Anatomy of the Phillips Curve: Micro Evidence and Macro Implications

Luca Gagliardone* Mark Gertler† Simone Lenzu‡ Joris Tielens§

November, 2023

Abstract

We develop a bottom-up approach to estimating the slope of the primitive form of the New Keynesian Phillips curve, which features marginal cost as the relevant real activity variable. Using quarterly microdata on prices, costs, and output from the Belgian manufacturing sector, we estimate dynamic pass-through regressions that identify the degrees of nominal and real rigidities in price setting. Our estimates imply a high slope for the marginal cost-based Phillips curve, which contrasts with the low estimates of the conventional unemployment or output gap-based formulations in the literature. We reconcile the difference by demonstrating that, although the pass-through of marginal cost into inflation is substantial, the elasticity of marginal cost with respect to the output gap is low. We also illustrate the advantage of a marginal cost-based Phillips curve for characterizing the transmission of supply shocks to inflation.

JEL: E30.

* New York University. Email: luca.gagliardone@nyu.edu; † New York University and NBER. Email: mark.gertler@nyu.edu; ‡ New York University, Stern School of Business. Email: slenzu@stern.nyu.edu; § National Bank of Belgium and KU Leuven. Email: joris.tielens@nbb.be. We thank Fernando Alvarez, Xavier Gabaix, Jonathon Hazell, Man Chon Iao, Virgiliu Midrigan, Alessio Moro, Carolin Pflueger, Ludwig Straub, Gianluca Violante, Olivier Wang, Ivan Werning, Raf Wouters, and seminar participants at various institutions for helpful comments and conversations. 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, or any other institution with which the authors are affiliated.

1 Introduction

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 $\widetilde{y}_t$ the output gap, the percentage difference between real output and its natural level. Then (what we will refer to as) the *conventional form* of the NKPC is given by:

$$\pi_t = \kappa \, \widetilde{y}_t + \beta \, \mathbb{E}_t \, \{\pi_{t+1}\} + u_t, \tag{1}$$

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 $\widetilde{y}_t$, 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 $\widetilde{y}_t$ 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 $\widetilde{y}_t$ and u_t. Given the absence of good instruments for $\widetilde{y}_t$, the estimation of κ using aggregate time-series data is problematic (Mavroeidis et al. 2014). Another identification issue involves trend inflation. The specification given by equation (1) presumes that trend inflation is constant. However, as emphasized 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 et al. (2020), McLeay and Tenreyro (2020), and Hazell et al. (2022).[1] Importantly, Hooper et al. (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 form* 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.[2] 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.[3] Moreover, even if all conditions that establish a proportional relationship are approximately met, it is crucial to recognize that the output gap-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 identification strategy to estimate the slope of the primitive form of the NKPC using microdata. We leverage a unique high-frequency dataset that provides information on production costs,

[1] Also relevant is Beraja et al. (2019), which uses regional data to identify wage Phillips curves.

[2] Indeed, it is for this reason that 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).

[3] These 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. They also apply to using an aggregate measure of real marginal cost such as the labor share.

prices, and quantities of production at the firm-product level to estimate dynamic pass-through regressions that account for both nominal and real rigidities in price setting within a general class of models of imperfect competition. The identification strategy and the granularity of the data enable us to tackle the issues that typically hinder identification with aggregate data.

Our estimates indicate that the slope of the marginal cost-based Phillips curve is steep, suggesting a substantial pass-through of marginal cost into prices and inflation. At the aggregate level, our estimates imply that, in a parsimonious framework, fluctuations in marginal cost alone can account for almost seventy percent of inflation dynamics without appealing to unobservable cost-push shocks or including lags of inflation as is often done.

These findings stand in stark contrast to the low estimates found in the literature employing the conventional output gap-based formulation of the NKPC (Rotemberg and Woodford 1997, Hazell et al. 2022). Using our firm-level data, we reconcile this difference by showing that the implied elasticity of marginal cost with respect to the output gap is low (at least for pre-pandemic data). In other words, the slope of the conventional NKPC does not stem from a limited transmission of fluctuations in marginal cost to inflation, but rather from the weak connection between movements in the output gap and marginal cost.

The paper proceeds as follows. In Section 2 we develop a theoretical framework that serves as the foundation of our estimation strategy. We start from first principles to derive a Phillips curve featuring both nominal price rigidities and a general form of strategic complementarities in price setting. Notably, our framework encompasses some of the leading models of imperfect competition, such as monopolistic competition with variable elasticity of demand (Kimball 1995), static oligopoly (Atkeson and Burstein 2008), and dynamic oligopoly (Wang and Werning 2022).

Section 3 provides an overview of our data. We collect administrative data on product-level output prices, quantities, and production costs for the Belgian manufacturing sector. Our data is similar to the series used by Amiti et al. (2019). The notable difference is that ours is recorded at a quarterly as opposed to annual frequency, which allows us to study the role of nominal rigidities in price setting. The data records market interactions between both domestic and foreign competitors over a two-decade period (1999—2019). We observe domestic firms' own prices (unit values), their competitors' prices, as well as different components of their variable costs, which directly map to the theoretical objects.

In Section 4, we outline our identification strategy. The conventional approach in the literature involves aggregating individual firm pricing decisions into a NKPC and then estimating its slope with aggregate data. Instead, we use our microdata to estimate dynamic pass-through regressions that identify both the degree of nominal and real rigidities from short-run co-movements in firm-level marginal costs and prices. By doing so, we relate to an earlier literature that estimates micro-level dynamic pass-through regressions of marginal cost into prices (Goldberg and Verboven 2001, Nakamura and Zerom 2010). Notably, in an environment with perfectly flexible prices, our econometric framework nests as a special case the long-run pass-through regressions estimated in Amiti et al. (2019) using annual data.

The use of firm-level data greatly strengthens identification. First, we observe different components of firms' variable costs, which we use to construct powerful instruments to tackle endogeneity and measurement issues. Second, the use of fixed effects allows us to address unobserved heterogeneity as well as potential concerns related to trends in output growth, trend inflation, and shifts in inflation expectations.

In Sections 5 and 6, we present the estimation results and assess their aggregate implications. We obtain sensible and robust estimates of the structural parameters governing firms pricing behavior, indicating a substantial degree of nominal rigidities (three to four quarters of stickiness) and underscoring a central role for strategic complementarities. These estimates imply an economically meaningful slope of the marginal cost-based Phillips curve, tightly estimated in the range of 0.05 to 0.07, even when accounting for empirically plausible degrees of macroeconomic complementarities. These estimates are an order of magnitude larger than the estimates of output gap-based or unemployment gap-based NKPC slopes available in the literature. Finally, we show that a parsimonious marginal cost-based PC tracks aggregate inflation dynamics well.

In Section 7, we reconcile our estimates with those of the conventional NKPC found in the literature. We first formalize the mapping between the marginal cost-based and output-based curves, which is mediated by the elasticity of marginal cost to output. Subsequently, we develop a framework to identify this elasticity and the slope of the output-based NKPC using firm-level data.

In Section 8, we show how to use the marginal cost-based PC to examine the impact of supply shocks on inflation. To illustrate this point, we study the transmission of identified oil shocks. In Section 9, we show that our empirical analysis is robust to extending the baseline Calvo setting to allow for menu costs in pricing. Section 10 concludes.

2 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 subject to nominal rigidity. Firms are granular. They internalize their impact on industry aggregates and are influenced by the pricing decisions of their competitors. This model generates a micro-founded New Keynesian Phillips curve, the slope of which is a function of the structural parameters that govern firms' pricing behavior.

2.1 Preferences and pricing behavior

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

Let P_{ft} be the price charged by each firm for a unit of its output, P_{it} the industry price index, φ_{ft} is 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:

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

We assume firms face nominal rigidities as in Calvo (1983).[4] Each period firms face a probability $(1-\theta)$ of being able to change their price, independent across time and across firms, with $\theta \in [0,1]$. Thus the price P_{ft} paid by consumers in any given period is either the (optimal) reset price set by a firm that is able to adjust, which we denote by P^o_{ft}, or the price charged in the previous period, P_{ft-1}.

The firms adjust their prices during the period in order to maximize expected profits. Their pricing decisions consider both the pricing choices made by competitors and the impact of their own price adjustments on their residual demand and the industry-wide price index. Additionally, 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 set by firms that are able to adjust is a weighted average of current and (expected) future nominal marginal costs and markups. 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^n_{ft} the nominal marginal cost of firm f. Then the optimal reset price P^o_{ft} solves the following profit maximization problem:

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

subject to the sequence of expected demand functions $\{\mathcal{D}_{ft+\tau}\}_{\tau\geq 0}$ in equation (2).

[4]In Section 9 we show that our identification approach and the estimated NPKC remain valid if the data-generating process features Ss-style price adjustments as in conventional menu cost models.

The FOC of the problem is:

$$\mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} \theta^{\tau} \Lambda_{t,\tau} \mathcal{D}_{ft+\tau} \left[\frac{P^o_{ft}}{P_{t+\tau}} - (1+\mu_{ft+\tau}) \frac{MC^n_{ft+\tau}}{P_{t+\tau}} \right] \right\} = 0, \tag{3}$$

where μ_{ft} denotes the desired log markup.

According to equation (3), the optimal reset price depends on the expected path of marginal cost 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. Moreover, in finding the optimal reset price, the firm factors in how its pricing decision today affects the expected path of the desired markups. The desired net markup is then given by the Lerner index:

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

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

2.2 Technology

Firms are heterogeneous in their production technologies. We assume that a unit of output of Y_{ft} is produced at a nominal marginal cost of:[5]

$$MC^n_{ft} = C_{it} \mathcal{A}_{ft} Y^{\nu_f}_{ft}, \tag{5}$$

where C_{it} denotes the nominal unit cost of the composite input factor (e.g., wages and intermediate goods) that is independent of the scale of production; $\mathcal{A}_{ft}$ is a firm-specific cost shifter that captures, among other idiosyncratic factors, heterogeneity in firm's production efficiency; ν_f is a firm-specific parameter that pins down the short-term returns to scale of firms production technology, which are given by $(1/(1+\nu_f))$.

Whereas in the empirical analysis we allow for non-constant returns to

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

scale at the firm level, to derive aggregate implications we focus on the aggregate constant returns to scale case (i.e., on average $\nu_f = 0$). This assumption rules out macroeconomic complementarities due to the feedback of firms' pricing behavior into their respective marginal cost (see e.g. Galí 2015).[6] In Appendix A.2, we present a general framework that allows for arbitrary aggregate returns to scale. In Section 5.1, 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.

2.3 The optimal reset price

We log-linearize the FOC in equation (3) around the symmetric steady state with zero inflation.[7] Denoting with lower-case letters the variables in logs, we obtain that the reset price satisfies:

$$p^o_{ft} = (1-\beta\theta)\mathbb{E}_t \left\{ \sum_{\tau=0}^{\infty} (\beta\theta)^\tau \left(\mu_{ft+\tau} + mc^n_{ft+\tau} \right) \right\}. \tag{6}$$

As we show in Appendix A.1, the log-linearized desired markup is a function that depends inversely on the gap between the firms' own reset price and the price of its competitors, which we denote by p^{-f}_{it}. Formally:

$$\mu_{ft} - \mu = -\Gamma \left(p^o_{ft} - p^{-f}_{it} \right) + u^\mu_{ft}, \tag{7}$$

where $\Gamma > 0$ denotes the markup elasticity with respect to prices and u^μ_{ft} is a shock to the desired markup. u^μ_{ft} is a firm demand shock that generally depends on the demand shifter φ_{ft} (equation 4 of the Appendix). Under weak assumptions, this relationship holds for standard imperfectly competitive frameworks, 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

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

[7] The choice of steady-state inflation is largely immaterial for our purposes but permits a lighter notation. We relax it in the empirical analysis, where we allow for sector/industry-specific trends.

of demand declines as its market share increases. Thus the presence of strategic complementarities in price-setting behavior implies that a relative price increase lowers a firm's desired markup, dampening the response of prices to movements in marginal cost.

