

NBER WORKING PAPER SERIES

ANATOMY OF THE PHILLIPS CURVE:
MICRO EVIDENCE AND MACRO IMPLICATIONS

Luca Gagliardone
Mark Gertler
Simone Lenzu
Joris Tielens

Working Paper 31382
<http://www.nber.org/papers/w31382>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2023, Revised May 2025

We thank Fernando Alvarez, Ester Faia, Jordi Gali, Xavier Gabaix, Jonathon Hazell, Man Chon Iao, Virgiliu Midrigan, Alessio Moro, Carolin Pflueger, David Rivers, Luminita Stevens, Ludwig Straub, Gianluca Violante, Olivier Wang, Ivan Werning, Raf Wouters, and seminar participants at various institutions for helpful comments and conversations and Siena Matsumoto for excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the National Bank of Belgium, the Eurosystem, any other institution with which the authors are affiliated, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Luca Gagliardone, Mark Gertler, Simone Lenzu, and Joris Tielens. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Anatomy of the Phillips Curve: Micro Evidence and Macro Implications
Luca Gagliardone, Mark Gertler, Simone Lenzu, and Joris Tielens
NBER Working Paper No. 31382
June 2023, Revised May 2025
JEL No. E31

ABSTRACT

We develop a bottom-up approach to estimate the slope of the primitive form of the New Keynesian Phillips curve, which features marginal cost as the real activity variable. Using quarterly micro data on prices, costs, and output, we estimate dynamic pass-through regressions that identify the slope as a function of primitive parameters. We find a high slope for the cost-based Phillips curve, which contrasts with the low estimates of the conventional output gap-based formulation found in the literature. We reconcile by showing that the output elasticity of marginal cost is low, at least during moderate inflation periods (e.g., pre-pandemic).

Luca Gagliardone
New York University
luca.gagliardone@nyu.edu

Simone Lenzu
New York University
slenzu@stern.nyu.edu

Mark Gertler
New York University
Department of Economics
and NBER
mark.gertler@nyu.edu

Joris Tielens
National Bank of Belgium
joris.tielens@nbb.be

Understanding the relation between inflation and real activity over the business cycle continues to be an important though unresolved matter in macroeconomics. At the heart of this inquiry lies the challenge of estimating the slope of the Phillips curve. To illustrate the issue, let us consider the New Keynesian version of the Phillips curve (NKPC), which is now the textbook formulation in the literature. Let π_t denote inflation and \tilde{y}_t the output gap, the percentage difference between real output and its natural level. Then (what we will refer to as) the *conventional formulation* of the NKPC is given by:

$$\pi_t = \kappa \tilde{y}_t + \beta \mathbb{E}_t \{ \pi_{t+1} \} + u_t, \quad (1)$$

where u_t is typically referred to as a cost-push shock, and β is a subjective discount factor, typically a parameter close to unity. The NKPC asserts that inflation depends positively on both \tilde{y}_t , which is interpreted as a measure of excess demand, and on expected future inflation. The main object of interest is κ , the slope coefficient on the output gap.

There are two interrelated sets of issues involved in uncovering κ . The first set revolves around the econometric identification of this parameter. First, as emphasized by McLeay and Tenreyro (2020), the output gap is an endogenous object. If the central bank acts to adjust \tilde{y}_t to stabilize π_t in response to positive cost-push shocks, the estimate of κ will be biased downward due to the negative correlation between \tilde{y}_t and u_t . Given the absence of good instruments for \tilde{y}_t , the estimation of κ using aggregate time-series data is challenging (Mavroeidis, Plagborg-Møller and Stock 2014) or requires additional assumptions.¹ Another identification issue involves trend inflation. The specification given by Equation (1) presumes that trend inflation is constant. However, as argued by Hazell et al. (2022) and Jørgensen and Lansing (2023), shifts in trend inflation may confound the identification of the Phillips curve. For instance, if trend inflation decreases as output declines, and the regression model does not account for this correlation, the estimate of κ will be upwardly biased.

These identification challenges have led researchers to employ regional data to estimate κ . Recent examples include Hooper, Mishkin and Sufi (2020), McLeay

¹Barnichon and Mesters (2020) and Lewis and Mertens (2022) show how to identify the slope of the Phillips curve using aggregate data when valid and powerful time-series instruments are available. In a similar vein, Section 7 presents a validation exercise using identified oil shocks. We note, however, that this approach cannot identify the underlying primitive parameters that determine the slope.

and Tenreyro (2020), and Hazell et al. (2022).² Importantly, Hooper, Mishkin and Sufi (2020) and Hazell et al. (2022) allow for time fixed effects to control for shifting trend inflation. In the latter study, this identification approach yields an astonishingly small estimate of κ , which suggests that the Phillips curve is “flat.” This view has become the conventional wisdom, at least for the pre-pandemic period.

The second set of considerations pertains to both the relevant measure of real activity that enters the Phillips curve and, consequently, the interpretation of the slope coefficient κ . In the underlying theory, firms set prices in response to current and anticipated movements in marginal cost. Thus, as emphasized by both Galí and Gertler (1999) and Sbordone (2002), the *primitive formulation* of the NKPC features real marginal cost (in percent deviations from trend) entering as the real activity variable. In fact, the conventional formulation of the NKPC in Equation (1) only holds under specific conditions that establish a proportional relationship between marginal cost and the output gap. Among other things, wages must be perfectly flexible.³ If these conditions are violated, then the output gap may not serve as an adequate proxy for real marginal cost, typically leading to a downward bias in the estimate of κ .⁴ Moreover, even if all conditions that establish a proportional relationship are approximately met, it is crucial to recognize that the output-based slope κ is ultimately the product of two parameters: the elasticity of inflation with respect to real marginal cost and the elasticity of marginal cost with respect to the output gap. The ability to separately identify the two coefficients is important for gaining a comprehensive understanding of inflation dynamics.

In this paper, we propose a novel empirical strategy to estimate the slope of the primitive formulation of the NKPC. The conventional estimation approach involves aggregating individual firm pricing decisions into an NKPC and then estimating its slope with aggregate data. Instead, we follow a bottom-up approach. We use micro data to estimate dynamic pass-through regressions that identify both the degree of nominal and real rigidities from short-run comovements in firm-level marginal costs

²Also relevant is Beraja, Hurst and Ospina (2019), which uses regional data to identify wage Phillips curves.

³For this reason, New Keynesian DSGE models with wage rigidity include the marginal cost-based Phillips curve in the system of equations as opposed to the conventional one (see Galí 2015 chapter 6 and the references therein).

⁴These considerations also extend to formulations of the conventional NKPC that utilize the unemployment gap as a measure of economic activity instead of the output gap.

and prices. We then use these estimates to recover the NKPC slope and compute the implied aggregate pass-through.

In Section 1, we develop a theoretical framework that serves as the foundation of our estimation strategy. Starting from first principles, we derive an expression for firms' optimal reset prices in an environment with nominal rigidities and imperfect competition. As is standard, firms' optimal reset price depend on the expected path of marginal cost over the period the firm expects its price to be fixed. Moreover, due to the presence of strategic complementarities, firms factor in the expected path of competitors' prices, which reduces the pass-through of marginal cost. The slope of the Phillips curve is then a function of the two parameters capturing the degree of nominal price rigidities and the strength of strategic complementarities in price setting.

We estimate these structural parameters using micro data, as described in Section 2. We collect administrative data on product-level output prices, quantities, and production costs for manufacturing firms in Belgium, which we use to construct granular proxies of firms' marginal costs and competitors' prices. Our data extends the database originally assembled by Amiti, Itskhoki and Konings (2019) in terms of cross-sectional and time-series coverage. Notably, our data is recorded at the quarterly (as opposed to annual frequency), which allows us to study the role of nominal rigidities in price setting at the business cycle frequency.

In Section 3 we map the theoretical model to the data to derive dynamic pass-through regressions that identify the structural parameters of interest. The use of micro data allows us to tackle the issues that hinder identification using aggregate data.⁵ By including in our model a set of fixed effects, we address unobserved heterogeneity and confounding factors stemming from trends in output growth, trend inflation, and shifts in inflation expectations. In addition, we can construct powerful instruments for marginal cost and competitors' prices to tackle endogeneity and measurement issues. Our approach relates to the literature on incomplete pass-through of marginal cost into prices (Goldberg and Verboven 2001; Nakamura

⁵Galí and Gertler (1999) originally estimated a marginal cost-based NKPC using aggregate data with the labor share as the measure of marginal cost. The use of micro data improves upon the identification as it addresses weak instruments concerns raised by Mavroeidis, Plagborg-Møller and Stock (2014) and allows us to deal with trends in costs and prices. Moreover, the micro data also provide us with a richer measure of marginal costs that accounts for intermediate input costs along with labor costs.

and Zerom 2010). In an environment with perfectly flexible prices, our dynamic pass-through framework nests as a special case the static pass-through regressions estimated in Amiti, Itskhoki and Konings (2019) using annual data.

In section 4 , we show that our analysis delivers sensible and robust estimates of the parameters governing firms' pricing behavior. We find a substantial degree of nominal rigidities (prices are fixed three to four quarters, on average) and a meaningful role for strategic complementarities (reducing the pass-through of marginal cost shocks by about half). These estimates imply an economically significant slope of the marginal cost-based NKPC, tightly estimated in the range of 0.05 to 0.07. They also imply a substantial aggregate pass-through from marginal cost to inflation. To show this, in Section 5 we construct a proxy for aggregate marginal cost, which we feed into our model. The model-implied inflation series tracks the actual PPI data well. Fluctuations in marginal cost alone can account for at least seventy percent of the variation in inflation, without appealing, as is often done, to unobservable cost-push shocks or including lags of inflation.

Our estimate of the slope of the cost-based NKPC differs markedly from existing estimates of the conventional output- or unemployment-based NKPC, which is estimated to be two to ten times smaller in magnitude (Rotemberg and Woodford 1997, Hazell et al. 2022). In Section 6, we show that these estimates are not inconsistent and can be reconciled with ours. We make standard assumptions that allow us to derive the output-based Phillips curve slope as the product of our marginal cost-based slope and the output elasticity of marginal cost. We then develop identification strategies to estimate this elasticity using micro data and retrieve the implied slope of the output-based NKPC. For our pre-pandemic sample, we find a low elasticity of marginal cost to changes in output, yielding point estimates of the output-based NKPC slope consistent with the literature. This suggests that the flat slope of the conventional NKPC reflects a weak link between the output gap and marginal cost over this sample period rather than a limited transmission of marginal cost fluctuations to inflation.

In Section 7, we conduct a model validation exercise using the cost-based NKPC to analyze the effects of supply shocks on inflation. By tracing the impact of identified oil shocks (Känzig 2021) on marginal cost and inflation, we show that the impulse-responses for inflation produced by a cost-based NKPC model—calibrated to

our micro-level estimates—closely match the empirical impulse-responses estimated in the data. This exercise validates our bottom-up estimation approach and, as we discuss, illustrates the usefulness of the cost-based NKPC for analyzing supply-side shocks.

1 Theoretical framework

This section presents the theoretical framework that underlies our empirical analysis. We formulate the minimum structure required to produce firm pricing equations that allow us to identify the slope of the aggregate Phillips curve. The framework features heterogeneous firms competing under imperfect competition and subject to nominal rigidity. In this environment, firms internalize the impact of their pricing decisions on industry aggregates and are in turn influenced by the pricing decisions of their competitors. This model generates a micro-founded cost-based New Keynesian Phillips curve, the slope of which is a function of the structural parameters that govern firms’ pricing behavior.

1.1 Preferences and pricing behavior

The economy is populated by heterogeneous producers (or firms), denoted by f , each operating in an industry $i \in \mathcal{I} = [0, 1]$. We denote by \mathcal{F}_i the set of firms competing in industry i . Each firm is measure zero relative to the economy as a whole but may be large relative to its industry. Hence, it takes the aggregate expenditure as given but internalizes the effect of its pricing decisions on the consumption and price index of its industry.

Let P_{ft} denote the price charged by each firm for a unit of its output, P_{it} the industry price index, φ_{ft} a firm-specific relative demand shifter, and Y_{it} the real industry output. For any industry i , we consider an arbitrary, invertible demand system that generates a residual demand function of the following form:⁶

$$\mathcal{D}_{ft} := d(P_{ft}, P_{it}, \varphi_{ft})Y_{it} \quad \forall f \in \mathcal{F}_i. \quad (2)$$

⁶The focus on invertible demand systems is a mild technical assumption that excludes scenarios where firms offer goods that are perfect substitutes but encompasses any demand system with an arbitrary (albeit finite) elasticity of substitution across goods

Firms adjust their prices during the period in order to maximize expected profits facing nominal rigidities as in Calvo (1983).⁷ Each period they face a probability $(1 - \theta) \in [0, 1]$ of being able to change their price, independent across time and across firms. Thus, the price P_{ft} paid by consumers to buy goods produced by firm f is either the (optimal) reset price if the firm is able to adjust, denoted by P_{ft}^o , or the price it charged in the previous period, P_{ft-1} .

When choosing P_{ft}^o , firms consider both their own costs, the pricing choices made by competitors, as well as the impact of their own price adjustments on their residual demand and on the industry-wide price index. Let $\Lambda_{t,\tau}$ denote the stochastic discount factor between time t and $t + \tau$, $TC_{ft} := TC(\mathcal{D}_{ft})$ the real total costs, and MC_{ft}^n the nominal marginal cost of firm f . Then the optimal reset price P_{ft}^o solves the following profit maximization problem:

$$\max_{P_{ft}^o, \{Y_{ft+\tau}\}_{\tau \geq 0}} \mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} \theta^\tau \left[\Lambda_{t,\tau} \left(\frac{P_{ft}^o}{P_{t+\tau}} \mathcal{D}_{ft+\tau} - TC(\mathcal{D}_{ft+\tau}) \right) \right] \right\},$$

subject to the sequence of expected demand functions $\{\mathcal{D}_{ft+\tau}\}_{\tau \geq 0}$ in Equation (2). Nominal rigidities generate forward-looking pricing behavior, as firms take into account that it might not be possible to adjust prices every period. As a result, the optimal reset price is a weighted average of current and expected future nominal marginal costs and markups. Denoting by μ_{ft} the desired log markup, the FOC of the problem is:

$$\mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} \theta^\tau \Lambda_{t,\tau} \mathcal{D}_{ft+\tau} \left[\frac{P_{ft}^o}{P_{t+\tau}} - (1 + \mu_{ft+\tau}) \frac{MC_{ft+\tau}^n}{P_{t+\tau}} \right] \right\} = 0. \quad (3)$$

Thus, the optimal reset price depends on the expected path of marginal cost and desired markups over the period the firm expects its price to be fixed, where θ^τ is the probability the firm expects its price to be fixed τ periods from now. The desired markup is given by the Lerner index:

$$\mu_{ft+\tau} := \ln \left(\frac{\epsilon_{ft+\tau}}{\epsilon_{ft+\tau} - 1} \right), \quad (4)$$

where $\epsilon_{ft+\tau} := -\frac{\partial \ln \mathcal{D}_{ft+\tau}}{\partial \ln P_{ft}^o}$ denotes the residual demand elasticity faced by f .

⁷In a companion paper (Gagliardone et al. 2025), we consider state-dependent pricing. We show that in normal times (i.e., in the absence of large aggregate shocks, as is generally the case during our sample period), Calvo provides a good approximation of firms' pricing decisions. See also Auclert et al. (2024) and references therein.

1.2 Technology

A unit of output of Y_{ft} is produced at a nominal marginal cost of:

$$MC_{ft}^n = C_{it} \mathcal{A}_{ft} Y_{ft}^{v_{ft}}, \quad (5)$$

where C_{it} denotes the nominal marginal unit cost of the composite input factor (e.g., wages and intermediate goods); \mathcal{A}_{ft} is a firm-specific cost shifter that affects the average unit cost of production and is inversely related to the firm's total factor productivity (TFP); v_{ft} is a firm-specific parameter that pins down short-run returns to scale in production (henceforth SR-RTS), given by $(1/(1 + v_{ft}))$.⁸

To derive the aggregate implications of the model, we assume that the economy displays constant returns to scale in the aggregate (i.e., $v_{ft} = 0$ on average). This assumption rules out macroeconomic complementarities due to the feedback of firms' pricing behavior on their respective marginal cost (see e.g. Galí 2015).⁹ We relax this assumption in Section 4.1. There we show that our estimates of the Phillips curve are robust as the empirical evidence is broadly consistent with the constant returns to scale assumption at both the sectoral and aggregate levels.¹⁰

1.3 The optimal reset price

We log-linearize the FOC in Equation (3) around the symmetric steady state with zero inflation.¹¹ Denoting the variables in logs with lower-case letters, we obtain that the reset price satisfies:

$$p_{ft}^o = (1 - \beta\theta) \mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} (\beta\theta)^\tau \left(\mu_{ft+\tau} + mc_{ft+\tau}^n \right) \right\}. \quad (6)$$

The log-linearized desired markup (in deviation from steady state markup μ_f) is a function that depends inversely on the log-difference between the firms' own reset

⁸This functional form is rather general and consistent with standard production technologies used in the literature (see, e.g., Hottman, Redding and Weinstein 2016). For instance, it nests Cobb-Douglas and CES as special cases.

⁹Macroeconomic complementarities can arise, for example, from roundabout production as in Basu (1995) or local input markets as in Woodford (2011).

¹⁰Near constant returns to scale also help reconcile our estimates of a steep cost-based Phillips curve with the flat output-based Phillips curve commonly found in the literature, as we discuss in Section 6.

¹¹The choice of the zero-inflation steady state permits simpler notation; but is largely immaterial for our purposes. We relax it in the empirical analysis, where we allow for sector/industry-specific trends.

price and its competitors' prices (p_{it}^{-f}):

$$\mu_{ft} - \mu_f = -\Gamma \left(p_{ft}^o - p_{it}^{-f} \right) + u_{ft}^\mu, \quad (7)$$

where $\Gamma > 0$ denotes the markup elasticity with respect to prices and u_{ft}^μ is a firm-specific demand shock to the desired markup that depends on the demand shifter φ_{ft} .¹² Under weak assumptions, the expression in Equation (7) holds for standard frameworks with imperfectly competitive firms, including monopolistic competition with variable elasticity of demand (Kimball 1995), static oligopoly (Atkeson and Burstein 2008) and dynamic oligopoly (Wang and Werning 2022). These frameworks share the property that, in equilibrium, a firm's elasticity of demand declines as its market share increases. Thus, the presence of strategic complementarities in price setting implies that a relative price increase lowers a firm's desired markup, dampening the response of prices to marginal cost.

Substituting the expression for $\mu_{ft+\tau}$ in the log-linearized first-order condition, we obtain the following forward-looking pricing equation:

$$p_{ft}^o = (1 - \beta\theta) \mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} (\beta\theta)^\tau \left((1 - \Omega)(mc_{ft+\tau}^n + \mu_f) + \Omega p_{it+\tau}^{-f} \right) \right\} + u_{ft}, \quad (8)$$

where $u_{ft} := (1 - \beta\theta)(1 - \Omega) \mathbb{E}_t \{ \sum_{\tau=0}^{\infty} (\beta\theta)^\tau u_{ft+\tau}^\mu \}$ captures the expected discounted value of the firm's future demand shocks. The parameter $\Omega := \frac{\Gamma}{1+\Gamma}$ captures the strength of strategic complementarities and impacts the firm's pricing policy by muting the price response to changes in marginal costs. If the demand elasticity is constant, as in the textbook New Keynesian model with monopolistically competitive firms, the desired markup is a constant. In this case, $\Omega = 0$ and the optimal pricing equation simplifies to the familiar formulation where the reset price exclusively depends on the current and future stream of marginal costs. Competitors' prices are then irrelevant.

