The Hitchhiker's Guide to Markup Estimation*

Maarten De Ridder[†] Basile Grassi[‡] Giovanni Morzenti[§]

July 29, 2022

Abstract

Is it feasible to estimate firm-level markups with commonly available datasets? Common methods to measure markups hinge on a production function estimation, but most datasets do not contain data on the quantity that firms produce. We use a tractable analytical framework, simulation from a quantitative model, and firm-level administrative production and pricing data to study the biases in markup estimates that may arise as a result. While the level of markup estimates from revenue data is biased, these estimates do correlate highly with true markups. They also display similar correlations with variables such as profitability and market share in our data. Finally, we show that imposing a Cobb-Douglas production function or simplifying the production function estimation may reduce the informativeness of markup estimates.

Keywords: Macroeconomics, Production Functions, Markups, Competition

*Marco Panunzi provides excellent research assistance. We would like to thank Jêrome Adda, Klaus Adam, Ariel Burstein, Vasco M. Carvalho, Thomas Chaney, Jan De Loecker, Francesco Decarolis, Ulrich Doraszelski, Jan Eeckhout, Xavier Gabaix, Kyle Herkenhoff, Chad Jones, Greg Kaplan, Marco Maffezzoli, Massimiliano Marcellino, Isabelle Mejean, Simon Mongey, Gianmarco Ottaviano, Fausto Panunzi, Jean-Marc Robin, Ricardo Reis, Ariell Reshef, John van Reenen and Frederic Warzynski for useful comments. We also thank seminar participants at Bocconi, Bank of Slovakia, Duke, Ecole Polytechnique, ENS Saclay, Erasmus, EUI, HEC Paris, LMU, Mannheim, Melbourne, Nottingham, Paris School of Economics, Paris Trade Seminar, LMU Munich, London Business School, London School of Economics, Tilburg, Toronto, Stockholm University, University Cergy-ESSEC, UC Irvine, University College London. We also thank the organizers and participants of the Aarhus Markups Workshop, BSE Summer Forum - FGE 2022, CAED 2021, CFM - Bristol Macro Workshop 2022, Kent Firms Workshop, Mannheim Firm in Macro Workshop 2022, SED 2021 and the T2M conference 2022 for useful comments and suggestions. We thank Isabelle Mejean for help with the French data. The access to the French data is provided by CASD which is supported by a public grant overseen by the French National Research Agency (ANR) as part of the 'Investissements d'Avenir' program (reference: ANR-10-EQPX-1, CASD).

[†] London School of Economics. m.c.de-ridder@lse.ac.uk

[‡] Bocconi University, CEPR and IGIER. basile.grassi@unibocconi.it

[§] Bocconi University. giovanni.morzenti@phd.unibocconi.it

1 Introduction

The markup of prices over marginal costs is a key variable in economics. In macroeconomics, firm-level markups impact the efficiency of resource allocation, the aggregate labor wedge or incentives to innovate, and they form a potential transmission channel of nominal shocks. In industrial organization, firm-level markups are an important metric for market power, and they are key to understand the welfare effects of regulations and (competition) policies. Recent evidence (e.g. De Loecker et al. 2020) of a rise in both the level and the dispersion of markups has furthermore caused markups to feature prominently in academic and policy debate.

While their economic importance is well-understood, it remains difficult to measure markups in the data. Prices are only observed in a small number of datasets, while marginal costs are rarely ever observed. To overcome this, much of recent empirical work relies on the markup estimator in Hall (1986, 1988). Hall derives an expression for markups from the first-order condition of cost-minimizing firms for a *variable input*. This is an input that firms set without adjustment costs or other intertemporal considerations, and for which the firm takes input prices as given. He shows that a firm's markup is the variable input's output elasticity, multiplied by the inverse of the input's revenue share. For firms with a markup of 1, the output elasticity exactly equals the revenue share. For firms with a greater markup, spending on the input as a fraction of revenue falls short of the output elasticity.

Markups therefore seem straightforward to estimate with income statement data on revenue and input spending, as long as the researcher knows the output elasticity of the variable input. Practitioners obtain that elasticity from an estimation of the production function, using techniques from the empirical industrial organization literature. These techniques typically assume, however, that firms are price takers – an assumption that is particularly unfortunate when estimating markups. A growing body of work therefore contests the assertion that production functions and markups are accurately estimated with these commonly used methods.

In this paper we therefore ask whether it is feasible to estimate markups and production functions when firms are price setters. We use a combination of theory, simulations and administrative micro data to address this question. Our purpose is twofold. First, we use a simple analytical framework to derive closed-form expressions for the biases that arise in production function and markup estimation when firms have market power. We particularly look at the bias that arises when researchers use revenue rather than output when estimating the production function. While output and revenue diverge when firms are price setters, most datasets only contain the latter. Second, we use simulations and an empirical exercise to quantify the resulting biases. Our simulations of a rich oligopolistic competition model allow us to scrutinize markup estimates in a setting where the true markup is known, while our firm-level price and production data allow us to explore the correlation between markup estimates from data on revenue and markup estimates from data on output.

Following De Loecker and Warzynski (2012), we estimate production function elasticities using the Ackerberg et al. (2015) two-stage GMM procedure. The procedure involves a first-stage regression to purge firm output of measurement error and transitory productivity shocks, followed by a second stage that identifies the production function by imposing structure on the productivity process to identify the true parameters. The latter is needed because productivity affects outputs and inputs, biasing a least squares estimation (e.g. Klette and Griliches 1996).

We focus on two prominent critiques of this commonly used methodology to estimate production functions and markups in settings when firms have market power. The first critique we study is that researchers often use revenue data as a measure of output to estimate production functions. Price-setting firms must reduce prices when raising output, such that a production function estimation with revenue as a proxy for output will understate actual output elasticities. A long literature, following Klette and Griliches (1996), derives that this may cause a bias in estimates of output elasticities. Bond et al. (2021) claim that this bias is severe enough to render the estimated markups uninformative of true markups. The scarcity of firm-level pricing data means that this critique has the potential to seriously limit future analysis of markups.

The second critique questions whether production functions can be estimated with the Ackerberg et al. (2015) two-stage GMM procedure when firms are price setters. Doraszelski and Jaumandreu (2020) show that in the first stage of the procedure, output can only be purged from measurement error and idiosyncratic productivity shocks if researchers observe markups. Given that the estimation of markups is the point of the exercise, this too suggests that firm-level markup estimation along the method in Hall (1986) is not feasible.

While both of these critiques have merit, we show that it is largely feasible to estimate markups and production functions. In an analytical framework, we first provide closed-form expressions for the biases caused by both the use of revenue to measure output as well as the omission of markups in the first-stage of the procedure. We then quantify these biases in two ways. We first estimate the production function for simulated firms from an oligopolistic competition model à la Atkeson and Burstein (2008) with endogenously heterogeneous markups. This enables us to compare estimates of output elasticities and markups with their true values. We then compare estimates of markups based on revenue and output from firm-level data on the universe of French manufacturing firms with at least 20 employees. The dataset contains balance sheet and income statement data for 2009 to 2019, as well as unit values of the products they sell. This enables us to empirically correlate markups from data on revenue and from data on quantities.

In our analytical framework, when output is observed, we show that imperfect competition can be a source of identification of the production function in the standard GMM procedure. When revenue is used as a measure of quantity, we show that the resulting biased elasticity can be used to recover information about the true markup. To be precise, revenue-based markup estimates correlate positively with true markups, although their average level contains little information about the average of true markups. This is because the bias in the estimated production function is an average of the demand elasticity across firms with the same production function. These demand elasticities are, in some static oligopoly models, equal to the inverse firm-level markup. It follows that, when firms have heterogeneous markups, the bias cannot cancel out variation in firm-level markups, though it may cancel out their *average*. This contrasts with the claim in Bond et al. (2021) that revenue-based markups are uninformative about true markups. We explain that this may hold on average, in the sense that the average revenue-based markup is usually not informative of the true average markup. In general, the average revenue-based markup can be higher or lower than the true average.

Our simulations show a strong correlation between true markups and estimated markups. In a perfect scenario where the researcher has data on the firm's output quantity and is able to correctly purge for measurement errors, markups and production function elasticities are estimated with precision. The correlation between estimated and true markups is 1, and both the level and standard deviation of markups is correctly identified. These results are robust to not controlling for markups in the first stage of the production function estimation, as we still find a correlation between estimated and true markups of 1. Not accounting for measurement error is more costly, with the correlation between estimated and true markups falling below 0.5.

In the more common scenario in which researchers lack data on prices and quantities, we still find high correlations between estimated and true markups. In our preferred specification we find a correlation of 0.80. When we calibrate our simulation's model with standard parameter values, revenue-based estimates of the markup slightly exceed their true value. We further show that markup *dispersion* – both in the cross-section and over time – is well-estimated.

Our empirical results validate these findings. While we do not know the true markups in that case, we do find moderately high correlations between markup estimates based on revenue and quantity data. We find a 0.3 correlation between revenue and quantity-based markups in our preferred specification, rising to 0.7 in first differences. We furthermore show that the regression coefficients relating estimated markups to profits, labor share and the market share are of the same sign and order of magnitude, irrespective of the specification considered. Trends in aggregate markups are also well-estimated with revenue data. Overall, we conclude that firm-level estimates of the markup along Hall (1986)'s methodology are informative of true markups.

Our results do imply that researchers should give careful consideration to the suitability of their data for the question at hand. When interested in the level of the markup, researchers need quantity data. When interested in dispersion, such as variation across firms or trends over time, revenue data may suffice. In additional results, we do note that assumptions on the functional form of the production function should not be restrictive. In particular, we show that it can be costly to assume a Cobb-Douglas production function when the true production function is more complex. This assumption has recently become common, as Cobb-Douglas markup estimates are robust to biases in production function estimates up to a sector fixed effect.[1]

We show that Cobb-Douglas estimates typically capture the average of true markups, but mismeasure their dispersion. Compared to the more flexible translog, which approximates any production function and nests Cobb-Douglas as a special case, we find less dispersion in markups in our simulations and French data. In a back-of-the-envelope exercise, we show that the welfare costs of markup dispersion are overstated by 140% when assuming Cobb-Douglas.

Related literature Our paper builds on a significant literature that estimates production functions. A simple regression of a set of (log) inputs on a firm's (log) output does not identify the production function elasticities because of unobserved differences in productivity across firms. Productivity directly affects output and indirectly affects inputs, such that a least squares estimation of a parametric production function is biased (e.g. Klette and Griliches 1996). Our analysis, in the spirit of the seminal work by Olley and Pakes (1996) and Levinsohn and Petrin (2003), uses a proxy regression to control unobservable productivity. In line with Ackerberg et al. (2015), we use materials as a variable input to proxy for unobservable productivity. We also closely relate to the literature – with Klette and Griliches (1996) as a seminal contribution – on the biases that arise when using revenue data to measure output in a production function estimation, although our focus is on markup estimation.

We focus on the feasibility of using estimates of production function elasticity to estimate markups at the firm level. This technique was pioneered by De Loecker and Warzynski (2012) to show that exporting firms have higher markups than do non-exporters. De Loecker et al. (2016) extend their methodology to multi-product firms.[2] De Loecker et al. (2020) apply the methodology to listed U.S. firms to show that estimated markups have increased sharply between 1980 and 2015, a result that has been confirmed for other countries by Díez et al. (2019) and that has sparked a rich discussion on the feasibility of the De Loecker and Warzynski (2012) methodology on accounting data (e.g., Traina 2018, Basu 2019, Syverson 2019). Baqaee and Farhi (2019) note that the rise of markups is driven by a reallocation of activity towards high-markup firms, in line with evidence of reallocation towards low-labor-share firms in Autor et al. (2020) and Kehrig and Vincent (2021). Hershbein et al. (2021) and Morlacco (2019) note that markup estimates are biased when firms have market power on the market for the flexible input.[3]

[1] If firms have the same production functions within sectors, an analysis of Cobb-Douglas log markups with industry effects is unaffected by bias in production function estimates (e.g. Peters 2020, Crouzet and Eberly 2019, Meier and Reinelt 2020). Cobb-Douglas is therefore frequently assumed in response to the Bond et al. (2021) critique.

[2] Because most datasets do not provide input allocations across products that firms produce, De Loecker et al. (2016) estimate production functions at the product level using data on single-product firms. They then estimate markups for multi-product firms with the estimated elasticities. We also rely on product-level data for prices and quantities, but aggregate this to the firm level when estimating markups.

[3] Several recent papers deploy markup estimates using the De Loecker and Warzynski (2012) methodology in specific applications. Burstein et al. (2020) show that markups in French data are either procyclical or countercyclical depending on the level of aggregation that is considered. Meier and Reinelt (2020) add that markups become more dispersed after monetary policy shocks, negatively affecting total factor productivity. Calligaris et al. (2018) find that

Estimating the production function when firms have market power is particularly challenging. As pointed out by Doraszelski and Jaumandreu (2019) and Brand (2019), one of the identifying assumptions in the procedure by Ackerberg et al. (2015) is that the demand function by firms is not affected by unobservables other than productivity.[4] This is true for cases such as perfect or monopolistic competition but not necessarily true for the case of oligopolistic competition. We analyze the bias arising from improperly accounting for demand (and therefore markup) heterogeneity in the estimation of the production function elasticities and markups.[5]

Our paper adds to the broader literature on the consequences of market power. Karabarbounis and Neiman (2014) find that the share of labor has fallen in most advanced economies over the last decades.[6] Barkai (2020) adds that, when accounting for the falling costs of capital, the capital share of income has also declined, leaving a rise in the profit share as the residual.[7] Increasing markups have been linked to low investments and a lack of entry (e.g. Gutiérrez and Philippon 2017, Eggertsson et al. 2021, Gutiérrez and Philippon 2022), low productivity growth (e.g. Aghion et al. 2019, De Ridder 2019), and industry concentration (e.g. Grullon et al. 2019, Autor et al. 2020). The effects of markup dispersion on welfare through misallocation have been quantified in (e.g.) Baqaee and Farhi (2019), Edmond et al. (2018), and Peters (2020).[8]

Outline The remainder of this paper proceeds as follows. Section 2 explains how we measure markups from the production function. Section 3 outlines our analytical framework, while Section 4 introduces the data. Section 5 presents simulations; the empirical exercise is presented in Section 6. Section 7 discusses Cobb-Douglas production functions, and Section 8 concludes.

2 From markups to production functions

Before discussing the conditions under which markups and production functions can be estimated, we explain how the two are related through the problem of a cost-minimizing firm. Following Hall (1986, 1988), we derive an expression for markups under the assumption that there exists an input that firms set statically, without intertemporal considerations; and that

markups have increased most in sectors with high digitization, a result confirmed at the firm level by De Ridder (2019). Crouzet and Eberly (2019) find a relationship between markups and a firm's intangible investment share. Pasqualini (2021) applies the De Loecker et al. (2016) methodology to estimate markups in the banking sector.

[4]The bias arising from a violation of this assumption (in particular on correlations between markups and demand determinants) is analyzed in Doraszelski and Jaumandreu (2020).

[5]Our estimation also requires sufficient variation in input prices for the variable input to allow separate identification of the variable input and productivity (e.g. Blundell and Bond 2000, Gandhi et al. 2020).

[6]Gutiérrez and Piton (2020) note that, outside of the U.S., the decline in the labor share is driven by housing.

[7]Neiman and Vavra (2021) note that unmeasured inputs would also appear as a rise in profits from this calculation. Van Vlokhoven (2019) estimates the cost of capital in a regression framework.

[8]Cavenaile et al. (2019) note that the rise of markups also incentives firms to invest in R&D. Bornstein (2018) shows that consumer demand has become less sensitive to price changes, a trend that might be driven by aging, while Neiman and Vavra (2021) consumption baskets are increasingly narrow. Anderson et al. (2018) analyze how markups vary over space and time and find significant differences in markups across regions.

firms are price-takers for that input. Further assume that output Y_{it} for firm i at time t obeys

$$Y_{it} = Y(V_{it}, \boldsymbol{K}_{it}, \Omega_{it}),$$

where V_{it} the static input, purchased at price W_t. Vector $\boldsymbol{K}_{it}$ contains all other inputs, while Ω_{it} is productivity. The cost-minimizing firm's first-order condition for V_{it} solves

$$\frac{1}{\lambda_{it}} = \frac{\partial Y(\cdot)}{\partial V_{it}} \frac{1}{W_t},$$

where λ_{it} is a Lagrange multiplier which measures marginal costs. Hall (1986, 1988)'s markup expression follows by multiplying both sides by the firm's price P_{it}. The right-hand side can be rewritten in terms of the output elasticity of V_{it}, multiplied by its inverse revenue share:

$$\mu_{it} = \left(\frac{\partial Y(\cdot)}{\partial V_{it}} \frac{V_{it}}{Y_{it}} \right) \frac{P_{it} Y_{it}}{W_t V_{it}}, \tag{1}$$

where $\mu_{it} \equiv P_{it}/\lambda_{it}$ is the markup. The expression yields the familiar result that an input's output elasticity equals its revenue share if markups are 1, while revenue shares fall short of the output elasticity when markups exceed 1. It follows that to estimate markups, the researcher needs data on revenue and input spending from the income statement, as well as an estimate of the output elasticity of V_{it}, $\alpha^v_{it} \equiv (\partial Y(\cdot)/\partial V_{it}) \cdot V_{it}/Y_{it}$. Estimating this elasticity under imperfect competition is therefore a primary empirical challenge in the estimation of markups.

3 Estimating production functions under imperfect competition

We next outline how the output elasticity of V_{it} is estimated. We describe the production function estimation under imperfect competition, which is the natural setting when estimating markups. We start from the ideal case where a researcher observes prices, such that output can be measured by quantity. We then discuss the bias that arises when prices are unobserved.

3.1 The model

We start by deriving an estimator for the output elasticity in a model where output is log-linear in a single input, while productivity is identically and independently distributed. This simplest possible environment enables us to derive clear closed-form solutions. In the appendix we show that our conclusions are robust to more realistic and less stringent assumptions.[9]

Consider an environment where firms produce their output Y_{it} using one input V_{it}. Firms are subject to total factor productivity shocks, denoted ω_{it} in logs, that are unobserved by the

[9] Appendix A.3 extends the results by allowing for a translog production function (A.3.1), multiple inputs (A.3.2), persistence in productivity (A.3.3), and all of these together (A.3.4).

econometrician but observed by the firm. Output is further subject to shocks η_{it}, which are not observed by either the econometrician or the firm. Shocks η_{it} capture measurement error and white-noise productivity shocks that are only observed after production decisions are made. Firms set V_{it} to minimize costs and share the Cobb-Douglas production function

$$y_{it} = \alpha v_{it} + \omega_{it} + \eta_{it}, \tag{2}$$

where lower caps denote log-deviations from the mean, and where the parameter α is the true output elasticity of v_{it} to be estimated.[10] Note that we have not made assumptions about the product market in which firms operate, and that we therefore allow for imperfect competition.

We assume that v_{it} is static and variable: spending on the input is entirely determined within the period and is not subject to any adjustment cost or other intertemporal constraints. We further assume, for now, that ω_{it} is independently and identically distributed (i.i.d.) across time and firms. Since v_{it} is a flexible input, we can use this input to compute markups, as explained in Section 2. Denoting the firm's price by P_{it}, true markups are given by $\mu_{it} = \alpha(P_{it}Y_{it})/(W_t V_{it})$.

3.2 Identification with price and quantity data

We first cover the estimation of α if revenue, prices – and therefore quantities – are observable. Our estimator for α builds on the two-stage GMM estimator of Ackerberg et al. (2015) to accommodate imperfect competition. The first stage purges the quantity of equation (2) of the measurement error and unobserved productivity shocks η_{it}. The second stage estimates the output elasticity α using an instrumental-variable generalized method of moments (IV-GMM). We first focus in Section 3.2.1 on the second stage – as it performs the actual production function identification. We then introduce measurement errors and the first stage in Section 3.2.2.

3.2.1 Identification

In the absence of measurement error, the production function simplifies to $y_{it} = \alpha v_{it} + \omega_{it}$. A least-square regression of input v_{it} on output y_{it} will be biased, as the unobserved productivity ω_{it} (the residual in the regression) affects firms' choice of v_{it}. Following the literature, we can construct an estimator to identify α by instrumenting v_{it} by v_{it-1}:

Definition 1 *The instrumental variable GMM (IV-GMM) estimator* $\widehat{\alpha} \in \mathbb{R}$ *is such that the moment* $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right]$ *is equal to zero where* $\widehat{\omega}_{it} = y_{it} - \widehat{\alpha} v_{it} = (\alpha - \widehat{\alpha}) v_{it} + \omega_{it}$.[11]

[10] To be precise, $x_{it} = \log X_{it} - \mathbb{E}\left[\log X_{it}\right]$ where $\mathbb{E}\left[\log X_{it}\right]$ is the limit of the empirical average across observations. This normalization allows us to get rid of any constant in the production function and ensures ω_{it} has mean zero.

[11] In the above definition, the expectation operator $\mathbb{E}$ denotes the limit of the empirical average across observations. We therefore study the asymptotic properties of the GMM estimator, which allows us to keep the argument as tractable as possible. Appendix A.1 derives the estimator for finite samples before deducing its asymptotic variance.

It is straightforward to solve for $\widehat{\alpha}$ in closed form by substituting $\widehat{\omega}_{it}$ into the moment condition:

$$(\alpha - \widehat{\alpha})\mathbb{E}\left[v_{it}v_{it-1}\right] = 0, \tag{3}$$

which uses the fact that that productivity ω_{it} is orthogonal to v_{it-1}, such that $\mathbb{E}\left[\omega_{it}v_{it-1}\right] = 0$. It follows that as long as v_{it-1} is a relevant instrument for v_{it}, that is $\mathbb{E}\left[v_{it}v_{it-1}\right]$ differs from zero, the only solution is that $\widehat{\alpha} = \alpha$. Our estimator $\widehat{\alpha}$ converges to the true elasticity α.

What ensures that the lagged variable input is a relevant instrument? As we have assumed – for now – that productivity is not persistent, autocorrelation in v_{it} comes from other sources.[12] The cost-minimizing firm's first-order condition for v_{it} summarizes the candidate drivers:

$$v_{it} = (1-\alpha)^{-1}\left(\omega_{it} + mc_{it} - w_t\right). \tag{4}$$

It follows that persistence in v_{it} has to either come from persistence in the input price w_t or from log marginal costs mc_{it}. Marginal costs equal P_{it}/μ_{it}, both of which are determined in equilibrium by the demand system and the strategic interactions among firms. Hence any persistence in output price or markups will contribute to persistence in the variable input and thus to identification of the production function. Persistence in input prices w_t is a source of persistence in variable inputs regardless of the mode of competition, providing a further source of identification of α (a point previously made by, e.g., Gandhi et al. 2020). We conclude that the parameters of the production function in our simple framework are identified under imperfect competition as long as there is persistent variation in markups, output prices or input prices.[13]

Appendix A.3 generalizes these basic identification results by allowing for translog production functions, multiple inputs, persistence in productivity, and all of these combined.[14] In Appendix A.1 we further derive the finite sample properties of the estimator and its asymptotic variance.

[12] When we generalize our setup in Appendix A.3.3 (where productivity is assumed to be persistent, e.g. a linear AR(1) process with persistence ρ), we show that the necessary condition for identification is to have autocorrelation in $\widetilde{v}_{it} = v_{it} - \rho v_{it-1}$. Persistence in productivity itself therefore does not aid identification.

[13] Note that this means that it is more straightforward to estimate the production function under imperfect competition than under perfect competition. Under perfect competition (where marginal costs equal prices), persistence in the variable input cannot come from the markup. If output prices are (e.g.) i.i.d., this means that the only source of persistence is the input price. Gandhi et al. (2020) provide a detailed investigation of this argument.

[14] In Appendix A.3.3 we do note that under persistent productivity, the identification of the production function parameters may only hold locally. In particular, we find that there are exactly two solutions to the GMM estimator under persistent productivity. One solution gives the true value of the parameters, while the second solution is a biased estimate of the true parameters. However, if $\mathbb{V}ar[v_{it-1}]$ is large compared to $\mathbb{V}ar[\omega_{it-1}]$ and $\mathbb{V}ar[\widetilde{v}_{it}]$, where $\widetilde{v}_{it} = v_{it} - \rho v_{it-1}$, then there exists a unique solution for $\widehat{\alpha}$ and $\widehat{\rho}$. This means that, if there is enough variation in the data, the parameters of the production function are globally identified. This is in line with a recent paper by Ackerberg et al. (2020), which shows that the two-stage estimator might have two solutions, rendering traditional numerical solvers unstable.

3.2.2 Adding the first stage

Thus far we have assumed that output is observed without measurement error and unobserved productivity shocks η_{it}. We next describe the estimation of α if this assumption is eased. Prior work pays specific attention to η_{it}, for two reasons. First, output measures in common datasets contain significant measurement error. In our empirical analysis, for example, we measure output by subtracting unit values from revenue, which are in turn obtained from surveys. As we discuss below, this diminishes the precision of the production function estimation. Second, the presence of η_{it} impedes the estimation of true productivity ω_{it} regardless of the ω_{it} process and can even impede the production function estimation if ω_{it} follows a non-linear dynamic process. We discuss the importance of both of these issues under imperfect competition, and demonstrate how a first-stage purging regression addresses these concerns.

