate output, that we will study in the following section.

The discussion above suggests that our entry-multiplier mechanism is likely to be more relevant, and more realistic, for negative productivity shocks than for positive ones, given the relative timing of entry and pricing decisions. For a positive shock, an unattractive feature of the model is that entry occurs before individual firms can adjust prices. One way to address this is to introduce a sunk cost that is lower than the price adjustment cost, as in Bilbiie, Ghironi, and Melitz (2007).

For negative productivity shocks, however, the same criticism has less force. If firms are stuck with prices that are too low and production scales too large, a larger fraction of them fail. After a large negative shock, more firms would survive if the decline in demand could be redistributed through lower individual sales, that is, through the intensive margin. But if such adjustment is not possible, disproportionately more firms exit. Price stickiness is only one possible representation of this failure of intensive-margin adjustment. More generally, the relevant friction is that firms cannot contract individual sales or raise prices sufficiently quickly after a large adverse shock. We therefore interpret price stickiness here as a tractable metaphor for a broader inability to adjust scale when a large negative profit shock would otherwise induce exit. Moreover, the difficulty of raising prices to stabilize individual production is likely to apply at the product level as well as at the firm level, so the exit mechanism emphasized here can also capture multi-product firms dropping products.[14]

We have simplified our model as much as possible to focus on the impact of supply shocks. In so doing, we have adopted a formulation of sticky prices that is too simple to adequately analyze the impact of demand (e.g. monetary) shocks in the presence of endogenous entry. We describe the optimal monetary policy in this simplest version of our model in Appendix A.2. It features a "divine coincidence" result that is analogous to models with no entry but similar price rigidities and policy tools (Blanchard and Gali, 2007): the central bank can replicate the efficient flexible-price level of output while at the same time also stabilizing inflation. However, we postpone an analysis of monetary policy until our full model with both prices and wages is developed in Section 3.

2.3 The "Standard" New Keynesian Model with Exogenous Entry

It is useful for comparison to review the standard New Keynesian (NK) model with exogenous entry: a fixed number of varieties $N_t^{\mathcal{N}} = N^{\circ}$, which we normalize to one. We use this notation of a superscript $\mathcal{N}$ (No entry) for all endogenous variables in this NK framework. Labor supply is still given by (9). The markup is similarly defined as $\mu_t^{\mathcal{N}} = A_t/w_t^{\mathcal{N}}$. This also pins down labor demand, with no aggregate productivity benefit to variety. Furthermore, there is now no distinction between producer and consumer prices. Since we normalize the mass of goods to 1, individual and aggregate variables coincide. The production function is $Y_t^{\mathcal{N}} = A_t L_t^{\mathcal{N}}$, and firm profit is $d_t^{\mathcal{N}} = \left[1 - \left(1/\mu_t^{\mathcal{N}}\right)\right] Y_t^{\mathcal{N}}$. This case with no entry can also be rationalized within the context of our earlier model where entry is possible but unprofitable as a result of a high entry cost f_E such that $d_t^{\mathcal{N}} < w_t^{\mathcal{N}} f_E$ for any realization of the TFP or monetary shocks.

Under *flexible prices*, optimal pricing implies a constant markup rule $\mu_t^{*\mathcal{N}} = \frac{\theta}{\theta-1}$. This implies that the wage is $w_t^{*\mathcal{N}} = \frac{\theta-1}{\theta} A_t$, with hours and consumption given by $L_t^{*\mathcal{N}} = 1$ and $Y_t^{*\mathcal{N}} = A_t L_t^{*\mathcal{N}}$.[15] Thus, output is supply-determined and the demand side only determines nominal variables: the price level. Hours are constant in this equilibrium, even though labor supply is endogenous, because income and substitution effects cancel out (a consequence of logarithmic utility in consump-

[14] See Argente et al (2018) for recent evidence on the cyclical relevance of this margin.
[15] Under the same optimal subsidy $s_w$and weight $\chi^{\mathcal{N}} = 1$ previously used.

tion).

With *fixed prices* we now have $Y_t^{\mathcal{N}} = A_t L_t^{\mathcal{N}} = \frac{M_t}{P^\circ}$, or equivalently $L_t^{\mathcal{N}} = \frac{M_t}{P^\circ}\frac{1}{A_t}$. Hours go up when productivity goes down in order to keep output constant at the demand-determined level.[16]This assumes no rationing, which would occur for a large enough drop in TFPA_t, in which case output would be supply-determined. [17]

2.4 Aggregate Demand Implications

We now contrast the implications of nominal rigidities between this standard New-Keynesian model and our simplest model with endogenous entry featuring the entry multiplier. We focus on the implications for aggregate output Y_t as a function of the exogenous technology TFP shock A_t. This is illustrated in Figure 1, with fixed prices in blue and flexible prices in red). Panel (a) on the left depicts the New-Keynesian case with exogenous entry. Under fixed prices, output is entirely demand determined, captured by the horizontal line $Y_t^{\mathcal{N}} = M_t/P^\circ$. Under flexible prices, money is neutral and output is supply-determined: it is the upward sloping line with slope $L^\circ = 1$. Thus, the output response is always greater under flexible prices. In particular, there is excess demand in response to a negative supply shock.[18]

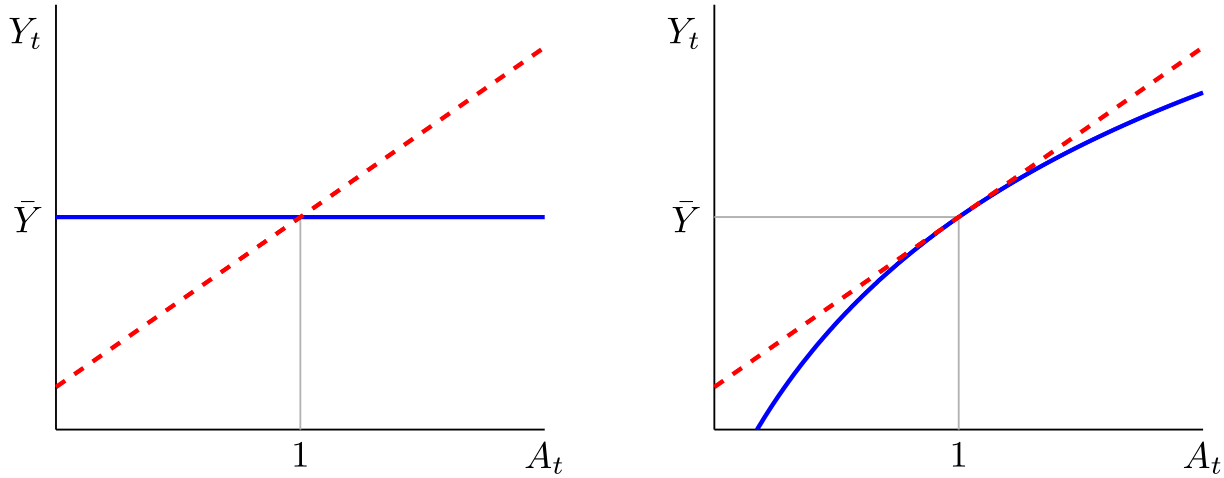
In our model with endogenous entry, aggregate output with nominal rigidities is no longer demand-determined. Even though individual prices p are fixed at p°, the welfare-relevant price index P is not fixed (because it also depends on entry). So monetary policy on its own no longer determines output. This is depicted in panel (b) on the right. In our simplest model, there are no first-order effects of nominal rigidities because—with endogenous entry—the constant labor subsidy is sufficient to eliminate the consumption-leisure distortion (just like with flexible prices). The only remaining distortion is to entry. But it is efficient in the steady state, leading to an envelope condition for output under fixed prices (a consequence of the neutrality proposition in Bilbiie (2021)).[19] This efficiency breaks when wage adjustment is no longer frictionless, as we highlight in the next section. There are then first order effects of nominal rigidities for output. And crucially, the predictions for the output-gap are reversed relative to the standard New-Keynesian model. In

[16]This effect on hours is due to income effects. Note that wages and profits are: $w_t^{\mathcal{N}} = A_t^{-\varphi}\frac{\theta-1}{\theta}\left(\frac{M_t}{P^\circ}\right)^{1+\varphi}$ and $D_t^{\mathcal{N}} = (M_t/P^\circ)\left(1 - A_t^{-(1+\varphi)}\frac{\theta-1}{\theta}(M_t/P^\circ)^{1+\varphi}\right)$. Wages are countercyclical and profits pro-cyclical conditional on supply shocks. In particular, wages go up and profits down in response to a bad shock. Agents work more because of the extra income effect of profits relative to the free-entry ($Y_t = w_t L_t$) case, whereby income and substitution effects cancel out. See Gali (1999) and a large literature thereafter discussing the evidence for this co-movement.

[17]This implies $\frac{M_t}{P^\circ}\frac{1}{A_t} < L_{\max}^\circ$ (the total time endowment), or equivalently (using $A^* = 1$), $A_t > \frac{M^\circ}{P^\circ}\frac{1}{L_{\max}^\circ} = \frac{1}{L_{\max}^\circ}$.

[18]This is a well-known property of the standard no-entry sticky-price model restated here as a benchmark.

[19]Note that a similar "envelope" result holds in the no-entry economy, but for *welfare* only: the allocations of consumption and hours are both distorted in the sticky equilibrium. In the entry model instead, they are both efficient.



(a) Exogenous Entry

(b) Endogenous Entry

Figure 1: The Output Gap: Fixed Price Y (solid blue) vs Flexible Price Y^* (red dash)

particular there is then a first-order demand shortage in response to a negative supply shock when entry is endogenous.

Howerver, even in this simplest case, output under sticky prices is always lower than under flexible prices – but only through second-order effects; this also generates welfare costs: as we show in Appendix A.2, despite the first-order neutrality there are second-order welfare costs of (squared) fluctuations in inflation and in the entry multiplier, i.e. the gap between the number of firms under sticky and flexible prices.[20]

3 The "Full" Heterogeneous-Firm NK Model with Endogenous Entry and Nominal Rigidities

We now develop our full model. On the "macro" side, we incorporate arbitrary levels of price and wage stickiness; and we allow for a higher intertemporal consumption elasticity $\sigma > 1$. And on the "micro" side, we introduce firm heterogeneity and selection. We show how such a model featuring the entry-multiplier described in the previous section replicates a series of business-cycle comovements. The addition of sticky wages generates first-order effects for the aggregate-demand amplification of supply shocks and positive co-movements between monetary policy and entry. Yet, we also show how sticky wages *on their own* cannot generate the key aggregate-demand am-

[20]It is worth emphasizing that in our economy with endogenous entry the "envelope" result holds for both output (consumption) and welfare: that is, the allocation of hours worked is efficient by virtue of the efficiency of entry combined with the $C_t = w_t L_t$ aggregate accounting: the wage is both the MRS and the MRT.

plification that arises with sticky prices. This is why we started with a simpler model with just that single nominal rigidity in order to highlight the key intuition underlying that mechanism. Nominal rigidities take the form of (Rotemberg) quadratic adjustment costs for both prices and wages, delivering standard Phillips curves for both prices and wages.[21] The full model derivations are relegated to Appendix A.[22]

Along with the more realistic modeling of price and wage rigidities, our full model also incorporates firm heterogeneity and selection. Although those forces play an important role in shaping the aggregate returns to scale of the economy, we show that they do not interact with the entry multiplier channel that we previously highlighted. Or put differently, selection is "orthogonal" to nominal rigidities, captured byt the difference between the flexible and sticky price equilibria. We incorporate firm heterogeneity in both productivity and quality.

3.1 Heterogeneous Firms

Firms have heterogeneous productivity (physical output per unit labor) and quality (quality-adjusted output per unit of physical output). We assume that those productivity and quality differences across firms do not change over time. All firms therefore face the same productivity shifter A_t over time that we modeled in the previous section with representative firms. Across firms, we normalize output units to equal output per unit labor. This normalization delivers the same reference price p_t for all firms. We then capture both productivity and quality differences terms across firms as an index $q > 0$, which we just call quality (though this could also reflect productivity differences). Our main restriction is that this index q does not vary over time: only aggregate productivity A_t fluctuates.[23]

Demand across varieties is still CES with elasticity θ, with quality-adjusted prices p_t/q and quality-adjusted output $qy_t(q)$ for a firm with quality q:

$$qy_t\left(q\right) = \left(\frac{p_t/q}{P_t}\right)^{-\theta} Y_t.$$

[21] We model wage stickiness in a standard way with a quadratic cost for nominal wages adjustments by a "labor union". The latter bundles the differentiated labor types of a unit mass of households, giving rise to a standard nonlinear " wage Phillips" curve (we follow the classic references on wage rigidity, Erceg et al, 2000; and Schmidt-Grohé and Uribe, 2006). In the limit as the adjustment cost increases, we recover the case of fixed nominal wages. We derive a full non-linear analytical solution for this case.

[22] In an online appendix A.1 we also derive a model extension with demand externalities that generate additional aggregate-demand amplification.

[23] Our framework also easily incorporates aggregate quality changes over time. This affects the inflation metrics – depending on whether the prices are adjusted for quality or not; but does not change any of the implications of our model in terms of the output gap and entry-multiplier responses to productivity shocks and the impact of monetary supply responses.

The price index P_t now aggregates over the quality-adjusted prices: $P_t^{1-\theta} = N_t \int_0^\infty (p_t/q)^{1-\theta} \, dG_S(q)$, where $G_S(q)$ is the distribution of quality for producing (surviving) firms. And similarly, output Y_t now aggregates over quality-adjusted units: $Y_t^{\frac{\theta-1}{\theta}} = N_t \int_0^\infty (q y_t(q))^{\frac{\theta-1}{\theta}} \, dG_S(q)$. Thus, the relative price $\rho_t \equiv p_t/P_t$ is no longer uniquely determined by the number of producers N_t due to endogenous selection. Given that common relative price across firms, aggregating sales across firms yields an alternative form for aggregate output $Y_t = \rho_t N_t \int_0^\infty y_t(q) \, dG_S(q)$.

We now add a per-period overhead production cost f_O in units of effective labor to the firm production function:[24]

$$y_t(q) = A_t l_t(q) - f_O.$$

This overhead cost is also reflected in the real profit per firm:

$$d_t(q) = \left(1 - \frac{1}{\mu_t}\right) \rho_t y_t(q) - \frac{w_t}{A_t} f_O. \tag{15}$$

However, note that operating profits $d_t(q) + w_t f_O / A_t$ remains unchanged. Given a constant markup μ_t across firms, those operating profits and firm sales $\rho_t y_t(q)$ are proportional to $q^{\theta-1}$. The modeling of quality for sales and profits (and all aggregate variables) is isomorphic to one with heterogeneous quality $A_t q$ and no quality differences. This clarifies how $A_t q$ captures heterogeneous quality with a common time-varying trend.

The overhead cost induces selection: Only firms with quality $q > q_{0t}$ will produce in period t, where the quality cutoff is determined by the zero profit cutoff (ZPC) condition $d_t(q_{0t}) = 0$. As we assumed in Section 2, firms must pay an entry cost f_E paid in labor units that does not change with aggregate productivity A_t. We assume that this cost is paid "ex ante" upon entry, before firms learn their idiosyncratic quality q, but after the aggregate TFP shock A_t is observed. Following Hopenhayn (1992), we model this "learning" as a draw from an exogenous distribution $G(q)$. A proportion $[1 - G(q_{0t})]$ of firms with high quality survive and select into production; and pay the overhead cost f_O "ex-post", yielding the distribution of quality among surviving firms $G_S(q) = G(q)/[1 - G(q_{0t})]$ with average profit:

$$\bar{d}_t \equiv \frac{1}{1 - G(q_{0t})} \int_{q_{0t}}^\infty d_t(q) dG(q).$$

There are N_{Et} entrants and $N_t = [1 - G(q_{0t})] N_{Et}$ producing firms such that the average ex-ante

[24]With CES demand, an overhead production cost is required to generate selection into production for the heterogeneous producers. We will also introduce an entry cost later in a similar way to the model in Section 2.

profit (prior to entry) is equal to the entry cost (free entry):

$$[1 - G(q_{0t})]\, \bar{d}_t = w_t f_E. \tag{16}$$

Note that we recover our representative firm model from the previous section by setting the overhead cost to $f_O = 0$. In this version, there is still firm heterogeneity, but that heterogeneity is immaterial as there is no selection when there is no overhead production cost: All entering firms with quality over the whole support of G produce in every period ($N_t = N_{Et}$).

We recover the same condition for labor demand, aggregating over firms:

$$L_t = N_{Et} \left\{ \left[\int_{q_{0t}}^{\infty} l_t(q) dG(q) \right] + f_E \right\}.$$

Without nominal rigidities, consumption C_t is still equal to output Y_t; so labor supply $\chi (L_t)^{\varphi} = [\theta/(\theta - 1)] (Y_t)^{-1/\sigma} w_t$ from (9) and aggregate accounting with free-entry $Y_t = w_t L_t$ from (6) remain unchanged. Aggregate production (7) changes to incorporate the overhead labor f_O:

$$Y_t = \rho_t A_t \left(L_t - N_{Et} f_E - N_t \frac{f_O}{A_t} \right),$$

where $L_t - N_{Et} f_E - N_t f_O / A_t$ represents the total hours devoted to production (not including overhead and entry).[25]

Aggregating real profit $d_t(q)$ from (15) yields an alternative representation for the average profit:

$$\bar{d}_t = \left(1 - \frac{1}{\mu_t} \right) \frac{Y_t}{N_t} - \frac{w_t}{A_t} f_O. \tag{17}$$

And using that same expression of real profit for the cutoff firm allows us to rewrite the cutoff profit condition $d_t(q_{0t}) = 0$ as: $q_{0t}^{\theta-1} = \frac{f_O}{\mu_t - 1} \frac{\rho_t^{\theta}}{Y_t}$.