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

$$p^o_{ft} = (1-\beta\theta)\mathbb{E}_t\left\{\sum_{\tau=0}^{\infty}(\beta\theta)^\tau\left((1-\Omega)(mc^n_{ft+\tau}+\mu)+\Omega p^{-f}_{it+\tau}\right)\right\} + u_{ft}, \tag{8}$$

where $u_{ft} := (1-\beta\theta)(1-\Omega)\mathbb{E}_t\left\{\sum_{\tau=0}^{\infty}(\beta\theta)^\tau u^\mu_{ft+\tau}\right\}$ is a firm demand shock. The parameter $\Omega := \frac{\Gamma}{1+\Gamma}$ captures the strength of strategic complementarities, which impacts the firm's pricing policy (8) by muting the price response to changes in marginal costs. If the elasticity of demand is constant—as it is in the textbook New Keynesian model with monopolistically competitive firms—so is the desired markup μ_{ft}. 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.

2.4 The New Keynesian Phillips Curve

As we show in Appendix A.2, the log-linear aggregate price index is given by:

$$p_t = (1-\theta)p^o_t + \theta p_{t-1}, \tag{9}$$

with p_t and p^o_t denoting the aggregate price indexes implied by equation (2), which average across firms and industries. Let mc^n_t denote the aggregate log-nominal marginal cost. Define the aggregate real marginal cost and aggregate inflation as $mc_t = mc^n_t - p_t$ 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:

$$p^o_t = (1-\beta\theta)\big((1-\Omega)(mc^n_t+\mu)+\Omega p_t\big) + \beta\theta\mathbb{E}_t p^o_{t+1} + \frac{\theta}{1-\theta}u_t, \tag{10}$$

where u_t is an aggregate cost-push shock, defined in the Appendix A.2. Combining equations (9) and (10) gives the *primitive* formulation of the NKPC curve:

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

which asserts that inflation depends on real marginal cost (in deviation from its steady state) and on expected future inflation. λ is the slope of the NKPC curve, defined by:

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

Two observations are worth noting. First, the primitive form 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. In contrast, the conventional formulation of the Phillips curve, displayed in equation (1), uses the output gap or unemployment to proxy for marginal cost. As we will discuss, the mapping between marginal cost and output gap is theoretically valid only under specific circumstances. Moreover, even when a proportionality between the two variables can be established, the elasticities of marginal cost to output gap and unemployment need not be equal to one. We return to these important points in Section 7.

Secondly, the slope of the NKPC, λ, is a function of the primitives governing firms' pricing behavior. As in standard New Keynesian models (e.g., Galí and Gertler 1999), high nominal rigidities and low discounting flatten the sensitivity of inflation to changes in real economic activity. Additionally, equation (12) shows how strategic complementarities also contribute to reducing the slope. Therefore, given a calibration of the discount factor β, estimates of the structural parameters θ and Ω pin down the slope of the Phillips curve.

Toward this end, we take the structural pricing equation (8) to the data. This exercise requires measures of prices and marginal costs, which we discuss in the next section. Notably, it is the use of firm-level data that permits the identification of the primitive parameters.

3 Data and measurement

We begin by introducing our dataset and highlighting its features that are relevant for measurement purposes. We then illustrate the procedure for constructing price and marginal cost measures using both product-level and firm-level data.

3.1 Data

We assemble a unique micro-level dataset that covers the manufacturing sector in Belgium between 1999 and 2019. A rare feature of our dataset is its ability to track, on a *quarterly basis*, product-level prices and quantities sold in the domestic market by both domestic and foreign producers, as well as information on production costs for domestic producers. Our dataset is compiled from four administrative sources: PRODCOM, international trade data, VAT declarations, and Social Security declarations.

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

We leverage detailed administrative data to measure firms' variable production costs at a quarterly frequency. Specifically, we obtain information on firms' purchases of intermediates (materials and services) from their VAT

[8]PRODCOM surveys all Belgian firms involved in manufacturing production with more than 10 employees, covering over 90% of production in each NACE 4-digit industry. The survey does not require firms to distinguish between production and sales to domestic and international customers. Therefore, we recover domestic values and quantities sold by combining information from PRODCOM with international trade data on firms' product-level exports (quantities and sales).

declarations. Additionally, we draw upon firms' Social Security declarations to obtain a measure of their labor costs (the wage bill).

After applying standard data cleaning filters, our final sample includes 4, 598 firms observed over 84 quarters (1999:Q1–2019:Q4), totaling 132, 915 observations. Appendix B provides detailed information on the data sources and data cleaning procedures. Table 1 presents summary statistics of our dataset. Several features of the data are worth noting.

First, our dataset covers the lion's share of domestic manufacturing production in Belgium, spanning the entire size distribution. The average firm in our dataset employs 74 employees (measured in full-time equivalents) and has a domestic turnover (sales) of €6 million. 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.[9] 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 lion's share of the firms in our sample specializes in only one manufacturing industry. Even for those firms 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, in line with the theoretical framework, we treat each industry as a separate firm.

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%

[9] The 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 Appendix B for sectors' definitions.

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.

3.2 Measurement

We now describe the measures of prices and marginal cost that map to the theoretical counterparts in Section 2. Appendix B 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). Consistent with the notation used in the theoretical framework, we use the subscript i to denote an industry, f to

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	98.22	86.57	100.00	100.00	100.00	100.00
Firm's market share within industry	1.72	0.06	0.22	0.53	1.36	6.57
Firm's market share within sector	0.21	0.01	0.02	0.05	0.13	0.70
Firm's market share within manufacturing	0.03	0.00	0.00	0.01	0.01	0.08
Number of consecutive quarters in sample	42.19	11.00	24.00	38.00	59.00	82.00

Notes. The summary statistics reported in this table refer to the sample of domestic producers in PRODCOM. The sample includes $4,598$ firms observed over 84 quarters (1999:Q1–2019:Q4), totaling $132,915$ observations.

denote a firm-industry pair, and t to denote time (quarters).[10] We denote by s_{ft} the revenue share of the firm in the industry.

We compute the change in firm prices P_{ft}/P_{ft-1}, using the most disaggregated level allowed by the data. For domestic producers, the finest level of aggregation is a firm×PC 8-digit product code level. For foreign competitors, it is the importing-firm×source country×PC 8-digit product code level.[11] 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:[12]

[10]Whenever a firm operates in multiple 4-digit industries, we treat each firm-industry pair as a separate unit in our sample. As we discussed, most firms operate in only one industry, and the main industry accounts for the lion's share of sales of multi-industry firms. Therefore, all our results are essentially unchanged if we restrict the sample to the main industry for each firm.

[11]In the raw customs data, products are measured using the more disaggregated CN 8-digit product classification. We map the CN product codes to PC 8-digit product codes using the official bridge tables available on the Eurostat web page. See Appendix B.1 for additional details.

[12]The Törnqvist index coincides with the Cobb-Douglas index whenever sales shares are

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

In the formula above, $\mathcal{P}_{ft}$ represents the set of 8-digit products manufactured by firm f, P_{pt} is the unit value of product p in $\mathcal{P}_{ft}$, and $\bar{s}_{pt}$ is a Törnqvist weight computed as the average of the sale shares between t and $t-1$: $\bar{s}_{pt} := \frac{s_{pt}+s_{pt-1}}{2}$. We then construct the time series of firms' prices (in levels) by concatenating quarterly changes.[13]

Using a similar approach, we construct the price index of competitors for each domestic firm by concatenating quarterly changes according to the following formula:

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

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, which is constructed by averaging the residual revenue share of competitors in the industry at time t (net of firm f revenues) with that at time $t-1$.[14] It is important to 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 that sell goods to Belgian customers.

Marginal costs. The general cost structure outlined in equation (5) implies that firms' nominal marginal costs are proportional to its average variable costs, as

constant over time. Allowing for time-varying shares is empirically relevant given that market shares of individual firms vary over time due, e.g., to changes in market conditions and entry or exit of firms

[13] Let t_f^0 denote the first quarter when f appears in our data. Starting from a base period P_{f0}, which we can normalize to one, prices are concatenated using the formula:

$$P_{ft} = P_{f0} \prod_{\tau = t_f^0 + 1}^{t} \left(\frac{P_{f\tau}}{P_{f\tau-1}}\right).$$

The normalization of the level of the firm's price index in the base year, P_{f0}, is one rationale for the inclusion of firm fixed effects in our empirical specifications.

[14] As with the firm's price index, the level of the price index of competitors is constructed by normalizing the first period to one and concatenating quarterly changes. Also, in this case, the normalization is immaterial for estimation purposes as our empirical model always includes firm fixed effects.

follows:

$$MC^n_{ft} = (1 + \nu_f)\frac{TVC_{ft}}{Y_{ft}}. \tag{14}$$

Taking logs of the formula, we obtain a proxy of firms' log nominal marginal costs that has a measurable counterpart in our data.

We measure total variable costs (TVC_{ft}) as the sum of intermediate costs (materials and services purchased) and labor costs (wage bill). 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.[15] Finally, returns to scale are not directly observable in the data. By applying a logarithmic transformation to equation (14), the inverse of the short-run returns to scale parameter, $\ln(1+\nu_f)$, enters as an additive term in our specifications. As we explain below, in the empirical analysis we control for this term using fixed effects.

4 Identification strategy

In this section, we present the identification strategy that enables us to take our theoretical framework to the data. We show how to connect theoretical reset prices to observed prices to obtain forward-looking pricing equations that have measurable counterparts. Within this framework, we estimate dynamic pass-through regressions, which identify the structural parameters that pin down the slope of the primitive NKPC.

[15] Specifically, 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 we have shown, the majority of the firms in our data operate in only one industry, and the sales of those firms that produce goods in multiple industries are typically concentrated mainly in their primary industry. Results are robust to defining $\bar{P}_{ft}$ as the price of the main industry or using other aggregation methods (such as an arithmetic average or a CES aggregator).

4.1 Mapping the model to the data

Consider a firm entering period t before knowing whether or not it will be able to change its price. Under the Calvo framework, the conditional expectation of the observed price given the information set at time t, $\mathcal{I}_t$, is given by:

$$\mathbb{E}\{p_{ft}|\mathcal{I}_t\} = (1-\theta)p^o_{ft} + \theta p_{ft-1}. \tag{15}$$

We define the sampling error $v_{ft} := p_{ft} - \mathbb{E}\{p_{ft}|p^o_{ft}, p_{ft-1}\}$ and map the conditional expectation to realized values as follows:

$$p_{ft} = (1-\theta)p^o_{ft} + \theta p_{ft-1} + v_{ft}.$$

Leveraging the rational expectations assumption, we replace for p^o_{ft} using the equation of the reset price (equation 8) to obtain to following population regression (up to a constant):

$$p_{ft} = (1-\theta)\left((1-\Omega)(mc^n_{ft})^\infty + \Omega(p^{-f}_{it})^\infty\right) + \theta p_{ft-1} + \varepsilon_{ft}, \tag{16}$$

where $(x_t)^T := (1-\beta\theta)\sum_{\tau=0}^{T-1}(\beta\theta)^\tau x_{t+\tau} + (\beta\theta)^T x_{t+T}$ denote the discounted present values of $x_t = \{mc^n_{ft}, p^{-f}_{it}\}$ up to time T. The residual ε_{ft} is given by

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

where e_{ft} denotes an expectational error such that $\mathbb{E}_t(e_{ft}) = 0$, and u_{ft} is the firm's demand shock that enters the equation for the reset price.

