INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

The Hitchhiker's Guide to Markup Estimation

Maarten De Ridder, Basile Grassi, Giovanni Morzenti

Working Paper n. 677

This Version: April 2025

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

The Hitchhiker's Guide to Markup Estimation: Assessing Estimates from Financial Data*

Maarten De Ridder† Basile Grassi‡ Giovanni Morzenti§

April 2025

Abstract

Macroeconomic outcomes depend on the distribution of markups across firms and over time, making firm-level markup estimates key for macroeconomic analysis. Methods to obtain these estimates require data on the prices that firms charge. Firm-level data with wide coverage, however, primarily comes from financial statements, which lack information on prices. We use an analytical framework to show that trends in markups or the dispersion of markups across firms can still be well-measured with such data. Measuring the average level of the markup does require pricing data, and we propose a consistent estimator for such settings. We validate the analytical results using simulations of a quantitative macroeconomic model and offer supporting evidence from firm-level administrative production and pricing data. Our analysis supports the use of financial data to measure trends in aggregate markups.

*We would like to thank the editor and referees, Jérôme 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, Ben Moll, Simon Mongey, Marco Ottaviani, Gianmarco Ottaviano, Fausto Panunzi, Michael Peters, Jean-Marc Robin, Ricardo Reis, Ariell Reshef, John van Reenen and Frederic Warzynski for useful comments. We also thank participants at numerous seminars and conferences. Stephan Hobler and Marco Panunzi provided excellent research assistance. The access to the French data is provided by CASD which is supported by a public grant 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 the macro-economy. Aggregate markups determine the labor and capital share in national income. Dispersion in markups across firms affects the efficiency of resource allocation. Variation in markups over the business cycle may explain real effects of nominal shocks. To understand macroeconomic outcomes, economists must therefore have a comprehensive picture of the distribution of firms' markups across the economy and over time. Yet neither prices nor marginal costs are observed in datasets at macroeconomists' disposal. This is because firm-level data with economy-wide coverage is nearly always derived from income statements and balance sheets, which at most have information on assets, revenue and costs.

A sprawling literature in macroeconomics and international trade has relied on markup estimates derived from such financial data, for example to quantify and test theories of imperfect competition à la Atkeson and Burstein (2008) and Kimball (1995).[1] A careful and quantitative assessment of the accuracy of firm-level markup estimates from financial data is, however, lacking.

This paper assesses the degree to which markups can be recovered from data on financial statements. We do so using an analytical framework, simulations of a quantitative macro model, and empirical analyses using French firm-level production and pricing data. We find that dispersion of markups across firms and trends over time can be well-estimated with financial data, provided that firms share a common production technology. This is the case if production is Cobb-Douglas and, more broadly, under empirically relevant conditions for more general production functions. Measuring average markups, however, requires data on prices, unless researchers employ methods that impose a demand system.

All parts of our analysis leverage the fact that, for cost-minimizing firms, markups equal the wedge between the elasticity of a firm's output with respect to a variable input–which firms set without adjustment costs–and that input's share in revenue (De Loecker and Warzynski 2012, Hall 1986, 1988). This is the production approach to markup estimation. As inputs' revenue shares are observed in financial statements, measuring the output elasticity is a main challenge when estimating firm-level markups. This usually involves estimating a production function, as is done in the pioneering paper by De Loecker and Warzynski (2012).

The paper starts with an analytical framework, in which we characterize the biases that arise from a key issue when estimating markups using the production approach: financial statements only detail firm revenues, not the quantity that they produce. This affects estimation of the variable input's output elasticity: if firms are price setters, revenue is not proportional to output, and hence revenue cannot be used to consistently estimate the production function (Klette and Griliches 1996). When

[1] This literature has been fuelled by rising aggregate markup estimates over time (e.g., De Loecker et al. 2020, Díez et al. 2021). These papers study the aggregate cost of markup (Baqaee and Farhi 2019; Edmond et al. 2023), the role of markups in inequality (Boar and Midrigan 2019), markup cyclicity (Hong 2017; Burstein et al. 2020), trade (Edmond et al. 2015; Gaubert and Itskhoki 2021), the monetary policy transmission (Baqaee et al. 2021; Chiavari et al. 2021; Meier and Reinelt 2022), inflation dynamics (Kouvavas et al. 2021), or price stickiness (Wang and Werning 2022; Mongey 2017).

revenue is used as a proxy for output, we show that the estimated output elasticities are biased by the correlation between firms' prices and the estimation's instruments, which often include last period's inputs. In simple models, that correlation is pinned down by an average of the price elasticity of demand. That average elasticity, in turn, determines average markups if firms are static profit maximizers. As a result, the estimated average of the markups is not informative about the true average. We show, however, that variation in markups across firms or over time can still be accurately measured, as long as firms have similar output elasticities. If firms differ in their output elasticities, revenue-based markup estimates correlate with true markups, and we show that the correlation approaches one if prices are orthogonal to the instruments.

These results appear at odds with the influential claim in Bond et al. (2021) that revenue-based markups are uninformative about true markups. We explain that their reasoning may hold on average, in the sense that the average revenue-based markup is usually not informative of the true average markup.[2]

The paper then validates these theoretical arguments through quantitative Monte Carlo simulations. The simulations enable us to scrutinize markup estimates from financial statements, where researchers lack data on prices, in a setting where true markups are known. We simulate a rich macroeconomic model of oligopolistic competition à la Atkeson and Burstein (2008) with endogenously heterogeneous markups. Firms differ in their fixed input and productivity, but share the same translog production function. For standard calibrations of the parameters and for various robustness calibrations, we find strong correlations between true markups and the various estimated markups. In a perfect scenario where one has data on the firms' quantity, markups are estimated with precision. In the practical scenario where researchers lack data on price and quantity, we still find a correlation of 0.94 between estimated and true markups. The high correlation is driven by the fact that in the full quantitative model, with various idiosyncratic and aggregate shocks, firms' prices and the instruments of the production function are close to orthogonal. Variation in markups – in the cross-section and over time – is also well-estimated, in line with the analytical results.

Finally, we compare revenue- and quantity-based markup estimates using administrative data on quantities and prices for French manufacturing firms. The data mostly validates the theoretical results. We find a 0.63 correlation between revenue and quantity markups, rising to 0.80 in first differences. These correlations are within sectors, consistent with the level of analysis in the theory and simulations. As in the simulations, we find that correlations between prices and the estimation instruments are low across all sectors. Binscatters show that revenue markups are good predictors of quantity markups across the domain. Regression coefficients that relate revenue and quantity

[2] Bond et al. (2021) write *"This approach uses the revenue elasticity for a flexible input, in place of the output elasticity"* and that the resultant markup is *"(..) identically equal to one, and therefore contains no useful information about markups."* Their conclusion and Appendix B.2. do note that it may be possible to study variation in markups without estimating production functions (and thus without data on quantity), employing data on the variable input's revenue share and firms' overall input usage.

markups to profits, labor- and market shares are also similar. Markup estimates do appear to be more dispersed when based on quantity, possibly due to several empirical complications: firms may not share identical production functions within sectors, for example, and functional forms for the production function and productivity process may be misspecified. For aggregate markups in France, we again find that the level is mismeasured when using revenue data. Trends over time, however, are well-estimated, and largely robust to different ways of aggregating markups.

We consider various extensions. If the variable input is subject to further input wedges, we show that markup estimates equal the product of markups and such wedges. The methods can also accommodate quality differences across firms.

Overall, we conclude that firm-level estimates of markups from the production approach are informative of true markups. However, 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 markups, one generally needs quantity data. Revenue-based markups should not be used, for example, to discipline parameters that govern aggregate markups. When interested in variation of markups across firms or trends over time, revenue data will likely suffice.

Related literature. We contribute to the large and growing literature that uses firm-level markups to understand the macroeconomic implications of imperfect competition. This literature relies on firm-level markup estimates across the entire economy and on long time windows to quantify theoretical models. Recent examples include Baqaee and Farhi (2019) and Edmond et al. (2023), who study the cost of markup dispersion, Boar and Midrigan (2019), who study the role of markups in inequality, or Hong (2017) and Burstein et al. (2020), who study markups over the business cycle.[3] We show that markup estimates from revenue data can be used to calibrate parameters relating to markup dispersion or relative markups across firms, but not for parameters that govern the average markup.

Methodologically, our paper relates most closely to work that estimates markups using the production approach. This method originates from the sector-level estimator in Hall (1986, 1988), who derives that the markup is equal to the wedge between a variable input's output elasticity and the input's revenue share.[4] In their seminal paper, De Loecker and Warzynski (2012) propose a generalized version of this estimator to measure firm-level markups, using production function estimation to obtain the output elasticity. If that elasticity were known, a variable input's revenue share suffices to measure markups, leaving the choice of a variable input as the sole challenge.[5] In practice, output

[3]Footnote 1 provides additional examples of recent papers.

[4]A significant body of the industrial organization literature, alternatively, derives markups from firms' first-order condition for pricing after estimating a demand system. This "demand approach", which builds on Berry et al. (1995), enables markup estimation at the product level. Recent contributions include Atalay et al. (2023), Döpper et al. (2022), Miller et al. (2023) and Grieco et al. (2024). Macroeconomists often prefer the production approach, as it does not impose a demand system.

[5]This methodology, deployed in recent papers including Burstein et al. (2020), Meier and Reinelt (2022), Calligaris et al.

elasticities must be estimated – for which Ackerberg et al. (2015)'s method is typically used. This two-stage procedure first purges output of measurement error and transitory productivity shocks, after which the elasticities are identified using instruments.

Our paper is closely related to recent work that criticizes the use of these production function estimation techniques when estimating markups. Bond et al. (2021) and Doraszelski and Jaumandreu (2021) point out that these techniques typically assume that firms are price takers – an unfortunate assumption when estimating markups. For price-taking firms, revenue is proportional to output, but price-setting firms must reduce prices when raising output. Revenue elasticities are thus generally lower than output elasticities (e.g., Klette and Griliches 1996), causing an underestimation of markups (De Loecker and Warzynski 2012). This does not mean quantity data is necessarily better, as quantities may be difficult to compare across firms due to product differentiation (De Loecker 2021).

A particularly strong critique is found in Bond et al. (2021), who claim that there is no information about true markups in estimates that rely on revenue to estimate output elasticities. We explain that this is only correct if the production function estimation delivers the revenue elasticity, which is typically not the case. The average revenue-based markup is not informative of the true average, but variation in markups is well-measured. This matters, as the paucity of price data as well, as the limited number of settings in which firms' quantities are comparable, means that Bond et al.'s claim seriously limited the scope for future analysis of markups.

Bond et al. (2021) do note that there are applications in which it is possible to study markup variation across firms without estimating a production function. This uses the fact that output elasticities are a function of inputs, and inputs can be added as controls in regressions. Our findings are complementary, as we conclude that variation in revenue-based markup estimates can be relied upon even in settings where it is not feasible to control for the output elasticities. Our results furthermore preserve the credibility of important prior contributions in macroeconomics and international trade that use revenue-based markup estimates to study variation in markups across firms and over time.[6]

A final issue is that production function estimates may be biased using Ackerberg et al. (2015)'s procedure if firms are price setters, even with quantity data.[7] Doraszelski and Jaumandreu (2019)

(2018) and De Ridder (2024), is a main alternative to markup estimators based on cost shares. A primary advantage is that it imposes little structure on the production function, which cost-share approaches do. One advantage of cost-share approaches is that they may be robust to factor-augmenting productivity – as is shown in Raval (2023a,b).

[6] De Loecker and Warzynski (2012) acknowledge that data on comparable quantities is needed to estimate average output elasticities and markups with precision. Our paper provides closed-form expressions for biases that arise in the absence of such data, as well as Monte Carlo and empirical evidence on the degree to which markup dispersion across firms and over time remains well-measured. De Loecker (2011) notes that specifying a demand system can also address biases in production function estimation due to missing firm-level price data. This solution is not suitable in a macroeconomic context as it reduces the generality of the conditions under which the markup estimates are obtained.

[7] Ackerberg et al. (2015)'s method is an example of the control function approach, developed to estimate production functions and productivity under perfect competition. The control function is a component of the purging regression, where measurement error is identified as the residual from regressing output on observables. To ensure productivity is not inadvertently purged, the observables must be an invertible function of productivity. The approach was pioneered by Olley and Pakes (1996) using investment to control for productivity. Levinsohn and Petrin (2003) modify this approach by proposing static controls (e.g., a variable input) to address the practical issue that observed investment is lumpy, and

and Brand (2019) point out that an identifying assumption in Ackerberg et al. (2015) is that demand faced by firms is not affected by unobservables other than productivity, which may not be true under oligopolistic competition. We propose a change to the procedure's first stage to address this.[8] A new paper by Ackerberg and De Loecker (2024) provides a detailed review of production function estimation under imperfect competition.

2 Analytical framework

In this analytical section, we first discuss how Hall (1986, 1988) and De Loecker and Warzynski (2012) relate markups to production functions. We follow the latter's more general derivation, which applies at the firm level. We then outline challenges to production function estimation with commonly available datasets if firms have market power. Finally, we derive expressions for the biases that arise when revenue is used to approximate output, or in case of measurement error.

2.1 From Markups to Production Functions

The seminal papers by Hall (1986, 1988) and De Loecker and Warzynski (2012) propose a method to estimate markups that relates markups to the production function. The idea is that markups can be inferred from the wedge between the output elasticity of a variable input and that input's share in revenue. An input is variable if firms choose its use every period to minimize costs, without intertemporal considerations or adjustment costs, taking the price of this input as given.

Formally, the output Y_{it} for firm i at time t is given by the production function $Y_{it} = Y(V_{it}, \boldsymbol{K}_{it}, \Omega_{it})$, where V_{it} is the variable input, purchased at price W_t. The vector $\boldsymbol{K}_{it}$ contains all other inputs, while Ω_{it} represents productivity. The first-order condition for the cost-minimizing firm with respect to V_{it} is given by $\frac{1}{\eta_{it}} = \frac{\partial Y_{it}}{\partial V_{it}} \frac{1}{W_t}$, where η_{it} is the Lagrange multiplier of the production function constraint and is equal to marginal costs. Multiplying both sides by price P_{it} yields the markup expression that we use throughout the analysis:

$$\mu_{it} = \alpha_{it} \frac{P_{it} Y_{it}}{W_t V_{it}}, \tag{1}$$

where $\mu_{it} \equiv P_{it}/\eta_{it}$ is the markup and $\alpha_{it} = \frac{\partial Y_{it}}{\partial V_{it}} \frac{V_{it}}{Y_{it}}$ is the output elasticity of V_{it}.

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

thus not an invertible function of productivity. Ackerberg et al. (2015) amend the procedure by identifying all coefficients in the second stage rather than the first, addressing collinearity between productivity and variable inputs. We elaborate in Section 2. A detailed discussion of the history of production function estimation, including a summary of practical issues with early methods, can be found in Ackerberg et al. (2015). De Loecker and Syverson (2022) provide a careful contemporaneous discussion, including a review of differences between control function approaches and dynamic panel approaches (e.g. Blundell and Bond 2000) for production function estimation.

[8] We do need sufficient variation in input prices so that inputs and productivity are not colinear (Blundell and Bond 2000, Gandhi et al. 2020). Further discussion on markup estimation with accounting data is found in Traina (2018), Morlacco (2019), Basu (2019), Syverson (2019), Yeh et al. (2022).

follows that to estimate markups, researchers need data on revenue and input spending from the income statement, as well as an estimate of α_{it}, the output elasticity of V_{it}. Estimating this elasticity under imperfect competition is thus a primary challenge in markup estimation.

It is worth noting that for inputs not conforming to the assumptions placed on V_{it} (i.e. adjustment costs for capital or imperfections in the labor market), that input's equivalent of equation (1) can be used to calculate the overall input wedge, as defined by Hsieh and Klenow (2009). We discuss this further in Section 2.4.

2.2 Estimating Markups with Revenue Data

We next introduce an analytical framework to analyze whether markups can be measured along (1) when estimating α_{it} using commonly available datasets. As these markups inform macroeconomic and international trade theories, it is crucial for data to have economy-wide coverage, preferably over long time horizons. Datasets meeting this bar are usually based on financial statements, where revenue is the sole measure of output. As firms' decisions influence prices under imperfect competition, however, revenue may poorly approximate output.

2.2.1 When do Revenue-Based Markups Measure True Markups?

We first examine how accurately revenue-based markup estimates measure true markups, showing that their correlation depends on the estimated output elasticity of the variable input. If the estimated elasticity matches the true output elasticity, true markups are recovered; if it matches the elasticity of revenue with respect to the variable input—the *revenue elasticity*—markup estimates are orthogonal to true markups; if it falls between these elasticities, markup estimates positively correlate with true markups.

To derive this, we introduce a simple demand system. Firms face a price-elasticity of demand $\frac{dy_{it}}{dp_{it}} = -1/\varepsilon_{it}$, where lower case letters denote the log deviation from a sample mean. The demand elasticity can be heterogeneous across firms and over time. Firms that maximize profits period-by-period in the face of this demand will charge a markup $\mu_{it} = (1-\varepsilon_{it})^{-1}$. This is the standard inverse elasticity rule that describes how firms set prices in static models of oligopolistic or monopolistic competition. The demand system gives us the elasticity of revenue $R_{it} = P_{it}Y_{it}$ with respect to V_{it} as $\frac{dr_{it}}{dv_{it}} = \frac{dp_{it}}{dy_{it}}\frac{dy_{it}}{dv_{it}} + \frac{dy_{it}}{dv_{it}} = (1-\varepsilon_{it})\alpha_{it}$.