Pareto quality distribution To obtain analytical expressions, we parameterize the distribution of quality draws to Pareto:

$$G(q) = 1 - \left(\frac{q_{\min}}{q} \right)^k,$$

which will also imply a Pareto distribution with shape $k/(\theta - 1)$ for firm size (sales). We assume that $k > (\theta - 1)$ in order to deliver a finite mean for firm size. $q_{\min} > 0$ is a lower bound for firm

[25] In our representative model, aggregate production was $Y_t = \rho_t A_t (L_t - N_t f_E)$, where $L_t - N_t f_E$ represented the total production hours. Alternatively, we can recover aggregate production function using firm production, aggregate sales, and labor demand.

quality.

Given a quality cutoff q_{0t}, a fraction $(q_{\min}/q_{0t})^k$ of entrants will produce:

$$N_t = \left(\frac{q_{\min}}{q_{0t}}\right)^k N_{Et}, \tag{18}$$

and the free entry condition (16) can then be written:

$$\left(\frac{q_{\min}}{q_{0t}}\right)^k \bar{d}_t = w_t f_E. \tag{19}$$

We can also use this to rewrite the zero cutoff profit condition as:

$$f_E = \left(\frac{q_{\min}}{q_{0t}}\right)^k \frac{\theta - 1}{k - (\theta - 1)} \frac{f_O}{A_t}. \tag{20}$$

We assume that the overhead cost f_O is high enough to generate selection with $q_{0t} > q_{\min}$. Combining this with our previous assumption on k yields a choice for that parameter that satisfies:

$$\theta - 1 < k < (\theta - 1)\left(1 + \frac{f_O}{A_t f_E}\right).$$

The parametrization of quality also yields an expression for the relative price from the CES price index:

$$\rho_t^{\theta-1} = N_{Et} \frac{k}{\theta - 1} \frac{A_t f_E}{f_O} q_{0t}^{\theta-1}. \tag{21}$$

Remark 2. **Selection-Stickiness Separation under Pareto**. A useful implication of the Pareto parametrization is that the cutoff q_{0t} is pinned down by exogenous TFP A_t through (20). In particular, the cutoff is independent of the markup μ_t and therefore of the price-setting block that closes the model. Selection is thus orthogonal to price stickiness in this parametrization.

This orthogonality is not a general property of models with selection: departures from Pareto can generate interactions between the firm-size distribution and the macroeconomic environment. Our point is instead that the entry-multiplier channel survives in a tractable environment with heterogeneous firms and endogenous selection. And as we show below, selection is itself a key determinant of how entry and the macroeconomic aggregates respond to TFP shocks.

Flexible price and wage benchmark

Under flexible prices and wages, firms choose their unconstrained optimal markup $\mu_t^* = \theta/(\theta-1)$. We solve our model using this condition, along with the new conditions for free-entry and zero-cutoff-profit (20) and (19), the number of producers (18), CES price index (21), and the average profit (17); along with the same labor supply (9) and aggregate accounting (6) conditions that carry over from our previous model. Note that this is essentially a static equilibrium where all endogenous variables are determined by the exogenous current period TFP A_t . And their solutions are all log-linear in that exogenous shock. We thus focus on those elasticities with respect to A_t:

$$
\begin{aligned}
\hat{q}_{0t}^* &= -\frac{1}{k}\hat{A}_t, \\
\hat{\rho}_t^* &= \left[\frac{1+\varphi}{\theta-1} - \frac{1}{k}\left(\varphi + \frac{1}{\sigma}\right)\right]\Phi\hat{A}_t, \\
\hat{w}_t^* &= \left(\varphi + \frac{1}{\sigma}\right)\left(\frac{\theta}{\theta-1} - \frac{1}{k}\right)\Phi\hat{A}_t, \\
\hat{L}_t^* &= \left(1 - \frac{1}{\sigma}\right)\left(\frac{\theta}{\theta-1} - \frac{1}{k}\right)\Phi\hat{A}_t, \\
\hat{Y}_t^* &= (1+\varphi)\left(\frac{\theta}{\theta-1} - \frac{1}{k}\right)\Phi\hat{A}_t, \\
\hat{N}_{Et}^* &= \left(1 - \frac{1}{\sigma}\right)\left(\frac{\theta}{\theta-1} - \frac{1}{k}\right)\Phi\hat{A}_t \\
\hat{N}_t^* &= \left[\left(1 - \frac{1}{\sigma}\right)\left(\frac{\theta}{\theta-1} - \frac{1}{k}\right)\Phi + 1\right]\hat{A}_t,
\end{aligned}
\tag{22}
$$

where $\Phi = \left(\varphi + \frac{\frac{\theta}{\sigma}-1}{\theta-1}\right)^{-1}$.

The flexible-price benchmark highlights how firm heterogeneity and selection change the representative firm environment of the previous section. First, aggregate output now displays increasing returns to aggregate productivity: $d\hat{Y}_t^*/d\hat{A}_t > 1$. There are two distinct sources for these increasing returns. The first is the standard increasing returns from variety and endogenous productivity, captured by $\frac{\theta}{\theta-1} > 1$. The second source is selection. A higher aggregate productivity shock lowers the cutoff q_{0t}, allowing lower-quality firms to produce while also changing the average quality of active producers. The strength of this selection effect depends on the Pareto shape parameter k. As k approaches its lower bound $\theta - 1$, the average size of active firms relative to the cutoff firm becomes arbitrarily large, so marginal changes in the cutoff have vanishing aggregate effects. At that lower bound, and with logarithmic utility in consumption, the increasing returns disappear: $d\hat{Y}_t^*/d\hat{A}_t = 1$.

Second, the model now has an active extensive margin of entry: $d\hat{N}_{Et}^*/d\hat{A}_t > 0$. This is the response of entrants, before selection into production. Its magnitude is governed by the same force that governs the response of hours worked $\hat{L}_t^*$. When $\sigma > 1$, the leisure-consumption substitution effect dominates the income effect of higher wage, and the households respond to higher TFP with more hours – just as it does in the no-entry model. This, in turn, raises the incentives for entry in our model. This force is already present in the homogeneous-firm entry model once we move away from logarithmic utility. When $\sigma = 1$, aggregate productivity has no effect on entry N_{Et}, even though it still affects the number of producing firms N_t through the cutoff q_{0t}. This invariance reflects the combination of an entry cost that is not indexed to aggregate productivity and the Pareto distribution of idiosyncratic quality.[26]

Finally, productivity also affects the intensive margin: $\left(d\hat{Y}_t^* - d\hat{N}_t^*\right)/d\hat{A}_t > 0$.Thus, in contrast to the stark homogeneous-firm, logarithmic-utility case used to isolate the entry multiplier in Section 2, the flexible-price response in the full environment operates through both entry and firm size. The analysis below therefore starts from a flexible-price benchmark in which both the intensive and extensive margins are already active, and asks how nominal rigidities distort both margins. We will then further discuss the contrast with the "standard" New-Keynesian model with no extensive margin (exogenous entry).

3.2 New Keynesian Block: Price and Wage Setting and Monetary Policy

We now extend the modeling of nominal rigidities. Price and wage adjustments are possible (prices are no longer fixed), but both are subject to Rotemberg-style adjustment costs. For price setting, we assume that all firms (including new entrants) pay an adjustment cost for resetting their price relative to last period's price p_{t-1}. We also follow the standard modeling of "wage stickiness" by assuming wage setting by a union that bundles the households' differentiated labor inputs. The union sets the nominal wage subject to similar adjustment costs. We detail the construction of the induced Phillips curves for price and wage inflation in Appendix A.3, along with all the equilibrium conditions for the full model (without any log-linearization). Similarly to producer-price inflation $1+\pi_t$, we define wage inflation relative to the rate of consumer-price inflation $1+\pi_t^c$: $\frac{1+\pi_{w,t}}{1+\pi_t^c} = \frac{w_t}{w_{t-1}}$. Monetary policy is set according to the same Taylor rule (12) as our representative firm model.

Log-linearizing around the same steady state with $A = 1$, $\varepsilon = 0$, $\pi = \pi_w = 0$, and $1 + i = \beta^{-1}$ delivers the linear equilibrium conditions outlined in Table 1. We use the same hat notation ($\hat{}$) to denote percent deviations from steady state: $\hat{X}_t = \ln X_t - \ln X$ for any endogenous variable, with

[26] This is the same force behind the invariance of entry to trade costs in the heterogeneous-firm environment analyzed by Arkolakis et al. (2012). Melitz and Redding (2015) show how departures from the Pareto parametrization break this invariance.

Table 1: Log-linearized Model, Summary

Static:	
Markup (4)	$\hat{\rho}_t = \hat{\mu}_t + \hat{w}_t - \hat{A}_t$
Aggregate accounting (6)	$\hat{Y}_t = \hat{w}_t + \hat{L}_t$
Average profit (17)	$\hat{\bar{d}}_t = (\theta - 1)\,\hat{\mu}_t + \hat{Y}_t - \hat{N}_t$
Number of surviving firms (18)	$\hat{N}_t = \hat{N}_{Et} - k\hat{q}_{0t}$
Free entry (19)	$\hat{\bar{d}}_t - k\hat{q}_{0t} = \hat{w}_t$
CES price index (21)	$(\theta - 1)\,\hat{\rho}_t = \hat{N}_{Et} + (\theta - 1)\,\hat{q}_{0t}$
Zero cutoff profit (20)	$k\hat{q}_{0t} + \hat{A}_t = 0$
Dynamic:	
Price Phillips-curve	$\pi_t = \beta E_t \pi_{t+1} - \psi\hat{\mu}_t$
Wage Phillips-curve	$\pi_{w,t} = \beta E_t \pi_{w,t+1} + \psi_w \left(\sigma^{-1}\hat{Y}_t + \varphi\hat{L}_t - \hat{w}_t\right)$
Euler bonds	$\hat{Y}_t = E_t\hat{Y}_{t+1} - \sigma\left(\hat{i}_t - E_t\pi^C_{t+1}\right)$
Taylor rule	$\hat{i}_t = \phi E_t \pi_{t+1} - \varepsilon_t$
Wage inflation	$\hat{w}_t = \hat{w}_{t-1} + \pi_{w,t} - \pi^C_t$
CPI inflation	$\pi_t = \pi^C_t + \hat{\rho}_t - \hat{\rho}_{t-1}$

Notes: The table displays all the log-linearized equilibrium conditions, around a zero-inflation steady state with an optimal subsidy.

the exception of rates: those are already in percentage points and are expressed as deviations from their steady-state value. The inflation rates are zero in steady state, so we drop the hats on the π variables. See appendix A.3 for the full derivation of the non-linear model with the functional forms for the cost of price adjustments and the optimal subsidy.

The first set of equations reproduce the set of equations that we used to solve for the flexible price without any log-linearization.[27] Although they are expressed in log-deviations, they do not involve any linearization except for the average profit. We label these equations "static" because they only involve current period variables.

The dynamic bloc includes the Phillips curves for prices and wages in log-linearized form: as anticipated, they are exactly identical to those of the standard NK model with no entry. Their slopes ψ and ψ_w respectively are inverse functions of the respective adjustment-cost parameter (a lower ψ means more stickiness, with full rigidity at $\psi = 0$ and full flexibility for $\psi \to \infty$).[28] Due to the nominal frictions, output is no longer equal to consumption ($Y_t = C_t$) in levels (see appendix A.3); but that relationship ($\hat{Y}_t = \hat{C}_t$) still holds in the log-linearized model. Thus, Table 1 just refers

[27] The labor supply equation does not show up here, it is incorporated into wage-Phillips curve in the dynamic bloc.

[28] Specifically, $\psi = (\theta - 1)/\kappa$ and $\psi_w = (1 + s_w)(\theta_w - 1)/\kappa_w$ where the κs are the adjustment cost parameters and s_w is the labor subsidy. The optimal subsidy offsets the markup distortions for both labor and intermediate goods. This is just a recalibration of the ψs; it will be relevant for our welfare analysis in section 4. In the log-linearized version, the ψs (slopes of the Philips curves) are a sufficient statistic for the impact of nominal rigidities – because the inflation costs are second order.

to the equilibrium output deviations $\hat{Y}_t$.

We now show how this log-linearized model has a reduced, 5-equation representation similar to the standard NK model (with both price and wage rigidities). However, endogenous entry now appears as a wedge in the model's key equations.

An NK, "Gap" Representation

Following the NK tradition, we rewrite our key model conditions in terms of "gaps" between the endogenous equilibrium variables under nominal rigidities and their flexible-price counterparts using tildes (˜): $\tilde{X}_t = \hat{X}_t - \hat{X}_t^*$, for any endogenous variable X. A key object, which acts as a composite reduced-form shock for the dynamics of the model, is r-star ($\hat{r}_t^*$): the "natural" real interest rate occurring under flexible prices and wages. This is obtained by replacing the equilibrium values under flexible prices (22) into the Euler equation for bonds, using the definition of CPI inflation in Table 1. In other words, this is the "real" with respect to PPI inflation π_t interest rate:

$$\hat{r}_t^* = (1 - \rho_a)\, \Gamma \hat{A}_t, \tag{23}$$

where

$$\Gamma = -\frac{(\sigma - 1)\varphi(\theta - 1) + k(\theta - \sigma)(1 + \varphi)}{k\left[(\theta - \sigma) + \sigma\varphi(\theta - 1)\right]} < 0$$

denotes the elasticity of r-star to a one-time TFP shock and ρ_a is the persistence of the TFP shock such that $E_t \hat{A}_{t+1} = \rho_a \hat{A}_t$.

This elasticity plays a very similar role in the standard NK model with exogenous entry. From here on out, we also extend this standard NK reference model from section 2 to incorporate the features that we have added to our endogenous entry model. Namely, we no longer restrict $\sigma = 1$ (the same restriction $\sigma \geq 1$ ensures that hours under flexible prices and wages respond positively to a TFP improvement), and we model both price and wage rigidity using the same adjustment cost process. The detailed derivation for this NK benchmark are relegated to Appendix B.

In this standard NK model, the elasticity of r-star to a one-time TFP shock is also negative: a fall in TFP raises the natural interest rate. For comparison, it is: $\Gamma^{\mathcal{N}} = -\left(1 + \varphi\right) / \left(1 + \sigma\varphi\right) < 0$. With and without endogenous entry, both elasticities are bounded between -1 and 0. Therefore, the dynamics of r-star are qualitatively similar in our model relative to the NK model; our mechanism does not rely on overturning the response of r-star. We thus obtain a system of equations in gaps that is strikingly similar to the standard NK model, with a few key differences:

$$\pi_t = \beta E_t \pi_{t+1} + \frac{\psi}{\theta}\tilde{Y}_t - \psi\frac{1}{\theta-1}\tilde{N}_{Et} \tag{24}$$

$$\pi_{w,t} = \beta E_t \pi_{w,t+1} + \frac{\psi_w}{\theta}\left(\frac{\theta}{\sigma} - 1 + (\theta-1)\varphi\right)\tilde{Y}_t \tag{25}$$

$$\hat{i}_t = \phi E_t \pi_{t+1} - \varepsilon_t \tag{26}$$

$$\tilde{Y}_t = E_t \tilde{Y}_{t+1} - \sigma\left(\hat{i}_t - E_t\pi_{t+1} - \hat{r}_t^*\right) - \sigma\frac{1}{\theta-1}\left(E_t\tilde{N}_{Et+1} - \tilde{N}_{Et}\right) \tag{27}$$

$$\frac{1}{\theta}\tilde{Y}_t = \frac{1}{\theta}\tilde{Y}_{t-1} + \pi_{w,t} - \pi_t - \left(\hat{A}_t - \hat{A}_{t-1}\right) + \frac{1}{\theta-1}\tilde{N}_{Et} - \frac{1}{\theta-1}\tilde{N}_{Et-1} \tag{28}$$

Like the standard NK model with both price and wage rigidity, this is a 5-equation model.[29] The main differences are that the price Phillips (24), IS (27), and wage growth (28) curves now have and additional additive term that incorporate the impact of endogenous entry. The TFP shock generates cyclical fluctuations in the familiar way: it impacts "r-star" $\hat{r}_t^*$, and also appears in the wage growth equation. However, its equilibrium effects change radically due to those additional endogenous entry terms. For instance, as in the standard NK model, r-star goes up with negative TFP. This normally leads to a positive output gap by inter-temporal substitution in (27); but—as we will show momentarily—the last term is endogenous due to entry and goes down, such that there is a recession (output falls below potential) instead. At the same time, on the supply side, the fall in number of firms acts as a shifter of the Phillips curve (24) at given consumption, so it is inflationary; this is akin to a fall in endogenous productivity. We first illustrate those differences quantitatively in a parameterized version of the model. We then derive the equilibrium in closed form in a special case and substantiate those insights beyond the specific chosen calibrated parameters.