4.2 Identification

Baseline model. To construct the sample analog of the population regression (16), we calibrate the discount factor $\beta = 0.99$, a standard value at the quarterly frequency, and truncate present values after $T = 8$ quarters, which is a sufficiently distant period to ensure that the discount factor $(\beta\theta)^\tau$ is approximately zero for $\tau > T$. We then augment the regression model to include sector-by-time fixed effects ($\alpha_{s\times t}$) and firm fixed effects (α_f), leading to the following dynamic pass-through regression:

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

The inclusion of fixed effects addresses several identification issues that generally complicate the identification of the NKPC with aggregate data. The sector-by-time fixed effects extend the theoretical framework to allow for sector-specific trends or time-varying steady states of the variables in the data. In addition, they address the concerns related to shifts in long-term inflation expectations discussed in Hazell et al. (2022). The firm fixed effects account for heterogeneity in the returns to scale in production (equation 14) and the normalization of the price level (footnote 13).

Model A nests the long-run model estimated in Amiti et al. (2019) as a special case when prices are flexible (i.e., $\theta \to 0$). As in long-run pass-through regressions, the ratio of the coefficients on marginal cost and competitors' prices identifies the degree of real rigidities Ω. Unlike long-run models, in short-run pass-through regressions the sum of the two coefficients $(1-\theta) \leq 1$ is pinned down by the degree of nominal rigidities. Moreover, the firm's lagged price enters the empirical specification as a control for the short-run dynamics of prices.

Instruments. We estimate Model A via Generalized Method of Moments (GMM).[16] To address potential endogeneity and measurement issues, we impose orthogonality conditions between the error term of the pricing equation and a set of supply-side instruments that closely relate to the ones used in Amiti et al. (2019).

Competitors' price index—Competitors' prices are jointly determined with a firm's own price and thus correlate with the firm's demand shocks entering the error

[16]The GMM estimation procedure follows the two-step approach. To ensure that our estimates are representative from a macroeconomic standpoint, we weigh observations using their Törnqvist weight, $\bar{s}_{ft}$, thereby ensuring that each firm is assigned the same weight as in the construction of aggregate price indexes. The moment conditions take the form $\mathbb{E}\{Z_{ft} \cdot \epsilon_{ft}\} = 0$, where the vector Z_{ft} includes the set of instruments, the lagged price control, and the unit vector. We cluster standard errors at the sector level to account for the potential correlation structure of error terms across firms in similar businesses. This choice is conservative but appropriate since it accounts for the possibility of correlated shocks within the sector.

term. To tackle this issue, we construct two instruments for p_{it}^{-f} that leverage variation in international trade prices.

Denote by $\mathcal{F}_{ki}^{\star}$ the set of international competitors of firm $f \in \mathcal{F}_i$ that operate in country k and industry i. The first instrument, denoted by $p_{it}^{\star EU}$, is a shifter of the price of Euro area international competitors. The rationale is that the average price charged by international competitors correlates with their marginal cost of production but not with demand shocks in Belgium. We use the Comext dataset from Eurostat to compute the (sales-weighted) average log price change that each Euro-area country k charges to the rest of the world for exported goods in industry i. We exclude Belgium from that average and denote it by $\Delta \bar{p}_{kit}^{-B}$. Then, we compute an index by averaging across all competitors j and all EU countries:

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

where the weight w_{jt} is obtained by normalizing the Törnqvist weight in formula (13) by the market share of EU competitors in industry i:

$$w_{jt} := \bar{s}_{jt}^{-f} \cdot \frac{\sum_{k \in EU} \sum_{j \in \mathcal{F}_{ki}^{\star}} \bar{s}_{jt}^{-f}}{\sum_{k} \sum_{j \in \mathcal{F}_{ki}^{\star}} \bar{s}_{jt}^{-f}}.$$

Finally, we concatenate $\Delta p_{it}^{\star EU}$ to obtain the instrument $p_{it}^{\star EU}$ in levels as before.

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 country k and Belgium:

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

where the weight w_{jt} is now scaled by the market share of non-EU competitors:

$$w_{jt} := \bar{s}_{jt}^{-f} \cdot \frac{\sum_{k \notin EU} \sum_{j \in \mathcal{F}_{ki}^{\star}} \bar{s}_{jt}^{-f}}{\sum_{k} \sum_{j \in \mathcal{F}_{ki}^{\star}} \bar{s}_{jt}^{-f}}.$$

Again, we concatenate $\Delta p_{it}^{\star F}$ to obtain the instrument $p_{it}^{\star F}$ in levels. Here the exclusion restriction requires that Euro exchange rates are orthogonal to domestic demand shocks.

Marginal cost—The primary concern with our proxy of marginal cost is measurement error, which can result in attenuation bias. Further, to the extent that firms operate with decreasing returns to scale in production, there could be concerns about the correlation of marginal cost with the firm-level demand shock in the error term. We address these issues using an instrument that leverages supply-side variation in marginal costs due to shifts in the price of intermediate inputs sourced from foreign suppliers, which we denote by $mc^{\star}_{ft}$:

$$mc^{\star}_{ft} = \sum_{k \in \mathcal{S}^{\star}_{f}} \omega_{i0} \cdot p^{\star}_{kt},$$

where $\mathcal{S}^{\star}_{f}$ denotes the set of international suppliers of firm f; ω_{i0} the share of imported inputs from supplier k in total variable costs of firm f, measured in the first period in which the firm appears in our dataset; and $p^{\star}_{kt}$ the log of supplier k's price.

Additionally, to improve the efficiency of our inference, we augment the set of instruments with a "long" lag of marginal cost (8 quarters), mc^{n}_{ft-8}. As shown by Montiel Olea and Plagborg-Møller (2021), in dynamic settings with persistent shocks, it is important to include lags in the set of instruments to produce robust estimates that have correct asymptotic coverage uniformly over the persistence in the data-generating process. Given our estimates of nearly constant returns to scale (see section 5.1), lagged marginal cost satisfies the exclusion restriction for a reasonable die-out rate of demand shocks. We confirm the validity of our instruments with formal tests in the next section.

5 Estimation results

Structural estimates. Column (A) in Table 2 presents the estimates of our baseline model. We begin by assessing the power of our instruments. In Panel a, we regress the present values of marginal cost and competitors' prices on the set of instruments, essentially producing what would be the first-stage regressions of a linear two-stage least squares model. 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 reject the hypothesis of weak identification at standard confidence levels. Moreover, 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. These findings highlight the benefits of estimating the slope using microdata, which overcomes the endogeneity and lack of instruments' power commonly encountered with aggregate time-series data (Mavroeidis et al. 2014).

Panel b reports the structural estimates for the degrees of nominal and real rigidities obtained estimating Model A via GMM. Our estimates indicate a substantial degree of price stickiness. We find a precisely estimated value of $\theta = 0.702$. Through the lens of a Calvo model, this implies that, on average, prices remain fixed for approximately three to four quarters. These estimates are remarkably consistent with the frequency of price adjustments measured by Nakamura and Steinsson (2008) from US PPI data and with the one obtained from Belgian PPI data. In section 9, we return to the interpretation of θ and consider its mapping to the observed frequency of price adjustment in a framework with menu costs.

Our estimates also reveal an economically meaningful role of strategic complementarities in the pass-through of shocks. The estimate of Ω is 0.556 and is precisely estimated. This estimate aligns with the one obtained by Amiti et al. (2019) in a long-run model with flexible prices, indicating that the pass-through from marginal costs and from competitors' prices are roughly of the same magnitude. However, it is important to stress how, in an environment with sticky prices and forward-looking pricing behavior, the short-run pass-through of marginal costs depends on both the degree of strategic complementarities and the degree of nominal rigidities. Specifically, the elasticity of a firm's own price to a permanent shock to marginal cost is given by $\frac{\partial p_{ft}}{\partial mc^n_{ft}} = (1 - \Omega)(1 - \theta)$, which is approximately equal to 0.135 at the estimated parameter values.

The slope of the primitive NKPC. Using the structural parameters estimates, we recover the slope of the marginal cost-based NKPC, presented in Panel c. We find an economically meaningful relation between fluctuations in marginal costs and aggregate inflation dynamics. The estimated slope is $\lambda = 0.057$, precisely estimated and statistically different from zero.

These estimates stand in stark contrast with the available estimates of the NKPC slope featuring the output gap or unemployment as a measure of real economic activity. These estimates typically display a magnitude that is two and a half to ten times smaller. For instance, Rotemberg and Woodford (1997) and Hazell et al. (2022) find a κ of 0.024 and 0.0062, respectively, for US data. In section 7, we return to this comparison and provide empirical evidence that helps reconcile why inflation appears to be much more responsive to marginal cost fluctuations than to changes in output or employment.

Alternative specifications. We now evaluate the robustness of our findings. We present estimates of the parameters of interest obtained from two alternative empirical models that rely on different assumptions and exploit different margins of variation in the data.

The first concern we address is related to a possible mismeasurement of the competitors' price index. Our baseline measure assumes that the set of competitors corresponds to all other firms operating in the same four-digit industry. However, it is possible that some relevant competitors might operate outside of the perimeter of the industry. To address this concern, we include a set of industry-by-time fixed effects, which agnostically absorb the present value of competitors' prices.[17] In addition, as with the sector-by-time fixed effects that we used earlier, these narrower fixed effects also control for trends. In this way, we obtain the following empirical model:

$$p_{ft} = (1-\theta)(1-\Omega)(mc^{n}_{ft})^{8} + \theta p_{ft-1} + \alpha_f + \alpha_{i\times t} + \varepsilon_{ft}, \qquad \text{(Model B)}$$

[17] Since most firms are small compared to their industry, over 90% of the variation in each firm's competitor price index occurs at the industry-year level.