1.4 The primitive New Keynesian Phillips curve

The log-linear aggregate price index is given by:

$$p_t = (1 - \theta)p_t^o + \theta p_{t-1}, \quad (9)$$

¹²See Section OA.1 of the Supplemental Appendix for derivations of the log-linearized markup and the expression for u_{ft}^μ under CES and Kimball demand systems.

with p_t and p_t^o denoting the aggregate price indices implied by the demand system. Let mc_t^n denote the aggregate log-nominal marginal cost, and define the aggregate real marginal cost and aggregate inflation as $mc_t = mc_t^n - p_t$ and $\pi_t = p_t - p_{t-1}$, respectively. Averaging the pricing equation in (8) across firms and industries and writing it in recursive form, we obtain an equation for the aggregate reset price:

$$p_t^o = (1 - \beta\theta)((1 - \Omega)(mc_t^n + \mu) + \Omega p_t) + \beta\theta\mathbb{E}_t p_{t+1}^o + \frac{\theta}{1 - \theta}u_t, \quad (10)$$

where u_t is an aggregate cost-push shock. Combining Equations (9) and (10) gives the primitive formulation of the NKPC curve:

$$\pi_t = \lambda \widehat{mc}_t + \beta \mathbb{E}_t \{\pi_{t+1}\} + u_t, \quad (11)$$

which asserts that inflation depends on real marginal cost in deviation from its steady state level, $\widehat{mc}_t := mc_t^n - p_t + \mu$, and on expected future inflation. The slope of the cost-based NKPC curve is given by:¹³

$$\lambda := \frac{(1 - \theta)(1 - \beta\theta)}{\theta}(1 - \Omega). \quad (12)$$

Two observations are worth noting. First, the primitive formulation of the Phillips curve in Equation (11) features the log-deviation of real marginal cost from its steady state as the relevant real activity variable driving inflation. In contrast, the conventional formulation of the Phillips curve, displayed in Equation (1), uses the output or unemployment gap as a proxy for marginal cost. As we will discuss, the mapping between marginal cost and the output gap is theoretically valid only under specific circumstances. Moreover, even when a proportionality between the two variables can be established, the elasticity of marginal cost to the output (or unemployment) gap is generally different from one. We return to these points in Section 6.

Secondly, the slope of the cost-based NKPC is a function of the primitives that govern firms' pricing behavior. As in standard New Keynesian models (e.g., Galí and Gertler 1999), high nominal rigidities and low discounting flatten the sensitivity

¹³In an environment with cross-industry heterogeneity in the parameters θ and Ω , aggregation across industries implies that the cost-based NKPC becomes $\pi_t = \lambda \cdot \widehat{mc}_t + Cov(\lambda_i, \widehat{mc}_{it}) + \beta\mathbb{E}_t \pi_{t+1}$, where $\lambda := \int \lambda_i di$, the slope in Equation (12), and $\lambda_i := \frac{(1 - \theta_i)(1 - \beta\theta_i)}{\theta_i}(1 - \Omega_i)$. Thus, aggregate pass-through also depends on the cross-sectional covariances $Cov(\widehat{mc}_{it}, \theta_i)$ and $Cov(\widehat{mc}_{it}, \Omega_i)$. This source of heterogeneity also matters in the presence of input-output linkages between industries (Rubbo 2023). While we abstract from these considerations, understanding their quantitative importance is an interesting avenue for future research.

of inflation to changes in real economic activity. Additionally, Equation (12) shows how strategic complementarities also contribute to reducing the slope. We take the structural pricing equation in (8) using micro data on prices and costs to identify the structural parameters θ and Ω and, given a calibration of the discount factor β , pin down the slope of the cost-based Phillips curve.

2 Data and measurement

2.1 Data

We assemble a micro-level dataset that covers the manufacturing sector in Belgium between 1999 and 2019, at the business cycle frequency. The dataset is compiled from administrative sources, extending and enriching the annual dataset used by Amiti, Itskhoki and Konings (2019). A unique feature of our dataset is its ability to track quarterly product-level prices and quantities sold in the domestic market by both domestic and foreign producers, as well as quarterly information on production costs for domestic producers.

The PRODCOM dataset allows us to observe domestic firms' quarterly sales and physical quantities sold for each narrowly defined (8-digit PC codes) manufacturing product. We use this highly disaggregated information to calculate domestic unit values (sales over quantities) at the firm-product level. We obtain similar data on foreign competitors from the administrative records of Belgian customs declarations. Specifically, for each manufacturing product sold by a foreign producer to a Belgian buyer, we observe quarterly sales and quantity sold for different products (8-digit CN codes), from which we compute unit values of foreign competitors in local markets.

We use detailed administrative data to measure firms' variable production costs. We obtain information on firms' quarterly purchases of intermediates (materials and services) from their VAT declarations. We draw upon firms' social security declarations to measure their labor costs (the wage bill) on a quarterly basis.

Sample properties—Our final sample includes 4,598 firms observed over 84 quarters (1999:Q1–2019:Q4), totaling 132,915 observations. We provide detailed information on the data sources and data cleaning procedures in the Supplemental Appendix OA.2. Table 1 presents summary statistics of our dataset. Several features are worth noting.

Table 1: Summary statistics

	Mean	5 st pctle	25 th pctle	Median	75 th pctle	95 th pctle
Number of industries within firm	1.10	1.00	1.00	1.00	1.00	2.00
Within firm revenue share of main industry	98.23	86.58	100.00	100.00	100.00	100.00
Firm's market share within industry	1.71	0.06	0.22	0.53	1.35	6.52
Firm's market share within sector	0.21	0.01	0.02	0.05	0.13	0.69
Firm's market share within manufacturing	0.03	0.00	0.00	0.01	0.01	0.08
Number of consecutive quarters in sample	42.21	11.00	24.00	38.00	59.00	82.00

Notes. This table reports summary statistics for sample of domestic producers in PRODCOM.

First, our dataset covers the lion's share of domestic manufacturing production in Belgium. The average firm in our dataset employs 74 employees (measured in full-time equivalents) and has a domestic turnover (sales) of 6 million Euros. The sales of the smallest firms in the sample are worth less than one-tenth of a thousandth of those generated by the largest producers.

Second, throughout the paper we adopt a narrow industry definition based on 4-digit NACE Rev.2 codes, the standard sector classification system in the European Union. Based on this classification, we sort firms into 169 manufacturing industries, distributed across 9 manufacturing sectors.¹⁴ This classification optimally balances a coherent definition of the industry (which is mostly precise if narrow) with the ability to identify an appropriate set of competitors (both domestic and foreign) competing to gain market share in Belgium. Table 1 shows that the vast majority of the firms in our sample specialize in only one manufacturing industry. Even for those firms

¹⁴The first four digits of the PRODCOM product classification coincide with the first four digits of the NACE Rev.2 classification and also to the first 4 digits of the CN product code classification used in the customs data. Following the official Eurostat classification system, we define manufacturing sectors by grouping 2-digit NACE Rev.2 codes, appropriately harmonized to account for changes in product classifications over time. See Section OA.2 of the Supplemental Appendix for sectors' definitions.

that operate in multiple industries, the contribution of the main industry to total firm revenues is, on average, 98% (median 100%). For the few multi-industry firms, we treat each industry as a separate firm in accordance with the theoretical framework.¹⁵

Third, the typical sector is characterized by a large number of firms with small market shares—the average within-industry share is approximately 1.5% on average, with a median of 0.5%—and a few relatively large producers. To the extent that these large firms internalize the effect of their pricing and production decisions on industry aggregates and strategically react to the pricing decisions of their competitors, the monopolistic competition benchmark would be a poor approximation. The theoretical framework introduced in the previous section explicitly accounts for this.

Fourth, although the largest firms have nontrivial market shares in their industries, they are small compared to the volume of economic activity of their macro sector (e.g., textile manufacturing or electrical equipment manufacturing) and, even more so, compared to the volume of economic activity in the whole manufacturing sector in Belgium. It is therefore reasonable to assume that even the largest producers do not internalize the effect of their pricing and production decisions on the aggregate economy.

Finally, our data allow us to observe a long time series of both prices and marginal costs. On average, we observe firms for approximately 10 consecutive years (42 quarters). This feature of the data is particularly important for identification purposes. As we discuss below, a long time series enables us to include unit fixed effects in our empirical models to control for time-invariant confounding factors without suffering from the classical Nickell bias that frequently complicates the estimation of dynamic panel models.

2.2 Measurement

We now describe how to map the theoretical counterparts to the data. We use product-level prices, firm-level production costs, and information on prices of competitors (firms that operate in the same 4-digit industry) to construct measurable counterparts of prices and reset prices, which vary at the firm-industry-quarter level.

¹⁵Because most firms operate in only one industry, and the main industry accounts for the lion's share of sales of multi-industry firms, all our results are essentially unchanged if we restrict the sample to the main industry for each firm.

Section OA.2 of the Supplemental Appendix provides a detailed description of the procedure used to construct all our variables.

Output prices—The key variable of interest is the domestic price of goods charged by firms in the local market (Belgium). We construct a firm-industry price index that varies at the same level as our reset prices. We use the subscript i to denote an industry, f to denote a firm-industry pair, and t to denote time (quarters). s_{ft} denotes the revenue share of the firm in the industry.

We compute the change in the firm-industry price index, P_{ft}/P_{ft-1} , using the most disaggregated level allowed by the data. For domestic producers, the finest level of aggregation is the firm×8-digit PC product code level. For foreign competitors, it is the importing-firm×source country×8-digit PC product code level.¹⁶ Approximately half of the domestic firms in our sample are multi-product firms, meaning they produce multiple 8-digit products within the same industry. For these entities, we compute the price change by aggregating changes in product-level prices using a Törnqvist index:¹⁷

$$P_{ft}/P_{ft-1} = \prod_{p \in \mathcal{P}_{ft}} (P_{pt}/P_{pt-1})^{\bar{s}_{pt}}.$$

In the formula above, \mathcal{P}_{ft} represents the set of 8-digit products manufactured by firm f , P_{pt} the unit value of product p , and \bar{s}_{pt} the product’s Törnqvist weight computed as the average of (within-firm) sale shares of the product between t and $t - 1$: $\bar{s}_{pt} := \frac{s_{pt} + s_{pt-1}}{2}$. Finally, we construct the time series of the price index by concatenating the quarterly price changes, starting from a firm-specific base year, as discussed in Section OA.2 of the Supplemental Appendix.¹⁸

Using a similar approach, we construct the price index of competitors for each

¹⁶In the raw customs data, products are measured using the more disaggregated CN 8-digit product classification. We map the CN product codes in the customs data to PC product codes used in PRODCOM using the official bridge tables available on the Eurostat web page. See the Supplemental Appendix OA.2 for additional details.

¹⁷Given that our measure of reset prices varies at the firm-industry level and our assumption that the elasticity of substitution is common across all firm’s products within an industry, we would obtain approximately the same parameter estimates running our models in the more granular dataset (with product-level price variation), as long as the product-level observations are weighted by the same Törnqvist weights, \bar{s}_{pt} , used in the aggregation.

¹⁸The normalization of the level of the price indices in the base year is one rationale for the inclusion of firm fixed effects in our empirical specifications.

domestic firm by concatenating quarterly price changes as follows:

$$P_{it}^{-f} / P_{it-1}^{-f} = \prod_{k \in \mathcal{F}_i / f} (P_{kt} / P_{kt-1})^{\bar{s}_{kt}^{-f}}. \quad (13)$$

Here, $\bar{s}_{kt}^{-f} := \frac{1}{2} \left(\frac{s_{kt}}{1-s_{ft}} + \frac{s_{kt-1}}{1-s_{ft-1}} \right)$ represents a Törnqvist weight, constructed by averaging the residual revenue share of competitors in the industry at time t (net of firm f revenues) with that at time $t-1$. Note that the set of domestic competitors for each Belgian producer, denoted as \mathcal{F}_i , includes not only other Belgian manufacturers operating in the same industry, but also foreign manufacturers selling the same goods to Belgian customers.

Marginal costs—The cost structure outlined in Equation (5) implies that a firm’s nominal log-marginal cost is equal to logarithm of average variable costs plus a term reflecting SR-RTS:

$$mc_{ft}^n = \ln(\text{TVC}_{ft} / Y_{ft}) + \ln(1 + v_{ft}). \quad (14)$$

Accordingly, we construct our empirical proxy of firms’ marginal costs using variation in average variable costs. We measure total variable costs (TVC_{ft}) as the sum of intermediate costs (materials and services purchased) and labor costs (wage bill). Intermediate input costs account, on average, for 75 percent of total variable costs. They are also the most volatile cost component, with a within-firm coefficient of variation that is more than twice as large as that of labor costs (1.77 vs 0.77). We obtain a firm-specific quantity index for domestic sales (Y_{ft}) by scaling a firm’s domestic revenues by its domestic price index, such that $Y_{ft} = (PY)_{ft} / \bar{P}_{ft}$. For single-industry firms, \bar{P}_{ft} coincides with the firm-industry price index P_{ft} , which was discussed earlier. For multi-industry firms, we aggregate industry prices P_{ft} by using as weights the firm-specific revenue shares of each industry.¹⁹

Returns to scale, which are not directly observable in the data, will enter the error term of our empirical models. In our baseline regression model, we assume Cobb-Douglas technologies. Under this assumption, we can control for the

¹⁹Specifically, we apply the Törnqvist weight of each (4-digit) industry bundle i produced by firm f in quarter t , which is defined as $(s_{fit} + s_{fit-1})/2$, where s_{fit} is the share of revenues of the firm coming from sales in industry i in total sales across industries. The choice of \bar{P}_{ft} has essentially no impact on our estimation results because, as discussed, the majority of the firms in our data operate in only one industry, and the sales of multi-industry firms are typically concentrated in a primary industry. In fact, our empirical results are robust to defining \bar{P}_{ft} as the price of the main industry or using other aggregation methods (e.g., an arithmetic average or a CES aggregator).

curvature of the production function using either industry or firm fixed effects. In the robustness section, we consider the possibility that SR-RTS vary over time and with the scale of production, as it would be the case under CES technologies.

3 Identification strategy

We now present the identification strategy behind the estimation of the structural parameters θ and Ω that determine the slope of the NKPC. We begin by mapping the theoretical model from Section 1 to the data and derive a dynamic pass-through model with a measurable counterpart. In doing so, we highlight the connections and distinctions between our dynamic pass-through and the static pass-through analyzed in previous studies. We then discuss the assumptions and instrumental variables underlying our identification approach.

3.1 Econometric framework for dynamic pass-through

Baseline model—Under our Calvo framework, given the firm’s information set at time t , we can express the conditional expectation of the observed price as:

$$\mathbb{E}\{p_{ft}|p_{ft}^o, p_{ft-1}\} = (1 - \theta)p_{ft}^o + \theta p_{ft-1}.$$

Starting from the above, we define the projection error $v_{ft} := p_{ft} - \mathbb{E}\{p_{ft}|p_{ft}^o, p_{ft-1}\}$ and use Equation (8) to solve out for p_{ft}^o . After substituting the expected values of prices and costs with their realizations, we obtain the following dynamic pass-through regression linking the observed price to short-run fluctuations in current and expected marginal cost and competitors’ prices:

$$p_{ft} = (1 - \theta)\left((1 - \Omega)\left((mc_{ft}^n)^\infty + \mu_f\right) + \Omega\left(p_{it}^{-f}\right)^\infty\right) + \theta p_{ft-1} + \varepsilon_{ft}, \quad (15)$$

where $(x_t)^\infty := (1 - \beta\theta) \sum_{\tau=0}^{\infty} (\beta\theta)^\tau x_{t+\tau}$ for $x_t = \{mc_{ft}^n, p_{it}^{-f}\}$ denotes the discounted present value of marginal cost and competitors’ prices, and ε_{ft} captures a composite residual:

$$\varepsilon_{ft} := v_{ft} + (1 - \theta)(1 - \beta\theta)e_{ft} + (1 - \theta)u_{ft}.$$

Here e_{ft} denotes a mean-zero expectational error—orthogonal to the realizations of the other variables under rational expectations—and u_{ft} captures the firm’s demand

shock that enters the reset price (Equation (8)).²⁰

The intuition underlying Equation (15) is the following: due to the presence of nominal rigidities, firms can adjust prices each period with probability $(1 - \theta)$. Conditional on being able to adjust, they do so by accounting for current and expected future changes in marginal costs and in their competitors' prices. The parameter Ω measures the relative strength of these two forces. Thus, in an environment with forward-looking pricing behavior and oligopolistic competition, the short-run pass-through of marginal costs depends on both the degree of nominal rigidity and strategic complementarities. Specifically, the elasticity of a firm's own price to a permanent shock to marginal cost is given by $\frac{\partial p_{ft}}{\partial mc_{ft}^o} = (1 - \Omega)(1 - \theta)$. As we explain below, the lagged price enters the empirical specification as a predetermined regressor that controls for short-run price dynamics.

The error correction framework and long-run pass-through—It is useful to highlight the connections and differences between our dynamic (short-run) pass-through model and the static (long-run) model used in previous literature.

Without loss of generality, prices and reset prices satisfy the long-run cointegrating relationship:

$$p_{ft} = p_{ft}^o + \eta_{ft}, \quad (16)$$

where $\eta_{ft} := p_{ft} - p_{ft}^o$ denotes a cointegration error. Adding and subtracting $(1 - \theta)p_{ft}^o + p_{ft-1}$ to Equation (8) and rearranging, we can express our dynamic pass-through regression in Equation (15) as an error-correction model:

$$\Delta p_{ft} = (1 - \theta)\Delta p_{ft}^o - (1 - \theta)\eta_{ft-1} + v_{ft}. \quad (17)$$

The first term, Δp_{ft}^o , captures variation in reset prices due to supply and demand shocks. The second term, $\eta_{ft-1} = p_{ft-1} - p_{ft-1}^o$, is the error-correction term. As in Engle and Granger (1987), this term controls for persistent deviations of prices and reset prices from their long-run cointegrating relation. Unless prices are fully flexible (i.e., $\theta = 0$), failing to account for this term in the regression model would lead to biased estimates.²¹ The firm's lagged price in the dynamic pass-through model in

²⁰Note that deviations from the rational expectation benchmark do not pose a threat to identification as long as our instruments for marginal costs and competitors' prices are orthogonal to the (possibly non-zero mean) forecast error e_{ft} .

²¹Combining the cointegrating relation in (16) and the error correction in (17), we obtain that the

Equation (15) serves precisely the purpose of the error correction term.

The omitted variable bias goes to zero as prices become flexible. Taking the limit of Equation (17) for $\theta \rightarrow 0$ and $\eta_{ft} \rightarrow v_{ft}$ obtaining:

$$\Delta p_{ft} = \Delta p_{ft}^o + \Delta \eta_{ft}. \quad (18)$$

Equation (18) represents a long-run (i.e., static) pass-through model, which corresponds to the time-differenced version of the cointegrating relation in (16). The orthogonality condition of this model, $Cov(\Delta p_{ft}^o, \Delta \eta_{ft}) = -\theta Var(\Delta p_{ft}^o) + Cov(\Delta p_{ft}^o, v_{ft}) = 0$, holds when the data is measured at low frequency (e.g., annual, as in Amiti, Itskhoki and Konings 2019) so that nominal rigidities can be ignored ($\theta \cong 0$). However, given the degree of stickiness of prices observed in the data—3 to 4 quarters according to our estimates—the orthogonality condition will not hold with high-frequency data (e.g., quarterly, as in our case). In this case, the dynamic pass-through model accounts for the error correction needed to identify the parameters of interest.

3.2 Empirical specification

We take the dynamic pass-through model in (15) to the data. Identification of the parameters θ and Ω requires us to address issues related to the measurement and endogeneity of the present values of prices and costs. The richness and granularity of our data allow us to tackle these issues through a combination of fixed effects and instrumental variables.

We map the population regression in (15) to the following sample analog:

$$p_{ft} = (1 - \theta) \left((1 - \Omega)(mc_{ft}^n)^8 + \Omega(p_{it}^{-f})^8 \right) + \theta p_{ft-1} + \alpha_f + \alpha_{s \times t} + \varepsilon_{ft}. \quad (\text{Model A})$$

We use the series of firm-level marginal cost and competitor's prices described in Section 2.2 to construct measurable counterparts of (realized) present discounted values. We calibrate the discount factor $\beta = 0.99$, a standard value for quarterly data, and truncate the present value sequences at eight-quarter leads.²²

cointegration error satisfies an ARMA process, $\eta_{ft} = \theta \eta_{ft-1} - \theta \Delta p_{ft}^o + v_{ft}$, with autocorrelation coefficient given by the degree of nominal rigidities.

²²Specifically, we have that $(x_t)^T := (1 - \beta\theta) \sum_{\tau=0}^{T-1} (\beta\theta)^\tau x_{t+\tau} + (\beta\theta)^T x_{t+T}$ for $x_t = \{mc_{ft}^n, p_{it}^{-f}\}$ and $T = 8$. For reasonable values of θ , this choice ensures that the discount factor $(\beta\theta)^\tau \approx 0$ for $\tau > T$. We verified that allowing for larger values of T has no impact on our estimation results.