3.3 Cyclical Implications: TFP Shocks Become True "Supply Shocks"

We quantitatively illustrate the response of all our endogenous variables to a negative TFP shock. We also discuss the implications of our model for demand shocks and the policymaker's use of aggregate-demand leverage (monetary policy) to respond to negative TFP shocks (supply-driven recession).

Consider a baseline parameterization with values that are commonly used in the New Keynesian literature. The elasticity of substitution between goods is $\theta = 3.8$, as in e.g. our previous work (Bilbiie et al, 2007, 2012); we assume the same elasticity of substitution for labor types ($\theta_w = 3.8$;

[29] Equation 28 uses the wage and CPI inflation conditions from Table 1 along with definition of the markup (4) to replace: $\pi_{w,t}^* - \pi_t^* = \hat{w}_t^* - \hat{\rho}_t^* - (\hat{w}_{t-1}^* - \hat{\rho}_{t-1}^*) = \hat{A}_t - \hat{A}_{t-1}$.

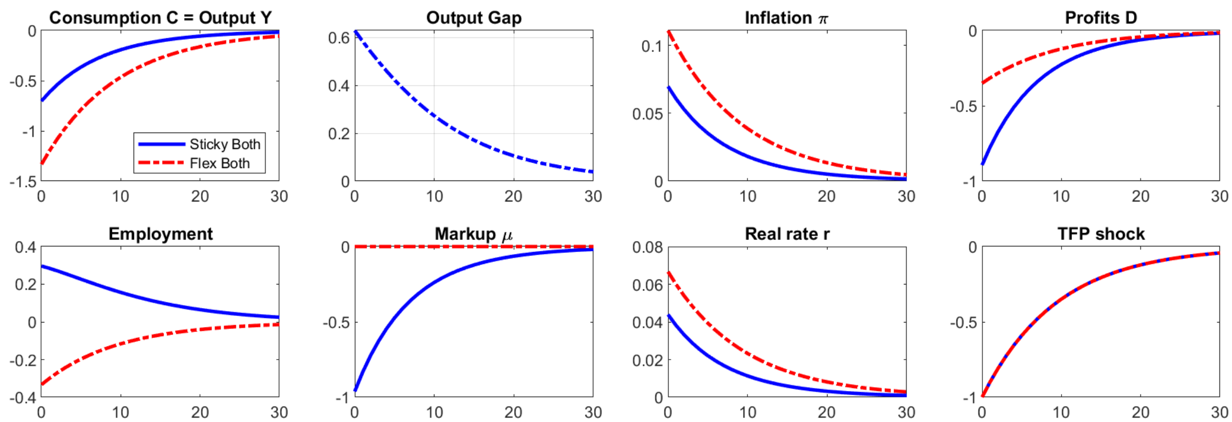
this parameter is largely inconsequential). The Pareto shape parameter is set in line with the trade literature to $k = 4$, see e.g. Melitz and Redding (2015); this implies a shape parameter for the size distribution of firms of $k/(\theta - 1) = 1.43$. The CRRA coefficient is $\sigma^{-1} = 0.5$, implying relatively low income effects on labor and thus satisfying the restriction needed for hours to increase in response to positive to TFP shocks generically;[30] Frisch labor supply elasticity is $\varphi = 1$. Finally, we assume an optimal subsidy eliminating the steady-state goods and labor market inefficiency: $1 + s_w = \frac{\theta}{\theta-1}\frac{\theta_w}{\theta_w-1}$.

For the nominal rigidity parameters, we adopt empirically standard Phillips-curve slopes consistent with medium-scale DSGE estimation and micro evidence on nominal rigidities. The price and wage Phillips curve slopes are set to $\psi = \psi_w = 0.01$ per quarter. Given the optimal labor subsidy, these translate into different values for the adjustment cost parameters, namely $\kappa = 280$ and $\kappa_w = \frac{\theta}{\theta-1}\frac{\theta_w}{\psi_w} = 515$, implying, notably, somewhat stickier wages than prices (our results are robust to variations around this magnitude). These values and their Calvo-equivalent implied frequencies (10 and 15 quarters, respectively) lie within the consensus range documented in Smets and Wouters (2007), Justiniano, Primiceri, and Tambalotti (2010, 2011, 2013), or Del Negro, Giannoni, and Schorfheide (2015). The Taylor rule response is $\phi = 1.5$.

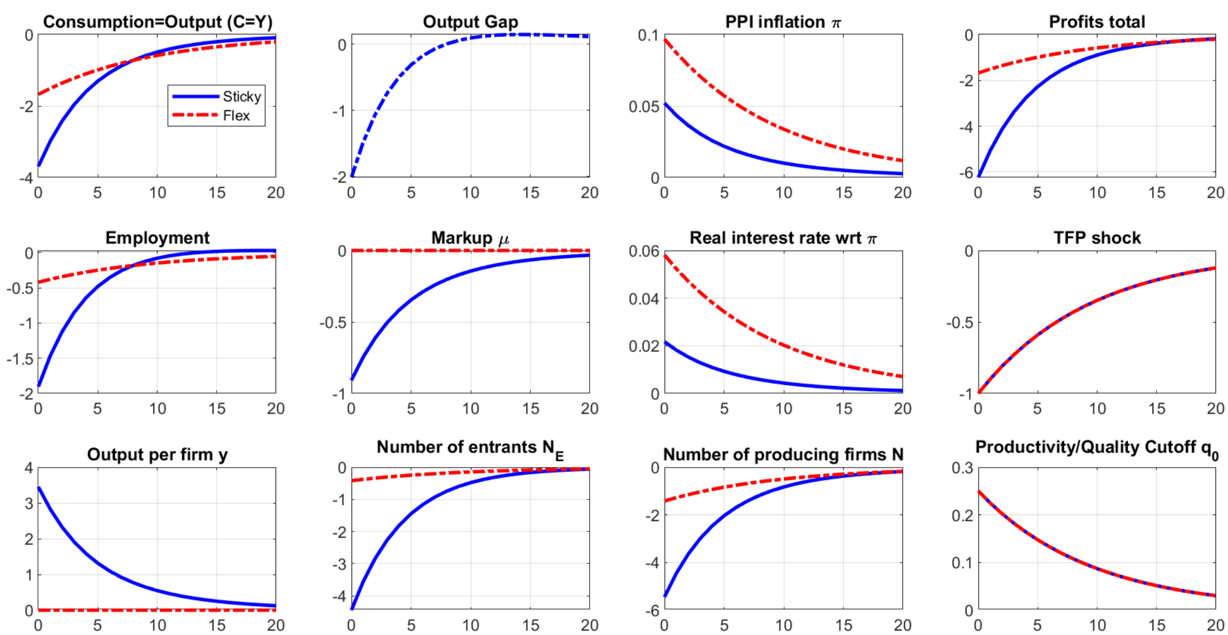
The associated impulse responses to a 1% fall in productivity with persistence $\rho_a = 0.9$ are shown in Figure 2. We use the same color convention to identify the flex-price equilibrium (dashed red) and its counterpart with sticky prices and wages (solid blue). The only exception is the output gap $\tilde{Y}_t$ (dashed blue). We start with the standard NK predictions (with both price and wage rigidities) in the top panel (a). Adding wage rigidities and adding adjustment costs (instead of fixed prices) does not change the implications we highlighted for the output gap: following a negative TFP shock, the output gap is positive, driven by the positive employment response under nominal rigidities.

The responses for our model with endogenous entry and selection are shown in the bottom panel (b). In sharp contrast, the response to the same negative TFP shock—persistent but transitory—is a negative output gap. As we previously discussed, the impact on output is now first order: In response to the 1% decline in productivity, the output gap on impact is roughly 2 percent and is associated with a 4 percent decline in entry: Lower productivity reduces profits and expected profitability, discouraging entry and shrinking productive capacity. Relative to the flexible-price equilibrium, the economy therefore experiences an aggregate-demand recession in response to a

[30]Since the parameter σ plays a dual role, this implies an elasticity of intertemporal substitution $\sigma = 2$; when viewed as capturing extra intertemporal forces from a more realistic dynamic model with investment, this number seems reasonable (see Woodford, 2003 for a similar argument justifying the same value in a model without investment).



(a) Exogenous Entry



(b) Endogenous Entry

Figure 2: Effects of a 1% TFP fall: flexible (red dash) vs sticky price and wage (solid blue)

supply disturbance.

In another sharp contrast with the standard NK predictions, the employment gap in our model is also negative. Interpreted through the lens of unemployment—defined by Galí (2011) as the wedge between sticky-wage and flexible-wage hours—a TFP-driven recession is associated with an increase in inefficient unemployment. This implication is consistent with a growing body of empirical evidence. VAR studies identifying transitory TFP shocks using utilization-adjusted productivity, news-shock decomposition, or sign restrictions find that adverse productivity shocks raise inflation while unemployment increases or remains flat and capacity utilization rises, indicating a contraction relative to the productive frontier rather than an expansion (Basu, Fernald, and Kimball 2006; Fernald and Wang 2016). Semi-structural state-space estimates of output gaps and DSGE-based measures of natural output deliver the same message: productivity improvements are associated with rising output gaps, while negative productivity shocks reduce them. In this sense, endogenous entry "transforms" TFP shocks into "true" supply shocks: they generate a negative co-movement between inflation and slack – typically associated in the standard NK model with ad-hoc cost-push or markup shocks.[31]

Indeed, for all variables common across our model with endogenous entry and the standard NK with exogenous entry, the impulse responses to TFP shocks in our model closely resemble those generated by markup shocks in the standard NK model, with one crucial exception: In our model, markups and profits fall following adverse productivity shocks, whereas markup shocks raise markups and imply countercyclical profits in the standard model. This distinction is empirically central. Profits are strongly pro-cyclical in the data, both unconditionally and conditional on supply disturbances. Aggregate profits rise following positive TFP shocks and fall following negative ones, and firm-level evidence shows that productivity improvements increase profits, survival, and entry, while productivity declines trigger exit (Basu et al. 2006; Fernald 2014; Syverson 2011; De Loecker and Eeckhout 2017; Gourio and Rudanko 2014; Bilbiie and Känzig 2023). By contrast, cost-push episodes such as large oil price shocks are stagflationary but are not associated with increases in aggregate profits; if anything, profits tend to fall as higher costs and lower quantities dominate any increase in markups (Barsky and Kilian 2004; Blanchard and Galí 2007; Bilbiie and Känzig 2023).

Finally, while our model predicts—like the standard NK model—that inflation rises following a negative TFP shock, inflation is lower than under flexible prices. Relative to the efficient allo-

[31] Endogenous entry thus provides a micro-foundation for endogenous cost-push shocks; in section 3.2, we show formally how entry appears as an endogenous cost-push, markup-shock wedge in the Phillips curve.

cation, there is therefore a deflationary gap associated with the increase in slack. This explains why, despite inflationary pressure, the appropriate policy response in our framework is expansionary: easing monetary policy closes the negative output gap by counteracting the endogenous contraction in productive capacity. We return to this point in the policy discussion below.

Although wage stickiness plays an important role in amplifying the quantitative effects of the entry multiplier, it is not sufficient on its own to generate aggregate-demand amplification of supply disruptions. In the next section 3.4, we prove analytically that price stickiness is necessary for our key responses (negative output gap to negative TFP shocks, and the entry multiplier): indeed, the model with sticky wages *only* behaves very much like the standard no-entry NK model.

Wage stickiness is also crucial for the model's behavior in response to demand shocks. Figure A.2 in Appendix A.5 reports the impulse responses to a temporary monetary expansion. With sticky prices but flexible wages, a demand shock lowers markups and profits, inducing firm exit—mirroring the standard NK result that profits are countercyclical to demand shocks. When both prices and wages are sticky, higher demand instead raises profitability and entry, increasing aggregate activity. Entry allows real wages to rise through a lower price index rather than higher nominal wages, dampening the fall in markups and restoring profits. We relegate the detailed discussion to the Appendix. In the following section, we derive a novel analytical condition for the relative degree of wage versus price stickiness required to obtain pro-cyclical entry.

To summarize, our model with endogenous entry delivers a combination of predictions for the impact of TFP shocks that is both empirically plausible and difficult to obtain in existing models: adverse productivity shocks generate stagflation—output gaps fall and inflation rises—while profits and entry are pro-cyclical. Neither standard New Keynesian TFP shocks nor cost-push shocks can replicate these joint dynamics. By endogenizing entry and productive capacity, our model reconciles stagflationary supply shocks with pro-cyclical profits, in line with the empirical evidence.

Our main finding is particularly important in the context of policy implications because endogenous entry overturns the standard New Keynesian prediction for the output gap. In our framework, adverse TFP shocks reduce profits and expected profitability, leading to a contraction in entry and in the mass of active producers. The resulting endogenous destruction of productive capacity implies that actual output falls *more* than flexible-price output, generating a negative output gap even though inflation rises. As a result, TFP-driven inflation is not a sign of overheating, and deflation following favorable productivity shocks can coexist with a positive output gap. Sticky prices combined with strong entry responses can therefore generate what might be described as "good deflation."

This logic has immediate implications for monetary policy. Central banks need not mechanically contract in response to inflation associated with productivity disturbances. The appropriate response depends on the source of inflation or disinflation—whether it reflects demand pressures or endogenous movements in productive capacity through entry. Our mechanism is thus crucially different from the "Keynesian supply shock" framework of Guerrieri et al. While both generate negative output gaps following adverse productivity shocks, the inflation response has the opposite sign. In Guerrieri et al., adverse supply shocks are deflationary at the aggregate level and thus observationally similar to demand contractions. In our model, adverse TFP shocks are inflationary because they trigger an endogenous contraction in the number of producers. The shock is therefore "supply–supply," not "supply–demand," a distinction with first-order consequences for policy design.

3.4 Closed-form Analytical Solution

To obtain a closed-form solution, we make two further assumptions: 1) Both Phillips curves are static ($\beta = 0$ in 24 and 25 only); and 2) the monetary policy rule perfectly neutralizes changes in expected inflation and thus keep the ex-ante real rate fixed, $\phi = 1$. These assumptions have been commonly adopted in the literature to derive analytical results for dynamic NK dynamic models; and they have been shown to be robust to forward-looking Phillips curves and standard Taylor rules (e.g., Bilbiie, 2025; Bilbiie and Känzig, 2023). Most importantly for our purpose, they do not substantially change the quantitative impulse responses, as we detail in Appendix A.4.

The Output Gap and Entry Multiplier

With these assumptions, the analytical solution of the model (in gaps) summarized by (24)-(28) yields:

$$\tilde{Y}_t = \sigma \frac{1}{\theta - 1} \tilde{N}_{Et} + \sigma \left(\epsilon_t + \Gamma \hat{A}_t \right), \tag{29}$$

$$\begin{aligned} \tilde{N}_{Et} = \Theta \tilde{N}_{Et-1} &+ \sigma \frac{\theta - 1}{\theta - \sigma} \Theta \left[\left(1 + \psi - \Psi_w\right) \epsilon_t - \epsilon_{t-1} \right] \\ &+ \frac{\theta - 1}{\theta - \sigma} \Theta \left[\left(\theta + \left(1 + \psi - \Psi_w\right) \sigma \Gamma \right) \hat{A}_t - \left(\theta + \sigma \Gamma \right) \hat{A}_{t-1} \right], \end{aligned} \tag{30}$$

where $\Psi_w = \psi_w \left[(\theta/\sigma) - 1 + (\theta - 1) \varphi \right] > 0$ denotes the slope of the wage Philips-curve and $\Theta \equiv \left[1 + \psi + \Psi_w \sigma / \left(\theta - \sigma \right) \right]^{-1} \geq 0$ is an index of overall price and wage stickiness that governs endogenous persistence. When both are flexible, $\Theta = 0$; it increases with stickiness reaching its upper-bound of 1 under fixed price and wage. The impact elasticities are then:

$$\frac{d\tilde{Y}_t}{d\hat{A}_t} = \frac{\sigma\theta}{\theta - \sigma}\Theta\left(1 + (1+\psi)\,\Gamma\right),$$
$$\frac{d\tilde{N}_{Et}}{d\hat{A}_t} = \frac{\theta - 1}{\theta - \sigma}\Theta\left[(1 + \psi - \Psi_w)\,\sigma\Gamma + \theta\right],$$

which directly leads to:

Proposition 3. *The output gap and entry multiplier response to a negative TFP shock are negative so long as prices are sticky "enough".*

The elasticities are both positive so long as $\psi < -(1+\Gamma)/\Gamma$, which implies a positive threshold for price stickiness (recall that lower ψ is inversely related to price stickiness).[32] In the limit as prices become flexible and only wages are sticky, our key results no longer hold: indeed, the model behaves very much like the standard, no-entry NK model where the output gap comoves negatively with TFP: $d\tilde{Y}_t/d\hat{A}_t = \frac{\sigma\theta}{\theta-\sigma}\Gamma < 0$. Relatedly, the entry multiplier disappears (the entry gap also comoves negatively with TFP): $d\tilde{N}_{Et}/d\hat{A}_t = \frac{\theta-1}{\theta-\sigma}\sigma\Gamma < 0$. These results, which extend to a quantitative version of the model, emphasize the key role of price stickiness for our mechanism. Sticky wages are instead essential for the next point.