Measurement error without a first-stage In the presence of measurement errors η_{it}, one can still use the IV-GMM estimator in Definition 1 to estimate the production function. This is the approach proposed by Blundell and Bond (2000). As long as the measurement error is uncorrelated with the firm's inputs, it will only increase the standard error with which α is estimated. In Appendix A.2 we show that the variance of the estimator is proportional to $\mathbb{E}\left[\omega_{it}^2\right] + \mathbb{E}\left[\eta_{it}^2\right]$ under this approach, while the estimator remains unbiased. The degree to which measurement error is problematic therefore depends on the variance of η_{it} and the size of the sample at hand. The residual from the production function estimation is the composite of η_{it} and ω_{it}, which means that researchers do lose the ability to observe true productivity in this case.

Measurement error can also impede consistency of the IV-GMM estimator if ω_{it} is persistent with non-linear autoregressive terms (Bond et al. 2021). With persistent productivity, the moment conditions of the IV-GMM estimator can be altered to consistently estimate α (Appendix A.3.3). For a linear AR(1) process of ω_{it}, the moment conditions are that lagged inputs v_{it-1} and estimated productivity $\widehat{\omega}_{it}$ are orthogonal to the residuals of the AR(1) process. For non-linear processes (e.g. quadratic, cubic), in the absence of measurement error, the additional moment conditions are that the higher-degree terms (e.g. $\widehat{\omega}_{it}^2$, $\widehat{\omega}_{it}^3$) are orthogonal to the residual.

Measurement error, however, contaminates the productivity estimates $\widehat{\omega}_{it}$. This means that moment conditions with, e.g., $\widehat{\omega}_{it}^2$, $\widehat{\omega}_{it}^3$ contain higher-order moments of the measurement error. This prevents the moment conditions from holding at the true value of the output elasticity.[15]

[15] For example, a common empirical assumption is that the productivity process is well-approximated by $\omega_{it} = \rho_1\omega_{it-1} + \rho_2\omega_{it-1}^2 + \xi_{it}$, where ξ_{it} are white-noise productivity shocks. In the presence of measurement error, the moment conditions $\mathbb{E}[v_{it-1}\widehat{\xi}_{it}] = 0, \mathbb{E}[\widehat{\omega}_{it-1}\widehat{\xi}_{it}] = 0$, and $\mathbb{E}[\widehat{\omega}_{it-1}^2\widehat{\xi}_{it}] = 0$, where $\widehat{\omega}_{it}$ is defined as before while $\widehat{\xi}_{it} \equiv \widehat{\omega}_{it} - \widehat{\rho}_1\widehat{\omega}_{it-1} - \widehat{\rho}_2\widehat{\omega}_{it-1}^2$, will not suffice to estimate the production function. The source of the problem is the non-linear moment condition $\mathbb{E}[\widehat{\omega}_{it-1}^2\widehat{\xi}_{it}] = 0$. To see this, consider the value of the moment at $\widehat{\alpha} = \alpha$:

$$\mathbb{E}[\widehat{\omega}^2\widehat{\xi}_{it}] = \mathbb{E}[(\omega_{it} + \eta_{it} + (\alpha - \widehat{\alpha})v_{it})^2\widehat{\xi}_{it}] = \mathbb{E}[\eta_{it}^2\widehat{\xi}_{it}] \neq 0.$$

It follows that the IV-GMM estimator does not estimate the production function parameters unless productivity follows a linear (dynamic) process.

Purging measurement error with a first stage The combination of higher standard errors, stringent assumptions on the dynamic process of ω_{it}, and the loss of direct estimates for true productivity ω_{it} form a case to purge observed output from measurement error. Ackerberg et al. (2015) do so in a first-stage regression for the case of perfect competition, we propose here a procedure that – deviating minimally from theirs– can do so under imperfect competition.

The aim of the purging regression is to separate η_{it} and ω_{it}, using the fact that firms only observe ω_{it} when deciding the quantity of inputs that they wish to deploy. The idea is that the demand for the variable input can therefore be expressed as a function of productivity: $v_{it} = v(\omega_{it}, \Xi_{it})$, where Ξ_{it} is a vector of all variables that determine v_{it} other than productivity. Under the assumption that v_{it} rises monotonically in ω_{it}, the demand function can be inverted, such that $\omega_{it} = v^{-1}(v_{it}, \Xi_{it})$. In our framework, the production function can therefore be written as

$$y_{it} = \alpha v_{it} + v^{-1}(v_{it}, \Xi_{it}) + \eta_{it}.$$

The fitted values of a non-parametric regression of y_{it} on v_{it} and Ξ_{it} therefore identify η_{it}. As long as the researcher correctly specifies the variables that determine the demand for v_{it}.

What variables are included in Ξ_{it} under imperfect competition? First-order condition (4) shows that these variables are factor prices and log marginal costs. Using the fact that marginal costs can be expressed in terms of prices and markups, observed output can be written as[16]

$$y_{it} = v_{it} - p_{it} + \log \mu_{it} + w_t + \eta_{it}. \tag{5}$$

This means that the researcher must run a regression of output on the variable input, prices, markups and time fixed-effects for w_t in order to purge measurement error. Under perfect competition, firms take prices as given and have log markups of 0.[17] Hence, a first-stage regression of output on v_{it} and a time fixed-effect is sufficient to purge output of measurement error.[18]

Under imperfect competition this is not sufficient, because firms have heterogeneous markups. As noted by Doraszelski and Jaumandreu (2020), the whole purpose of the exercise is to estimate these. We propose resolving this by including controls for the markup in the first stage

[16]Note that the expression of the marginal cost $MC_{it} = P_{it}/\mu_{it}$ in log deviation from its mean mc_{it} is equal to $p_{it} - \log \mu_{it}$ up to a constant $\mathbb{E}[\log \mu_{it}]$, which we include in the first stage.

[17]In the more general multi-input, non Cobb-Douglas case, the first-order condition of the cost-minimization problem is not linear in inputs and cannot be inverted analytically. Nevertheless, the functional relationship between productivity and inputs, price and markups is well defined and can be approximated by a polynomial of inputs.

[18]Under perfect competition, firms are price takers, which implies that markups equal to one and prices are orthogonal to firms' choices. The production function, after substituting the expression for productivity, reduces to $y_{it} = v_{it} + w_t + \eta_{it} - p_{it}$. The last two terms are orthogonal to input usage, v_{it}, and input price w_t.

of the procedure. Note that when doing so, we only need to know that there is a structural relationship between markup and controls; we do not need to know the parameters that govern this relationship. In our simulations and empirical sections, we assume that markups are determined by a firm's market share, which holds in a large set of models. We therefore include prices and market shares as controls in the first stage of our baseline two-stage estimator.

In summary, we propose that researchers use a two-step procedure in order to estimate the production function given by equation (2) under imperfect competition. First, quantity is purged from measurement error in a regression of observed quantity, y_{it}, on the variable input v_{it}, the output price p_{it}, controls for the markup μ_{it} such as market share, and time fixed-effects for w_t. The fitted values of output, true quantity, are then used to construct moment $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right]$, a function of $\widehat{\alpha}$. A numerical solver can then find the $\widehat{\alpha}$ that makes this moment equal to zero. As discussed above, this value is an asymptotically consistent estimator of the true parameter α.

3.3 Identification with revenue data

We have thus far assumed that researchers observe the quantity that firms produce, with or without measurement error. Most firm-level datasets do not contain such data and instead only contain data on revenue. This section studies the bias in estimates of the output elasticity that result from solely relying on revenue data causes, as well as the consequences of the bias for the resulting markup estimates.

In our framework, we can derive the bias arising from the use of revenue as a measure of output. Revenue is the product of quantity and price; $r_{it} = y_{it} + p_{it}$ gives revenue in log-deviations from the mean. Inserting the production function (2) for y_{it} yields

$$r_{it} = y_{it} + p_{it} = \alpha v_{it} + \omega_{it} + p_{it},$$

where α remains the parameter of interest and where we abstract from measurement error to focus on the bias caused by using revenue as a measure for output. The following definition captures the IV-GMM estimator in Definition 1 when revenue is used in place of quantity.

Definition 2 *(Revenue IV-GMM estimator) The estimator is* $\widehat{\alpha} \in \mathbb{R}$ *such that the moment* $\mathbb{E}[\widehat{\mathrm{tfpr}}_{it} v_{it-1}]$ *is equal to zero where* $\widehat{\mathrm{tfpr}}_{it} = p_{it} + y_{it} - \widehat{\alpha} v_{it} = (\alpha - \widehat{\alpha}) v_{it} + p_{it} + \omega_{it}$.

Let us show that the revenue IV-GMM estimator is biased. Solving for the $\widehat{\alpha}$ such that $0 = \mathbb{E}[\widehat{\mathrm{tfpr}}_{it} v_{it-1}] = (\alpha - \widehat{\alpha})\mathbb{E}\left[v_{it} v_{it-1}\right] + \mathbb{E}\left[p_{it} v_{it-1}\right]$, yields the following unique solution as long as the lagged variable input is a relevant instrument:

$$\widehat{\alpha} = \alpha + \frac{\mathbb{E}\left[p_{it} v_{it-1}\right]}{\mathbb{E}\left[v_{it} v_{it-1}\right]}, \tag{6}$$

It follows that the IV-GMM estimator delivers a biased estimate of the true α if prices and lagged variable inputs are correlated. Using revenue rather than quantity to measure output creates an omitted variable bias, because the revenue-production function has prices in the residual.

Under imperfect competition, it is probable that p_{it} will correlate with lagged variable inputs, such that $\mathbb{E}[p_{it}v_{it-1}]$ differs from zero.[19] Note that there are no model-free constraints on either the size or sign of the covariance. If firms face persistent aggregate demand shocks and decreasing returns to scale, for example, positive shocks drive up marginal costs and prices, causing a *positive* correlation between prices and lagged variable inputs. Conversely, firms with downward-sloping demand curves reduce prices to sell additional output, causing a *negative* correlation. The estimates of α can therefore be smaller, larger or equal to the true output elasticity. Equally, the ensuing markup estimates may overstate, understate or equal true markups.

Revenue-based markups and oligopolies We next show that despite the bias from a non-zero covariance $\mathbb{E}[p_{it}v_{it-1}]$, revenue-based markup estimates are still informative about true markups. As we are interested in how the bias affects markups, we focus on the correlation between prices and lagged inputs that is caused by the downward-sloping demand functions. To do so, we add a demand side to our baseline framework. To keep assumptions minimal, we assume a very general invertible demand system, where a firm's demand depends on prices of all firms. We abstract from aggregate shocks that alter price-quantity relationships across periods. Hence we assume that the vector of quantities produced by all firms, $Y = \{Y_{it}\}$, is a function of the price vector $P = \{P_{it}\}$ such that $Y = D_t(P)$. A log-linear approximation yields

$$p_{it} = -\sum_j d_{ijt} y_{jt}, \tag{7}$$

where d_{ijt} is the cross-elasticity of firm i's price with respect to firm j's quantity.[20] With this demand system, we can write (6) as

$$\widehat{\alpha} = \alpha \left(1 - \sum_j \frac{\mathbb{E}\left[d_{ijt}(v_{jt} + \frac{\omega_{jt}}{\alpha})v_{it-1}\right]}{\mathbb{E}[v_{it}v_{it-1}]} \right).$$

It follows that the bias due to the use of revenue data is equal to one minus the weighted average of demand elasticities and cross-elasticities among the firms sharing the same production function. This biased estimate of the production function can be used to estimate a firm-level markup based on revenue data, $\widehat{\mu}_{it}^R = \widehat{\alpha}\frac{P_{it}Y_{it}}{W_t V_{it}}$. It follows that revenue-based markup equal:

$$\widehat{\mu}_{it}^R = \mu_{it} \left(1 - \sum_j \frac{\mathbb{E}\left[d_{ijt}(v_{jt} + \frac{\omega_{jt}}{\alpha})v_{it-1}\right]}{\mathbb{E}[v_{it}v_{it-1}]} \right),$$

[19] The correlation is zero in case there are no aggregate shocks and firms are atomistic price takers – but that assumption is counterintuitive in contexts where markups are a variable of interest.

[20] In appendix A.4, we derive this approximation formally. We implicitly assume a static demand system, where only current period quantities affect current prices.

where we use the Hall (1986, 1988) formula to substitute for the true markup, $\mu_{it} = \alpha \frac{P_{it}Y_{it}}{W_t V_{it}}$. The revenue-based markups are equal to the true markups up to a constant. They have equal variances, and the correlation between the (log) revenue markup and the (log) true markup is equal to one.[21]

A special case: static oligopoly While the correlation between true markups and revenue-based estimates of the markup is one, more assumptions are needed to understand the effect of using revenue data on estimates of the various moments of the markup, such as dispersion and average levels. In this section, we explore this for the simple special case in which firms play a static oligopoly game in which $d_{ijt} = 0$ while $d_{iit} \neq 0$ and $d_{iit} \neq d_{jjt}$. Thus, the firm's demand is only determined by its own supply, and firms face heterogeneous price-elasticities of demand. Formally, the demand system for this special case, which is the equivalent of equation (7), is such that $p_{it} = -d_{it} y_{it}$ where, with some abuse of notation, $d_{it} = d_{iit}$. When firms maximize profit under these assumptions, they will charge a markup $\mu_{it} = \frac{1}{1-d_{it}}$ over their marginal cost while, as before, cost minimization yields $\mu_{it} = \alpha \frac{P_{it}Y_{it}}{W_t V_{it}}$.

Under these assumptions, the IV-GMM estimator on revenue equals the output elasticity times the *average* inverse markup among firms in the sample:[22]

$$\widehat{\alpha} = \mathbb{E}\left[\alpha(1-d_{it})\frac{v_{it}v_{it-1}}{\mathbb{E}\left[v_{it}v_{it-1}\right]}\right] = \mathbb{E}\left[\alpha\mu_{it}^{-1}\frac{v_{it}v_{it-1}}{\mathbb{E}\left[v_{it}v_{it-1}\right]}\right],$$

where the last equality uses the profit-maximizing condition. Note that the estimator is also equal to the average of revenue elasticities among the firms sharing the same production function. Indeed, the revenue elasticity of firm i at t, in this simple case, is given by $\frac{\partial r_{it}}{\partial v_{it}} = \frac{\partial y_{it}}{\partial v_{it}} + \frac{\partial p_{it}}{\partial y_{it}}\frac{\partial y_{it}}{\partial v_{it}} = \alpha\mu_{it}^{-1}$. It follows that whenever firms have heterogeneous revenue elasticities, the IV-GMM estimator on revenue is not equal to the firm-level revenue elasticity.

Turning to the resultant markup estimates using the Hall (1986, 1988) formula, we have

$$\widehat{\mu}_{it}^{R} \equiv \widehat{\alpha}\frac{P_{it}Y_{it}}{P_t^V V_{it}} = \mathbb{E}\left[\mu_{it}^{-1}\frac{v_{it}v_{it-1}}{\mathbb{E}[v_{it}v_{it-1}]}\right]\mu_{it}. \tag{8}$$

As in the previous case, the revenue-based markup estimates equal the true markups up to a constant. For our simple demand system, this constant is equal to the weighted average of inverse markup among firms sharing the same production function. It follows that, as soon as there exist two firms i and j such that their markups are different $\mu_{it} \neq \mu_{jt}$, the estimated revenue markup $\widehat{\mu}_{it}^{R}$ is different from one for either firm.

[21] The result that revenue and quantity markups perfectly correlate depends on the Cobb-Douglas assumption that output-elasticity is a constant. In Appendix A.5 we discuss what happens if the output elasticity is not constant (in the more general case of a translog production function) and show that the main insights remain.

[22] Here we assumed that $\mathbb{E}\left[d_{it}\omega_{it}v_{it-1}\right] = 0$. This assumption is satisfied (for example) when, conditional on v_{it-1}, productivity ω_{it} and demand elasticity d_{it} are orthogonal. This assumption is also satisfied when conditional on d_{it}, ω_{it} and v_{it-1} are orthogonal. This assumption is in place merely to clarify the argument.

Note that there is one case where revenue-based markup estimates do not contain any useful information about true markups. This is when firms compete monopolistically and have identical price-elasticities of demand such that $p_{it} = -\gamma y_{it}$. This assumption is satisfied by constant elasticity of substitution (CES) preferences with atomistic firms if the aggregate price index is fixed. Under these assumptions, the revenue estimator equals the revenue elasticity with respect to the variable input $\widehat{\alpha} = \alpha(1-\gamma) = \frac{\partial y_{it}}{\partial v_{it}}(1 + \frac{\partial p_{it}}{\partial y_{it}}) = \frac{\partial r_{it}}{\partial v_{it}}$. Both the revenue elasticity and the true markup are equal across firms, where the latter is equal to $(1-\gamma)^{-1}$. It follows that the revenue markup is equal to one $\widehat{\mu}_{it}^R = (1-\gamma)^{-1}(1-\gamma) = 1$, as in Bond et al. (2021). When markups are identical across firms sharing the same production function, revenue markups do not contain any information on the true markup.[23]

If firms do not have homogeneous markups, however, estimates of the markup from revenue data will generally not equal one. This is because in models with heterogeneous markups (e.g. Atkeson and Burstein (2008), Kimball (1995) or Klenow and Willis (2016)), demand elasticities differ across firms, while we estimate a single output elasticity of the variable input α. The key intuition comes from the fact that when the production function is estimated for a set of firms, the estimated elasticity equals an *average* revenue elasticity, which is not the same across firms. Variation in revenue-based markup estimates therefore does reflect variation in true markups.

Average levels of revenue-based markups When it comes to the average level of the markup, the bias arising from revenue data is also problematic. Given (8), we can write the average of the revenue-based markup estimates as

$$\mathbb{E}\left[\widehat{\mu}_{it}^R\right] = \mathbb{E}\left[\mu_{it}^{-1}\frac{\upsilon_{it}\upsilon_{it-1}}{\mathbb{E}[\upsilon_{it}\upsilon_{it-1}]}\right]\mathbb{E}\left[\mu_{it}\right],$$

which is equal to one up to a Jensen's inequality.[24] It follows that markup estimates based on revenue data carry little information about the average true markups in this demand system.

It is clear that the average markup is not identified with revenue data. Meanwhile, analyses of variation, such as trends over time or cross-sectional dispersion, can be performed well.

4 Data

We use administrative data on French manufacturing firms both to quantify our simulations and to empirically analyse the properties of markup estimates. We combine two main datasets. The FARE dataset (*Fichier Approaché des Résultats d'Esane*) provides a detailed balance sheet

[23] Another interesting/pathological case is when there is only one observation in the sample. Heuristically, assuming that the above formula remains valid, everything is as if the expectation operators dropped, leading to a revenue markup of one.

[24] Here we are making a similar assumption to the one in footnote 22.

and income statement, while the EAP survey (*Enquête Annuelle de Production*) provides data on both revenues and the quantities of products that firms ship, which we use to obtain a proxy for prices. FARE covers the universe of non-financial French firms and originates from filings to the tax administration (DGFiP). EAP is based on a product-level statistical survey by the statistical office (INSEE) which exhaustively covers manufacturing firms with at least 20 employees or revenue in excess of 5 million euros, and a representative sample of smaller firms.[25]

With the exception of prices, we obtain all variables for the production function estimation from FARE. Revenue is a firm's total sales (including exports),[26] the wage bill (measured as the sum of wages and social security payments), capital (measured by fixed tangible assets on the balance sheet),[27] expenditure on purchased services and expenditure on purchased materials. Materials are defined as physical intermediate goods and raw materials that firms purchase from others. We use NACE Rev. 2 industry codes and define industries j (at which firms have the same production technologies) at the two-digit level. Market share is defined as the ratio of the firm's revenue over total revenue of all firms in FARE in the 5-digit industry in a given year.

We obtain data on prices from EAP. EAP is a product-level dataset detailing a firm's revenue and quantity produced across 10-digit industries.[28] We define a product as the combination of a 10-digit product code and a unit of account.[29] We drop around one-third of firm-products without quantity data. For each combination of a firm and a product we calculate a price as the ratio of revenue over the quantity of the product sold. We then standardize this price by dividing it by the revenue-weighted average price of the product across the entire sample.[30] The firm's price in a year is then given by the sales-weighted average of standardized prices across the products that it produces. We define quantity as the ratio of revenue over this price.

To deflate input variables we use two-digit industry deflators from EU-KLEMS.[31] This is consistent with the assumption that firms operate on competitive input markets with equal prices across the two-digit level. Revenue is deflated with the gross output deflator, material inputs

[25] Smaller firms are re-sampled annually. Because our production function estimation requires lagged instruments, small firms are not included unless they were randomly sampled for two consecutive years. Our data should therefore be seen as exhaustive of manufacturing firms with at least 20 employees or 5 million euros in revenue.

[26] Data on domestic sales is also available separately, but because we do not have data on the fraction of inputs that account for exports we cannot rely on data on domestic sales to estimate the production function.

[27] We do not rely on the perpetual inventory method because that would require a guess for the firm's initial value of capital. Because our data only cover 11 years, this would lead to a particularly large measurement error (see, e.g., Collard-Wexler and De Loecker 2020). Data on investments are furthermore missing from FARE in 2008. For 2009 to 2019, the correlation between balance sheet capital and estimates of capital from the perpetual inventory method have a correlation of 0.92 to 0.99, depending on the assumed rate of depreciation.

[28] EAP separately reports data for various models of production based on the degree to which the producer is a subcontractor or subcontracts production. We define revenue and quantities for a product as the sum of revenues and quantities over all modes of production for a product in a given firm year.

[29] Examples of units of accounts are kilos, tons or pieces. We combine units of accounts and product codes, as firms that use different units of accounts for the same product might produce relatively heterogeneous goods

[30] As a robustness check we standardize prices using the revenue-weighted average price at the 8-digit sector level. The resulting firm-level prices have a 0.89 correlation with prices standardized at the 10-digit product level.

[31] At the time of writing, the most recent year for EU-KLEMS deflators is 2017. To deflate 2019 variables we extrapolate the price index using the sector's average inflation in other years.

Table 1: Summary Statistics

Variable	Mean	St. Dev.	Median	10th Pct.	90th Pct.	Observations
Revenue	16,911	66,723	3,045	544	31,346	175,538
Quantity	14,845	62,121	1,891	236	27,458	175,538
Wage Bill	3,346	12,865	830	194	6,505	175,538
Capital	8,343	35,803	869	114	13,635	175,538
Purchased Materials	7,561	29,763	1,017	116	13,730	175,538
Purchased Services	4,253	22,880	755	120	7,388	175,538
Standardized Price	9.45	89.29	1.23	0.77	6.25	175,538

Note: Summary statistics for French manufacturing firms from 2009 to 2019. Data are obtained from FARE (balance sheet and income statement variables) and EAP (normalized prices). Nominal values are deflated using two-digit EU-KLEMS deflators and are expressed in thousands of 2010 euros. Quantity is measured as firm-deflated revenue. The dataset contains 26,143 unique firms across 206 (19) sectors at the five (two) digit level. All variables are winsorized at their 1% tails.

and purchased services are deflated using the intermediate input deflator. Wages and the capital stock are deflated using the GDP deflator.

We drop firms with missing, zero or negative revenue, material purchases, service purchases, wage bills or capital.[32] We drop firms without price data in EAP, which restricts the sample to manufacturing firms. We also drop firms with fewer than two employees, as the number of single-employee firms has grown rapidly over our sample due to a regulatory change. To treat for outliers in the remaining sample we winsorize sales, quantity, prices, material and service inputs, and capital at the 1% level within two-digit industries. The resulting sample contains 157,277 firm-years for 26,143 unique firms across 206 five-digit sectors. Summary statistics are provided in Table 1. Table 2 describes the two-digit sectors in our analysis.

5 Simulation

In this section we estimate production functions and markups in a setting where their true values are known. To do so, we estimate the production function for a set of simulated firms in a rich macroeconomic model. Firms are heterogeneous in their productivity, the quantity of a fixed input at their disposal, and therefore the market share that they achieve. Heterogeneous market shares cause differences in markups across firms, determined endogenously as a consequence of oligopolistic competition.

5.1 Model

We analyze a single sector. A sector is defined as a collection of firms that have the same structural parameters of their production function and that face the same prices on input markets.

[32]We calculate market share before restricting the sample.