Table 2: Estimation Results

Model:	(A)		(B)	(C)
	Panel a: First stages			
Endog. Var.	$(mc^n_{ft})^8$	$(p^{-f}_{it})^8$	$(mc^n_{ft})^8$	mc^n_{ft}
mc^n_{ft-8}	0.131	0.017	0.127	0.286
	(0.025)	(0.012)	(0.018)	(0.061)
$mc^{\star}_{ft}$	0.068	-0.007	0.045	0.046
	(0.025)	(0.025)	(0.030)	(0.029)
$p^{\star EU}_{it}$	0.141	0.584		
	(0.052)	(0.052)		
$p^{\star F}_{it}$	0.128	0.585		
	(0.045)	(0.054)		
p_{ft-1}	0.220	0.135	0.252	0.305
	(0.049)	(0.035)	(0.028)	(0.014)
Cragg-Donald F	444		971	2690
Kleibergen-Paap F	15.5		29.4	15.5
Hansen-Sargan J	5.919		1.053	5.086
	Panel b: Structural estimates			
θ	0.702		0.685	0.706
	(0.025)		(0.011)	(0.011)
Ω	0.556		0.605	0.466
	(0.074)		(0.093)	(0.088)
ρ				0.800
				(0.017)
	Panel c: Slope of the Phillips curve			
λ	0.057		0.059	0.067
	(0.018)		(0.020)	(0.016)
Firm FE	y		y	y
Sect. x time	y			
Ind. x time			y	y

Notes. This table presents the empirical estimates of models A, B, and C. For each model, Panel a reports the estimates of linear regressions of the endogenous variables on the exogenous instruments and controls. Panel b reports the GMM estimates of the structural parameters. Panel c reports the slope of the Phillips curve (λ) implied by the estimated parameters. The discount factor is calibrated to $\beta = 0.99$. All models are estimated using the complete sample ($N = 132,915$). In all regressions, observations are weighted using Törnqvist weights. Robust standard errors (reported in parenthesis) are clustered at the sector level.

where $\alpha_{i\times t}$ is an industry-by-time fixed effect. Column (B) presents the first-stage regression estimates, structural parameters estimates, and the implied NKPC slope (corresponding to Panels a, b, and c, respectively). Both the estimated degree of price stickiness and the degree of strategic complementarities are notably stable and precisely estimated. Consequently, the implied slope of the Phillips curve is nearly identical to the one of our baseline model.

A valuable feature of both Model A and Model B is that they do not impose stringent constraints on how firms form expectations about the dynamics of future marginal costs and industry prices. The flip side of this flexibility is that the estimating equations are highly nonlinear because θ enters the estimating equation both as a coefficient in front of the present values and lagged prices as well as in the construction of the discounted present values, which might be demanding on the data. 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 $\rho < \frac{1}{\beta\theta}$. This allows us to estimate the following system of linear equations:

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

where the pass-through of transitory shocks to marginal cost into prices is:

$$\Psi^{mc} := (1-\theta)(1-\Omega)\frac{1-\beta\theta}{1-\beta\theta\rho}.$$

Column (C) presents the estimation results. These estimates are in line with the ones obtained from the nonlinear GMM specifications, but even more precisely estimated. They imply a price elasticity to a transitory increase in marginal cost of approximately $\frac{\partial p_{ft}}{\partial mc^n_{ft}} = \Psi^{mc} = 0.093$.

5.1 Robustness to returns to scale

We derived our benchmark model under the assumption that short-run returns to scale are constant on average. However, if the economy exhibits aggregate decreasing (increasing) returns to scale, firms' price adjustments in response to changes in economic activity would be more modest (amplified), resulting in a flatter (steeper) slope of the Phillips curve (see, e.g., Galí 2015). We now investigate the importance of this channel for our results.

In the general case with arbitrary aggregate returns to scale ($\nu \neq 0$), the slope of the Phillips curve can be expressed as follows:[18]

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

where the additional term $\Theta := \frac{1}{1+\gamma\nu(1-\Omega)}$ captures the role of macroeconomic complementarities that stem from decreasing returns, with ν inversely related to short-run average returns to scale and γ denoting the elasticity of substitution across goods within industries.

In Appendix B.3 we provide empirical evidence indicating that our results are robust to empirically reasonable departures from the assumption of aggregate constant short-run returns to scale. Following Lenzu et al. (2023), we use data on physical quantities sold and inputs used to perform a gross output production function estimation that gives us sector-specific estimates of returns to scale. Our estimates indicate that the returns to scale of the different sectors, and consequently in the aggregate, are close to unity. Specifically, the sectoral estimates range from 0.86 to 1.02, while the aggregate returns to scale are estimated to be approximately 0.96. This implies a value of ν of approximately 0.04.

Calibrating γ to 4 to obtain a gross aggregate steady-state markup between 1.3 and 1.4, and using our baseline estimate of $\Omega = 0.55$, we obtain a value of $\Theta = 0.94$. Thus macroeconomic complementarities imply a reduction of the slope of seven percent relative to our baseline estimate (from 0.06 to 0.056), which is well within the confidence bounds of our baseline estimates.

[18] See Appendix A.2 for derivations.

6 Aggregate Inflation dynamics

In this section, we assess the capacity of our estimated model to capture the aggregate times series of inflation for the Belgian manufacturing sector.

To derive an expression for aggregate inflation, we use the equation for the price index (equation 9) and the equation for the reset price (equation 10). We then close the model by assuming that nominal marginal cost follows a random walk with drift.[19] We therefore obtain the following reduced-form expression for quarterly inflation (see Appendix A.3 for derivations):

$$\pi_t = \tilde{\lambda}\left(mc_t^n - p_{t-1}\right) + \alpha + \theta u_t, \tag{17}$$

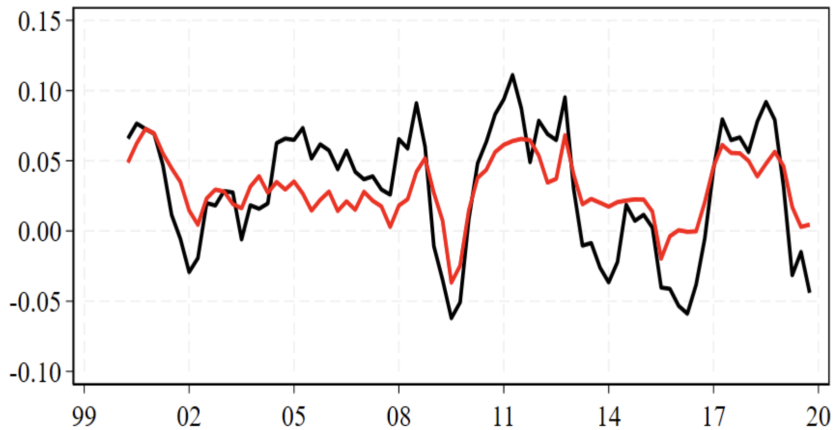
where $\tilde{\lambda} \equiv \tilde{\lambda}(\theta, \Omega)$ is an analytical function of the structural parameters, α captures trend inflation, and u_t is the aggregate cost-push shock. According to equation (17), quarterly inflation is increasing in current nominal marginal cost scaled by the lagged price level, consistent with the theory presented earlier. As before, the sensitivity of inflation depends upon the primitive pricing parameters, θ and Ω. We combine lags of equation (17) to derive an expression in terms of year-over-year inflation, which depends on a four-quarter moving average of nominal marginal cost, scaled by the price level:

$$\pi_t^{\text{y-y}} = \sum_{\tau=0}^{3} \tilde{\lambda}(1-\tilde{\lambda})^{\tau}(mc_{t-\tau}^n - p_{t-4}) + \alpha^{\text{y-y}}. \tag{18}$$

The black line in Figure 1 plots year-over-year producer-price inflation for the Belgian manufacturing sector from PRODCOM. The red line in Figure 1 depicts the model-implied inflation series. The difference between the black and red lines is the component of inflation due to the cost-push shock u_t.

[19]This 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, with 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 the model-implied one. The black line represents manufacturing producer price inflation in the data. The red line is the model-implied manufacturing producer price inflation constructed as in equation (18).

As we can see, this parsimonious model effectively tracks the broad swings in Belgian manufacturing inflation over our sample period. It accounts for almost seventy percent of the variation in inflation ($R^2 = 0.68$) with a correlation of 0.8. Particularly noteworthy is its ability to capture the drop in inflation during the 2008 financial crisis and the sharp run-up in 2016 and subsequent decline. Additionally, the model successfully captures the consistent decline in inflation from 2011 to 2016, although it does not fully capture its amplitude.

Note that, within our framework, unobservable cost-push shocks account for a much lower fraction of inflation volatility than typically found in the quantitative literature.[20] In addition, we purposely chose to compare the data against the simplest possible framework. For instance, we did not account for other forces that would further help 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.

[20]For example, in Primiceri et al. (2006), cost-push shocks arising from variation in the desired price and wage markups account for about 70% of inflation volatility.

7 Reconciliation with the conventional NKPC

In this section, we reconcile our estimates of a high slope 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 cost, prices, and the output gap at the firm level. Under these assumptions, the output gap-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 gap-based slope, which is accounted for by a low elasticity of marginal cost to changes in output.

7.1 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 marginal cost through the impact of labor demand on wages (see e.g., Galí 2015).

In particular, 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 suppose industry real wages are flexible and increasing in current industry output y_{it}. In addition, firm marginal product of labor depends inversely on firm output y_{ft} and positively on firm productivity z_{ft}, where the latter may contain both an aggregate and an idiosyncratic component:

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

In the equation above, σ^w denotes the elasticity of real wages with respect to industry output and the parameter ν varies inversely with the short-run returns to scale in production. The presence of industry output captures the influence of general equilibrium effects on 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 shocks.

Without loss of generality, we can write log firm output as the sum of log industry output and idiosyncratic supply (ϵ^s_{ft}) and demand (ϵ^d_t) shocks:

$$y_{ft} = y_{it} + \epsilon^s_{ft} + \epsilon^d_{ft}.$$

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

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 level of y_{it} in the equilibrium with flexible prices and wages, such that the desired markup is constant. The natural level y^*_{ft} is defined similarly by also taking into account the idiosyncratic firm supply shock:

$$y^*_{ft} := y^*_{it} + \epsilon^s_{ft}.$$

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

$$\widehat{mc}_{ft} = \sigma^y (y_{ft} - y^*_{ft}) - \sigma^w \epsilon^d_{ft}, \tag{19}$$

where the coefficient $\sigma^y := \sigma^w + \nu$ is the elasticity of marginal cost with respect to the output gap. The error term $\sigma^w \epsilon^d_{ft}$ accounts for the fact the wages depend only on the industry component of firm demand and not the idiosyncratic component.

To derive a pricing equation in terms of output that allows us to identify σ^y and therefore κ, we rearrange equation (19) and substitute for mc^n_{ft} into Model A from Section 4.2. 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:

$$p_{ft} = \Psi^y \cdot \sigma^y y^n_{ft} + \Psi^p p^{-f}_{it} + \theta p_{ft-1} + \alpha_f + \alpha_{s\times t} + \varepsilon^p_{ft}, \qquad \text{(Model D)}$$

where the coefficients Ψ^y and Ψ^p depend on the persistence of shocks:

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

The error term $\varepsilon^p_{ft} := (1-\sigma^w)\varepsilon^d_{ft} - \sigma^y y^*_{ft}$ depends on the idiosyncratic demand shock and the (unobservable) natural level of output.

A complementary way to identify σ^y is to rewrite equation (19) in terms of nominal marginal cost. Then, taking first differences we obtain:

$$\Delta mc^n_{ft} = \sigma^y \Delta y^n_{ft} + \alpha_f + \alpha_{s\times t} + \varepsilon^{mc}_{ft}, \qquad \text{(Model E)}$$

where the price level is absorbed into the sector-by-time fixed effects and the error term $\varepsilon^{mc}_{ft} := -\sigma^w \Delta\epsilon^d_{ft} - \sigma^y \Delta y^*_{ft}$. Unlike Model D, which directly maps into the dynamic pass-through framework developed in section 4.2, Model E allows us to directly estimate the elasticity of interest from contemporaneous changes of marginal cost and output.