The Response of Entry to Monetary Policy

We previously mentioned that our initial model from Section 2 with only sticky prices could not adequately capture the impact of monetary shocks – because in that simplest version monetary expansions are contractionary for entry. We now highlight how the addition of sticky wages addresses this issue. We can solve for the threshold value of wage stickiness beyond which the response of entry to a monetary easing becomes positive:

Proposition 4. *The response of entry to an interest rate cut is positive when wages are sticky "enough":*

$$\frac{d\hat{N}_{Et}}{d\varepsilon_t} = \frac{\theta - 1}{\frac{\theta}{\sigma}\left(1 - \frac{\Psi_w}{1+\psi}\right)^{-1} - 1} > 0 \qquad when \qquad \Psi_w < 1 + \psi.$$

[32] This requirement and ensuing result for the output-gap response also extend to the full dynamic solution beyond the impact elasticity. It is also a sufficient condition for the sign of the entry multiplier: $\psi < -(1+\Gamma)/\Gamma \implies (1 + \psi - \Psi_w)\,\sigma\Gamma + \theta > 0$. The output gap is zero under flexible wages, for any price-stickiness, as then $\Theta = 0$: a generalization of the envelope result in Section 2. Nevertheless, the entry multiplier is $\tilde{N}_{Et} = -(\theta - 1)\left(\epsilon_t + \Gamma\hat{A}_t\right)$, which generalizes the result in Section 2 by incorporating selection through k and general labor and inter-temporal elasticities through Γ.

The intuition for the first part is that when wages are flexible, the increase in labor demand translates directly into an increase in the nominal wage – with the ensuing negative consequences for firm profitability and exit. However, this counterfactual prediction is overturned once we add enough wage stickiness. Then, the increase in labor demand is accommodated by increased entry. This provides workers with real consumption gains at a given nominal wage – and thus substitutes for the sharp increase in nominal wages that is no longer feasible. In other words, when both nominal wages and individual prices are sticky, the consumption good price index P_t must fall to induce a rise in the real wage. This occurs via entry, dampening the fall in markups. Recall than when wages are sticky, entry is inefficiently low. The positive monetary shock can then boost real output by offsetting this inefficiency.

Then, so long as prices and wages are sticky "enough" (in the sense that propositions 3 and 4 are satisfied):

Proposition 5. *A monetary expansion closes the output gap:*

$$d\epsilon_t|_{d\tilde{Y}_t=0} = \Gamma d\left(-\hat{A}_t\right) - \frac{1}{\theta-1} d\tilde{N}_{Et} = \left(\Gamma + \frac{1}{1+\psi}\right) d\left(-\hat{A}_t\right) > 0$$

Focusing on the first equality, the first term $\Gamma d\left(-\hat{A}_t\right)$, on its own, would imply that a contraction (increase in the interest rate) is required in order to track r-star – which is going up with the adverse TFP shock. This delivers the same prediction as in the standard NK model. However, the second term $-1/(\theta-1)\, d\tilde{N}_t$ is key in overturning this prediction, driven by the reduction in entry and associated fall in endogenous productivity in response to adverse TFP. Thus, the demand-relevant r-star (with respect to inflation in the price index of the final good, π^c) is going down, and this requires a cut in interest rates – thus overturning the prediction of the standard NK model.[33] This intuition is sharpest in the extreme case of fixed prices and wages that we cover in the next section.

The implications we just derived are in sharp contrast to the standard NK model (of which we include a short recap and a novel analytical illustration in Appendix B). In that standard model, a negative TFP shock induces a positive output gap, which can be closed by tracking r-star, entailing a monetary contraction. Furthermore, this standard r-star-tracking policy is equivalent to completely stabilizing an "average" composite inflation measure. We now explore this equivalence in our

[33] In the standard NK model, a monetary contraction is required to close the output gap (see Appendix B for details):

$$d\epsilon_t^{\mathcal{N}}\Big|_{d\tilde{Y}_t^{\mathcal{N}}=0} = \Gamma^{\mathcal{N}} d\left(-\hat{A}_t\right) < 0.$$

model, where the required gap-closing policy instead entails a monetary expansion.

Divine Coincidence

In the standard NK model with sticky prices and wages, Woodford (2003) and Gali (2015) show that a modified version of divine coincidence holds: the central bank can close the output gap by stabilizing an appropriately aggregated measure of price and wage inflation, namely:

$$\bar{\pi}_t \equiv \frac{\psi_w}{\psi_w + \psi}\pi_t + \frac{\psi}{\psi_w + \psi}\pi_{w,t}.$$

This is no longer true in our model due to the entry multiplier. Writing the aggregate Phillips curve for this inflation measure by aggregating (24) and (25) we obtain:

$$\begin{aligned}\bar{\pi}_t &= \beta E_t \bar{\pi}_{t+1} + \bar{\psi}\left(\sigma^{-1} + \varphi - (1+\varphi)\frac{1}{\theta}\right)\tilde{Y}_t - \bar{\psi}\hat{\mu}_t \\ &= \beta E_t \bar{\pi}_{t+1} + \bar{\psi}\left(\sigma^{-1} + \varphi\frac{\theta - 1}{\theta}\right)\tilde{Y}_t - \bar{\psi}\frac{1}{\theta - 1}\tilde{N}_{Et},\end{aligned}$$

where we define $\bar{\psi} \equiv \frac{\psi_w \psi}{\psi_w + \psi}$. This shows how stabilizing that measure of average inflation still induces an output gap that will be proportional to the entry multiplier $\tilde{N}_{Et}$. In our model, monetary policy should target wage inflation: $\pi_{w,t} = 0$ closes the output gap. However, closing the output gap is no longer optimal for welfare. We describe this in further detail in section 4.

In our model, closing the output gap does not restore efficiency due a combination of the inflationary costs and distortions associated with entry. This inefficiency result for the output gap is reminiscent of a "cost-push" shock, traditionally motivated as an exogenous shock to the markup. In our model, the markup also varies endogenously with the endogenous entry response. It induces a similar response for the output gap as an exogenous "cost-push" shock.[34] However, in the standard NK model with exogenous entry, the optimal monetary response to such a stagflationary "cost-push" shock (resulting from an exogenous increase in the markup) is *tightening*. This worsens the output gap but raises welfare by reducing inflation. In the next section, we show how this response of monetary policy is reversed in our model with endogenous entry. Closing the output gap with monetary *easing* in response to the stagflationary TFP shock will improve welfare.

[34]As an (endogenous) time-varying real distortion, this has a similar flavor to the implications of real wage rigidities studied by Blanchard and Gali (2010).

4 Welfare and Policy Implications

In this section we analyze the welfare implications of our model combining TFP and monetary policy responses. This requires the full derivation for the time path of the nonlinear equilibrium response under both price and wage rigidities. Due to the complexities associated with this problem, we focus on the case of fixed price and wages and derive results for the direction of the welfare responses. Thus, we impose $\psi = \psi_w = 0$. As we described in Section 2, the resulting Euler equation under the fixed-real rate is (13), $\rho_t Y_t^{-\frac{1}{\sigma}} = \exp(-\epsilon_t)$. Solving for output, cutoff, and surviving firms yields:

$$Y_t = q_0 \left(\frac{\theta A_t - (\theta - 1)}{\theta f_O} \right)^{\frac{\sigma}{\theta - \sigma}} (A_t)^{-\frac{1}{k}\frac{\sigma(\theta-1)}{\theta-\sigma}} \exp(\frac{\sigma\theta}{\theta - \sigma}\epsilon_t), \tag{31}$$

$$L_t = \frac{\theta}{\theta - 1} Y_t^{1-\frac{1}{\sigma}} \exp(\epsilon_t), \tag{32}$$

$$q_{0t} = (A_t)^{-\frac{1}{k}} q_0,$$

$$N_t = Y_t^{1-\frac{1}{\sigma}} \frac{k - (\theta - 1)}{k} \frac{\theta A_t - (\theta - 1)}{\theta f_O} \exp(\epsilon_t),$$

where the steady-state cutoff is, evaluating 20 at steady state:

$$q_0 = \left(\frac{\theta - 1}{k - (\theta - 1)} \frac{f_O}{f_E} \right)^{\frac{1}{k}} q_{\min}.$$

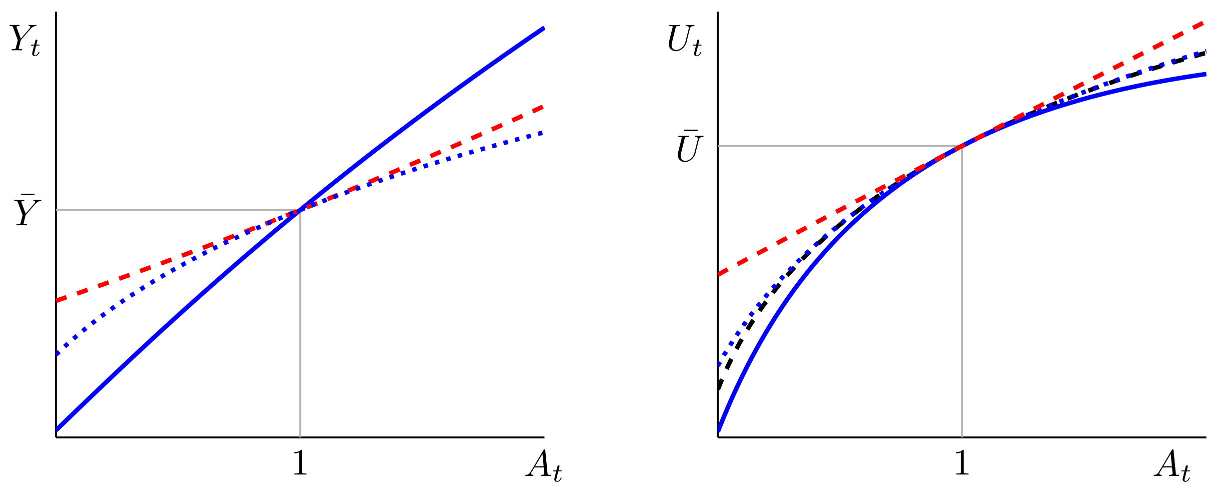
Once again, due to our Pareto parametrization, we recover the feature that selection does not interact with monetary policy (the ε_t shock does not affect the cutoff); though we clearly see how selection (via the shape of the Pareto parameter k) affects all the equilibrium variables. Of particular appeal is the loglinearized solution for output (a special case of the one we previously derived, evaluated under $\psi = \psi_w = 0$):

$$\hat{Y}_t = \sigma \frac{\theta - 1}{\theta - \sigma} \left(\frac{\theta}{\theta - 1} - \frac{1}{k} \right) \hat{A}_t + \frac{\sigma\theta}{\theta - \sigma} \epsilon_t.$$

This makes clear that the response of output to both TFP and monetary easing is positive, and the same two effects we identified in the flex-price equilibrium (increasing returns vs selection) modulate the magnitude of the TFP elasticity.

Welfare implications of alternative policies

Panel (a) on the left in Figure 3 plots the nonlinear solution of our model for output Y_t as a function of the TFP shock A_t. This is the counterpart to panel (b) of Figure 1 for our initial model with just fixed prices and log-utility ($\sigma = 1$). The flex-price curve is in red dash. In our calibration, our choices for θ, σ, and k are such that the increasing returns are substantial (the elasticity to TFP is 3.44 under the baseline calibration), yet the output curve is quasi-linear and looks very similar to our previous model with log-utility. Most importantly, the sticky price curve (solid blue) looks very different as it no longer exhibits the envelope property for output of the previous simple model. There are first-order effects of nominal rigidities, as we have previously discussed. Thus, the first order negative output gap for adverse TFP shocks that we have highlighted (the difference between the two curves). As we have emphasized, these predictions for the output gap reverse the predictions of the standard NK model in the panel (a) counterpart from Figure 1. Before discussing the dotted-blue curve in panel (a), we turn to the normative implications of the output gap between flex and sticky prices with no monetary policy intervention.



(a) Output (b) Utility

Figure 3: The Output Gap and Utility $U(C_t, L_t)$: Flexible Prices (red dash), Sticky Prices (no intervention: solid blue, U-max: dotted blue, gap-closing: black dash)

Panel (b) on the right plots nonlinear (lifetime) utility using the full model solution. Since the equilibrium is efficient under flexible prices and wages (by virtue of the constant labor subsidy), the flex-price utility curve (dashed red) in that equilibrium is an upper envelope of utility under nominal rigidity for any given monetary policy intervention. The solid blue curve again depicts the sticky price case with no monetary intervention. This highlights the inefficient expansion (positive output gap) in response to positive technology shocks: there is too much entry relative to the efficient flex-price case (the entry multiplier); this increases output, but the associated labor market distortion dominates and yields lower welfare.

The dashed black curve represents the outcome with a monetary policy designed to close the output gap. (By construction, the output curve for this case overlaps with the flex-price red-dashed curve.) As we have shown in the previous section, this is achieved via monetary easing in response to adverse TFP shocks – and conversely via a monetary contraction in response to positive TFP shocks. Evidently, this monetary intervention is associated with welfare gains relative to the laissez-faire monetary policy (solid blue). This is not an immediate consequence of inefficiently low output under negative TFP shocks and inefficiently high output under positive shocks. There is also the response of hours (dis-utility of work) that shows up in utility but not in output. Neither is efficient under laissez-faire monetary policy. Thus, correcting only the inefficiency in output is not guaranteed to improve welfare; However, we show this is indeed the case:

Proposition 6. *A monetary policy expansion improves welfare when starting from a situation with below-trend TFP associated with a negative output gap; conversely, a monetary contraction improves welfare with above-trend TFP associated with a positive output gap.*

We prove this formally in Appendix C.1 by replacing the equilibrium expressions derived in (31) in the utility function and deriving comparative statics with respect to a monetary expansion in a low-TFP state. A monetary contraction is instead optimal in an (above-trend) TFP-driven expansion. This result for monetary policy overturns a key prediction of the standard NK model. In that model, a monetary tightening is required to close the (positive) output gap—thereby also improving welfare—in response to negative TFP shocks. And conversely, the standard NK model requires monetary easing to close the negative output gap in response positive TFP shocks.[35] Appendix C.1 also contains a proof for this welfare result for the standard NK case and reviews the intuition for it.

Finally, we derive the welfare-maximizing monetary policy (a simple, static version of a Ramsey problem) by maximizing utility subject to the equilibrium values derived under fixed prices and wages in (31). This delivers our final proposition:

[35]In that standard NK model, this gap-closing policy would track r-star. In our economy, tracking r-star would uniformly deteriorate welfare relative to no intervention: it induces a contraction (r-star still goes up in our model) when the economy needs an expansion. This is one of the sharpest policy lessons of our model: simply tracking r-star can dramatically backfire in an economy with endogenous entry. Of course, in our model the gap could be closed by tracking the *true demand-relevant natural rate* (rather than the standard PPI-based r-star), which entails a monetary expansion; due to measurement issues, this welfare-relevant object that includes endogenous productivity via the variety effect is not directly observable.

Proposition 7. *Optimal monetary policy is countercyclical, requiring an expansion in a TFP-driven recession.*

$$\begin{aligned}\exp(\epsilon_t^{opt}) &= \left(\frac{\theta}{\theta-1}\right)^{-\frac{\varphi}{1+\varphi}} Y_t^{-\left(1-\frac{1}{\sigma}\right)\frac{\varphi}{1+\varphi}} \\ &= \left(\left(\frac{\theta}{\theta-1}\right)^{\sigma\frac{\theta-1}{\theta-\sigma}} \left(\theta A_t - (\theta-1)\right)^{\frac{\sigma-1}{\theta-\sigma}} \left(A_t\right)^{-\frac{1}{k}(\theta-1)\frac{\sigma-1}{\theta-\sigma}}\right)^{-\frac{\varphi}{1+\varphi\sigma\frac{\theta-1}{\theta-\sigma}}}\end{aligned}$$

Furthermore, to first order, this is identical to the monetary easing required to close the output gap ϵ_t^{opt}*. Thus, under the optimal policy, the output gap is stabilized to a linear approximation.*

Optimal policy requires easing "automatically" whenever there is a recession (the first line) and also in reduced form, whenever there is a negative TFP shock (the second line). This and the proof of the second statement can be seen most clearly in log-linearized form:

$$\begin{aligned}d\epsilon_t^{opt} &= -\left(1-\frac{1}{\sigma}\right)\frac{\varphi}{1+\varphi}d\hat{Y}_t, \\ &= -\frac{(\sigma-1)}{(\theta-\sigma)\,\varphi^{-1}+\sigma\,(\theta-1)}\left(\theta-\frac{\theta-1}{k}\right)d\hat{A}_t, \\ &= (\Gamma+1)\left(-d\hat{A}_t\right).\end{aligned}$$

This shows that $d\epsilon_t^{opt}$ is identical to the gap-closing value $d\epsilon_t|_{d\tilde{Y}_t=0}$ we derived in Proposition 5. However, this equivalence holds *only to first order*. The gap-closing policy is still inefficient due to higher-order terms, for two reasons. First, the inefficient change in entry. And second, when prices are not fixed, there is also inflation.[36] Both of these welfare costs feature analytically—even when the output gap is closed—in the second-order approximation to the welfare function we derive in Appendix A.2.

The dotted blue curves in Figure 3 plot output and welfare under this optimal policy. Even though the output gap is closed to first order under that optimal policy $\tilde{Y}_t^{opt} = 0$, it is still negative to higher order and the policy is still "second best" due to the welfare costs associated with the change in entry.[37] Panel (a) highlights how the output gap is closed to first order, but not to higher order. Panel (b) highlights how the gap-closing policy is optimal to first order. In fact, in our calibration, even the higher order differences between that policy and optimal one are barely noticeable.