Table 2: Sectors (two-digit) in the EAP-FARE Dataset

Description	NACE code	Observations
Manufacturing of ...		
... textiles	13	6,716
... wearing apparel	14	5,200
... leather and related products	15	2,256
... wood and products of wood and cork, except furniture	16	9,599
... paper and paper products	17	6,511
... printing and reproduction of recorded media	18	8,589
... chemicals and chemical products	20	8,498
... rubber and plastic products	22	17,939
... other non-metallic mineral products	23	13,850
... basic metals	24	4,471
... fabricated metal products, except machinery and equipment	25	26,693
... computer, electronic and optical products	26	6,401
... electrical equipment	27	7,575
... machinery and equipment n.e.c.	28	16,738
... motor vehicles, trailers and semi-trailers	29	5,493
... other transport equipment	30	889
... furniture	31	10,844
Other manufacturing	32	5,094
Repair and installation of machinery and equipment	33	12,182

Demand We choose a market structure that allows firms to have heterogeneous markups that are determined by a combination of structural parameters and their market share. Following Atkeson and Burstein (2008), we implement this by assuming that firms compete in a double-nested CES demand system. The sector consists of a discrete number N markets, where a market is defined as a group of firms that compete oligopolistically with one another. Output across markets, which are indexed by h, is aggregated to the sector level along

$$Y_t = \left[\sum_{h=1}^{N} Y_{ht}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{9}$$

where σ denotes the elasticity of substitution across market-level goods. Market-level output Y_{ht} is the aggregate of firm-level output across the N_h firms that operate in h along

$$Y_{ht} = \left[\sum_{i=1}^{N_h} Y_{iht}^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{10}$$

where Y_{iht} denotes the output of firm i and where ε denotes the elasticity of substitution across firm-level goods within a market. Following Atkeson and Burstein (2008), we assume that $\varepsilon > \sigma$, reflecting that it is easier to substitute goods across firms than across markets. The double-nested CES system gives rise to the standard demand function for firm i's output:

$$Y_{iht} = \left(\frac{P_{iht}}{P_{ht}}\right)^{-\varepsilon} Y_{ht}, \quad \text{where} \quad Y_{ht} = P_{ht}^{\sigma} D_t, \tag{11}$$

where aggregate demand D_t is exogenous and where P_{ht} is the usual CES market price index:

$$P_{ht} = \left(\sum_{i=1}^{N_h} P_{iht}^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}. \tag{12}$$

The price-setting firm internalizes that P_{ht} increases when it raises its own prices. Following Burstein et al. (2020), we assume for tractability that firms do not internalize that it may induce an increase in the aggregate demand D_t. As such, we assume that firms behave as if markets are atomistic (as in Atkeson and Burstein 2008), despite the actual setup featuring a finite number of markets. Under Cournot competition, the resultant profit-maximizing markup reads as

$$\mu_{iht} = \frac{\varepsilon}{\varepsilon - 1} \left(1 - \frac{\frac{\varepsilon}{\sigma} - 1}{\varepsilon - 1} s_{iht} \right)^{-1}, \tag{13}$$

where market share s_{iht} is defined as the firm's share in market revenue:

$$s_{iht} = \frac{P_{iht} Y_{iht}}{P_{h(i)t} Y_{h(i)t}}, \tag{14}$$

where subscript $h(i)$ indicates the market in which firm i operates.[33] The firm's markup ranges from $\varepsilon/(\varepsilon - 1)$ for a firm whose market share approaches zero to $\sigma/(\sigma - 1)$ for a monopolist, which is higher than the small firm's markup, given the assumption that $\varepsilon > \sigma$.

Technology Firms produce using a variable input V_{iht} and a fixed input K_{iht}, with log-inputs respectively denoted by v_{iht} and k_{iht}. The production function for log output y_{iht} is translog:

$$y_{iht} = \omega_{it} + \gamma \alpha v_{iht} + \gamma (1 - \alpha) k_{iht} + \gamma \frac{\alpha (1 - \alpha)}{2} \frac{\phi - 1}{\phi} \left(v_{iht}^2 + k_{iht}^2 - 2 k_{iht} v_{iht} \right), \tag{15}$$

where ω_{it} is the log of (hicks-neutral) total factor productivity, γ measures the degree of returns to scale, α determines the weight of the variable input in the production function, while ϕ approximates the elasticity of substitution between the flexible and the fixed input. Our log production function (15) is motivated by the following generalized constant elasticity of substitution production function:

$$Y_{iht} = \Omega_{iht} \left(\alpha V_{iht}^{\frac{\phi - 1}{\phi}} + (1 - \alpha) K_{iht}^{\frac{\phi - 1}{\phi}} \right)^{\frac{\phi}{\phi - 1} \gamma},$$

where $\Omega_{iht} \equiv \exp \omega_{iht}$. In Appendix B we show that this production function converges to the Cobb-Douglas production function as $\phi \to 1$ and that an approximation of the production

[33] Bertrand competition, where the firm's first order condition takes prices rather than quantities as given, yields a similar expression. See Atkeson and Burstein (2008), Grassi (2017), and Burstein et al. (2020) for an elaborate discussion. For the purpose of these simulations we require that markups are determined by demand elasticities and market share, which is is the case for both Cournot and Bertrand competition.

function around $\phi = 1$ yields translog function (15). We specify the production function with a constant degree of homogeneity (γ) such that the model admits an analytical expression for (size-dependent) marginal costs, facilitating the calculation of the equilibrium.

Equilibrium We consider the following partial equilibrium. Given an exogenous sequence for variable input prices W_t, aggregate demand D_t, productivities ω_{iht} and fixed factors k_{iht}, the equilibrium is defined as a sequence of markups μ_{iht}, prices P_{iht}, log marginal costs mc_{iht}, market shares s_{iht}, log variable inputs v_{iht},and log outputs y_{iht} and sector price indices P_{ht} such that output follows from demand function (11), sector price indices follow from (12), markups are set along equation (13), market share is given by (14), prices are the product of marginal costs and markups, variable inputs are in line with the firm's first-order condition and marginal costs follow from (15), for all i, h and t. Derivations are provided in Appendix B.

5.2 Calibration

We simulate the behavior of 1440 firms, which is the average number of firms in two-sector industries in the EAP data. We divide these firms into 180 markets and simulate the economy for 40 periods.[34] There are 13 parameters, each of which we calibrate externally. The parameters are summarized in Table 3. In calibrating the model, we are constrained by the fact that the true values of many parameters (such as those of the production function and the productivity process) are in fact the object of our empirical analysis. Our approach is therefore to assume reasonable values in line with the literature as an example of a possible quantification.

There are two aggregate shocks: aggregate demand D_t and prices of the variable input W_t. We assume both series follow a log-linear first-order autoregressive process with persistence ρ_{PY} and ρ_W, respectively, with shocks $\xi_{PY} \sim N(0, \sigma_{PY})$ and $\xi_W \sim N(0, \sigma_W)$. Fluctuations in aggregate demand ensure that the relationship between output and market share vary over time. Fluctuations in the price of the variable input ensure that firms' lagged productivity and lagged variable inputs are not co-linear after conditioning on the fixed inputs, which is needed to be able to separately identify the productivity process and the production function parameters, as discussed in Section 3. To calibrate the process for the price of the flexible inputs, we estimate an autoregressive process for the price index of intermediate inputs from sector-level manufacturing data in EU-KLEMS. We run simple autoregressive regressions for the log of the index, and find an autoregressive coefficient ρ^w of 0.87 at the two-digit sector level when controlling for industry- and year fixed effects. Residuals have a standard deviation σ^w of 0.06.[35] For aggre-

[34] Recall that markets define the level at which firms compete. By modeling many small markets rather than a small number of large markets we reduce the computational complexity of the simulation. An appropriate calibration of the productivity process ensures that markets have realistic levels of concentration and markup dispersion.

[35] Appendix Table A2 presents AR(1) coefficients for various specifications, which suggest a narrow range of 0.86 to 0.90 for the AR(1) coefficient and 0.042 to 0.046 for the standard deviation of the shocks.

Table 3: Parameter Calibration for Simulation

Parameters	Value	Description
α	0.4	Share of variable input
γ	0.8	Returns to scale
ϕ	1.1	Elasticity of substitution
σ	1.1	Demand elasticity across markets
ε	10	Demand elasticity across firms in a market
N, N_h	180, 8	Number of markets and firms per market
ρ^w, σ^w	0.87 , 0.06	AR(1) persistence and std. dev. of W_t
ρ^D, σ^D	0.78 , 0.19	AR(1) persistence and std. dev. of $P_t^{-\sigma} Y_t$
$\rho^\omega, \sigma^\omega$	0.70 , 0.10	AR(1) persistence and std. dev. of firm-level ω_{it}
ρ^k, σ^k	0.66 , 0.66	AR(1) persistence and std. dev. of firm-level k_{it}
$\tilde{\sigma}^\eta$	0.095	std. dev. meas. error on output

gate demand P_t we estimate a similar autoregressive process, using the detrended sector-level nominal value added as the dependent variable.[36] We find a high degree of persistence in aggregate demand, with a ρ^D of 0.78, while the residuals have a standard deviation of 0.19.

There are two sources of firm heterogeneity in the model: the firm's log-endowment of the fixed input k_{iht} and the firm's log-total factor productivity ω_{it}. Both evolve exogenously through linear first-order autoregressive processes with persistence ρ_k and ρ_ω, respectively, and are subject to innovations $\xi_k \sim N(0, \sigma_k)$ and $\xi_\omega \sim N(0, \sigma_\omega)$. Both sources of firm heterogeneity are similar in that firms with either higher productivities or higher values for the exogenous fixed input have, *ceteris paribus*, greater output. They are different in that the fixed input is observable, while productivity is not. To calibrate the persistence and volatility of the fixed factor, we run autoregressive regressions on the log of capital in the EAP data. We find a persistence parameter ρ^k of 0.66 and a volatility of shocks σ_k of 0.66.[37] A particular challenge is the estimation of the persistence ρ^ω and volatility σ^ω of the productivity process. To obtain these empirically requires knowledge of the parameters of the production function, which is the objective of our analysis. We take the pragmatic approach of calibrating ρ^ω and σ^ω in line with common values of the literature, and check that these values are in line with our findings in Section 4. We set ρ^ω to 0.6, in line with Decker et al. (2020), and set productivity volatility σ^ω to 0.1.

When calibrating the production function, we think of purchased materials as v_{iht} and a composite of all other factors as k_{iht}. We calibrate α to 0.4 to match the average ratio of material purchases over revenue in EAP-FARE, which is 0.38. We calibrate returns-to-scale parameter γ to 0.8 in order to have modest decreasing returns to scale, in line with the estimate by Basu and Fernald (1997). We assume an elasticity of substitution ϕ of 1.1, as purchased materials include intermediate inputs from other firms, which can substitute for in-house production.

[36] We detrend $P_t^{-\sigma} Y_t$ using nominal GDP to account both for trend increases in prices and for aggregate growth in order to obtain a stationary nominal series. Results are similar when detrending with the GDP deflator.

[37] Appendix Table A3 presents the AR(1) estimates for capital.

We introduce measurement error in observed quantity y_{iht}, denoted by η_{iht}, after computing the equilibrium. We assume that $\eta_{iht} \sim N(0, \sigma_y \tilde{\sigma}_\eta)$, where σ_y is the standard deviation of true output across all firm-years in the sector, and $\tilde{\sigma}_\eta$ is a scalar that determines the magnitude of measurement error relative to the standard deviation of true output. We calibrate $\tilde{\sigma}_\eta$ to 0.095, in line with the relative variance of output and fitted values of a regression of output on prices, market share, time fixed effects and a third-degree polynomial in the firms' inputs in EAP.

5.3 Results

5.3.1 Production function estimation

We now take the simulated firm-level data on revenue, output and inputs and use it to estimate markups along the two-step iterative GMM procedure. We estimate various alternative specifications of the procedure. To match the approaches typically followed in the literature, in all specifications we make the assumption rat the researcher correctly assumes that the variable input is v_{iht}. This is therefore also the variable that we use to calculate markups after estimating the production function. The moment conditions are that the residual of the AR(1) process for productivity, ξ_{iht}, is orthogonal to the lagged variable input and to the current fixed input.[38]

In order to establish that it is feasible to estimate the production function parameters and markups in our setup, we first estimate a preferred specification. The preferred specification has three components, each of which we deviate from in subsequent specifications. Firstly, we estimate our preferred specification using quantity as the measure of a firm's output; hence, we assume that the researcher perfectly observes the prices that firms set. Secondly, we estimate the preferred specification using a theoretically valid first-stage regression. As shown in Section 3, to correctly control for ω_{iht} in the regression that purges measurement error η_{iht}, the control variables must account for the log of marginal costs. To do so, we include as additional controls both the log of the price and the firm's market share in the first stage, where the latter is the proxy for the markup.[39] Combined with a third-order expansion of the inputs v_{it} and k_{it}, this should allow us to identify the measurement error η_{iht} with reasonable precision. Thirdly, the preferred specification estimates a production function of the translog form, in line with (15):

$$y_{iht} = \beta_v v_{iht} + \beta_k k_{iht} + \beta_{vv} v_{iht}^2 + \beta_{kk} k_{iht}^2 + \beta_{vk} k_{iht} v_{iht} + \omega_{iht}, \tag{16}$$

where elasticities satisfy the following relations with the true production parameters (α, γ, η):

$$\beta_v = \gamma\alpha, \quad \beta_k = \gamma(1-\alpha), \quad \beta_{vv} = \gamma \tfrac{\alpha(1-\alpha)}{2} \tfrac{\phi-1}{\phi}, \quad \beta_{kk} = \beta_{vv}, \quad \beta_{vk} = -2\beta_{vv}.$$

[38]We add a constant to the second-stage auto-regressive productivity estimation, so that no constant needs to be added to the production function itself. It is straightforward to show that estimating a constant in the production function or estimating a constant in the AR process is equivalent.

[39]We do not include a polynomial of market share to control for non-linearities in the markup-market share relationship because the correlation between market share and its square exceeds 0.99.

Table 4: Estimated Translog Production Function Parameters

Coefficients	True	Quantity			Revenue		
		Full	Basic	None	Full	Basic	None
$\beta_v = \alpha\gamma$	0.32	0.32	0.32	0.38	0.39	0.39	0.47
		(0.01)	(0.011)	(0.018)	(0.013)	(0.014)	(0.021)
$\beta_k = (1-\alpha)\gamma$	0.48	0.47	0.47	0.45	0.28	0.28	0.25
		(0.006)	(0.006)	(0.01)	(0.007)	(0.007)	(0.011)
$\beta_{vv} = \gamma\frac{\alpha(1-\alpha)}{2}\frac{\phi-1}{\phi}$	0.009	0.01	0.008	0.021	0.021	0.02	0.036
		(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
$\beta_{kk} = \beta_{vv}$	0.009	0.009	0.007	0.011	-0.001	-0.001	0.002
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\beta_{vk} = -2\beta_{vv}$	-0.017	-0.019	-0.016	-0.029	-0.026	-0.024	-0.039
		(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
Implied avg. elasticity	0.309	0.312	0.31	0.324	0.32	0.319	0.334
(Std. dev.)	(0.019)	(0.021)	(0.018)	(0.042)	(0.041)	(0.04)	(0.071)

Note: Estimated production-function coefficients for different specifications. The top panel presents production-function estimates. The bottom panel presents elasticities of the variable input v on output (measured in terms of quantity or revenue). The first column presents true values for comparison. Bootstrapped standard errors are in parentheses. Full, Basic and None describe the specification of the first-stage regressions. Full first stages include a third-order expansion in the production-function inputs, time fixed effects and additional controls for log price and market share. Basic first stages do not include the additional controls. Columns headed None do not deploy a first stage and therefore estimate markups on output variables that include measurement error η_{iht}.

Note that we do not impose these theoretical restrictions when estimating (16). We then consider various 'imperfect' specifications of the two-stage GMM procedure and see how the production function and markup estimates change.

5.3.2 Elasticity estimates and markups

The estimates of the translog production function parameters are presented in Table 4. Coefficients in the column titled 'True' are directly calculated from the deep production function parameters (α, γ, ϕ). The three subsequent columns present estimates of the production function where output is measured in quantities while the final three columns present estimates where revenue is used. Bootstrapped standard errors are in parentheses.

The preferred specification is presented in the second column, where 'Full' indicates that the first stage includes log price and market share controls. The estimates show that the preferred specification is able to identify the parameters of the production successfully. All coefficients are within one tenth of a decimal point of their true value. Bootstrapped standard errors are generally small and coefficients are highly significant.[40] The markups which arise from these

[40]Standard errors differ from zero because (1) the first stage approximates the implicit relationship between productivity and inputs through a third-order polynomial and (2) market share proxies imperfectly for markups.

Table 5: Overview - Translog Log Markup Estimates

	Correlation ln $\widehat{\mu}_{iht}$ with true markup	Log Markup Moments			
		Mean	St. Dev.	Median	IQR
True values	1.00	0.258	0.065	0.213	0.367
Quantity					
Full first stage (preferred)	1.00	0.266	0.064	0.220	0.346
Basic first stage	1.00	0.263	0.067	0.218	0.370
No first stage	0.66	0.300	0.102	0.225	0.621
Revenue					
Full first stage	0.69	0.286	0.108	0.206	0.713
Basic first stage	0.73	0.286	0.105	0.208	0.677
No first stage	0.40	0.312	0.201	0.184	1.543

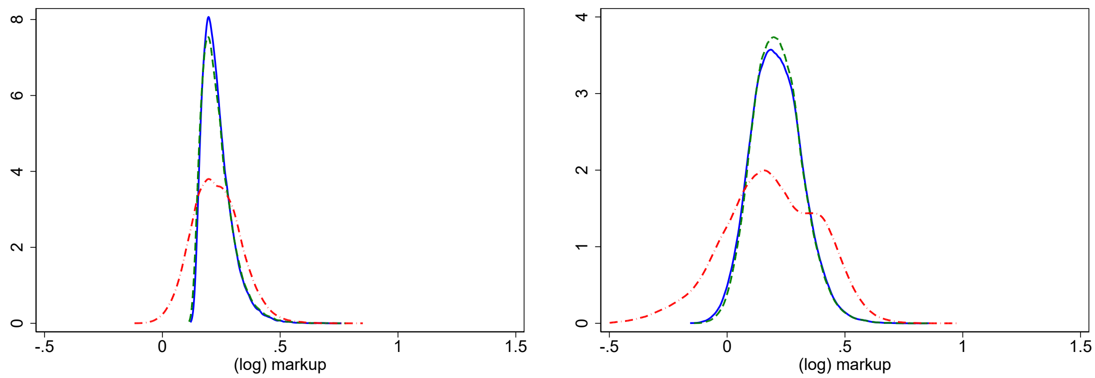
Note: The first column presents correlations of estimated markups with true values. Full first stages include a third-order expansion in production inputs and additional controls for log price and market share. Basic first stages do not include the additional controls.

production function estimates are summarized in the second row of Table 5. Results for the preferred specification with a translog production function, quantity data and a full first stage are closely in line with true markups. The estimated markups have a correlation of 1.00 with the true markup, although the level and standard deviation of the markup are estimated with slight error. This is in line with the modest differences between the true and estimated production function parameters in Table 4.

Revenue versus quantity We next deviate from the preferred specification by using revenue instead of quantity data to estimate the production function. In the fifth column of Table 4 we report a significantly higher estimate for β_v, which increases from 0.32 to 0.39. The sixth column, where the first-stage regression also does not control for price, finds an identical β_v. Conversely, the linear coefficient for the fixed input, β_k, falls from 0.47 to 0.28. The increase in β_v might be surprising, because the revenue elasticity of an input should fall short of the quantity elasticity when demand curves are sloping downward. Recall, however, the result in (6) that revenue-based coefficients can be biased upwards, downwards or be unaffected, depending on the correlation between prices and inputs. Only in the absence of time-fixed effects are the elasticities biased downwards by the inverse average markup. Our simulated firms are subject to aggregate demand shocks, which create a positive correlation between input usage and prices under diminishing returns to scale. Indeed, prices have a 0.42 correlation with variable inputs in the simulation. Controlling for time fixed effects, the correlation is negative - as expected.

The bottom panel of Table 5 compares markup estimates based on the revenue data. Average markups are overestimated, in line with the overestimation of β_v. We find that revenue-based markups are highly informative of true markups, with a point correlation of 0.73 between the preferred estimate and the revenue-based counterpart. These results confirm that the revenue-

Figure 1: Distribution of Simulated (Log) Markup by Output Variable, First-Stage Specification



(a) Quantity (b) Revenue

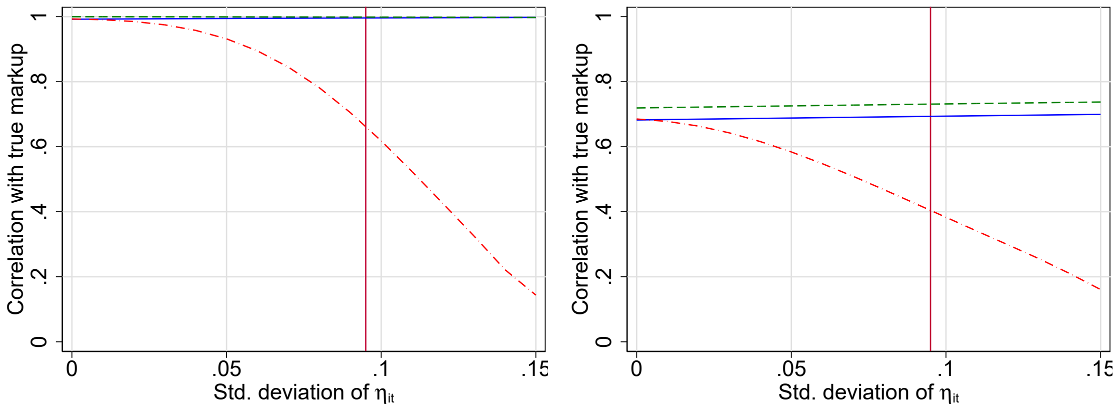
Notes: The figure plots the distribution of simulated markups (in logs). Blue (solid) lines present markups estimated with a first stage that includes price and market share controls. Green (dashed) lines present specifications that exclude price and market share controls. Red (dash-dotted) lines present results from specifications without a first stage.

based estimates of the production function elasticities are *not* the revenue elasticities of an input. If they were, Bond et al. (2021) show that the log markups should equal 0 and be uninformative of true markups. Rather, the revenue-based elasticities are *biased* estimates of output elasticities of the inputs. This bias may cause the log-markup to *average* 0 up to a Jensen's inequality, or to have a further bias due to the presence of time effects such as demand shocks.

It is important to note that the bias in revenue-based markup estimates depends on the calibration. Our baseline calibration is in line with prior work, and the high correlation between revenue-based markup estimates and true markups is robust to many alternative calibrations. As an example, Table A4 and A5 in Appendix C present markup estimates from a calibration with higher returns to scale and higher material share. While revenue markups still correlate well with true markups, the bias in their average level is negative rather than positive.

First stage In a second deviation from the preferred specification, we compare production function estimates with different first stages. Results so far include a third-order expansion of the inputs as well as additional control variables for price and market share. We next consider a 'basic' first stage where we drop the control variables for price and market share. The resulting first-stage specification is frequently used in markup estimations (e.g. De Loecker and Warzynski 2012, De Loecker et al. 2020). We find similar estimates for the parameters in the production function when using the basic first stage. The linear coefficients β_v and β_k are unaffected for both quantity and revenue-based estimations, although all of the higher-order terms are slightly underestimated. Looking at the correlation with the true (log) markup, Table 5 shows that markups from the basic first stage again have a correlation of 1.00 with true markups. The moments of the markup distribution are similar to the full-first-stage markups.

Figure 2: Relationship between Variance of Measurement Error and Markup Correlations



(a) Quantity

(b) Revenue

Notes: The figures plot the correlation between the true markups and estimates of the markups under different calibrations of σ^{η}, which measures the fraction of observed output that is due to unobserved productivity shocks or measurement error. The vertical line is at the baseline calibration in the simulations. Blue (solid) and green (dashed) lines are estimates from the full and basic first-stage estimations, respectively. Red (dash-dotted) lines do not purge for η_{it} in a first stage, as in Blundell and Bond (2000).

In contrast, we find that the accuracy of the production function estimates declines sharply when no first stage is conducted. In this case, the production function estimation becomes a simple IV-GMM estimation with lagged variable inputs used to instrument for current variable inputs. The unidentified measurement error causes an overestimation of the output elasticity of the variable input. This causes an overestimation of the average log markup in Table 5 and a significant reduction (from 1 to 0.66) in the correlation between estimated and true markups. Figure 4, which plots the kernel densities of the markup estimates, illustrates the poor performance of the estimator without a first stage (red, dash-dotted). The variance of the estimates is much larger than that of the full (solid-blue) and basic (green-dash) first-stage estimates.

The driver of the poor performance of the no-first-stage estimates is clear from Figure 2. It presents the correlation between true markups and the various markup estimates, for various calibrations of σ^{η}. In simulations without measurement error or transitory productivity shocks ($\sigma^{\eta} = 0$), all estimators perform similarly and the quantity-based markups have perfect correlations with true markups. The performance of the no-first-stage estimates quickly decays, however, as correlations approach zero when errors make up 15% of observed output ($\sigma^{\eta} = 0.15$).

These results are in line with the predictions in Section 3. While estimates without a first stage are unbiased, they come with greater variance. The simulations suggest that it is greatly preferable to include a first stage to purge transitory productivity shocks and measurement errors, even if one does not have good controls for the markup – as was the case in our basic first stage.