We include a vector of sector-by-time fixed effects, $\alpha_{s \times t}$. These fixed effects help us address several issues that typically complicate the identification of the NKPC with aggregate data. First, they allow us to extend the theoretical framework to incorporate sector-specific trends and time-varying steady states of the variables in the model. Second, the inclusion of sector-by-time fixed effects helps us address concerns related to shifts in long-term inflation expectations (Hazell et al. 2022). Third, these fixed effects soak up variation in prices driven by demand shocks common across firms within a sector, which could generate a spurious correlation between marginal cost and prices due to general equilibrium effects.

Our empirical model also includes a vector of firm fixed effects, α_f , which helps us address both measurement and endogeneity issues. First, firm fixed effects absorb variation in steady-state markups. These could be heterogeneous across firms because, for instance, firms might be producing goods of different quality. Second, given our empirical measure of nominal marginal costs, the residuals of our baseline empirical model will capture the present discounted value of firms' SR-RTS. The firm fixed effects control for this possible source of omitted variable bias to the extent that production technologies are time-invariant.²³ Finally, by controlling for these fixed effects, the normalization of firm-level prices becomes immaterial (see footnote 18).

Despite the inclusion of this rich set of fixed effects, both measurement and endogeneity issues may still threaten identification. First, firm-level marginal costs and competitors' prices are subject to measurement error, potentially leading to attenuation bias. More importantly, some of the variation in these variables might be due to demand-side factors and thus correlate with the error term. For marginal cost, this occurs if firms' short-run cost schedules depend on the scale of production (i.e., if SR-RTS differ from unity) or if firms face locally upward-sloping input supply curves (i.e., if input prices adjust with demand). Similarly, competitors' prices are jointly determined with a firm's own price and, therefore, might correlate with the firm's demand shocks.

To address these issues, we construct a set of instruments (Z_{ft}) and estimate Model A via Generalized Method of Moments (GMM) imposing moment conditions of the form $\mathbb{E}\{Z_{ft} \cdot \varepsilon_{ft}\} = 0$.²⁴

²³In Section 4.1, we present a robustness exercise that explicitly accounts for possible variation in this variable and show that our estimates are essentially unchanged.

²⁴The GMM estimation procedure follows a two-step approach, iterating over guesses of θ and Ω

Instrument for marginal cost—Our baseline instrument for marginal cost leverages variation in the persistent component of firms’ total factor productivity. We construct a firm-level total factor productivity index using information on firm-level physical output and input demands (labor, capital, and intermediate inputs). In logs:

$$TFPQ_{ft} = y_{ft} - f(l_{ft}, k_{ft}, m_{ft}; \vartheta),$$

where $f(\cdot)$ denotes the firm’s gross-output production function and ϑ a vector of parameters determining the output elasticities with respect to different inputs. Variation in TFPQ captures changes in firms’ technical efficiency, which is a fundamental component of their marginal cost, as noted in Equation (5). We use four-quarters lagged technical efficiency ($TFPQ_{ft-4}$) as our instrument for marginal cost.

This instrument is relevant if technical efficiency affects marginal costs and if technical productivity is persistent. As we demonstrate below, the instrument is strong. The identifying assumption that guarantees that its relevant is that the persistent component of technical efficiency is orthogonal to the current and future demand shocks captured in the error term. Given the timing of our instrument, this exclusion restriction holds if *either* of the following conditions is satisfied: (i) our technical productivity measure is uncorrelated with demand shocks, or (ii) demand shocks are i.i.d. or sufficiently transitory after removing both a permanent component of demand (absorbed by the firm fixed effects) and industry trends (absorbed by the industry-by-time fixed effects). In Section 4.2, we present a battery of empirical tests suggesting that both conditions are likely satisfied in the data.

For each firm in our sample, we recover firm-level technical efficiency as a residual from a gross-output production function estimation. For our baseline instrument, we parameterized firms’ production functions assuming Cobb-Douglas technologies. We adopt the estimation strategy developed in Lenzu, Rivers and Tielens (2024) to estimate the parameter vector ϑ , allowing elasticities to vary

until convergence. In each iteration, we compute the present discounted values of marginal costs and competitors’ prices, $(mc_{ft}^n)^8$ and $(p_{it}^{-f})^8$, based on the current guess of θ . To ensure the estimates are macroeconomically representative, we weight observations using their Törnqvist weight, \bar{s}_{ft} , so that each firm contributes proportionally to its role in the construction of aggregate price index. Standard errors are clustered at the sector level to account for potential correlation of error terms across firms within similar industries.

across sectors.²⁵ We demonstrate robustness to alternative parameterizations of the production function in Section 4.1.

Instruments for competitors' prices—Building on Amiti, Itskhoki and Konings (2019), we construct two instruments for competitors' prices that leverage variation in international trade prices.

Let us denote by \mathcal{F}_{ki}^* the set of international competitors of firm f that are located in country k that sell products in industry i . The first instrument, denoted by $p_{it}^{\star EU}$, is a shifter of the price of Euro area international competitors. We use the COMEXT dataset from Eurostat to compute the (sales-weighted) average log price of goods in industry i that a given Euro-area country, k , charges to different export destinations around the world.²⁶ We exclude Belgium as well as all other Euro-area countries from the list of export destinations to make sure that variation in the instrument does not pick up demand shocks that are correlated across Belgium and other neighboring countries. We compute an index by averaging across all competitors j from EU countries:

$$\Delta p_{it}^{\star EU} = \sum_{k \in EU} \sum_{j \in \mathcal{F}_{ki}^*} w_{jt} \cdot \Delta \bar{p}_{kit}^{-B},$$

where the weight is obtained by normalizing the Törnqvist weight in formula (13) by the market share of EU competitors in industry i in Belgium: $w_{jt} := \bar{s}_{jt}^{-f} \cdot (\sum_{k \in EU} \sum_{j \in \mathcal{F}_{ki}^*} \bar{s}_{jt}^{-f}) / (\sum_k \sum_{j \in \mathcal{F}_{ki}^*} \bar{s}_{jt}^{-f})$. Finally, we concatenate $\Delta p_{it}^{\star EU}$ to obtain the instrument (in levels), $p_{it}^{\star EU}$. The rationale for this instrument is that the average price charged by international competitors outside the EU correlates with their marginal cost of production but not with demand shocks in Belgium.

The second instrument, denoted by $p_{it}^{\star F}$, is a shifter of the price of non-EU competitors that leverages variation in bilateral exchange rates (Δe_{kt}) between the currency used by country k and the Euro:

$$\Delta p_{it}^{\star F} = \sum_{k \notin EU} \sum_{j \in \mathcal{F}_{ki}^*} w_{jt} \cdot \Delta e_{kt},$$

where the weight w_{jt} is now scaled by the market share of non-EU competitors: $w_{jt} := \bar{s}_{jt}^{-f} \cdot (\sum_{k \notin EU} \sum_{j \in \mathcal{F}_{ki}^*} \bar{s}_{jt}^{-f}) / (\sum_k \sum_{j \in \mathcal{F}_{ki}^*} \bar{s}_{jt}^{-f})$. As before, we concatenate $\Delta p_{it}^{\star F}$ to obtain the instrument $p_{it}^{\star F}$ in levels. Here, the exclusion restriction requires that non-EU

²⁵The estimation procedure and results are detailed in Section OA.2 of the Supplemental Appendix.

²⁶As in Amiti, Itskhoki and Konings (2019), we consider the following set of EU countries $k \in \{\text{Austria, Germany, Spain, Finland, France, Greece, Ireland, Italy, Netherlands, and Portugal}\}$.

exchange rates are orthogonal to domestic demand shocks.

4 Estimation results

Instrument relevance and over-identification test—We begin by assessing the power of our instruments. In Panel A in Table 2, we regress the present values of marginal cost and competitors’ prices on our set of instruments, essentially producing what would be the first-stage regressions of a linear two-stage least squares model. The first two columns refer to the first stage of our baseline model, Model A. As we can see, all coefficients have the expected signs and are statistically significant. The high values of the Cragg-Donald and Kleibergen-Paap F-statistics indicate that we can confidently reject the hypothesis of weak instruments at standard confidence levels. Importantly, the low test statistics for the Hansen-Sargan over-identification test indicate that our instruments also satisfy the exclusion restrictions required by the moment conditions. Further evidence supporting instrument validity is presented below.

Baseline estimates of θ and Ω —Column (1) in Panel B reports the structural estimates for the degrees of nominal and real rigidities. Our estimates indicate a substantial degree of price stickiness, with a precisely measured estimate of $\theta = 0.711$. Through the lens of a Calvo model, this implies that, on average, prices remain fixed for approximately three to four quarters. This result aligns remarkably well with the average frequency of price adjustments measured by Nakamura and Steinsson (2008) from US PPI data and with those obtained from Belgian PPI data (0.72).

Our estimates also indicate an economically meaningful role for strategic complementarities in the pass-through of shocks. The estimate of Ω is 0.570 and is precisely estimated. This value is consistent with the findings in Amiti, Itskhoki and Konings (2019), suggesting that the pass-through of firms’ own marginal costs and of competitors’ prices are roughly of the same magnitude.

Pass-through and the slope of the cost-based NKPC—These estimates imply an elasticity of a firm’s own price to a permanent shock to marginal cost $\frac{\partial p_{f,t}}{\partial mc_{f,t}^p} = (1 - \Omega)(1 - \theta)$ of approximately 0.125. At the aggregate level, we find an economically

Table 2: Estimation results

<i>Panel A: First stage regressions</i>				
	Model A		Model B	Model C
Dep. Var →	$(mc_{ft})^8$	$(p_{it}^{-f})^8$	$(mc_{ft})^8$	mc_{ft}
	(1)	(2)	(3)	(4)
$TFPQ_{ft-4}$	-0.124 (0.021)	0.041 (0.023)	-0.111 (0.024)	-0.261 (0.004)
p_{ft}^{*EU}	0.246 (0.172)	0.535 (0.181)		
p_{ft}^{*F}	0.254 (0.173)	0.542 (0.190)		
p_{ft-1}	0.244 (0.031)	0.122 (0.033)	0.270 (0.028)	0.324 (0.004)
<i>Panel B: Structural estimates</i>				
	Model A	Model B	Model C	Model A-U
	(1)	(2)	(3)	(4)
θ	0.711 (0.014)	0.679 (0.026)	0.710 (0.019)	0.714 (0.065)
Ω	0.570 (0.059)	0.502 (0.152)	0.432 (0.170)	0.519 (0.084)
ρ^{mc}			0.747 (0.092)	
ϱ				0.708 (0.021)
<i>Panel C: Slope of the Phillips curve</i>				
λ	0.052 (0.007)	0.077 (0.038)	0.069 (0.030)	0.056 (0.027)
<i>Test statistics</i>				
Cragg-Donald F	939.489	1778.129	5775.577	
Kleibergen-Paap F	93.093	137.779	101.657	
Hansen-Sargan J	4.523	0.662	0.214	4.501
$H_0 : \theta = \varrho$ p-val				0.900

Notes. Panel A reports the estimates of linear regressions of marginal costs and competitors' prices on our instruments and controls. The standard errors of the first stage regression (reported in parenthesis) are block-bootstrapped to account for estimation error in the estimates needed to construct the present values. Panels b and c report the estimates of the structural parameters and slope of the NKPC (λ), respectively. GMM robust standard errors are clustered at the sector level. Models A and A-U include sector-by-time fixed effects and firm fixed effects. Models B and C include industry-by-time fixed effects and firm fixed effects.

meaningful relationship between fluctuations in marginal costs and aggregate inflation dynamics, even after accounting for the role of imperfect competition. Using Equation (12), we estimate the slope of the cost-based NKPC $\lambda = 0.052$ (Panel C, Column (1)), statistically significant and precisely estimated.

Our estimate differs markedly from existing estimates of the NKPC slope that use the output gap or unemployment as measures of real economic activity. These estimates typically range from two to ten times smaller than ours. For instance, Rotemberg and Woodford (1997) and Hazell et al. (2022) find a κ of 0.024 and 0.006, respectively, for US data. In Section 6, we revisit this comparison and demonstrate that the two estimates are, in fact, not inconsistent. We provide empirical evidence that helps reconcile why inflation appears to respond much more strongly to fluctuations in marginal cost than output (or unemployment).

4.1 Robustness analysis

Measurement of competitors’ prices and movements in industry demand—Our baseline measure for the competitors’ price index considers the set of relevant competitors to consist of other firms, both domestic and international, operating within the same four-digit industry as firm f . However, some relevant competitors may operate outside this industry boundary. To address this concern, we incorporate industry-by-time fixed effects, which absorb the present value of competitors’ prices without requiring prior assumptions about the relevant price index:²⁷

$$p_{ft} = (1 - \theta)(1 - \Omega)(mc_{ft}^n)^8 + \theta p_{ft-1} + \alpha_f + \alpha_{i \times t} + \varepsilon_{ft}. \quad (\text{Model B})$$

Note that the inclusion of narrowly defined industry-by-time fixed effects also captures more granular demand variation than the sector-by-time fixed effects in our baseline model may miss.

Column (3) of Panel A in Table 2 presents the first-stage regression estimates. The structural parameter estimates and implied NKPC slope are reported in Column (2) in Panels b and c, respectively. Both the estimated degree of price stickiness and strategic complementarities are close to their baseline values and precisely estimated.

²⁷Over 90 % of the variation in each firm competitors’ price index occurs at the industry-year level, since the vast majority of firms are small compared to the industry.

The implied slope of the Phillips curve is somewhat higher than our baseline estimate ($\lambda = 0.077$), suggesting an even stronger pass-through of fluctuations in marginal costs into prices.

Measurement of present values—A valuable feature of both Model A and B is that they do take a stand on the dynamic governing the evolution of marginal costs. The flip side of this flexibility is that the estimating equations are data-demanding and highly nonlinear. In particular, θ enters both as a coefficient in front of the present values and lagged prices as well as in the construction of the discounted present values. To address this concern, we assume that marginal cost, in deviations from its industry trend, follows a first-order auto-regressive process with persistence parameter $\rho < \frac{1}{\beta\theta}$ and estimate the following system of linear equations:

$$\begin{aligned} p_{ft} &= \Psi^{mc} \cdot mc_{ft}^n + \theta p_{ft-1} + \alpha_f + \alpha_{i \times t} + \varepsilon_{ft}, \\ mc_{ft}^n &= \rho^{mc} mc_{ft-1}^n + \alpha_f + \alpha_{i \times t} + \epsilon_{ft}^{mc}, \end{aligned} \quad (\text{Model C})$$

where $\Psi^{mc} := (1 - \theta)(1 - \Omega) \frac{1 - \beta\theta}{1 - \beta\theta\rho^{mc}}$ measures the price elasticity to a transitory shock to marginal costs, given the persistence of these shocks. Column (4) in Panel A presents the first-stage regression estimates. The structural parameter estimates and the implied NKPC slope are reported in Column (3) in Panels b and c, respectively. These estimates are in line with those obtained from the nonlinear GMM specifications, but even more precisely estimated. The implied price elasticity to a transitory increase in marginal cost of approximately $\Psi^{mc} = 0.103$ and, as with Model B, the estimated slope of the NKPC is somewhat higher than our baseline ($\lambda = 0.069$).

Unrestricted model—As explained in the previous section, the firm's lagged price enters the empirical specification of Model A as a control for short-run price dynamics with coefficient θ . On the one hand, imposing this theoretical restriction tightens the inference of the structural parameters. On the other hand, one might be concerned about the endogeneity of this variable and the possible biased that it might introduce in the estimate of θ . This would be the case to the extent that, after netting out industry-time fixed effects, the idiosyncratic demand shocks subsumed in the error term somehow display some degree of persistency (see the discussion in Section 4.2).

To address this concern, we estimate an unrestricted variant of Model A:

$$p_{ft} = (1 - \theta) \left((1 - \Omega)(mc_{ft}^n)^8 + \Omega(p_{it}^{-f})^8 \right) + \rho p_{ft-1} + \alpha_f + \alpha_{sxt} + \varepsilon_{ft}, \quad (\text{Model A-U})$$

which allows the coefficient in front of p_{ft-1} to possibly differ from θ .

We present the estimates obtained from Model A-U in Column (4), Panels b and c, of Table 2.²⁸ We find that estimates of both $\theta = 0.714$ and $\Omega = 0.519$ (and therefore the NKPC slope) are almost identical to those obtained from the restricted model in Column (1). Additionally, the estimated coefficient on lagged prices is $\rho = 0.708$. We cannot reject the null hypothesis that $\rho = \theta$ with a p-value of 0.90. Accordingly, the unconstrained model delivers an estimate of the NKPC slope that is essentially the same as in our baseline model ($\lambda = 0.056$). These results suggest that the estimates from the baseline model are robust to the possibility that p_{ft-1} is endogenous and also lend strong empirical support to the restrictions imposed by the economic theory.

Flexible production functions and variable SR-RTS—Our measure of marginal cost does not account for variation in firms’ SR-RTS, which are not directly observable in the data. If firms’ technologies are well approximated by a Cobb-Douglas production function, this variation would be absorbed by firm or industry fixed effects. However, if SR-RTS vary with the scale of production—as is generally the case under a CES production function—not accounting for this source of variation could lead to estimation bias.

As a robustness exercise, we estimate production functions using a Translog specification. This flexible functional form, a second-order approximation to CES, allows us to capture output elasticities with respect to different inputs, varying across firms and over time.²⁹ We then use the estimated output elasticities of labor and intermediate inputs to recover firm-specific and time-varying estimates of the sensitivity of marginal cost to output— $\hat{v}_{ft} = (\hat{\vartheta}_{ft}^l + \hat{\vartheta}_{ft}^m)^{-1} - 1$ in Equation (14)—and produce a measure of marginal cost that accounts for heterogeneity in firm-level SR-RTS.

Column (1) in Table 3 reports the estimates of our baseline model including the logarithm of the four-quarters lagged \hat{v}_{ft-4} as a control for SR-RTS. In Column (2),

²⁸The first stage regression is the same for Models A and Model A-U, reported in Panel A, Columns (1) and (2).

²⁹See Section OA.2 of the Supplemental Appendix for details on the estimation of Translog production functions and associated TFPQ instrument.

we directly account for variation in SR-RTS in the construction of our measure of nominal marginal cost, $mc_{ft}^n = \ln(\text{TVC}_{ft}/Y_{ft}) + \ln(1 + \hat{v}_{ft})$. In Column (3), we replace our baseline TFPQ instrument (recovered assuming Cobb-Douglas technologies) with a TFPQ index constructed assuming Translog technologies. Across all these alternative specifications, the pass-through coefficients and the implied NKPC slope remain broadly aligned with our baseline model estimates, both in terms of magnitude and statistical significance.

Macroeconomic complementarities—In Section 1, we derived the equation of the cost-based NKPC under the assumption of constant aggregate SR-RTS. In Section OA.1 of the Supplemental Appendix, we consider a more general framework with time-invariant, but possibly decreasing aggregate SR-RTS. In this case, the NKPC slope can be expressed as:

$$\lambda = \frac{(1 - \theta)(1 - \beta\theta)}{1 + \gamma\nu(1 - \Omega)}(1 - \Omega)\Theta,$$

where the additional term $\Theta := \frac{1 - \theta}{1 + \gamma\nu(1 - \Omega)} < 1$ captures the role of macroeconomic complementarities that stem from decreasing returns. Here ν is a parameter inversely related to the average of SR-RTS and γ denotes the elasticity of substitution across goods within industries. Thus, if the economy exhibits aggregate decreasing returns to scale ($\nu < 1$), the slope of the NKPC would be flatter, resulting in a more modest sensitivity of inflation to changes in real economy activity (see, e.g., Galí 2015).

In Section OA.2 of the Supplemental Appendix, we present estimates of returns to scale for different sectors, which suggest that SR-RTS are close to unity at both the sector level and in the aggregate. Given our estimates of ν and Ω , and a reasonable calibration of γ , we calculate a value of Θ of approximately 0.941.³⁰ Accordingly, we conclude that macroeconomic complementarities would lead to a modest reduction of the NKPC slope, well within the confidence bounds of our baseline estimates.