The above results apply for fixed prices and wages. One might worry that in a model with *sticky*

[36] In our initial model in Section 2, we characterized both inefficiencies using a second-order approximation to the welfare function.

[37] Appendix A.2 also derives the second order welfare costs for this optimal policy.

(not fixed) prices and wages, a gap-closing policy would be inflationary and thus welfare-reducing. It is important to understand why this is not a concern in our model, even though it is a legitimate worry. Indeed, in the standard NK model, a monetary easing will close the output gap in response to a cost-push (markup) shock—but generically *worsen* welfare because of its inflationary effect: the optimal policy in the standard NK model in response to a stagflationary cost-push shock is a monetary contraction, trading off its inflation-reduction benefit against a worsening of the output gap. Our main result is that even though the (TFP) supply shock is now stagflationary, the welfare-maximizing monetary response is reversed: it amounts to *easing* instead of contracting.

That a monetary easing in response to a negative TFP shock *does not create extra inflation* is a subtle point that requires some elaboration. Prima facie, this makes it sounds like a downward-sloping Phillips curve. However, this is not about the PC slope, but rather about an endogenous, entry-driven shift of the curve when demand shifts. A negative TFP reduces both productivity and profits and implies fewer entrants, fewer varieties, and lower markups. Inflation rises due to these variety and markup effects. But, crucially when the central bank eases, demand and profitability rise, which boosts entry. Variety expands, markups increase, and consequently inflation declines again. The mechanism behind "easing curtails inflation" is there: not a slope reversal, but an endogenous "supply restoration" channel. In other words, it is not that the Phillips curve itself flips but that policy-induced entry shifts the effective supply curve endogenously.

Figure C.1 in the Appendix C.2 substantiates these points by plotting the impulse-responses of all variables under the gap-closing policy (with black dash and crosses), along with the same baseline cases previously pictured. First, trivially, wage inflation is zero (the gap-closing policy is the one that fully stabilizes wage inflation to first order in our model). Second, more subtly, PPI inflation under the gap-closing policy is *even lower* than no monetary response (the main and robust finding is that it is not significantly higher; it only happens to be slightly lower under our baseline parameterization). So even going back to the more extreme intervention from Figure 3 of closing the output gap entirely (optimal is to intervene less), does not generate inflation. And this is despite an exogenous monetary expansion by itself being ceteris paribus still inflationary, as we illustrated in Figure A.2 in Appendix A.5. Indeed, note that easing "too much" – in this exercise, beyond what is needed to close the output gap – would then induce higher inflation just like we show for the case of monetary easing on its own.

5 Conclusion

The qualitative predictions of our framework for an adverse (negative) TFP shock, set against the standard New Keynesian model with and without cost-push shocks, can be summarized compactly as follows. All three cases predict an ensuing rise in inflation (driven by worse technology in the TFP cases, and by an exogenous markup wedge in the cost-push case) even under flexible prices – and price rigidity dampens but does not eliminate this upward pressure. The cases diverge sharply, however, on the output gap and on the appropriate policy response.

In the standard NK model, an adverse TFP shock opens a *positive* output gap, counterfactually pairing adverse FTP shocks with an overheating economy. The prescribed response is to tighten, which lowers inflation and closes the gap without trade-off. Cost-push shocks were introduced precisely to repair this counterfactual, delivering the stagflationary comovement of higher inflation and a negative output gap. But this does not address the NK model's prediction for monetary policy: tightening is still the optimal response for welfare with cost-push shocks. In that case, it involves a trade-off of widening the output gap in return for lower inflation.

Our model with endogenous entry delivers the same stagflationary comovement—inflation rises and the output gap turns negative—but without this trade-off for the monetary policy response: *easing* raises the output gap while inflation falls. In addition, our model predicts that profits and entry are procyclical, consistent with the evidence. This resolves an inherent tension in New Keynesian models, where cost-push, markup shocks can generate stagflation but only at the cost of countercyclical profits.

In our framework, an adverse TFP shock thus becomes, in both its positive and its normative implications, a true supply shock. This is the paper's central message: endogenous entry allows productivity shocks to generate stagflationary dynamics together with procyclical profits—a combination existing models struggle to reconcile—and it implies that monetary easing, not tightening, is the right response to adverse TFP shocks. The associated demand expansion increases profitability, induce entry, and raises output by expanding the economy's productive base. This channel is absent in representative-firm models and highlights the importance of firm dynamics for policy transmission.

Even though our model features firm heterogeneity and endogenous selection, we show how those can be embedded within a modeling framework that closely resembles the structure of New Keynesian models – with entry and selection appearing as wedges in the Phillips curves and resource constraints. This structure allows us to deliver closed-form analytical results while preserving the familiar logic and transparency of the New Keynesian literature.

More broadly, our results suggest that firm dynamics are not a second-order feature for monetary policy analysis. By endogenizing the creation and destruction of productive capacity, entry transforms the interpretation of inflation, slack, and the appropriate policy response to supply disturbances. Productivity-driven disinflation need not signal weakness, and productivity-driven inflation need not signal overheating. Understanding which margin of adjustment is operative—intensive or extensive—is therefore essential for both positive and normative analysis.

References

Aghion, P., and P. Howitt., 1992, "A Model of Growth through Creative Destruction." Econometrica, 60(2), 323–351

Alexopoulos, M. 2011, " Read All About It! What Happens Following a Technology Shock?" American Economic Review, 101(4), 1144–1179.

Argente, D., M. Lee, and S. Moreira, 2018, "Innovation and product reallocation in the great recession," Journal of Monetary Economics

Auerbach, A., Y. Gorodnichenko and D. Murphy, 2021, "Inequality, Fiscal Policy and COVID19 Restrictions in a Demand-Determined Economy," European Economic Review

Baqaee, D. and E. Farhi, 2020, " Supply and Demand in Disaggregated Keynesian Economies with an Application to the Covid-19 Crisis," American Economic Review

Barsky, R. and L. Kilian 2004: "Oil and the Macroeconomy Since the 1970s." Journal of Economic Perspectives.

Basu, S., J. G. Fernald, and M. S. Kimball, 2006. "Are Technology Improvements Contractionary?," American Economic Review, 96(5), 1418-1448

Bergin, P. and G. Corsetti, 2008, "The extensive margin and monetary policy," Journal of Monetary Economics, 55(7), pp. 1222–1237

Bilbiie F. 2008. "Limited Asset Market Participation, Monetary Policy, and (Inverted) Aggregate Demand Logic". Journal of Economic Theory 140, 162-196.

Bilbiie, F. 2020, "The New Keynesian Cross," Journal of Monetary Economics

Bilbiie, F. 2021, "Monetary Neutrality With Sticky Prices and Free Entry," Review of Economics and Statistics

Bilbiie, F. 2025, "Monetary Policy and Heterogeneity: An Analytical Framework," Review of Economic Studies

Bilbiie, F. O., F. Ghironi, and M. Melitz, 2007, " Monetary Policy and Business Cycles with Endogenous Entry and Product Variety," NBER Macroeconomics Annual, vol. 22, pp. 299-353

Bilbiie, F. O., F. Ghironi, and M. Melitz, 2012, "Endogenous Entry, Product Variety and Business Cycles," Journal of Political Economy,120(2), 304-345

Bilbiie, F. O., F. Ghironi, and M. Melitz, 2019, "Monopoly Power and Endogenous Product Variety: Distortions and Remedies" American Economic Journal: Macroeconomics.

Bilbiie, F. O., I. Fujiwara, and F. Ghironi, 2014, "Optimal Monetary Policy with Endogenous Product Variety," Journal of Monetary Economics, 2014, vol. 64, pp. 1-20.

Bilbiie, F. O. and D. Känzig 2023: "Greed? Profits, Inflation, and Aggregate Demand." NBER Working Paper 31618

Blanchard, O. and J. Galí, 2007, “Real Wage Rigidities and the New Keynesian Model”, Journal of Money, Credit, and Banking

Blanchard, O., and J. Galí. 2010. "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment." American Economic Journal: Macroeconomics 2 (2): 1–30.

Broer, T., N. J. Hansen, P. Krusell, and E. Oberg, 2020 “The New Keynesian transmission mechanism: a heterogeneous agent perspective,” Review of Economic Studies

Cacciatore, M. and G. Fiori, 2016, “The Macroeconomic Effects of Goods and Labor Markets Deregulation,” Review of Economic Dynamics

Campbell, J., 1998, “Entry, Exit, Embodied Technology, and Business Cycles.” Review of Economic Dynamics, 1(2), 371-408

Cantore, C., M. León-Ledesma, P. McAdam, A. Willman, 2014, “Shocking Stuff: Technology, Hours, And Factor Substitution,” Journal of the European Economic Association, 12(1), 108-128

Cantore, C., F. Ferroni, and M. León-Ledesma, 2021, "The Missing Link: Monetary Policy and the Labor Share ." Journal of the European Economic Association, 19(3), 1592-1620.

Cesa-Bianchi, A. and A. Ferrero, 2021, “The Transmission of Keynesian Supply Shocks,” Mimeo.

Chaney, T., 2008 "Distorted Gravity: The Intensive and Extensive Margins of International Trade." American Economic Review, 98(4), 1707-1721

Chari, V.V., P. Kehoe, and E. McGrattan, 2008. “Are structural VARs with long-run restrictions useful in developing business cycle theory?,” Journal of Monetary Economics, 55(8), 1337-1352

Chatterjee, S., and R. Cooper, 2014, “ Entry and Exit, Product Variety and the Business Cycle,” Economic Inquiry

Christiano, L., M. Eichenbaum, and C. Evans, 1997, “Sticky Price and Limited Participation Models of Money: A Comparison”, European Economic Review

Christiano, L., 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., M. Eichenbaum, and R. Vigfusson, 2003, “ What Happens after a Technology Shock?” Mimeo

Clementi, G. L., and B. Palazzo. 2016. “Entry, Exit, Firm Dynamics, and Aggregate Fluctuations.” American Economic Journal: Macroeconomics, 8 (3): 1-41

Colciago, A. and F. Etro, 2010 “Endogenous Market Structures and the Business Cycle” The Economic Journal

Colciago, A. and R. Silvestrini, 2022, “Monetary Policy, Productivity, and Market Concentration,” European Economic Review

Cook, D., 2001, “ Time to Enter and Business Cycles,” Journal of Economic Dynamics and Control 25 (8): 1241-1261

Cooke, D. and T. Damjanovic, 2020, “Monetary Policy and Welfare with Heterogeneous Firms and Endogenous Entry,” mimeo

De Loecker, J., and J. Eeckhout, 2017 "The Rise of Market Power and the Macroeconomic Implications." NBER Working Paper No. 23687.

Devereux, M. B., A. C. Head, and B. J. Lapham, 1996, " Aggregate Fluctuations with Increasing Returns to Specialization and Scale," Journal of Economic Dynamics and Control 20: 627-656.

Dixit, A. K., and J. E. Stiglitz, 1977, " Monopolistic Competition and Optimum Product Diversity," American Economic Review 67: 297-308.

Dixon, H. and A. Savagar, 2020, "Firm Entry, Excess Capacity and Endogenous Productivity," European Economic Review

Edmond, C., V. Midrigan, and D. Xu, 2019, "How costly are markups?," Journal of Political Economy

Erceg, C. J., D. W. Henderson, and A. T. Levin, 2000, " Optimal monetary policy with staggered wage and price contracts". Journal of Monetary Economics 46 (2), 281-313.

Ethier, W., 1982, "National and international returns to scale in the modern theory of international trade," American Economic Review, 72, 389-405.

Etro, F. and L. Rossi, 2015, "New-Keynesian Phillips curve with Bertrand competition and endogenous entry," Journal of Economic Dynamics and Control

Fernald, John G., 2014, "Productivity and Potential Output Before, During, and After the Great Recession." NBER Macroeconomics Annual (29), 1-51

Fernald, John G., and J. Christina Wang, 2016, "Why Has the Cyclicality of Productivity Changed? What Does It Mean?" Annual Review of Economics, 8, 465-496

Fornaro, L. and M. Wolf, 2020, "The Scars of Supply Shocks," Mimeo

Galí, J. 1999, " Technology, Employment and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" American Economic Review 89/1, 249–7

Galí, J., 2011, " The Return of the Wage Phillips Curve", Journal of the European Economic Association, 9(3), 436–461.

Galí, J. 2015, "Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications", Princeton University Press, Second Edition

Ghironi, F. and M. J. Melitz., 2005, "International Trade And Macroeconomic Dynamics With Heterogeneous Firms," Quarterly Journal of Economics, 120(3), 865-915

Gourio, F., and L. Rudanko., 2014, "Customer Capital." The Review of Economic Studies, 81(3), 1102-1136

Guerrieri, V., G. Lorenzoni, I. Werning, and L. Straub, 2020 "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages," American Economic Review

Gutierrez, G., C. Jones and T. Philippon, 2021, "Entry Costs and Aggregate Dynamics," mimeo

Hamano M. and F. Zanetti, 2017, "Endogenous Product Turnover and Macroeconomic Dynamic," Review of Economic Dynamics

Hamano M. and F. Zanetti, 2022, "Monetary Policy, Firm Heterogeneity, and Product Variety," European Economic Review

Hamano, M., 2025, "Heterogeneity in Tastes, Productivities, and Macroeconomic Volatility." Mimeo

Hamano, M., and F. Pappadà., 2023 "Exchange Rate Policy and Firm Heterogeneity." IMF Economic Review, 71(3), 759-790

Helpman, E., M. J. Melitz, and S. R. Yeaple. "Export Versus FDI with Heterogeneous Firms." American Economic Review, 94(1), 300-316

Hopenhayn, H. A., 1992, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." Econometrica, 60(5), 1127-1150

Jaimovich, N., 2007, “Firm Dynamics and Markup Variations: Implications for Multiple Equilibria and Endogenous Economic Fluctuations,” Journal of Economic Theory, 137:1, 300-325.

Jaimovich, N., and M. Floetotto, 2008, Firm Dynamics, Markup Variations and the Business Cycle, Journal of Monetary Economics, 55:7, 1238-1252.

Justiniano, A., G. E. Primiceri, and A. Tambalotti, 2010. "Investment Shocks and Business Cycles." Journal of Monetary Economics, 57(2), 132–145.

Justiniano, A., G. E. Primiceri, and A. Tambalotti, 2011. "Investment Shocks and the Relative Price of Investment." Review of Economic Dynamics, 14(1), 101–121.

Justiniano, A., G. E. Primiceri, and A. Tambalotti, 2013. "Is There a Trade-Off between Inflation and Output Stabilization?" American Economic Journal: Macroeconomics, 5(2), 1–31

Kaplan, G. and P. Zoch, 2020, “Markups, Labor Market Inequality and the Nature of Work,” Mimeo

Kim, J., 2004. “What determines aggregate returns to scale?,” Journal of Economic Dynamics and Control, 28(8), 1577-1594

Krugman, P., 1980, “Scale Economies, Product Differentiation, and the Pattern of Trade,” American Economic Review, 70:5, 950-59

Lee, Y. and T. Mukoyama, 2018. “A model of entry, exit, and plant-level dynamics over the business cycle,” Journal of Economic Dynamics and Control, 96(C), 1-25.

Lewis, V., 2009, “ Business Cycle Evidence on Firm Entry.” Macroeconomic Dynamics

Lewis, V. and C. Poilly, 2012, “Firm Entry, Markups and the Monetary Transmission Mechanism,” Journal of Monetary Economics 59(7), 670-685.

Melitz, M., 2003, “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” Econometrica 71(6), pp. 1695-1725

Melitz, M. J., and S. J. Redding, 2015, "New Trade Models, New Welfare Implications." American Economic Review, 105(3), 1105–1146

Michelacci,C., L. Paciello, and A. Pozzi, 2019, “The extensive margin of aggregate demand,” Review of Economic Studies

Romer, P., 1986, Increasing Returns and Long-Run Growth, Journal of Political Economy, Vol. 94, No. 5. pp. 1002-1037

Romer, P., 1987, Growth Based on Increasing Returns Due to Specialization, American Economic Review, Vol. 77, No. 2, pp. 56-62

Schmitt-Grohé, S. and M. Uribe, 2006, “ Optimal fiscal and monetary policy in a medium-scale macroeconomic model”. NBER Macroeconomics Annual 2005, vol. 20, 383-462.

Syverson, C., 2011, "What Determines Productivity?" Journal of Economic Literature, 49(2), 326-365

Woodford, M., 2003, Interest and prices: foundations of a theory of monetary policy, Princeton University Press

Woodford, M., 2020 “Effective Demand Failures and the Limits of Monetary Stabilization Policy,” American Economic Review.

Appendix

A Derivations and Complete General Model Outline

In this Appendix, we first provide some key derivations for the simplest benchmark model and then outline the full general model with both sticky prices and wages.

A.1 Second-order solution in the simplest model

To understand what drives the key result that completely overturns the propagation of supply shocks in the no-entry New Keynesian model, we compare the two free-entry equilibria using a second-order approximation around the point $Y = Y^*$. This leads to our Proposition 8. Let temporarily small letters denote log deviations, and small letters *with a tilde* denote deviations as a share of the steady-state value, i.e. $\tilde{x}_t \equiv (X_t - X)/X$. Of course, when focusing on a first order approximation this distinction is immaterial, since to first order $\tilde{x}_t \simeq x_t$; but to second order the two are different ($\tilde{x}_t \simeq x_t + \frac{1}{2}x_t^2$).