Table 6: Correlations across Simulated Specifications - Log Markups

	True	Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
				Pearson Correlation			
True	1.00	1.00	0.69	1.00	0.73	0.66	0.40
Full First Stage - Quantity	1.00	1.00	0.75	0.99	0.79	0.72	0.48
Full First Stage - Revenue	0.69	0.75	1.00	0.66	1.00	1.00	0.94
Basic First Stage - Quantity	1.00	0.99	0.66	1.00	0.70	0.62	0.36
Basic First Stage - Revenue	0.73	0.79	1.00	0.70	1.00	0.99	0.92
No First Stage - Quantity	0.66	0.72	1.00	0.62	0.99	1.00	0.95
No First Stage - Revenue	0.40	0.48	0.94	0.36	0.92	0.95	1.00
				Spearman Rank Correlation			
True	1.00	0.99	0.65	1.00	0.69	0.61	0.39
Full First Stage - Quantity	0.99	1.00	0.72	0.98	0.76	0.69	0.48
Full First Stage - Revenue	0.65	0.72	1.00	0.60	1.00	1.00	0.94
Basic First Stage - Quantity	1.00	0.98	0.60	1.00	0.64	0.56	0.33
Basic First Stage - Revenue	0.69	0.76	1.00	0.64	1.00	0.99	0.92
No First Stage - Quantity	0.61	0.69	1.00	0.56	0.99	1.00	0.96
No First Stage - Revenue	0.39	0.48	0.94	0.33	0.92	0.96	1.00

Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage. All markups are expressed in log. Simulated data.

5.3.3 Markup Correlations

In the final analysis of the simulated data, we examine how markup estimates correlate across specifications and with key variables such as profits and the labor share. Table 6 presents the correlations of markups across markup specifications. The table shows that the correlations between specifications are generally high, and often of similar magnitudes as the correlation between markup estimates and the true markup. The correlation between the preferred estimate, with quantity to measure output and the full first stage, and other markups from the two-stage procedure, are at least 0.79. This is the case for both the Pearson and the rank correlations. Figure 3 provides a graphical illustration by means of a binned scatter plot between the preferred specification and the most commonly used empirical specification, with revenue data and the basic first stage. The right-hand figure, which plots the relationship between the markup estimates in first-differences, confirms that these are tightly linked. Analysis that studies trends in markups over time therefore seems particularly feasible with revenue-based markups.

Table 7 then runs a number of canonical regressions on the relationship between markups and other variables. The idea is to check whether these regressions are qualitatively similar when using different markup estimates. For each specifications s, we run the following

$$x_{it} = \chi(\ln \widehat{\mu}_{it}^{s}) + \varphi_i + \psi_t + \epsilon_{it}, \tag{17}$$

where respectively φ_i and ψ_t denote firm- and time fixed effects, and where x_{it} denotes some variable of interest. We estimate this regression using a firm's profit rate (ratio of operating

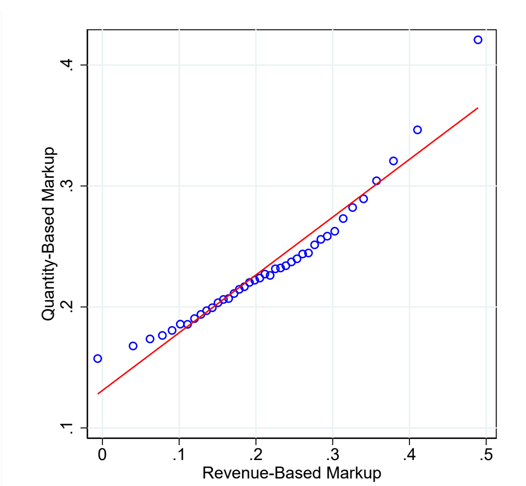
Table 7: Simulated Relation between Markup and Other Variables by Markup Specification

		Quantity			Revenue		
	True	Full	Basic	None	Full	Basic	None
Profit Rate	0.0196***	0.0195***	0.0198***	0.0259***	0.0260***	0.0252***	0.0344***
	(0.00005)	(0.00006)	(0.00005)	(0.00010)	(0.00008)	(0.00007)	(0.00020)
R-squared	0.936	0.922	0.946	0.825	0.863	0.881	0.709
Materials Share	-0.0196***	-0.0195***	-0.0198***	-0.0259***	-0.0260***	-0.0252***	-0.0344***
	(0.00005)	(0.00006)	(0.00005)	(0.00010)	(0.00008)	(0.00007)	(0.00020)
R-squared	0.936	0.922	0.946	0.825	0.863	0.881	0.709
Market Share (%)	0.0614***	0.0616***	0.0615***	0.0871***	0.0855***	0.0820***	0.122***
	(0.00008)	(0.00008)	(0.00007)	(0.00016)	(0.00013)	(0.0001)	(0.00055)
R-squared	0.997	0.995	0.996	0.940	0.965	0.976	0.808

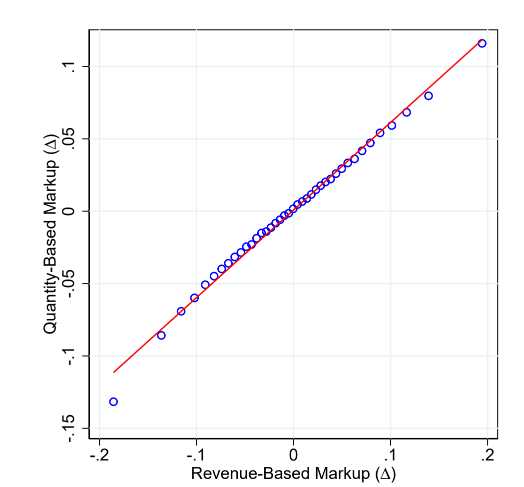
Note: Each entry gives the OLS coefficient with the cursive variable as the dependent variable and the markup series in the column header as the explanatory variable. The markup estimation's first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage ("None"). All markups are expressed in logs, they are normalized to have standard deviation equal to 1, and they are computed on simulated data. Firm-clustered standard errors in parentheses. *** denotes significance at the 1% level. All specifications include time- and firm fixed effects. 63,624 observations.

profits over sales), material cost share (ratio of variable-input spending over sales), and market share as dependent variables. We divide markup estimates by their standard deviations to ease comparison of columns. The table confirms that all markup estimates do reasonably well at retrieving the OLS coefficient χ from the true markup. The relationship is best estimated using quantity-based markups and a first stage, and worst when not including a first stage.

Figure 3: Binned Scatter Plot for Simulated Quantity and Revenue-Based Markups



(a) Log-Markups



(b) Log-Differenced Markups

Notes: The figures plot the linear relationship and binned scatter plot between quantity-based markups (full first stage) and revenue-based markups (basic first stage) in simulated data. Log-markups are used in figure (a), log-differenced markups in figure (b). Regression coefficients for the linear fit are 0.48 and 0.61, respectively.

6 Empirics

This section describes the results from the production function and markup estimation on the French EAP-FARE manufacturing data. We start by assessing the elasticities of quantity and revenue with respect to materials. We then compare the levels and dispersion of markups from various specifications, and assess the correlation between the various markup estimates. Finally, we look at how estimated key relationships such as the markup-market share and the markup-profit relationship depend on production function specifications.

6.1 Production Function Estimates

In line with our simulations, we estimate the production function in six specifications comprised of the combinations of quantity or revenue as output measures, and either the full first stage (with price and market share controls), the basic first stage (with only the polynomial in inputs), or an absent first stage. We assume that log output y_{it} follows a translog production function of the log inputs materials m_{it}, the wage bill l_{it}, capital k_{it} and services o_{it}:

$$y_{it} = \beta_m m_{it} + \beta_l l_{it} + \beta_k k_{it} + \beta_o o_{it} + \sum_{h\in\{m,l,k,o\}} \sum_{j\in\{m,l,k,o\}} \beta_{hj} h_{it} j_{it}, \tag{18}$$

which we estimate for each two-digit sector. Following Burstein et al. (2020) we assume that materials involve no adjustment costs and therefore correspond to the variable input v_{iht} in Section 3.[41] To estimate markups, we are therefore interested in the output elasticity:

$$\alpha^m_{it} = \beta_m + 2 \cdot \beta_{mm} m_{it} + \beta_{mo} o_{it} + \beta_{ml} l_{it} + \beta_{mk} k_{it}.$$

Table 8 presents the estimated material elasticities α^m_{it} for each of our specifications. Specifications in the first three columns use quantity as the measure of output, while the final three columns use revenue. Rows present averages for α^m_{it} within the two-digit sector for which the production function was estimated, with standard deviations in parentheses.

Our preferred specification, which uses the full first stage and quantity data, yields an average output elasticity of 0.62. In line with the notion that firms face downward-sloping demand curves, we find that the revenue-based output elasticity of materials is usually lower than a quantity-based one. For our preferred specification we find higher average elasticities than the revenue-based counterpart in 18 out of 19 industries. On average, the quantity-based output elasticity exceeds the revenue-based elasticity by 50%.

[41] This is more appropriate than assuming that labor is freely set, especially for France (e.g., Caselli et al. (2021).

Table 8: Estimated Translog Material-Output Elasticity by Sector and Specification

	Quantity			Revenue		
NACE	Full	Basic	None	Full	Basic	None
Avg.	0.62	0.53	0.57	0.42	0.42	0.42
	(0.34)	(0.22)	(0.28)	(0.13)	(0.13)	(0.14)
13	0.53	0.46	0.53	0.44	0.44	0.43
	(0.18)	(0.17)	(0.2)	(0.12)	(0.13)	(0.11)
14	0.69	0.57	0.66	0.40	0.40	0.41
	(0.22)	(0.18)	(0.22)	(0.11)	(0.11)	(0.12)
15	0.44	0.40	0.40	0.42	0.39	0.44
	(0.12)	(0.13)	(0.10)	(0.12)	(0.12)	(0.15)
16	0.59	0.54	0.59	0.49	0.49	0.48
	(0.14)	(0.15)	(0.13)	(0.12)	(0.12)	(0.14)
17	0.55	0.55	0.55	0.47	0.47	0.45
	(0.15)	(0.16)	(0.16)	(0.10)	(0.10)	(0.10)
18	0.37	0.37	0.35	0.31	0.31	0.30
	(0.15)	(0.16)	(0.15)	(0.09)	(0.09)	(0.10)
20	0.95	0.79	0.92	0.49	0.49	0.50
	(0.38)	(0.26)	(0.37)	(0.12)	(0.11)	(0.13)
22	0.62	0.55	0.62	0.47	0.47	0.46
	(0.15)	(0.15)	(0.15)	(0.11)	(0.11)	(0.11)
23	0.50	0.45	0.54	0.42	0.41	0.44
	(0.12)	(0.08)	(0.12)	(0.14)	(0.13)	(0.17)
24	0.65	0.71	0.65	0.45	0.45	0.45
	(0.20)	(0.15)	(0.23)	(0.17)	(0.17)	(0.20)
25	0.45	0.42	0.44	0.38	0.38	0.38
	(0.21)	(0.16)	(0.18)	(0.14)	(0.14)	(0.14)
26	1.15	0.65	1.03	0.42	0.42	0.40
	(0.47)	(0.32)	(0.42)	(0.09)	(0.10)	(0.10)
27	0.68	0.62	0.66	0.48	0.48	0.50
	(0.27)	(0.17)	(0.23)	(0.11)	(0.11)	(0.21)
28	0.97	0.56	0.51	0.47	0.48	0.48
	(0.60)	(0.11)	(0.32)	(0.10)	(0.10)	(0.09)
29	0.55	0.95	0.50	0.54	0.54	0.53
	(0.23)	(0.36)	(0.20)	(0.14)	(0.13)	(0.14)
31	1.07	0.70	0.85	0.40	0.40	0.39
	(0.18)	(0.11)	(0.27)	(0.09)	(0.09)	(0.09)
32	0.58	0.43	0.63	0.35	0.35	0.36
	(0.22)	(0.25)	(0.21)	(0.09)	(0.10)	(0.10)
33	0.28	0.33	0.30	0.33	0.32	0.33
	(0.10)	(0.09)	(0.13)	(0.11)	(0.10)	(0.11)

Note: The table presents estimated elasticities of materials on output (measured in terms of quantity or revenue) from the estimation of translog production functions. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects (“Full”), only the polynomial and time fixed effects (“Basic”) or no first stage ("None"). Translog specifications have heterogeneous elasticities within industries, with standard deviations presented in brackets. Industry codes refer to two-digit NACE codes. Industry names are provided in Table 2.

Table 9: Overview - Log Markup Estimates

	Mean	St. Dev.	Median	25th Pct.	75th Pct.	Observations
Quantity data						
Full first stage	0.39	0.50	0.36	0.17	0.65	157,485
Basic first stage	0.30	0.33	0.28	0.12	0.48	157,485
No first stage	0.33	0.45	0.32	0.15	0.56	157,485
Revenue data						
Full first stage	0.11	0.17	0.10	0.01	0.21	157,485
Basic first stage	0.11	0.17	0.10	0.01	0.21	157,485
No first stage	0.10	0.21	0.10	0.00	0.20	157,485

Note: All markups are expressed in log. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage ("None"). Data for 2009-2019 from EAP-FARE.

The estimated elasticities depend only modestly on the specification of the first stage. Notably, we only find minor differences between the (average) estimated elasticities between the specification with the full first stage and the specification with no first stage at all. This means that the first-stage purging of the production function estimation procedure has only modest effect on the output used for the production function estimation. The columns with the "basic" first stage that omits price or market share controls seem to have a slight downward bias in the estimated coefficients when using quantity to measure output.[42]

The results in Table 8 show that industries differ significantly in the elasticity of output with respect to materials. Industries with low elasticities include NACE industry 33 (repairs) while those with high elasticities include NACE industry 31 (manufacturing of furniture) and 28 (machinery and equipment). The standard deviations furthermore show that there is sizable heterogeneity in elasticities across firms within industries.

6.2 Markups

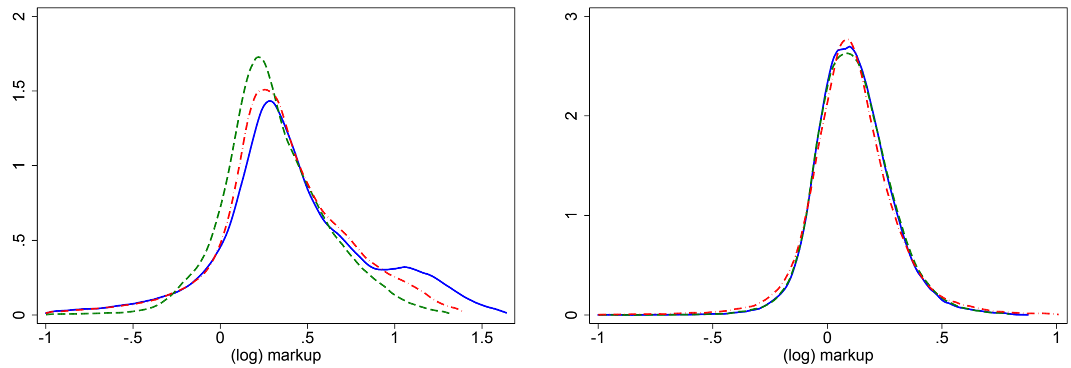
We next compute markups along the Hall (1986, 1988) equation using the estimated firm-level elasticities. In the remaining analysis we focus on the log of markups. To treat for outliers, we trim the bottom and top of the distribution at the 1.5% level for each specification. To facilitate comparison, we focus on the non-trimmed sample. This leaves 157,485 firm-year observations.

6.2.1 Levels and Distribution

Summary statistics are provided in Table 9. A clear pattern emerges from the table. First, the table shows that markups estimated from revenue data are consistently lower than markups

[42]The bias is negligible when using revenue to measure output. This is expected because, when using revenue, prices should be added to the right-hand side of the first stage equation. In the full-first stage, prices enter with a coefficient of -1 (see equation 5), and therefore cancel out.

Figure 4: Distribution of (Log) Markup by Output Variable and First-Stage Specification



(a) Quantity (b) Revenue

Notes: The figure plots the distribution of markups (in logs). Blue lines present markups estimated with a first stage that includes price and market share controls. Green lines present specifications that exclude price and market share controls. Red lines present results from specifications without a first stage.

estimated from quantity data. Average revenue-based markups are around 0.11 in logs and close to one in levels, in line with the bias described in Section 3.3 and in Bond et al. (2021). The average of the log markup of the preferred specification, translog in quantity with the full first stage, is 0.39.

A second pattern in Table 9 is that the distributions of markups across first-stage specifications are reasonably similar when holding the output variable constant. Most published moments are within 10 log-points of each other. An illustration of how similar these distributions are is provided in Figure 4. It plots the full (blue-solid), basic (green-dashed) and no (red-dash-dotted) first-stage specifications. The left figure plots the kernel densities for quantity, while the right figure plots them for revenue. There are some differences for quantity-based markup estimates, which match the downward bias in the elasticity estimates for the "basic" first stage. The distributions are overall similar in terms of mean, median and standard deviation. It therefore seems that the choice of the output variable has a much larger effect on the estimated markups than the exact specification of the first stage.

6.2.2 Correlations across Markup Estimates

The correlation between markups from various specifications is presented in Table 10. The top panel presents the Pearson correlations over the entire sample. Correlations are generally lower than in the simulations. This is expected, as the data include multiple sectors, while the simulation contains a single sector. Nevertheless, we find consistently positive correlations across the specifications. Like before, these correlations are particularly strong when comparing markups with production functions that use the same output variable. For quantity, for

Table 10: Correlations across Specifications - Log Markups

	Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
			Pearson Correlations			
Full first stage - Quantity data	1.00	0.28	0.60	0.29	0.67	0.28
Full first stage - Revenue data	0.28	1.00	0.43	0.98	0.31	0.74
Basic first stage - Quantity data	0.60	0.43	1.00	0.46	0.58	0.31
Basic first stage - Revenue data	0.29	0.98	0.46	1.00	0.32	0.73
No first stage - Quantity data	0.67	0.31	0.58	0.32	1.00	0.22
No first stage - Revenue data	0.28	0.74	0.31	0.73	0.22	1.00
			Spearman Rank Correlations			
Full first stage - Quantity data	1.00	0.33	0.71	0.34	0.81	0.33
Full first stage - Revenue data	0.33	1.00	0.49	0.99	0.37	0.86
Basic first stage - Quantity data	0.71	0.49	1.00	0.51	0.73	0.41
Basic first stage - Revenue data	0.34	0.99	0.51	1.00	0.38	0.84
No first stage - Quantity data	0.81	0.37	0.73	0.38	1.00	0.32
No first stage - Revenue data	0.33	0.86	0.41	0.84	0.32	1.00
			Spearman Rank Correlations - Within Sectors			
Full first stage - Quantity data	1.00	0.50	0.74	0.50	0.76	0.52
Full first stage - Revenue data	0.50	1.00	0.62	0.99	0.51	0.86
Basic first stage - Quantity data	0.74	0.62	1.00	0.63	0.75	0.54
Basic first stage - Revenue data	0.50	0.99	0.63	1.00	0.52	0.86
No first stage - Quantity data	0.76	0.51	0.75	0.52	1.00	0.46
No first stage - Revenue data	0.52	0.86	0.54	0.86	0.46	1.00

Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage. All markups are expressed in log. Data for 2009-2019 from EAP-FARE.

example, the correlation between markups estimated using the full first stage and no first stage is 0.67 for quantity-based markups, and 0.74 for revenue-based markups. The correlation between markups from full-first-stage quantity and basic-first-stage estimations is now 0.60.

We next calculate Spearman rank correlations, in the middle panel of Table 10. This tests whether alternative production function estimations preserve the rank of the markup estimates. Without exception, the Spearman correlations exceed the Pearson correlations. The correlation between quantity-based full- and basic-first-stage markup estimates rises to 0.71, while the revenue-quantity markup correlation rises to 0.33. This means that analyses relying more on markup rank than on dispersion are more likely to be robust to flaws in the production function estimation. When we additionally control for sector fixed effects, the revenue-quantity correlation rises to 0.50. While that is lower than correlations within the same output variable, it still suggests that revenue-based markups are informative about true markups.

A further illustration of the clear relationship between quantity and revenue-based markup estimates is provided in Figure 5. It contains a binned scatter plot that relates these in log-levels (left) and log-differences (right). Both show an excellent linear fit between both series, with the linear fit approaching a 45-degree line when markups are analyzed in first-differences. Table 11 expands on this by showing that there is generally a high correlation between markups from various specifications. The table finds a strong correlation in first differences between translog

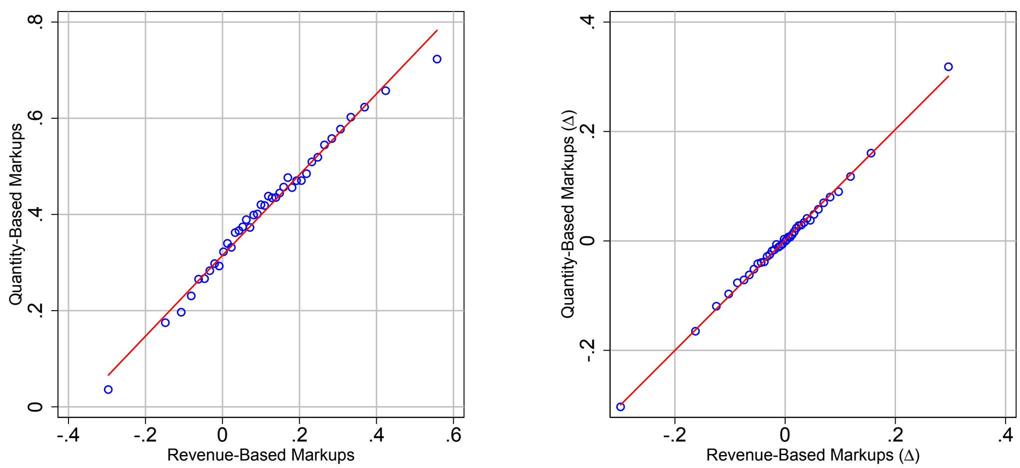
Table 11: Correlations across Specifications for Log-Differenced Markups

	Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
			Pearson Correlations			
Full first stage - Quantity data	1.00	0.44	0.51	0.45	0.43	0.42
Full first stage - Revenue data	0.44	1.00	0.65	0.99	0.52	0.78
Basic first stage - Quantity data	0.51	0.65	1.00	0.68	0.58	0.53
Basic first stage - Revenue data	0.45	0.99	0.68	1.00	0.54	0.77
No first stage - Quantity data	0.43	0.52	0.58	0.54	1.00	0.41
No first stage - Revenue data	0.42	0.78	0.53	0.77	0.41	1.00
			Spearman Rank Correlations			
Full first stage - Quantity data	1.00	0.68	0.81	0.68	0.77	0.71
Full first stage - Revenue data	0.68	1.00	0.81	0.99	0.76	0.91
Basic first stage - Quantity data	0.81	0.81	1.00	0.82	0.85	0.77
Basic first stage - Revenue data	0.68	0.99	0.82	1.00	0.78	0.91
No first stage - Quantity data	0.77	0.76	0.85	0.78	1.00	0.72
No first stage - Revenue data	0.71	0.91	0.77	0.91	0.72	1.00

Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage. All markups are expressed in log difference. Data for 2009-2019 from EAP-FARE.

markups estimated with revenue and quantity-based markups, with the preferred specifications displaying a Pearson and Spearman correlation of 0.44 and 0.68, respectively.[43]

Figure 5: Binned Scatter Plot for Relationship between Quantity and Revenue-Based Markups

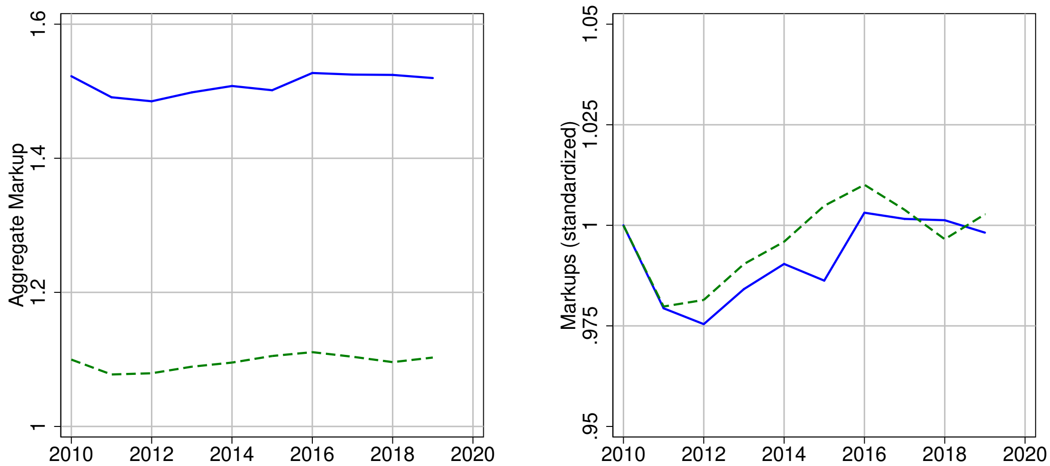


(a) Log-Markups (b) Log-Differenced Markups

Notes: The figures plot the binned scatter plot between quantity-based markups (full first stage) and revenue-based markups (basic first stage). Log-markups are used in figure (a), log-diff. markups in figure (b). Regression coefficients for the linear fit are 0.80 and 0.97, respectively.

[43] The first-differenced correlations are unaffected by sector fixed effects.