4.2 Threats to identification

We now address possible threats to our identification strategy concerning our instrument for marginal cost. The exclusion restriction on $TFPQ_{ft-4}$ is violated

³⁰We calibrate γ to 4 to obtain a gross aggregate steady-state markup between 1.3 and 1.4. The sectoral estimates of SR-RTS range from 0.93 to 0.98. The aggregate returns to scale are estimated to be approximately 0.965, which implies a value of ν of approximately 0.036.

Table 3: Robustness exercises and assessment of identification threats

	SR-RTS control (1)	mc^n with SR-RTS (2)	Translog TFPQ (3)	CU K (4)	CU K&L (5)	$TFPQ_{t-8}$ (6)
<i>Panel A: Structural estimates</i>						
θ	0.711 (0.013)	0.712 (0.014)	0.711 (0.014)	0.711 (0.014)	0.711 (0.014)	0.702 (0.019)
Ω	0.567 (0.061)	0.552 (0.059)	0.549 (0.059)	0.570 (0.059)	0.565 (0.060)	0.664 (0.091)
<i>Panel B: Slope of the Phillips curve</i>						
λ	0.052 (0.007)	0.054 (0.005)	0.054 (0.005)	0.052 (0.007)	0.052 (0.007)	0.044 (0.015)
<i>Test statistics</i>						
Cragg-Donald F	969.598	971.037	844.290	925.500	891.652	399.520
Kleibergen-Paap F	99.521	93.762	62.420	90.866	99.396	10.102
Hansen-Sargan J	4.462	4.917	4.616	4.524	4.651	4.033

Notes. This table reports various robustness tests and empirical tests on the validity of our instruments. All regressions build on our baseline specification (Model A). In Columns (1)–(3), we include firm-time varying SR-RTS in the control set, account for firm-time varying SR-RTS in the construction of our measure of nominal marginal cost, and replace our baseline TFPQ instrument with one that assumes Translog production functions. In Columns (4) and (5), we adjust our productivity instrument to account for variable capacity utilization in capital and in both capital and labor. In Column (6) we lag our TFPQ instrument by eight quarters instead of four quarters.

if lagged technical productivity correlates with current or future demand shocks subsumed in the error term ($\text{Cov}(TFPQ_{ft-4}, \varphi_{ft+\tau}) \neq 0$ for any $\tau \geq 0$). A possible concern is that our technical productivity measure could be sensitive to demand via adjustments in capacity utilization. However, even if this problem were to exist, for it to invalidate our identifying assumption, demand shocks must remain sufficiently serially correlated after removing a permanent component (absorbed into the firm-fixed effects) and industry trends (absorbed into the industry-time fixed effects). Put differently, our estimates could be biased if *both* (i) demand shocks move TFPQ and (ii) demand shocks display a quantitatively relevant persistence in the data. We discuss each issue in turn and conclude that neither seems to pose a threat to our identification scheme.

Variable capacity utilization—Data limitations prevent us from accounting for

capacity utilization in the construction of our TFPQ instrument for the full sample.³¹ Nevertheless, we have access to a measure of capacity utilization for a subsample of our dataset, which we use to perform the following empirical tests.

First, to assess whether our marginal cost measure is affected by past variation in capacity utilization, we regress mc_{ft}^n on capacity utilization lagged four quarters (as our instrument). We find a small and statistically insignificant elasticity (estimate 0.011, SE 0.052). We obtain a similarly small and statistically insignificant coefficient on lagged capacity utilization when this variable is included as a control in the first-stage regression for marginal cost (estimate 0.036, SE 0.025) and essentially no impact on the first-stage coefficient attached to our TFPQ instrument. Together, these tests suggest that the predictive power of our instrument for mc_{ft}^n does not appear to reflect variation in capacity utilization.

Second, we construct two “purified” TFPQ instruments that adjust inputs (capital and labor) for variation in capacity utilization and re-estimate our baseline model using the two utilization-adjusted TFPQ variables as instruments.³² The results are reported in Table 3. In Column (4), the TFPQ instrument is constructed by adjusting only the capital stock for utilization; in Column (5), both capital and labor are adjusted.³³ In both cases, the regressions pass the weak instrument and overidentification tests with test statistics similar to those of our baseline specification. More importantly, the estimates of the pass-through coefficients and NKPC slope are essentially unaffected, suggesting that unobserved demand-driven movements in (lagged) capacity utilization are unlikely to be a concern for identification.

Persistence of demand shocks—Next, we address the possibility that the unanticipated demand shocks in the error term might display persistent serial correlation. The difficulty with producing direct empirical tests that speak to this

³¹See Basu, Fernald and Kimball (2006) for a discussion on variable capacity utilization as a demand-driven source of variation in the Solow residual, as well as for procedures to account for it in the construction of TFP residuals when information on labor hours is available.

³²We discuss in Section OA.3 of the Supplemental Appendix a procedure to predict capacity utilization for observations when this information is not directly observable.

³³As further support for our TFP instrument, in Section OA.3 of the Supplemental Appendix, we present a falsification test that evaluates whether our purified TFP measure—adjusted to remove cyclical utilization effects—responds to high-frequency monetary or oil shocks when aggregated across firms. Reassuringly, the results show small and statistically insignificant effects on TFP, consistent with theoretical predictions.

issue is that these shocks are not directly observable. However, two pieces of evidence suggest that the serial correlation of demand shocks is weak and short-lived and, therefore, unlikely to introduce bias in our estimates.

A first piece of evidence comes from inspecting the regression residuals. The idea is that if demand shocks are serially correlated, so will be the residuals that are a function of such shocks. We find that the autocorrelation coefficient is economically small and marginally significant at a one-quarter lag (point estimate -0.09 , p-value 0.09) and it becomes economically and statistically insignificant already at the second-quarter lag (point estimate -0.01 , p-value 0.69). These results suggest that demand shocks (net of fixed effects) are likely highly transitory.

As a second exercise, we show in Column (6) of Table 3 that our results are robust to lagging our TFP instrument eight quarters, as opposed to four. Both the parameter estimates and the slope are essentially unchanged. To the extent that the serial correlation decays over time, these results suggest that the identifying variation in our instrument is unlikely capturing persistent demand shocks.

Other threats to identification—Other threats to identification include persistent measurement errors and forward-looking investments in productivity. Regarding the first point, the concern is that mismeasurement of output and input prices—if persistent—could create a mechanical correlation between the lagged TFPQ instrument, our marginal cost measure, and the regression residual. As is often the case, directly addressing this form of non-classical measurement error is challenging. However, the short-lived serial correlation in the residuals and the robustness with the further-lagged instrument discussed above speak to this issue and offer useful evidence to alleviate this concern.

Regarding the second point, one might worry that firms invest in productivity-enhancing activities today in anticipation of future demand changes. While this is possible, existing evidence indicates that technical productivity responds to such investments only after a significant lag. For example, Lenzu, Rivers and Tielens (2024) finds that technical productivity reacts to R&D expenditures with a three-year delay.

5 Aggregate inflation dynamics

In this section, we evaluate the ability of our estimated model to capture the aggregate time series of inflation for the Belgian manufacturing sector. To derive an expression for aggregate inflation, we use the equation for the price index in (9) and the equation for the reset price in (10). We then close the model by assuming that nominal marginal cost follows a random walk with drift.³⁴ We therefore obtain the following reduced-form expression for quarterly inflation (see Appendix A for derivations):

$$\pi_t = \tilde{\lambda} (mc_t^n - p_{t-1}) + \alpha + \theta u_t, \quad (19)$$

where α captures trend inflation and u_t is the aggregate cost-push shock. The parameter $\tilde{\lambda}$ in Equation (19) is a reduced-form slope, capturing the contemporaneous pass-through of fluctuations in aggregate real marginal cost—defined as nominal marginal cost scaled by the lagged price level—into inflation, taking into account the persistence of cost shocks. Similar to the cost-based NKPC slope (λ), this reduced-form slope is a functional equation in the primitive pricing parameters θ , Ω , and β . When evaluated at our estimated parameter values, $\tilde{\lambda}$ equals 0.22.

Iterating Equation (19) backward, we derive an expression for year-over-year inflation as a function of a four-quarter moving average of nominal marginal cost scaled by the price level:

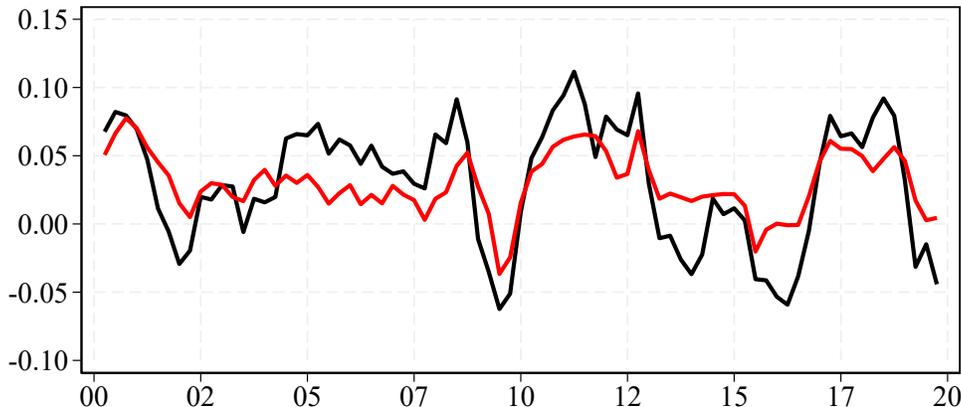
$$\pi_t^{y-y} = \sum_{\tau=0}^3 \tilde{\lambda} (1 - \tilde{\lambda})^\tau (mc_{t-\tau}^n - p_{t-4}) + \alpha^{y-y}. \quad (20)$$

The black line in Figure 1 plots year-over-year producer-price inflation (PPI) for the Belgian manufacturing sector. The red line in Figure 1 depicts a model-implied inflation series computed according to Equation (20). Through the lens of our model, the difference between the black and red lines is the component of inflation due to cost-push shocks, u_t .

As we can see, this parsimonious model effectively tracks the broad swings in

³⁴This assumption is consistent with the empirical evidence. To show this, we first construct aggregate marginal cost, mc_t^n , as a weighted average (with Törnqvist weights) of firm marginal costs mc_{ft}^n . Then, we regress mc_t^n on its one-quarter lag, instrumenting the latter with a two-quarter lag to reduce downward bias due to measurement error. We find that the estimated autoregressive coefficient is $\hat{\rho}^{mc} = 0.987$ (0.015), with Newey-West standard errors in brackets. Additionally, the Dickey-Fuller test does not reject the null hypothesis of unit root with $Z = -1.639$ and p-value = 0.463. Notice that this estimate is different, although consistent, from those in Table 2, as those estimates should be interpreted as the persistence of deviations from trend due to the inclusion of time fixed effects.

Figure 1: Aggregate inflation dynamics



Notes. This figure compares the inflation dynamics in the data to that implied by the model. The black line represents manufacturing PPI inflation obtained by aggregating price changes from PRODCOM. The red line is the model-implied manufacturing PPI inflation constructed as in Equation (20).

PPI inflation over our sample period. It accounts for nearly seventy percent of the variation in inflation ($R^2 = 0.68$) with a correlation coefficient of 0.8. It is noteworthy that the model captures the drop in inflation during the 2008 financial crisis and the sharp run-up in 2016 followed by a subsequent decline. Additionally, the model successfully captures the consistent decline in inflation from 2011 to 2016, although it does not fully capture the amplitude.

Note also that, within our framework, unobservable cost-push shocks account for a much smaller fraction of inflation volatility than is typically found in the quantitative literature.³⁵ In addition, we purposely chose to compare the data against the simplest possible framework. For instance, we did not account for other factors that could further rationalize inflation dynamics, such as lag-dependence in inflation, deviations from rational expectations, or imperfect information (see, e.g., Galí and Gertler 1999, Jørgensen and Lansing 2023, Gabaix 2020). Incorporating these forces in future research may further enhance our understanding of the relationship between inflation dynamics and real economic activity.

³⁵For example, in Primiceri, Schaumburg and Tambalotti (2006), cost-push shocks arising from variation in the desired price and wage markups account for about 70 percent of the volatility of inflation.

6 Reconciliation with the conventional NKPC

In this section, we reconcile our high slope estimates for the marginal cost-based curve with the low estimates of the conventional formulation. Following the literature, we make assumptions that allow us to establish a log-linear relationship between marginal costs, prices, and the output gap at the firm level. Under these assumptions, the output-based Phillips curve slope (κ) is the product of the marginal cost-based slope (λ) and the output elasticity of marginal cost (σ^y):

$$\kappa = \lambda \cdot \sigma^y.$$

We then develop two different identification approaches to estimate σ^y from micro-level data and retrieve κ . Consistent with the literature, we find a low output-based slope, which is explained by a low elasticity of marginal cost with respect to changes in output.

6.1 Marginal cost and the output gap at the firm level

To begin, we derive a log-linear relation between firm-level marginal cost and the output gap, similar to the one typically assumed at the aggregate level to obtain the conventional formulation of the Phillips curve. In doing so, we allow for general equilibrium effects that affect firms' costs through the impact of labor demand on wages (see e.g., Galí 2015).

We assume real wages are determined in general equilibrium at the industry level. Accordingly, we can express firm-level log real marginal cost, mc_{ft} , as a function of the industry real wage, $w_{it} - p_t$, and firm-level marginal product of labor, mpn_{ft} :

$$mc_{ft} = (w_{it} - p_t) - mpn_{ft}.$$

Next, as in the benchmark NK model, we assume that industry real wages are flexible and increasing in current industry output, $(w_{it} - p_t) = \sigma^w y_{it}$, with elasticity σ^w . The presence of industry output captures the general equilibrium feedback of aggregate demand on firms' marginal cost. We assume that labor supply is industry-specific, which implies that σ^w is independent of whether industry output is driven by aggregate or industry-specific shocks. In addition, we assume the firm's marginal product of labor depends positively on firm productivity, z_{ft} —which may contain

both an aggregate and an idiosyncratic component—and inversely on firm output, y_{ft} , with elasticity given by SR-RTS (ν) (homogeneous across firms and time-invariant, for simplicity). We can thus express firm-level real marginal cost as:

$$mc_{ft} = \sigma^w y_{it} + z_{ft} + \nu y_{ft}.$$

Without loss of generality, we write the logarithm of a firm's output as the sum of log industry output and idiosyncratic supply (ϵ_{ft}^s) and demand (ϵ_{ft}^d) shocks:

$$y_{ft} = y_{it} + \epsilon_{ft}^s + \epsilon_{ft}^d.$$

The supply and demand shocks are linear in the idiosyncratic component of firm's productivity, z_{ft} , and in the firm's idiosyncratic demand shifter, φ_{ft} , respectively.

Finally, we define the natural levels of industry and firm output, y_{it}^* and y_{ft}^* . As is conventional, we define y_{it}^* as the equilibrium level of output under flexible prices and wages, where the desired markup is constant. The natural level, y_{ft}^* , is similarly defined, also taking into account idiosyncratic supply shocks:

$$y_{ft}^* := y_{it}^* + \epsilon_{ft}^s.$$

Under these assumptions, we can express the deviation of real firm marginal cost from steady state, \widehat{mc}_{ft} , as a constant-elasticity function of the firm-level output gap:

$$\widehat{mc}_{ft} = \sigma^y (y_{ft} - y_{ft}^*) - \sigma^w \epsilon_{ft}^d, \quad (21)$$

where the coefficient $\sigma^y := \sigma^w + \nu$ represents the elasticity of marginal cost with respect to the output gap. The error term $\sigma^w \epsilon_{ft}^d$ accounts for the fact that wages depend on the industry component (but not on the idiosyncratic component) of firm demand.

To derive a pricing equation in terms of output that allows us to identify σ^y and therefore κ , we rearrange Equation (21) and substitute for mc_{ft}^n into Model A introduced in Section 3.2. As in Model C, we then postulate that nominal output and the competitors' price index, in deviations from their trends, follow first-order autoregressive processes. This leads to an empirical model that directly relates firm-level prices and output:

$$p_{ft} = \Psi^y \cdot \sigma^y y_{ft}^n + \Psi^p p_{it}^{-f} + \theta p_{ft-1} + \alpha_f + \alpha_{s \times t} + \epsilon_{ft}^p, \quad (\text{Model D})$$

where Ψ^y and Ψ^p depend on the persistence of shocks to output and prices:

$$\Psi^y := (1 - \theta)(1 - \Omega) \frac{1 - \beta\theta}{1 - \beta\theta\rho^y} \quad \text{and} \quad \Psi^p := (1 - \theta)\Omega \frac{1 - \beta\theta}{1 - \beta\theta\rho^p}.$$

Note that, in contrast to our cost-based pass-through regressions, the error term in Model D, $\varepsilon_{ft}^p := (1 - \sigma^w)\varepsilon_{ft}^d - \sigma^y y_{ft}^*$, captures both the firm's idiosyncratic demand shocks and idiosyncratic supply shocks, which affect the firm's (unobservable) natural level of output.

A complementary way to identify σ^y is to rewrite Equation (21) in terms of nominal marginal cost. Then, taking first differences, we obtain:

$$\Delta mc_{ft}^n = \sigma^y \Delta y_{ft}^n + \alpha_f + \alpha_{s \times t} + \varepsilon_{ft}^{mc}, \quad (\text{Model E})$$

where the price level is absorbed into the sector-by-time fixed effects and the error term $\varepsilon_{ft}^{mc} := -\sigma^w \Delta \varepsilon_{ft}^d - \sigma^y \Delta y_{ft}^*$. Unlike Model D, which directly maps into the dynamic pass-through framework developed in Section 3.2, Model E allows us to directly estimate the elasticity of interest from contemporaneous changes of marginal cost and output.

6.2 Identification of σ^y and κ

We take Model D and Model E to the data to identify the elasticity σ^y and thereby recover the slope of the output-based NKPC, κ . To do so, we use firms' nominal value added (revenues minus costs of intermediate inputs) as a measure of firm-level nominal output, y_{ft}^n . The identification of σ^y requires us to isolate variation in y_{ft}^n that is orthogonal to both the firm-level natural level of output and idiosyncratic demand shocks, as both enter the error terms ε_{ft}^{mc} and ε_{ft}^p .

To tackle this issue, we follow the literature that estimates output and unemployment gap-based NKPCs by exploiting shifts in aggregate demand.³⁶ We recover industry sensitivities to high-frequency monetary policy shocks. We then construct a Bartik-style instrument to improve the power of the aggregate shocks while also allowing us to include sector-by-time fixed effects to control for higher-level movements in demand.

For each industry i , we estimate the sensitivity to an aggregate demand shock

³⁶See, e.g., Barnichon and Mesters (2020), McLeay and Tenreyro (2020), and Hazell et al. (2022).

Table 4: Estimates of the output-based slope

<i>Panel A: Structural estimates</i>		
	<i>Model D</i>	<i>Model E</i>
	(1)	(2)
σ^y	0.406 (0.099)	0.112 (0.026)
ρ^y	0.880 (0.039)	
ρ^p	0.912 (0.004)	
<i>Panel B: Slope of the output-based Phillips curve</i>		
κ	0.021 (0.005)	0.006 (0.001)
<i>Test statistics</i>		
Cragg-Donald F	191.954	484.808
Kleibergen-Paap F	23.385	62.453

Notes. This table presents the empirical estimates of models D and E. The output-based NKPC slope, κ , is obtained as the product of the estimate of σ^y and the estimate of λ from Model A. All models are estimated using the complete sample. Robust standard errors (reported in parenthesis) are clustered at the industry-by-time level. Models D and E include sector-by-time fixed effects and firm fixed effects.

by projecting firm-level nominal value-added output on the monetary policy shock.³⁷ We lag the shock to reflect its delayed effect on real activity, which peaks at four quarters:

$$y_{ft}^n = \alpha_f + \psi_i MS_{t-4} + \epsilon_{ft}^m.$$

We then obtain our demand-side instrument by interacting the aggregate money shock with the estimated sensitivity: $y_{ft}^{IV} := \hat{\psi}_i \cdot MS_{t-4}$. This shifter is orthogonal to both aggregate and idiosyncratic supply shocks as well as idiosyncratic demand shocks. However, it picks up movements in firms' output due to general equilibrium effects as it captures common demand shocks at the industry level. Moreover, unlike the aggregate monetary policy surprises, it is a powerful instrument as it leverages

³⁷Monetary policy shocks are constructed following Gürkaynak, Sack and Swanson (2005) as the log change in the price of overnight index swaps within a narrow window around ECB monetary policy announcements. The time series of aggregate money shocks (MS) are taken from Altavilla et al. (2019).

variation in the high-frequency surprises interacted with the cross-industry response to these shocks.