Proposition 8. *To second order, output under flexible and sticky prices is, respectively:*

$$\begin{aligned} \tilde{y}_t^* &\simeq a_t + \frac{1}{2}a_t^2, \\ \tilde{y}_t &\simeq a_t + \frac{1}{2}(1-\theta)\, a_t^2. \end{aligned} \tag{A.1}$$

Therefore, the "output gap" is:

$$\tilde{y}_t - \tilde{y}_t^* \simeq -\frac{1}{2}\theta a_t^2. \tag{A.2}$$

But the output gap response is always negative, due to the second-order effect. There is also an asymmetry: output increases by less in response to positive shocks, but falls by more in response to negative shocks. For large negative shocks in particular, the response under sticky prices can be substantially larger.

Dissecting the Mechanism. The key to understanding these second-order (concavity) effects lies in the equilibrium dependence of aggregate output to the number of intermediate inputs $Y(N)$. As we already noted, N itself is a linear function of A. In other words, N is (linearly) amplified through our entry multiplier. Y is then amplified further through second-order effects. In particular, consider the "aggregate production function":

$$Y_t = N_t^{\frac{\theta}{\theta-1}} A_t \left(\frac{\bar{L}}{N_t} - f_E \right). \tag{A.3}$$

A second-order approximation around the steady-state equilibrium yields:

$$\begin{aligned} \tilde{y}_t &\simeq \frac{\bar{L} - \theta f_E N}{(\theta-1)\left(\bar{L} - f_E N\right)} n_t + \frac{1}{2} \frac{(2-\theta)\,\bar{L} - \theta f_E N}{(\theta-1)^2 \left(\bar{L} - f_E N\right)} n_t^2 \\ &= \frac{\theta}{\theta-1} \frac{1 - \frac{N}{N^*}}{\theta - \frac{N}{N^*}} n_t - \frac{1}{2} \frac{\theta}{(\theta-1)^2} \frac{\theta + \frac{N}{N^*} - 2}{\theta - \frac{N}{N^*}} n_t^2 \\ &= -\frac{1}{2} \frac{\theta}{(\theta-1)^2} n_t^2. \end{aligned} \tag{A.4}$$

The first linear term drops out given that $N = N^* = \bar{L}/\theta f_E$ (The *steady-state* value is the same as in the efficient flexible-price equilibrium). Thus, the output gap is always zero to a first order. Intuitively, entry implies an adjustment mechanism such that if demand is too high and therefore profits too low (due to sticky prices), some firms exit. This reduces aggregate output through the variety effect, and since variety provision is efficient with CES preferences and flexible prices, output is the same as in the flex-price level to a first order (although the number of varieties is inefficiently small); in other words, the individual per-firm labor demand shifts, but the number of firms moves so as to offset the effect on aggregate labor demand to first order. This is a more general case of the local neutrality result in response to monetary shocks first emphasized in Bilbiie (2021).[38]

The amplification of exit (lower N) to a negative productivity shock (lower A) thus generates amplification for the fall of real output Y. This effect, which operates through the concavity of consumption in the number of intermediates, is increasing (in absolute value) with the benefit of variety $(\theta-1)^{-1}$.[39] It is thus determined by the same parameter governing the amplification of entry itself. Naturally, when the benefit of variety (the degree of increasing returns to specialization) vanishes there is no curvature of output in the number of varieties. The degree of increasing returns to specialization is crucial for the balance between the extensive and intensive margin adjustment that becomes distorted under sticky prices. More intensive-margin adjustment would be desirable but is unfeasible, and this distortion is *less* important when goods are closer substitutes:

[38] We show in Appendix A.2 that this first-order irrelevance holds generally for arbitrary price stickiness.

[39] In Appendix A.2, we provide an alternative interpretation (for an arbitrary degree of price stickiness). We take a second-order approximation to household utility (Woodford, 2003, Chapter 6) delivering a loss function in squared inflation and the gap of the number of firms from its flex-price level (equation (A.6) therein); replacing the latter equilibrium expressions we obtain the equivalent of Proposition 8 above.

θ larger, less returns to scale (less benefit of variety), less distortion. To summarize, θ determines both entry amplification and the concavity distortion, but has opposite effects on these two forces. Overall, the net effect of θ is to amplify the difference between flexible and sticky-price allocations.

A.2 Stabilization Policy Implications and Second-order approximation to utility

The simplest, stripped-down version of our model with only sticky prices is unsuitable for studying monetary policy or demand shocks: a monetary expansion induces a counterfactual prediction of increased exit. The reason is a well-known feature of the standard NK model: markups and profits are countercyclical to demand shocks. As demand goes up, labor demand shifts up, increasing the real wage and real marginal cost and eroding margins; with free entry, this leads to exit. We show in Section 3 that extending the model to introduce wage rigidity solves this issue and its monetary policy implications—while also providing a quantitatively powerful amplification mechanism.. Here, we nevertheless prove a "divine coincidence" result analogous to fixed-entry economies (Blanchard and Gali, 2007): the central bank can replicate the efficient flexible-price level of output while at the same time *also* stabilizing inflation (in a version where prices are not fixed but arbitrarily sticky).[40]

This can be seen directly by replacing the free-entry condition written for arbitrarily sticky prices, as a function of the markup (8) into the aggregate accounting equation (7), obtaining (assuming log utility in consumption without loss of generality, so that hours are constant at $\bar{L}$):

$$Y_t = \frac{1}{\mu_t} \left[\frac{1}{f_E} \left(1 - \frac{1}{\mu_t} \right) \right]^{\frac{1}{\theta - 1}} \left(A_t \bar{L} \right)^{\frac{\theta}{\theta - 1}} . \tag{A.5}$$

It follows directly that stabilizing the markup at its flexible-price level $\mu^* = \theta / (\theta - 1)$ and thus eliminating inflation in individual prices, delivers the flexible-price level of real activity Y_t^* in Table 1, or vice versa: there is no conflict between the two objectives, i.e. "divine coincidence".

This can be further illustrated by taking a second-order approximation to household utility, following e.g. Woodford (2003, Chapter 6). The derivation, described in the Appendix A.2 (for the benchmark case of log utility in consumption that isolates our channel) delivers the quadratic loss function:

$$\mathcal{L}_t \simeq -\frac{1}{2} \left[\kappa \pi_t^2 + \frac{\theta}{(\theta - 1)^2} \left(n_t - n_t^* \right)^2 \right] , \tag{A.6}$$

capturing the costs of (squared) individual-prices inflation (κ is the Rotemberg adjustment-cost

[40] We are grateful to an anonymous referee who suggested both emphasizing this property and the connection with the second-order welfare approximation.

coefficient) and the gap of the number of firms relative to the flexible-price level. The latter is a sufficient statistic for the welfare loss, due to the local neutrality result emphasized above: output is equal to the first order to flexible-price output, but it is different to the second order because of the extensive-margin concavity effects discussed above. Replacing the equilibrium expressions of n_t (under fixed prices, so for $\pi_t = 0$) and n_t^*, we obtain exactly the second-order approximation of the output gap in Proposition 8 above.

This welfare criterion can be used to assess the implications of different, suboptimal policy rules. For example, the suboptimal rule of fixing the money supply (or, with fixed prices, the real interest rate) "costs" $\mathcal{L}_t = -\frac{1}{2}\theta a_t^2$ in the endogenous-entry model; while in the fixed-entry model (where the loss function is readily derived in e.g. Woodford, 2003; or Gali, 2015), it is $\mathcal{L}_t^N = -\frac{1}{2}\left(1+\varphi\right) a_t^2$, where φ is the inverse labor elasticity . This illustrates that the models are not directly comparable (they are not nested in one another); the welfare costs are determined by different features of the economy—the benefit of variety and elasticity of substitution, in the former case; and labor elasticity, in the latter. Furthermore, while the fixed-monetary-policy rule is suboptimal in response to negative productivity shocks and thus costly in both economies, its underlying implications are radically different: a positive output gap under fixed entry, but a negative output gap in our free-entry economy. This is the key takeaway of our benchmark model.

Second-order Approximation

Note that we approximate around the steady state of the flex-price equilibrium (which is the same as for the sticky-price equilibrium) with $N^* = \frac{A\bar{L}}{f_E \theta}$ and $C^* = \left(N^*\right)^{\frac{1}{\theta-1}}\left(A\bar{L} - N^* f_E\right) = \left(N^*\right)^{\frac{1}{\theta-1}} \frac{\theta-1}{\theta} A\bar{L}$. A second-order approximation to utility around this steady state (which, by virtue of free entry, is efficient) delivers:

$$\hat{U}_t \equiv U\left(C_t, L_t\right) - U\left(C, L\right) \simeq U_C C \frac{C_t - C}{C} + U_L L \frac{L_t - L}{L} + \frac{1}{2} U_{CC} C^2 \left(\frac{C_t - C}{C}\right)^2 + U_{LL} L^2 \left(\frac{L_t - L}{L}\right)^2$$
$$= U_C C \left[c_t + \frac{1-\sigma^{-1}}{2} c_t^2\right] + U_L L \left[l_t + \frac{1+\varphi}{2} l_t^2\right] + t.i.p + O\left(\| \zeta \|^3\right),$$

where small letters denote log-deviations from steady state $c_t \equiv \log \frac{C_t}{C}$ and we used

$$\frac{C_t - C}{C} \simeq c_t + \frac{1}{2} c_t^2,$$

and same for L_t. Finally, $t.i.p$ are terms independent of policy and $O\left(\| \zeta \|^3\right)$ groups all terms of order 3 or higher.

Next, note that we focus here on the simple case of logarithmic utility in consumption $\sigma = 1$, implying that hours worked are always fixed in equilibrium (regardless of price stickiness and regardless of monetary policy). Therefore, the second term in the approximation drops out (this allows us to focus on the entry channel that is of the essence here). Furthermore, the term in squared consumption deviations c_t^2 also drops out since $\sigma = 1$, so we are left with the linear term—that we nevertheless need to approximate to *second* order. Taking a second-order approximation of the aggregate production function/resource constraint:

$$C_t = \left(1 - \frac{\kappa}{2}\pi_t^2\right) N_t^{\frac{1}{\theta-1}} \left(A_t L_t - N_t f_E\right)$$

around the steady-state using that hours are always constant in equilibrium and denoting by $\delta_t = -\ln\left(1 - \frac{\kappa}{2}\pi_t^2\right)$ the inflation welfare cost (which we then approximate to second order below):

$$\begin{aligned} C_t \simeq\ & C - N^{\frac{1}{\theta-1}} \left(L - N f_E\right) \delta_t + \left(\left(\frac{1}{\theta-1}\right) N^{\frac{1}{\theta-1}-1} \left(\bar{L} - N f_E\right) - N^{\frac{1}{\theta-1}} f_E\right) \left(N_t - N\right) \\ & + N^{\frac{1}{\theta-1}} L \left(A_t - A\right) + \frac{1}{2} \left(\begin{array}{c} \frac{1}{\theta-1}\left(\frac{1}{\theta-1} - 1\right) N^{\frac{1}{\theta-1}-2} \left(\bar{L} - N f_E\right) \\ -\frac{1}{\theta-1} N^{\frac{1}{\theta-1}-1} f_E - \frac{1}{\theta-1} N^{\frac{1}{\theta-1}-1} f_E \end{array} \right) \left(N_t - N\right)^2 \\ & + \frac{1}{\theta-1} N^{\frac{1}{\theta-1}-1} \bar{L} \left(N_t - N\right)\left(A_t - A\right) \end{aligned}$$

and writing with percentage deviations (recall $A = 1$ by normalization):

$$\begin{aligned} \frac{C_t - C}{C} =& -\delta_t + \frac{1}{\theta-1} \frac{N^{\frac{\theta}{\theta-1}}\left(\frac{\bar{L}}{N} - \theta f_E\right)}{C} \frac{N_t - N}{N} + \frac{N^{\frac{1}{\theta-1}} L}{C} \frac{A_t - A}{A} \\ & + \frac{1}{2}\frac{1}{\theta-1} \frac{N^{\frac{\theta}{\theta-1}}\left(\left(\frac{1}{\theta-1} - 1\right)\left(\frac{\bar{L}}{N} - f_E\right) - 2 f_E\right)}{C} \left(\frac{N_t - N}{N}\right)^2 + \frac{1}{\theta-1} \frac{N^{\frac{1}{\theta-1}} \bar{L}}{C} \frac{N_t - N}{N} \frac{A_t - A}{A} \end{aligned}$$

Use the steady-state, replacing $N = \frac{\bar{L}}{f_E \theta}$ and noticing that the 1st order term disappears:

$$\frac{C_t - C}{C} = -\delta_t + \frac{\theta}{\theta-1} \frac{A_t - A}{A} - \frac{1}{2} \frac{\theta}{(\theta-1)^2} \left(\frac{N_t - N}{N}\right)^2 + \frac{\theta}{(\theta-1)^2} \frac{N_t - N}{N} \frac{A_t - A}{A}$$

Finally, using the second-order approximation of the cost of inflation $\delta_t \simeq \frac{1}{2}\kappa\pi_t^2$ and the expression of the flexible-price number of firms $n_t^* = a_t$ and ignoring terms independent of policy and of order higher than 2, the loss function is proportional to

$$-\frac{1}{2}\left(\kappa\pi_t^2 + \frac{\theta}{(\theta-1)^2}\left(n_t - n_t^*\right)^2\right).$$

A.3 Nonlinear general model with both price and wage stickiness

Firms face nominal rigidity in the form of a quadratic cost of adjusting prices over time (Rotemberg, 1982). Specifically, the real cost (in units of the composite basket) of output-price inflation volatility around a steady-state level of inflation equal to 0 facing firm q is:

Each firm faces:

$$pac_t(q) = \frac{\kappa}{2}\left(\frac{p_t(q)}{p_{t-1}(q)} - 1\right)^2 \rho(q) y_t^D(q), \tag{A.7}$$

where $\rho_t(q) \equiv p_t(q)/P_t$ denotes the real (relative) price of firm q's output. This expression is interpreted as the amount of marketing materials that the firm must purchase when implementing a price change. We assume that this basket has the same composition as the consumption basket. The cost of adjusting prices is proportional to the real revenue from output sales, $(p_t(q)/P_t)\, y_t^D(q)$, where $y_t^D(q)$ is firm q's output demand.

Firms face demand for their output from consumers and firms themselves when they change prices. In each period, there is a mass N_t of firms producing and setting prices in the economy. [41]

The total demand for the output of firm q is thus

$$qy_t(q) = \left(\frac{p_t/q}{P_t}\right)^{-\theta} Y_t. \tag{A.8}$$

where $Y_t. = C_t + PAC_t = C_t + N_t pac_t(q)$

Then, the objective function of a firm choosing its price is equal to the firm q's real profit in period t (distributed to households as dividend) *net* of the price adjustment cost; this can be written as

$$d_t^R(q) = \rho_t(q)\, y_t^D(q) - w_t l_t(q) - \frac{\kappa}{2}\left(\frac{p_t(q)}{p_{t-1}(q)} - 1\right)^2 \rho_t(q)\, y_t^D(q).$$

At time t, firm q chooses $l_t(q)$ and $p_t(q)$ to maximize $d_t^R(q)$ subject to $y_t(q) = y_t^D(q)$, taking w_t, P_t, C_t, PAC_t, and A_t as given. Letting $\lambda_t(q)$ denote the Lagrange multiplier on the constraint $y_t(q) = y_t^D(q)$, the first-order condition with respect to $l_t(q)$ yields:

$$\lambda_t(q) = \frac{w_t}{qA_t}.$$

The shadow value of an extra unit of output is simply the firm's marginal cost, common across all

[41] When a new firm sets the price for the first time, we appeal to symmetry across firms and interpret the $t-1$ price in the expression of the price adjustment cost for that firm as the notional price that the firm would have set at time $t-1$ if it had been producing in that period. See Bilbiie et al (2007) for a justification of this assumption—which is consistent with the original Rotemberg (1982) setup—and a relaxation. Note that symmetry of the equilibrium will imply $p_{t-1}(q) = p_{t-1} \forall q$.

firms in the economy.

Maximizing the PDV of net profits (discounted wth the SDF, i.e. the marginal rate of intertemporal substitution of the household $\Lambda_{t,t+1}$) we obtain, imposing symmetry of the equilibrium $p_t(q) = p_t$ and denoting $\pi_t = \frac{p_t}{p_{t-1}} - 1$, after some substitutions the Phillips curve corresponding to the pricing decision is:

$$(1+\pi_t)\,\pi_t = \beta E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\sigma}}\frac{Y_{t+1}}{Y_t}\frac{N_t}{N_{t+1}}\left(1+\pi_{t+1}\right)\pi_{t+1}\right] + \frac{\theta}{\kappa}\left[\frac{1}{\mu_t} - \frac{\theta-1}{\theta}\left(1-\frac{\kappa}{2}\pi_t^2\right)\right].$$

Notice that by virtue of the assumption of heterogeneity in product quality, price stickiness does not interact with selection. Note that the total GDP of the economy, inclusive of the adjustment cost, is $Y_t = \left(1-\frac{\kappa}{2}\pi_t^2\right)^{-1} C_t$ and we used this in rewriting the Phillips curve.