With true markups delivered by the demand system, we can now derive the correlation between true markups and revenue-based markup estimates. These estimates use equation (1), where the output elasticity α_{it} is replaced by the estimated elasticity on revenue data, $\widehat{\alpha}_{it}$. The literature has developed techniques to estimate parameters of a production function with firm data. When such techniques use revenue in place of quantity data, the production function parameters (and hence $\widehat{\alpha}_{it}$) are biased (Klette and Griliches 1996). The bias is a function of the joint distribution of inputs and of the elasticities of both demand and output, as we show in Section 2.2.2. Hence, revenue-based markup estimates read as

$$\widehat{\mu}_{it}^{R} \equiv \widehat{\alpha}_{it}\frac{P_{it}Y_{it}}{W_tV_{it}} = \frac{\widehat{\alpha}_{it}}{\alpha_{it}}\alpha_{it}\frac{P_{it}Y_{it}}{W_tV_{it}} = \frac{\widehat{\alpha}_{it}}{\alpha_{it}(1-\varepsilon_{it})} = \frac{\widehat{\alpha}_{it}}{\alpha_{it}}\mu_{it}. \tag{2}$$

This elucidates the relationship between true and estimated markups, and true and estimated output elasticities. The final equality confirms that if the estimated elasticity $\widehat{\alpha}_{it}$ equals the true elasticity, revenue-based markups recover true markups. The penultimate equality shows that the revenue-based estimates $\widehat{\mu}_{it}^{R}$ can correlate with true markups μ_{it} as long as $\widehat{\alpha}_{it}$ is different from the revenue elasticity $\alpha_{it}(1-\varepsilon_{it})$. In such cases, $\mathbb{C}ov\left[\log\mu_{it}, \log\widehat{\mu}_{it}^{R}\right] \neq 0$. However, if markups are computed using the true revenue elasticity $\alpha_{it}(1-\varepsilon_{it})$ in place of α_{it}, we have

$$\widehat{\mu}_{it}^{RE} \equiv (1-\varepsilon_{it})\alpha_{it}\frac{P_{it}Y_{it}}{W_tV_{it}} = (1-\varepsilon_{it})\mu_{it} = \frac{1-\varepsilon_{it}}{1-\varepsilon_{it}} = 1.$$

That is, the markup estimates would be identically equal to one and would be orthogonal to true markups: $\mathbb{C}ov\left[\log\mu_{it}, \log\widehat{\mu}_{it}^{RE}\right] = 0$, as in Bond et al. (2021).

2.2.2 Estimating the Output Elasticity with Revenue Data

We next show that, in practice, estimating a parametric production function using revenue instead of quantity yields neither the revenue nor the output elasticity, as long as firms have heterogeneous markups. Rather, the estimate of the output elasticity $\widehat{\alpha}_{it}$ that one obtains when using revenue instead of quantity data suffers from an omitted variable bias. When quantity is used to estimate the production function, the output elasticity can be consistently estimated using standard tools, as we discuss in the appendix.[9] Below, use a simple analytical framework to derive the omitted variable bias that arises when using revenue instead.

Consider a set of firms in a single sector. Output Y_{it} is log-linear in a single variable input V_{it}, while productivity is identically and independently distributed across firms and time. Productivity is unobserved by the researcher but observed by the firm. Firms set V_{it} and share the same Cobb-Douglas production function

$$y_{it} = \alpha v_{it} + \omega_{it}, \tag{3}$$

where lower caps denote log-deviations from the mean, and where α is the true output elasticity of v_{it} to be estimated.[10] This simple environment allows us to keep the argument as transparent as possible and to derive clear closed-form solutions. Despite the simplicity, the intuitions extend to more general models that are standard in the literature. In Section 2.2.4, for example, we study

[9] We show that an IV-GMM estimator is consistent and recovers the output elasticity for the simple analytical framework (Appendix A.1) and in more general cases (Appendix A.4).

[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 sample average. This normalization gets rid of constants in the production function and ensures that ω_{it} is mean zero.

the case of a translog production function where the output elasticity is heterogeneous across firms. Appendix A.4 furthermore extends the results by allowing for multiple inputs (A.4.2), persistence in productivity (A.4.3), and all of these together (A.4.4).[11]

Turning to the estimation of α, a least-squares regression of input v_{it} on output y_{it} will be biased even if output is observed. This is because productivity affects a firm's input choices and is unobserved to the econometrician, and is thus in the residual. The production function estimation literature, such as Blundell and Bond (2000) or Ackerberg et al. (2015), identifies α by instrumenting v_{it} by its lag v_{it-1}. In our setup, since productivity is i.i.d., v_{it-1} is not in the same information set and thus is orthogonal to ω_{it}.[12] The instrument therefore satisfies the exclusion restriction. For instrument relevance, we also need v_{it} to be persistent over time. This might arise through persistence in input price W_t. Gandhi et al. (2020) note that under perfect competition, this is the sole driver of persistence, so that long time samples are required for identification. We show that under imperfect competition (the natural setting when estimating markups), it is easier to obtain persistence in v_{it}, as a firm's persistent set of competitors affects its demand for inputs (Appendix A.1). Hence the output elasticity α can be identified.

When estimating the production function with revenue data, one obtains a biased estimate of α. To show this, let us construct an instrumental variable estimator based on the generalized method of moments (IV-GMM) when revenue is used in place of quantity. We focus on infinite samples to abstract from finite-sample variation. Hence, $\mathbb{E}\left[x_{it}\right]$ denotes the limit in probability of the sample average of a variable x_{it} as the sample size goes to infinity.[13] We thus focus on consistency of the estimator. Revenue is quantity times price, such that $r_{it} = y_{it} + p_{it}$ is revenue in log-deviations from the mean. Furthermore, inserting production function (3) for y_{it} yields $r_{it} = y_{it} + p_{it} = \alpha v_{it} + \omega_{it} + p_{it}$, where α remains the parameter of interest. If a researcher were to use an IV-GMM estimator that is consistent for quantity, but uses revenue as the dependent variable instead, that estimator would be defined as follows:

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

Without loss of generality we treat $\widehat{\alpha}$ as a non-random real number. Formally, $\widehat{\alpha}$ is a random variable which is almost surely equal to a constant. Solving for $\widehat{\alpha}$ such that $0 = \mathbb{E}[\widehat{\mathrm{tfpr}}_{it} v_{it-1}] = (\alpha -$

[11] Appendix A.4.3 notes that if productivity has persistence ρ, identification may only hold locally. In particular, there are exactly two solutions to the IV-GMM estimator. One matches the true parameters, while the second is a biased estimate of the true parameters. This is in line with recent work by Ackerberg et al. (2020), who show independently that the two-stage estimator may have two solutions, rendering standard numerical solvers unstable. However, in our framework, we show that if $\mathbb{V}ar[v_{it-1}]$ is large compared to $\mathbb{V}ar[\omega_{it-1}]$ and $\mathbb{V}ar[v_{it} - \rho v_{it-1}]$, then there exists a unique solution for $\widehat{\alpha}$ and $\widehat{\rho}$. Hence, if there is enough variation in the data, the production function is globally identified.

[12] Appendix A.4.3 notes that when productivity is persistent, e.g. when it is AR(1), v_{it-1} is still a valid instrument if the moment condition is for v_{it-1} to be orthogonal to innovations of productivity. A similar argument applies there: v_{it-1} is not in the same information set as the time t innovations.

[13] By the weak law of large numbers, under independence of the x_{it}, $\mathbb{E}\left[x_{it}\right] \equiv \text{plim} \frac{1}{N} \sum_{it} x_{it}$ also denotes the expected value of x_{it}. Appendix A.2 derives the estimator for a finite size sample of firms.

$\widehat{\alpha})\mathbb{E}\left[v_{it}v_{it-1}\right] + \mathbb{E}\left[p_{it}v_{it-1}\right]$, yields a unique solution as long as the lagged variable input is a relevant instrument (that is, $\mathbb{E}\left[v_{it}v_{it-1}\right] \neq 0$):

$$\widehat{\alpha} = \alpha + \frac{\mathbb{E}\left[p_{it}v_{it-1}\right]}{\mathbb{E}\left[v_{it}v_{it-1}\right]}. \tag{4}$$

The estimator is clearly not consistent if prices and lagged variable inputs covary, i.e. $\mathbb{E}\left[p_{it}v_{it-1}\right] \neq 0$. Using revenue to measure output creates an omitted variable bias: the revenue-production function has prices in the residual, as first pointed out by Klette and Griliches (1996) and discussed in De Loecker et al. (2016).

Under perfect competition, the correlation between price and lagged inputs is zero, as firms are atomistic price takers. Under imperfect competition, it is probable that p_{it} will correlate with lagged variable inputs, such that $\mathbb{E}\left[p_{it}v_{it-1}\right]$ differs from zero. 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.

2.2.3 Revenue-Based Markup Estimates

We next show that, even when biased, revenue-based markup estimates are still informative about true markups. The bias in the estimated elasticity in equation (4) is, in part, determined by the demand system – and so to show this, we re-introduce a demand side to our baseline framework. We assume a general invertible demand system, where a firm's demand depends on prices of all firms. Formally, the vector of quantities produced by all firms, $\{Y_{it}\}$, is a function of the price vector $\{P_{it}\}$ such that $\{Y_{it}\} = D_t(\{P_{it}\})$. A log-linear approximation yields

$$p_{it} = -\textstyle\sum_j \varepsilon_{ijt} y_{jt}, \tag{5}$$

where ε_{ijt} is the cross-elasticity of firm i's price to firm j's quantity. For now, we abstract from aggregate shocks that alter price-quantity relationships across periods, and hence focus on the bias caused by downward-sloping demand curves.

With this demand system, the revenue elasticity of the variable input, taking other firms' output as given, is $\frac{dr_{it}}{dv_{it}} = \alpha(1-\varepsilon_{iit})$. The variable input v_{it} can be used to compute markups along eq. (1); as firms share a common output elasticity α, true markups equal $\mu_{it} = \alpha(P_{it}Y_{it})/(W_t V_{it})$. The estimated elasticity, substituting demand system (5) into eq. (4) and using production function (3), is given by

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

The ratio of the true output elasticity and the revenue-based estimated elasticity is equal to one minus the weighted average of inverse demand elasticities and cross-elasticities among the firms sharing the same production function. Importantly, the estimated elasticity $\widehat{\alpha}$ differs, generally, from the revenue elasticity $\alpha(1-\varepsilon_{iit})$. This implies that the revenue markup is different from one, as in eq. (2). Indeed, firm-level markup estimates based on revenue data, $\widehat{\mu}_{it}^R = \widehat{\alpha}\frac{P_{it}Y_{it}}{W_tV_{it}}$, are

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

This shows that revenue-based markup estimates are equal to true markups up to a constant, as the second term in parenthesis is not firm-specific (the $\mathbb{E}$ takes an average over i). The true and estimated revenue log markups have equal variances and the correlation between the estimates and true markups is one.

The result that revenue and quantity markups perfectly correlate depends on the Cobb-Douglas assumption that output elasticity is constant. The bias is a constant, and thus does not cancel out variation in markups. In Section 2.2.4 we discuss the case of variable output elasticities and show that the core insights remain: variation in the bias does not cancel out variation in markups.

Case I: heterogeneous demand elasticities. The result that revenue-based estimates of the markup perfectly covary is only useful if markups are variable across firms. In our simple demand system, in which firms set prices to maximize contemporaneous profits, firms will have heterogeneous markups if they are subject to heterogeneous price elasticities of demand. For this case, it is straightforward to derive that all dispersion in the markup is preserved, but mean markups may be sufficiently biased such that *no information* about the true average remains.

To see this, start from the demand system in equation (5), with the additional assumption that for all pairs of firms i,j with $i\neq j$, $\varepsilon_{ijt}=0$ while $\varepsilon_{iit}\neq 0$ and $\varepsilon_{iit}\neq\varepsilon_{jjt}$. Thus, demand is determined by firms' own supply, and firms face heterogeneous demand elasticities. Formally, the demand system is such that $p_{it}=-\varepsilon_{it}y_{it}$ where, with some abuse of notation, we denote $\varepsilon_{it}\equiv\varepsilon_{iit}$ as the own inverse price elasticity. When firms maximize profits, they charge a markup $\mu_{it}=(1-\varepsilon_{it})^{-1}$.

Under these assumptions, revenue IV-GMM estimators give the *average* revenue elasticity among firms sharing a production function. This differs from each firm's own revenue elasticity, because firms have different demand elasticities: [14]

$$\widehat{\alpha} = \mathbb{E}\left[\alpha(1-\varepsilon_{it})\frac{v_{it}v_{it-1}}{\mathbb{E}\left[v_{it}v_{it-1}\right]}\right] \quad \neq \quad \frac{\partial r_{it}}{\partial v_{it}} = \alpha(1-\varepsilon_{it}). \tag{6}$$

Turning to the resultant markup estimates along equation (1) and for markups $\mu_{it}=(1-\varepsilon_{it})^{-1}$

[14] This assumes that $\mathbb{E}\left[\varepsilon_{it}\omega_{it}v_{it-1}\right]=0$. This assumption is satisfied (for example) when, conditional on v_{it-1}, productivity ω_{it} and elasticity ε_{it} are orthogonal or, alternatively, when conditional on ε_{it}, ω_{it} and v_{it-1} are orthogonal. We make this assumption merely to clarify the argument.

that maximize profits, we get:

$$\widehat{\mu}_{it}^{R} \equiv \widehat{\alpha}\frac{P_{it}Y_{it}}{W_t V_{it}} = \mathbb{E}\left[\mu_{it}^{-1}\frac{v_{it}v_{it-1}}{\mathbb{E}[v_{it}v_{it-1}]}\right]\mu_{it}. \tag{7}$$

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. Given the assumption that two firms i and j have different markups, the estimated revenue markup $\widehat{\mu}_{it}^{R}$ is different from one. However, the average of the estimated revenue markup is not informative about the average true markup. Indeed, the average estimated revenue markup can be written as

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

which equals one up to a Jensen's inequality. Revenue markups thus carry no information about the true average in this demand system. We conclude that using revenue data instead of quantity data does not allow us to recover the level of markups but allows us to recover variation in markups, if variation exists.

Case II: homogenous demand elasticities. 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. In that case, the revenue estimator equals the revenue elasticity $\widehat{\alpha} = \alpha(1-\gamma) = \frac{\partial y_{it}}{\partial v_{it}}(1+\frac{\partial p_{it}}{\partial y_{it}}) = \frac{\partial r_{it}}{\partial v_{it}}$. Both the revenue elasticity and the true markup are equal across firms, where the latter is equal to $(1-\gamma)^{-1}$. It follows that revenue-based markup estimates are identically equal to one $\widehat{\mu}_{it}^{R} = (1-\gamma)^{-1}(1-\gamma) = 1$, as in Bond et al. (2021). If markups are identical across firms, revenue markups thus do not yield any information on true markups.

2.2.4 Beyond Constant Output Elasticities

The analysis thus far assumes that firms have equal output elasticities. We next study the more general case where the output elasticity depends on firms' input choices, by assuming that the production function is translog. The translog production function nests Cobb-Douglas and is a second-order approximation of any production function. In our one-input environment, the function is given by

$$y_{it} = \alpha v_{it} + \beta v_{it}^2 + \omega_{it}, \quad \text{so that} \quad \alpha_{it} \equiv \frac{\partial y_{it}}{\partial v_{it}} = \alpha + 2\beta v_{it}. \tag{8}$$

As for the case where firms have equal output elasticities, the bias from using revenue instead of quantity depends on the correlation between prices and the instruments in the production function estimation, which are typically the lags of v_{it} and v_{it}^2. As we derive in Appendix A.5, the revenue IV-GMM estimator gives

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

If prices are orthogonal to the instruments, the second term disappears and the estimates are consistent. If either the lagged variable input or its square is correlated with prices, both parameters are biased, as is the estimated elasticity $\widehat{\alpha}_{it} = \widehat{\alpha} + 2\widehat{\beta} v_{it}$. Inserting this into equation (1) gives the revenue markup:

$$\widehat{\mu}_{it}^R = (1 + f(v_{it}))\, \mu_{it} \quad \text{where} \quad f(v_{it}) = \frac{(\widehat{\alpha} - \alpha) + 2(\widehat{\beta} - \beta) v_{it}}{\alpha + 2\beta v_{it}},$$

which equals true markups multiplied by a function of firms' own input usage. If prices are orthogonal to the instruments, markups are accurately estimated.

As for the case of constant output elasticities, we can impose a demand system to shed more light on what might drive a correlation between prices and instruments, and how that affects both the level and dispersion of markup estimates.

Case I: heterogeneous demand elasticities. If firms have heterogeneous demand elasticities and face the previously specified demand system, $p_{it} = -\varepsilon_{it} y_{it}$, variation in markups can be accurately measured if both v_{it} and the instruments are orthogonal to ε_{it}. In that case, we get $(\widehat{\alpha}, \widehat{\beta}) = (1 - \mathbb{E}[\varepsilon_{it}])(\alpha, \beta)$. The markup estimates equal $\mu_{it}^R = (1 - \mathbb{E}[\varepsilon_{it}])\mu_{it}$ and have a correlation of 1 with true markups. Derivations are provided in Appendix A.5. The average markup will be uninformative of the true average, however, as in the case of constant output elasticities.