Figure 6: Aggregate Markups



(a) Aggregate Markup - Level

(b) Aggregate Markup - Index

Notes: The figures plot the aggregate markup based on quantity data (blue-solid) and revenue data (green-dashed). The left figure plots the level of both estimates, while the right figure is an index where the aggregate markup in each year is divided by the level in 2010. Aggregate markups are the harmonic average of firm-level markups, weighted by sales share.

6.2.3 Aggregate Markups

In the next analysis, we scrutinize the implications of our firm-level markup estimates for the behavior of the aggregate markup. We define the aggregate markup as the sales-weighted harmonic average of the firm-level markup:

$$\mathcal{M}_t \equiv \left(\sum_{i\in I_t} s_{it}\mu_{it}^{-1}\right)^{-1},$$

where I_t is the set of firms in the data at time t, while s_{it} denotes firm i's share in aggregate sales at time t. In a broad set of models, this harmonic average determines the distortion in factor prices from market power (Grassi 2017, Edmond et al. 2018, Burstein et al. 2020).

We calculate the aggregate markup for the quantity-based markups, using the full first stage, as well as the revenue-based markups, using the basic first stage. Figure 6 plots the results. The left figure, which plots $\mathcal{M}_t$, confirm that the revenue-based markup estimates are sizably smaller than the quantity-based markups. Aggregate revenue-based markups average around 1.1, while aggregate quantity-based markups average around 1.5. The right figure, however, which plots $\mathcal{M}_t/\mathcal{M}_{2010}$, again shows that the revenue-based markup estimates preserve useful information about quantity markups. The figure shows that adjusted for the average, revenue-based markups follow similar aggregate movements as quantity-based markups. Both of the estimates show a decline around the Eurocrisis in 2011 and 2012, followed by an upward trend from 2013. In Appendix Figure A1, we show that aggregate markup estimates from revenue and quantity data also track each other closely at the sector level.

Table 12: Relation between Markup and Explanatory Variables by Markup Specification

	Quantity			Revenue		
	Full	Basic	None	Full	Basic	None
Profit rate χ	0.0514***	0.0937***	0.0764***	0.0944***	0.0972***	0.0662***
	(0.00194)	(0.00338)	(0.00286)	(0.00264)	(0.00254)	(0.00279)
R-squared	0.091	0.239	0.155	0.447	0.456	0.236
Labor share χ	-0.0145***	-0.0316***	-0.0346***	-0.0291***	-0.0286***	-0.0253***
	(0.000921)	(0.00123)	(0.00130)	(0.000701)	(0.000646)	(0.00111)
R-squared	0.021	0.068	0.078	0.104	0.096	0.084
Material share χ	-0.0298***	-0.0543***	-0.0314***	-0.0602***	-0.0636***	-0.0393***
	(0.00142)	(0.00251)	(0.00167)	(0.00212)	(0.00212)	(0.00184)
R-squared	0.071	0.181	0.063	0.408	0.436	0.187
Market share (%) χ	0.0134***	0.0282***	0.0330***	0.0234***	0.0217***	0.0193***
	(0.00259)	(0.00368)	(0.00327)	(0.00347)	(0.00325)	(0.00268)
R-squared	0.005	0.006	0.007	0.006	0.006	0.006

Note: Each entry gives the OLS estimate using the cursive variable as the dependent variable and the markup series in the column header as the regressor. Markup estimates use a first-stage regression with a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time effects ("Basic") or no first stage ("None"). Markups are expressed in logs and standardized to have a unit standard deviation. Markups are winsorized at 1.5% tails. Firm-clustered standard errors in parentheses. **, *** denote significance at the 5, 1% level. Data for 2009-2019 from EAP-FARE. All regressions include time & firm fixed effects. Observations: 157,277.

6.2.4 Correlations of Markup Estimates and Other Variables

Next we assess whether relationships between markups and key variables depend on the markup specification. To do so, we estimate the same regressions as on the simulated data along (17). We estimate this regression using a firm's profit rate (defined as the ratio of operating profits over sales), labor share (defined as the ratio of its wagebill over sales), material cost share (defined as the ratio of materials purchased over sales), and market share (defined as its share in revenue at the 5-digit sector level) as dependent variables. Our aim is not to causally estimate the relationship between these variables and markups, but rather to see how the correlation between these variables and markups depends on how the markup was estimated.

Results are presented in Table 12. Each row presents regression coefficients for a particular explanatory variable (described in italics), while each column contains results for a specific markup specification. Before describing differences across specifications, note that all relationships in the table run in the expected direction. Firms with higher markup estimates are more profitable, have lower labor shares, lower material shares, and greater market shares. This is the case irrespective of whether revenue or quantity data was used to estimate the production function elasticities, and the relationships are all significant at the 1% level. Looking more carefully at the specifications, we see that estimated βs do differ across specifications, both when changes are made to the first stage or when quantity or revenue is used. The estimated βs tend to be smaller for quantity-based markups than for revenue-based markups. This is in line with the finding in Table 9 that there is more dispersion in the quantity-based markup estimates; a higher variance of the markup mechanically reduces the estimated βs holding everything else equal. Overall, however, the results in Table 12 suggest that relationships between markups and key relationships are qualitatively robust to using imperfect first-stage regressions or revenue-based markup estimates. This further supports our derivation in Section 3 that these estimates contain useful information about a firm's true markup.

7 Alternative: assuming a Cobb-Douglas production function

In our final analysis, we discuss the validity of a common remedy to difficulties in markup estimation: assuming a Cobb-Douglas production function. Because all firms in an industry share the same output elasticities, markups under Cobb-Douglas are pinned down by the inverse revenue share of v_{it} up to a sector fixed effect. This makes the markup estimates robust to biases in production function estimation. Actual production functions are likely not of Cobb-Douglas form, however, as evidenced by the variation in within-sector output elasticities (see Table 8). Below we study the implications of assuming a Cobb-Douglas production function in that case.

7.1 Analytical results

To structure the analysis, we return to the analytical framework in Section 3, where output is produced with the single (variable) input v_{it}. In line with our simulations and empirical analysis we assume that the true production function is translog, i.e., $y_{it} = \alpha v_{it} + \beta v_{it}^2$. Under these assumptions, the IV-GMM estimator from Definition 1 is misspecified. A researcher that estimates a Cobb-Douglas production function would erroneously assume that the output elasticity of v_{it} does not vary with v_{it}. The misspecified GMM estimator is:

Definition 3 *(Misspecified GMM estimator) the GMM estimator is $\widehat{\alpha} \in \mathbb{R}$ such that the moment $\mathbb{E}\left[\widehat{\omega}_{it}^{Mis} v_{it-1}\right]$ is equal to zero, where $\widehat{\omega}_{it}^{Mis} = y_{it} - \widehat{\alpha} v_{it} = (\alpha - \widehat{\alpha}) v_{it} + \beta v_{it}^2 + \omega_{it}$.*

The productivity estimate that is used in the moment condition, $\widehat{\omega}_{it}^{Mis}$, is the sum of true productivity and an additional term that differs from zero under translog ($\beta \neq 0$). The unique solution to the estimator – as long as v_{it-1} is indeed a relevant instrument for v_{it} – is given by

$$\widehat{\alpha} = \alpha + \beta \frac{\mathbb{E}[v_{it}^2 v_{it-1}]}{\mathbb{E}[v_{it} v_{it-1}]}.$$

It follows that the Cobb-Douglas estimate of the output elasticity suffers from a bias from the omitted variable v_{it}^2. This bias contaminates the resulting markup estimates $\widehat{\mu}_{it}^{Miss} \equiv$

Table 13: Estimated Cobb-Douglas Material-Output Elasticity by Sector and Specification

NACE	Quantity			Revenue		
	Full	Basic	None	Full	Basic	None
All (average)	0.50	0.46	0.46	0.44	0.70	0.34
13	0.28	0.13	0.28	0.14	0.14	0.25
14	0.31	0.20	0.30	0.34	0.02	0.02
15	0.25	0.20	0.22	0.09	0.09	0.11
16	0.56	0.50	0.55	0.51	0.51	0.46
17	0.41	0.41	0.42	0.54	0.55	0.36
18	0.30	0.29	0.29	1.04	0.02	0.22
20	0.36	0.19	0.36	0.65	0.64	0.32
22	0.57	0.54	0.56	0.44	3.38	0.43
23	0.50	0.45	0.51	0.38	0.38	0.37
24	0.65	0.76	0.65	0.60	0.60	0.44
25	0.41	0.42	0.40	0.38	0.38	0.35
26	0.64	1.00	0.61	0.29	0.29	0.29
27	0.44	0.45	0.44	0.72	0.73	0.38
28	0.28	0.66	0.33	0.35	0.35	0.35
29	1.42	0.74	0.22	0.50	0.50	0.45
31	1.08	0.66	1.04	0.39	0.39	0.36
32	0.38	0.31	0.44	0.17	0.17	0.23
33	0.33	0.35	0.33	0.31	0.31	0.30

Note: The table presents estimated output elasticities of materials from the estimation of Cobb-Douglas production functions. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects (“Full”), only the polynomial and time fixed effects (“Basic”) or no first stage ("None"). Translog elasticities are given in Table 8.

$\widehat{\alpha}(P_{it}Y_{it})/(W_{it}V_{it})$. The true markup is $\mu_{it} = (\alpha + 2\beta v_{it})\,(P_{it}Y_{it})/(W_{it}V_{it})$, which means that

$$\frac{\widehat{\mu}_{it}^{Miss}}{\mu_{it}} = \frac{\alpha + \beta \frac{\mathbb{E}[v_{it}^2 v_{it-1}]}{\mathbb{E}[v_{it} v_{it-1}]}}{\alpha + 2\beta v_{it}}. \tag{19}$$

As the true output elasticity of v_{it} varies under a translog production function, the correlation between true markups and the mismeasured Cobb-Douglas markups will not equal one. Taking logs and averages, we can see that the Cobb-Douglas markup estimates do preserve information about the average of true markups. The average (log) estimated markup is equal to the average true markup up to a Jensen's inequality:

$$\mathbb{E}\left[\log(\widehat{\mu}_{it}^{Miss})\right] = \mathbb{E}\left[\log \mu_{it}\right] + \log\left(\alpha + \beta \mathbb{E}\left[v_{it}\frac{v_{it}v_{it-1}}{\mathbb{E}[v_{it}v_{it-1}]}\right]\right) - \mathbb{E}\left[\log(\alpha + 2\beta v_{it})\right].$$

The variance of the misspecified markup is different from the true markup variance for two reasons. First, the output elasticity is not constant across firms under a translog production function. Second, the covariance between markup and the output elasticity can be non-zero:

$$\mathbb{V}ar\left[\log(\widehat{\mu}_{it}^{Miss})\right] = \mathbb{V}ar\left[\log \mu_{it}\right] + \mathbb{V}ar\left[\log(\alpha + 2\beta v_{it})\right] - 2\mathbb{C}ov\left[\log \mu_{it}, \log(\alpha + 2\beta v_{it})\right]. \tag{20}$$

The variance of the estimated markups can therefore be larger or smaller than the true variance.

Table 14: Overview - Log Markup Estimates from Cobb-Douglas Production Functions

	Mean	St. Dev.	Median	25th Pct.	75th Pct.	Observations
Quantity data						
Full first stage	0.25	0.57	0.19	-0.16	0.62	157,485
Basic first stage	0.17	0.60	0.20	-0.15	0.55	157,485
No first stage	0.19	0.56	0.15	-0.19	0.54	157,485
Revenue data						
Full first stage	0.13	0.62	0.08	-0.23	0.46	157,485
Basic first stage	0.08	1.10	0.08	-0.32	0.54	157,485
No first stage	-0.11	0.62	-0.09	-0.37	0.22	157,485

Note: All markups are expressed in log. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full") or only the polynomial and fixed effects ("Basic"). Data for 2009-2019 from EAP-FARE.

7.2 Empirical results

Turning to the empirical analysis, we estimate a Cobb-Douglas production function on the production data. Rather than the full translog production function (18), we now estimate

$$y_{it} = \beta_m^{cd} m_{it} + \beta_l^{cd} l_{it} + \beta_k^{cd} k_{it} + \beta_o^{cd} o_{it}$$

by industry for the six specifications described above. Cobb-Douglas markups are then given by the product of the estimated β_m^{cd} and the inverse revenue share of materials.

7.2.1 Production function and markup estimates

The estimates for elasticities β_m^{cd} are listed in Table 13. The average elasticities range from 0.34 to 0.70. As expected from the analytical derivations in Section 7.1, this is similar to the range of average elasticities under translog (0.42 to 0.62). Compared to elasticities under translog, the main difference is that the Cobb-Douglas elasticities are constant within sectors. While the Cobb-Douglas elasticities do not match the average translog elasticities of each sector precisely, the rank of the estimates is well-maintained. Notice, however, that the within-sector variation in output elasticities is substantial under translog. This means that non-linear production function terms differ from zero; a Cobb-Douglas production function is misspecified.

Turning to the markup estimates, Table 14 provides summary statistics. These differ from their translog counterparts (Table 9) in two ways. First, averages are somewhat lower, in line with the lower output elasticities. Second, and perhaps more importantly, they differ strongly in standard deviations and 75th/25th percentiles. For each specification, Cobb-Douglas markups have higher standard deviations. The 25th percentile of log markups is also negative in all specifications. If these estimates are correct, a significant portion of firms is selling below marginal costs. Most specifications also show higher 75th percentiles for markups. Overall, it is clear that

Table 15: Correlations across Specifications - Cobb-Douglas Markups

	Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
			Pearson Correlations			
Full first stage - Quantity data	1.00	0.59	0.81	0.41	0.82	0.66
Full first stage - Revenue data	0.59	1.00	0.58	0.18	0.59	0.58
Basic first stage - Quantity data	0.81	0.58	1.00	0.43	0.77	0.72
Basic first stage - Revenue data	0.41	0.18	0.43	1.00	0.42	0.62
No first stage - Quantity data	0.82	0.59	0.77	0.42	1.00	0.67
No first stage - Revenue data	0.66	0.58	0.72	0.62	0.67	1.00
			Spearman Rank Correlations			
Full first stage - Quantity data	1.00	0.63	0.83	0.53	0.85	0.81
Full first stage - Revenue data	0.63	1.00	0.58	0.53	0.63	0.75
Basic first stage - Quantity data	0.83	0.58	1.00	0.50	0.79	0.82
Basic first stage - Revenue data	0.53	0.53	0.50	1.00	0.55	0.70
No first stage - Quantity data	0.85	0.63	0.79	0.55	1.00	0.82
No first stage - Revenue data	0.81	0.75	0.82	0.70	0.82	1.00

Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage. All markups are expressed in log. Data for 2009-2019 from EAP-FARE.

Cobb-Douglas estimates are more dispersed in the French data.

To see what drives differences in dispersion, note that all variation in Cobb-Douglas markups within sectors comes from the inverse material revenue share. The variance of log-markups is

$$\mathbb{V}ar\left(\log \widehat{\mu}_{it}^{cd}\right) = \mathbb{V}ar\left(p_{it} + y_{it} - p_t^m - m_{it}\right).$$

Conversely, because the translog production function admits firm-specific elasticities, we have

$$\mathbb{V}ar(\log \widehat{\mu}_{it}^{tl}) = \mathbb{V}ar\left(\log \widehat{\alpha}_{it}^m\right) + \mathbb{V}ar\left(p_{it} + y_{it} - p_t^m - m_{it}\right) + 2 \cdot \mathbb{C}ov\left(\log \widehat{\alpha}_{it}^m,\ p_{it} + y_{it} - p_t^m - m_{it}\right).$$

Hence, the lower dispersion of markups in the translog specification implies that there is a negative correlation between a firm's revenue-over-materials share and the elasticity of its output with respect to materials. In other words, firms that have relatively high revenue compared to their spending on material inputs, on average have lower output elasticities of materials.

7.2.2 Markup Correlations

The correlation between Cobb-Douglas markups from various specifications is presented in Table 15. The top panel presents Pearson correlations while the bottom panel presents the rank correlations. The table generally shows high correlations across specifications. With few exceptions, correlations exceed 0.5. Correlations are particularly high for markups based on the same output variable. For quantity, for example, the correlation between markups estimated using the full first stage and no first stage is 0.82. Correlations across output variables are lower.

Table 16: Correlations between Translog and Cobb-Douglas-based Markup Estimates

		Translog Markup Estimates					
		Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
		Pearson Correlations					
Cobb-Douglas	Full first stage - Quantity data	0.27	0.53	0.44	0.55	0.30	0.32
	Full first stage - Revenue data	0.11	0.54	0.32	0.57	0.13	0.32
	Basic first stage - Quantity data	0.21	0.48	0.31	0.51	0.14	0.28
	Basic first stage - Revenue data	0.05	0.21	0.13	0.21	0.09	0.11
	No first stage - Revenue data	0.34	0.55	0.33	0.57	0.38	0.34
	No first stage - Revenue data	0.04	0.48	0.18	0.50	0.04	0.27
		Pearson - first differences					
	Δ Cobb Douglas (any)	0.28	0.59	0.48	0.60	0.34	0.44
		Spearman Rank Correlations					
Cobb-Douglas	Full first stage - Quantity data	0.28	0.59	0.48	0.60	0.34	0.44
	Full first stage - Revenue data	0.20	0.61	0.44	0.64	0.23	0.46
	Basic first stage - Quantity data	0.31	0.54	0.41	0.57	0.25	0.40
	Basic first stage - Revenue data	0.14	0.36	0.28	0.37	0.20	0.25
	No first stage - Quantity data	0.36	0.62	0.38	0.63	0.44	0.47
	No first stage - Revenue data	0.15	0.68	0.32	0.69	0.17	0.52
		Spearman Rank - first differences					
	Δ Cobb Douglas (any)	0.53	0.69	0.63	0.71	0.53	0.61

Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects (“Full”), only the polynomial and time fixed effects (“Basic”) or no first stage. All markups are expressed in log. Data for 2009-2019 from EAP-FARE.

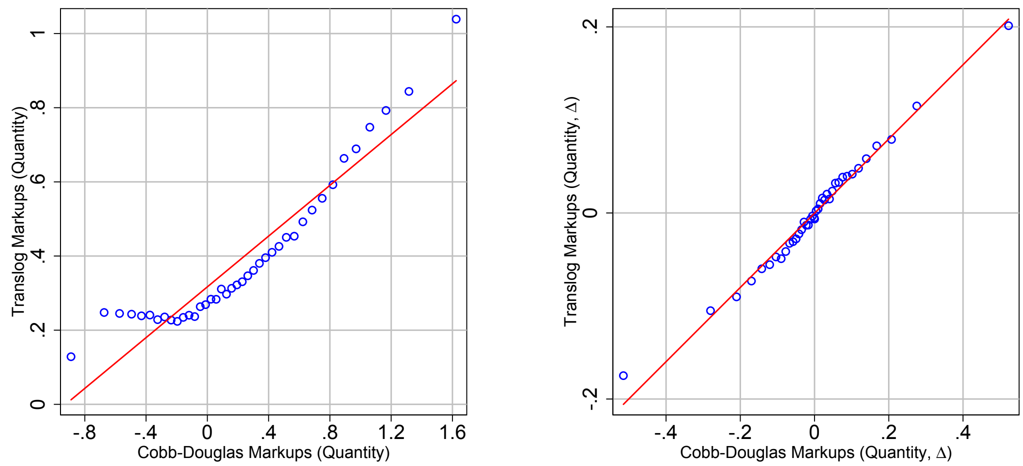
The correlation between markups using the full first stage from quantity and revenue data, for example, is 0.59. While that is lower than correlations within the same output variable, it still shows that revenue-based markups are informative about true markups. Note that in first differences, correlations across Cobb-Douglas markup estimates equal 1.

Table 16 shows that, while Cobb-Douglas markup estimates correlate well across specifications, the correlation between Cobb-Douglas and translog-based estimates of the markup have lower correlations. Correlations between Cobb-Douglas estimates from revenue data and translog estimates on quantity data appear almost orthogonal. The binned scatter plot for the preferred specification, with quantity data and the full first stage, confirm the poorer fit (Figure 7). While correlations are higher when markups are studied in first-differences, they remain considerably lower than the correlations across the various translog specifications.

7.3 Application: Costs of Markup Dispersion

The overstatement of markup dispersion in the Cobb-Douglas specifications is particularly important because researchers frequently assume a Cobb-Douglas production function to ‘prevent’ having to estimate a production function: when studying log markups, by taking industry fixed effects to correct for output elasticities researchers can analyze within-industry markup dispersion simply from observing revenue and material expenditures. Our translog results

Figure 7: Binned Scatter Plot for Relationship between Translog and Cobb-Douglas Markups



(a) Log-Markups

(b) Log-Differenced Markups

Notes: The figures plot the linear relationship and binned scatter plot between translog markups and Cobb-Douglas markups (both full first stage) Log-markups are used in figure (a), log-differenced markups in figure (b). Regression coefficients for the linear fit are 0.32 and 0.38, respectively.

however, suggest that the Cobb-Douglas assumption is rejected in the French data. This is important for a number of applications - for example when analyzing the effect of heterogeneous markups on allocative efficiency and productivity. The idea is that firms with higher markups produce inefficiently little because they raise prices above marginal costs. As the Cobb-Douglas production function estimates overstate the degree of markup dispersion, a researcher relying on these estimates would therefore overstate the degree of misallocation in the economy.[44]

By how much a researcher would overstate the costs of markups when using Cobb-Douglas estimates would depend on assumptions about the demand system. As a back-of-the-envelope exercise, we perform the simple misallocation cost calculation in Peters (2020) for the case of

Table 17: Costs of Markup Dispersion by Production Function Estimate

	St. Dev. of Log Markups	Dispersion costs
Translog - Quantity	0.43	8.57%
Translog - Revenue	0.16	1.30%
Cobb-Douglas - Quantity	0.56	13.5%
Cobb-Douglas - Revenue	0.80	21.2%

Notes: Example calculation of how the costs of markup dispersion change when using alternative production function estimates using the formula in Peters (2020). Data for the EAP-FARE sample (2009-2019). Dispersion costs are expressed as $(1-\widetilde{\mathcal{M}})\cdot 100\%$. Markups from quantity data are estimated using the full first stage that includes price as a control; markups from revenue data are estimated using the basic first stage that does not include price as a control. All markups are trimmed.

[44] A researcher could equally *underestimate* the costs of markup dispersion in case the ratio of sales over material spending is positively correlated with the output elasticity of materials. The point here is that the Cobb-Douglas production function assumption is not without loss of generality.

innovation-driven markup dispersion with a Cobb-Douglas aggregator. He shows that the ratio of true aggregate productivity and productivity under allocative efficiency is given by

$$\widetilde{\mathcal{M}} = \frac{\exp\left(\sum_{i\in I} \ln \mu_i^{-1}\right)}{\sum_{i\in I} \mu_i^{-1}},$$

where I denotes the set of all firms. The denominator and numerator are equal if firms have homogeneous markups, such that aggregate productivity equals its efficient benchmark. Table 17 presents the results. Assuming that the translog quantity markups are correct, the true reduction in aggregate total factor productivity because of markup dispersion is 8.57%. A researcher that would use Cobb-Douglas estimates from revenue data would measure the cost of dispersion at 21.2% – an overstatement of 140%.

8 Conclusion

This paper provides an assessment of the validity of the ratio estimator of firm-level markups. We start by deriving the conditions under which the commonly used two-stage iterative GMM estimator is able to consistently estimate the parameters of the production function. Using an analytical framework, we assess the feasibility of estimating markups from accounting data, and derive the biases from not observing prices and markups when estimating a production function. We confirm the insights we glean with simulations from a rich macro model and empirical data on prices and production for French manufacturing firms.

We find that the use of revenue rather than quantity data to estimate production functions affects the level of the estimated markups, but has only modest effects on dispersion. The correlation between markups from quantity and revenue data ranges from 0.3 to 0.5 in log-levels and 0.7 in log-differences. The correlation between various markup estimates and variables such as market share, profitability and the labor share is also similar across the use of revenue or quantity data. We find significant improvement in estimates of the markups when production functions are estimated with a first-stage purging regression, instead of with simple IV-GMM.

Practically, we conclude that if a researcher is faced with imperfect data, then it depends on individual applications whether the analysis can proceed. Optimally, production functions for markups should be estimated with quantity rather than revenue data. In the absence of data on prices however, researchers that are interested in the dispersion or correlations of markups should think twice before assuming a Cobb-Douglas production function. Conversely, in applications where researchers are interested in the average level of the markup, revenue data may not be appropriate. Revenue data may be used to estimate trends of markups (as differences over time are a part of dispersion), provided the researcher is willing to assume that the production function parameters do not change over time.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Ackerberg, Daniel, Garth Frazer, Yao Luo, Yingjun Su et al., "Under-Identification of Structural Models Based on Timing and Information Set Assumptions," *Working Paper*, 2020.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li, "A Theory of Falling Growth and Rising Rents," *Federal Reserve Bank of San Francisco Working Paper*, 2019.

Anderson, Eric, Sergio Rebelo, and Arlene Wong, "Markups Across Space and Time," *National Bureau of Economic Research Working Paper*, 2018.