We estimate Models D and E via GMM. For Model D, we calibrate θ and Ω to our baseline estimates. We then estimate the pass-through equation jointly with the AR(1) dynamics for output and competitors' prices. Table 4 presents the resulting estimates for the output elasticity of marginal cost and the implied estimates for the slope of the output-based Phillips curve. For Model D, we find a value of $\sigma^y = 0.406$ and $\kappa = 0.021$. For Model E, we find even smaller estimates, $\sigma^y = 0.112$ and $\kappa = 0.006$.

The low sensitivity of marginal cost to movements in output is consistent, for example, with a high degree of wage rigidity and an economy not operating too close to full capacity. These low estimates of κ corroborate the findings in previous literature, which conclude that the slope of the output-based and unemployment-based NKPC appears to be flat. Indeed, our point estimates overlap closely with those in Rotemberg and Woodford (1997) and Hazell et al. (2022) who find slope coefficients of 0.024 and 0.006.

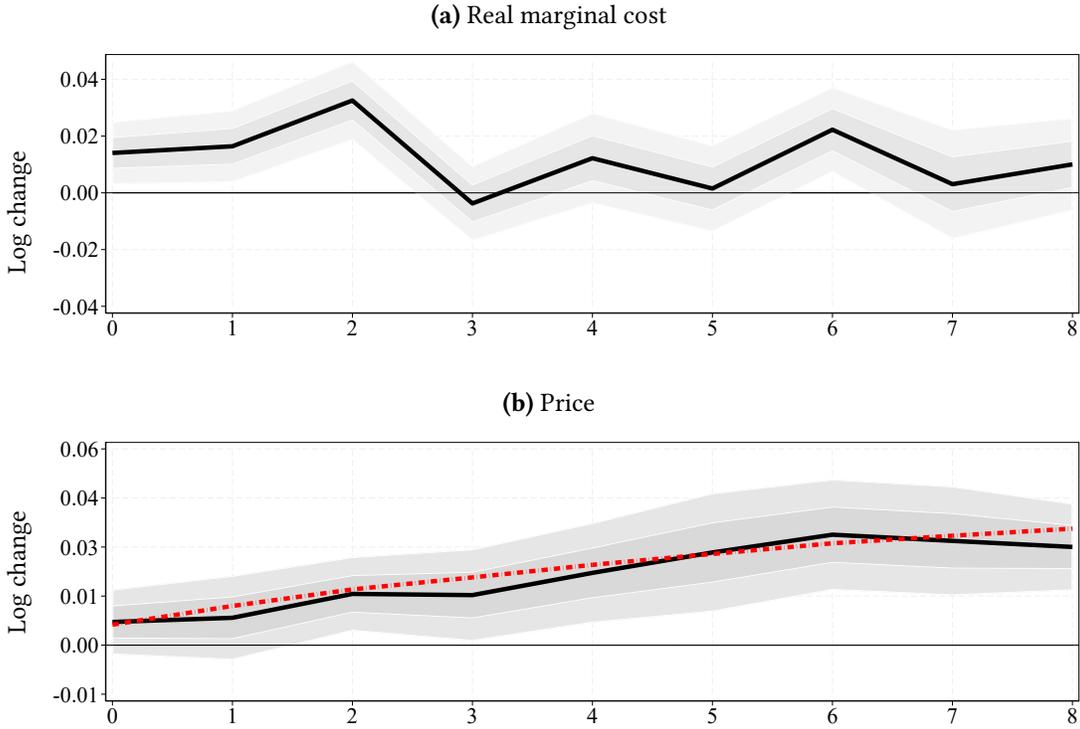
In sum, our analysis suggests that the pass-through from marginal costs to prices is high, as the micro-estimates indicate, but the flatness of the conventional NKPC is likely due to a low sensitivity of marginal cost to output. These considerations call for further theoretical and empirical work to elucidate the relationship between output and marginal cost, especially considering the possibility that the elasticity connecting the two could be time-varying and possibly nonlinear.

7 Tracing the effect of supply shocks on inflation

In this section, we study the pass-through of identified supply shocks to inflation. This exercise serves two purposes: it illustrates the advantage of a marginal cost-based Phillips curve for characterizing the transmission of supply shocks to inflation; it also provides an alternative estimate of the NKPC slope via impulse-response matching (e.g., Barnichon and Mesters 2020), validating the one obtained from our micro-level pass-through regressions.

Tracing the effect of supply shocks—The discussion in the previous section highlighted that it is challenging to assess the impact of such shocks on inflation using

Figure 2: Dynamic effects of oil shocks on real marginal costs and prices



Notes. This figure shows the IRF of real marginal cost and price level to aggregate oil shocks estimated via local linear projections. The dark (light) gray shaded areas are 68 (95) percent confidence bands obtained from Newey-West standard errors with four quarters of correlation. In Panel B, the red line represents the model-based IRF of prices, calculated by feeding the path of real marginal cost into a New Keynesian model featuring a cost-based NKPC. The x-axis measures quarters since the aggregate oil shock.

the output-based NKPC curve, without relying on a fully specified macroeconomic model to derive the natural level of output. The cost-based NKPC curve does not suffer from this. The impact of supply shocks on marginal cost is measurable, which implies that the cost-based NKPC can be used to quantify the pass-through of supply shocks to inflation.

We consider oil shocks as prototype supply shocks. Following Känzig (2021), we measure oil shocks as unexpected movements in oil price futures the day after OPEC meetings and estimate the empirical impulse response functions (IRF) of aggregate real marginal cost and price level to these identified shocks via local linear projections:

$$x_{ft+h} - x_{ft-1} = a_f + b_h^x OS_{t-1} + \epsilon_{ft+h},$$

for $x \in \{mc^r, p\}$ and $h = 0, \dots, 8$ quarters. We normalize the oil shock so that b_1^x represents the effect on impact of a one-standard deviation shock to oil prices (a 15.7 percent increase in Brent crude oil price). As in our pass-through regressions, observations are weighted using Törnqvist sales weights.

The estimates are reported in Figure 2. On average, firms' real marginal costs rise by approximately 1.5 to 3 percent within the first three quarters in response to a one-standard-deviation shock to oil prices, before gradually returning to their pre-shock level (Panel A). The oil shock also has a significant effect on firms' prices (Panel B, black line). Consistent with the presence of nominal rigidities, the price response is delayed but persistent, peaking at approximately 3 percent increase after six quarters.

Estimation via impulse-response matching—These empirical results allow us to validate our estimate of the slope of the NKPC. We feed the path of real marginal cost shocks (with perfect foresight) to an NK model featuring a cost-based Phillips curve. We then estimate the slope of the NKPC by minimizing the distance between the empirical impulse responses of prices, $\{\hat{b}_h^p\}_{h=0}^8$, and the corresponding model's impulse-responses of prices, $\{g_h^p(\lambda)\}_{h=0}^8$:

$$\lambda^{IRF} = \arg \min_{\lambda} (\hat{\mathbf{b}}^p - \mathbf{g}^p(\lambda))' W (\hat{\mathbf{b}}^p - \mathbf{g}^p(\lambda)),$$

where the weighting matrix W is a diagonal matrix whose elements are the reciprocals of the variances of empirical IRFs estimates. Standard errors are calculated using the delta method (Mertens and Ravn 2011).

The model accurately reflects the dynamic effects of the shocks on the price level, both in terms of magnitude and persistence (Figure 2 Panel B, red dotted line). The model's impulse responses consistently lie within the confidence bands of those estimated in the data. Importantly, the slope of the cost-based NKPC that allows us to match the IRFs is $\lambda^{IRF} = 0.042$ (SE 0.005), which is close to and within the confidence bands of the estimate obtained from our micro-level pass-through model ($\hat{\lambda} = 0.052$).

8 Concluding remarks

We use disaggregated data to identify the slope of the primitive form of the New Keynesian Phillips curve, which features marginal cost as the relevant measure

of economic activity. Our findings reveal a high pass-through of marginal cost into prices, supported by both the analysis of the microdata and by the ability of the marginal cost-based Phillips curve to track aggregate inflation dynamics. Additionally, we show that in the pre-pandemic period, a low elasticity of marginal cost to output helps reconcile the low sensitivity of aggregate inflation to output found in the literature with the high pass-through of marginal cost documented in this paper.

Our analysis is based on data from manufacturing firms, where output and inflation volatility are typically higher than in other sectors. However, because we are measuring an elasticity with respect to real activity, our results are likely relevant for other sectors as well, such as retail. These sectors tend to exhibit sluggish output movements but also more stable inflation dynamics.

The ability of the marginal cost-based NKPC to capture inflation dynamics should not be of narrow interest. Indeed, most quantitative macroeconomic (DSGE) models feature some form of wage rigidity to fit the data, which requires the cost-based formulation of the Phillips curve to capture the transmission of shocks into prices.³⁸ Our analysis suggests that there is room for improvement in different directions.

A first direction for improvement is the measure of marginal cost. DSGE models typically feature labor as the only variable input. However, accounting for the cost of intermediate inputs is pivotal to the success of our empirical measure, as intermediates represent both the largest and most volatile component of production costs. To this point, research has shown that intermediate goods price shocks were among the most important drivers of the recent inflation surge (e.g., Di Giovanni et al. 2022). Developing models that incorporate how intermediates factor into firms' costs is an active and important direction for future research (see, e.g., Rubbo 2023).

Secondly, further research is needed to understand the primitive drivers of the elasticity of marginal cost with respect to output, both at the micro level and at the aggregate level, and how it may evolve over time. Our findings indicate that this elasticity is low during normal times, likely due to two factors working in combination. First, at the micro level, firms' marginal cost schedules appear relatively insensitive to changes in output, reflecting near-constant short-run returns to scale.

³⁸See, e.g., Chapter 6 in Galí (2015) and the references therein.

Second, wage rigidity likely mutes general equilibrium effects at the aggregate or industry level. However, the recent inflation surge has demonstrated that the sensitivity of marginal cost to output can rise rapidly and substantially when the economy faces large shocks. Such shocks may push the economy up against its capacity constraints, whether due to labor market tightness or bottlenecks in the supply of intermediate goods (e.g., Boehm and Pandalai-Nayar 2022 and Comin, Johnson and Jones 2023). The net effect is a sharp increase in marginal cost due to the rapid increase in input prices in the face of high aggregate demand relative to capacity.

Finally, the elasticity of inflation with respect to marginal cost can itself vary in response to large aggregate shocks, as it depends on the frequency of price adjustments. Our evidence indicates that this frequency was stable during the pre-pandemic period, consistent with the Calvo model and implying a stable elasticity. However, recent data shows that the price adjustment frequency rose significantly throughout the post-pandemic inflation surge, leading to a jump in the elasticity of inflation with respect to marginal cost. Our companion paper, Gagliardone et al. (2025), suggests that over this recent period, state-dependent pricing models more accurately capture both micro- and macro cost-price dynamics.

References

- Altavilla, Carlo, Luca Brugnolini, Refet S Gürkaynak, Roberto Motto, and Giuseppe Ragusa.** 2019. “Measuring euro area monetary policy.” *Journal of Monetary Economics*, 108: 162–179.
- Amiti, Mary, Oleg Itskhoki, and Jozef Konings.** 2019. “International shocks, variable markups, and domestic prices.” *The Review of Economic Studies*, 86(6): 2356–2402.
- Atkeson, Andrew, and Ariel Burstein.** 2008. “Pricing-to-market, trade costs, and international relative prices.” *American Economic Review*, 98(5): 1998–2031.
- Auclert, Adrien, Rodolfo Rigato, Matthew Rognlie, and Ludwig Straub.** 2024. “New Pricing Models, Same Old Phillips Curves?” *The Quarterly Journal of Economics*, 139(1): 121–186.

- Barnichon, Regis, and Geert Mesters.** 2020. “Identifying modern macro equations with old shocks.” *The Quarterly Journal of Economics*, 135(4): 2255–2298.
- Basu, Susanto.** 1995. “Intermediate Goods and Business Cycles: Implications for Productivity and Welfare.” *American Economic Review*, 85(3): 512.
- Basu, Susanto, John G Fernald, and Miles S Kimball.** 2006. “Are technology improvements contractionary?” *American Economic Review*, 96(5): 1418–1448.
- Beraja, Martin, Erik Hurst, and Juan Ospina.** 2019. “The aggregate implications of regional business cycles.” *Econometrica*, 87(6): 1789–1833.
- Boehm, Christoph E, and Nitya Pandalai-Nayar.** 2022. “Convex supply curves.” *American Economic Review*, 112(12): 3941–3969.
- Calvo, Guillermo A.** 1983. “Staggered prices in a utility-maximizing framework.” *Journal of Monetary Economics*, 12(3): 383–398.
- Comin, Diego A, Robert C Johnson, and Callum J Jones.** 2023. “Supply chain constraints and inflation.” National Bureau of Economic Research.
- Di Giovanni, Julian, Sebnem Kalemli-Özcan, Alvaro Silva, and Muhammed A Yildirim.** 2022. “Global supply chain pressures, international trade, and inflation.” National Bureau of Economic Research.
- Engle, Robert F, and Clive WJ Granger.** 1987. “Co-integration and error correction: representation, estimation, and testing.” *Econometrica: journal of the Econometric Society*, 251–276.
- Gabaix, Xavier.** 2020. “A behavioral New Keynesian model.” *American Economic Review*, 110(8): 2271–2327.
- Gagliardone, Luca, Mark Gertler, Simone Lenzu, and Joris Tielens.** 2025. “Micro and macro cost-price dynamics in normal times and during inflation surges.” *Working paper*.
- Galí, Jordi.** 2015. *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- Galí, Jordi, and Mark Gertler.** 1999. “Inflation dynamics: A structural econometric analysis.” *Journal of Monetary Economics*, 44(2): 195–222.
- Goldberg, Pinelopi Koujianou, and Frank Verboven.** 2001. “The evolution of price dispersion in the European car market.” *The Review of Economic Studies*, 68(4): 811–848.

- Gürkaynak, Refet S, Brian P Sack, and Eric T Swanson.** 2005. “Do actions speak louder than words? The response of asset prices to monetary policy actions and statements.” *International Journal of Central Banking*, 1(1): 55–93.
- Hazell, Jonathon, Juan Herreno, Emi Nakamura, and Jón Steinsson.** 2022. “The slope of the Phillips Curve: evidence from US states.” *The Quarterly Journal of Economics*, 137(3): 1299–1344.
- Hooper, Peter, Frederic S Mishkin, and Amir Sufi.** 2020. “Prospects for inflation in a high pressure economy: Is the Phillips curve dead or is it just hibernating?” *Research in Economics*, 74(1): 26–62.
- Hottman, Colin J, Stephen J Redding, and David E Weinstein.** 2016. “Quantifying the sources of firm heterogeneity.” *The Quarterly Journal of Economics*, 131(3): 1291–1364.
- Jørgensen, Peter Lihn, and Kevin J Lansing.** 2023. “Anchored inflation expectations and the slope of the phillips curve.”
- Känzig, Diego R.** 2021. “The macroeconomic effects of oil supply news: Evidence from OPEC announcements.” *American Economic Review*, 111(4): 1092–1125.
- Kimball, Miles S.** 1995. “The quantitative analytics of the basic neomonetarist model.” *Journal of Money, Credit, and Banking*, 27(4): 1241–1277.
- Lenzu, Simone, David Rivers, and Joris Tielens.** 2024. “Financial Shocks, Productivity, and Prices.” *Working paper*.
- Lewis, Daniel J, and Karel Mertens.** 2022. “Dynamic Identification Using System Projections and Instrumental Variables.”
- Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H Stock.** 2014. “Empirical evidence on inflation expectations in the New Keynesian Phillips Curve.” *American Economic Journal: Journal of Economic Literature*, 52(1): 124–188.
- McLeay, Michael, and Silvana Tenreyro.** 2020. “Optimal inflation and the identification of the Phillips curve.” *NBER Macroeconomics Annual*, 34(1): 199–255.
- Mertens, Karel, and Morten O Ravn.** 2011. “Understanding the aggregate effects of anticipated and unanticipated tax policy shocks.” *Review of Economic dynamics*, 14(1): 27–54.
- Nakamura, Emi, and Dawit Zerom.** 2010. “Accounting for incomplete pass-through.” *The Review of Economic Studies*, 77(3): 1192–1230.
- Nakamura, Emi, and Jón Steinsson.** 2008. “Five facts about prices: A reevaluation of menu cost models.” *The Quarterly Journal of Economics*, 123(4): 1415–1464.

- Primiceri, Giorgio, Ernst Schaumburg, and Andrea Tambalotti.** 2006. “Intertemporal disturbances.” *NBER Working Paper No. 12243*.
- Rotemberg, Julio J, and Michael Woodford.** 1997. “An optimization-based econometric framework for the evaluation of monetary policy.” *NBER Macroeconomics Annual*, 12: 297–346.
- Rubbo, Elisa.** 2023. “Networks, Phillips curves, and monetary policy.” *Econometrica*, 91(4): 1417–1455.
- Sbordone, Argia M.** 2002. “Prices and unit labor costs: a new test of price stickiness.” *Journal of Monetary Economics*, 49(2): 265–292.
- Wang, Olivier, and Iván Werning.** 2022. “Dynamic oligopoly and price stickiness.” *American Economic Review*, 112(8): 2815–49.
- Woodford, Michael.** 2011. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Anatomy of the Phillips Curve: Micro Evidence and Macro Implications

L. Gagliardone M. Gertler S. Lenzu J. Tielens

Appendix

A Derivations of aggregate inflation dynamics

Consider the system of equations given by the aggregate price index (Equation (9)), the aggregate reset price (Equation (10), ignoring the intercept), and a random walk dynamic for aggregate marginal cost:

$$\begin{aligned} p_t &= (1 - \theta)p_t^o + \theta p_{t-1}, \\ p_t^o &= (1 - \beta\theta)((1 - \Omega)mc_t^n + \Omega p_t) + \beta\theta \mathbb{E}_t p_{t+1}^o + \frac{\theta}{1 - \theta} u_t, \\ mc_t^n &= mc_{t-1}^n + \varepsilon_t^{mc}. \end{aligned} \quad (\text{A.1})$$

We use the method of undetermined coefficients to guess and verify a solution for the system (A.1), which takes the following form:

$$\begin{aligned} p_t^o - p_{t-1} &= \Xi(mc_{t-1}^n + \varepsilon_t^{mc} - p_{t-1}) + \frac{\theta}{1 - \theta} u_t \\ \pi_t &= \tilde{\lambda}(mc_{t-1}^n + \varepsilon_t^{mc} - p_{t-1}) + \theta u_t, \end{aligned}$$

where Ξ and $\tilde{\lambda}$ are the coefficients to be determined. Subtracting p_{t-1} , the equations become:

$$\begin{aligned} p_t^o - p_{t-1} &= (1 - \beta\theta)((1 - \Omega)(mc_t^n - p_{t-1}) + \Omega \pi_t) + \beta\theta \mathbb{E}_t (p_{t+1}^o - p_t + \pi_t) + \frac{\theta}{1 - \theta} u_t, \\ \pi_t &= (1 - \theta)(p_t^o - p_{t-1}) + \theta u_t \end{aligned}$$

The expectation is computed as:

$$\mathbb{E}_t(p_{t+1}^o - p_t) = \Xi \mathbb{E}_t(mc_t^n + \varepsilon_t^{mc} - p_t) = (1 - \tilde{\lambda})\Xi(mc_t^n - p_{t-1})$$

Plugging the guessed solution into the system gives the following restrictions on the parameters:

$$\begin{aligned}\Xi &= (1 - \beta\theta)(1 - \Omega + \Omega\tilde{\lambda}) + \beta\theta((1 - \tilde{\lambda})\Xi + \tilde{\lambda}), \\ \tilde{\lambda} &= (1 - \theta)\Xi.\end{aligned}$$

We select the solution of system (A.1) characterized by having exactly one eigenvalue larger than one in modulus. Using our structural estimates, $\theta = 0.7$ and $\Omega = 0.52$ (median values across the different models in Table 2), we obtain that an aggregate reduced-form pass-through coefficient of $\tilde{\lambda} = 0.22$.