We should note, however, that v_{it} is not necessarily orthogonal to the demand elasticity ε_{it}. In macroeconomic models such as Atkeson and Burstein (2008) and Kimball (1995), for example, large firms have lower demand elasticities. This means that the correlation between revenue-based markup estimates and true markups may be below 1. Even in those settings, however, the prices and instruments may be close to orthogonal because prices are determined by a combination of returns to scale, a firm's direct competitors, aggregate demand conditions, and factor prices. The extent to which revenue markups and true markups correlate is thus a quantitative question that we answer with simulations in Section 4 and empirically in Section 5. In both, we find high correlations between revenue and true markups, and negligible correlations between prices and instruments.

Case II: homogeneous demand elasticities. Revenue markups are uninformative of true markups, $\widehat{\mu}_{it}^R = 1$, if firms have homogeneous demand elasticities (i.e. $p_{it} = -\gamma y_{it}$). The logic extends from the constant output elasticity case: the average output elasticity from revenue data is biased by the demand elasticity, which also determines the markup. Appendix A.5 provides a formal derivation.

2.2.5 Analyzing Markup Variation without Estimating Production Functions

We should note that there are research designs in which variation in markups can be analyzed without estimating the production function, rendering a lack of quantity data inconsequential (De Loecker and Warzynski 2012, Bond et al. 2021). This still leverages the insight from the production approach to markup estimation that markups are given by the product of the inverse revenue share of a variable input, $P_{it}Y_{it}/(W_tV_{it})$, and that input's output elasticity, α_{it}. If the output elasticity is some function of observables $\mathbf{h}_{it}$, such that $\alpha_{it} = \alpha(\mathbf{h}_{it})$, the output elasticity can be controlled for and variation in the input's revenue share can be used to assess variation in markups. For example, if markups are such that $\ln \mu_{it} = \gamma + \beta X_{it} + \epsilon_{it}$ and β is of interest, one can substitute-out the markup and run

$$\ln P_{it}Y_{it}/(W_tV_{it}) = -\ln \alpha(\mathbf{h}_{it}) + \gamma + \beta X_{it} + \epsilon_{it}.$$

This is feasible, as for many production functions (including translog), the output elasticity is a function of inputs. As $P_{it}Y_{it}/(W_tV_{it})$ is observed in the data, adding controls for inputs in a sufficiently flexible functional form enables one to estimate β without knowledge of the production function. This does have limitations. It is infeasible to control for input usage, e.g., if one is interested in the relationship between markups and firm size. The same holds for other analyses that macroeconomists often perform with markups, such as measuring misallocation, or within-between firm decompositions of markup trends over time.

2.3 Markups, Productivity, and Measurement Errors

After studying how the use of revenue instead of quantity data affects the estimation of markups, we now focus on the case where output data is available but imperfectly measured. Specifically, we now assume that output is observed as $\widetilde{y}_{it} = \alpha v_{it} + \omega_{it} + \eta_{it}$, where $\widetilde{y}_{it}$ is observed quantity (while y_{it} is true quantity), ω_{it} is the productivity observed by the firm, and η_{it} is measurement error and white noise productivity shocks that firms observe after making production decisions.

Prior work pays specific attention to η_{it}, for three reasons. First, observed output can contain significant measurement error. In our empirics, for example, we measure output by dividing revenue by unit values, which are in turn obtained from surveys. Second, the presence of η_{it} impedes estimating true productivity ω_{it}: even if the production function parameters are known, one can only recover productivity with measurement error, $\omega_{it} + \eta_{it}$. Third, measurement errors complicate production function estimation when ω_{it} follows a nonlinear process.

Below, we explain that one can either estimate the production function with greater standard errors, or purge measurement error in a first-stage regression.

Retaining Measurement Errors. In the presence of measurement errors η_{it}, one can still use the standard IV-GMM estimator to estimate the production function (Blundell and Bond 2000). This is the procedure proposed by Doraszelski and Jaumandreu (2019, 2021). In Appendix A.3 we show that, in our simple framework, the estimator remains consistent. Note, however, that for a finite-sized sample, the estimator's asymptotic variance is proportional to $\mathbb{E}\left[\omega_{it}^2\right] + \mathbb{E}\left[\eta_{it}^2\right]$. This means that the estimator's variance increases in measurement errors' variance.

The main drawback is that this procedure cannot recover productivity. To see this, note that productivity is measured as the difference between output and the product of all inputs and their estimated output elasticities. For our framework, this is $\widetilde{y}_{it} - \alpha v_{it} = \omega_{it} + \eta_{it}$. This correlates with true productivity, but the correlation fades as the ratio of variance of measurement error to productivity increases.

In Appendix A.3, we furthermore discuss that 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, Ackerberg and De Loecker 2024). In our quantitative analysis we further explore the impact of abstracting from measurement errors in estimating the production function and markups.

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

The purging regression aims 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 $\omega_{it} = v^{-1}(v_{it}, \Xi_{it})$. This function is often called the control function, as in Olley and Pakes (1996), Levinsohn and Petrin (2003), or Ackerberg et al. (2015). Previous literature on the control function approach is summarized in footnote 7. In our framework, the observed output can therefore be written as $\widetilde{y}_{it} = \alpha v_{it} + v^{-1}(v_{it}, \Xi_{it}) + \eta_{it}$. The fitted values of a non-parametric regression of $\widetilde{y}_{it}$ on v_{it} and Ξ_{it} therefore identify η_{it}, as long as Ξ_{it} contains all variables that determine the demand for v_{it}.[15] The fitted values can then be used to construct moment $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right]$, a function of $\widehat{\alpha}$. The $\widehat{\alpha}$ that sets this moment to zero is a consistent estimator of the true α.

What variables are included in Ξ_{it} under imperfect competition? As noted by De Loecker and Warzynski (2012) and De Loecker and Syverson (2022), the determinants of v_{it} depend on the setting that is considered, including the competition that firms face. In our framework, one can readily find the variables in Ξ_{it} from the firm's cost-minimization problem. Inverting the demand function gives $\omega_{it} = (1-\alpha)v_{it} - mc_{it} + w_t$. It follows that factor prices and (controls for) log marginal costs should be

[15] Our vector Ξ_{it} corresponds to the vector z in De Loecker and Warzynski (2012).

included in Ξ_{it}. Given that marginal costs can be expressed in terms of prices and markups, observed output can be written as[16]

$$\widetilde{y}_{it} = v_{it} - p_{it} + \log \mu_{it} + w_t + \eta_{it}. \tag{9}$$

To purge for measurement error, researchers must thus regress observed output on the variable input, prices, markups, and time-fixed effects for w_t.[17] Under perfect competition, firms are price takers and have log markups of 0. Hence, a first-stage regression of output on v_{it} and time fixed effects is sufficient to purge for measurement error.[18] Under imperfect competition this is not sufficient, as firms have heterogeneous markups. As noted by Doraszelski and Jaumandreu (2019, 2021), controlling for markups is infeasible – as the whole purpose of the exercise is to estimate these markups. This is the so-called inversion problem. We propose resolving this by including price and controls for the markup in the first stage of the procedure. Note that when controlling for markups, 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. One potential control, which we use for the rest of the paper, is market share. We do so because this is consistent with the growing literature in macroeconomics and international trade that builds on Atkeson and Burstein (2008) or Kimball (1995), where markups are determined by market share. When disciplining such models with firm-level markup estimates, inserting market share in a first stage is thus internally consistent. Note that market share is not a perfect control for markup and demand conditions in every case – in many industrial organization models, market share does not control for markups (see, e.g., seminal contributions in empirical IO such as Berry et al. 1995 or Foster et al. 2008).

In summary, the production approach can still be used to estimate markups when output is observed with error. Measurement error affects only estimation of the output elasticity, not the approach in general. One strategy is to purge error in the first stage of a two-stage procedure. Alternatively, the Blundell and Bond (2000) estimator can be used as before, at the expense of greater standard errors. From Section 3, we evaluate these methods with simulations and empirical data.

2.4 Theoretical Extensions

Finally, we use the framework to cover two extensions. Readers initially interested in the main argument may opt to skip these and proceed to Section 3.

[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] For general multi-input, non-Cobb-Douglas cases, the cost-minimization first-order condition is not linear in inputs and cannot be inverted analytically. Nevertheless, the relationship between ω_{it} and inputs, price and markups is well-defined and can be approximated by a polynomial of inputs.

[18] As firms are price takers and have a markup of one, the observed output (after substituting the expression for productivity) under perfect competition reduces to $\widetilde{y}_{it} = v_{it} + w_t + \eta_{it} - p_{it}$. The last two terms are orthogonal to inputs, v_{it}, and input price w_t.

2.4.1 Wedges

Macroeconomists often study markups in conjunction with other distortions that create wedges between an input's marginal products and its factor price. Below, we first discuss the extent to which our tools allow any kind of input wedge to be identified, and then discuss how wedges affect the measurement of markups.

Following Hsieh and Klenow (2009), we introduce two wedges in the framework: a non-markup output wedge τ_{it} and an input wedge λ_{it}. In the context of our analytical framework, these affect the firm's perceived profits $\tilde{\pi}_{it}$ as follows: $\tilde{\pi}_{it} = \tau_{it} P_{it} Y_{it} - \lambda_{it} W_t V_{it}$. Both are wedges because they prevent firms from equating marginal revenue products to factor prices. Rather, optimizing firms equate marginal revenue products to the factor price net of wedges: $\frac{\partial P_{it} Y_{it}}{\partial V_{it}} = W_t \frac{\lambda_{it}}{\tau_{it}}$.

If the variable input is subject to the additional input wedge λ_{it}, our tools identify the complete wedge $\lambda_{it}\mu_{it}$ rather than the markup. Inserting the input price distortion into the cost-minimization problem in Section 2.1, the markup equation (1) in the presence of the input price distortion reads

$$\mu_{it}\lambda_{it} = \alpha_{it} \frac{P_{it} Y_{it}}{W_t V_{it}}. \tag{10}$$

Thus, if markups are measured along (1), the estimated markups are the product of the true markup and the input wedge.[19] Researchers solely interested in measuring markups should thus, if they observe multiple variable inputs, calculate markups using the input that is the least likely to be subject to input wedges.

Output wedge τ_{it} cannot be estimated using our tools and will not distort the measurement of markups, assuming researchers observe actual revenue $P_{it} Y_{it}$ and not $\tau_{it} P_{it} Y_{it}$. Intuitively, (1) is derived from firms' cost-minimization problem rather than from profit maximization, and revenue does not appear in the former. This is why τ_{it} does not appear in (10), either. Output wedges can distort markup estimation if researchers observe $P_{it} Y_{it} \widehat{\tau}_{it}$, where $\widehat{\tau}_{it}$ is an output wedge that directly affects measured revenue. The markup estimate is $\mu_{it} \widehat{\tau}_{it} \lambda_{it} = \alpha_{it} \frac{P_{it} Y_{it} \widehat{\tau}_{it}}{W_t V_{it}}$.

In practice, revenue is usually well-measured and output wedges are implicit, e.g. they reflect the fact that firms face financing or regulatory constraints that prevent them from achieving their optimal scale. Output wedges are therefore less likely to obfuscate markup estimation than input wedges are.[20]

Finally, note that wedges on inputs other than V_{it} do not affect the measurement of markups, but these wedges can be estimated using our tools. To see this, consider the case where the firm maximizes profits $\pi_{it} = P_{it} Y_{it} - W_t V_{it} - \lambda_{it}^K P_t^K K_{it}$, where K_{it} is an additional input, with factor price

[19] Morlacco (2019) uses this property to identify markdowns in the input market for French manufacturing firms. Hashemi et al. (2022) note that if firms have constant markups and when output elasticities are replaced by revenue elasticities, (1) measures input wedges rather than markups.

[20] It is tempting to think that (1) measures the product of all output and input wedges as the marginal product is $\alpha Y_{it}/V_{it}$; hence, the wedge between the marginal product and W_t is $\alpha Y_{it}/(W_t V_{it})$. That wedge does not equal (1), however, as (1) has revenue rather than output in the numerator.

P_t^K, which is subject to the input wedge λ_{it}^K. The new wedge does not alter markup estimates because equation (1) is derived from cost minimization with respect to V_{it}, in which terms with K_{it} drop. One could, however, use the markup equation to identify λ_{it}^K if it was the object of interest. After using the production function estimation tools to identify output elasticity α_{it}^K, the wedge could be calculated as $\lambda_{it}^K = \frac{\alpha_{it}^K}{\alpha_{it}} \frac{P_t^K K_{it}}{W_t V_{it}}$.

2.4.2 Quality Differences

An alternative concern, even with quantity data, is that firms within an industry may produce goods of varying quality. If these firms do not share a common production function, one must estimate separate production functions by (often unobserved) quality. A more common assumption, however, is that quality raises demand, as consumers derive greater utility from better products.[21] We consider the case where quality affects only demand, and where it also impacts costs.

Quality as a demand shifter. If quality affects demand but not marginal costs, as in endogenous growth models, quality differences have limited effects on production function and markup estimation. To see this, assume that demand is a function of price p_{it} and quality q_{it}; i.e. $y_{it} = d(p_{it}, q_{it})$. For ease of exposition, keep the simple production function $y_{it} = \alpha v_{it} + \omega_{it}$, with i.i.d. productivity ω_{it}. If output is measured without error, the standard IV-GMM estimator that relies on moment condition $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0$, where $\widehat{\omega}_{it} = y_{it} - \widehat{\alpha} v_{it}$, still delivers a consistent estimate of α. This is obvious, as the proof for consistency of this estimator makes no assumptions on the demand system that firms face (see Appendix A.1).

Turning to the case where output is measured with an error η_{it}, quality could affect identification of the production function, but only if researchers purge the error using the two-stage procedure. If researchers deploy an estimator that does not purge errors (i.e. Blundell and Bond 2000), standard moment conditions still deliver a consistent estimate of α. This is unchanged from the previous results as, again, the consistency derivations do not rely on assumptions on demand.

Quality differences may be more problematic if researchers purge η_{it}. Doing so requires that firms' demand for the variable input is an invertible function of productivity, $v_{it} = v(\omega_{it}, \Xi_{it})$. If quality affects firms' variable input demand as part of Ξ_{it}, one must control for quality to assure that the function can be inverted, $\omega_{it} = v^{-1}(v_{it}, \Xi_{it})$, to run the purging regression $\widetilde{y}_{it} = \alpha v_{it} + v^{-1}(v_{it}, \Xi_{it}) + \eta_{it}$.

In our analytical framework, additional quality controls are only necessary if quality affects firms' markups. This is because $v^{-1}(v_{it}, \Xi_{it}) = (1-\alpha)v_{it} - p_{it} + \log \mu_{it} + w_t$, as we show in Section 2.3 without making assumptions on demand. Markups may depend on quality (e.g., because high-income households demand higher-quality products), and have lower demand elasticities (Nord 2023). The first-stage regression of observed quantity on v_{it}, price p_{it}, and time-fixed effects for w_t, then needs

[21] Endogenous growth models, for instance, posit that innovation improves product quality, enhancing long-run welfare (e.g., Grossman and Helpman 1991, Aghion et al. 2022).

additional controls for quality to cover the markup μ_{it}. If quality is constant over time – as is commonly assumed in demand estimations following Nevo (2001) – adding firm-fixed effects to the first stage suffices to consistently purge η_{it}. In settings where quality's effect on markups is less easily controlled for, it is prudent to use Blundell and Bond (2000)'s approach instead.

Quality that raises costs and demand. As firms within an industry have common parameters of the production function, quality differences that raise the cost of production are akin to lower productivity: higher-quality products require more inputs to produce the same output quantity. Below we show that when researchers apply standard production function methods, the implicit assumption is that quality-adjusted productivity $\omega_{it} - q_{it}$ has the same stochastic process as ω_{it}.[22] To see this, assume that y_{it} is still a firm's quantity of output, unadjusted for quality. The production function now reads as $y_{it} = \omega_{it} - q_{it} + \alpha v_{it}$.

If output is observed without measurement error, the IV-GMM estimator is consistent if the moment condition that held for productivity in fact holds for the sum of quality and productivity. Formally, the moment condition $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0$ will now have $\widehat{\omega}_{it} = \omega_{it} - q_{it} - (\widehat{\alpha} - \alpha) v_{it}$. The solution to the estimator is given by

$$\widehat{\alpha} = \alpha - \frac{\mathbb{E}\left[q_{it} v_{it-1}\right]}{\mathbb{E}\left[v_{it} v_{it-1}\right]}.$$

The IV-GMM estimator is therefore consistent if quality is orthogonal to the instrument, $\mathbb{E}\left[q_{it} v_{it-1}\right] = 0$, which is the assumption that we usually impose on technological productivity, ω_{it}. If one applies the estimator to firms with heterogeneous cost-raising qualities, one thus assumes that the stochastic process of quality-adjusted productivity is the same as the process assumed for ω_{it} under homogeneous quality. This holds regardless of whether quality affects demand.

If output is measured with error, the IV-GMM estimator can be applied at the expense of greater standard errors. For a two-step estimator, our derivations for cases in which quality is just a demand shifter can readily be applied. There is only a change in labels: where the first stage had to control for productivity, it must now control for quality-adjusted productivity. Inverting variable input demand, we get $\omega_{it} - q_{it} = (1 - \alpha) v_{it} - p_{it} + \log \mu_{it} + w_t$. The purging regression is therefore unchanged from the case in which quality is only a demand shifter.