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, December 2008, *98* (5), 1998–2031.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," *Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Baqaee, David Rezza and Emmanuel Farhi, "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," *Econometrica*, July 2019, *87* (4), 1155–1203.

Barkai, Simcha, "Declining Labor and Capital Shares," *The Journal of Finance*, 2020, *75* (5), 2421–2463.

Basu, Susanto, "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence," *Journal of Economic Perspectives*, August 2019, *33* (3), 3–22.

_ **and John G Fernald**, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, April 1997, *105* (2), 249–83.

Blundell, Richard and Stephen Bond, "GMM Estimation with Persistent Panel Data: an Application to Production Functions," *Econometric reviews*, 2000, *19* (3), 321–340.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch, "Some Unpleasant Markup Arithmetic: Production Function Elasticities and their Estimation from Production Data," *Journal of Monetary Economics*, 2021, *121*, 1–14.

Bornstein, Gideon, "Entry and Profits in an Aging Economy: The Role of Consumer Inertia," *Working Paper*, 2018.

Brand, James, "Estimating Productivity and Markups Under Imperfect Competition.," *Working Paper*, 2019.

Burstein, Ariel, Vasco M. Carvalho, and Basile Grassi, "Bottom-up Markup Fluctuations," NBER Working Papers 27958, National Bureau of Economic Research, Inc October 2020.

Calligaris, Sara, Chiara Criscuolo, and Luca Marcolin, "Mark-ups in the Digital Era," *OECD Working Paper*, 2018.

Caselli, Mauro, Lionel Nesta, and Stefano Schiavo, "Imports and Labour Market Imperfections: Firm-Level Evidence from France," *European Economic Review*, 2021, *131*, 103632.

Cavenaile, Laurent, Murat Alp Celik, and Xu Tian, "Are Markups too High? Competition, Strategic Innovation, and Industry Dynamics," *Working Paper*, 2019.

Collard-Wexler, Allan and Jan De Loecker, "Production Function Estimation and Capital Measurement Error," *National Bureau of Economic Research Working Paper*, 2020.

Crouzet, Nicolas and Janice C Eberly, "Understanding weak capital investment: The role of market concentration and intangibles," Technical Report, National Bureau of Economic Research 2019.

De Loecker, Jan and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, May 2012, *102* (6), 2437–71.

De Ridder, M., "Market Power and Innovation in the Intangible Economy," Cambridge Working Papers in Economics 1931, Faculty of Economics, University of Cambridge March 2019.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "Changing business dynamism and productivity: Shocks versus responsiveness," *American Economic Review*, 2020, *110* (12), 3952–90.

Doraszelski, Ulrich and Jordi Jaumandreu, "Using Cost Minimization to Estimate Markups," 2019.

_ and _ , "The Inconsistency of De Loecker and Warzynski's (2012) Method to Estimate Markups and Some Robust Alternatives," 2020.

Díez, Federico, Jiayue Fan, and Carolina Villegas-Sanchez, "Global Declining Competition," CEPR Discussion Papers 13696, C.E.P.R. Discussion Papers April 2019.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How Costly Are Markups?," Working Paper 24800, National Bureau of Economic Research July 2018.

Eggertsson, Gauti B, Jacob A Robbins, and Ella Getz Wold, "Kaldor and Piketty's facts: The Rise of Monopoly Power in the United States," *Journal of Monetary Economics*, 2021, *124*, S19–S38.

Gandhi, Amit, Salvador Navarro, and David A Rivers, "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

Grassi, Basile, "IO in I-O: Size, Industrial Organization and the Input-Output Network Make a Firm Structurally Important," *Working Paper*, 2017.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, "Are US industries becoming more concentrated?," *Review of Finance*, 2019, *23* (4), 697–743.

Gutiérrez, Germán and Sophie Piton, "Revisiting the Global Decline of the (Non-Housing) Labor Share," *American Economic Review: Insights*, 2020, *2* (3), 321–38.

Gutiérrez, Germán and Thomas Philippon, "Declining Competition and Investment in the U.S," NBER Working Papers 23583, National Bureau of Economic Research, Inc July 2017.

_ and _ , "How EU Markets Became More Competitive Than US Markets: A Study of Institutional Drift," *Forthcoming in the Journal of the European Economics Association*, 2022.

Hall, Robert E, "Market Structure and Macroeconomic Fluctuations," *Brookings papers on economic activity*, 1986, *1986* (2), 285–338.

Hall, Robert E., "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 1988, *96* (5), 921–947.

Hershbein, Brad, Claudia Macaluso, and Chen Yeh, "Monopsony in the U.S. Labor Market," *Working Paper*, 2021.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Kehrig, Matthias and Nicolas Vincent, "The Micro-Level Anatomy of the Labor Share Decline," *The Quarterly Journal of Economics*, 2021, *136* (2), 1031–1087.

Kimball, Miles S., "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 1995, *27* (4), 1241–1277.

Klenow, Peter J. and Jonathan L. Willis, "Real Rigidities and Nominal Price Changes," *Economica*, July 2016, *83* (331), 443–472.

Klette, Tor Jakob and Zvi Griliches, "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous," *Journal of Applied Econometrics*, 1996, *11* (4), 343–361.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, 01 2020, *135* (2), 561–644.

_ , Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups, and Trade Reform," *Econometrica*, 2016, *84* (2), 445–510.

Meier, Matthias and Timo Reinelt, "Monetary Policy, Markup Dispersion, and Aggregate TFP," *Working Paper*, 2020.

Morlacco, Monica, "Market Power in Input Markets: Theory and Evidence from French Manufacturing," *Working Paper*, 2019.

Neiman, Brent and Joseph S Vavra, "The Rise of Niche Consumption," *Forthcoming in American Economic Journal: Macroeconomics*, 2021.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, *64* (6), 1263–97.

Pasqualini, Andrea, "Markups, Markdowns and Bankruptcy in the Banking Industry," *Working Paper*, 2021.

Peters, Michael, "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 2020, *88* (5), 2037–2073.

Syverson, Chad, "Macroeconomics and Market Power: Context, Implications, and Open Questions," *Journal of Economic Perspectives*, August 2019, *33* (3), 23–43.

Traina, James, "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," Working Paper 17, Stigler Cente February 2018.

Van Vlokhoven, Has, "Estimating the Cost of Capital and the Profit Share," *Working Paper*, 2019.

A Theory Appendix

A.1 Finite Sample Estimator and its Asymptotic Variance

In this section we derive the estimator for a finite sample. We also use this derivation to compute the asymptotic variance of the GMM estimator. First, let us define the estimator for a finite sample.

Definition: The GMM estimator is $\widehat{\alpha}$ such that $\sum_{i,t} \widehat{\omega}_{it} v_{it-1} = 0$ with $\widehat{\omega}_{it} = y_{it} - \widehat{\alpha} v_{it} = (\alpha - \widehat{\alpha}) v_{it} + \omega_{it}$.

Second, to solve for the estimator, we need to find the value of $\widehat{\alpha}$ such that $\sum_{i,t} \widehat{\omega}_{it} v_{it-1} = (\alpha - \widehat{\alpha}) \sum_{i,t} v_{it} v_{it-1} + \sum_{i,t} \omega_{it} v_{it-1} = 0$. As long as $\sum_{i,t} v_{it} v_{it-1} \neq 0$, the unique $\widehat{\alpha}$ that solves this equation is

$$\widehat{\alpha} = \alpha + \frac{\sum_{i,t} \omega_{it} v_{it-1}}{\sum_{i,t} v_{it} v_{it-1}}$$

whose limit is α when the sample size increases, given that $\mathbb{E}[\omega_{it} v_{it-1}] = 0$.

Finally, let us derive the asymptotic variance of the GMM estimator. Using the (finite sample) expression of the estimator, we have

$$\sqrt{n}(\widehat{\alpha} - \alpha) = \frac{\sqrt{n}\frac{1}{n} \sum_{i,t} \omega_{it} v_{it-1}}{\frac{1}{n} \sum_{i,t} v_{it} v_{it-1}}.$$

By the (weak) law of large numbers, $\frac{1}{n} \sum_{i,t} v_{it} v_{it-1} \xrightarrow{p} \mathbb{E}[v_{it} v_{it-1}]$, and, by the central limit theorem, $\sqrt{n}\frac{1}{n} \sum_{i,t} \omega_{it} v_{it-1} \xrightarrow{d} \mathcal{N}\left(0, \mathbb{E}\left[\omega_{it}^2 v_{it-1}^2\right]\right)$. The Slutsky theorem implies $\sqrt{n}(\widehat{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}\left(0, \frac{\mathbb{E}\left[\omega_{it}^2 v_{it-1}^2\right]}{\mathbb{E}[v_{it} v_{it-1}]^2}\right)$; that is,

$$\mathbb{V}ar\left[\widehat{\alpha}\right] \sim \frac{\mathbb{E}\left[\omega_{it}^2\right] \mathbb{E}\left[v_{it-1}^2\right]}{\sqrt{n}\mathbb{E}[v_{it} v_{it-1}]^2}.$$

A.2 With measurement errors

As in the baseline framework, assume that firms produce y_{it} using the single variable input v_{it} while being subject to idiosyncratic productivity shocks ω_{it}. Furthermore, assume that the firms' output is observed subject to measurement error, or equivalently, that unexpected productivity shocks that occur after input v_{it} is set. The measurement error is log-additive and denoted by η_{it}. All firms produce along

$$y_{it} = \alpha v_{it} + \omega_{it} + \eta_{it},$$

where y_{it} denotes observed output or output inclusive of the unexpected productivity shocks. We assume that measurement errors at time t are independent of the past value of the variable input; that is, $\mathbb{E}\left[\eta_{it}v_{it-1}\right]=0$. If the econometrician ignores the presence of these measurement errors, the GMM estimator is defined as follows:

Definition 4 *The GMM estimator is* $\widehat{\alpha}\in\mathbb{R}$ *such that the moment* $\mathbb{E}\left[(\widehat{\omega_{it}+\eta_{it}})v_{it-1}\right]$ *is equal to zero where* $\widehat{\omega_{it}+\eta_{it}}=y_{it}-\widehat{\alpha}v_{it}=(\alpha-\widehat{\alpha})v_{it}+\omega_{it}+\eta_{it}$.

The GMM estimator is characterized by:

$$\mathbb{E}\left[(\widehat{\omega_{it}+\eta_{it}})v_{it-1}\right]=(\alpha-\widehat{\alpha})\mathbb{E}[v_{it}v_{it-1}]=0,$$

where we use the fact that $\mathbb{E}\left[\widehat{\omega}_{it}v_{it-1}\right]=0$. As was the case in the baseline framework, the GMM estimator $\widehat{\alpha}$ of the variable input's output elasticity is equal to α as long as $\mathbb{E}[v_{it}v_{it-1}]\neq 0$. The estimator remains unbiased and identified as the additional measurement error only increases the variance of the composite error term $\omega_{it}+\eta_{it}$ in the production function. This point is known and has been discussed, for example, in Blundell and Bond (2000).

If the single-stage GMM estimator is consistent, then why bother purging output from measurement error? There are two advantages to purging. The first is that the increase in the variance of the composite error term $\omega_{it}+\eta_{it}$ in the production function raises the standard errors of the production function estimation. Indeed, a similar derivation to the one in Appendix A.1 yields that the asymptotic variance of estimator is

$$\mathbb{V}ar\left[\widehat{\alpha}\right]\sim\frac{\mathbb{E}\left[v_{it-1}^2\right]}{n\mathbb{E}[v_{it}v_{it-1}]^2}\left(\mathbb{E}\left[\omega_{it}^2\right]+\mathbb{E}\left[\eta_{it}^2\right]\right),$$

which increases in measurement error variance. The second advantage is that purging allows the econometrician to identify true productivity ω_{it}, which is relevant in many applications.[45]

A.3 Extensions

We now show that the identification results of our estimator is robust to several extensions that are common in practical applications. We study the case of the translog production function, the case of several inputs, and the case of AR(1) productivity. We discuss the case with all of these extensions together in the appendix.

A.3.1 Translog Production Function

We first ease the assumption that output is log-linear by replacing the Cobb-Douglas production function with a translog specification:

$$y_{it}=\alpha v_{it}+\beta v_{it}^2+\omega_{it}.$$

[45]A further benefit of purging output from measurement error is that it allows more sophisticated persistent productivity processes than the linear AR(1).

The other assumptions are unchanged. Our aim is to identify the parameters α and β, to be able to calculate the size-dependent output elasticity of the variable input for the calculation of the true markup $\mu_{it} = (\alpha + 2\beta v_{it})(P_{it}Y_{it})/(W_t V_{it})$. The least-squares estimation of the production function suffers from the same bias as before, which we address by instrumenting v_{it} and v_{it}^2 by their respective lags. Econometrically, estimating the more sophisticated translog production is therefore simply akin to estimating a multivariate GMM regression with instrumental variables. Formally, we define the estimator as:

Definition 5 *The GMM estimator is a pair* $(\widehat{\alpha}, \widehat{\beta})$ *such that* $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0$ *and* $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}^2\right] = 0$ *where* $\widehat{\omega}_{it} = y_{it} - \widehat{\alpha} v_{it} - \widehat{\beta} v_{it}^2 = (\alpha - \widehat{\alpha}) v_{it} + (\beta - \widehat{\beta}) v_{it}^2 + \omega_{it}$.

It is again straightforward to solve for the estimator $(\widehat{\alpha}, \widehat{\beta})$ in our parsimonious setting. It involves solving the system of linear equations implied by the moment conditions:

$$\begin{array}{l} \mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0 \\ \mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}^2\right] = 0 \end{array} \iff \begin{array}{l} (\alpha - \widehat{\alpha})\mathbb{E}[v_{it} v_{it-1}] + (\beta - \widehat{\beta})\mathbb{E}[v_{it}^2 v_{it-1}] = 0 \\ (\alpha - \widehat{\alpha})\mathbb{E}[v_{it} v_{it-1}^2] + (\beta - \widehat{\beta})\mathbb{E}[v_{it}^2 v_{it-1}^2] = 0 \end{array} .$$

This system can be rewritten in matrix form with $V(B - \widehat{B}) = 0$ where

$$B - \widehat{B} = \begin{pmatrix} \alpha - \widehat{\alpha} \\ \beta - \widehat{\beta} \end{pmatrix} \qquad \text{and} \qquad V = \begin{pmatrix} \mathbb{E}[v_{it} v_{it-1}] & \mathbb{E}[v_{it}^2 v_{it-1}] \\ \mathbb{E}[v_{it} v_{it-1}^2] & \mathbb{E}[v_{it}^2 v_{it-1}^2] \end{pmatrix}.$$

As long as the determinant of V is not zero, the GMM estimator on translog is identified and asymptotically consistent such that $\widehat{\alpha} = \alpha$ and $\widehat{\beta} = \beta$. This is the case as long as v_{it} and its square are not colinear and when the lagged values of v_{it} and v_{it}^2 are relevant instruments.

A.3.2 Several Inputs

In the next extension we assume that firms produce with two inputs, a variable input v_{it} and another input k_{it}. We assume that the additional input is, in the terminology of the production function literature, dynamic. This means that firms face adjustment costs and other intertemporal constraints when setting k_{it}, which leads firms to choose k_{it} before observing contemporaneous productivity. The production function in logs reads $y_{it} = \alpha v_{it} + \beta k_{it} + \omega_{it}$ and we are interested in estimating the parameters (α,β). Because k_{it} is set before productivity is observed, we only need to instrument the variable input with its lag. The estimation is therefore akin to a GMM regression with one endogenous and one exogenous variable. The estimator can be defined as follows:

Definition 6 *The GMM estimator is a pair* $(\widehat{\alpha}, \widehat{\beta})$ *such that* $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0$ *and* $\mathbb{E}\left[\widehat{\omega}_{it} k_{it-1}\right] = 0$, *where* $\widehat{\omega}_{it} = y_{it} - \widehat{\alpha} v_{it} - \widehat{\beta} k_{it} = (\alpha - \widehat{\alpha}) v_{it} + (\beta - \widehat{\beta}) k_{it} + \omega_{it}$.

Solving for the estimator $(\widehat{\alpha}, \widehat{\beta})$ implies solving for the following system of equations, defined by the moment conditions:

$$\begin{array}{l}\mathbb{E}\left[\widehat{\omega}_{it}v_{it-1}\right]=0 \\ \mathbb{E}\left[\widehat{\omega}_{it}k_{it-1}\right]=0\end{array} \iff \begin{array}{l}(\alpha-\widehat{\alpha})\mathbb{E}[v_{it}v_{it-1}]+(\beta-\widehat{\beta})\mathbb{E}[k_{it}v_{it-1}]=0 \\ (\alpha-\widehat{\alpha})\mathbb{E}[v_{it}k_{it-1}]+(\beta-\widehat{\beta})\mathbb{E}[k_{it}k_{it-1}]=0\end{array}.$$

This system can be rewritten in matrix form, with $V(B-\widehat{B})=0$, where

$$B-\widehat{B}=\begin{pmatrix}\alpha-\widehat{\alpha} \\ \beta-\widehat{\beta}\end{pmatrix} \qquad \text{and} \qquad V=\begin{pmatrix}\mathbb{E}[v_{it}v_{it-1}] & \mathbb{E}[k_{it}v_{it-1}] \\ \mathbb{E}[v_{it}k_{it-1}] & \mathbb{E}[k_{it}k_{it-1}]\end{pmatrix}.$$

Note that if the input k_{it} is perfectly correlated with the variable inputs v_{it} (and hence also variable), the matrix V will not be of full rank leading to non-identification of the estimator $(\widehat{\alpha},\widehat{\beta})$. However, if k_{it} is not variable, it is thus not perfectly correlated with v_{it}, which means that the determinant of V can be different from zero. As long as the determinant of V is not zero, the GMM estimator is identified and asymptotically consistent such that $\widehat{\alpha}=\alpha$ and $\widehat{\beta}=\beta$.

A.3.3 Persistent Productivity

In the final extension we assume that total factor productivity follows a first-order autoregressive (AR1) process in logs. The production function is still $y_{it}=\alpha v_{it}+\omega_{it}$, while the productivity process is $\omega_{it}=\rho\omega_{it-1}+\xi_{it}$. We would like to show the properties of GMM estimator $(\widehat{\alpha},\widehat{\rho})$ using v_{it-1} and $\widehat{\omega}_{it-1}$ as an instrument for v_{it} and $\widehat{\omega}_{it}$, where productivity is fitted based on a guess for α, because the true level of productivity is unobserved. The estimator is now defined as

Definition 7 *The GMM estimator is a pair* $(\widehat{\alpha},\widehat{\rho})$ *such that* $\mathbb{E}\left[\widehat{\xi}_{it}v_{it-1}\right]=0$ *and* $\mathbb{E}\left[\widehat{\xi}_{it}\widehat{\omega}_{it-1}\right]=0$*, where* $\widehat{\omega}_{it}=y_{it}-\widehat{\alpha}v_{it}=(\alpha-\widehat{\alpha})v_{it}+\omega_{it}$ *and* $\widehat{\xi}_{it}=\widehat{\omega}_{it}-\widehat{\rho}\widehat{\omega}_{it-1}=\xi_{it}+(\alpha-\widehat{\alpha})(v_{it}-\rho v_{it-1})+(\rho-\widehat{\rho})\omega_{it-1}+(\rho-\widehat{\rho})(\alpha-\widehat{\alpha})v_{it-1}$.

In practice, this estimator is solved for iteratively. Because the fitted productivity $\widehat{\omega}_{it}$ depends on α, the econometrician iterates over potential output elasticities $\widehat{\alpha}$ until the moment conditions for both productivity and the variable input are satisfied. This is why De Loecker and Warzynski (2012) label this procedure "iterative GMM". The estimator, $(\widehat{\alpha},\widehat{\rho})$, is characterized by the following system of equations defined by the moment conditions:

$$\begin{array}{l}\mathbb{E}\left[\widehat{\xi}_{it}v_{it-1}\right]=0 \\ \mathbb{E}\left[\widehat{\xi}_{it}\widehat{\omega}_{it-1}\right]=0\end{array} \iff \begin{array}{l}(\alpha-\widehat{\alpha})\mathbb{E}\left[(v_{it}-\rho v_{it-1})v_{it-1}\right]+(\rho-\widehat{\rho})\mathbb{E}\left[\omega_{it-1}v_{it-1}\right]+(\alpha-\widehat{\alpha})(\rho-\widehat{\rho})\mathbb{E}\left[v_{it-1}^2\right]=0 \\ (\alpha-\widehat{\alpha})\mathbb{E}\left[(v_{it}-\rho v_{it-1})\omega_{it-1}\right]+(\rho-\widehat{\rho})\mathbb{E}\left[\omega_{it-1}^2\right]+(\alpha-\widehat{\alpha})(\rho-\widehat{\rho})\mathbb{E}\left[v_{it-1}\omega_{it-1}\right]=0\end{array} \tag{1}$$

In general, the above system of equations admits two solutions. One is the true solution with $\widehat{\alpha}=\alpha$ and $\widehat{\rho}=\rho$, while the other solution converges to (α,ρ) as variation in the data increases. We leave the full and formal discussion of this case in Appendix A.3.3. To understand the essence of the argument consider the following proof sketch, when $\widehat{\alpha}$ and $\widehat{\rho}$ are not too far from α and ρ, respectively, the terms of the form $(\widehat{\alpha}-\alpha)(\widehat{\rho}-\rho)$ are of second order. In this case, the system characterizing the estimator $(\widehat{\alpha},\widehat{\rho})$ reduced locally to the matrix equation $V(B-\widehat{B})=0$ where

$$B-\widehat{B}=\begin{pmatrix}\alpha-\widehat{\alpha}\\ \rho-\widehat{\rho}\end{pmatrix} \text{ and } V=\begin{pmatrix}\mathbb{E}[(v_{it}-\rho v_{it-1})v_{it-1}] & \mathbb{E}[\omega_{it-1}v_{it-1}]\\ \mathbb{E}[(v_{it}-\rho v_{it-1})\omega_{it-1}] & \mathbb{E}[\omega^2_{it-1}]\end{pmatrix}.$$

As long as the determinant of V is not zero, the GMM estimator is locally identified and asymptotically consistent.

Below, we show that the GMM estimator is globally identified and asymptotically consistent as long as there is enough variation in the data. The GMM estimator with AR(1) productivity (Definition 7) is characterized by the system of equations

$$\begin{cases}\mathbb{E}\left[\widehat{\xi}_{it}v_{it-1}\right]=0\\ \mathbb{E}\left[\widehat{\xi}_{it}\widehat{\omega}_{it-1}\right]=0\end{cases}\iff\begin{cases}\mathbb{E}\left[\widehat{\xi}_{it}v_{it-1}\right]=0\\ (\alpha-\widehat{\alpha})\mathbb{E}\left[\widehat{\xi}_{it}v_{it-1}\right]+\mathbb{E}\left[\widehat{\xi}_{it}\omega_{it-1}\right]=0\end{cases}\iff\begin{cases}\mathbb{E}\left[\widehat{\xi}_{it}v_{it-1}\right]=0\\ \mathbb{E}\left[\widehat{\xi}_{it}\omega_{it-1}\right]=0\end{cases}$$

$$\iff\begin{cases}\mathbb{E}\left[\xi_{it}v_{it-1}\right]+(\alpha-\widehat{\alpha})\mathbb{E}\left[(v_{it}-\rho v_{it-1})v_{it-1}\right]+(\rho-\widehat{\rho})\mathbb{E}\left[\omega_{it-1}v_{it-1}\right]+(\alpha-\widehat{\alpha})(\rho-\widehat{\rho})\mathbb{E}\left[v^2_{it-1}\right]=0\\ \mathbb{E}\left[\xi_{it}\omega_{it-1}\right]+(\alpha-\widehat{\alpha})\mathbb{E}\left[(v_{it}-\rho v_{it-1})\omega_{it-1}\right]+(\rho-\widehat{\rho})\mathbb{E}\left[\omega^2_{it-1}\right]+(\alpha-\widehat{\alpha})(\rho-\widehat{\rho})\mathbb{E}\left[v_{it-1}\omega_{it-1}\right]=0\end{cases}$$

$$\iff\begin{cases}g+aX+bY+cXY=0\\ h+dX+eY+fXY=0,\end{cases}$$

where $X=\alpha-\widehat{\alpha}$, $Y=\rho-\widehat{\rho}$, and, $a=\mathbb{E}\left[(v_{it}-\rho v_{it-1})v_{it-1}\right]$, $b=\mathbb{E}\left[\omega_{it-1}v_{it-1}\right]$, $c=\mathbb{E}\left[v^2_{it-1}\right]$, $d=\mathbb{E}\left[(v_{it}-\rho v_{it-1})\omega_{it-1}\right]$, $e=\mathbb{E}\left[\omega^2_{it-1}\right]$, $f=\mathbb{E}\left[v_{it-1}\omega_{it-1}\right]=b$, $g=\mathbb{E}\left[\xi_{it}v_{it-1}\right]$, $h=\mathbb{E}\left[\xi_{it}\omega_{it-1}\right]$. Let us look at the asymptotic where $g=0$ and $h=0$. Assuming $c\neq 0$, we get

$$\begin{cases}aX+bY+cXY=0\\ dX+eY+fXY=0\end{cases}\iff\begin{cases}X=0\\ Y=0\end{cases}\text{ or }\begin{cases}X=-\frac{bd-ae}{cd-af}\\ Y=\frac{bd-ae}{ce-bf}\end{cases}\quad\text{if } cd-af\neq 0 \text{ and } ce-bf\neq 0.$$

It follows that there are two global solutions for the GMM estimator with AR(1):

$$\begin{cases}\widehat{\alpha}=\alpha\\ \\ \widehat{\rho}=\rho\end{cases}\text{ or }\begin{cases}\widehat{\alpha}=\alpha-\frac{bd-ae}{cd-af}=\alpha-\sqrt{\frac{\mathbb{V}ar[\omega_{it-1}]}{\mathbb{V}ar[v_{it-1}]}}\frac{Corr(\widetilde{v}_{it},v_{it-1})-Corr(\widetilde{v}_{it},\omega_{it-1})Corr(\omega_{it-1},v_{it-1})}{Corr(\widetilde{v}_{it},\omega_{it-1})-Corr(\widetilde{v}_{it},v_{it-1})Corr(\omega_{it-1},v_{it-1})}\\ \\ \widehat{\rho}=\rho+\frac{bd-ae}{ce-bf}=\rho+\sqrt{\frac{\mathbb{V}ar[\widetilde{v}_{it}]}{\mathbb{V}ar[v_{it-1}]}}\frac{Corr(\widetilde{v}_{it},v_{it-1})-Corr(\widetilde{v}_{it},\omega_{it-1})Corr(\omega_{it-1},v_{it-1})}{1-Corr(\omega_{it-1},v_{it-1})^2}\end{cases}$$

where $\widetilde{v}_{it}\equiv v_{it}-\rho v_{it-1}=\frac{1}{1-\alpha}\left(\xi_{it}+mc_{it}-\rho mc_{it-1}+w_t-\rho w_{t-1}\right)$.[46] The GMM estimator admits (exactly) two possible solutions. One solution provides the true value of the parameters, while the second solution is unrelated to the true parameters. However, if $\mathbb{V}ar[v_{it-1}]$ is large compared to $\mathbb{V}ar[\omega_{it-1}]$ and $\mathbb{V}ar[\widetilde{v}_{it}]$ (that is, their ratio goes to infinity while keeping fixed the correlation structure), then there is a unique solution for $\widehat{\alpha}$ and $\widehat{\rho}$. To conclude, if there is enough variation in the data, the GMM estimator is identified.