Replacing $p_t = \tilde{\lambda}mc_t^n + (1 - \tilde{\lambda})p_{t-1} + \theta u_t$ in the guessed solution, adding back the intercept, and rearranging we obtain Equation (19).

Anatomy of the Phillips Curve: Micro Evidence and Macro Implications

Supplemental Appendix

L. Gagliardone M. Gertler S. Lenzu J. Tielens

OA.1 Theory derivations

In this section, we derive the key equations presented in Section 1 of the paper. We first derive the markup functions under two prominent frameworks featuring imperfect competition: Dynamic oligopoly with nested CES preferences and Kimball preferences. We then present the aggregation steps followed to derive the cost-based NKPC.

OA.1.1 Derivation of the markup function

Dynamic oligopoly with nested CES preferences—Assume that there is a continuum of industries (indexed by i) and a finite number of firms N within each industry. Each firm is indexed by f (or j). Within each industry, firms compete à la Bertrand. In this environment, the price index for each industry P_{it} and the aggregate price index P_t are defined, respectively, as:

$$P_{it} := \left(\frac{1}{N} \sum_{f=1}^N (\varphi_{fit} P_{fit})^{1-\gamma} \right)^{\frac{1}{1-\gamma}} ; \quad P_t := \left(\int_{i \in I} (\varphi_{it} P_{it})^{1-\sigma} di \right)^{\frac{1}{1-\sigma}},$$

where φ_{fit} is a firm-specific relative demand shifter (firm appeal), and φ_{it} is an industry-specific demand shifter (relative across industries). In what follows, the subscript i is dropped when redundant, and we normalize the steady-state price level to simplify the notation. The demand function for firm $f \in \mathcal{F}_i$ takes a nested CES form, with the elasticity of substitution across industries $\sigma > 1$ and the elasticity of substitution within industries $\gamma > \sigma$:

$$\mathcal{D}_{ft+\tau} = \left(\frac{\varphi_{ft+\tau} P_{ft}^o}{\varphi_{it+\tau} P_{it+\tau}} \right)^{-\gamma} \left(\frac{\varphi_{it+\tau} P_{it+\tau}}{P_{t+\tau}} \right)^{-\sigma} Y_{t+\tau}. \quad (\text{OA.1})$$

Firms internalize the dynamic effect of their choices on the industry price index and on industry demand. Therefore, the residual elasticity of demand faced by firm f takes the following form:

$$\epsilon_{ft+\tau} := -\frac{\partial \ln \mathcal{D}_{ft+\tau}}{\partial \ln P_{ft}^o} = \gamma - (\gamma - \sigma) \frac{\partial p_{it+\tau}}{\partial p_{ft}^o}. \quad (\text{OA.2})$$

We can further characterize the determinants of the residual demand elasticity. First, the price index of competitors of firm f is defined as:

$$P_{it}^{-f} := \left(\frac{1}{N-1} \sum_{j \neq f}^{N-1} (\varphi_{jit} P_{jit})^{1-\gamma} \right)^{\frac{1}{1-\gamma}}.$$

It follows that $P_{it}^{1-\gamma} = \frac{N-1}{N} (P_{it}^{-f})^{1-\gamma} + \frac{1}{N} (\varphi_{ft} P_{ft}^o)^{1-\gamma}$. Next, we can express the derivative of the price index in period $t + \tau$ with respect to the firms' reset price in period t as follows:

$$\frac{\partial P_{it+\tau}}{\partial P_{ft}^o} = P_{it+\tau}^\gamma \left[\left(\frac{N-1}{N} \right) (P_{it+\tau}^{-f})^{-\gamma} \frac{\partial P_{it+\tau}^{-f}}{\partial P_{ft}^o} + \left(\frac{1}{N} \right) (\varphi_{ft})^{1-\gamma} (P_{ft}^o)^{-\gamma} \right].$$

Multiplying both sides by $\frac{P_{ft}^o}{P_{it+\tau}}$, and defining the competitors' reaction function $\zeta_{ft+\tau} := \frac{\partial P_{it+\tau}^{-f}}{\partial P_{ft}^o}$, we obtain:

$$\begin{aligned} \frac{\partial p_{it+\tau}}{\partial p_{ft}^o} &= \zeta_{ft+\tau} \left(\frac{N-1}{N} \right) \left(\frac{P_{it+\tau}^{-f}}{P_{it+\tau}} \right)^{1-\gamma} + \frac{1}{N} \left(\frac{\varphi_{ft+\tau} P_{ft}^o}{P_{it+\tau}} \right)^{1-\gamma} \\ &= \zeta_{ft+\tau} (1 - s_{ft+\tau}) + s_{ft+\tau}, \end{aligned}$$

where $s_{ft+\tau} := \frac{1}{N} \frac{P_{ft}^o \mathcal{D}_{ft+\tau}}{P_{it} Y_{it+\tau}} = \frac{1}{N} \left(\frac{\varphi_{ft+\tau} P_{ft}^o}{P_{it+\tau}} \right)^{1-\gamma}$ denotes the within-industry revenue share of firm f , and $Y_{it+\tau} := \varphi_{it+\tau}^{\gamma-\sigma} \left(\frac{P_{it+\tau}}{P_{t+\tau}} \right)^{-\sigma} Y_{t+\tau}$ is the industry demand. Replacing the expression for $\frac{\partial p_{it+\tau}}{\partial p_{ft}^o}$ into Equation (OA.2), we have that the within-industry elasticity of demand faced by firm f is given by:

$$\epsilon_{ft+\tau} = \gamma - (\gamma - \sigma) \left[\zeta_{ft+\tau} (1 - s_{ft+\tau}) + s_{ft+\tau} \right]. \quad (\text{OA.3})$$

The intuition behind this expression is straightforward. The stronger the reaction of competitors to a firm's price change—captured by $\zeta_{ft+\tau}$ —the lower the residual elasticity of demand. A low residual elasticity of demand, in turn, implies that the firm can sustain a higher markup in equilibrium. This result mirrors the one in the

dynamic oligopoly environment in Wang and Werning (2022) and it nests a number of static environments featuring imperfectly competitive firms. In a static oligopoly, $\epsilon_{ft+\tau} = 0$ for $\tau > 0$. In Atkeson and Burstein (2008) static Nash oligopoly, $\epsilon_{ft+\tau} = 0$ for $\tau > 0$ and $\zeta_{ft+\tau} = 0$ for all τ s. Under monopolistic competition, $N \rightarrow \infty$, which implies $\zeta_{ft+\tau} \rightarrow 0$ and $s_{ft+\tau} \rightarrow 0$.

We now use this result to derive the expression for the log-linearized desired markup in Equation (7) in the paper. As is standard, we log-linearize around a symmetric Nash steady state (Atkeson and Burstein, 2008).¹ Log-linearizing the elasticity in (OA.3) around the steady state, we obtain the steady state residual demand elasticity:

$$\epsilon = \gamma - (\gamma - \sigma) \frac{1}{N},$$

which corresponds to the expression in Atkeson and Burstein (2008). In this model, the desired markup is given by the Lerner index:

$$\mu_{ft+\tau} := \ln(\epsilon_{ft+\tau} / (\epsilon_{ft+\tau} - 1)).$$

Log-linearizing this expression and substituting the expression for steady-state residual demand elasticity, we obtain the expression for the log-linearized desired markup (in deviation from steady state) in Equation (7):

$$\mu_{ft+\tau} - \mu_f = -\Gamma \left(p_{ft}^o - p_{ft+\tau}^{-f} \right) + u_{ft+\tau}^\mu,$$

where $\mu_f = \mu$ for all f s in the symmetric steady state, $\Gamma := \frac{(\gamma - \sigma)(\gamma - 1)}{\epsilon(\epsilon - 1)} \frac{N - 1}{N} > 0$ denotes the markup elasticity with respect to prices, and the firm-specific markup shock:

$$u_{ft}^\mu := -\frac{(\gamma - \sigma)(\gamma - 1)}{\epsilon(\epsilon - 1)} \ln \varphi_{ft} + \frac{\gamma - \sigma}{\epsilon(\epsilon - 1)} \frac{N - 1}{N} \zeta_{ft}, \quad (\text{OA.4})$$

captures residual variation in the markup that depends on the demand shifters and changes in the slope of competitors' reaction function.

¹ The symmetry assumption is standard in the literature (e.g., Midrigan (2011) and Alvarez and Lippi (2014)), which eases the notation but is largely immaterial for our estimation purposes. Relaxing this assumption would imply firm-specific steady-state demand elasticities, ϵ_f . In this case, the estimates of the parameters of our pricing equations should be interpreted as average across firms. The assumption of Nash steady state, also standard in the literature, implies that $\zeta_{j,\tau} = 0$ in the steady state for all j s and τ s. This comes with some loss of generality, but two points can be made. First, as shown by Wang and Werning (2022), one can write a "behavioral" model with the weaker assumption that $\mathbb{E}\{\zeta_{j,\tau}\} = 0$ for all j s and τ s, which delivers, under specific values for the elasticities σ and γ , a pass-through of shocks to marginal cost into prices qualitatively similar to that produced by the Nash model. Second, these considerations also apply to our empirical analysis, as we directly estimate the parameters (Γ , in particular) rather than the underlying elasticities.

Finally, log-linearizing the industry price index and ignoring constants, we obtain:

$$p_{it} = \frac{N-1}{N} p_{it}^{-f} + \frac{1}{N} (\ln \varphi_{ft} + p_{ft}^o).$$

Substituting this expression in Equation (6) for the markup and rearranging we obtain the dynamic pricing equation in Equation (8):

$$p_{ft}^o = (1 - \beta\theta) \mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} (\beta\theta)^\tau \left((1 - \Omega)(mc_{ft+\tau}^n + \mu_f) + \Omega p_{it+\tau}^{-f} + (1 - \Omega) u_{ft+\tau}^\mu \right) \right\}, \quad (\text{OA.5})$$

where, as in the paper, $\Omega := \frac{\Gamma}{1+\Gamma}$. This parameter denotes the relative weight on the competitors' price index (p_{it}^{-f}) and captures the importance of strategic complementarities. When Ω is close to one, firms are not strategic and only look at their marginal cost when resetting prices. In particular, $\Omega \rightarrow 0$ as $N \rightarrow \infty$, which is the monopolistic competition case. The error term in Equation (8) is:

$$u_{ft} := (1 - \beta\theta)(1 - \Omega) \mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} (\beta\theta)^\tau u_{ft+\tau}^\mu \right\}, \quad (\text{OA.6})$$

which is therefore a firm-specific demand shock that depends on the expected discounted future demand shocks.

Monopolistic competition with Kimball preferences—Assume that the industry output Y_{it} is produced by a unitary measure of perfectly competitive firms using a bundle of differentiated intermediate inputs Y_{ft} , $f \in i$. The bundle of inputs is assembled into final goods using the Kimball aggregator:

$$\int_0^1 \Upsilon \left(\frac{Y_{ft}}{Y_{it}} \right) df = 1,$$

where $\Upsilon(\cdot)$ is strictly increasing, strictly concave, and satisfies $\Upsilon(1) = 1$. Taking as given the industry demand Y_{it} , each firm minimizes costs subject to the aggregate constraint:

$$\min_{Y_{ft}} \int_0^1 \varphi_{ft} P_{ft} Y_{ft} df \quad \text{s.t.} \quad \int_0^1 \Upsilon \left(\frac{Y_{ft}}{Y_{it}} \right) df = 1,$$

where φ_{ft} is a firm-specific relative demand shifter (firm appeal). Denoting by ψ the Lagrange multiplier of the constraint, the associated first-order condition is:

$$\varphi_{ft} P_{ft} = \psi \Upsilon' \left(\frac{Y_{ft}}{Y_{it}} \right) \frac{1}{Y_{it}}. \quad (\text{OA.7})$$

Define implicitly the industry price index P_{it} as:

$$\int_0^1 \phi \left(\Upsilon'(1) \frac{\varphi_{ft} P_{ft}}{P_{it}} \right) df = 1,$$

where $\phi := \Upsilon \circ (\Upsilon')^{-1}$. Evaluating the first-order condition (OA.7) at symmetric prices, $\varphi_{ft} P_{ft} = P_{it}$, we get $\psi = \frac{P_{it} Y_{it}}{\Upsilon'(1)}$. Replacing for ψ , we obtain:

$$\frac{\varphi_{ft} P_{ft}}{P_{it}} = \frac{1}{\Upsilon'(1)} \Upsilon' \left(\frac{Y_{ft}}{Y_{it}} \right). \quad (\text{OA.8})$$

Therefore, the demand function faced by firms when resetting prices is:

$$\mathcal{D}_{ft+\tau} = \left[(\Upsilon')^{-1} \left(\Upsilon'(1) \frac{\varphi_{ft} P_{ft}^o}{P_{it+\tau}} \right) \right] \left(\frac{P_{it+\tau}}{P_{t+\tau}} \right)^{-\sigma} Y_{t+\tau}.$$

Taking logs of Equation (OA.1.1) and differentiating, we obtain the following expression for the residual elasticity of demand:

$$\epsilon_{ft+\tau} := -\frac{\partial \ln \mathcal{D}_{ft+\tau}}{\partial \ln P_{ft}^o} = -\frac{\Upsilon' \left(\frac{Y_{ft+\tau}}{Y_{it+\tau}} \right)}{\Upsilon'' \left(\frac{Y_{ft+\tau}}{Y_{it+\tau}} \right) \cdot \left(\frac{Y_{ft+\tau}}{Y_{it+\tau}} \right)}. \quad (\text{OA.9})$$

We now use this result to derive the expression for the log-linearized desired markup in Equation (7) in the paper, under monopolistic competition with Kimball preferences. As above, for ease of exposition, we focus on the symmetric steady state. Denote the steady-state residual demand elasticity by $\epsilon = -\frac{\Upsilon'(1)}{\Upsilon''(1)}$. Then the derivative of the residual demand elasticity $\epsilon_{ft+\tau}$ in (OA.9) with respect to $\frac{Y_{ft+\tau}}{Y_{it+\tau}}$, evaluated at the steady state, is given by:

$$\epsilon' = \frac{\Upsilon'(1) (\Upsilon'''(1) + \Upsilon''(1)) - (\Upsilon''(1))^2}{(\Upsilon''(1))^2} \leq 0. \quad (\text{OA.10})$$

The equation above holds with equality if the elasticity is constant (e.g., under CES preferences). Also in this model, the desired markup is given by the Lerner index. Log-linearizing the Lerner index around the steady state and using Equation (OA.10), we have that, up to a first-order approximation, the log-markup (in deviation from the steady state) is equal to:

$$\mu_{ft+\tau} - \mu_f = \frac{\epsilon'}{\epsilon(\epsilon - 1)} (y_{ft+\tau} - y_{it+\tau})$$

Finally, log-linearizing the demand function (OA.1.1) and using it to replace the log difference in output, we obtain:

$$\mu_{ft+\tau} - \mu_f = -\Gamma \left(p_{ft}^o - p_{it+\tau} \right) + u_{ft}^\mu,$$

where, in the case of Kimball preferences, the sensitivity of the markup to the relative price is given by $\Gamma := \frac{\epsilon'}{\epsilon(\epsilon-1)} \frac{1}{\Gamma''(1)}$ and $u_{ft}^\mu := -\Gamma \varphi_{ft}$ captures residual variation in the markup that depends on the demand shifters. Note that, because there is a continuum of firms within an industry, we have that $p_{it+\tau} = p_{it+\tau}^{-f}$ without loss of generality. Substituting into the pricing equation in (6) and rearranging leads to the expression in Equation (7) in the paper. Finally, following the same steps as the previous section, we obtain $\Omega := \frac{\Gamma}{1+\Gamma}$ and the corresponding mapping to the dynamic pricing equation in Equation (8).

OA.1.2 Aggregation and the cost-based New Keynesian Phillips curve

Suppose $N < \infty$ and order firms in each industry from 1 to N .² The aggregate price index (in log-linear terms) is:

$$p_t = \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^N p_{fit} \right) di,$$

(In the paper, we dropped the industry subscript for ease of notation.) Denote by A_{ft}^\star for $f \in \{1, \dots, N\}$ the set of industries in which the f -th firm can adjust. The price index can then be rewritten as:

$$p_t = \frac{1}{N} \sum_{f=1}^N \left(\int_{i \in I/A_{ft}^\star} p_{fit-1} di + \int_{i \in A_{ft}^\star} p_{fit}^o di \right),$$

where we are using the fact that firms that cannot adjust set their price to their $t-1$ level, whereas firms that can adjust set it to the optimal reset price.

Since A_{ft}^\star has measure $1 - \theta$, and the identity of firms that adjust is an i.i.d. draw from the total population of firms, using the law of large numbers for each

²Letting $N \rightarrow \infty$, all results hold under Kimball preferences. Note also that the same argument goes through with minor modifications, but heavier notation, for $N_i \neq N$ for a non-zero measure of industries. In general, heterogeneity of the parameters can be accommodated by repeating the same argument for each group of homogeneous industries with non-zero measure and then taking weighted averages of different industries. See, e.g., Appendix C2 in Wang and Werning (2022).

$f = \{1, \dots, N\}$ across industries we have that:³

$$\frac{1}{N} \sum_{f=1}^N \int_{i \in I/A_{f,t}^*} p_{fit-1} di = \theta \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^N p_{fit-1} \right) di = \theta p_{t-1}$$

and

$$\frac{1}{N} \sum_{f=1}^N \int_{i \in A_{f,t}^*} p_{fit}^o di = (1 - \theta) \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^N p_{fit}^o \right) di.$$

Defining the average reset price in the economy:

$$p_t^o := \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^N p_{fit}^o \right) di,$$

we obtain equation Equation (9) in the paper:⁴

$$p_t = \theta p_{t-1} + (1 - \theta) p_t^o.$$

Next, we replace the aggregate reset price, p_t^o , with an expression that depends on aggregate marginal costs and prices. Using the definition of firm-level marginal cost in Equation (5), and allowing for arbitrary returns to scale ($v_{ft} \leq 0$), we obtain the following expression for the logarithm of firm-level nominal marginal cost:

$$mc_{fit}^n = c_{it} + a_{fit} + v_{ft} y_{fit}.$$

The average marginal cost in the industry is $mc_{it}^n := \frac{1}{N} \sum_{f=1}^N mc_{fit}^n$, implying:

$$mc_{it}^n = c_{it} + a_{it} + v y_{it}.$$

where we defined aggregate returns to scale $v := \frac{1}{N} \sum_{f=1}^N v_{ft}$ (same for all t and constant across industries, for simplicity). Combining the two equations above and subtracting the (log) industry price index from both sides, we obtain an expression that relates real marginal costs to cost shifters and output:

$$mc_{fit} = mc_{it} + (a_{fit} - a_{it}) + v(y_{fit} - y_{it}).$$

We use the demand function to express the log output deviation, $y_{fit} - y_{it}$, in terms

³The i.i.d. assumption implies that: $\int_{i \in B \subseteq [0,1]} p_{fit} di = Pr(B) \int_{i \in I} p_{fit} di$. Notice also that $\int_{i \in [0,1]} \left(\frac{1}{N} \sum_{f=1}^N p_{it}^{-f} \right) di = \int_{i \in [0,1]} \left(\frac{1}{N} \sum_{f=1}^N \left[\frac{N}{N-1} p_{it} - \frac{1}{N-1} p_{fit} \right] \right) di = p_t$.