3 Data

The remainder of the paper scrutinizes the theoretical predictions from Section 2 with rich quantitative simulations and empirical data. We use data on French manufacturing firms, both to quantify

[22] There is a separate body of work on how to consistently measure "technological productivity" ω_{it} in the presence of quality differences. Key contributions include Verhoogen (2008), Kugler and Verhoogen (2012), and De Roux et al. (2021). Hahn (2024) provides a recent review. An alternative setup, where quality raises both demand and input prices, is studied in De Loecker et al. (2016).

the simulations and for the empirical analysis. We combine FARE (*Fichier Approaché des Résultats d'Esane*), which has detailed financial statements, with EAP (*Enquête Annuelle de Production*), which provides product-level revenues and quantities. The sample is from 2009 to 2019. FARE covers the universe of non-financial firms and originates from filings to the tax administration (DGFiP). EAP is a product-level survey by the statistical office (INSEE), covering all manufacturing firms with at least 20 employees or 5 million euros in revenue. We detail the data construction in Appendix G.

We obtain quantities and prices from EAP, which details a firm's revenue and quantity produced at the level of 10-digit products. For each firm-product we calculate the ratio of revenue over the quantity of the product sold. We then standardize this unit value by dividing it by the revenue-weighted average price of the 10-digit product across the sample. As some firms produce multiple products, we define a firm's price as the sales-weighted average of standardized prices across the goods that it produces. Quantity is the ratio of revenue over this price.[23]

We obtain the remainder of the variables from FARE. These are revenue, wage bill, capital, spending on purchased services and spending on purchased materials (the latter including physical intermediate goods and raw materials). Our sample is the intersection of FARE and EAP. We drop firms with missing, zero or negative values for any of the variables, and deflate nominal variables using EU-KLEMS deflators.[24] Summary statistics are given in Appendix Tables 12 and 13.

4 Simulation

We first assess markup estimates in rich quantitative Monte Carlo simulations in a leading macroeconomic model, to scrutinize the estimates in a setting where true markups are known. We use a model that nests and generalizes the analytical framework and that adds an explicit, rich, demand system. We quantify the model using standard parameters and various robustness calibrations in order to assess the quantitative robustness of the analytical conclusions in broad settings.

The simulated model is based on Atkeson and Burstein (2008), where firms face double-nested CES demand and compete à la Cournot. The profit-maximizing markup for firm i in market h at time t is a function of a firm's market share s_{iht}:

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

where ε is the elasticity of substitution within narrow markets, and σ is the elasticity of substitution across markets. Under the assumption that goods are easier to substitute within markets than across markets (e.g. because goods within a market are more similar), this yields that markups rise in firms'

[23]As a robustness check we standardize prices using the revenue-weighted average price at the 8-digit sector level. These firm-level prices have a 0.89 correlation with our baseline prices.

[24]As in the analytical model, we assume that firms within sectors face identical factor prices. De Loecker et al. (2016) offer a markup estimation method for firms with heterogeneous input prices.

market shares.

Firms produce using two inputs one variable, v_{iht}, and one fixed, k_{iht} whose endowment follows an AR(1) process. Firms differ in the quantity of the fixed input at their disposal and in their productivity ω_{iht}, which is also AR(1). Given the fixed input and the productivity, firms choose the variable input to minimize costs. Firms combine these inputs using a translog production function:

$$y_{iht} = \omega_{iht} + \gamma\alpha v_{iht} + \gamma(1-\alpha)k_{iht} + \gamma\frac{\alpha(1-\alpha)}{2}\frac{\phi-1}{\phi}\left(v_{iht}^2 + k_{iht}^2 - 2k_{iht}v_{iht}\right). \tag{11}$$

The translog production function is a second-order approximation of any production function. Function (11) approximates a CES production function with elasticity of substitution ϕ, degree of homogeneity γ, and variable input weight α. It nests the analytical framework's production function if $\alpha = \phi = 1$ and $\gamma \in (0,1)$.

The economy is subject to two aggregate shocks: aggregate demand D_t, which scales up demand at given prices, and factor price W_t, which affects demand for the variable inputs. Both are exogenous AR(1) processes. Finally, the econometrician observes output $\widetilde{y}_{iht}$, which is the sum of y_{iht} and measurement error η_{iht}.

Market share, markups, quantity, and input usage are endogenous and determined in equilibrium. We provide detailed derivations of the production function, the demand system, and equilibrium factor demand in Appendix B, which also details how parameters are chosen based on data from Section 3.

We perform 200 Monte Carlo simulations. Each simulation has 1600 firms, the average firm count in two-digit industries in EAP. We divide firms into 180 markets, the level at which firms compete, and simulate the economy for 40 periods.

4.1 Estimation

Turning to the estimation procedure, we assume that researchers correctly estimate a translog production function with an AR(1) process for productivity

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

where ω_{iht} is productivity which follows $\omega_{iht} = \rho\omega_{iht-1} + \xi_{iht}$, while βs and ρ are to be estimated. We then use the v_{iht}, to estimate markups along equation 1. The output elasticity of v_{iht} is $\frac{\partial y_{iht}}{\partial v_{iht}} = \beta_v + 2\beta_{vv} v_{iht} + \beta_{vk} k_{iht}$, which varies across firms. The estimated production function is consistent with the model's true one (equation 11). The βs satisfy the following relations with the true production parameters: $\beta_v = \gamma\alpha, \beta_k = \gamma(1-\alpha), \beta_{vv} = \gamma\frac{\alpha(1-\alpha)}{2}\frac{\phi-1}{\phi}, \beta_{kk} = \beta_{vv}, \beta_{vk} = -2\beta_{vv}$. Importantly, we do not impose these theoretical relations in the estimation.

We estimate the production function and compare resultant markups across four specifications. The baseline specification uses observed output $\widetilde{y}_{iht}$ in two stages. First, $\widetilde{y}_{iht}$ is purged of measurement error by regressing it on a third-order polynomial of inputs $viht$ and k_{iht}, time-fixed effects, price, and market share. As explained in Section 2.3, the residual from this regression consistently captures measurement error. Purged output is then the dependent variable in an IV-GMM estimation using past inputs as instruments. Appendix C elaborates.

The second specification uses revenue instead of quantity. We run the first stage without price and market share, as researchers with only revenue data would be unable to control for prices. Additionally, this first stage matches the one frequently used in prior work, following Ackerberg et al. (2015) (ACF below).

In the third and fourth specifications, we estimate the production function without a first-stage regression. The third specification uses quantity and follows Blundell and Bond (2000), which is an application of Arellano and Bond (1991) and Blundell and Bond (1998). Essentially, this specification uses lagged first-differences as instruments for equations in levels, in addition to the usual lagged levels as instruments for equations in first-differences. This is the specification recommended by Doraszelski and Jaumandreu (2019, 2021). In the fourth specification, we run the same estimation using revenue as the dependent variable.

4.2 Results

We assess the quality of quantity- and revenue-based markup estimates in two steps. Section 4.2.1 provides a detailed assessment of the production function and markup estimates for a preferred calibration based on the micro data. Section 4.2.2 shows that the assessment is robust to significant parameter changes.

4.2.1 Main Calibration

In the main calibration we quantify the model using parameter values that are either standard in the literature or that directly come from the French data described in Section 3. Full details on the calibration are provided in Appendix B.1.

Production function estimates. We first estimate the production function using both the two-stage baseline specification with quantity data, and the two-stage "ACF" specification with revenue data. Table 1 presents the estimated production function, showing averages across the 200 Monte Carlo simulations. The baseline specification identifies the production function with precision, with all coefficients equal to or within a standard deviation of their true value. Estimates are also consistent across simulations, as indicated by low standard deviations. The average elasticity of output to the variable input, which determines the average estimate of the markup, equals 0.294, and is close to the true average of 0.295.

Table 1: Estimated Production Function Parameters

	$\beta_v = \alpha\gamma$	$\beta_k = (1-\alpha)\gamma$	$\beta_{vv=\gamma\frac{\alpha(1-\alpha)}{2}\frac{\phi-1}{\phi}}$	$\beta_{kk} = \beta_{vv}$	$\beta_{vk} = -2\beta_{vv}$
True value	0.32	0.48	0.009	0.009	-0.017
Quantity	0.32	0.48	0.009	0.009	-0.018
(Baseline)	(0.002)	(0.001)	(.0004)	(.0005)	(0.001)
Revenue	0.29	0.31	0.005	-0.001	-0.016
(ACF)	(0.002)	(0.003)	(.0022)	(.0037)	(0.005)

NOTE: "True value" coefficients are directly calculated from the calibrated parameters (α, γ, ϕ). "Baseline": A two-stage IV-GMM on observed quantity. "ACF": A two-stage IV-GMM on revenue. Production function coefficients are averaged across 200 Monte Carlo simulations. Standard deviations across the Monte Carlo simulations are given in parentheses.

The "ACF" specification does not identify the production function coefficients, as expected when using revenue instead of quantity data. The average estimate for β_v of 0.29 is slightly lower than the true value of 0.32. The other coefficients affecting the elasticity of output to the variable input, β_{vv} and β_{vk}, are both closer to zero than their true value, although the higher standard errors show that these coefficients are estimated with considerable variability. More strikingly, the coefficient for the fixed input, β_k, falls from 0.48 to 0.31, and the one for the fixed input squared, β_{kk}, flips sign. The latter are not used to compute the output elasticity of v, so they do not affect the average estimate of the markup. The average estimated output elasticity of v is 0.296, close to the true value of 0.295.

These estimates are consistent with our analytical results. Section 2.2.2 showed that revenue-based estimates of the production function can be biased upwards, or downwards, or be unaffected, depending on the correlation between prices and inputs. Section 2.2.3 showed that downward-sloping demand curves cause a downward bias on the estimated elasticity in the absence of demand shocks. Our simulated firms *are* subject to aggregate demand shocks, which create a positive correlation between input usage and prices under diminishing returns to scale, limiting the impact of the bias coming from downward-sloping demand.

Markup estimates. We next use these coefficients to estimate markups. The results are summarized in Table 2. We calculate the correlations with true (log) markups for each simulation and present the average, with standard deviations in parentheses. The baseline estimates are close to true markups. The correlation is close to one, and the mean, standard deviation, median and interquartile range are estimated to within a tenth of a decimal point. The slight deviations between true markups and estimated markups are in line with the modest differences between the true and estimated coefficients in Table 1, and may be caused, for example, by the fact that the first-stage regression approximates the implicit relationship between productivity and inputs through a third-order polynomial.

The bottom panel of Table 2 describes the markup estimates based on revenue data. These are highly informative of true markups, with a correlation of 0.94 between the true markup and the revenue-based "ACF" markups. These results again show that the revenue-based estimates of the

Table 2: Overview - Markup Estimates

	Correlation with true markups	Log Markup Moments			
		Mean	St. Dev.	Median	IQR
True markups	1.00	0.23	0.07	0.21	0.09
Quantity (baseline)	1.00 (0.01)	0.22 (.016)	0.07 (.001)	0.21 (.016)	0.09 (.002)
Revenue (ACF)	0.94 (0.06)	0.23 (.020)	0.07 (.011)	0.21 (.022)	0.08 (.015)

NOTE: The first column presents estimates' correlations with true markups. Subsequent columns show moments of the estimated (log) markup distribution. Standard deviations across 200 Monte Carlo simulations are in parentheses. "Baseline": A two-stage IV-GMM on observed quantity. "ACF": A two-stage IV-GMM on revenue.

production function elasticities are *not* the revenue elasticities of an input. If they had been, log markups should equal 0 and be uninformative of true markups (Section 2.2.1). Rather, the revenue-based elasticities are *biased* estimates of output elasticities of the inputs.

From Section 2.2.4, it follows that a high correlation between true markups and revenue-based estimates requires that prices have low correlations with the production function instruments. Appendix Table 6 confirms that correlations are low in the simulations, and that the low correlations are consistent with the data.

Figure 1 provides a graphical illustration of the high correlation between revenue-based markup estimates and true markups. It presents a binned scatter plot between the baseline and "ACF" estimates in log-levels (left-hand panel) and in log first-differences (right-hand panel). The plots confirm that quantity-based and revenue-based markups are tightly linked over the entire distribution, especially in first-differences. The regression coefficients are 0.85 and 0.99, respectively, and the linear fit in first-differences resembles a perfect 45-degree line.

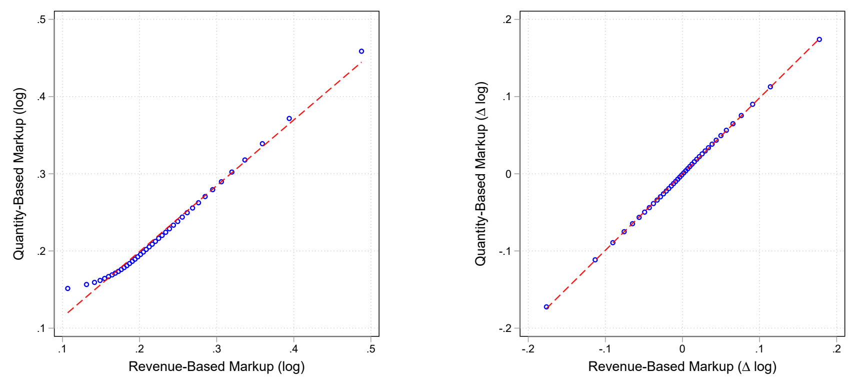
Markup correlations. We next examine how markup estimates correlate with variables from seminal regressions such as markups on the profits rate (operating profit over sales), the material share (ratio of variable-input spending over sales) and market share. For both revenue and quantity-based estimates, we run

$$x_{it} = \chi(\ln \widehat{\mu}_{it}) + \varphi_i + \psi_t + \epsilon_{it}, \tag{12}$$

where φ_i and ψ_t denote firm- and time effects. We divide markup estimates by their standard deviations to ease comparison of rows. The idea is to check whether these regressions are similar for markups based on revenue and quantity data. If so, this supports prior work using revenue markups in such regressions.

Table 3 presents the results, confirming that revenue-based markup estimates do well at retriev-

Figure 1: Binned Scatter Plot for Simulated Quantity and Revenue Markups



NOTE: The figures plot the relationship between quantity-based markups ("baseline") and revenue-based markups ("ACF") in simulated data. Log-markups are used in panel (a), log-differenced markups in panel (b). Linear regression coefficients are 0.85 and 0.99, respectively. The scatters are averages across Monte Carlo simulation.

ing the OLS coefficient χ for true markups. Table 3 shows that if researchers are unaware of biases from the use of revenue data, relationships between markups and other variables remain well-estimated. Market share's R^2 is particularly high, as expected in an Atkeson and Burstein (2008) model.

As we note in Section 2.2.5, there is an alternative approach to avoid biases in regressions such as these. As the output elasticity of v_{it} is pinned down by the interaction of sector-fixed effects and input usage, one could add these as controls to the regression, as long as they are not the object of interest. The remaining markup variation then comes from v_{it}'s revenue share. That is the approach advocated by Bond et al. (2021) and De Loecker and Warzynski (2012).

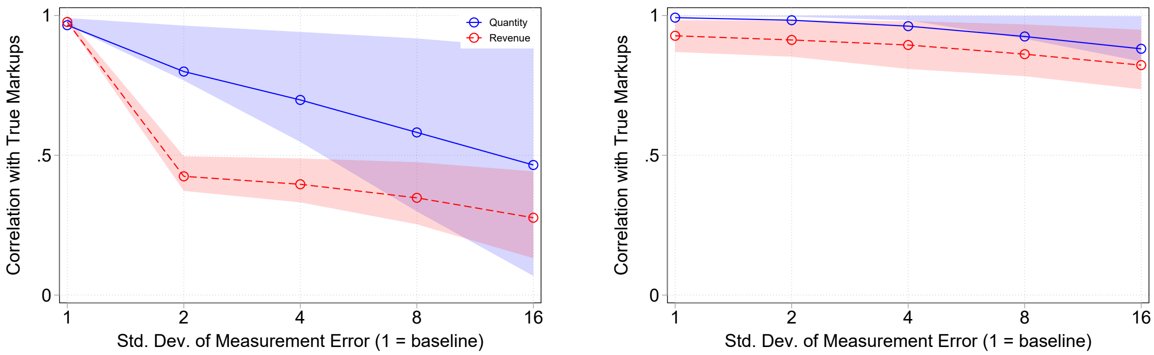
Measurement error. Thus far we have purged output from measurement error through a first-stage

Table 3: Simulated Relation between Markup Estimates and Other Variables

	Profit Rate	Material Share	Market Share
True Markups	0.0192***	-0.0192***	0.0683***
	(0.0001)	(0.0001)	(0.0001)
R-Squared	0.901	0.901	0.996
Quantity (Baseline)	0.0197***	-0.0197***	0.0700***
	(0.0001)	(0.0001)	(0.0001)
R-Squared	0.899	0.899	0.995
Revenue (ACF)	0.0216***	-0.0216***	0.0758***
	(0.0001)	(0.0001)	(0.0001)
R-Squared	0.903	0.903	0.966

NOTE: OLS coefficients. Explanatory variables are in the column headers. "Baseline": IV-GMM on quantity. "ACF": IV-GMM on revenue. Details in Section 4.1. Markups are normalized to have unit standard deviations. Firm-clustered standard errors in parentheses. *** denotes 1% level significance. All specifications include time- and firm-fixed effects. OLS coefficients, standard errors and R^2s are averages across the Monte Carlo simulations.