[46] Note that $Corr(\widetilde{v}_{it},\omega_{it-1})=Corr(mc_{it}-\rho mc_{it-1}+w_t-\rho w_{t-1},\omega_{it-1})$. Intuitively, if input price and marginal cost ($=P_{it}/\mu_{it}$) are uncorrelated with past value of productivity, this correlation will be equal to zero.

A.3.4 Full Proof

In this appendix we study the production function estimator for an arbitrary number of inputs, an arbitrary functional form (Cobb-Douglas or Translog) and an AR(1) productivity process.

Specifically, we assume the output of firm i at time t is such that $y_{it} = X'_{it}\beta + \omega_{it}$, where $\beta \in \mathbb{R}^N$ is a vector of parameters to be estimated, and, $X_{it} \in \mathbb{R}^N$ is a vector of inputs that can contain monomes and products of several inputs. This formulation nests the Cobb-Douglas and Translog case. For example, a two-inputs, v_{it}, m_{it} translog production function is modeled by $X_{it} = (v_{it}, m_{it}, v_{it}^2, m_{it}^2, v_{it}m_{it})'$ with parameters $\beta = (\beta_v, \beta_m, \beta_{v^2}, \beta_{m^2}, \beta_{vm})'$. We further assume that the (log) productivity ω_{it} follows an AR(1) process, that is, $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$. The GMM based estimator that we study here is defined as follows:

Definition 8 *The GMM estimator is $\widehat{\beta} \in \mathbb{R}^N$ and $\widehat{\rho} \in \mathbb{R}$ such that the moments $\mathbb{E}\left[X_{it-1}\widehat{\xi}_{it}\right]$ and $\mathbb{E}\left[\widehat{\omega}_{it-1}\widehat{\xi}_{it}\right]$ are equal to zero where $\widehat{\omega}_{it} = y_{it} - X'_{it}\widehat{\beta} = X'_{it}(\beta - \widehat{\beta}) + \omega_{it}$ and $\widehat{\xi}_{it} = \widehat{\omega}_{it} - \widehat{\rho}\widehat{\omega}_{it-1} = (X_{it} - \rho X_{it-1})'(\beta - \widehat{\beta}) + X'_{it-1}(\beta - \widehat{\beta})(\rho - \widehat{\rho}) + \omega_{it-1}(\rho - \widehat{\rho}) + \xi_{it}$*

In the remainder of this appendix we study the condition under which the above estimator admits solutions. To this end, let us study the following system of equations, which defined the estimator and whose unknowns are $\widehat{\beta}$ and $\widehat{\rho}$:

$$\begin{cases} \mathbb{E}\left[X_{it-1}\widehat{\xi}_{it}\right] = 0 \\ \mathbb{E}\left[\widehat{\omega}_{it-1}\widehat{\xi}_{it}\right] = \mathbb{E}\left[X_{it-1}\widehat{\xi}_{it}\right]'(\beta - \widehat{\beta}) + \mathbb{E}\left[\omega_{it-1}\widehat{\xi}_{it}\right] = 0 \end{cases} \iff \begin{cases} \mathbb{E}\left[X_{it-1}\widehat{\xi}_{it}\right] = 0 \\ \mathbb{E}\left[\omega_{it-1}\widehat{\xi}_{it}\right] = 0 \end{cases} \iff$$

$$\begin{cases} \mathbb{E}\left[X_{it-1}\widetilde{X}'_{it}\right](\beta - \widehat{\beta}) + \mathbb{E}\left[X_{it-1}X'_{it-1}\right](\beta - \widehat{\beta})(\rho - \widehat{\rho}) + \mathbb{E}\left[X_{it-1}\omega_{it-1}\right](\rho - \widehat{\rho}) = 0 \\ \mathbb{E}\left[\omega_{it-1}\widetilde{X}'_{it}\right](\beta - \widehat{\beta}) + \mathbb{E}\left[\omega_{it-1}X'_{it-1}\right](\beta - \widehat{\beta})(\rho - \widehat{\rho}) + \mathbb{E}\left[\omega_{it-1}^2\right](\rho - \widehat{\rho}) = 0 \end{cases},$$

where we use $\mathbb{E}\left[X_{it-1}\xi_{it}\right] = 0$ and $\mathbb{E}\left[\omega_{it-1}\xi_{it}\right] = 0$, and, where we denote $\widetilde{X}_{it} = X_{it} - \rho X_{it-1}$. Note that the first line of the above system of equations corresponds to N equations, while the second line is just a scalar equation. We have $N+1$ equations with unknown $(\widehat{\beta}, \widehat{\rho}) \in \mathbb{R}^{N+1}$. In general, this system of equations has multiple solutions, as in the case of one input.

Heuristically, when $(\widehat{\beta}, \widehat{\rho})$ is not too far from the true value (β, ρ), the terms in $(\beta - \widehat{\beta})(\rho - \widehat{\rho})$ are of second order. Ignoring these terms leads to the following reduced system which ca be written in matrix form:

$$\begin{cases} \mathbb{E}\left[X_{it-1}\widetilde{X}'_{it}\right](\beta - \widehat{\beta}) + \mathbb{E}\left[X_{it-1}\omega_{it-1}\right](\rho - \widehat{\rho}) = 0 \\ \mathbb{E}\left[\omega_{it-1}\widetilde{X}'_{it}\right](\beta - \widehat{\beta}) + \mathbb{E}\left[\omega_{it-1}^2\right](\rho - \widehat{\rho}) = 0 \end{cases} \iff \begin{pmatrix} \mathbb{E}\left[X_{it-1}\widetilde{X}'_{it}\right] & \mathbb{E}\left[X_{it-1}\omega_{it-1}\right] \\ \mathbb{E}\left[\omega_{it-1}\widetilde{X}'_{it}\right] & \mathbb{E}\left[\omega_{it-1}^2\right] \end{pmatrix} \begin{pmatrix} \beta - \widehat{\beta} \\ \rho - \widehat{\rho} \end{pmatrix} = 0$$

which admits a unique solution $(\widehat{\beta}, \widehat{\rho}) = (\beta, \rho)$ as long as the $(N \times N)$ matrix $\begin{pmatrix} \mathbb{E}\left[X_{it-1}\widetilde{X}'_{it}\right] & \mathbb{E}\left[X_{it-1}\omega_{it-1}\right] \\ \mathbb{E}\left[\omega_{it-1}\widetilde{X}'_{it}\right] & \mathbb{E}\left[\omega_{it-1}^2\right], \end{pmatrix}$ is invertible. We conclude that the GMM estimator is locally identified and unbiased.

A.4 Approximation of Demand System

In this appendix we show how to approximate the demand system specified by $Y = D(P)$ or $P = D^{-1}(Y)$. Note that both of these demand systems allows for differentiated goods across firms. For the former case, let us define the function $D_{it}(P)$ such that $Y_{it} = D_{it}(P)$. Around some symmetric equilibrium, (P_0^*, Y_0^*), at the first-order, we have, for all i, t

$$y_{it} = \log Y_{it} - \log Y_0^* \approx \sum_{jt} \frac{\partial \log D_{it}}{\partial \log P_{jt}} (\log P_{jt} - \log P_0^*) = \sum_{jt} J_{ijt} p_{jt},$$

where the matrix whose element are J_{ijt} is the Jacobian of the log of the demand D. Inverting this system of equations yields that for all i, $p_{it} = \sum_{jt} d_{ijt} y_{jt}$, where d_{ijt} are the elements of the inverse of the Jacobian matrix of the (log) demand D. For this case, when the demand is specified by $Y = D(P)$, we need to assume that the Jacobian of $\log D$ is invertible.

For the case where the demand is given by the inverse demand directly, $P = D^{-1}(Y)$, let us define the function D_{it}^{-1} such that $P_{it} = D_{it}^{-1}(Y)$. A first-order approximation around a symmetric equilibrium (P_0^*, Y_0^*) yields

$$p_{it} = \log P_{it} - \log P_0^* \approx \sum_{jt} \frac{\partial \log D_{it}^{-1}}{\partial \log Y_{jt}} (\log Y_{jt} - \log Y_0^*) = \sum_{jt} d_{ijt} y_{jt},$$

where, here, the d_{ijt} are the elements of the Jacobian matrix of the (log) inverse demand D^{-1}.

These formulations are useful when deriving the markup of firms of static oligopolistic Cournot or Bertrand competition. Under Bertrand (that is, when firms take other firm's prices as given), the profit of firm i at time t can be written as $\Pi_{it} = P_{it} Y_{it} - C_{it}(Y_{it}) = P_{it} D_{it}(P) - C_{it}(D_{it}(P))$, where $C_{it}(Y_{it})$ is the total cost of producing Y_{it} units. Under Bertrand, firms maximize their profit by setting their price P_{it}, while taking others' price as given. The first-order condition of this profit maximization problem yields that the markup is $\mu_{it} \equiv \frac{P_{it}}{\frac{\partial C_{it}}{\partial Y_{it}}} = \left(1 + \left(\frac{\partial \log D_{it}}{\partial \log P_{it}}\right)^{-1}\right)^{-1}$. Similarly, under Cournot competition, the profit of firm i at time t can be written as $\Pi_{it} = P_{it} Y_{it} - C_{it}(Y_{it}) = D_{it}^{-1}(Y) Y_{it} - C_{it}(Y_{it})$. Under Cournot, firms choose their quantity, taking other firm's quantity as given, which implies that the markup is $\mu_{it} \equiv \frac{P_{it}}{\frac{\partial C_{it}}{\partial Y_{it}}} = \left(1 + \frac{\partial \log D_{it}^{-1}}{\partial \log Y_{it}}\right)^{-1}$. To conclude, in most static oligopolistic competition models the firm-level markup can be written as $\mu_{it} = (1 + d_{iit})^{-1}$.[47]

[47] Under Cournot, we always have $\mu_{it} = (1 + d_{iit})^{-1}$, while under Bertrand, we further need to assume that d_{iit} the diagonal term of the Jacobian matrix of the (log) inverse demand D^{-1} is equal to $\left(\frac{\partial \log D_{it}}{\partial \log P_{it}}\right)^{-1}$.

A.5 Revenue Markup and Translog Production Function

We next compare markups from revenue and quantity production functions in a more general framework with a translog production function. The main intuition remains valid: the bias of the estimator on revenue Data are equal to the *average* demand elasticity among firms sharing the same production function.

Assume that the production function is $y_{it} = \alpha v_{it} + \beta v_{it}^2 + \omega_{it}$, while we maintain the other assumptions of our baseline framework. Let us study the bias implied by the use of revenue data in place of quantity data. Following the same logic as above, the coefficients of the production function estimated on revenue are such that

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + V^{-1} \begin{pmatrix} \mathbb{E}[p_{it} v_{it-1}] \\ \mathbb{E}[p_{it} v_{it-1}^2] \end{pmatrix}, \qquad \text{with} \qquad V = \begin{pmatrix} \mathbb{E}[v_{it} v_{it-1}] & \mathbb{E}[v_{it}^2 v_{it-1}] \\ \mathbb{E}[v_{it} v_{it-1}^2] & \mathbb{E}[v_{it}^2 v_{it-1}^2] \end{pmatrix}.$$

As in the Cobb-Douglas case, these estimates are biased. The above equation is the translog equivalent of equation (6) where the correlation of the instruments (lagged variable inputs and lagged variable inputs squared) with the output price is the case of the bias.

In the case of a translog production function, the true markup is such that $\mu_{it} = (\alpha + 2\beta \log V_{it}) \frac{P_{it} Y_{it}}{W_t V_{it}}$, and, the revenue markup is thus $\widehat{\mu}_{it}^R = \frac{\widehat{\alpha} + 2\widehat{\beta} \log V_{it}}{\alpha + 2\beta \log V_{it}} \mu_{it}$. As pointed out by Bond et al. (2021) and as in the Cobb-Douglas case, if we assume homogeneous inverse demand elasticities among firms in the sample (that is for all i we have $p_{it} = -\gamma y_{it}$), the revenue markup is equal to one.[48] However, in general the revenue markup is different from one and contains information on the true markup. To see this formally, we assume again that inverse demand elasticities are heterogeneous among firms, such that for all i by $p_{it} = -d_{iit} y_{it}$ where there is at least one pair (i, j) such that $d_{iit} \neq d_{jjt}$. As above, the true markup is given by $\mu_{it} = (1 - d_{iit})^{-1}$. In this heterogeneous inverse demand elasticity case, we have

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} = \left(I - \mathbb{E}\left[X_{it-1} X_{it}' \right]^{-1} \mathbb{E}\left[d_{iit} X_{it-1} X_{it}' \right] \right) \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

where X_{it} is vector $(v_{it}, v_{it-1}^2)'$ and I is the identity matrix. Hence, revenue markups satisfy

$$\widehat{\mu}_{it}^R = \left[1 - (\alpha + 2\beta \log V_{it})^{-1} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}' \left(\mathbb{E}\left[d_{iit} X_{it} X_{it-1}' \right] \mathbb{E}\left[X_{it} X_{it-1}' \right]^{-1} \right) \begin{pmatrix} 1 \\ 2 \log V_{it} \end{pmatrix} \right] (1 - d_{iit})^{-1}. \quad (2)$$

This markup is in general different from one for at least some firms. To see that clearly, let us further assume that the inverse demand elasticities are independent of the variable input usage and its square, such that, for any $n, m \in \mathbb{N}$, $\mathbb{E}\left[d_{iit} v_{it}^n v_{it-1}^m \right] = \mathbb{E}\left[d_{iit} \right] \mathbb{E}\left[v_{it}^n v_{it-1}^m \right]$.With these

[48] When $p_{it} = -\gamma y_{it}$, the vector $V^{-1} \begin{pmatrix} \mathbb{E}[p_{it} v_{it-1}] \\ \mathbb{E}[p_{it} v_{it-1}^2] \end{pmatrix} = \gamma \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ and the revenue markup becomes $\widehat{\mu}_{it}^R = (1 - \gamma) \frac{\alpha + 2\beta \log V_{it}}{\alpha + 2\beta \log V_{it}} (1 - \gamma)^{-1} = 1$.

assumptions in place, one can show that $\widehat{\alpha} = \alpha(1 - \mathbb{E}[d_{iit}])$ and $\widehat{\beta} = \beta(1 - \mathbb{E}[d_{iit}])$. The revenue markup is equal to $\widehat{\mu}_{it}^R = (1 - \mathbb{E}[d_{iit}])(1 - d_{iit})^{-1}$ which is different from one since there exist a pair (i, j) such that $d_{iit} \neq d_{jjt}$. As for the Cobb-Douglas case, the bias is determined by an average of the inverse demand elasticities.

In the translog case, the average revenue markup is $\mathbb{E}\left[\log \widehat{\mu}_{it}^R\right] = \mathbb{E}\left[\log(\mu_{it})\right] + \mathbb{E}\left[\log \frac{\widehat{\alpha}+2\widehat{\beta}\log V_{it}}{\alpha+2\beta\log V_{it}}\right]$. Let us assume that the inverse demand elasticities are heterogeneous across firms in the sample. From equation (2), we can see that the average of the log revenue markup is equal to zero up to a Jensen-like inequality:

$$\mathbb{E}\left[\log \widehat{\mu}_{it}^R\right] = -\mathbb{E}\left[\log(1-d_{iit})\right] + \mathbb{E}\left[\log\left(1-(\alpha+2\beta\log V_{it})^{-1}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}'\left(\mathbb{E}\left[d_{iit}X_{it}X_{it-1}'\right]\mathbb{E}\left[X_{it}X_{it-1}'\right]^{-1}\right)\begin{pmatrix}1\\ 2\log V_{it}\end{pmatrix}\right)\right].$$

When the inverse demand elasticities are homogeneous, $\forall i, d_{iit} = \gamma$, then the average log revenue markup is exactly zero. The relationship between the average revenue and true markup now depends on the distribution of the variable input $\log V_{it}$ and the extent of the bias in the production function estimation. Importantly, the variance of the revenue markup is different from the variance of the true markup and also depends on the distribution of inputs and the covariance of input and the true markup. Finally, the correlation between the revenue and the true markup is no longer equal to one. To gauge the information content of the revenue markup under translog, we rely on the simulations.

B Derivation of the Translog Production Function

In this appendix we derive the translog approximation of the CES production function and show that it nests the Cobb-Douglas production function. We specify a CES production function with homogeneity of degree γ:

$$Y_{iht} = \Omega_{iht}\left(\alpha[V_{iht}]^{\frac{\eta-1}{\eta}} + (1-\alpha)[K_{iht}]^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}\gamma}.$$

Cobb-Douglas derivation The generalized CES production function nests the Cobb-Douglas production function as $\eta \to 1$. To see this, note that

$$\ln y_{iht} = \omega_{iht} + \frac{\eta}{\eta-1}\gamma \ln\left[\alpha[V_{iht}]^{\frac{\eta-1}{\eta}} + (1-\alpha)[K_{iht}]^{\frac{\eta-1}{\eta}}\right].$$

The limit of this function for $\eta \to 1$ is

$$\begin{aligned}
\lim_{\eta\to1} \ln y_{iht} &= \omega_{iht} + \gamma \lim_{\eta\to1} \frac{\ln\left[\alpha[V_{iht}]^{\frac{\eta-1}{\eta}} + (1-\alpha)[K_{iht}]^{\frac{\eta-1}{\eta}}\right]}{\frac{\eta-1}{\eta}}, \\
&= \omega_{iht} + \gamma \lim_{\eta\to1} \frac{V_{iht}^{\frac{\eta-1}{\eta}} \ln V_{iht}\frac{\alpha}{\eta^2} + K_{iht}^{\frac{\eta-1}{\eta}} \ln K_{iht}\frac{(1-\alpha)}{\eta^2}}{1/\eta^2\left(\alpha[V_{iht}]^{\frac{\eta-1}{\eta}} + (1-\alpha)[K_{iht}]^{\frac{\eta-1}{\eta}}\right)}, \\
&= \omega_{iht} + \gamma\left[\ln V_{iht}\alpha + \ln K_{iht}(1-\alpha)\right],
\end{aligned}$$

where the second step follows from l'Hopital's rule. In levels, this yields the Cobb-Douglas production function with returns to scale γ:

$$Y_{iht} = \Omega_{iht}\left(V_{iht}^{\alpha} K_{iht}^{1-\alpha}\right)^{\gamma}.$$

Translog derivation The function implies the translog production function (15) up to a first-order approximation around $\eta = 1$. To see this, start from

$$\ln y_{iht} = \omega_{iht} + \frac{\eta}{\eta-1}\gamma \ln\left[\alpha[V_{iht}]^{\frac{\eta-1}{\eta}} + (1-\alpha)[K_{iht}]^{\frac{\eta-1}{\eta}}\right],$$

$$\ln y_{iht} = \omega_{iht} + \frac{\eta}{\eta-1}\gamma \ln\left[\alpha[V_{iht}]^{\frac{\eta-1}{\eta}}\left(1 + \frac{(1-\alpha)}{\alpha}\left[\frac{K_{iht}}{V_{iht}}\right]^{\frac{\eta-1}{\eta}}\right)\right],$$

$$\ln y_{iht} = \omega_{iht} + \frac{\eta}{\eta-1}\gamma \ln\left[\alpha[V_{iht}]^{\frac{\eta-1}{\eta}}\right] + \frac{\eta}{\eta-1}\gamma \ln\left[1 + \frac{1-\alpha}{\alpha}\left(\frac{K_{iht}}{V_{iht}}\right)^{\frac{\eta-1}{\eta}}\right].$$

Then moving the α back into the log term:

$$\ln y_{iht} = \omega_{iht} + \gamma v_{iht} + \frac{\eta}{\eta-1}\gamma \ln\left[\alpha + (1-\alpha)\left(\frac{K_{iht}}{V_{iht}}\right)^{\frac{\eta-1}{\eta}}\right].$$

Consider the final term. Rewriting yields

$$f(x) = \frac{\eta}{\eta-1}\gamma \ln\left[1 + (1-\alpha)\left(\left(\frac{K_{iht}}{V_{iht}}\right)^{\frac{\eta-1}{\eta}} - 1\right)\right],$$

$$f(x) = \frac{\gamma}{x} \ln\left[1 + (1-\alpha)\left(B^x - 1\right)\right],$$

where $B = K_{iht}/L_{iht}$ and $x = (\eta-1)/\eta$, such that our approximation is around $x \to 0$. Rewriting, we get

$$
\begin{aligned}
f(x) &= \frac{\gamma}{x}\ln\left[1+(1-\alpha)\left((\exp\left(\text{x}\ln\text{B}\right)-1\right)\right],\\
&\approx \frac{\gamma}{x}\ln\left[1+(1-\alpha)\left(\text{x}\ln\text{B}-\frac{x^2[\ln B]^2}{2}\right)\right],\\
&\approx \frac{\gamma}{x}\left[(1-\alpha)\left(\text{x}\ln\text{B}-\frac{x^2[\ln B]^2}{2}\right)-\frac{(1-\alpha)^2}{2}\left(\text{x}\ln\text{B}-\frac{x^2[\ln B]^2}{2}\right)^2\right].
\end{aligned}
$$

Given that we are approximating the function up to a first order we remove higher order terms, such that the final equation simplifies to

$$
f(x)=\frac{\gamma}{x}\left[(1-\alpha)\text{x}\ln\text{B}+\alpha\frac{1-\alpha}{2}x^2[\ln B]^2\right].
$$

Hence, the first-order approximation of the generalized CES production function reads

$$
y_{iht}=\omega_{iht}+\gamma\ln V_{iht}+\gamma(1-\alpha)\ln\left(\frac{K_{iht}}{V_{iht}}\right)+\gamma\alpha\frac{1-\alpha}{2}\frac{\eta-1}{\eta}\left[\ln\left(\frac{K_{iht}}{V_{iht}}\right)\right]^2.
$$

Grouping terms and denoting $x\equiv\ln X$:

$$
y_{iht}=\omega_{iht}+\gamma\alpha v_{iht}+\gamma(1-\alpha)k_{iht}+\gamma\alpha\frac{1-\alpha}{2}\frac{\eta-1}{\eta}\left(v_{iht}^2+k_{iht}^2-2k_{iht}v_{iht}\right),
$$

which is the translog production function (15) with homogeneity of degree γ.

Variable input demand We next derive the demand for the variable input for the translog production function. The firms' cost minimization problem involves minimizing costs W_tV_{iht} subject to the production function (15). Note that the output elasticity of the variable input is

$$
\frac{\partial y_{iht}}{\partial v_{iht}}=\gamma\alpha\left(1+[1-\alpha]\frac{\eta-1}{\eta}\left[v_{iht}-k_{iht}\right]\right),
$$

such that the first order condition of the cost minimization problem is:

$$
W_t=\lambda_{iht}\frac{Y_{iht}}{V_{iht}}\gamma\alpha\left(1+[1-\alpha]\frac{\eta-1}{\eta}\ln\left[\frac{V_{iht}}{K_{iht}}\right]\right)
$$

where λ_{iht} is the Lagrange multiplier. Inverting the first-order condition and inserting that marginal costs MC_{iht} equal λ_{iht}, we obtain (3).