The "wage stickiness" part is standard—wage setting is done by an union bundling the differentiated labor inputs of households, setting the nominal wage subject to adjustment costs—and its details are unaffected by the introduction of firm entry. For the sake of space, we do not report all derivations of this block but refer to Erceg et al (2000) and Schmitt-Grohé and Uribe (2006). We first provide the nonlinear equations (used to derive the loglinearized equations in text).The disutility of labor can now be written as $\frac{1}{1+\varphi}\left(\frac{L_t}{1-\frac{\kappa_w}{2}\pi_{w,t}^2}\right)^{1+\varphi}$ where we already replaced labor supply using labor market clearing $L_t/\left(1-\frac{\kappa_w}{2}\pi_{w,t}^2\right)$ where L_t is total labor demand and the denominator is related to the labor cost of adjusting nominal wages paid by the union. Denote the wage inflation rate by:

$$1+\pi_{w,t} = \frac{w_t}{w_{t-1}}\left(1+\pi_t\right)$$

The optimality condition for each union setting wages for a differentiated labor type subject to a downward sloping labor demand with elasticity θ_w, and a Rotemberg adjustment costs paid in labor units κ_w is:

$$\begin{aligned}\frac{\pi_{w,t}\left(\pi_{w,t}+1\right)}{1-\frac{\kappa_w}{2}\pi_{w,t}^2} &= \beta \mathrm{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\sigma}}\frac{L_{t+1}}{L_t}\frac{\pi_{w,t+1}\left(\pi_{w,t+1}+1\right)^2}{1-\frac{\kappa_w}{2}\pi_{w,t+1}^2}\right] \\ &+ \frac{\theta_w-1}{\kappa_w}\left[\frac{\theta_w}{\theta_w-1}\frac{1}{w_t C_t^{-\frac{1}{\sigma}}}\left(\frac{L_t}{1-\frac{\kappa_w}{2}\pi_{w,t}^2}\right)^{\varphi} - (1+s_w)\right],\end{aligned}$$

where $1+s_w$ is a labor subsidy set by a fiscal authority to eliminate both distortions in the goods and labor market.

The full model is described by Table A1. In the simulations, we set $\theta_w = \theta$, $\kappa_w = \kappa$ and an op-

Table A.1: Benchmark Model, Summary

1. Pricing	$\rho_t = \mu_t \frac{w_t}{A_t}$
2. Price PC	$(1+\pi_t)\,\pi_t = \beta E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\sigma}} \frac{Y_{t+1}}{Y_t} \frac{N_t}{N_{t+1}} (1+\pi_{t+1})\,\pi_{t+1} \right] + \frac{\theta}{\kappa} \left[\frac{1}{\mu_t} - \frac{\theta-1}{\theta} \left(1 - \frac{\kappa}{2}\pi_t^2\right) \right]$
3. Wage PC	$\frac{\pi_{w,t}(\pi_{w,t}+1)}{1-\frac{\kappa_w}{2}\pi_{w,t}^2} = \beta \mathrm{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\sigma}} \frac{L_{t+1}}{L_t} \frac{\pi_{w,t+1}(\pi_{w,t+1}+1)^2}{1-\frac{\kappa_w}{2}\pi_{w,t+1}^2} \right] + \frac{\theta_w - 1}{\kappa_w} \left[\frac{\theta_w}{\theta_w - 1} \frac{1}{w_t C_t^{-\frac{1}{\sigma}}} \left(\frac{L_t}{1-\frac{\kappa_w}{2}\pi_{w,t}^2} \right)^{\varphi} - (1+s_w) \right]$
4. Euler bonds	$C_t^{-\frac{1}{\sigma}} = \beta E_t \left(\frac{1+i_t}{1+\pi_{t+1}^C} C_{t+1}^{-\frac{1}{\sigma}} \right)$
5. Taylor rule	$1 + i_t = \beta^{-1} (1 + E_t \pi_{t+1})^{\phi} \exp(-\varepsilon_t)$
6. Wage inflation	$1 + \pi_{w,t} = \frac{w_t}{w_{t-1}} (1+\pi_t)$
7. Agg. acct.	$C_t + \frac{\kappa}{2}\pi_t^2 Y_t + N_{Et} w_t f_E = w_t L_t + N_t \bar{d}_t$
8. Profits	$\bar{d}_t = \left(1 - \frac{1}{\mu_t}\right) \frac{Y_t}{N_t} - \frac{w_t}{A_t} f_O$
9. Free entry	$w_t f_E = (1 - G(q_{0t}))\, \bar{d}_t = \left(\frac{q_{\min}}{q_{0t}} \right)^k \bar{d}_t$
10. CPI inflation	$\frac{1+\pi_t}{1+\pi_t^C} = \frac{\rho_t}{\rho_{t-1}}$
11. Variety effect	$\rho_t = \left(\frac{k}{k-(\theta-1)} \right)^{\frac{1}{\theta-1}} q_{0t} N_t^{\frac{1}{\theta-1}}$
12. Survivors	$N_t = (1 - G(q_{0t}))\, N_{Et} = \left(\frac{q_{\min}}{q_{0t}} \right)^k N_{Et}$
13. Exit Cutoff	$q_{0t} = \left(\frac{f_O}{\mu_t - 1} \frac{1}{Y_t} \right)^{\frac{1}{\theta-1}} \rho_t^{\frac{\theta}{\theta-1}}$
14. Output	$C_t = \left(1 - \frac{\kappa}{2}\pi_t^2\right) Y_t$

Notes: The table displays all the nonlinear equilibrium conditions.

timal subsidy eliminating the steady-state inefficiency stemming from both market power sources $1 + s_w = \frac{\theta}{\theta-1}\frac{\theta_w}{\theta_w-1}$.

Note that the price adjustment cost is second order and it does not appear in the log-linearized version around a zero-inflation steady state.

Finally, aggregate accounting, free entry and the profits definition imply that the "aggregate production function" takes the form:

$$Y_t = A_t \rho_t \left(L_t - N_{Et} f_E - N_t \frac{f_O}{A_t} \right);$$

in other words, labor used in production equals total labor minus labor used to pay for the entry fost (for all entrants) and labor used to pay for the fixed overhead cost in efficiency units (for all producting firms)—i.e., the labor market clears by Walras' law.

A.4 Baseline vs Analytical Calibration

In Figure A.1 we repeat the experiment of Figure 1, comparing the baseline calibration $\beta = .99$ with the static-PCs calibration used in the closed-form analytical solution $\beta = 0$. As the figure makes clear, the responses are barely distinguishable for all real variables; differences naturally appear for inflation (and hence the real rate), since such a sharp change in β essentially amounts to a recalibration of the long-run Phillips curve's slope.

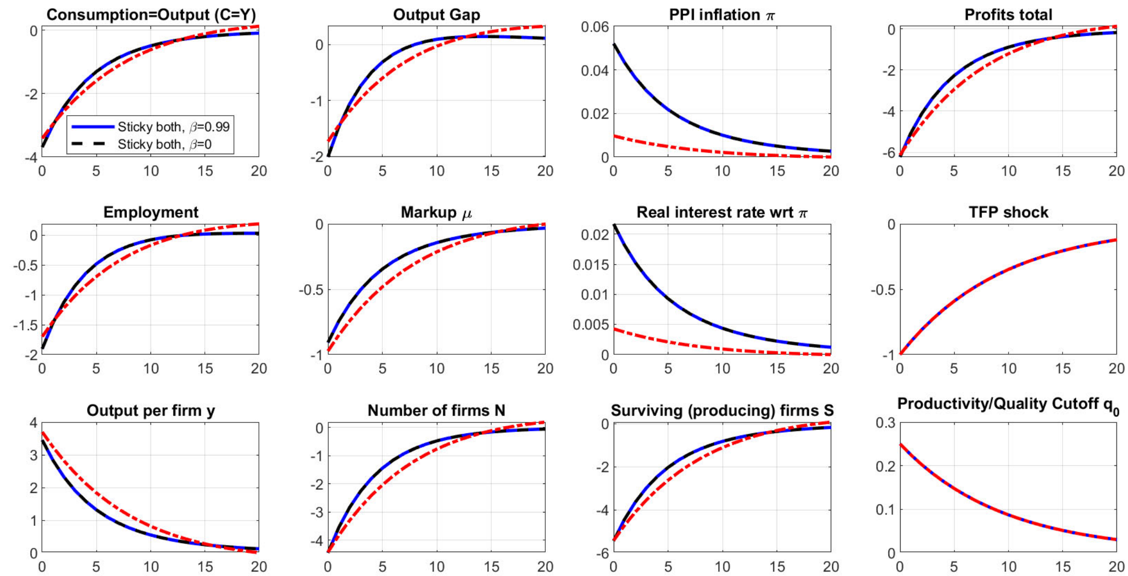


Figure A.1: Effects of a 1% productivity fall: $\beta = 0$ (red dash) vs $\beta = 0.99$ (solid blue).

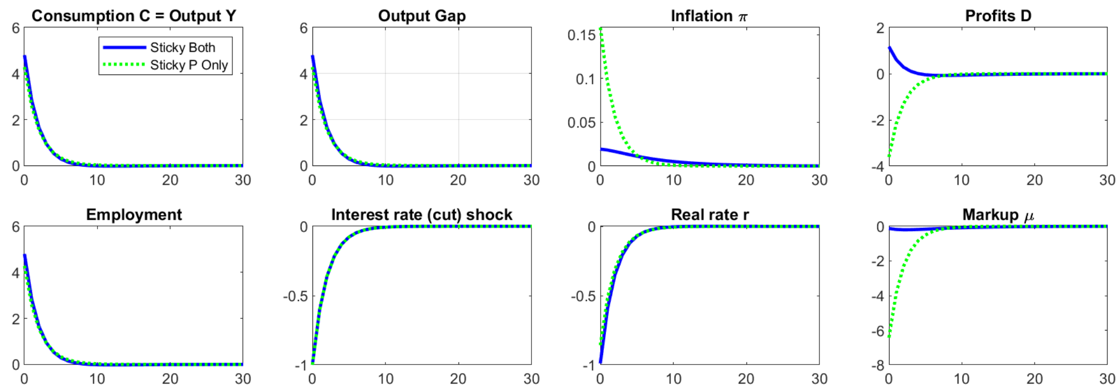
A.5 Demand shocks: with vs without entry and sticky wages

In Figure A.2, we report the impulse responses to a 1% interest rate cut for the no-entry NK (panel a) and our entry (panel) models, respectivly—without sticky wages (only sticky prices) in dashed green, and with sticky both prices and wages in solid blue.

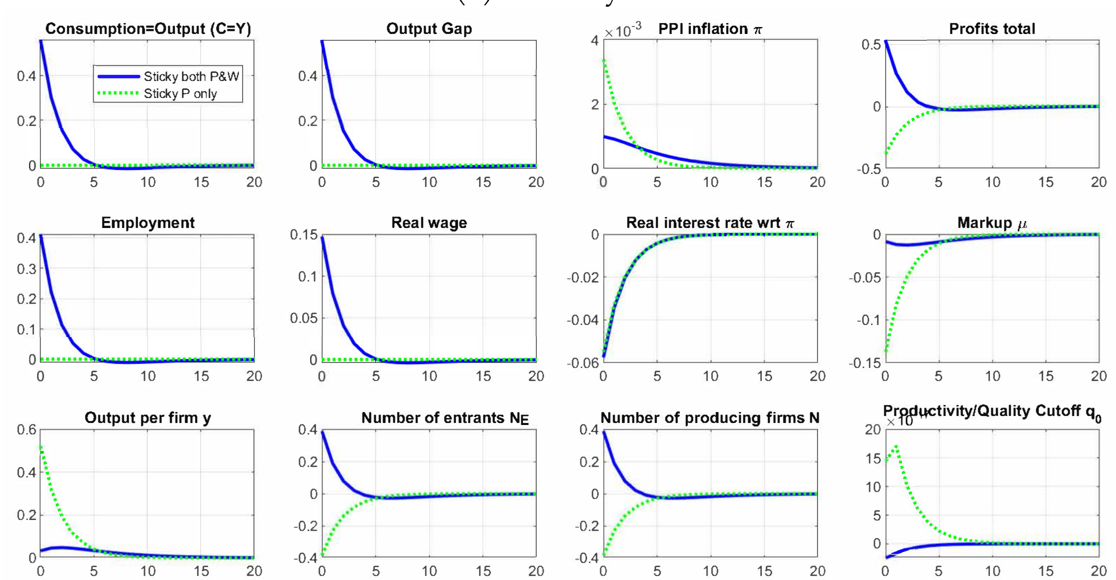
In our model with entry (panel b) but sticky prices only, we see how the demand shock induces exit. As we previously discussed (and shown analytically), this counterfactual prediction is driven by the sharp drop in the markup, and hence profits (per firm).

[42] In the no-entry NK model (upper panel a), the countercyclicality of markups in response to demand shocks implies that profits are countercyclical too under flexible wages. This is a well-known issue of sticky-price (only) New Keynesian models, for which wage stickiness is an equally well-known solution making profits procyclical; we see in panel a that the markup and profits are countercyclical in the sticky-price-only case, but the markup gets dampened and profits become procyclical under wage stickiness.

[42]The cylical properties of profits have nontrivial (and sometimes perverse) aggregate-demand implications throgh distributional mechanisms in models with heterogeneous agents, see e.g. Bilbiie (2008, 2020). Wage stickiness helps alleviate some of those issues, as discussed recently e.g. by Broer et al (2020).



(a) No Entry NK



(b) Endogenous Entry NK

Figure A.2: Effects of a 1% rate cut: Sticky P only (green dots) vs Sticky P & W (solid blue)

B Standard NK Model (Exogenous Entry) with Sticky Prices and Wages

The first 8 equations in our Table 1 are exactly the same as in the standard, no-entry NK model, modulo the right modifications: in particular, the first equation (4) would have $\hat{\rho}_t = 0$ (no benefit of variety), implying simply that the markup is the inverse of real marginal cost; in the last equation, CPI inflation would this be the same as producer inflation ($\pi_t = \pi_t^C$). In the second equation (6), aggregate accounting and the CRS production function would simply imply $\hat{Y}_t = \hat{A}_t + \hat{L}_t$.[43] The third equation (17) would simply be the definition of profits (with no change in entry,$\hat{N}_t = 0$), which are irrelevant for the equilibrium outcome.

We now show (mostly review) how, under DRS $\hat{Y}_t = \hat{A}_t + \alpha \hat{L}_t$, and with the same fixed real rate rule used in text: 1. the output gap goes up with negative TFP; 2. it can be closed by tracking r-star, which entails a monetary contraction; and 3. this tracking r-star policy is equivalent to completely stabilizing the "average" composite inflation measure (this last point is to the best of our knowledge novel). The natural interest rate is:

$$r_t^{*\mathcal{N}} = (1 - \rho_a)\, \Gamma^{\mathcal{N}} \hat{A}_t$$

where

$$\Gamma^{\mathcal{N}} = -\frac{1 + \frac{\varphi + \alpha}{1-\alpha}}{1 + \sigma \frac{\varphi+\alpha}{1-\alpha}}$$

so it co-moves negatively to TFP. Note that the response of hours in the flexible equilibrium is $\sigma^{-1} \hat{Y}_t^* + \varphi \hat{L}_t^* = \hat{w}_t^* = \hat{A}_t + (\alpha - 1)\, \hat{L}_t^*$ so $\hat{L}_t^* = \left(1 - \sigma^{-1}\right) \frac{1}{\alpha \sigma^{-1} + 1 - \alpha + \varphi} \hat{A}_t$, which shows that $\sigma| > 1$ is the restriction needed for hours to respond positively to TFP, with the same intuition as before (income effects dominated by substitution effects). Writing in gaps (ignore superscript $\mathcal{N}$ when obvious to lighten notation)

Take difference assume $\phi = 1$

$$\pi_{w,t} - \pi_t = \beta \left(E_t \pi_{w,t+1} - E_t \pi_{t+1}\right) + \psi_w \left(\left(\sigma^{-1} + \frac{\varphi}{1-\alpha} \right) \tilde{Y}_t - \hat{w}_t \right) - \psi \left(\hat{w}_t + \frac{\alpha}{1-\alpha} \tilde{Y}_t \right)$$

Demand side fully determines output gap in this version of the standard NK model:

$$\tilde{Y}_t^{\mathcal{N}} = E_t \tilde{Y}_{t+1}^{\mathcal{N}} + \sigma \left(\varepsilon_t + r_t^{*\mathcal{N}} \right) \rightarrow \tilde{Y}_t^{\mathcal{N}} = \sigma \left(\epsilon_t + \Gamma^{\mathcal{N}} \hat{A}_t \right)$$

[43] Or equivalently from the income side: $\hat{Y}_t = \left(1 - \theta^{-1}\right) \left(\hat{w}_t + \hat{L}_t \right) + \theta^{-1} \hat{d}_t$, with $\hat{d}_t = \hat{Y}_t + (\theta - 1)\, \hat{\mu}_t$.