Figure 2: Effect of First-Stage Purging Regression on Markup Estimates



Without First-Stage Purging (Blundell-Bond)

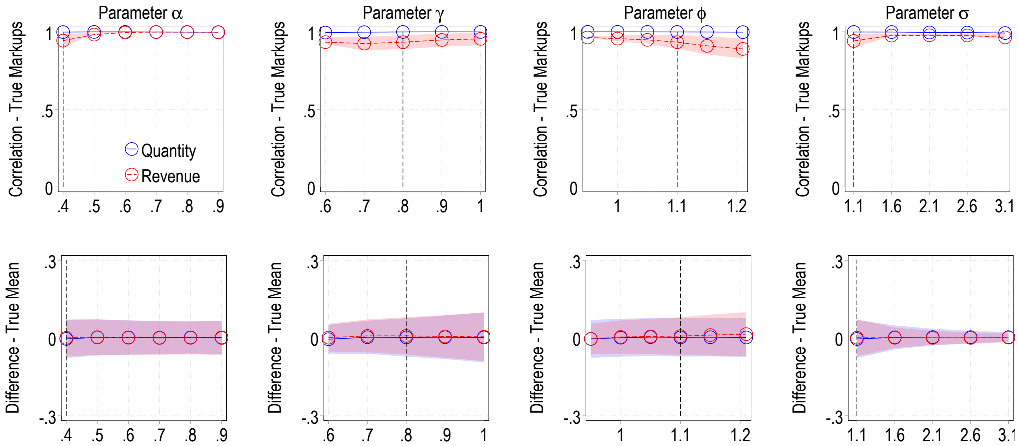
With First-Stage Purging (Baseline/ACF)

NOTE: Figures plot the average correlation between true markups and quantity (blue-solid) or revenue (red-dashed) based markup estimates by standard deviation of measurement error (normalized to 1 for the main calibration). Confidence intervals are the interquantile range. The left-hand panel uses the single-stage Blundell and Bond (2000) estimator; the right-hand panel uses the two-stage baseline and ACF estimator.

regression. There are settings where purging is not feasible, however, for example when the variables that determine firms' variable input demand are unknown or unobserved (Doraszelski and Jaumandreu 2019, 2021), or when there are quality differences between products (see Section 2.4.2). As explained in Section 2.3, the Blundell and Bond (2000) estimator consistently estimates the production function without purging, at the expense of greater asymptotic variance. We can use the simulations to assess the impact of this higher variance on the markup estimates at various levels of measurement error. In our calibration, 9.5% of the variation in quantity comes from measurement error, consistent with the French data (see Appendix B.1). Below we estimate markups using the Blundell and Bond estimator with the baseline measurement error, and also subsequently when the error's standard deviation is 2, 4, 8 or 16 times larger.

Figure 2 presents the results. At the baseline measurement error, Blundell and Bond (2000)'s estimator in the left-hand panel performs very well. The average correlation of markup estimates is close to one, regardless of whether quantity (blue-solid) or revenue (red-dashed) data is used. For higher levels of measurement error, the two-stage procedures in the right-hand panel perform better. Even when the error's standard deviation is 16 times the baseline level, the quantity estimates have a correlation with true markups of 0.88, and the "ACF" revenue estimates have a correlation with true markups of 0.82. Not purging for the error lowers those correlations to 0.47 and 0.28, respectively. It is thus advisable to use a two-stage procedure in settings with significant measurement error. The Blundell and Bond alternative performs well in settings with low measurement error.

Figure 3: Revenue and Quantity Markups for Alternative Parameters



NOTE: Figures plot quantity-based (Baseline; blue-solid) and revenue-based (ACF; red-dashed) markup estimates. Vertical lines give baseline values for each column's parameter. Upper figures: correlation between estimates and true markups, interquantile range in shaded areas. Bottom figures: average deviation of estimated markups' mean from true mean; shaded areas give standard deviations of that deviation across Monte Carlo simulations.

4.2.2 Robustness

Results for the main calibration are robust to substantial changes in the parameters. We perturb all the production function parameters (α, γ, ϕ), and the elasticity of substitution across markets (σ). We vary the variable input's share in production (α) from 0.4 to 0.9. We raise this parameter because in the baseline (α of 0.4) the average output elasticity is 0.3, on the low end of the sector averages of 0.3 to 0.8 that we find in Section 5 below. We vary returns to scale (γ) from 0.6 to 1.0, symmetrically around the baseline calibration of 0.8. Elasticity of substitution ϕ is 1.10 in the baseline, and we vary this parameter from 0.95 to 1.2.[25] Finally, $\sigma > 1$ governs the dispersion of markups. We raise it from the standard value in the literature of 1.1 to 3.1, which reduces true markup dispersion by 78%.

Figure 3 presents the results. The top plots show that the correlation between markups and quantity-based estimates is close to 1 across estimations, while the revenue-based estimates have a correlation with true markups of at least 0.88. This shows that results in Section 4.2.1 are robust to alternative parametrizations.

[25] Oberfield and Raval (2021) find a 0.8 elasticity of substitution. We use a higher lower bound, because for low values the Atkeson and Burstein (2008) model becomes hard to solve, as marginal costs are highly non-linear in productivity. There is no degradation in the estimators' performance.

5 Empirics

We now turn to markup estimation in the EAP-FARE data. True markups are unknown, but we can still assess whether markup estimates based on revenue data are similar to estimates based on quantity data. There are a series of challenges. First, quantities are approximated by the ratio of revenue over relative prices, as explained in Section 3. Second, firms are unlikely to have identical production functions and, more generally, the translog production function and the AR(1) productivity process may not accurately reflect true production processes. The errors that this causes are specific to each estimation and may differ for quantity and revenue. Despite these complications, we show below that revenue- and quantity-based markups are highly correlated within industries, especially in first-differences, and have reasonably similar variances. We also show that revenue and quantity-based markups have similar relationships with key variables such as the profit rate, labor, materials, and the market share.

5.1 Production Function Estimation

In contrast to the single sector in the simulations and the analytical framework, the empirical analysis has 18 two-digit manufacturing industries, summarized in Table 13. We assume that firms only share parameters of the production function and productivity process within industries, and thus perform a separate estimation by industry. Specifically, we assume that log output y_{it} is a translog function of the log of materials m_{it}, the wage bill l_{it}, capital k_{it} and service inputs o_{it}:

$$y_{it} = \omega_{it} + \beta^I_m m_{it} + \beta^I_l l_{it} + \beta^I_k k_{it} + \beta^I_o o_{it} + \sum_{\{h,j\} \in \{m,l,k,o\}} \sum \beta^I_{hj} h_{it} j_{it}, \quad \forall i \in I \tag{13}$$

where productivity ω_{it} follows AR(1) process $\omega_{it} = \rho^I \omega_{it-1} + \xi_{it}$ for two-digit sector I. We assume that materials, m_{it}, correspond to variable input v_{it} in Section 2. To estimate markups, we are therefore interested in output elasticity α^m_{it}:

$$\alpha^m_{it} \equiv \partial y_{it} / \partial m_{it} = \beta^I_m + 2\beta_{mm} m_{it} + \beta^I_{mo} o_{it} + \beta^I_{ml} l_{it} + \beta^I_{mk} k_{it}.$$

Note that firms within an industry do not have the same output elasticities, α^m_{it}, as the elasticity depends on the level of each input that the firm uses.

5.2 Results

As in the simulations, we estimate production functions using a two-stage procedure with either quantity (baseline) or revenue (ACF) as the dependent variable.[26] Estimated output elasticities av-

[26] The appendix also provides results for the single-stage Blundell and Bond (2000) estimator.

erage 0.54 across all sectors in the baseline estimation, with a standard deviation of 0.29. Table 10 lists separate averages by two-digit industry. The average output elasticity of materials is 0.40 using revenue, with a standard deviation of 0.14. Average quantity-based elasticities are higher than revenue-based elasticities in 16 of 18 industries. On average, quantity elasticities exceed revenue-based elasticities by 38%.

Markup Estimates. We next use these estimated firm-level elasticities to compute markups along equation (1). 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, and perform the analysis on firms for which all estimates fall within the non-trimmed sample, leaving 147,403 observations.[27]

To make the empirical results comparable to the simulations and the analytical framework, we compare revenue- and quantity-based markups within industries. Table 4 presents the estimates' moments. Each moment is calculated at the sector level and subsequently averaged across industries. The table presents unweighted averages, but results are similar when sectors are weighted by size.

We find an average log markup of 0.37 using quantity data and 0.13 for revenue data. That gap wedge is consistent with the lower estimated output elasticities from revenue data. In contrast to the simulation, we also find a gap between the standard deviation of revenue-based markups and quantity-based markups. On average, quantity-based markups have a standard deviation of 0.23, while revenue-based markups have a standard deviation of 0.16. The medians and percentiles show that the lower dispersion of revenue-based markups is not driven by the tails. In the analytical framework the standard deviations would be equal, because the production function is Cobb-Douglas. In Table 2's simulation, however, revenue markups had similar standard deviations for sizeable deviations from Cobb-Douglas. It is possible that various empirical complications explain the wedge: the production function and productivity process are unlikely to be exactly equal across firms within an industry, and may also not be exactly equal to equation (13). Below, we show that despite these limitations, overall correlations between revenue and quantity-based markups are high.

Markup Correlations Table 5 presents the correlation between revenue- and quantity-based markup estimates in the data. The main panel summarizes sector-level correlations, which are directly comparable to the simulations in Table 2. The average simple correlation for log markups is 0.61, while the Spearman rank correlation is 0.62. The high within-sector correlations are visible in the majority of sectors: median correlations exceed average correlations, and consistently exceed 0.65 for levels

[27] The sample reduction is largely because of high variance of the Blundell and Bond (2000) elasticities. These are negative for a non-negligible fraction of firms, for whom log markups are not defined.

Table 4: Summary Statistics - Within-Sector Moments of the Log Markup Estimates

	Mean	St. Dev.	Median*	25th Pct.*	75th Pct.*
Quantity (baseline)	0.366	0.232	-0.011	-0.149	0.137
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Revenue (ACF)	0.128	0.163	-0.015	-0.113	0.099
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)

NOTE: Moments are calculated at the sector level and then averaged across the sectors in the data. Parentheses present bootstrapped standard errors of the moment. "Baseline": two-stage IV-GMM on quantity. "ACF": two-stage IV-GMM on revenue. 147,704 obs. *: Medians and percentiles are expressed in deviation of sector averages.

and 0.84 for first-differenced markups.[28] Turning to first-differences, the correlations increase to 0.80 for the simple correlation and 0.83 for the rank correlation. This is close to the 0.94 correlation in the simulations, despite the complications of estimating production functions in the data.

The final column of Table 5 also present correlations when all firms are pooled. As expected, combining heterogeneous sectors lowers the correlations: the Pearson correlation is 0.34 while the rank correlation is 0.43. Yet the correlation of growth in markups remains high: 0.80 for the Pearson correlation and 0.84 for rank correlations. Growth in revenue markups thus seems a good predictor of growth in quantity-based markups, even when comparing firms across sectors.

To provide an alternative illustration of the extent to which revenue-based markups successfully predict in quantity-based markups, we next assess whether the linear regression of quantity-based markups on revenue-based markups has a coefficient close to 1. If so, $\widehat{\mu}_{it}^{R}$ will on average equal quantity-based markup estimates up to a sector effect. We plot results in binned scatter plots in Figure 4, both in levels (left) and first-differences (right). The linear regression coefficients equal 0.89 and 0.92, respectively. Both show an excellent fit between the estimates across values of the revenue-based markups, with the linear fit nearing a 45° line for first-differences. This is quantitatively similar to the corresponding simulation plots (Figure 1), where regression coefficients were 0.85 and 0.99.

Finally, we assess whether relationships between markups and key variables depend on the markup

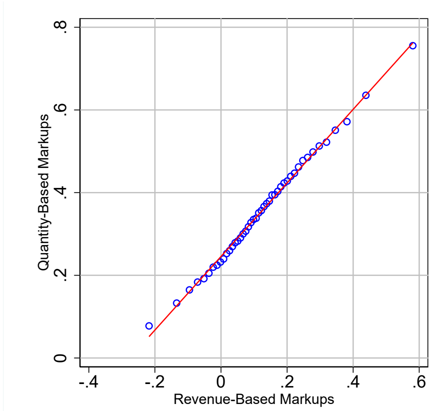
Table 5: Correlation between Quantity- and Revenue-Based Log Markup Estimates

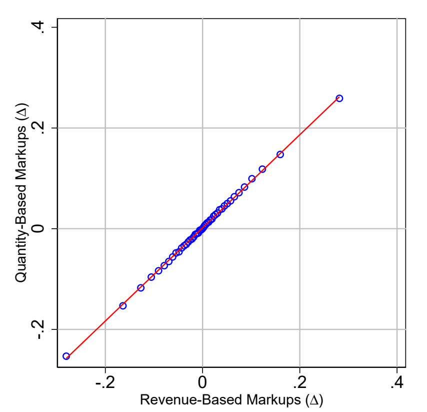
	Summary of Within-Sector Correlations					Pooled
	Average	St. Dev.	Median	25th Pct.	75th Pct.	Overall
Pearson Correlation	0.61	(0.23)	0.65	0.41	0.78	0.34
Pearson Corr. First Diff.	0.80	(0.16)	0.84	0.70	0.94	0.79
Rank Correlation	0.62	(0.23)	0.68	0.39	0.79	0.43
Rank Corr. First Diff.	0.83	(0.12)	0.84	0.71	0.95	0.84

NOTE: The table presents correlations between revenue- and quantity-based markup estimates. Correlations are calculated at the two-digit sector in the left-hand panel, and for all observations pooled in the right-hand panel. The summary statistics are averages of the correlations across sectors. 147,704 observations.

[28]Specifications based on the Blundell and Bond (2000) estimator in Appendix E are qualitatively similar with, in general, lower positive correlations. This might be due to the fact that prices, and thus quantities, may be subject to considerable error in our data. As we show in Section 4.2, two-stage estimators are more robust in such settings.

Figure 4: Binned Scatter Plot: Quantity- versus Revenue-Based Markups





Notes: The figures plot the binned scatter plot between quantity-based markups (baseline) and revenue-based markups (ACF). Regression coefficients for the linear fit are 0.89 and 0.91, respectively.

specification, as we did for the simulations in Table 3. We regress revenue- or quantity-based markups on either the profit rate (the ratio of operating profits over sales), labor share (the ratio of its wage bill over sales), material cost share (the ratio of materials purchased over sales), and market share.

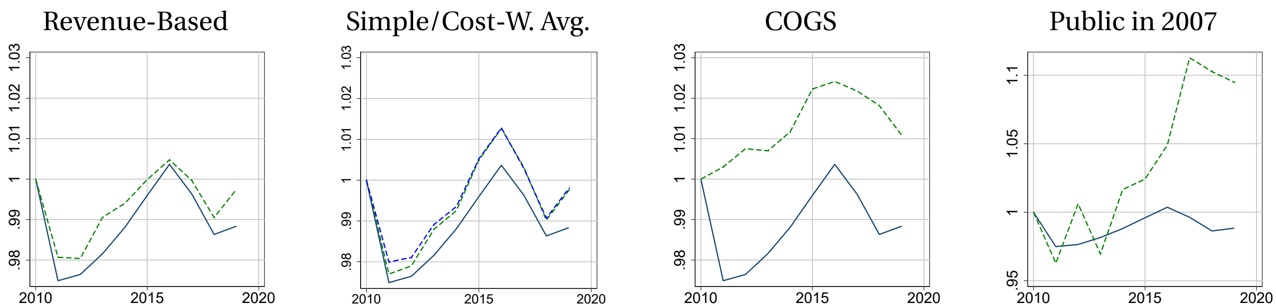
Results are presented in Table 6. Rows present regression coefficients for a markup estimate, columns contain the regressor. We pool all observations and normalize the markups to have unit standard deviations. All relationships run in the expected direction: high markup firms are more profitable, have lower labor shares, lower material shares, and greater market shares, irrespective of whether revenue or quantity markups are used. Moreover, the relationships are all significant at the 1% level. Overall, Table 6 suggests that estimates of relationships between markups and key variables are qualitatively robust to the use of revenue-based markup estimates. This lends credibility to prior work using revenue-based markups in regressions and further supports our conclusion that these estimates contain useful information about firms' true markups.

Table 6: Simulated Relation between Markup Estimates and Other Variables

	Profit Rate	Labor Share	Material Share	Market Share
Quantity (baseline)	0.150***	-0.050***	-0.096***	0.070***
	(0.002)	(0.001)	(0.001)	(0.006)
R-Squared	0.357	0.098	0.312	0.005
Revenue (ACF)	0.111***	-0.027***	-0.077***	0.031***
	(0.001)	(0.001)	(0.001)	(0.004)
R-Squared	0.502	0.075	0.501	0.005

NOTE: Each entry provides the OLS estimate using the variable in the column header as the dependent variable and the markup series in the row as the regressor. Firm-clustered standard errors in parentheses. *** denotes significance at 1% level, respectively. All regressions include time- & firm-fixed effects. Observations: 147,704.

Figure 5: Trend in Aggregate Markups in France (index = 1 in 2010)



NOTE: Solid-blue lines plot aggregate (harmonic revenue-weighted) markups based on quantity. Green dashed lines, from left to right, plot: aggregate markups based on revenue; aggregate markups based on quantity that are simple rather than harmonic weighted; aggregate markups based on quantity with COGS as the variable input; aggregate markups based on quantity data for publicly listed firms. Blue-dash-dotted lines in the second plot: aggregate markups based on quantity that are cost weighted. Markup levels are not shown; all are indexed to 1.0.

6 Aggregate Markups

Finally, we discuss the robustness of trends in aggregate markups. De Loecker et al. (2020) show a significant rise in average firm-level markups since the 1980s for U.S. Compustat firms. This influential result has raised several concerns, mostly about the use of Compustat: (i) revenue is used to proxy for quantity, our main subject of analysis; (ii) they measure trends in simple revenue-weighted averages rather than harmonic or cost-weighted averages, which may not be welfare-relevant aggregate markups; (iii) they assume Cost of Goods Sold (COGS) is the variable input, which may be too broad (Basu 2019); and (iv) Compustat covers public firms, which are unrepresentative of the entire economy.