7.2 Identification of σ^y and κ

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

To tackle this issue, we follow the literature that estimates output and unemployment gap-based NKPCs by exploiting shifts in aggregate demand. We use high-frequency monetary policy shocks for the Euro area. We then construct a Bartik-style instrument that allows us to improve the power of the aggregate shocks and be able to include sector-by-time fixed effects.

In detail, for each industry i, we estimate the sensitivities to aggregate demand shock (ψ_i) by projecting firm-level nominal value-added output on the lagged monetary policy shock (MS_{t-1}):[21]

$$y^n_{ft} = \alpha_f + \psi_i MS_{t-1} + \epsilon^m_{ft},$$

We then obtain our demand-side instrument by interacting the aggregate money shock with the estimated sensitivity,

$$y^{IV}_{ft} := \hat{\psi}_i \cdot MS_{t-1}.$$

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 both 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 3 presents the resulting estimates for the output elasticity of marginal cost and the implied estimates for the slope of the output gap-based Phillips curve. For model D, we find a value of $\sigma^y = 0.355$ and $\kappa = 0.020$. For model E, we find even smaller estimates, $\sigma^y = 0.213$ and $\kappa = 0.012$. The low sensitivity of marginal cost to movements in output is consistent, for example, with a high degree of wage rigidity observed in the data (Alvarez et al. 2006). These low estimates of κ corroborate the findings in previous literature, which concluded that the slope of the output gap-based and unemployment-based NKPC appears to be flat.

Taken together, the evidence presented in this section helps us reconcile our Phillips curve estimates with the estimates of the conventional output-based

[21] Monetary policy shocks are constructed following Gürkaynak et al. (2005) as the log-change in the price of overnight index swaps in a narrow window around ECB monetary policy announcements. The time series of aggregate money shocks are taken from Altavilla et al. (2019).

Table 3: Estimates of the output gap-based slope

Model:	**(D)**	**(E)**
	Panel a: Estimates	
σ^y	0.355	0.213
	(0.149)	(0.045)
ρ^y	0.853	
	(0.061)	
ρ^p	0.911	
	(0.005)	
Cragg-Donald F	86	351
Kleibergen-Paap F	10.6	43.5
	Panel b: Slope of the output gap PC	
κ	0.020	0.012
	(0.008)	(0.003)
Firm FE	y	y
Sect x time FE	y	y

Notes. This table presents the empirical estimates of models D and E. All models are estimated using the complete sample. Observations are weighted using Törnqvist weights. Robust standard errors (reported in parenthesis) are clustered at the industry-by-time level. κ is obtained from the estimates of σ^y and calibrating $\lambda = 0.057$.

Phillips curve available in the literature. Specifically, 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. Note that the low marginal cost to output elasticity also suggests that demand-side shocks likely generate weaker contemporaneous inflationary pressures than supply-side shocks.

These considerations call for further theoretical and empirical work focusing on understanding the structural relationship between output and marginal cost, particularly so given that the elasticity connecting the two could be time-varying and possibly nonlinear.

8 Oil shocks, marginal cost and inflation

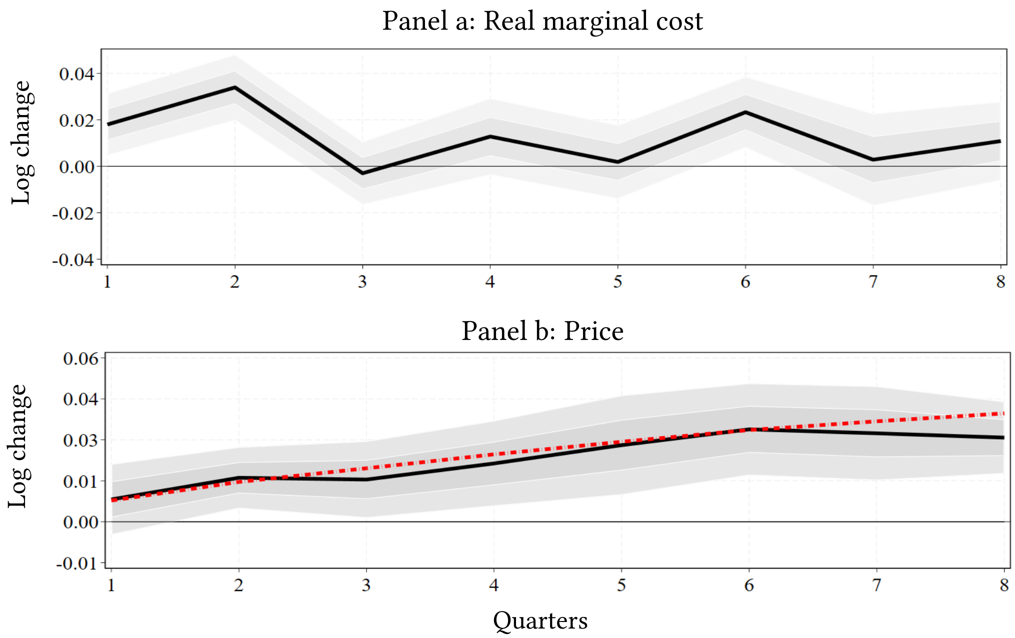
The recent surge in inflation has reignited the debate concerning the role of supply shocks in driving price dynamics. The discussion in the previous section highlighted that it is challenging to assess the impact of such shocks on inflation through the lenses of the output-based Phillips curve, without relying on a fully specified macroeconomic model that explicitly addresses the endogeneity of the natural level of output. The marginal cost-based Phillips curve does not suffer from this as the impact of supply shocks on marginal cost is measurable, and it can therefore be used to quantify the pass-through of supply shocks to prices.

To illustrate this point, we estimate the effect of identified oil price shocks on real marginal cost and inflation. Subsequently, we contrast the empirical impulse response functions with the theoretical generated by our model, calibrated to match to our earlier estimates. Following Känzig (2021), we measure oil shocks as the unexpected movements in oil price futures the day after an OPEC meeting.

Panel a in Figure 2 shows the response of aggregate real marginal cost to a one-standard-deviation shock that increases the price of Brent crude oil by roughly fifteen percent. In response to this shock, real marginal cost rises by two to two and a half percent during the first two quarters. It then gradually returns to its pre-shock level. The black line in Panel b shows that the oil shock also has substantial effects on the price level, which displays a delayed but significant and rather persistent three percent increase.

In Panel b, we assess the ability of our estimated marginal cost-based Phillips curve to reproduce the inflation response to the oil shock. We perform the following exercise. We assume firms have perfect foresight and feed into our marginal cost-based Phillips curve a path of the expected real marginal cost that matches the one generated by the impulse response of marginal cost in Panel a. We then compute the implied price dynamic and plot it in Panel b (red dotted line). As we can see, the model performs well in capturing the inflation dynamics induced by the oil shock. The model-based impulse-response function consistently lies within the confidence bands of the impulse response estimated in the data.

Figure 2: Dynamic effects of oil shocks



Notes. This figure shows the impulse response function of real marginal cost and price level to aggregate oil shocks estimated via local linear projections. The plot reports the coefficients b_h from the regressions $x_{ft+h} - x_{ft-1} = a_f + b_h OS_{t-1} + \epsilon_{ft+h}$ for $x \in \{mc, p\}$ and $h = 1, \cdots, 8$ quarters. The impact is normalized to a 15.7% increase in Brent crude oil price (one standard deviation). The dark (light) gray shaded areas are 68 (95) percent confidence bands obtained from Newey-West standard errors with four quarters of correlation. The red line is the model-based response of prices calculated by feeding in the path of marginal cost (with perfect foresight) to a Phillips curve, calibrated with $\lambda = 0.057$ and $\beta = 0.99$. All the regressions are weighted using Törnqvist weights.

This exercise provides additional validation of our empirical estimates and also demonstrates the utility of the marginal cost-based Phillips curve in tracing the effects of supply shocks on inflation.

9 Extension to menu costs

Our baseline framework presumes that firm pricing behavior is time dependent á la Calvo. With menu costs, the degree of nominal rigidities may differ from the frequency of price adjustment due to selection in price setting, as we elaborate shortly. Under certain reasonable conditions, we that our firm-level pricing regressions correctly identify the degree of nominal rigidities, and therefore our estimates of the slope of the Phillips curve remain valid.

In the conventional menu cost framework, the fixed cost of adjustment gives rise to an endogenous inaction region around the target price that is bounded by "Ss bands." As a result, price adjustments can be broken down into two components: shifts in the reset price given the adjustment frequency (the intensive margin) and shifts in the adjustment frequencies that correspond to shifts in the Ss bands (the extensive margin). As discussed by Caballero and Engel (2007), the extensive margin gives rise to a selection effect, wherein firms farthest away from their target price are more likely to adjust. The selection effect implies that, for a given price adjustment frequency, there will be greater price flexibility in the menu cost framework compared to the corresponding Calvo setup, which features only intensive margin adjustments.

Despite these differences, Auclert et al. (2022) argues that, when aggregate shocks are not too large, there exists an approximate observational equivalence between models with Calvo rigidities and canonical models with menu costs (Golosov and Lucas 2007; Nakamura and Steinsson 2010).[22] In particular, a Calvo model calibrated with a "fictitious" degree of nominal rigidities $\tilde{\theta}$ serves as a good approximation of canonical menu cost models calibrated using the frequency of

[22]The authors show that, up to a first-order approximation of the price response, one can derive two "virtual hazard rates" (i.e., fictitious probabilities of keeping the price fixed between any two consecutive periods) that exactly replicate both the intensive and extensive margins of adjustments. It follows that, quantitatively, an average of the two virtual hazard rates can be used to calibrate a time-dependent model so that the aggregate response of the price level is similar to that of a menu cost model. Moreover, the average virtual hazard rate is approximately flat for a wide range of calibrations, which implies that the adjustment probability declines geometrically.

price adjustments in the data $(1-\theta)$.[23] In particular, when the equivalence result holds, we can express the conditional expectation of a firm's price change as:

$$\mathbb{E}\{p_{ft} - p_{ft-1} | \mathcal{I}_t\} \approx \underbrace{(1-\theta)(p^o_{ft} - p_{ft-1})}_{\text{Calvo term}} + \underbrace{(\theta - \tilde{\theta})(p^o_{ft} - p_{ft-1})}_{\text{Selection term}}, \tag{20}$$

The selection term captures the fact that adjusting firms are not a random sample of the population, but are exactly those whose reset price is farthest from their price in the previous period. Rearranging equation (20) leads to a population regression equivalent to equation (15) in Section 4.1, but with $\tilde{\theta}$ replacing θ:

$$\mathbb{E}\{p_{ft} | \mathcal{I}_t\} \approx (1-\tilde{\theta}) p^o_{ft} + \tilde{\theta} p_{ft-1}. \tag{21}$$

Next, combining equation (21) with the expression for the optimal reset price in (8) leads to the following generalized Phillips curve under menu costs:

$$\pi_t \approx \tilde{\lambda}\, \widehat{mc}_t + \beta\, \mathbb{E}_t\{\pi_{t+1}\} + u_t, \qquad \text{where} \quad \tilde{\lambda} := \frac{(1-\tilde{\theta})(1-\beta\tilde{\theta})}{\tilde{\theta}}(1-\Omega).$$

Note that this generalized slope takes the same form as under Calvo, but with $\tilde{\theta}$ replacing θ. Since $\tilde{\theta} \leq \theta$ because of selection, the slope is steeper. Therefore, when the approximate equivalence holds, our empirical methodology remains valid because it correctly identifies the degree of nominal rigidities and hence the slope of the generalized Phillips curve.