Table B.1: Loglinearized NK Model, in Gaps

Common equations:	
1. Pricing	$-\hat{\mu}_t = \tilde{w}_t + \frac{\alpha}{1-\alpha}\tilde{Y}_t$
2. Price PC	$\pi_t = \beta E_t \pi_{t+1} + \psi\left(\tilde{w}_t + \frac{\alpha}{1-\alpha}\tilde{Y}_t\right)$
3. Wage PC	$\pi_{w,t} = \beta E_t \pi_{w,t+1} + \psi_w\left(\left(\sigma^{-1} + \frac{\varphi}{1-\alpha}\right)\tilde{Y}_t - \tilde{w}_t\right)$
4. Euler bonds	$\tilde{Y}_t = E_t \tilde{Y}_{t+1} - \sigma\left(\hat{i}_t - E_t \pi_{t+1} - \hat{r}^*_t\right)$
5. Taylor rule	$\hat{i}_t = \phi E_t \pi_{t+1} - \varepsilon_t$
6. Wage inflation	$\tilde{w}_t = \tilde{w}_{t-1} + \pi_{w,t} - \pi_t - \left(w^*_t - w^*_{t-1}\right)$

Notes: The table displays all the loglinearized equilibrium conditions of the no-entry NK model, in "gap" notation, around a zero-inflation steady state with an optimal subsidy.

The Phillips curve and wage equation deliver:

$$\tilde{w}_t - \tilde{w}_{t-1} + \left(w^*_t - w^*_{t-1}\right) = \beta\left(E_t\tilde{w}_{t+1} - \tilde{w}_t + \left(w^*_{t+1} - w^*_t\right)\right) + \left(\psi_w\left(\sigma^{-1} + \frac{\varphi}{1-\alpha}\right) - \psi\frac{\alpha}{1-\alpha}\right)\tilde{Y}_t - \left(\psi_w + \psi\right)\tilde{w}_t$$

$$\beta E_t\tilde{w}_{t+1} - \left(1 + \beta + \psi_w + \psi\right)\tilde{w}_t + \tilde{w}_{t-1} = \left(\psi\frac{\alpha}{1-\alpha} - \psi_w\left(\sigma^{-1} + \frac{\varphi}{1-\alpha}\right)\right)\tilde{Y}_t + \left(w^*_t - w^*_{t-1}\right) - \beta\left(E_t w^*_{t+1} - w^*_t\right)$$

Use static PC

$$\left(1 + \psi_w + \psi\right)\tilde{w}_t = \tilde{w}_{t-1} + \left(\psi_w\left(\sigma^{-1} + \frac{\varphi}{1-\alpha}\right) - \psi\frac{\alpha}{1-\alpha}\right)\tilde{Y}_t - \left(w^*_t - w^*_{t-1}\right)$$

This determines the wage gap, but output gap already given by demand side, given r-star. When TFP falls, r-star goes up, and the output gap goes up. "Cure" this by contracting:

$$d\tilde{Y}^{\mathcal{N}}_t = 0 \rightarrow d\epsilon^{*\mathcal{N}}_t = \Gamma^{\mathcal{N}} d\left(-\hat{A}_t\right),$$

increase interest rates (negative ϵ) in response to a negative TFP shock because it engenders an output-gap expansion.

Finally, note that this "tracking r-star" policy is exactly equivalent to a policy stabilizing the average inflation measure defined in text, since in this model modified divine coincidence holds (closing the output gap stabilizes average inflation). It can be easily shown by aggregating the two

PCs (and it is a textbook result known from Woodford 2003 and Gali 2015) that in this model:

$$\bar{\pi}_t = \beta E_t \bar{\pi}_{t+1} + \bar{\psi}\left(\sigma^{-1} + \frac{\varphi + \alpha}{1-\alpha}\right)\tilde{Y}_t.$$

As we saw, all these policy are very different in our endogenous-entry model.

C Welfare and inflation implications of easing and gap-closing policy

C.1 Proof that monetary easing increases utility in low-TFP state

Suppose we start from a realization of TFP under steady state, $\bar{A} < 1$ which leads to a recession under fixed prices and wages. Under the calibration used in the welfare figure, the solution for output (consumption) derived in 31 becomes, denoting $M_t = \exp(\epsilon_t)$ for ease of interpretation:

$$\begin{aligned}
Y_t &= \left(\theta \bar{A} - (\theta - 1)\right)^{\frac{1}{\frac{\theta}{\sigma}-1}} \bar{A}^{-\frac{1}{k}\frac{\theta-1}{\frac{\theta}{\sigma}-1}} \left(\frac{\theta}{\theta-1} M_t\right)^{\frac{\frac{\theta}{\sigma}}{\frac{\theta}{\sigma}-1}} \\
L_t &= M_t \frac{\theta}{\theta-1} Y_t^{1-\frac{1}{\sigma}} = \left(\theta \bar{A} - (\theta-1)\right)^{\frac{\sigma-1}{\theta-\sigma}} \bar{A}^{-\frac{\theta-1}{k}\frac{\sigma-1}{\theta-\sigma}} \left(\frac{\theta}{\theta-1} M_t\right)^{\sigma\frac{\theta-1}{\theta-\sigma}}
\end{aligned}$$

where the second equation is the solution for hours worked.

The change in utility in response to a monetary expansion will have two terms, one from utility of consumption the second from disutility of labor. The first term of $dU\left(.\right)$is:

$$\begin{aligned}
Y_t^{-\frac{1}{\sigma}} \frac{dY_t}{dM_t} &= Y_t^{-\frac{1}{\sigma}} \left(\theta \bar{A} - (\theta-1)\right)^{\frac{1}{\frac{\theta}{\sigma}-1}} \bar{A}^{-\frac{1}{k}\frac{\theta-1}{\frac{\theta}{\sigma}-1}} \left(\frac{\theta}{\theta-1}\right)^{\sigma\frac{\theta}{\theta-\sigma}} \sigma \frac{\theta}{\theta-\sigma} \left(M_t\right)^{\sigma\frac{\theta}{\theta-\sigma}-1} \\
&= Y_t^{1-\frac{1}{\sigma}} \sigma \frac{\theta}{\theta-\sigma} \frac{1}{M_t}
\end{aligned}$$

The second term is

$$\begin{aligned}
-L_t^{\varphi} \frac{dL_t}{dM_t} &= -L_t^{\varphi} \left(\theta \bar{A} - (\theta-1)\right)^{\frac{\sigma-1}{\theta-\sigma}} \bar{A}^{-\frac{\theta-1}{k}\frac{\sigma-1}{\theta-\sigma}} \left(\frac{\theta}{\theta-1}\right)^{\sigma\frac{\theta-1}{\theta-\sigma}} \sigma \frac{\theta-1}{\theta-\sigma} \left(M_t\right)^{\sigma\frac{\theta-1}{\theta-\sigma}-1} \\
&= -L_t^{1+\varphi} \sigma \frac{\theta-1}{\theta-\sigma} \frac{1}{M_t}
\end{aligned}$$

This needs to be evaluated at steady-state M:

$$M = \left(\frac{\theta}{\theta-1}\right)^{\frac{\frac{\theta}{\sigma}-1}{\theta-1}\frac{1}{1+\varphi-\left(1-\frac{1}{\sigma}\right)\frac{\theta}{\theta-1}}-1}$$

Collecting terms

$$\begin{aligned}\frac{dU_t}{dM_t} &= Y_t^{1-\frac{1}{\sigma}}\sigma\frac{\theta}{\theta-\sigma}\frac{1}{M_t} - L_t^{1+\varphi}\sigma\frac{\theta-1}{\theta-\sigma}\frac{1}{M_t} \\ &= \sigma\frac{\theta-1}{\theta-\sigma}\frac{1}{M_t}\left(\frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}} - L_t^{1+\varphi}\right)\end{aligned}$$

Evaluate sign of paranthetic term, show it is larger than 0 (derivative of utility is then positive under the parameter condition (2)):

$$\begin{aligned}\frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}} - L_t^{1+\varphi} &= \frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}} - \left(M_t\frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}}\right)^{1+\varphi} \\ &= \frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}}\left(1-(M_t)^{1+\varphi}\left(\frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}}\right)^{\varphi}\right)\end{aligned}$$

Show

$$(M_t)^{1+\varphi}\left(\frac{\theta}{\theta-1}Y_t^{1-\frac{1}{\sigma}}\right)^{\varphi} < 1$$

Use

$$Y_t^{1-\frac{1}{\sigma}} = \left(\theta\bar{A}-(\theta-1)\right)^{\frac{\sigma-1}{\theta-\sigma}}\bar{A}^{-\frac{1}{k}\frac{(\theta-1)(\sigma-1)}{\theta-\sigma}}\left(\frac{\theta}{\theta-1}M_t\right)^{\frac{\theta(\sigma-1)}{\theta-\sigma}}$$

$$(M_t)^{1+\varphi}\left(\frac{\theta}{\theta-1}\left(\theta\bar{A}-(\theta-1)\right)^{\frac{\sigma-1}{\theta-\sigma}}\bar{A}^{-\frac{1}{k}\frac{(\theta-1)(\sigma-1)}{\theta-\sigma}}\left(\frac{\theta}{\theta-1}M_t\right)^{\frac{\theta(\sigma-1)}{\theta-\sigma}}\right)^{\varphi} < 1$$

$$(M_t)^{1+\varphi\frac{(\theta-1)\sigma}{\theta-\sigma}}\left(\frac{\theta}{\theta-1}\right)^{\varphi\frac{(\theta-1)\sigma}{\theta-\sigma}}\left(\theta\bar{A}-(\theta-1)\right)^{\varphi\frac{\sigma-1}{\theta-\sigma}}\bar{A}^{-\frac{1}{k}\frac{(\theta-1)(\sigma-1)}{\theta-\sigma}\varphi} < 1$$

$$M^{SS} = \left(\frac{\theta}{\theta-1}\right)^{\frac{\frac{\theta}{\sigma}-1}{\theta-1}\frac{1}{1+\varphi-\left(1-\frac{1}{\sigma}\right)\frac{\theta}{\theta-1}}-1}$$

$$\left(\left(\frac{\theta}{\theta-1}\right)^{\frac{\frac{\theta}{\sigma}-1}{\theta-1}\frac{1+\varphi\frac{(\theta-1)\sigma}{\theta-\sigma}}{1+\varphi-\left(1-\frac{1}{\sigma}\right)\frac{\theta}{\theta-1}}-1-\varphi\frac{(\theta-1)\sigma}{\theta-\sigma}+\varphi\frac{(\theta-1)\sigma}{\theta-\sigma}}\right)\left(\theta\bar{A}-(\theta-1)\right)^{\varphi\frac{\sigma-1}{\theta-\sigma}}\bar{A}^{-\frac{1}{k}\frac{(\theta-1)(\sigma-1)}{\theta-\sigma}\varphi} < 1$$

$$\left(\left(\frac{\theta}{\theta-1}\right)^{\frac{\frac{\theta}{\sigma}-1}{\theta-1}\frac{1+\varphi\frac{(\theta-1)\sigma}{\theta-\sigma}}{1+\varphi-\left(1-\frac{1}{\sigma}\right)\frac{\theta}{\theta-1}}-1}\right)\left(\theta\bar{A}-(\theta-1)\right)^{\varphi\frac{\sigma-1}{\theta-\sigma}}\bar{A}^{-\frac{1}{k}\frac{(\theta-1)(\sigma-1)}{\theta-\sigma}\varphi} < 1$$

$$\frac{\frac{\frac{\theta}{\sigma}-1}{\theta-1}-1+\frac{\theta\left(1-\frac{1}{\sigma}\right)}{\theta-1}}{1+\varphi-\left(1-\frac{1}{\sigma}\right)\frac{\theta}{\theta-1}}$$

Simplifying, the term in the first bracket is in fact 1 (the exponent is 0), so it becomes:

$$\left(\theta\bar{A}-(\theta-1)\right)^{\varphi\frac{\sigma-1}{\theta-\sigma}}\ \bar{A}^{-\frac{1}{k}\frac{(\theta-1)(\sigma-1)}{\theta-\sigma}\varphi}<1$$

$$\left(\left(\theta\bar{A}-(\theta-1)\right)\bar{A}^{-\frac{\theta-1}{k}}\right)^{\varphi\frac{\sigma-1}{\theta-\sigma}}<1$$

As long as the power is positive (and, obviously, so is the term in brackets), it is therefore sufficient to show that $\left(\theta\bar{A}-(\theta-1)\right)\bar{A}^{-\frac{\theta-1}{k}}<1$. Since $\theta>1$, $k>\theta-1$ and, crucially, $\bar{A}<1$ it can be shown by standard calculus that this inequality always holds. (Define $J\left(\bar{A}\right)=\bar{A}^{\frac{\theta-1}{k}}-\left(\theta\bar{A}-(\theta-1)\right)$. For the bracketed term to be positive, the relevant domain is $\bar{A}\in(\frac{\theta-1}{\theta},1]$. Over this domain, J is globally concave since $J''(\bar{A})=\frac{\theta-1}{k}(\frac{\theta-1}{k}-1)\bar{A}^{\frac{\theta-1}{k}-2}<0$. Because the endpoints satisfy $J(1)=0$ and $J(\frac{\theta-1}{\theta})>0$, global concavity guarantees $J(\bar{A})>0$ strictly inside the interval.)

The exact same derivation—with opposite sign—shows that a monetary contraction improves welfare with $\bar{A}>1$, i.e. in a TFP-driven expansion.

In the No-Entry NK model, the opposite holds. We can prove this using the expressions from Section 2 (for the log-utility case, but easily generalizable), reproducing them here $Y_t^{\mathcal{N}}=\frac{M_t}{P^\circ}$ and $L_t^{\mathcal{N}}=\frac{1}{A_t}\frac{M_t}{P^\circ}$ and evaluating utility starting from steady-state money balances $\frac{M}{P^\circ}=1$; The change in utility starting at $\bar{A}$ is

$$\frac{dU_t^{\mathcal{N}}}{dM_t}=Y_t^{-\frac{1}{\sigma}}\frac{1}{P^\circ}-L_t^{\varphi}\frac{1}{P^\circ}\frac{1}{A_t}=\left(\frac{M_t}{P^\circ}\right)^{-\frac{1}{\sigma}}\frac{1}{P^\circ}-\left(\frac{M_t}{P^\circ}\frac{1}{\bar{A}}\right)^{\varphi}\frac{1}{P^\circ}\frac{1}{\bar{A}}=1-\left(\frac{1}{\bar{A}}\right)^{1+\varphi}<0 \text{ when } \bar{A}<1$$

Conversely, utility increases with money when starting in a high-TFP (recession-inducing, in the no-entry model) state.

As we recall from the left panel of Figure 1 and the discussion therein, a low TFP realization triggers an expansion in that model, which is obtained by working more hours. Since an "envelope" result applies to that model too for welfare, the effect on utility is unambiguously negative; but since a monetary expansion increases output even further (with a further increase in hours), it deteriorates welfare—this is why a contraction is in fact needed in the no-entry NK model.

C.2 Gap-closing expansionary monetary policy is not necessarily inflationary

In Figure C.1, we report the impulse responses to a 1% TFP fall. We first reproduce the two cases in the benchmark panel b of Figure 2, and add a third line, corresponding to a policy that—instead of following a Taylor rule—acts as though to close the output gap, which in our model is the same (to first order) as stabilizing wage inflation. This is represented with black dash-crosses. The main points are that this is indeed an easing (the real rate goes down) but it does not translate into much higher inflation because of the mechanism emphasized in text (the smaller fall in entry visible in the respective panel keeps inflation in check). Nevertheless, an expansion larger than this one *would* imply higher inflation (in both prices and wages).

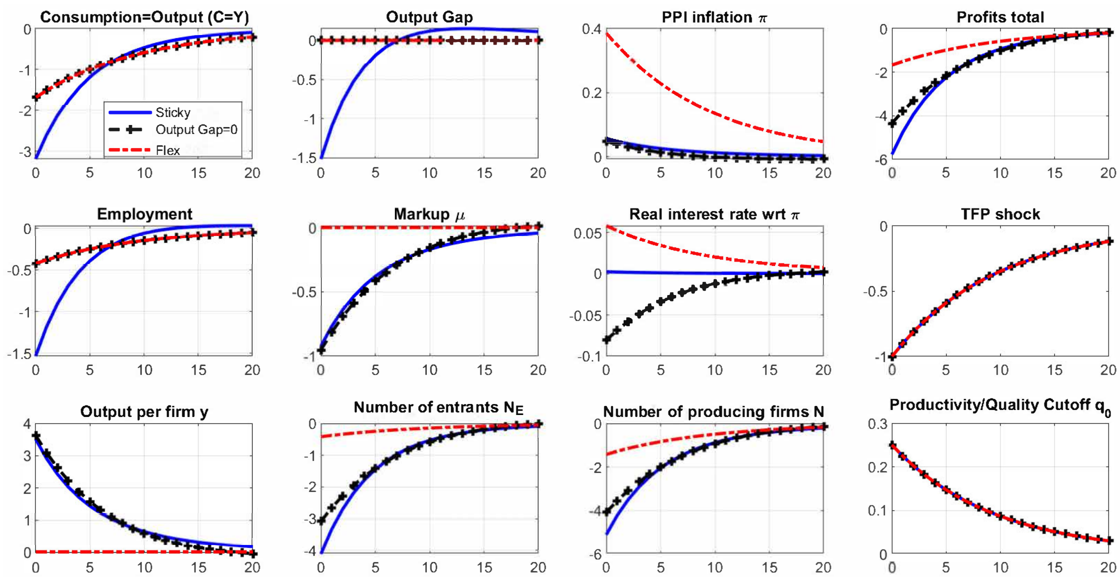


Figure C.1: Effects of a 1% TFP fall: flexible (red dash) vs sticky price and wage (solid blue). Black dash-crosses: gap-closing policy (zero wage inflation).