These choices are defensible with the data limitations faced by De Loecker et al. (2020), but we are able to analyse the quantitative importance of these critiques for aggregate markup trends in France. For the baseline, we define the aggregate markup as the sales-weighted harmonic average of our main quantity-based markup estimates, $\left(\sum_{i \in I_t} s_{it}\mu_{it}^{-1}\right)^{-1}$, where I_t is the set of sampled firms at time t, while s_{it} denotes firm i's share in aggregate sales. This measure is the welfare-relevant measure of the aggregate markups in a broad set of models (see, e.g., Edmond et al. 2023; Grassi 2017). We then deviate from this baseline in four ways: (i) using the revenue-based markups of the "ACF" specification, (ii) using a simple or cost-weighted average rather than a harmonic sales-weighted average, (iii) using the sum of materials and labor expenditure, a proxy for COGS, to estimate production functions and markups, and (iv) computing the aggregate markup using our baseline quantity-based markups only for firms that were publicly listed in 2007 – which is the final year for which we observe listed status.

Figure 5 shows the resulting time series for aggregate markups, where we investigate trends by normalizing each series to one in 2010. The first panel shows that quantity-based and revenue-based aggregate markups follow similar dynamics: they exhibit a decline around the Eurocrisis in 2011 and 2012, followed by an upward trend. Their average level is different, the aggregate revenue-based markups averaging 1.08, while aggregate quantity markups average 1.45. The second panel shows similar movements for the cost-weighted average, the simple sales-weighted average, and our baseline aggregate markup. The third panel shows that when COGS is used as the variable input, markups increase around 3%-points more over the sample. The average of the COGS aggregate markup is lower than the baseline, at 1.30. The fourth panel shows that the aggregate quantity-based markup for public firms jumps by around 20 points in the later part of the sample. Note, however, that our sample contains an average of only 38 public firms per year, whose aggregate markup averages 1.81.

We conclude that trends in markups of French firms between 2010 and 2019 are robust to the use of revenue and the type of aggregation, while both of these changes strongly affect the level of the aggregate markup.

7 Conclusion

This paper assesses the feasibility of estimating markups from widely used data in the macroeconomics and international trade literature. Practically, we conclude that it depends on the research design whether an analysis can proceed with revenue data. If feasible, production functions for markups should be estimated with quantity data. Given the paucity of such data, we show that revenue data can suffice for researchers interested in dispersion of markups across firms. Revenue data may also be used to estimate trends, if one is willing to assume a constant production function over time. Conversely, in applications that focus on the average level of the markup, revenue data is not appropriate.

There are a number of caveats. Our markup estimates are contingent upon correct specification of the production function. While both our simulation and empirical analysis rely on the translog function, which is a second-order approximation of any production function, approximation error will affect markup estimates. Our approach to measure markups will also capture other input wedges, such as markdowns due to monopsony power, if the variable input is subject to them. A quantification of markups' contribution to overall input wedges, for example with external markup estimates, is an exciting avenue for future research.

References

Ackerberg, D., Frazer, G., Luo, Y., Su, Y., et al. (2020). Under-identification of structural models based on timing and information set assumptions. *Working Paper*.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Ackerberg, D. A. and De Loecker, J. (2024). Production function identification under imperfect competition. Technical report, Mimeo, University of Leuven.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P., and Li, H. (2022). Good rents versus bad rents: R&d misallocation and growth. Technical report, Working Paper.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application. *The Review of Economic Studies*, 58(2):277–297.

Atalay, E., Frost, E., Sorensen, A. T., Sullivan, C. J., and Zhu, W. (2023). Scalable demand and markups. Technical report, National Bureau of Economic Research.

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

Baqaee, D., Farhi, E., and Sangani, K. (2021). The supply-side effects of monetary policy. Working Paper 28345, National Bureau of Economic Research.

Baqaee, D. R. and Farhi, E. (2019). Productivity and Misallocation in General Equilibrium. *The Quarterly Journal of Economics*. qjz030.

Basu, S. (2019). Are price-cost markups rising in the united states? a discussion of the evidence. *Journal of Economic Perspectives*, 33(3):3–22.

Basu, S. and Fernald, J. G. (1997). Returns to scale in u.s. production: Estimates and implications. *Journal of Political Economy*, 105(2):249–83.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Blundell, R. and Bond, S. (2000). Gmm estimation with persistent panel data: an application to production functions. *Econometric reviews*, 19(3):321–340.

Boar, C. and Midrigan, V. (2019). Markups and inequality.

Bond, S., Hashemi, A., Kaplan, G., and Zoch, P. (2021). Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics*, 121:1–14.

Brand, J. (2019). Estimating productivity and markups under imperfect competition.

Burstein, A., Carvalho, V. M., and Grassi, B. (2020). Bottom-up Markup Fluctuations. NBER Working Papers 27958, NBER.

Calligaris, S., Criscuolo, C., and Marcolin, L. (2018). Mark-ups in the digital era. *OECD*.

Carvalho, V. M. and Grassi, B. (2019). Large firm dynamics and the business cycle. *American Economic Review*, 109(4):1375–1425.

Chiavari, A., Morazzoni, M., and Smirnov, D. (2021). Heterogeneous markups cyclicality and monetary policy. Technical report, Pompeu Fabra University.

Collard-Wexler, A. and De Loecker, J. (2020). Production function estimation and capital measurement error. *National Bureau of Economic Research Working Paper*.

De Loecker, J. (2011). Recovering markups from production data. *International Journal of Industrial Organization*, 29(3):350–355.

De Loecker, J. (2021). Comment on (un) pleasant... by bond et al (2020). *Journal of Monetary Economics*, 121:15–18.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications*. *Quarterly Journal of Economics*, 135(2):561–644.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

De Loecker, J. and Syverson, C. (2022). An industrial organization perspective on productivity. In *Handbook of industrial organization*, volume 4, pages 141–223. Elsevier.

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102(6):2437–71.

De Ridder, M. (2024). Market power and innovation in the intangible economy. *American Economic Review*, 114(1):199–251.

De Roux, N., Eslava, M., Franco, S., and Verhoogen, E. (2021). Estimating production functions in differentiated-product industries with quantity information and external instruments.

Decker, R. A., Haltiwanger, J., Jarmin, R., and Miranda, J. (2020). Changing business dynamism and productivity: Shocks vs responsiveness. *American Economic Review*, 110(12):3952–90.

Döpper, H., MacKay, A., Miller, N., and Stiebale, J. (2022). Rising markups and the role of consumer preferences. *Working Paper*, (22-025).

Doraszelski, U. and Jaumandreu, J. (2019). Using cost minimization to estimate markups.

Doraszelski, U. and Jaumandreu, J. (2021). Reexamining the De Loecker & Warzynski (2012) method for estimating markups. CEPR Discussion Papers 16027, CEPR.

Díez, F. J., Fan, J., and Villegas-Sánchez, C. (2021). Global declining competition? *Journal of International Economics*, 132:103492.

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, markups, and the gains from international trade. *American Economic Review*, 105(10):3183–3221.

Edmond, C., Midrigan, V., and Xu, D. Y. (2023). How costly are markups? *Journal of Political Economy*, 131(7):1619–1675.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1):394–425.

Gandhi, A., Navarro, S., and Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016.

Gaubert, C. and Itskhoki, O. (2021). Granular comparative advantage. *Journal of Political Economy*, 129(3):871–939.

Grassi, B. (2017). Io in i-o: Size, industrial organization and the input-output network make a firm structurally important. *Working Paper*.

Grieco, P. L., Murry, C., and Yurukoglu, A. (2024). The evolution of market power in the us automobile industry. *The Quarterly Journal of Economics*, 139(2):1201–1253.

Grossman, G. M. and Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European economic review*, 35(2-3):517–526.

Hahn, N. (2024). Product differentiation and quality in production function estimation.

Hall, R. E. (1986). Market structure and macroeconomic fluctuations. *Brookings papers on economic activity*, 1986(2):285–338.

Hall, R. E. (1988). The relation between price and marginal cost in u.s. industry. *Journal of Political Economy*, 96(5):921–947.

Hashemi, A., Kirov, I., and Traina, J. (2022). The production approach to markup estimation often measures input distortions. *Economics Letters*, 217:110673.

Hong, S. (2017). Customer Capital, Markup Cyclicality, and Amplification.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Kimball, M. S. (1995). The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking*, 27(4):1241–1277.

Klette, T. and Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11:343–361.

Kouvavas, O., Osbat, C., Reinelt, T., and Vansteenkiste, I. (2021). Markups and inflation cyclicality in the euro area. Working Paper Series 2617, European Central Bank.

Kugler, M. and Verhoogen, E. (2012). Prices, plant size, and product quality. *The Review of Economic Studies*, 79(1):307–339.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2):317–341.

Meier, M. and Reinelt, T. (2022). Monetary Policy, Markup Dispersion, and Aggregate TFP. *The Review of Economics and Statistics*, pages 1–45.

Miller, N., Osborne, M., Sheu, G., and Sileo, G. (2023). Technology and market power: The united states cement industry, 1974-2019. *Georgetown McDonough School of Business Research Paper*, (4041168).

Mongey, S. (2017). Market Structure and Monetary Non-neutrality. Working paper.

Morlacco, M. (2019). Market power in input markets: Theory and evidence from french manufacturing. *Working Paper*.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342.

Nord, L. (2023). Shopping, demand composition, and equilibrium prices. *Working Paper*.

Oberfield, E. and Raval, D. (2021). Micro data and macro technology. *Econometrica*, 89(2):703–732.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–97.

Raval, D. (2023a). A flexible cost share approach to markup estimation. *Economics Letters*, 230:111262.

Raval, D. (2023b). Testing the production approach to markup estimation. *Review of Economic Studies*, 90(5):2592–2611.

Syverson, C. (2019). Macroeconomics and market power: Context, implications, and open questions. *Journal of Economic Perspectives*, 33(3):23–43.

Traina, J. (2018). Is aggregate market power increasing? production trends using financial statements. Working Paper 17, Stigler Cente.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *The Quarterly Journal of Economics*, 123(2):489–530.

Wang, O. and Werning, I. (2022). Dynamic oligopoly and price stickiness. *American Economic Review*, 112(8):2815–49.

Yeh, C., Macaluso, C., and Hershbein, B. (2022). Monopsony in the us labor market. *American Economic Review*, 112(7):2099–2138.

“The Hitchhiker’s Guide to Markup Estimation”
Appendix - For Online Publication Only

A Theory Appendix

In this appendix, we first outline how the output elasticity of V_{it} is estimated. We start from the ideal case in which a researcher observes prices, such that output can be measured by quantity (A.1). The main text discusses the case where price is not observed. We then discussed the small sample properties of the estimator (A.2) and the case of measurement errors (A.3). We also show that the main results extend to more general frameworks (A.4) with a translog production function (A.4.1), several inputs (A.4.2), with persistent productivity (A.4.3), and with all of this together (A.4.4). We generalize our results on markup estimation when revenue is used in place of quantity for translog production functions (A.5).

A.1 Identification with Price and Quantity Data

We here cover the estimation of α if revenue, prices, and 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 (3) 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 focus in this section on the second stage – as it performs the actual production function identification. We introduce measurement errors and a first stage in Section A.3.

Absent measurement error, the production function is $y_{it} = \alpha v_{it} + \omega_{it}$. Least-square regressions of v_{it} on output y_{it} will be biased, as 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 2 *The instrumental var. 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 0 where* $\widehat{\omega}_{it} = y_{it} - \widehat{\alpha} v_{it} = (\alpha - \widehat{\alpha}) v_{it} + \omega_{it}$.

It is straightforward to solve for $\widehat{\alpha}$ in closed form by substituting $\widehat{\omega}_{it}$ into the moment condition:[29] $(\alpha - \widehat{\alpha})\mathbb{E}\left[v_{it} v_{it-1}\right] = 0$, 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 $\widehat{\alpha}$ thus converges to the true elasticity α.

[29] 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.

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.[30] The cost-minimizing firm's first-order condition for v_{it} summarizes the candidate drivers: $v_{it} = (1-\alpha)^{-1}(\omega_{it} + mc_{it} - w_t)$.

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.[31]

A.2 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.

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

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.3 Adding 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 occur

[30] When we generalize our setup in Appendix A.4.3 (where productivity is persistent, e.g. a linear AR(1) process with coeff. 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.

[31] 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).

after input v_{it} is set. The measurement error is log-additive and denoted by η_{it}. All firms produce along $\widetilde{y}_{it} = \alpha v_{it} + \omega_{it} + \eta_{it}$, where $\widetilde{y}_{it}$ denotes observed output. 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 IV-GMM estimator is defined as follows:

Definition 3 *The IV-GMM estimator is $\widehat{\alpha} \in \mathbb{R}$ so that moment $\mathbb{E}\left[(\widehat{\omega_{it} + \eta_{it}}) v_{it-1}\right]$ is equal to zero, where $\widehat{\omega_{it} + \eta_{it}} = \widetilde{y}_{it} - \widehat{\alpha} v_{it} = (\alpha - \widehat{\alpha}) v_{it} + \omega_{it} + \eta_{it}$.*

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 $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0$. 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 is still consistent and 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).

There are three advantages to purging the observed quantity from measurement errors. 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.2 yields that the asymptotic variance of the estimator is $\mathbb{V}ar\left[\widehat{\alpha}\right] \sim \frac{\mathbb{E}[v_{it-1}^2]}{n\mathbb{E}[v_{it} v_{it-1}]^2}\left(\mathbb{E}\left[\omega_{it}^2\right] + \mathbb{E}\left[\eta_{it}^2\right]\right)$, 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.

Third, measurement error can also impede the 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 have to be slightly altered to consistently estimate α (see Appendix A.4.3). 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 innovation 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 AR(1) innovation. 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. 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. 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.

A.4 Extensions

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

A.4.1 Translog Production Function

We first ease the assumption that output is log-linear by replacing the Cobb-Douglas production function with a translog specification: $y_{it} = \alpha v_{it} + \beta v_{it}^2 + \omega_{it}$. The other assumptions are unchanged. We assume that quantity is observed in this section. We leave the discussion of unobserved quantity and its implications for markup estimation to the dedicated Appendix A.5. Our aim is to identify α and β, in order to calculate size-dependent output elasticities of the variable input for the calculation of true markups $\mu_{it} = (\alpha + 2\beta v_{it})(P_{it}Y_{it})/(W_t V_{it})$. 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 translog production involves estimating a multivariate GMM regression with instrumental variables. Formally, we define the estimator as:

Definition 4 *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 remains simple to solve for the estimator $(\widehat{\alpha}, \widehat{\beta})$ in our parsimonious setting. This involves solving the system of equations implied by the moment conditions:

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

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

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

If the determinant of V is not zero, the GMM estimator on translog is identified and consistent, i.e. $\widehat{\alpha} = \alpha$ and $\widehat{\beta} = \beta$.

A.4.2 Several Inputs

In this extension, we assume that firms produce with two inputs: a variable input v_{it} and another input k_{it}. We assume that k_{it} is, in the terminology of the production function literature, dynamic. This means that firms face adjustment costs and other inter-temporal constraints when setting k_{it}, which leads firms to choose k_{it} before observing current productivity, i.e. $\mathbb{E}\left[\omega_{it} k_{it}\right] = 0$. The production function in logs reads $y_{it} = \alpha v_{it} + \beta k_{it} + \omega_{it}$. We are interested in estimating (α,β). As k_{it} is set before productivity is observed, we only need to instrument the variable input with its lag. The estimation is therefore a GMM regression with one endogenous and one exogenous variable. For quantity, the estimator is

Definition 5 *The GMM estimator is a pair* $(\widehat{\alpha}, \widehat{\beta})$ *so that* $\mathbb{E}\left[\widehat{\omega}_{it} v_{it-1}\right] = 0$ *and* $\mathbb{E}\left[\widehat{\omega}_{it} k_{it}\right] = 0$*, 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:

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

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

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

As long as the determinant of V is not zero, the GMM estimator is identified and asymptotically consistent such that $\widehat{\alpha} = \alpha$ and $\widehat{\beta} = \beta$.

Using Revenue Instead of Quantity. When revenue, denoted r_{it} in log, is used as a proxy for quantity, the estimator can be defined as follows:

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

The estimator is the solution of the following system of equations:

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

which admits a unique solution if the determinant of V is not zero. If the latter is satisfied, this unique solution estimator is $\widehat{B} = B + V^{-1} P$, where we denote the vector $P = \left(\mathbb{E}\left[p_{it} v_{it-1}\right], \mathbb{E}\left[p_{it} k_{it}\right]\right)'$. As in the simple framework, the bias comes from the correlation of price with the instruments. This estimator will be asymptotically non-consistent if either $\mathbb{E}\left[p_{it} v_{it-1}\right]$ or $\mathbb{E}\left[p_{it} k_{it}\right]$ are different from zero.

A.4.3 Persistent Productivity

In this extension, we assume that total factor productivity follows a first-order autoregressive (AR1) process in logs. We assume that quantity is observed. We leave the discussion on the case when revenue is used in place of quantity to the general proof in Appendix A.4.4. The production function is still $y_{it} = \alpha v_{it} + \omega_{it}$, while the productivity process is $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$. Below we define the GMM estimator $(\widehat{\alpha}, \widehat{\rho})$ using v_{it-1} and $\widehat{\omega}_{it-1}$ as an instrument for v_{it} and $\widehat{\omega}_{it}$.