⁴Notice that, up to a first-order approximation around the symmetric steady state, $p_t = \theta p_{t-1} + (1 - \theta) p_t^o$ also holds with Kimball preferences.

of log prices. In the case of CES preferences (see Equation (OA.1)), we obtain:

$$mc_{fit} = mc_{it} + (a_{fit} - a_{it}) - \gamma v (p_{fit}^o - p_{it}) - \gamma v \ln \varphi_{fit},$$

where γ denotes the within-industry elasticity of substitution.⁵

We then proceed with the following steps: we first manipulate Equation (OA.5) to express the reset price in recursive form, then decompose firm-level nominal marginal cost into firm-level real marginal cost and the industry price index prices, and finally use Equation (OA.1.2) to replace for firm-level real marginal cost:

$$\begin{aligned} p_{fit}^o &= (1 - \beta\theta) \left((1 - \Omega)(mc_{fit}^n + \mu_f) + \Omega p_{it}^{-f} + (1 - \Omega)u_{fit}^\mu \right) + \beta\theta \mathbb{E}_t p_{fit+1}^o \\ &= (1 - \beta\theta)\Theta \left((1 - \Omega)\widehat{mc}_{it} + \Omega p_{it}^{-f} + (1 - \Omega)(1 + \gamma v)p_{it} + (1 - \Omega)u_{fit}^\mu \right) \\ &\quad + \beta\theta \mathbb{E}_t p_{fit+1}^o + (1 - \beta\theta)\Theta(1 - \Omega) \left(a_{fit} - a_{it} - \gamma v \ln \varphi_{fit} \right), \end{aligned}$$

where $\Theta := \frac{1}{1 + \gamma v(1 - \Omega)}$ captures macroeconomic complementarities due to aggregate returns to scale in production. By averaging across firms and industries, we have that the aggregate reset price is given by:

$$p_t^o = (1 - \beta\theta) \left((1 - \Omega)\Theta \widehat{mc}_t + p_t \right) + \beta\theta \mathbb{E}_t p_{t+1}^o + \frac{\theta}{1 - \theta} u_t,$$

where $u_t := \frac{(1 - \theta)(1 - \beta\theta)}{\theta} (1 - \Omega)\Theta \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^N u_{fit}^\mu \right) di$ is an aggregate cost-push shock and $\left(a_{fit} - a_{it} + \gamma v \ln \varphi_{fit} \right)$ is such that $\int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^N (a_{fit} - a_{it} + \gamma v \ln \varphi_{fit}) \right) di = 0$. This follows from the i.i.d. assumption on price adjustments, which implies that the average productivity of resetting firms coincides with the unconditional average. Subtracting p_t from both sides and using the log-linearized price index:

$$\begin{aligned} p_t^o - p_t &= (1 - \beta\theta)(1 - \Omega)\Theta \widehat{mc}_t + \beta\theta (\mathbb{E}_t p_{t+1}^o - p_t) + \frac{\theta}{1 - \theta} u_t \\ \Rightarrow \frac{\theta}{1 - \theta} \pi_t &= (1 - \beta\theta)(1 - \Omega)\Theta \widehat{mc}_t + \beta\theta \mathbb{E}_t \left(\frac{\theta}{1 - \theta} \pi_{t+1} + \pi_{t+1} \right) + \frac{\theta}{1 - \theta} u_t \end{aligned}$$

Rearranging one obtains the marginal cost-based Phillips curve:

$$\pi_t = \lambda \Theta \widehat{mc}_t + \beta \mathbb{E}_t \pi_{t+1} + u_t$$

where $\lambda := \frac{(1 - \theta)(1 - \beta\theta)}{\theta} (1 - \Omega)$ is the slope. The equation above highlights that macroeconomic complementarities also mediate the pass-through of marginal cost to prices via Θ . Under the assumption of constant aggregate returns to scale, we have

⁵A similar expression holds under monopolistic competition with Kimball preferences. In this case, γ is replaced with the corresponding elasticity of relative output to relative prices, $1/\gamma''(1)$.

that $\Theta = 1$, and the Phillips curve simplifies to Equation (11). This condition is exactly met when $\nu = 0$, but also when $\Omega = 1$.

OA.2 Data and measurement

OA.2.1 Data sources and data cleaning

We use information from PRODCOM to compute the quarterly change in product- and firm-level prices and to define the boundaries of markets (industries) in which firms compete. PRODCOM is a large-scale survey commissioned by Eurostat and administered in Belgium by the national statistical office. The survey is designed to cover at least 90% of domestic production value within each manufacturing industry (4-digit NACE codes) by surveying all firms operating in the country with (a) a minimum of 20 employees or (b) total revenue above 4.5 million Euros (European Commission 2014). Firms are required to disclose, on a monthly basis, product-specific physical quantities (e.g., volume, kg., m^2 , etc.) of production sold and the value of production sold (in Euros) for all their manufacturing products. Products are defined in PRODCOM by an 8-digit PC code (e.g., 10.83.11.30 is “Decaffeinated coffee, not roasted”, 10.83.11.50 is “Roasted coffee, not decaffeinated”, and 10.83.11.70 is “Roasted decaffeinated coffee”). Industries are defined by the first four digits of the product codes (e.g., 10.83 is “Processing of tea and coffee”). Sectors are defined by the first two digits of the product codes (e.g., AC is “Manufacture of food products, beverages, and tobacco products”). The sector definitions follow the NACE Rev.2 classification. For the most part, the industry definition also follows Rev.2. The only exceptions are industries under NACE Rev.1 (i.e., before 2008) that do not directly map into a NACE Rev.2 industry; these are assigned a fictitious, unique 4-digit industry code.

In the raw data, there are approximately 4,000 product headings distributed across 13 manufacturing sectors. The PC product codes have been revised several times between 1999 and 2019, with a substantial overhaul in 2008. We use the conversion tables provided by Eurostat and firm-specific information on firms’ product baskets to harmonize the 8-digit product codes across consecutive

quarters and harmonize 4-digit industry codes over time.⁶ In most cases, the conversion tables provide a unique mapping of the 8-digit product codes across consecutive years. In a limited number of cases, the mapping is many-to-one, one-to-many, or many-to-many. The many-to-one mapping is straightforward, while the one-to-many and many-to-many mappings could be problematic. We are able to handle most of these cases using information on the basket of products produced by each firm.⁷ In a limited number of cases (less than 0.1% of the sample), we do not have sufficient information to resolve the uncertainty regarding the mapping. We drop these observations from the sample. Table OA.1 reports the list of manufacturing sectors and their 2-digit PC codes.

We aggregate monthly information at the quarterly level and construct product-level prices (unit values) by dividing product-level sales by product-level quantities sold. As explained in the paper, we are interested in domestic prices, i.e., prices charged by producers in Belgium. PRODCOM does not require firms to separately report production and sales to domestic and international customers. Therefore, we recover domestic values and quantities sold by combining information from PRODCOM with data on firms' product-level exports (quantities and sales) available through Belgian Customs (for extra-EU trade) and the Intrastat Inquiry (for intra-EU trade).⁸ We use the official conversion tables provided by Eurostat to map the CN product code classification used in the international trade data to the PRODCOM product code classification.⁹ In most cases, the CN-to-PC conversion involves either a one-to-one or many-to-one mapping, which poses no issues. We drop observations that involve one-to-many and many-to-many mappings. These account for less than 5% of the observations and production value.

⁶The official conversion tables are available at <https://ec.europa.eu/eurostat/ramon>. The harmonization of the industry code essentially consists of harmonizing the NACE Rev.1 industry, used before 2008, to the NACE Rev.2 industry codes, used from 2008.

⁷For example, consider a case where the official mapping indicates that product 11.11.11.11 in year t could map to either 22.22.22.21 or 22.22.22.22 in year $t + 1$. Suppose two firms, f_1 and f_2 , report in period t sales of product 11.11.11.11 in year t . If f_1 reports *only* sales of 22.22.22.21 and f_2 *only* reports sales of 22.22.22.22 in year $t + 1$ we infer that we should map 11.11.11.11 to 22.22.22.21 for the former and 11.11.11.11 to 22.22.22.22 for the latter.

⁸In constructing our measure of domestic sales, we address issues related to carry-along trade, which might overstate the amount of production by firms that import products destined for immediate sales.

⁹The first six digits of the CN product classification codes correspond to the World HS classification system.

We apply the following filters and data manipulations to the PRODCOM data set. First, we retain firms' observations in a given quarter only if there was positive production reported for at least one product in that quarter. This avoids large jumps in quarterly values due to non-reporting for some months by certain firms. In the rare cases when a firm reports positive values but quantities are missing, we impute the quantity sold from the average value-to-quantity ratio in the months where both values and quantities are reported. Second, we require firms to file VAT declarations and Social Security declarations (as explained below); these two data sources are needed to measure firms' marginal costs.

We use the customs microdata to obtain information on international competitors selling manufacturing products in Belgium. For each domestic firm, the merged Customs-Intrastat data reports the quantity purchased (in kg) and sales (converted to Euros) of different manufacturing products (about 10,000 distinct CN product headings) purchased by Belgian firms from each foreign country. As is standard when dealing with customs data, we define a foreign competitor as a foreign country-domestic buyer pair. For each foreign competitor, we aggregate the product-level sales and quantity sold at the quarterly level (the reporting is monthly in the raw data) and compute quarterly prices (unit values) by taking the ratio of the two.¹⁰

We leverage data from two administrative sources to measure firms' total production (turnover) and variable production costs on a quarterly basis. Belgian firms file VAT declarations to the tax authority that contain information on the total sales of the enterprise as well as information on purchases of raw materials and other goods and services that entail VAT-liable transactions, including domestic and international transactions. The coverage of the VAT declarations is almost universal, with a limited number of exceptions that affect the reporting of sole proprietorships and self-employed individuals, and therefore mostly do not apply to the firms surveyed by PRODCOM.¹¹ We obtain information on employment and

¹⁰Some CN codes change over time (although to a lesser extent than PC codes). We use the official conversion tables, available on the Eurostat website, to map CN product codes across consecutive years. We make adjustments only if the code change is one-to-one between two years. We do not account for changes in PC codes that involve splitting into multiple codes or multiple PC codes combining into one code. Effectively, these changes in the PC codes are treated as if new products are generated.

¹¹Enterprises file their VAT declarations online, either monthly or quarterly, depending on certain

Table OA.1: List of manufacturing sectors

Sector	Sector definition	NACE Rev.2 2-digits codes
CA	Food products, beverages and tobacco products	10–12
CB	Textiles, apparel, leather and related products	13–15
CC	Wood and paper products, and printing	16–18
CE	Chemicals and chemical products	20
CG	Rubber and plastics products, and other non-metallic mineral products	22–23
CH	Basic metals and fabricated metal products, except machinery and equipment	24–25
CJ	Electrical equipment	27
CK	Machinery and equipment n.e.c.	28
CL	Transport equipment	29–30

Notes. This table reports the list of manufacturing sectors in our sample and the corresponding 2-digit NACE Rev.2 codes.

labor costs (wage bill) from the Social Security declarations filed quarterly by each Belgian firm with the Department of Social Security of Belgium.

We sum firms' quarterly expenditures on intermediates and labor to obtain a measure of total variable costs, which we use to construct firms' marginal costs. We multiply these costs by the ratio of total manufacturing sales (from PRODCOM) to total sales (from the VAT declarations) to adjust for the fact that some firms also have production outside manufacturing.¹²

Sample selection—We apply the following data-cleaning steps to address missing values and outliers. (i) We focus on manufacturing industries defined by the NACE Rev.2 industry codes from 10 to 30, excluding from our sample all PRODCOM product headings that correspond to mining, quarrying, and industrial services

size-based thresholds. Small enterprises (turnover \leq 2.5 million Euros excl. VAT) can choose to file either monthly or quarterly. Large enterprises file monthly. In the case of multiple plants or establishments under one VAT identifier, the declaration is filed as a single document for that VAT identifier. We aggregate all monthly declarations to the quarterly level. At this reporting frequency, VAT declarations tend to reflect the sales of output produced in the previous quarter. For this reason, we use one-quarter leads in VAT declarations to construct the measure of firm-level value added used in the regressions discussed in Section 6.

¹²As mentioned below, we conservatively drop observations referring to firms whose manufacturing sales are less than seventy percent of total sales. In the remaining sample, the ratio has a mean of 0.94 and a median of 0.97, confirming PRODCOM's extensive coverage.

related to other manufacturing activities. (ii) We drop observations referring to firms whose sales from manufacturing products (as measured in PRODCOM) are lower than seventy percent of total firm-level sales (as reported in the VAT declarations). This ensures that our sample includes firms whose real activity is primarily, if not entirely, in manufacturing. (iii) As is standard, we exclude firms that operate in the “Coke and refined petroleum products” sector and the “Pharmaceuticals, medicinal chemical, and botanical products” sector, whose output prices are frequently privately bargained or determined in international markets. We drop firms operating in the “Computer, electronic and optical products” and “Electrical equipment” sectors due to the small number of domestic producers operating in these sectors. Finally, we exclude firms operating in the “Other manufacturing and repair and installation of machinery and equipment” sector, a residual grouping consisting of firms producing diverse and varied products for which it is difficult to define an appropriate set of competitors. (iv) We keep only observations for which we are able to compute product-level price index, the corresponding quantity index, competitors’ price index, and marginal costs. (v) We drop observations for which the quarter-to-quarter change of either the firm-level price index or marginal costs is greater than 100% in absolute value. (vi) For each firm-industry pair that enters our dataset intermittently, we only retain the longest continuous time period. This approach ensures that each time series used in the estimation is gapless, avoiding the need to interpolate data and make assumptions about prices and marginal costs when information is missing. (vii) Finally, we exclude firm-industry pairs with a continuous time series shorter than 8 quarters from our dataset. This minimum duration requirement helps mitigate potential Nickell bias in our model estimation.

OA.2.2 Construction of price indices

We construct a set of indices that capture price changes in manufacturing goods at various levels of aggregation (firm-industry, firm, industry, individual manufacturing sector, and whole manufacturing sector).

Firm-industry price index—The main variable of interest is the price of domestically sold manufacturing products at the firm-industry level, P_{ft} , for both domestic and foreign producers. We construct this variable using information on

price changes at the most disaggregated level allowed by the data.

Due to repeated revisions of PRODCOM product codes, a consistent 8-digit product code taxonomy does not exist across the entire sample period. Therefore, we compute the sequence of price changes across consecutive time periods (t and $t+1$) by mapping the product codes at $t+1$ to their corresponding codes at t , aggregating them at the firm-industry level, and recovering the time series of the firm-industry price index (in levels) by concatenating quarterly price changes.

Specifically, denote by \mathcal{P}_{ft} the set of products manufactured by firm f and by P_{pt} the price (unit value) of a given product $p \in \mathcal{P}_{ft}$. We first compute the gross price change for each product, $P_{pt}/P_{p,t-1}$. In doing so, we utilize the official PRODCOM harmonization tables to account for changes in product codes between consecutive quarters and eliminate product-level observations with abnormally large price fluctuations within a given quarter ($P_{pt}/P_{p,t-1} > 3$ or $P_{pt}/P_{p,t-1} < 1/3$) to mitigate the impact of outliers. We then construct a Törnqvist index measuring the quarterly firm-industry price changes:

$$P_{ft}/P_{f,t-1} = \prod_{p \in \mathcal{P}_{ft}} (P_{pt}/P_{p,t-1})^{\bar{s}_{pt}}, \quad (\text{OA.11})$$

where \bar{s}_{pt} is a Törnqvist weight computed as the average of the sale shares between t and $t-1$: $\bar{s}_{pt} := \frac{s_{pt} + s_{p,t-1}}{2}$.¹³ Finally, the time series of firm-industry price levels, P_{ft} , is constructed by concatenating the Törnqvist index:

$$P_{ft} = P_{f0} \prod_{\tau=t_f^0+1}^t (P_{f\tau}/P_{f,\tau-1}). \quad (\text{OA.12})$$

Here t_f^0 denotes the firm-industry's base year. That is the first quarter when f appears in our data. P_{f0} is the level of the price index in the base year. We construct P_{f0} as follows. First, using product-level prices, we compute the average price of the products produced by f in period t_f^0 relative to the average price of products in the industry in the same period, and multiply this relative price by the aggregate price

¹³This index accounts for the presence of multi-product firms by averaging across products produced by the same firm in a given industry. The Törnqvist weights give larger weights to those products that account for a larger share of the firm's turnover.

index $P_{t_f^0}$:

$$P_{f0} = P_{t_f^0} \cdot \frac{\prod_{p \in \mathcal{P}_{f t_f^0}} (P_{p t_f^0})^{s_{p t_f^0}}}{\prod_{p \in \mathcal{F}_{i t_f^0}} (P_{p t_f^0})^{s_{p t_f^0}}}. \quad (\text{OA.13})$$

The aggregate price index is initialized to one in the first quarter of our sample (1999:Q1) and is constructed recursively by concatenating a Törnqvist index of industry-level prices. As discussed in the paper, the choice of the firm-industry price index is immaterial for our empirical analysis, as level effects are absorbed by the firm-industry fixed effects included in our empirical specifications.

Firm price index—As discussed in the paper, the vast majority of firms in our data operate in only one (4-digit) industry, implying that the firm-industry price index, P_{ft} , and the firm price index, \bar{P}_{ft} , coincide. However, in a limited number of cases, it becomes necessary to construct a firm's price index that aggregates across different firm-industry price indices. In doing this, we construct the firm-level price index \bar{P}_{ft} following a method similar to the one described above. Specifically, we construct a Törnqvist index that aggregates across price changes of the individual (4-digit) industry bundles $i \in I_f$ produced by firm f in quarter t : $\bar{P}_{ft}/\bar{P}_{ft-1} = \prod_{i \in I_f} (P_{fit}/P_{fit-1})^{\bar{s}_{fit}}$, with Törnqvist weights defined as $\bar{s}_{fit} := (s_{fit} + s_{fit-1})/2$, where s_{fit} is the share of sales of industry i in the firms' total sales (across manufacturing industries). To obtain the price index in levels, \bar{P}_{ft} , we concatenate the quarterly price changes starting from a base year. The procedure to construct the base years is the same as the one discussed above (Equation OA.13). Note that for single-industry firms the price index \bar{P}_{ft} coincides with the firm-industry price index P_{ft} in Equation (OA.12).

Competitors price index—Using a similar approach, we construct the competitors' price index for each domestic firm. We start by computing quarterly price changes: $P_{it}^{-f}/P_{it-1}^{-f} = \prod_{k \in \mathcal{F}_i/f} (P_{kt}/P_{kt-1})^{\bar{s}_{kt}^{-f}}$, with $\bar{s}_{kt}^{-f} := \frac{1}{2} \left(\frac{s_{kt}}{1-s_{ft}} + \frac{s_{kt-1}}{1-s_{ft-1}} \right)$. Here s_{kt-1} is a Törnqvist weight, averaging the residual revenue share of industry competitors (excluding firm f) at times t and $t-1$. We concatenate these changes starting from a base year, which is the geometric average of firm sales in the industry. Note that the set of domestic competitors for each Belgian producer, denoted in the paper by \mathcal{F}_i , includes not only other Belgian manufacturers operating in the same industry,

but also foreign manufacturers that belong to the same industry and sell to Belgian customers.

Industry, sector, and aggregate price index—We construct the industry-level, sector-level, and aggregate (manufacturing) price indices by aggregating quarterly firm-level price changes. The formula to construct the percentage change in these price indices is analogous to the one in Equation (OA.11), where now the Törnqvist weights assigned to each firm-industry price change, P_{ft}/P_{ft-1} , capture the (weighted) average market shares of the firm in its own industry, sector, or manufacturing, respectively. Once again, the levels of the indices are constructed by concatenating quarterly changes in their respective Törnqvist indices.

OA.2.3 Estimation of firm-level technical efficiency (TFPQ)

We construct our productivity index, log-technical productivity (TFPQ), as a residual from a gross-output production function. In logarithms:

$$y_{ft} = TFPQ_{ft} + f(l_{ft}, k_{ft}, m_{ft}; \vartheta),$$

where y_{ft} denotes firm-level physical output produced and $f(\cdot)$ the log-gross output production function aggregating labor (l_{ft}), capital (k_{ft}), and intermediate inputs (m_{ft}). ϑ is a vector collecting the parameters that pin down the elasticities of output with respect to different inputs.

We recover the vector ϑ via a production function estimation following the estimation developed by Lenzu, Rivers and Tielens (2024). The approach builds on the two-stage estimation routine in Gandhi, Navarro and Rivers (2020), augmented to control for differences in output quality and market power in the product market. We outline the assumptions and steps behind the estimation procedure.

Setup and estimation routine—We assume that a firm’s technical efficiency is the sum of two components $TFPQ_{ft} := \omega_{ft} + z_{ft}$. The first, ω_{ft} , represents the persistent component of productivity. It is observable by the firm when it makes production decisions at the beginning of each period t and it evolves as a first-order Markov process. The second component, z_{ft} , is a non-persistent shock, realized after period- t input and pricing decisions have been made.

Denote by I_{ft} the firm’s information at the beginning of period t . Exploiting the

Markovian nature of the persistent component of TFPQ, we have that $\omega_{jt} = h(\omega_{jt-1}) + \xi_{jt}$, where $h(\omega_{jt-1}) = \mathbb{E}[\omega_{jt} \mid \omega_{jt-1}]$ and ξ_{jt} is an unanticipated productivity “innovation” such that $\mathbb{E}[\xi_{jt} \mid \mathcal{I}_{jt-1}] = 0$.