Finally, we calculate the average frequency of price changes using PPI microdata and obtain an estimate for $(1-\theta) = 0.32$. As this number closely matches our estimates of $\tilde{\theta} \approx 0.7$, selection does not appear to play a major role in our sample. This is further confirmed by evidence on the observed kurtosis of price changes (Alvarez et al. 2016). Using again PPI data, we calculate the kurtosis to be 5.4, which is in line with the kurtosis produced by standard Calvo models (around 6) and larger than canonical menu cost models (between 1 and 3).

[23] See also Gertler and Leahy (2008) for conditions under which an exact equivalence result holds.

10 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 a relevant measure of economic activity. We observe a high pass-through of marginal cost into prices, as evidenced by both the microdata and the ability of the marginal cost-based Phillips curve to track aggregate inflation dynamics. We have also shown that a low elasticity of marginal cost to output can reconcile the low sensitivity of aggregate inflation to output (or employment) with the high pass-through of marginal cost.

Though our analysis is based on pre-pandemic data, it also offers useful insights for the current surge in inflation. First, recent research has shown that supply-side shocks are an important driver of the current inflation surge (e.g., Di Giovanni et al. 2022). We illustrated with the example of oil shocks how, unlike the conventional formulation, the primitive NKPC provides a convenient way of examining the transmission of supply shocks to inflation.

Secondly, recent research suggests that the sensitivity of inflation to output might have increased. For example, Benigno and Eggertsson (2023) argue that the slope of the Phillips curve has risen. Our analysis suggests that a candidate explanation is an increase in the elasticity of marginal cost with respect to output, due for example to supply-side constraints. Understanding the primitive drivers of this elasticity and how it may evolve over time is a fruitful topic for future research.

More broadly, our framework characterizes inflation dynamics conditional on the path of marginal cost. A logical next step is to endogenize the behavior of marginal cost and then use microdata to estimate this relationship. As we discussed, wage rigidity and/or other labor market frictions rule out a simple log-linear relation between real marginal cost and the output gap. When both nominal wages and output prices are set for multiple periods, output will still influence marginal cost, but not in the simple way implied by the conventional NKPC formulation.[24] Accordingly, modeling the evolution of marginal cost in a way that also factors in wage dynamics is an important task for the future.

[24] See Lorenzoni and Werning (2023) for a recent discussion.

References

ALTAVILLA, C., L. BRUGNOLINI, R. S. GÜRKAYNAK, R. MOTTO, AND G. RAGUSA (2019): "Measuring euro area monetary policy," *Journal of Monetary Economics*, 108, 162–179.

ALVAREZ, F., H. LE BIHAN, AND F. LIPPI (2016): "The real effects of monetary shocks in sticky price models: a sufficient statistic approach," *American Economic Review*, 106, 2817–2851.

ALVAREZ, L. J., E. DHYNE, M. HOEBERICHTS, C. KWAPIL, H. LE BIHAN, P. LÜNNEMANN, F. MARTINS, R. SABBATINI, H. STAHL, P. VERMEULEN, ET AL. (2006): "Sticky prices in the euro area: a summary of new micro-evidence," *Journal of the European Economic Association*, 4, 575–584.

AMITI, M., O. ITSKHOKI, AND J. KONINGS (2019): "International shocks, variable markups, and domestic prices," *The Review of Economic Studies*, 86, 2356–2402.

ATKESON, A. AND A. BURSTEIN (2008): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 98, 1998–2031.

AUCLERT, A., R. D. RIGATO, M. ROGNLIE, AND L. STRAUB (2022): "New Pricing Models, Same Old Phillips Curves?" Tech. rep., National Bureau of Economic Research.

BASU, S. (1995): "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare," *American Economic Review*, 85, 512.

BENIGNO, P. AND G. B. EGGERTSSON (2023): "It's baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve," Tech. rep., National Bureau of Economic Research.

BERAJA, M., E. HURST, AND J. OSPINA (2019): "The aggregate implications of regional business cycles," *Econometrica*, 87, 1789–1833.

CABALLERO, R. J. AND E. M. ENGEL (2007): "Price stickiness in Ss models: New interpretations of old results," *Journal of Monetary Economics*, 54, 100–121.

CALVO, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 12, 383–398.

DI GIOVANNI, J., Á. KALEMLI-ÖZCAN, A. SILVA, AND M. A. YILDIRIM (2022): "Global supply chain pressures, international trade, and inflation," Tech. rep., National Bureau of Economic Research.

Gabaix, X. (2020): “A behavioral New Keynesian model,” *American Economic Review*, 110, 2271–2327.

Galí, J. (2015): *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press.

Galí, J. and M. Gertler (1999): “Inflation dynamics: A structural econometric analysis,” *Journal of Monetary Economics*, 44, 195–222.

Gertler, M. and J. Leahy (2008): “A Phillips curve with an Ss foundation,” *Journal of Political Economy*, 116, 533–572.

Goldberg, P. K. and F. Verboven (2001): “The evolution of price dispersion in the European car market,” *The Review of Economic Studies*, 68, 811–848.

Golosov, M. and R. E. Lucas (2007): “Menu costs and Phillips curves,” *Journal of Political Economy*, 115, 171–199.

Gürkaynak, R. S., B. P. Sack, and E. 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, 55–93.

Hazell, J., J. Herreno, E. Nakamura, and J. Steinsson (2022): “The slope of the Phillips Curve: evidence from US states,” *The Quarterly Journal of Economics*, 137, 1299–1344.

Hooper, P., F. S. Mishkin, and A. Sufi (2020): “Prospects for inflation in a high pressure economy: Is the Phillips curve dead or is it just hibernating?” *Research in Economics*, 74, 26–62.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016): “Quantifying the sources of firm heterogeneity,” *The Quarterly Journal of Economics*, 131, 1291–1364.

Jørgensen, P. L. and K. J. Lansing (2023): “Anchored inflation expectations and the slope of the phillips curve,” in *Federal Reserve Bank of San Francisco Working Paper 2019-27*.

Känzig, D. R. (2021): “The macroeconomic effects of oil supply news: Evidence from OPEC announcements,” *American Economic Review*, 111, 1092–1125.

Kimball, M. S. (1995): “The quantitative analytics of the basic neomonetarist model,” *Journal of Money, Credit, and Banking*, 27, 1241–1277.

Lenzu, S., D. Rivers, and J. Tielens (2023): "Financial Shocks, Productivity, and Prices," *Working paper.*

Lorenzoni, G. and I. Werning (2023): "Inflation is conflict," Tech. rep., National Bureau of Economic Research.

Mavroeidis, S., M. Plagborg-Møller, and J. H. Stock (2014): "Empirical evidence on inflation expectations in the New Keynesian Phillips Curve," *American Economic Journal: Journal of Economic Literature*, 52, 124–188.

McLeay, M. and S. Tenreyro (2020): "Optimal inflation and the identification of the Phillips curve," *NBER Macroeconomics Annual*, 34, 199–255.

Montiel Olea, J. L. and M. Plagborg-Møller (2021): "Local projection inference is simpler and more robust than you think," *Econometrica*, 89, 1789–1823.

Nakamura, E. and J. Steinsson (2008): "Five facts about prices: A reevaluation of menu cost models," *The Quarterly Journal of Economics*, 123, 1415–1464.

——— (2010): "Monetary non-neutrality in a multisector menu cost model," *The Quarterly Journal of Economics*, 125, 961–1013.

Nakamura, E. and D. Zerom (2010): "Accounting for incomplete pass-through," *The Review of Economic Studies*, 77, 1192–1230.

Primiceri, G., E. Schaumburg, and A. Tambalotti (2006): "Intertemporal disturbances," *NBER Working Paper No. 12243.*

Rotemberg, J. J. and M. Woodford (1997): "An optimization-based econometric framework for the evaluation of monetary policy," *NBER Macroeconomics Annual*, 12, 297–346.

Sbordone, A. M. (2002): "Prices and unit labor costs: a new test of price stickiness," *Journal of Monetary Economics*, 49, 265–292.

Wang, O. and I. Werning (2022): "Dynamic oligopoly and price stickiness," *American Economic Review*, 112, 2815–49.

Woodford, M. (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 for Online Publication

A Derivations

This section provides additional information and derivations of the key equations presented in Section 2. We begin showing how the markup function in the paper maps to the markup functions under two prominent frameworks featuring imperfect competition. We then present the aggregation steps followed to derive the Phillips curve.

A.1 Derivation of 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 indexes for each industry P_{it} and the aggregate price index P_t are defined, respectively, as:

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

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 elasticity of substitution within industries $\gamma > \sigma$:

$$\mathcal{D}_{ft+\tau} = \left(\frac{\varphi_{ft+\tau}P^o_{ft}}{\varphi_{it+\tau}P_{it+\tau}}\right)^{-\gamma}\left(\frac{\varphi_{it+\tau}P_{it+\tau}}{P_{t+\tau}}\right)^{-\sigma} Y_{t+\tau}. \tag{1}$$

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:

$$\epsilon_{ft+\tau} := -\frac{\partial \ln \mathcal{D}_{ft+\tau}}{\partial \ln P^o_{ft}} = \gamma - (\gamma - \sigma)\frac{\partial p_{it+\tau}}{\partial p^o_{ft}}. \tag{2}$$

We can further characterize the derivative above. First, the price index of competitors of firm f is defined as:

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

It follows that $P^{1-\gamma}_{it} = \frac{N-1}{N}\left(P^{-f}_{it}\right)^{1-\gamma} + \frac{1}{N}\left(\varphi_{ft}P^o_{ft}\right)^{1-\gamma}$. Next, using the definition of the industry price index P_{it} and denoting by $\zeta_{ft+\tau} := \frac{\partial p^{-f}_{it+\tau}}{\partial p^o_{ft}}$, its derivative with respect to the firms' reset price is given by:

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

Multiplying both sides by $\frac{P^o_{ft}}{P_{it+\tau}}$, we obtain:

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

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

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

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 is. 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)'s static Nash oligopoly, $\epsilon_{ft+\tau} = 0$ for $\tau > 0$ and $\zeta_{ft+\tau} = 0$ for all τs. Under monopolistic competition, $N \to \infty$, which implies $\zeta_{ft+\tau} \to 0$ and $s_{ft+\tau} \to 0$.

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).[25] Log-linearizing the elasticity in (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):

$$\mu_{ft+\tau} - \mu = -\Gamma\left(p^{o}_{ft} - p^{-f}_{it+\tau}\right) + u^{\mu}_{ft+\tau},$$

[25]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 that 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 the literature, implies that $\zeta_{j,\tau} = 0$ at the steady state for all js and τs. 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 js and τs that delivers, under specific values for the elasticities σ and γ, a pass-through of shocks to marginal cost into prices that is qualitatively the same as the one produced by the Nash model. Secondly, these considerations also apply to our empirical analysis, as we directly estimate the parameters (Γ, in particular) rather than the underlying elasticities.

where $\Gamma := \frac{(\gamma-\sigma)(\gamma-1)}{\epsilon(\epsilon-1)}\frac{N-1}{N} > 0$ denotes the markup elasticity with respect to prices and

$$u^{\mu}_{ft} := -\frac{(\gamma-\sigma)(\gamma-1)}{\epsilon(\epsilon-1)}\ln\varphi_{ft} + \frac{\gamma-\sigma}{\epsilon(\epsilon-1)}\frac{N-1}{N}\zeta_{ft}, \tag{4}$$

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

Finally, using these expressions, we can show how to obtain the pricing equation (8). Log-linearizing the industry price index and ignoring constants, we get:

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

Substituting in equation (6) for the markup and rearranging, we obtain:

$$p^{o}_{ft} = (1-\beta\theta)\mathbb{E}_t\left\{\sum_{\tau=0}^{\infty}(\beta\theta)^{\tau}\left((1-\Omega)(mc^{n}_{ft+\tau}+\mu) + \Omega p^{-f}_{it+\tau} + (1-\Omega)u^{\mu}_{ft+\tau}\right)\right\}, \tag{5}$$

where, as in the paper, $\Omega := \frac{\Gamma}{1+\Gamma}$. This parameter denotes the relative weight on the price index of competitors (p^{-f}_{it}) 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:

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

which is therefore a firm-specific shock that depends on the expectation of future demand shifters.