Definition 7 *The GMM estimator is a pair* $(\widehat{\alpha}, \widehat{\rho})$ *so 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* $\widehat{\xi}_{it} = \widehat{\omega}_{it} - \widehat{\rho}\widehat{\omega}_{it-1} = \xi_{it} + (\alpha - \widehat{\alpha})(v_{it} - \rho v_{it-1}) + (\rho - \widehat{\rho})\omega_{it-1} + (\rho - \widehat{\rho})(\alpha - \widehat{\alpha})v_{it-1}$.

The estimator, $(\widehat{\alpha}, \widehat{\rho})$, is characterized by the following system of equations defined by the moment conditions:

$$\begin{aligned}
&\mathbb{E}\left[\widehat{\xi}_{it} v_{it-1}\right] = 0 \\
&\mathbb{E}\left[\widehat{\xi}_{it}\widehat{\omega}_{it-1}\right] = 0 \quad \Longleftrightarrow \\
&(\alpha - \widehat{\alpha})\mathbb{E}\left[(v_{it} - \rho v_{it-1})v_{it-1}\right] + (\rho - \widehat{\rho})\mathbb{E}\left[\omega_{it-1}v_{it-1}\right] + (\alpha - \widehat{\alpha})(\rho - \widehat{\rho})\mathbb{E}\left[v_{it-1}^2\right] = 0 \\
&(\alpha - \widehat{\alpha})\mathbb{E}\left[(v_{it} - \rho v_{it-1})\omega_{it-1}\right] + (\rho - \widehat{\rho})\mathbb{E}\left[\omega_{it-1}^2\right] + (\alpha - \widehat{\alpha})(\rho - \widehat{\rho})\mathbb{E}\left[v_{it-1}\omega_{it-1}\right] = 0.
\end{aligned}$$

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. Below we formally discussed this case, but first, 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

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

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 (Def. 7) is characterized by the system of equations

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

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

where $X = \alpha - \widehat{\alpha}$, $Y = \rho - \widehat{\rho}$, and, $a = \mathbb{E}\left[(v_{it} - \rho v_{it-1})v_{it-1}\right]$, $b = \mathbb{E}\left[\omega_{it-1}v_{it-1}\right]$, $c = \mathbb{E}\left[v_{it-1}^2\right]$, $d =$

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

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

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 $\widetilde{v}_{it}\equiv v_{it}-\rho v_{it-1}=\frac{1}{1-\alpha}\left(\xi_{it}+mc_{it}-\rho mc_{it-1}+w_t-\rho w_{t-1}\right)$.[32]

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.

A.4.4 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 moments and products of several inputs. This formulation nests the Cobb-Douglas and Translog case. For example, a two-inputs v_{it},k_{it} translog production function is modelled by $X_{it}=(v_{it},k_{it},v_{it}^2,k_{it}^2,v_{it}k_{it})'$ with parameters $\beta=(\beta_v,\beta_k,\beta_{vv},\beta_{kk},\beta_{vk})'$. 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 estimator that we study here is defined as follows:

Definition 8 *The GMM estimator is $\widehat{\beta}\in\mathbb{R}^N$ and $\widehat{\rho}\in\mathbb{R}$, so 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 $\widehat{\xi}_{it}=\widehat{\omega}_{it}-\widehat{\rho}\widehat{\omega}_{it-1}=(X_{it}-\rho X_{it-1})'(\beta-\widehat{\beta})+X_{it-1}'(\beta-\widehat{\beta})(\rho-\widehat{\rho})+\omega_{it-1}(\rho-\widehat{\rho})+\xi_{it}$.*

The remainder of this appendix studies 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 $\widehat{\rho}$:

[32] 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 values of productivity, this correlation will be equal to zero.

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

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

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 can be written in matrix form:

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

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

Using Revenue Instead of Quantity. The estimator for revenue is as follows:

Definition 9 *The GMM estimator is* $\widehat{\beta}\in\mathbb{R}^N$ *and* $\widehat{\rho}\in\mathbb{R}$ *such that the moments* $\mathbb{E}\left[X_{it-1}\widehat{\varsigma}_{it}\right]$ *and* $\mathbb{E}\left[\widehat{\mathrm{tfpr}}_{it-1}\widehat{\varsigma}_{it}\right]$ *are equal to zero where* $\mathrm{tfpr}_{it}\equiv\omega_{it}+p_{it}$, $\widehat{\mathrm{tfpr}}_{it}=r_{it}-X'_{it}\widehat{\beta}=X'_{it}(\beta-\widehat{\beta})+\mathrm{tfpr}_{it}$ *and* $\widehat{\varsigma}_{it}=\widehat{\mathrm{tfpr}}_{it}-\widehat{\rho}\widehat{\mathrm{tfpr}}_{it-1}=(X_{it}-\rho X_{it-1})'(\beta-\widehat{\beta})+X'_{it-1}(\beta-\widehat{\beta})(\rho-\widehat{\rho})+\mathrm{tfpr}_{it-1}(\rho-\widehat{\rho})+p_{it}-\rho p_{it-1}+\xi_{it}$.

Using the same notation as above and $\widetilde{p}_{it}=p_{it}-\rho p_{it-1}$, the system of equations that characterized the above estimator is given by

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

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

where we use the fact that prices at $t-1$ are unrelated to the innovation ξ_{it} and thus that $\mathbb{E}\left[\mathrm{tfpr}_{it-1}\,\xi_{it}\right]=\mathbb{E}\left[\omega_{it-1}\xi_{it}\right]+\mathbb{E}\left[p_{it-1}\xi_{it}\right]=0$. In general, this system of equations admits multiple solutions, as we show in the case of one input in Appendix A.4.3. For a heuristic proof, we abstract from the higher order terms in $(\beta-\widehat{\beta})(\rho-\widehat{\rho})$ which we consider small when $\widehat{\beta}$ and $\widehat{\rho}$ are not too far from their true values. In that case, the system of equations that characterized the estimator can be written in a matrix form as $W(B-\widehat{B})+R=0$, where $B-\widehat{B}=\left(\beta-\widehat{\beta},\rho-\widehat{\rho}\right)'$;

$$R=\left(\begin{array}{c}\mathbb{E}\left[X_{it-1}\widetilde{p}_{it}\right]\\ \mathbb{E}\left[\text{tfpr}_{it-1}\,\widetilde{p}_{it}\right]\end{array}\right)\text{ and, }W=\left(\begin{array}{cc}\mathbb{E}\left[X_{it-1}\widetilde{X}'_{it}\right] & \mathbb{E}\left[X_{it-1}\omega_{it-1}\right]\\ \mathbb{E}\left[\omega_{it-1}\widetilde{X}'_{it}\right] & \mathbb{E}\left[\omega^2_{it-1}\right],\end{array}\right)+\left(\begin{array}{cc}0 & \mathbb{E}\left[X_{it-1}p_{it-1}\right]\\ 0 & \mathbb{E}\left[p_{it-1}\,\text{tfpr}_{it-1}\right]\end{array}\right),$$

which has a solution $\widehat{B}=B+W^{-1}R$. As in the simple framework, the bias is due to the correlation of price (adjusted for persistence), $\widetilde{p}_{it}$, with past input X_{it-1} and tfpr_{it-1} collected in the vector R.

A.5 Revenue Markup and Translog Production Function

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

Assume that the production function is $y_{it}=\alpha v_{it}+\beta v^2_{it}+\omega_{it}$, 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 (especially as in Appendix A.4.2), the coefficients of the production function estimated on revenue are such that

$$\left(\begin{array}{c}\widehat{\alpha}\\ \widehat{\beta}\end{array}\right)=\left(\begin{array}{c}\alpha\\ \beta\end{array}\right)+V^{-1}\left(\begin{array}{c}\mathbb{E}[p_{it}v_{it-1}]\\ \mathbb{E}[p_{it}v^2_{it-1}]\end{array}\right),\qquad\text{with}\qquad V=\left(\begin{array}{cc}\mathbb{E}[v_{it}v_{it-1}] & \mathbb{E}[v^2_{it}v_{it-1}]\\ \mathbb{E}[v_{it}v^2_{it-1}] & \mathbb{E}[v^2_{it}v^2_{it-1}]\end{array}\right).$$

As for Cobb-Douglas, the estimates are biased. The above equation is the translog equivalent of (4). Correlation of instruments with output prices causes bias.

In the case of a translog production function, the true markup is such that $\mu_{it}=(\alpha+2\beta\log V_{it})\frac{P_{it}Y_{it}}{W_tV_{it}}$, and, the revenue markup is thus $\widehat{\mu}^R_{it}=\frac{\widehat{\alpha}+2\widehat{\beta}\log V_{it}}{\alpha+2\beta\log V_{it}}\mu_{it}$. 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.[33] However, in general, revenue markups are different from one and contain information on true markups. To see this formally, 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

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

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

$$\widehat{\mu}^R_{it}=\left[1-(\alpha+2\beta\log V_{it})^{-1}\left(\begin{array}{c}\alpha\\ \beta\end{array}\right)'\left(\mathbb{E}\left[d_{iit}X_{it}X'_{it-1}\right]\mathbb{E}\left[X_{it}X'_{it-1}\right]^{-1}\right)\left(\begin{array}{c}1\\ 2\log V_{it}\end{array}\right)\right](1-d_{iit})^{-1}.\qquad(14)$$

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

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 assumptions in place, one can show that $\widehat{\alpha} = \alpha(1 - \mathbb{E}\left[d_{iit}\right])$ and $\widehat{\beta} = \beta(1 - \mathbb{E}\left[d_{iit}\right])$. The revenue markup is equal to $\widehat{\mu}_{it}^R = (1 - \mathbb{E}\left[d_{iit}\right])(1 - d_{iit})^{-1}$, which is different from one since there exists 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 (14), we can see that the average of the log revenue markup is equal to zero up to a Jensen-like inequality:

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

When the inverse demand elasticities are homogeneous, $\forall i, d_{iit} = \gamma$, then the average log revenue markup is exactly zero. In general, 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 and Parametrization of the Simulated Model

B.1 Model and Parametrization

We analyze a single sector, defined as a collection of firms that have the same structural production function parameters and that face the same input prices.

Demand. We choose a market structure where firms 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 continuum of markets, where a market is defined as a finite number of firms that compete oligopolistically with one another. In this setup, the demand faced by market h, Y_{ht}, satisfies at sector price index P_{ht}:

$$P_{ht} = Y_{ht}^{-\frac{1}{\sigma}} D_t^{\frac{1}{\sigma}} \quad \text{with} \quad Y_{ht} = \left[\sum_{i=1}^{N_h} Y_{iht}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{15}$$

where σ denotes the elasticity of substitution across markets, D_t is exogenous aggregate demand, and market-level output Y_{ht} is the aggregate of firm-level output across the N_h firms that operate in h. Y_{iht} denotes the output of firm i and ε is the elasticity of substitution across goods within a market. Following Atkeson and Burstein (2008), we assume $\varepsilon > \sigma$, reflecting that it is easier to substitute goods across firms than across markets. The inverse demand function for firm *i* is

$$P_{iht} = \left(\frac{Y_{iht}}{Y_{ht}} \right)^{-\frac{1}{\varepsilon}} P_{ht}, \tag{16}$$

where P_{iht} is the price of firm *i* and P_{ht} satisfies the market-level inverse demand (equation 15). Under Cournot competition, firm i in market h maximizes profit by choosing its quantity taking other firms' quantity as given – subject to the inverse demand given by the above equation (16). The quantity-setting firm internalizes that Y_{ht} increases and P_{ht} decreases when it raises its own quantities according to equation (15). The resultant profit-maximizing markup reads as

$$\mu_{iht} = \frac{\varepsilon}{\varepsilon - 1} \left(1 - \frac{\frac{\varepsilon}{\sigma} - 1}{\varepsilon - 1} s_{iht} \right)^{-1} \quad \text{with} \quad s_{iht} = \frac{P_{iht} Y_{iht}}{P_{ht} Y_{ht}}, \tag{17}$$

where s_{iht} is the market share in terms of revenue.[34] Markups range from $\varepsilon/(\varepsilon - 1)$ for a firm whose market share approaches zero, to $\sigma/(\sigma - 1)$ for a monopolist.

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:

$$y_{iht} = \omega_{it} + \gamma \alpha v_{iht} + \gamma (1 - \alpha) k_{iht} + \gamma \frac{\alpha(1 - \alpha)}{2} \frac{\phi - 1}{\phi} \left(v_{iht}^2 + k_{iht}^2 - 2 k_{iht} v_{iht} \right), \tag{18}$$

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. When $\phi = 1$, this production function nests the Cobb-Douglas specification. Our log production function (18) is motivated by an approximation around $\phi = 1$ of the constant elasticity of substitution production function $Y_{iht} = e^{\omega_{iht}} (\alpha V_{iht}^{\frac{\phi-1}{\phi}} + (1 - \alpha) K_{iht}^{\frac{\phi-1}{\phi}})^{\frac{\phi}{\phi-1}\gamma}$ (see Appendix B.2).

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 $W_t V_{iht}$ subject to the

[34] The Bertrand markup is similar (e.g. Atkeson and Burstein 2008, Grassi 2017, Burstein et al. 2020).

production function (18). Note that the output elasticity of the variable input is

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

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

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

, where MC_{iht} is the marginal cost which can rewritten as

$$V_{iht} = \left(\frac{MC_{iht}}{W_t}\right)\gamma\alpha\left(1 + [1-\alpha]\frac{\phi-1}{\phi}\ln\left[\frac{V_{iht}}{K_{iht}}\right]\right)Y_{iht}, \tag{19}$$

Marginal costs. As firms face an exogenous sequence of the fixed input K_{iht}, marginal costs can be derived from the production function (18) and optimal demand for the variable input (19). Inserting the latter into the former, we get $y_{iht} = \omega_{iht} + \gamma\alpha\ln\left[\left(\frac{MC_{iht}}{W_t}\right)\gamma\alpha\left(1-[1-\alpha]\frac{\phi-1}{\phi}\ln\left[\frac{K_{iht}}{V_{iht}}\right]\right)Y_{iht}\right] + \gamma(1-\alpha)k_{iht} + \gamma\alpha\frac{1-\alpha}{2}\frac{\phi-1}{\phi}\left[\ln\left(\frac{K_{iht}}{V_{iht}}\right)\right]^2$. Isolating log marginal costs on the left-hand side, we can express the log marginal costs $mc_{iht} \equiv \ln MC_{iht}$as

$$mc_{iht} = \ln\left[\frac{W_t}{\gamma}Y_{iht}^{\frac{1-\alpha\gamma}{\alpha\gamma}}\Omega_{iht}^{-\frac{1}{\alpha\gamma}}K_{iht}^{\frac{\alpha-1}{\alpha}}\right] - \ln\left(1+[1-\alpha]\frac{\phi-1}{\phi}\ln\left[\frac{V_{iht}}{K_{iht}}\right]\right) + \frac{1-\alpha}{2}\frac{\phi-1}{\phi}\left(\ln\left[\frac{V_{iht}}{K_{iht}}\right]\right)^2. \tag{20}$$

Equilibrium. We consider an 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}, output Y_{iht}, log marginal costs mc_{iht}, market shares s_{iht}, log variable inputs v_{iht}, and market-level output Y_{ht} and price P_{ht} such that price is equal to markup times marginal cost, the demand is satisfies (equations 15 and 16), quantities are set to maximize profit (equations 17), and the variable input is chosen to minimize cost (equations 19 and 20 in Appendix B.2).

Calibration. We perform 200 Monte Carlo simulations. In each simulation, we model the behavior of 1600 firms, which is the average number of firms in two-sector industries in the EAP data. We divide these firms into 180 markets, the level at which firms compete, and simulate the economy for 40 periods.

There are 13 parameters, each of which we calibrate externally. The parameters are summarized in Table 7. 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 interest in 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 sources of firm heterogeneity: the firm's log-endowment of the fixed input k_{iht} and

Table 7: 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 $P_t^{-\sigma} Y_t$
$\rho^\omega, \sigma^\omega$	0.60 , 0.20	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

the firm's log-total factor productivity ω_{it}. Both evolve exogenously through log-linear AR(1) 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 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 log capital in the data. We find a persistence parameter ρ^k of 0.66 and a volatility of shocks σ_k of 0.66. We set ρ^ω to 0.6 and set volatility σ^ω to 0.2, in line with Decker et al. (2020) and Carvalho and Grassi (2019).

There are two aggregate shocks: aggregate demand D_t and the variable input price W_t. We assume both series follow a log-linear AR(1) process with persistence ρ_D and ρ_W, respectively, and shocks $\xi_D \sim N(0, \sigma_D)$ and $\xi_W \sim N(0, \sigma_W)$. To calibrate the process for the variable inputs price, we estimate an AR(1) process for the price index of intermediate inputs from sector-level manufacturing data in EU-KLEMS. We run simple AR(1) 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. For aggregate demand D_t we estimate a similar autoregressive process, using the detrended sector-level nominal value added as the dependent variable.[35] 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.

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.

We introduce measurement error in observed quantity $\widetilde{y}_{iht}$, denoted by η_{iht}, after computing the

[35]We detrend D_t using nominal GDP to account both for increases in prices and real output to obtain a stationary nominal series. Results are similar when detrending with the GDP deflator.

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 quantity on prices, market share, time-fixed effects and a third-degree polynomial in the firms' inputs in EAP.