The estimation routine for TFPQ consists of two steps.

The first step of the estimation strategy is based on a transformation of the firm’s first-order condition for intermediate inputs, which relates observed input shares for intermediate inputs (s_{ft}^m) to the elasticity of output for intermediate inputs and firm-level markup chosen by the firm:

$$s_{ft}^m = \left(\frac{\partial}{\partial m_{ft}} f(k_{ft}, l_{ft}, m_{ft}; \cdot) \right) - \mu_{ft} - z_{ft}, \quad (\text{OA.14})$$

where, as in our theoretical framework, the desired log-markup μ_{ft} is given by a Lerner index that depends on the elasticity of demand at the optimum, $\epsilon_{ft} = \left(\frac{\partial Y_{ft}}{\partial P_{ft}} \frac{P_{ft}}{Y_{ft}} \right)$.

As in our theoretical model, we assume an invertible demand system that generates the following residual demand function for goods produced by firm f :

$$\mathcal{D}_{ft} := d(P_{ft}, P_{it}, \varphi_{ft}) Y_{it},$$

which implies that quantity demanded depends on the firm’s price P_{ft} , a firm-specific demand shock φ_{ft} (assumed to be multiplicative), and a vector of industry-time specific shifters α_{it} capturing industry expenditures. Using this result, we can approximate ϵ_{ft} as a function of the firm’s price and industry and time shifters (p_{ft}, α_{it}) and re-write Equation (OA.14) as:

$$s_{ft}^m = \ln \left(\frac{\partial}{\partial m_{ft}} f(k_{ft}, l_{ft}, m_{ft}; \cdot) \right) - \mu(p_{ft}, \alpha_{it}) - z_{ft} \quad (\text{OA.15})$$

$$= n(k_{ft}, l_{ft}, m_{ft}, p_{ft}, \alpha_{it}) - z_{ft}, \quad (\text{OA.16})$$

Thus, following Blum et al. (2024), we recover the ex-post productivity shock, \widehat{z}_{ft} , and a term combining the output elasticity of intermediate inputs and the desired markup by estimating a regression of intermediate input shares on a polynomial in $(k_{ft}, l_{ft}, m_{ft}, p_{ft})$ and fixed effects. Note that, as in Akerberg, Caves and Frazer (2015), the output elasticity of intermediate inputs is recovered in the second stage, along with the rest of the production function.

The second step of the estimation procedure recovers the production function and productivity. Define firm output net of the ex-post productivity shock as $\mathcal{Y}_{ft} \equiv$

$y_{ft} - z_{ft} = f(k_{ft}, l_{ft}, m_{ft}; \cdot) + \omega_{ft}$. We recover the empirical counterpart of this as a difference between the firm's output and the \hat{z}_{ft} recovered from the step of the estimation procedure in (OA.16).

Exploiting the Markovian property of ω_{ft} , we can re-write \mathcal{Y}_{ft} as:

$$\mathcal{Y}_{ft} = h(\underbrace{\mathcal{Y}_{ft-1} - f(k_{ft-1}, l_{ft-1}, m_{ft-1}; \cdot)}_{h(\omega_{ft-1})}) + f(k_{ft}, l_{ft}, m_{ft}; \cdot) + \xi_{ft}.$$

where $h(\omega_{ft-1})$ is a control function that controls for the persistency of productivity. As is standard (Akerberg, Caves and Frazer (2015); Gandhi, Navarro and Rivers (2020)), we model the control function as polynomial in inputs, $h(\omega_{ft-1}) = \sum_{0 < a \leq 3} \psi_a \omega_{ft-1}^a$, and construct the following recursive equation:

$$\mathcal{Y}_{ft} = f(k_{ft}, l_{ft}, m_{ft}; \cdot) + \sum_{0 < a \leq 3} \psi_a (\mathcal{Y}_{ft-1}(\psi, \vartheta) - f(k_{ft-1}, l_{ft-1}, m_{ft-1}; \cdot))^a.$$

We need to address possible bias arising from the fact that firm-level output is measured in quantity but, as we explain below, firm-level inputs are measured as expenditures deflated by their respective industry price indices.¹⁴ To do so, we follow the approach in De Loecker et al. (2016) and augment the regression equation with a second control function of (output) prices and market shares to correct for the bias to obtain the following estimating equation:

$$\begin{aligned} \mathcal{Y}_{ft} = & f(k_{ft}, l_{ft}, m_{ft}; \cdot) - cf(p_{ft}, s_{it}; \phi) + \\ & \sum_{0 < a \leq 3} \psi_a (\mathcal{Y}_{ft-1}(\psi, \vartheta) - f(k_{ft-1}, l_{ft-1}, m_{ft-1}; \cdot) - cf(p_{ft-1}, s_{it-1}; \phi))^a, \end{aligned} \quad (\text{OA.17})$$

which depends on the production parameter vector ϑ —the object of interest—and the ancillary parameter vectors ψ and ϕ .

We estimate Equation (OA.17) via GMM. As is standard, we assume that capital is pre-determined at the beginning of each period t , while labor and intermediate inputs are flexibly chosen period-by-period. Under these assumptions $(k_{ft}, l_{jt}, \mathcal{Y}_{ft-1})$ are orthogonal to ξ_{ft} and they can be used as instruments for themselves. However, since l_{ft}, m_{ft}, p_{ft} and s_{ft} are chosen in period t , they are correlated with ξ_{ft} . We therefore use their one-year lag $l_{ft-1}, m_{ft-1}, p_{ft-1}$ and s_{ft-1} as instruments.

The following set of moment conditions identifies parameters (ϑ, ψ, ϕ) :

$$\mathbb{E}[\xi_{ft} \cdot k_{ft}^{\tau_k} l_{ft-1}^{\tau_l} m_{ft-1}^{\tau_m}] = 0 \quad \mathbb{E}[\xi_{ft} \cdot \mathcal{Y}_{ft-1}^a] = 0 \quad \mathbb{E}[\xi_{ft} \cdot p_{ft}^{\tau_p} s_{ft}^{\tau_s}] = 0$$

¹⁴See De Loecker et al. (2016) for a discussion of the sources and implications of this bias.

Table OA.2: Estimates of output elasticities and returns to scale

Sector	Output elasticities			Returns to scale	
	Labor (ϑ^l)	Intermediates (ϑ^m)	Capital (ϑ^k)	Short-run ($\vartheta^l + \vartheta^m$)	Long-run ($\vartheta^l + \vartheta^m + \vartheta^k$)
CA	0.214	0.026	0.752	0.966	0.992
CB	0.200	0.028	0.764	0.964	0.992
CC	0.230	0.033	0.731	0.960	0.994
CE	0.217	0.039	0.720	0.937	0.976
CG	0.219	0.008	0.768	0.987	0.996
CH	0.188	0.028	0.776	0.963	0.991
CJ	0.177	0.036	0.783	0.960	0.996
CK	0.247	0.054	0.684	0.932	0.986
CL	0.243	0.036	0.725	0.968	1.004
Aggregate	0.212	0.026	0.753	0.965	0.992

Notes. This table reports the production function estimates for different sectors. The first three columns report the estimated output elasticities of labor, intermediate inputs, and capital. The last two columns report the short-run and long-run returns to scale. Each row refers to a different manufacturing sector. The last row is the average of the sectoral estimates.

where $0 \leq \tau_k + \tau_l + \tau_m \leq 2$, $0 \leq \tau_p + \tau_s \leq 2$, and $\tau_k, \tau_l, \tau_m, \tau_p, \tau_s \geq 0$.

Measurement and implementation—We perform the production function estimation using yearly data. To construct our baseline TFPQ instrument, we assume firms’ technologies can be approximated by a Cobb-Douglas production function, allowing output elasticities to vary between sectors (2-digit NACE):

$$TFPQ_{ft} = y_{ft} - \widehat{\vartheta}^l \cdot l_{ft} - \widehat{\vartheta}^k \cdot k_{ft} - \widehat{\vartheta}^m \cdot m_{ft}. \quad (\text{OA.18})$$

where $Y_{ft} = \frac{(PY)_{ft}}{P_{ft}}$ is the firm-level quantity index constructed by deflating firm-level sales by the firm-level price index. We measure labor services using the total quarterly wage bill from the Social Security dataset and intermediate costs using the quarterly expenses in materials and services reported in firms’ VAT declarations. We construct a quarterly measure of capital services following the perpetual inventory method, using data on quarterly investments in fixed tangible assets from firms’ VAT declarations. We deflate labor, capital, and intermediate inputs using the corresponding industry-level deflators.

Table OA.2 presents the estimates of the output elasticities and returns to scale

for individual manufacturing sectors and for the aggregate economy. The latter is the average of the sectoral estimates. SR-RTS are the sum of the elasticities of variable inputs (labor and intermediates). Consistent with the findings in previous studies, our estimates indicate returns to scale in the ballpark of unity for most sectors and, therefore, in the aggregate.¹⁵

Alternative production functions and variable SR-RTS—We construct an alternative productivity index by modeling firms’ technologies with a Translog production function:

$$TFPQ_{ft}^{Translog} = y_{ft} - \sum_{x=l,k,m} \widehat{a}_x \cdot x_{ft} - \sum_{x=l,k,m} \sum_{x'=l,k,m} \widehat{a}_{xx'}/2 \cdot x_{ft} x'_{ft}.$$

As in our baseline estimation, we estimate the production function parameters \widehat{a}_s separately for each sector and recover firm-time specific input elasticities, $\widehat{\vartheta}_{ft}^x = \widehat{a}_x + \widehat{a}_{xx'}/2 \cdot x'_{ft} + \widehat{a}_{xx'} \cdot x_{ft}$ of capital, labor, and intermediate inputs.

Using these, we obtain an estimate of firm-time specific short-run returns to scale, $\widehat{\vartheta}_{ft}^l + \widehat{\vartheta}_{ft}^m$, and construct a measure of marginal cost that directly accounts for this variation, $MC_{ft}^n = AVC_{ft}^{(1+\widehat{v}_{ft})}$, where $\widehat{v}_{ft} = (\widehat{\vartheta}_{ft}^l + \widehat{\vartheta}_{ft}^m)^{-1} - 1$. These variables are used to perform the robustness exercises reported in Table 3.

OA.3 Adjustment of TFPQ for variable capacity utilization

As a robustness exercise for our marginal cost instrument, we construct two “purified” TFPQ indices that account for variability in capacity utilization. Toward this purpose, we gather supplementary data on capacity utilization from the Business Survey administered by the National Bank of Belgium. The survey covers a subset of our data, asking firms to report the percentage of their production capacity utilized in any given quarter. After matching the Business Survey to our final regression sample, we are able to gather information on capacity utilization for 485 firms and 18,422 firm-industry-quarter observations.

Construction of TFPQ adjusted for capacity utilization—Because information

¹⁵See Gandhi, Navarro and Rivers (2020) and Lenzu, Rivers and Tielens (2024) and references therein.

on capacity utilization is available for only a small subset of our dataset, we rely on predictive regressions to estimate firm-level capacity when it is not directly observable. First, for firms where capacity utilization data is available for some periods, missing values are imputed using the capital-to-labor ratio, intermediates-to-labor ratio, intermediates-to-sales ratio, current and past investments in physical assets, percentage change in the firm’s sales, percentage change in the firm’s wage bill, firm fixed effects, and industry-by-time fixed effects:

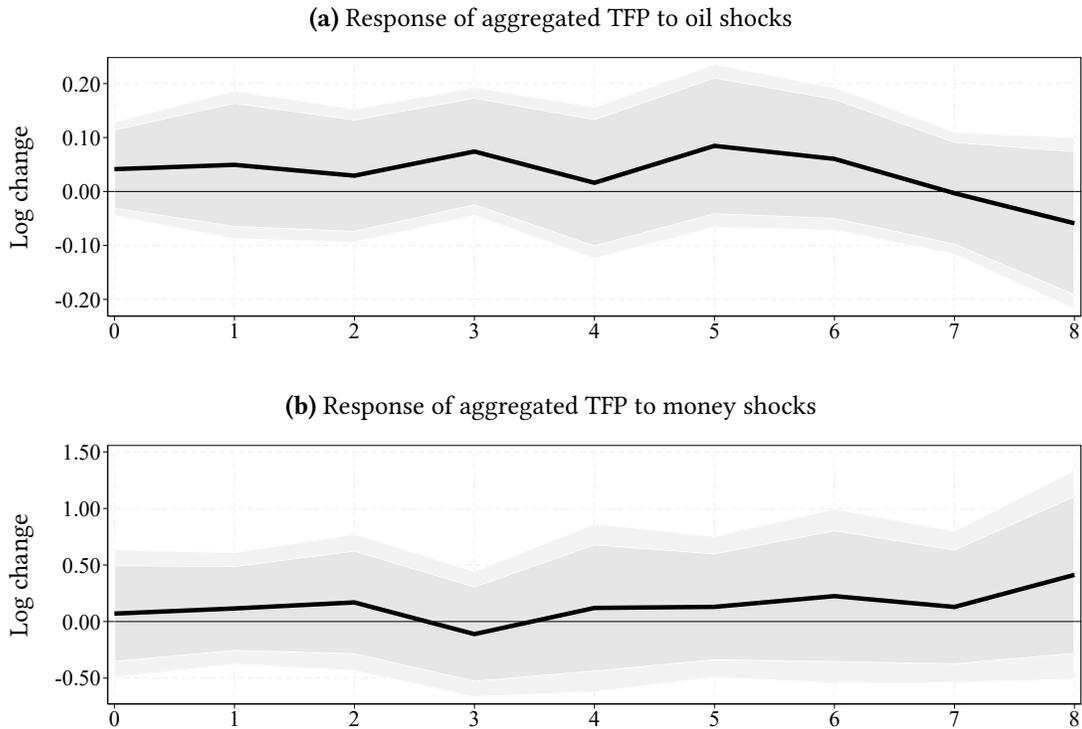
$$\begin{aligned} \widehat{CU}_{ft} = & \underbrace{\beta_1}_{-0.042} (K_{ft}/WL_{ft}) + \underbrace{\beta_2}_{0.091} (PM_{ft}/WL_{ft}) + \underbrace{\beta_3}_{-0.045} (PM_{ft}/PY_{ft}) + \\ & \underbrace{\beta_4}_{0.004} (\ln Invest_{ft}) + \underbrace{\beta_5}_{0.006} (\ln Invest_{ft-1}) + \underbrace{\beta_6}_{0.006} (\Delta \ln PY_{ft}) + \\ & \underbrace{\beta_7}_{0.059} (\Delta \ln WL_{it}) + \iota_f + \iota_{it}. \end{aligned}$$

For firms that never appear in the Business Survey sample, we impute capacity utilization using information on the firm’s capital-to-labor ratio, intermediates-to-labor ratio, intermediates-to-sales ratio, current and past investments in physical assets, percentage change in industry sales, percentage change in the firm’s sales, percentage change in the firm’s wage bill, a second-order polynomial in the firm’s age, and industry-by-calendar quarter fixed effects:

$$\begin{aligned} \widehat{CU}_{ft} = & \underbrace{\beta_1}_{-0.018} (K_{ft}/WL_{ft}) + \underbrace{\beta_2}_{0.050} (PM_{ft}/WL_{ft}) + \underbrace{\beta_3}_{-0.043} (PM_{ft}/PY_{ft}) + \\ & \underbrace{\beta_4}_{0.007} (\ln Invest_{ft}) + \underbrace{\beta_5}_{0.006} (\ln Invest_{ft-1}) + \underbrace{\beta_6}_{0.010} (\Delta \ln PY_{it}) + \\ & \underbrace{\beta_7}_{0.018} (\Delta \ln PQ_{ft}) + \underbrace{\beta_8}_{0.065} (\Delta \ln WL_{ft}) + \underbrace{\beta_9}_{0.001} (age_{ft}) + \underbrace{\beta_{10}}_{-0.001} (age_{ft}^2) + \iota_{qi}. \end{aligned}$$

We then construct two “purified” TFPQ instruments by adjusting either capital or both capital and labor for capacity utilization. The capital- and capital-labor adjusted TFPQ measures are constructed as in Equation (OA.18) but scaling the measure of capital and labor in the production function as follows: $k_{ft}^{CU} = \ln(\widehat{CU}_{ft} \cdot K_{ft})$ and $l_{ft}^{CU} = \ln(\widehat{CU}_{ft} \cdot L_{ft})$.

Figure OA.1: Dynamic effects of oil and money shocks



Notes. The dark (light) gray shaded areas are 90 (95) percent confidence bands obtained from Newey-West standard errors with four quarters of correlation.

Falsification tests using money and oil shocks—We perform the following falsification test to assess the sensitivity of our purified technical productivity measure (adjusting both capital and labor for capacity utilization).

We aggregate our firm-level measure and check whether this variable responds to high-frequency monetary shocks and oil shocks. The idea is that these high-frequency demand and supply shocks may trigger adjustments in capacity utilization, but—to the extent that our firm-level productivity index has been adjusted for this variation—the aggregate “purified” TFPQ index should reflect technical productivity, which is a slow-moving variable. To test this, we estimate the following local linear projection models:

$$TFPQ_{t+h} - TFPQ_{t-1} = b_0^{OS} + b_h^{OS} OS_{t-1} + \epsilon_{ft+h}$$

$$TFPQ_{t+h} - TFPQ_{t-1} = b_0^{MS} + b_h^{MS} MS_{t-1} + \epsilon_{ft+h}$$

for $h = 0, \dots, 8$ quarters. As in the paper, oil shocks (OS) are measured as in Känzig (2021) (unexpected movements in oil price futures the day after OPEC meetings) and money shocks (MS) are constructed as in Gürkaynak, Sack and Swanson (2005) (log change in the price of overnight index swaps within a narrow window around ECB monetary policy announcements).

Figure OA.1 plots the empirical IRF coefficients over different horizons. In Panel A, we plot the response to oil shocks. We normalize the oil shock so that b_1^{OS} represents the effect on impact of a one-standard deviation shock to oil prices (a 15.7 percent increase in Brent crude oil price). In Panel B, we plot the response to monetary shocks. The coefficient represents the effect of an annualized shock of 100 basis points. Reassuringly, both shocks have no explanatory power on our aggregate “purified” TFPQ measure at any horizon.

References

- Akerberg, Daniel A, Kevin Caves, and Garth Frazer.** 2015. “Identification properties of recent production function estimators.” *Econometrica*, 83(6): 2411–2451.
- Alvarez, Fernando, and Francesco Lippi.** 2014. “Price setting with menu cost for multiproduct firms.” *Econometrica*, 82(1): 89–135.
- Atkeson, Andrew, and Ariel Burstein.** 2008. “Pricing-to-market, trade costs, and international relative prices.” *American Economic Review*, 98(5): 1998–2031.
- Blum, Bernardo S, Sebastian Claro, Ignatius Horstmann, and David A Rivers.** 2024. “The ABCs of firm heterogeneity when firms sort into markets: The case of exporters.” *Journal of Political Economy*, 132(4): 1162–1208.
- De Loecker, Jan, Pinelopi K. Goldberg, Amit K Khandelwal, and Nina Pavcnik.** 2016. “Prices, Markups and Trade Reform.” *Econometrica*, 84(2): 445–510.
- European Commission.** 2014. “Quality Report on PRODCOM 2014 Annual Data.” PRODCOM Working Group.
- Gandhi, Amit, Salvador Navarro, and David A. Rivers.** 2020. “On the Identification of Gross Output Production Functions.” *Journal of Political Economy*, 128(8): 2973–3016.

- Gürkaynak, Refet S, Brian P Sack, and Eric T Swanson.** 2005. “Do actions speak louder than words? The response of asset prices to monetary policy actions and statements.” *International Journal of Central Banking*, 1(1): 55–93.
- Känzig, Diego R.** 2021. “The macroeconomic effects of oil supply news: Evidence from OPEC announcements.” *American Economic Review*, 111(4): 1092–1125.
- Lenzu, Simone, David Rivers, and Joris Tielens.** 2024. “Financial Shocks, Productivity, and Prices.” *Working paper*.
- Midrigan, Virgiliu.** 2011. “Menu costs, multiproduct firms, and aggregate fluctuations.” *Econometrica*, 79(4): 1139–1180.
- Wang, Olivier, and Iván Werning.** 2022. “Dynamic oligopoly and price stickiness.” *American Economic Review*, 112(8): 2815–49.