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:[26]

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

[26] For simplicity we now abstract from taste shocks.

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:

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

Denoting by ψ the Lagrange multiplier of the constraint, the first-order condition of the problem is:

$$P_{ft} = \psi \Upsilon'\left(\frac{Y_{ft}}{Y_{it}}\right) \frac{1}{Y_{it}} \tag{7}$$

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

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

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

$$\frac{P_{ft}}{P_{it}} = \frac{1}{\Upsilon'(1)} \Upsilon'\left(\frac{Y_{ft}}{Y_{it}}\right). \tag{8}$$

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

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

Taking logs of equation (A.1) and differentiating, we get the residual elasticity of demand:

$$\epsilon_{ft+\tau} := -\frac{\partial \ln \mathcal{D}_{ft+\tau}}{\partial \ln P^o_{ft}} = -\frac{\Upsilon'\left(\frac{Y_{ft+\tau}}{Y_{it+\tau}}\right)}{\Upsilon''\left(\frac{Y_{ft+\tau}}{Y_{it+\tau}}\right) \cdot \left(\frac{Y_{ft+\tau}}{Y_{it+\tau}}\right)} \tag{9}$$

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 $\epsilon = -\frac{\Upsilon'(1)}{\Upsilon''(1)}$ and by by ϵ' the derivative of the residual demand elasticity $\epsilon_{ft+\tau}$ in (9) with respect to $\frac{Y_{ft+\tau}}{Y_{it+\tau}}$, evaluated at the steady state:

$$\epsilon' = \frac{\Upsilon'(1)\left(\Upsilon'''(1) + \Upsilon''(1)\right) - \left(\Upsilon''(1)\right)^2}{\left(\Upsilon''(1)\right)^2} \leq 0. \tag{10}$$

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 Learner index around the steady state and using equation (10), we have that, up to a first-order approximation, the log-markup (in deviation from steady state) is equal to:

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

Finally, log-linearizing the demand function (A.1) and using it to replace the log difference in output, we obtain:

$$\mu_{ft+\tau} - \mu = -\Gamma\left(p^o_{ft} - p_{it+\tau}\right)$$

where, in the case of Kimball preferences, the sensitivity of the markup to the relative price is given by $\Gamma := \frac{\epsilon'}{\epsilon(\epsilon-1)}\frac{1}{\Upsilon''(1)}$.

Notice that, without loss of generality, $p_{it+\tau} = p^{-f}_{it+\tau}$ because of the continuum of firms within an industry. Substituting into the pricing equation (6) and rearranging leads to the expression equation (7).

Finally, following the same steps as the previous section, we obtain $\Omega := \frac{\Gamma}{1+\Gamma}$ and the corresponding mapping to the pricing equation in (8).

A.2 Aggregation and the Phillips Curve

Suppose $N < \infty$ and order firms in each industry from 1 to N.[27] [28] 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 $A^{\star}_{ft}$ for $f \in \{1, \ldots, N\}$ the set of industries in which the f-th firm can adjust. The price index can then be rewritten as:

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

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 their price to their optimal reset price.

Since $A^{\star}_{ft}$ 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 $f = \{1, \ldots, N\}$ across industries we have that:[29]

$$\frac{1}{N} \sum_{f=1}^{N} \int_{i \in I / A^{\star}_{ft}} p_{fit-1} di = \theta \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^{N} p_{fit-1} \right) di = \theta p_{t-1}$$

and

$$\frac{1}{N} \sum_{f=1}^{N} \int_{i \in A^{\star}_{ft}} p^{o}_{fit} di = (1-\theta) \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^{N} p^{o}_{fit} \right) di.$$

Defining $p^{o}_{t} := \int_{i \in I} \left(\frac{1}{N} \sum_{f=1}^{N} p^{o}_{fit} \right) di$ as the average reset price in the economy, we

[27] Notice 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 for example Wang and Werning (2022), appendix C2.

[28] Letting $N \longrightarrow \infty$, all results hold under Kimball preferences.

[29] The 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 $\int_{i \in [0,1]} \left(\frac{1}{N} \sum_{f=1}^{N} p^{-f}_{it} \right) di = \int_{i \in [0,1]} \left(\frac{1}{N} \sum_{f=1}^{N} \left[\frac{N}{N-1} p_{it} - \frac{1}{N-1} p_{fit} \right] \right) di = p_t$.

obtain:

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

which is equation (9) in the paper.[30]

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), under aggregate decreasing returns to scale, in log-terms we have that:

$$mc_{fit}^n = c_{it} + a_{fit} + \nu y_{fit}.$$

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

$$mc_{it}^n = c_{it} + a_{it} + \nu y_{it},$$

where ν is the average return to scale in production. Combining the two equations above and subtracting the (log) industry price index on both sides, we obtain an expression that relates real marginal costs to cost shifters and output:

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

We use the demand function to express the log output deviation, $y_{fit} - y_{it}$, in terms of log prices. In the case of CES preferences (see equation (1)), we obtain:

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

where γ denotes the within-industry elasticity of substitution.[31]

We then proceed with the following steps in order: we manipulate equation (5) to express the reset price in recursive form, decompose firm-level nominal marginal cost into firm-level real marginal cost and the industry price index prices,

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

[31]A 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, $1/\Upsilon''(1)$.

and finally use equation (A.2) to replace for firm-level real marginal cost:

$$\begin{aligned} p^o_{fit} &= (1-\beta\theta)\left((1-\Omega)(mc^n_{fit}+\mu)+\Omega p^{-f}_{it}+(1-\Omega)u^{\mu}_{fit}\right)+\beta\theta\mathbb{E}_t p^o_{fit+1} \\ &= (1-\beta\theta)\Theta\left((1-\Omega)\widehat{mc}_{it}+\Omega p^{-f}_{it}+(1-\Omega)(1+\gamma\nu)p_{it}+(1-\Omega)u^{\mu}_{fit}\right) \\ &\quad +\beta\theta\mathbb{E}_t p^o_{fit+1}+(1-\beta\theta)\Theta(1-\Omega)\left(a_{fit}-a_{it}-\gamma\nu\ln\varphi_{fit}\right), \end{aligned}$$

where $\Theta := \frac{1}{1+\gamma\nu(1-\Omega)}$ captures macroeconomic complementarities due to aggregate returns to scale in production.

Finally, averaging across firms and industries, we have that the aggregate reset price is given by:

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

where $u_t := \frac{(1-\theta)(1-\beta\theta)}{\theta}(1-\Omega)\Theta\int_{i\in I}\left(\frac{1}{N}\sum_{f=1}^{N}u^{\mu}_{fit}\right)di$ is an aggregate cost-push shock and $\left(a_{fit}-a_{it}+\gamma\nu\ln\varphi_{fit}\right)$ is such that $\int_{i\in I}\left(\frac{1}{N}\sum_{f=1}^{N}(a_{fit}-a_{it}+\gamma\nu\ln\varphi_{fit})\right)di=0$. This follows from the i.i.d. assumption on price adjustments, which implies that the average firm-level shifter of resetting firms coincides with the unconditional average.

Subtracting p_t from both sides and using the log-linearized price index:

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

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

A.3 Derivations of inflation dynamics

Ignoring the intercept, the system of equations is given by:

$$
\begin{aligned}
p_t^o &= (1-\beta\theta)((1-\Omega)mc_t^n + \Omega p_t) + \beta\theta\mathbb{E}_t p_{t+1}^o + \frac{\theta}{1-\theta}u_t, \\
p_t &= (1-\theta)p_t^o + \theta p_{t-1}, \\
mc_t^n &= mc_{t-1}^n + \varepsilon_t^{mc}.
\end{aligned} \tag{11}
$$

We guess and verify using the method of undetermined coefficients that the solution is of the form:

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

where Ξ and $\tilde{\lambda}$ are the coefficients to be determined. Plugging the guessed solution into the system gives the following restrictions on the parameters:

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

We select the solution that implies that system (11) has exactly one eigenvalue larger than one in modulus. This gives the following values for the parameters in terms of primitives:

$$
\begin{aligned}
\Xi &= \frac{\beta\theta(2-\Omega(1-\theta)-\theta) + \Omega(1-\theta) - 1}{2\beta(1-\theta)\theta} \\
&\quad + \frac{\sqrt{(-\Omega(1-\theta)(1-\beta\theta) - \beta(2-\theta)\theta + 1)^2 + 4\beta(1-\Omega)(1-\theta)\theta(1-\beta\theta)}}{2\beta(1-\theta)\theta}, \\
\tilde{\lambda} &= (1-\theta)\Xi.
\end{aligned}
$$

Rearranging the guessed solution for $p_t = \tilde{\lambda} mc_t^n + (1-\tilde{\lambda})p_{t-1} + \theta u_t$ and adding back the intercept, we obtain equation (17).

B Data and Measurement

B.1 Data sources and data cleaning

In this section we describe the different administrative sources used to assemble our micro-level data.

We use the information in PRODCOM to compute the quarterly change in product- and firm-level prices and to define the boundary 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. Processing of tea and coffee is "Processing of tea and coffee"). Sectors are defined by the first two digits of the product codes (AC is "Manufacture of food products, beverages, and tobacco products"). The industry and sector definitions follow the NACE classification as the first four digits of PRODCOM codes are identical to the first four digits of the NACE classification.

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.[32] 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. The one-to-many and many-to-many could be problematic. We are able to deal with most of these cases using information on the basket of products produced by each firm.[33] 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 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 the product-level sales by the product-level quantity sold. As explained in the paper, we are interested in domestic prices, that is prices charged by producers in Belgium. PRODCOM does not require firms to separately report distinguishing between 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 through Intrastat Inquiry (for intra-EU trade).[34] We use the official conversion tables provided by Eurostat to map the CN product codes classification used in the international trade data to the PRODCOM product code classification.[35] In the majority of the cases, the CN-to-PC conversion involves

[32]The official conversion tables are available at `https://ec.europa.eu/eurostat/ramon`. The harmonization of the industry code essentially consists in harmonizing the to NACE rev. 1 industry, used before 2008, to the NACE rev. 2 industry codes, used from 2008.

[33]For example, consider a case when 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 map 11.11.11.11 to 22.22.22.22 for the latter.

[34]Importantly, in constructing our measure of domestic sales we address issues related to carry-along-trade, which might overstate the amount of production of firms that import products that are destined for immediate sales.