Marginal costs As firms face an exogenous sequence of the fixed input K_{iht}, marginal costs can be derived from the production function (15) and optimal demand for the variable input

(3). Inserting the latter into the former, we get

$$y_{iht} = \omega_{iht} + \gamma\alpha \ln \left[\left(\frac{MC_{iht}}{W_t} \right) \gamma\alpha \left(1 - [1-\alpha] \frac{\eta - 1}{\eta} \ln \left[\frac{K_{iht}}{V_{iht}} \right] \right) Y_{iht} \right]$$
$$+ \gamma(1-\alpha)k_{iht} + \gamma\alpha \frac{1-\alpha}{2} \frac{\eta-1}{\eta} \left[\ln \left(\frac{K_{iht}}{V_{iht}} \right) \right]^2 .$$

Isolating log marginal costs on the left-and side yields (4).

For the calculation of the equilibrium, we use the fact that firms within the sector are subject to the same sequence of factor prices. For a given price of a unit of the variable input W_t, the firm solves the static cost-minimization problem by choosing optimal variable input demand. This yields the first-order condition:

$$V_{iht} = \left(\frac{MC_{iht}}{W_t} \right) \gamma\alpha \left(1 + [1-\alpha] \frac{\phi - 1}{\phi} \ln \left[\frac{V_{iht}}{K_{iht}} \right] \right) Y_{iht}, \quad (3)$$

where, from inserting optimal variable demand into the production function, log marginal costs $mc_{iht} \equiv \ln MC_{iht}$ can be expressed as:

$$mc_{iht} = \ln \left[\frac{W_t}{\gamma} Y_{iht}^{\frac{1-\alpha\gamma}{\alpha\gamma}} \Omega_{iht}^{-\frac{1}{\alpha\gamma}} K_{iht}^{\frac{\alpha-1}{\alpha}} \right] - \ln \left(1 + [1-\alpha] \frac{\phi-1}{\phi} \ln \left[\frac{V_{iht}}{K_{iht}} \right] \right) + \frac{1-\alpha}{2} \frac{\phi-1}{\eta} \left(\ln \left[\frac{V_{iht}}{K_{iht}} \right] \right)^2 . \quad (4)$$

C Additional Tables and Figures

Table A1: Estimation of AR(1) process for intermediate input prices

	(1)	(2)	(3)	(4)
Auto-regressive coefficient (ρ^w)	0.900***	0.871***	0.865***	0.868***
	(0.009)	(0.011)	(0.014)	(0.014)
St. Dev. of shocks (σ^w)	0.046	0.042	0.042	0.045
Controls	None	Year F.E.	Year F.E. & Ind. F.E.	Time Pol. & Ind. F.E.
Observations	798	798	798	798
R-squared	0.922	0.936	0.918	0.908

Note: Results from auto-regressions for intermediate input price indices (log) at the two-digit level. Data from EU-KLEMS for France, 1995-2016. Standard errors in parentheses. *, ** and *** denote statistical significance at the 10, 5 and 1% level, respectively. Time Pol. refers to the inclusion of a third-degree polynomial for time as a control.

Table A2: Estimation of AR(1) process for detrended nominal value added

	(1)	(2)	(3)	(4)
Auto-regressive coefficient (ρ^w)	0.999***	1.001***	0.677***	0.708***
	(0.00479)	(0.00412)	(0.0285)	(0.0257)
St. Dev. of shocks (σ^w)	0.166	0.140	0.419	0.390
Controls	None	Year F.E.	Year F.E. & Ind. F.E.	Time Pol. & Ind. F.E.
Observations	798	798	798	798
R-squared	0.982	0.987	0.709	0.608

Note: Results from auto-regressions for nominal sector-level value added (log) at the two-digit level, detrended with nominal GDP. Data from EU-KLEMS for France, 1995-2016. Standard errors in parentheses. *, ** and *** denote statistical significance at the 10, 5 and 1% level, respectively. Time Pol. refers to the inclusion of a third-degree polynomial for time as a control.

Table A3: Estimation of AR(1) process for fixed input using capital

	(1)	(2)	(3)	(4)
Auto-regressive coefficient (ρ^w)	0.988***	0.656***	0.656***	0.651***
	(0.000)	(0.008)	(008)	(0.002)
St. Dev. of shocks (σ^w)	0.215	0.215	0.662	11.79
Controls	None	Year F.E.	Year F.E. &	Ind-Year F.E. & Firm. F.E. Firm F.E.
Observations	160,124	160,124	160,124	160,124
R-squared	0.987	0.490	0.490	0.493

Note: Results from auto-regressions for French firms using EAP-FARE data for 2009-2019. Data on 27,857 firms. Standard errors in parentheses are clustered by firm. *, ** and *** denote statistical significance at the 10, 5 and 1% level, respectively. Industry fixed effects are at the 5-digit level.

Table A4: Estimated Translog production function parameters with alternative calibration

Coefficients	True	Quantity			Revenue		
		Full	Basic	None	Full	Basic	None
$\beta_v = \alpha\gamma$	0.77	0.75	0.73	0.3	0.73	0.7	0.73
		(0.02)	(0.018)	(0.227)	(0.016)	(0.011)	(0.059)
$\beta_k = (1-\alpha)\gamma$	0.33	0.34	0.37	0.65	0.2	0.23	0.21
		(0.016)	(0.015)	(0.146)	(0.013)	(0.009)	(0.036)
$\beta_{vv} = \gamma\frac{\alpha(1-\alpha)}{2}\frac{\phi-1}{\phi}$	0.011	0.007	0.004	-0.057	0.003	-0.001	0.001
		(0.003)	(0.003)	(0.031)	(0.002)	(0.002)	(0.008)
$\beta_{kk} = \beta_{vv}$	0.011	0.008	0.006	-0.016	0.001	-0.001	0.001
		(0.002)	(0.002)	(0.011)	(0.001)	(0.001)	(0.003)
$\beta_{vk} = -2\beta_{vv}$	-0.021	-0.016	-0.01	0.068	-0.008	-0.002	-0.004
		(0.004)	(0.004)	(0.039)	(0.004)	(0.002)	(0.01)
Implied avg. elasticity	0.715	0.714	0.713	0.684	0.712	0.711	0.73
(Std. dev.)	(0.024)	(0.017)	(0.01)	(0.126)	(0.008)	(0.003)	(0.003)

Note: Estimated production function coefficients for different specifications. The top panel presents production function estimates. The alternative calibration sets returns to scale $\gamma = 1.1$ and variable input share $\alpha = 0.7$, as well as demand elasticity across firms within markets $\epsilon = 7$ and demand elasticity across markets $\sigma = 1.7$ as in Burstein et al. (2020). The bottom panel presents elasticities of the variable input v on output (measured in terms of quantity or revenue). Standard deviations are presented in brackets. The first column presents true values for comparison. Bootstrapped standard errors are in parentheses. Full, Basic and None describe the specification of the first-stage regressions. Full first stages include a third-order expansion in the production function inputs, time fixed effects and additional controls for log price and market share. Basic first stages do not include the additional controls. Columns headed None do not deploy a first stage and therefore estimate markups on output variables that include measurement error η_{iht}.

Table A5: Translog Log Markup Estimates with alternative calibration

	Correlation $\ln \widehat{\mu}_{iht}$ with true markup	Log Markup Moments			
		Mean	St. Dev.	Median	IQR
True values	1.00	0.256	0.048	0.210	0.270
Quantity					
Full first stage (preferred)	0.98	0.255	0.049	0.216	0.262
Basic first stage	0.93	0.254	0.053	0.221	0.272
No first stage	0.13	0.196	0.208	0.345	0.860
Revenue					
Full first stage	0.91	0.253	0.054	0.222	0.279
Basic first stage	0.81	0.251	0.060	0.230	0.306
No first stage	0.84	0.277	0.058	0.253	0.266

Note: The alternative calibration sets returns to scale $\gamma = 1.1$ and variable input share $\alpha = 0.7$, as well as demand elasticity across firms within markets $\epsilon = 7$ and demand elasticity across markets $\sigma = 1.7$ as in Burstein et al. (2020). The first column present correlations of estimated markups with true values. Full first stages include a third-order expansion in production inputs and additional controls for log price and market share. Basic first stages do not include the additional controls.

Table A6: Change in Bias under Alternative Calibrations

Coefficients	True	Translog - Revenue (Full)			CD - Quantity (Full)		
		$N_k = 16$	$N_k = 8$	$N_k = 4$	$\phi = 1.05$	$\phi = 1.1$	$\phi = 1.3$
Production Function							
$\beta_v = \alpha\gamma$	0.32	0.38	0.39	0.33	0.32	0.31	0.29
		(-0.022)	(-0.013)	(-0.015)	-0.003	-0.003	-0.002
$\beta_k = (1-\alpha)\gamma$	0.48	0.33	0.28	0.23	0.49	0.49	0.52
		(-0.015)	(-0.007)	(-0.008)	-0.001	-0.001	-0.001
$\beta_{vv} = \gamma\frac{\alpha(1-\alpha)}{2}\frac{\phi-1}{\phi}$	0.009	0.018	0.021	0.013			
		(-0.004)	(-0.003)	(-0.003)			
$\beta_{kk} = \beta_{vv}$	0.009	0.003	-0.001	-0.009			
		(-0.002)	(-0.001)	(-0.001)			
$\beta_{vk} = -2\beta_{vv}$	-0.017	-0.023	-0.026	-0.018			
		(-0.005)	(-0.003)	(-0.003)			
Implied avg. elasticity	0.298	0.306	0.32	0.306	0.32	0.31	0.29
(Std. dev.)	(0.019)	(0.035)	(0.041)	(0.028)			

Note: Estimated production function coefficients for different specifications. The top panel presents production function estimates. The bottom panel presents elasticity of the variable input v with respect to output (quantity or revenue depending on the specification). The first column presents true values for comparison. Bootstrapped standard errors are in parentheses.

Table A7: Overview - Productivity and Log Markup Estimates on different specifications

	Correlation			Markup Moments (diff with true)			
	Markup	Prod.	Error	Mean	St. Dev.	Median	IQR
True	1.00	1.00	1.00	0.00	0.00	0.00	0.00
Translog - Revenue (Full FS)							
Sector with $N_k = 16$	0.61	0.76	1.00	0.02	0.04	-0.01	0.45
Sector with $N_k = 8$	0.69	0.65	1.00	0.03	0.04	-0.01	0.35
Sector with $N_k = 4$	0.97	0.52	0.99	-0.04	0.01	-0.06	0.15
Cobb-Douglass - Quantity (Full FS)							
Sector with $\phi = 1.05$	0.93	1.00	1.00	0.01	0.02	0.02	0.06
Sector with $\phi = 1.1$	0.82	0.99	1.00	0.02	0.04	0.04	0.16
Sector with $\phi = 1.3$	0.57	0.94	1.00	0.04	0.18	0.10	0.66

Note: Table of moments of estimated productivity and markups. The first two columns present correlations of estimated productivity and markups with the true values. For *Revenue TL Full* and *Quantity CD Full* markup moments are reported as difference with the true ones, since every sector has different markup values.

Table A8: Correlations across Simulated Specifications - Log Markups

	True	Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
			Pearson Correlation Cobb-Douglas Production Function				
True	1.00	0.82	0.82	0.82	0.82	0.82	0.82
Full First Stage - Quantity	0.82	1.00	1.00	1.00	1.00	1.00	1.00
Full First Stage - Revenue	0.82	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Quantity	0.82	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Revenue	0.82	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Quantity	0.82	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Revenue	0.82	1.00	1.00	1.00	1.00	1.00	1.00
			Pearson Correlation Translog Production Function				
True	1.00	1.00	0.69	1.00	0.73	0.66	0.40
Full First Stage - Quantity	1.00	1.00	0.75	0.99	0.79	0.72	0.48
Full First Stage - Revenue	0.69	0.75	1.00	0.66	1.00	1.00	0.94
Basic First Stage - Quantity	1.00	0.99	0.66	1.00	0.70	0.62	0.36
Basic First Stage - Revenue	0.73	0.79	1.00	0.70	1.00	0.99	0.92
No First Stage - Quantity	0.66	0.72	1.00	0.62	0.99	1.00	0.95
No First Stage - Revenue	0.40	0.48	0.94	0.36	0.92	0.95	1.00
			Spearman Correlation Cobb-Douglas Production Function				
True	1.00	0.76	0.76	0.76	0.76	0.76	0.76
Full First Stage - Quantity	0.76	1.00	1.00	1.00	1.00	1.00	1.00
Full First Stage - Revenue	0.76	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Quantity	0.76	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Revenue	0.76	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Quantity	0.76	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Revenue	0.76	1.00	1.00	1.00	1.00	1.00	1.00
			Spearman Correlation Translog Production Function				
True	1.00	0.99	0.65	1.00	0.69	0.61	0.39
Full First Stage - Quantity	0.99	1.00	0.72	0.98	0.76	0.69	0.48
Full First Stage - Revenue	0.65	0.72	1.00	0.60	1.00	1.00	0.94
Basic First Stage - Quantity	1.00	0.98	0.60	1.00	0.64	0.56	0.33
Basic First Stage - Revenue	0.69	0.76	1.00	0.64	1.00	0.99	0.92
No First Stage - Quantity	0.61	0.69	1.00	0.56	0.99	1.00	0.96
No First Stage - Revenue	0.39	0.48	0.94	0.33	0.92	0.96	1.00

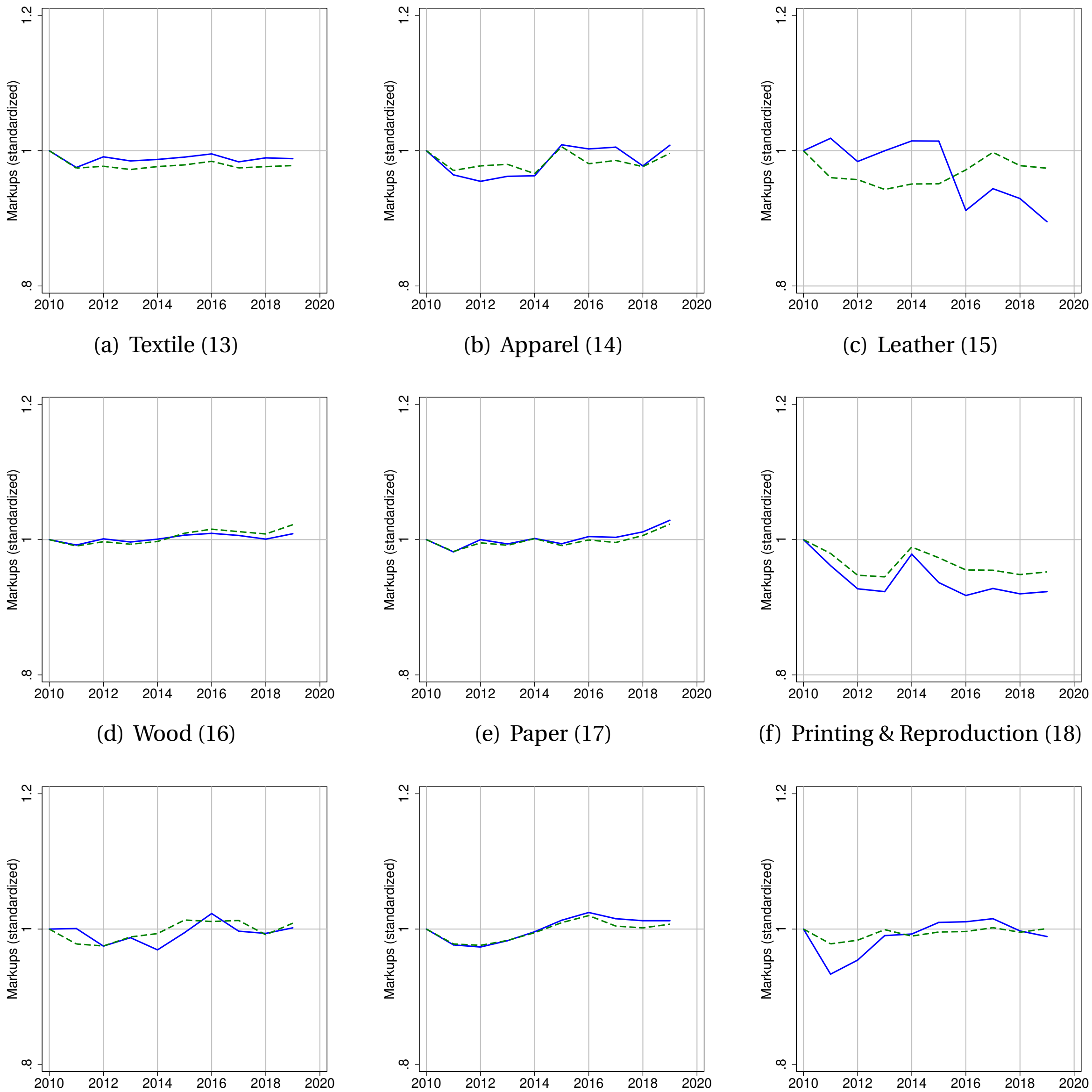
Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage. All markups are expressed in log. Simulated data.

Table A9: Correlations across Simulated Specifications - Log-Differenced Markups

	True	Full - Q	Full - R	Basic - Q	Basic - R	None - Q	None - R
	Pearson Correlation Cobb-Douglas Production Function						
True	1.00	0.89	0.89	0.89	0.89	0.89	0.89
Full First Stage - Quantity	0.89	1.00	1.00	1.00	1.00	1.00	1.00
Full First Stage - Revenue	0.89	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Quantity	0.89	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Revenue	0.89	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Quantity	0.89	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Revenue	0.89	1.00	1.00	1.00	1.00	1.00	1.00
	Pearson Correlation Translog Production Function						
True	1.00	1.00	0.80	1.00	0.83	0.78	0.49
Full First Stage - Quantity	1.00	1.00	0.83	0.99	0.86	0.81	0.54
Full First Stage - Revenue	0.80	0.83	1.00	0.78	1.00	1.00	0.91
Basic First Stage - Quantity	1.00	0.99	0.78	1.00	0.81	0.75	0.46
Basic First Stage - Revenue	0.83	0.86	1.00	0.81	1.00	0.99	0.89
No First Stage - Quantity	0.78	0.81	1.00	0.75	0.99	1.00	0.93
No First Stage - Revenue	0.49	0.54	0.91	0.46	0.89	0.93	1.00
	Spearman Correlation Cobb-Douglas Production Function						
True	1.00	0.86	0.86	0.86	0.86	0.86	0.86
Full First Stage - Quantity	0.86	1.00	1.00	1.00	1.00	1.00	1.00
Full First Stage - Revenue	0.86	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Quantity	0.86	1.00	1.00	1.00	1.00	1.00	1.00
Basic First Stage - Revenue	0.86	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Quantity	0.86	1.00	1.00	1.00	1.00	1.00	1.00
No First Stage - Revenue	0.86	1.00	1.00	1.00	1.00	1.00	1.00
	Spearman Correlation Translog Production Function						
True	1.00	1.00	0.79	1.00	0.82	0.76	0.52
Full First Stage - Quantity	1.00	1.00	0.83	0.99	0.85	0.80	0.57
Full First Stage - Revenue	0.79	0.83	1.00	0.76	1.00	1.00	0.92
Basic First Stage - Quantity	1.00	0.99	0.76	1.00	0.79	0.73	0.48
Basic First Stage - Revenue	0.82	0.85	1.00	0.79	1.00	0.99	0.90
No First Stage - Quantity	0.76	0.80	1.00	0.73	0.99	1.00	0.94
No First Stage - Revenue	0.52	0.57	0.92	0.48	0.90	0.94	1.00

Note: Each cell presents the pairwise correlation between the markup in the row and the column header. The first-stage regression includes a third-degree polynomial of inputs, price and market share controls, and time fixed effects ("Full"), only the polynomial and time fixed effects ("Basic") or no first stage. All markups are expressed in log. Simulated data.

Figure A1: Aggregate Markups - Sector Level



(a) Textile (13)

(b) Apparel (14)

(c) Leather (15)

(d) Wood (16)

(e) Paper (17)

(f) Printing & Reproduction (18)

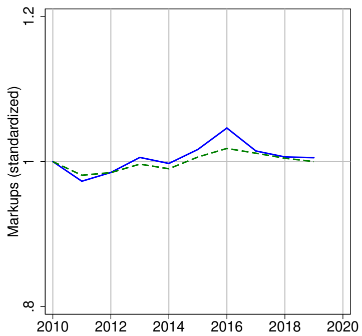
(g) Chemicals & Prod. (20)

(h) Rubber & Plastic (22)

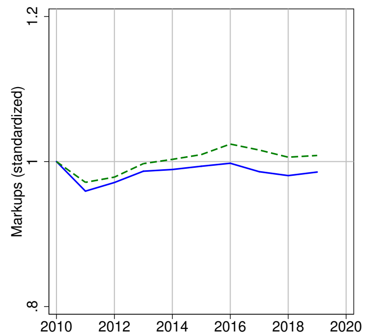
(i) Other Non-Metal, Minerals (23)

Notes: The figures plot the aggregate markup based on quantity data (blue-solid) and revenue data (green-dashed). The plots are an index where the aggregate markup in each year is divided by the level in 2010. Aggregate markups are the harmonic average of firm-level markups, weighted by sales share. two-digit NACE code in brackets.

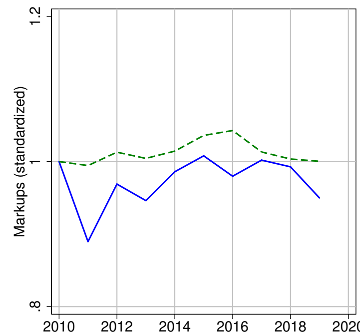
Figure A1: Aggregate Markups - Sector Level (Continued)



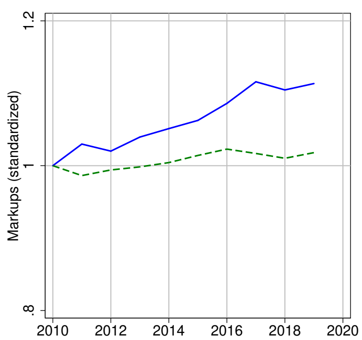
(a) Basic Materials (24)



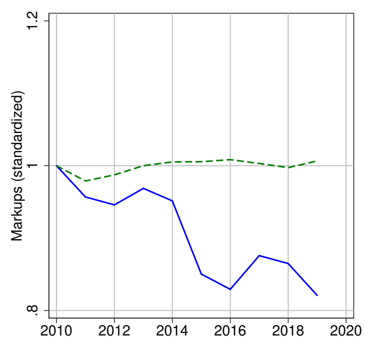
(b) Fabricated Metals (25)



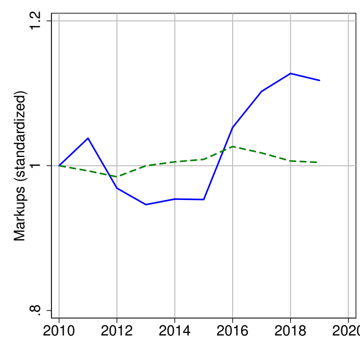
(c) Computers, Electr., Optic. (26)



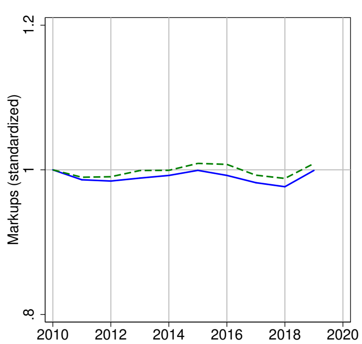
(d) Electrical Equipment (27)



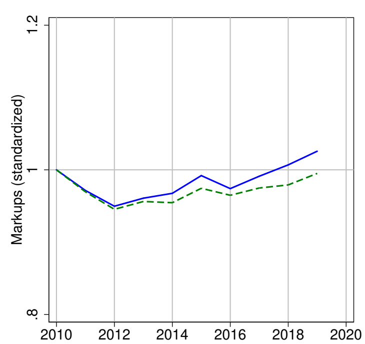
(e) Machinery and Equipment (28)



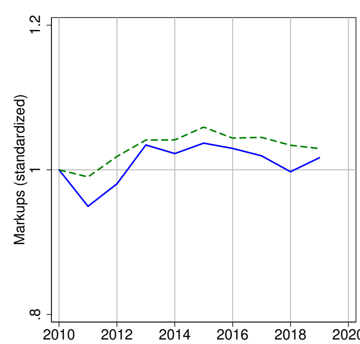
(f) Motor Vehicles & Trailers (29)



(g) Furniture (31)



(h) Other Manufacturing (32)



(i) Repairs & Installation (33)

Notes: The figures plot the aggregate markup based on quantity data (blue-solid) and revenue data (green-dashed). The plots are an index where the aggregate markup in each year is divided by the level in 2010. Aggregate markups are the harmonic average of firm-level markups, weighted by sales share. Two-digit NACE code in brackets.