B.2 Translog Approximation

This appendix derives the translog production function (equation 18) as an approximation of a CES production function around $\phi = 1$ with homogeneity of degree γ that is $Y_{iht} = \Omega_{iht}(\alpha[V_{iht}]^{\frac{\phi-1}{\phi}} + (1-\alpha)[K_{iht}]^{\frac{\phi-1}{\phi}})^{\frac{\phi}{\phi-1}\gamma}$. To see this, rearrange terms

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

where we move the α back into the log term for the last equality. Rewriting the final term yields $\frac{\phi}{\phi-1}\gamma \ln\left[1 + (1-\alpha)\left(\left(\frac{K_{iht}}{V_{iht}}\right)^{\frac{\phi-1}{\phi}} - 1\right)\right]$; let us then define $f(x) = \frac{\gamma}{x}\ln\left[1 + (1-\alpha)\left(B^x - 1\right)\right]$, where $B = K_{iht}/V_{iht}$ and $x = (\phi-1)/\phi$, such that our approximation is around $x \to 0$. Taking a second-order approximation yields

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

where we remove higher-order terms given that we are approximating the function up to a second order. Hence, the first-order approximation of the generalized CES production function reads $y_{iht} = \omega_{iht} + \gamma \ln V_{iht} + \gamma(1-\alpha)\ln\left(\frac{K_{iht}}{V_{iht}}\right) + \gamma\alpha\frac{1-\alpha}{2}\frac{\phi-1}{\phi}\left[\ln\left(\frac{K_{iht}}{V_{iht}}\right)\right]^2$. After rearranging terms and denoting by small cap letters the log of a variable, $x \equiv \ln X$, we have the translog production function (18) with homogeneity of degree γ: $y_{iht} = \omega_{iht} + \gamma\alpha v_{iht} + \gamma(1-\alpha)k_{iht} + \gamma\alpha\frac{1-\alpha}{2}\frac{\phi-1}{\phi}\left(v_{iht}^2 + k_{iht}^2 - 2k_{iht}v_{iht}\right)$.

C Implementation of the Estimation

This appendix describes how we implement the production function estimation. Let us assume that the observed output of firm i at time t is such that $\widetilde{y}_{it} = X'_{it}\beta + \omega_{it} + \eta_{it}$, 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. We assume further that η_{it} is a measurement error shock, such that actual output is $y_{it} = \widetilde{y}_{it} - \eta_{it} = X'_{it}\beta + \omega_{it}$. Productivity, ω_{it}, follows a Markov process. We assume that we have access to a sample of observed output $\widetilde{y}_{it}$ and input usage X_{it}. Here we assume that we additionally observe price p_{it} and controls for markups s_{it}.

The first stage consists of purging the observed output from measurement errors. As explained in Section 2.3, we do so by running a regression of measure output $\widetilde{y}_{it}$ on time-fixed effects, a polynomial of inputs usage of some order (second or third in practice), price p_{it}, and the additional controls s_{it}.[36] We then compute an estimate of the measurement errors $\widehat{\eta}_{it}$ as the difference of observed output $\widetilde{y}_{it}$ and the fitted value $\widehat{y}_{it}$, which we interpret as true output.

The estimator of the parameters $\widehat{\beta}$ is delivered by a numerical algorithm that makes moments equal to zero. These moments are computed as follows. For a given guess of parameters $\widehat{\beta}$, we compute $\widehat{\omega}_{it} = y_{it} - \widehat{\beta}X_{it}$. We then estimate the Markov process: in the case of AR(1), we obtain $\widehat{\rho}$ by an OLS regression. The estimates of the innovation of the AR(1) process are given by $\widehat{\xi}_{it} = \widehat{\omega}_{it} - \widehat{\rho}\widehat{\omega}_{it-1}$. The set of instruments Z_{it} is chosen from the vector of input usage X_{it} or its lag for dynamic and static input, respectively (see Appendix A.4.2 for an example). Finally, we compute the moments as $\sum_{i,t} \widehat{\xi}_{it} Z_{it}$.

D Convergence

We assess the convergence properties of the estimators. In particular, we examine the speed at which our markup estimates converge to their true values as sample size increases. In the baseline calibration, we simulate 1600 firms per sector, matching the average sector size in France. However, many sectors have fewer firms in practice. To determine if markups can be reliably estimated in smaller samples, we repeat our estimations for samples ranging from 150 to 1600 firms per year, increasing in increments of 50. Results indicate that precision in estimating the average log markup improves significantly up to approximately 500-600 firms, beyond which additional increases in sample size yield limited improvements. Since most sectors in administrative datasets surpass this threshold, accurately estimating markups appears feasible with administrative datasets that are typically available.

[36] When price and quantity are not observed, revenue is used in place of observed output; we do not include the extras controls p_{it} and s_{it} in the first-stage regression.

E Additional Tables and Figures

Table 8: Summary Statistics - Within-Sector Moments of the Log Markup Estimates

	Mean	St. Dev.	Median*	25th Pct.*	75th Pct.*
Quantity	-0.003	0.536	0.04	-0.312	0.354
	(0.008)	(0.006)	(0.008)	(0.011)	(0.009)
Revenue	0.072	0.210	-0.026	-0.150	0.124
	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)

NOTE: Moments are calculated at the sector level and then averaged across the sectors in the data. Single-stage markup estimates. Parentheses present bootstrapped standard errors of the moment. 147,704 obs. *: Medians and percentiles are expressed in deviation of sector averages.

Table 9: Correlation between Quantity and Revenue-Based Markups Log Markups

	Summary of Within-Sector Correlations					Pooled
	Average	St. Dev.	Median	25th Pct.	75th Pct.	Overall
Pearson Correlation	0.40	(0.35)	0.45	0.19	0.62	0.12
Pearson Corr. First Diff.	0.55	(0.29)	0.58	0.36	0.78	0.41
Rank Correlation	0.40	(0.29)	0.50	0.25	0.65	0.12
Rank Corr. First Diff.	0.60	(0.17)	0.60	0.40	0.80	0.54

NOTE: The table presents correlations between single-stage revenue- and quantity-based markup estimates. Correlations are calculated at the two-digit sector in the left-hand panel, and for all observations pooled in the right-hand panel. The summary statistics are averages of the correlations across sectors. 147,704 observations.

Table 10: Estimated Material-Output Elasticity for Various Specifications by Sector

Industry	NACE	Quantity		Revenue	
		Baseline	BB-Q	ACF	BB-R
All (average)	-	0.54	0.38	0.40	0.39
		(0.29)	(0.30)	(0.15)	(0.14)
Manufacturing of textiles	13	0.45	0.22	0.41	0.44
		(0.23)	(0.14)	(0.16)	(0.13)
Manufacturing of wearing apparel	14	0.46	0.36	0.27	0.28
		(0.25)	(0.16)	(0.20)	(0.14)
Manufacturing of leather and related products	15	0.38	0.34	0.30	0.26
		(0.22)	(0.20)	(0.19)	(0.14)
Manufacturing of wood and products of wood	16	0.58	0.64	0.47	0.49
		(0.15)	(0.24)	(0.14)	(0.14)
Manufacturing of paper and paper products	17	0.52	0.35	0.46	0.45
		(0.19)	(0.09)	(0.12)	(0.09)
Manufacturing of printing and reproduction	18	0.34	0.19	0.41	0.41
		(0.20)	(0.12)	(0.13)	(0.13)
Manufacturing of chemicals and chemical products	20	0.68	0.38	0.48	0.40
		(0.40)	(0.24)	(0.16)	(0.12)
Manufacturing of rubber and plastic products	22	0.60	0.41	0.45	0.49
		(0.18)	(0.11)	(0.13)	(0.15)
Manufacturing of other non-metallic mineral products	23	0.53	0.50	0.39	0.34
		(0.13)	(0.25)	(0.15)	(0.10)
Manufacturing of basic metals	24	0.62	0.72	0.43	0.43
		(0.19)	(0.15)	(0.19)	(0.17)
Manufacturing of fabricated metal products	25	0.42	0.28	0.37	0.31
		(0.19)	(0.41)	(0.15)	(0.11)
Manufacturing of computer-, electronic products	26	0.71	0.26	0.39	0.34
		(0.31)	(0.31)	(0.13)	(0.11)
Manufacturing of electrical equipment	27	0.62	0.42	0.46	0.48
		(0.22)	(0.17)	(0.14)	(0.14)
Manufacturing of machinery and equipment	28	0.37	0.53	0.43	0.36
		(0.21)	(0.38)	(0.13)	(0.09)
Manufacturing of motor vehicles	29	0.81	0.65	0.52	0.53
		(0.21)	(0.22)	(0.17)	(0.14)
Manufacturing of furniture	31	1.15	0.31	0.39	0.38
		(0.15)	(0.34)	(0.10)	(0.06)
Manufacturing of other	32	0.43	0.24	0.30	0.30
		(0.28)	(0.14)	(0.13)	(0.12)
Repair and installation of machinery and equipment	33	0.31	0.14	0.31	0.31
		(0.13)	(0.14)	(0.11)	(0.11)

NOTE: Estimated elasticities of materials on output from the estimation of translog production functions. The headers, "Baseline", "BB-Q", "ACF" and "BB-R", refers to different specifications. "Baseline": IV-GMM on observed quantity. "BB-Q" and "BB-R": dynamic panel estimators. "ACF": IV-GMM on revenue. See Section 5.1 for details. 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 13.

Table 11: Correlation of Instruments and Price in the Simulations

Instruments	v_{iht-1}	k_{iht}	v^2_{iht-1}	k^2_{iht}	$v_{iht-1}k_{iht}$
Instrument's correlation with $\tilde{p}_{it}$	.133	-.177	-.126	.169	.041
	(.004)	(.003)	(.004)	(.003)	(.002)

NOTE: $\tilde{p}_{it} \equiv p_{it} - \rho p_{it-1}$ where ρ is the persistence of productivity. The inputs are denoted by v for variable, k for fixed. The table presents average correlations across the Monte Carlo, simulations while parentheses give the standard deviation.

F Correlation Prices and Instruments

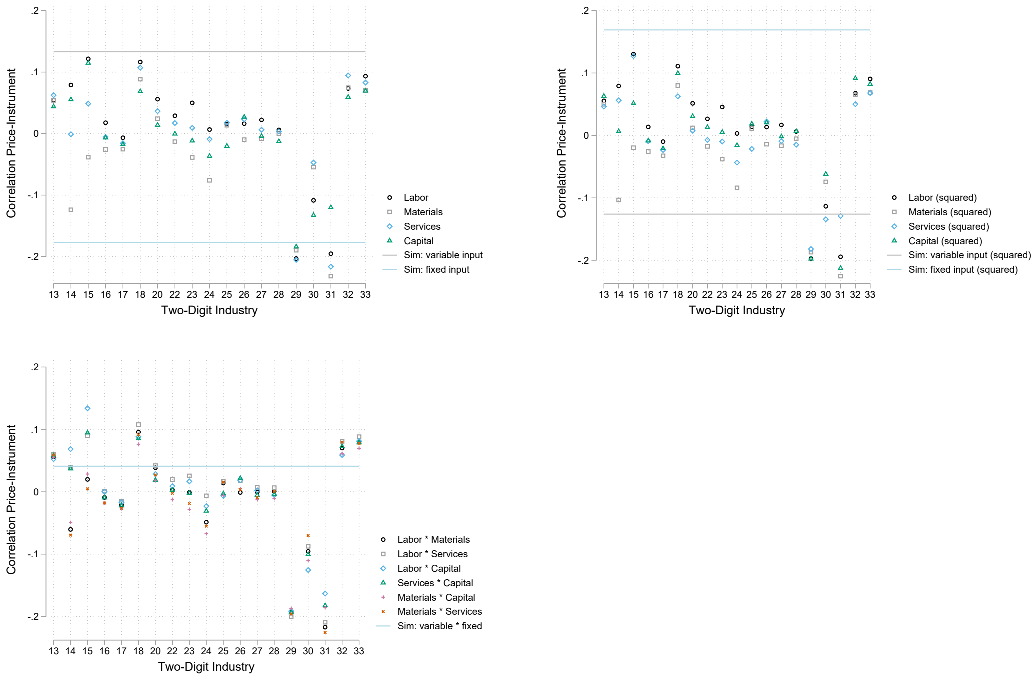
In this appendix we present the correlation between prices and the production function instruments. Section 2.2.4 shows that revenue-based markup estimates for the translog production function will be close to true markups if the correlations between prices and the instruments are low. That result is derived for i.i.d. productivity. If productivity is linear first-order autoregressive, correlation between the instruments and $\tilde{p}_{it} \equiv p_{it} - \rho p_{it-1}$ (where ρ is the persistence of productivity) determines the correlation between true markups and revenue-based markups. Hence, all correlations here are between $\tilde{p}_{it}$ and the production function estimation instruments.

Table 11 presents the correlations for the baseline simulation. The instruments for the variable input are its lags, as the identification assumption is that innovations in productivity are orthogonal to the lagged variable input. The instruments for the fixed factor are just its contemporaneous value, as the fixed factor is determined exogenously and thus uncorrelated with productivity innovations to begin with. We calculate the correlations for each of the Monte Carlo repetitions and present the average correlations in the table. The table shows that the correlations between prices and the instruments are generally close to zero, ranging from 0.041 for the interaction term to -0.177 for the square of the fixed factor.

The correlations in the simulations are similar to the correlations in the data. In the data there are more inputs (and thus instruments); since we estimate the production function separately for every sector, the correlations are plotted in Figure 6. Horizontal axes give the two-digit industry codes. From left to right, the figures contain correlations between price innovations and instruments for linear inputs; for quadratic inputs; and for interacted inputs. Each dot presents a sector level correlation. Horizontal lines, for reference, give the correlations in the simulations.

The figure shows that the empirical correlations between price innovations and instruments are either similar to or closer to zero than the simulations. As a result, in the absence of empirical complications such as misspecification of the production function or heterogeneous production functions within sectors, the correlations between revenue-based markup estimates and quantity-based markup estimates in the data should be similar to the simulations.

Figure 6: Effect of First-Stage Purging Regression on Markup Estimates



NOTE: The figure plots the correlation between $\tilde{p}_{it} \equiv p_{it} - \rho p_{it-1}$ where ρ is the persistence of productivity, and instruments in the production function estimation. The instruments are l_{it}, k_{it}, m_{it-1}, o_{it-1}, l_{it}^2, k_{it}^2, m_{it-1}^2, o_{it-1}^2 respectively for the inputs l_{it}, k_{it}, m_{it}, o_{it}, l_{it}^2, k_{it}^2, m_{it}^2, o_{it}^2, where inputs are denoted by l for labor, k for capital, o for service, and m for material. The horizontal axes give the two-digit industry for which correlations are calculated. Legends give the input for which the instrument's correlation is plotted.

G Data Construction

The dataset is constructed from two sources. The first is a panel of the universe of French firms, which we obtain from Burstein et al. (2020).

The panel is based on FARE, which is an administrative dataset based on data from the tax office DGFiP. It contains detailed balance-sheet- and income statement information. We also drop firms with fewer than two employees, as the number of single-employee firms grew rapidly over our sample due to a regulatory change. The FARE data contains all of the variables for the production function estimation with the exception of quantities and prices.

For revenue, we use total sales (including exports). This is REDI_R310 in FARE. The wage bill is the sum of wage payments and social security contributions (REDI_R216 and REDI_R217). Materials, the variable input in our empirical section, is the sum of intermediate inputs and raw materials, adjusted for inventories (REDI_R210, REDI_R212, REDI_R211, and REDI_213). Services include other purchases, such as marketing and insurance (REDI_214). The capital stock is measured as fixed tangible assets (IMMO_CORP in FARE). Market share is defined as a firm's total share in sales in its five-digit NACE industry, which we calculate before outlier treatments are applied. We drop firms with missing, zero or negative revenue, material purchases, service purchases, wage bills, capital or prices and winsorize variables within two-digit industries, and deflate nominal variables using EU-KLEMS deflators. As in the analytical model, we assume firms within sectors face equal factor prices.

We merge the panel with the administrative survey *Enquête Annuelle de Production* (EAP). EAP is a product-level survey by the statistical office (INSEE) and is the French counterpart of the European PRODCOM survey. It details revenue and quantity sold by 10-digit product code for the universe of manufacturing firms with at least 20 employees or 5 million euros in revenue. The sample is from 2009 to 2019. We merge EAP to FARE using the *siren* code, which is a common firm identifier.

The baseline sample is the intersection of the EAP and the FARE dataset. We include all firm-products unless the firm only acted as an outsourcer or designer, and drop around one-third of firm-products without quantity data. We define a product as a combination of a 10-digit product code and a unit of account, because products reported in different units may not be comparable. For each firm-product we calculate the ratio of revenue over the quantity of the product sold. We then standardize this unit value by dividing it by the revenue-weighted average price of the 10-digit product across the sample. As some firms produce multiple products, we define a firm's price as the sales-weighted average of standardized prices across the goods that it produces. Quantity is the ratio of revenue over this price.[37] Summary statistics are given in Tables 12 and 13.

Table 12: Summary Statistics

Variable	Mean	St. Dev.	Median	10th Pct.	90th Pct.	Observations
FARE						
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
EAP						
Quantity	14,845	62,121	1,891	236	27,458	175,538
Standardized Price	9.45	89.29	1.23	0.77	6.25	175,538

NOTE: Nominal values are in thousands of 2010 euros, deflated using EU-KLEMS deflators. Revenue is deflated with the gross output deflator; purchased inputs are deflated using the intermediate input deflator. Wages and capital tock are deflated using the GDP deflator. Capital is measured as fixed tangible assets. [38] The data contains 26,143 unique firms across 206 (19) sectors at the five (two) digit level.

[37]As a robustness check we standardize prices using the revenue-weighted average price at the 8-digit sector level. These firm-level prices have a 0.89 correlation with our baseline prices.

Table 13: Sectors (two-digit) in the EAP-FARE Dataset

Manufacturing of ...	NACE code	Observations
... 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	32	5,094
Repair and installation of machinery and equipment	33	12,182