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

either a one-to-one or many-to-one mapping, which poses no issues. We drop the observations that involve one-to-many and many-to-many mappings. These account for less than 5% of the observations and production value.

We apply the following filters and data manipulations to the PRODCOM data. First, we only keep firms' observations in a given quarter if there was a positive production reported for at least one product in the quarter. This avoids large jumps in the quarterly values due to non-reporting for some months by some firms. In the rare cases when a firm reports positive values but quantities are missing, we impute 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.

The second important use of international trade data is to obtain information on international competitors selling their 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, 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.[36]

We leverage data from two administrative sources to measure firms' total production (turnover) and variable production costs at a quarterly frequency. Belgian firms file VAT declarations to the Belgian tax authority that contain information on the total sales of the enterprise as well as information on purchases

[36] Some CN codes change over time (although to a smaller extent relative to the PC codes). We use the official conversion tables, also available on the Eurostat website, to map CN product codes across consecutive years. We only make an adjustment if the code is a one-to-one change between two years. We do not take into account 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.

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 proprietorship and self-employed and therefore mostly do not apply to the firms surveyed by PRODCOM.[37] We obtain information on employment and labor costs (wage bill) from the Social Security declarations filed on a quarterly basis by each Belgian firm to the Department of Social Security of Belgium.

We sum firm-quarter level expenses on intermediates and labor to obtain a measure of total variable costs, which we use in the construction of firms' marginal costs. We multiply these costs by the ratio of total manufacturing sales (from PRODCOM) over total sales (from the VAT) to adjust for the fact that some firms also have some production outside manufacturing.[38]

Finally, we apply the following data-cleaning steps to deal with missing values and outliers. (i) We focus on manufacturing industries, defined by the NACE rev.2 2-digit codes 15–36, dropping from our sample all product headings that correspond to mining and quarrying and all product codes corresponding to industrial services. (ii) We drop observations referring to firms whose sales from manufacturing products (as measured in PRODCOM) is 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 standard, we exclude firms that operate in the "Coke and refined petroleum products" sector and "Pharmaceuticals, medicinal chemical, and

[37]Enterprises file their VAT declaration online, either on a monthly or a quarterly frequency, depending on some size-based thresholds. Smaller enterprises (turnover $<$ 2.5M euros excl. VAT) can choose to file at the monthly or quarterly frequency. Larger enterprises file monthly. In the case of multiple plants or establishments under one VAT identifier, the declaration is filed as a single file for that VAT identifier. We aggregate all monthly declarations at the quarterly level. At this reporting frequency, VAT declarations tend to reflect the sales of output produced 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 7.

[38]As mentioned below, we conservatively drop observations referring to firms whose manufacturing sales are lower 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 the extensive coverage of PRODCOM.

botanical products" sector whose output prices are frequently privately bargained or determined in international markets. We also exclude firms operating in the "Other manufacturing and repair and installation of machinery and equipment" sector, which is a residual grouping that consists 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 indexes, the corresponding quantity indexes, competitors' price indexes, 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) Finally, for each firm-industry pair that enters our dataset discontinuously we keep only the longest continuous time-spell. This ensures that each time series used in the estimation has no gaps, which would force us to interpolate making assumptions about prices and marginal costs when the data is not recorded.

Table 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
CI	Computer, electronic and optical products	26
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.

B.2 Construction of price indexes

We construct a set of indexes that capture price changes in manufacturing goods at different 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 prices changes at the most disaggregated level allowed by the data.

Due to repeated product code revisions, it does not exist a consistent 8-digit product code taxonomy across the entire sample period.[39] 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, aggregate them at the firm-industry level, and recover 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 price change for each product, P_{pt}/P_{pt-1}, appropriately accounting for any change in product codes. In the construction of the product-level price changes, we drop product-level observations with abnormally large price jumps in a given quarter ($P_{pt}/P_{pt-1} > 3$ or $P_{pt}/P_{pt-1} < 1/3$). Then, we construct firm-industry price change as a Törnqvist index:

$$P_{ft}/P_{ft-1} = \prod_{p \in \mathcal{P}_{ft}} (P_{pt}/P_{pt-1})^{\bar{s}_{pt}}, \tag{12}$$

where $\bar{s}_{pt}$ is a Törnqvist weight computed as the average of the sale shares between t and $t-1$: $\bar{s}_{pt} := \frac{s_{pt}+s_{pt-1}}{2}$.[40] Finally, we use the sequence of quarterly price changes

[39] See Appendix B.1 for additional information on the data.

[40] This 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, $\bar{s}_{pt}$, give larger weights to those produces that account for a larger share of firms turnover.

to construct the time series of firm-industry's prices (in levels):

$$P_{ft} = P_{f0} \prod_{\tau=t_f^0+1}^{t} \left(P_{f\tau}/P_{f\tau-1}\right), \tag{13}$$

where t_f^0 denotes the first quarter when f appears in our data and P_{f0} is the price level in that quarter. We normalize P_{f0} to one for all firm-industry pairs f in our dataset. As discussed in the paper, this normalization is immaterial for our empirical analysis as any level-effects are absorbed by the firm-industry fixed effects included in all 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, which implies that the firm-industry price index, P_{ft}, and the firm price index, $\bar{P}_{ft}$, coincide. Yet, in a in a limited number of cases, it becomes necessary to construct a firm's price index that aggregates across different firm-industry price indexes. In doing this, we construct the firm-level price index $\bar{P}_{ft}$ following method similar to the one outlined above. Specifically, we construct a Törnqvist index that aggregates across price changes of 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). We then concatenate the quarterly price changes above to obtain the price index $\bar{P}_{ft}$, normalizing the level of the price index to one in the first quarter when the firm first appears in our dataset. Note that for single-industry firms the price index $\bar{P}_{ft}$ coincides with with the firm-industry price index P_{fit} in (13).

Competitors price index. Using a similar approach, we construct the price index of competitors for each domestic firm. We start 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 $\bar{s}_{kt}^{-f} := \frac{1}{2}\left(\frac{s_{kt}}{1-s_{ft}} + \frac{s_{kt-1}}{1-s_{ft-1}}\right)$ denoting 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

$t-1$. We then concatenate the changes normalizing the level of the price index in the first period to one. Also, in this case, the normalization is immaterial for estimation purposes as our empirical model always includes firm fixed effects. 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 indexes by aggregating quarterly firm-level price changes. The formula to construct the percentage change in these price indexes is analogous to the one in (12), where now the Törnqvist weights assigned to each firm-industry price change, P_{ft}/P_{ft-1}, captures the (weighted) average market shares of the f in its own industry, sector, or manufacturing, respectively. Once again, the level of the indexes is constructed by concatenating changes and normalizing the level of the price index to one for the first observation in the time series.

B.3 Estimates of returns to scale

We estimate the elasticities that determine the returns to scale of production (both short- and long-run) by performing production function estimations. We consider the following log-production function: $y_{ft} = \ln A_{ft} + f(l_{ft}, m_{ft}, k_{ft}; \gamma)$. Here, y_{ft} denotes firm-level output (physical quantity) produced by firm f in period t, and A_{ft} captures a firm's technical efficiency (TFPQ). $f(\cdot)$ is the log-gross output production function, which we model as a Cobb-Douglas aggregate of labor (l_{ft}), intermediates (m_{ft}), and capital (k_{ft}). The vector of structural parameters to be estimated is denoted by $\gamma := \gamma^l, \gamma^m, \gamma^k$, which collects the output elasticities of the different inputs.

Following Lenzu et al. (2023), we construct a firm-level quantity index by deflating firm-level sales by the firm-level price index: $Y_{jt} = \frac{(PY)_{ft}}{\bar{P}_{ft}}$. Labor services are measured using the wage bill, and intermediates costs are measured as the total value of materials and services used in production. A measure of the capital stock is constructed from investments in fixed assets using the perpetual inventory method. We deflate labor, capital, and intermediate inputs using the corresponding industry-level producer price deflators.

We estimate the production function separately for each sector, following the approach developed in Lenzu et al. (2023), which combines the structural approach developed in Gandhi et al. (2020) with the control function approach developed by De Loecker et al. (2016) to control for differences in input quality across firms. This approach identifies the production function parameters by addressing the simultaneity bias that derives from the correlation between input choices and unobserved (to the econometrician) productivity (Marschak and Andrews Jr. (1944)), and it solves the identification problem that affects the estimates of the output elasticities of flexible inputs.[41] In line with the rest of our analysis, we perform the production estimation for each industry by weighting observations using within-industry sales-based Törnqvist weights.

Table 2 presents the estimates of the output elasticities and returns to scale for individual manufacturing sectors and for the aggregate economy.[42] The latter is obtained as a sales-weighted average of the sectoral estimates. For our purposes, the key estimates are the ones regarding the elasticities of variable inputs, whose sum pins down the *short-run returns to scale* and determines the strength of macroeconomic complementarities. Consistent with the previous studies (see, e.g., Lenzu et al. (2023) and the references therein), our estimates indicate returns to scale in the ballpark of unity for most sectors and, therefore, in the aggregate.

[41]The details of the estimation routine are provided in the Appendix of Lenzu et al. (2023).

[42]We are unable to perform the production function estimation for a handful of the sector "Computer, electronic and optical products" (CI) due to its small sample size.

Table 2: Estimates of output elasticities and returns to scale

Sector	Output elasticities			Returns to scale	
	Labor (γ^l)	Intermediates (γ^m)	Capital (γ^k)	Long-run ($\gamma^l + \gamma^m + \gamma^k$)	Short-run ($\gamma^l + \gamma^m$)
CA	0.248	0.770	0.094	1.112	1.018
CB	0.201	0.748	0.061	1.010	0.949
CC	0.253	0.729	0.040	1.022	0.982
CE	0.080	0.794	0.124	0.999	0.874
CG	0.242	0.717	0.129	1.088	0.958
CH	0.250	0.721	0.147	1.119	0.972
CJ	0.322	0.646	0.120	1.088	0.967
CK	0.197	0.707	0.180	1.084	0.904
CL	0.149	0.796	0.075	1.020	0.945
Aggregate	0.209	0.749	0.104	1.062	0.958

Notes. This table reports the within-sector average production function elasticities estimated following the approach in Lenzu et al. (2023), as described above. The first column indicates the manufacturing sector. The subsequent three columns report the estimates obtained from a quantity production function estimation. The following two columns report the long-run and short-run returns to scale. Each row corresponds to a different manufacturing sector. The last row is a sales-weighted average of the sectoral estimates.

References

Alvarez, F. and F. Lippi (2014): "Price setting with menu cost for multiproduct firms," *Econometrica*, 82, 89–135.

Atkeson, A. and A. Burstein (2008): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 98, 1998–2031.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016): "Prices, Markups and Trade Reform," *Econometrica*, 84, 445–510.

European Commission (2014): "Quality Report on PRODCOM 2014 Annual Data," PRODCOM Working Group.

Gandhi, A., S. Navarro, and D. A. Rivers (2020): "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, 128, 2973–3016.

Lenzu, S., D. Rivers, and J. Tielens (2023): "Financial Shocks, Productivity, and Prices," *Working paper.*

Marschak, J. and W. H. Andrews Jr. (1944): "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 143–205.

Midrigan, V. (2011): "Menu costs, multiproduct firms, and aggregate fluctuations," *Econometrica*, 79, 1139–1180.

Wang, O. and I. Werning (2022): "Dynamic oligopoly and price stickiness," *American Economic Review*, 112, 2815–49.