Measuring Markups with Revenue Data

Ivan Kirov
Paolo Mengano
James Traina*

November 2022

Abstract

We propose a method that generates unbiased and consistent markup estimates using only revenue data. Building on standard production function estimators, our method requires flexibly modeling markups as a specified function of observables and fixed effects and modifying physical productivity process assumptions into revenue productivity process assumptions. We show that it solves the omitted price bias without imposing additional assumptions on the demand side or the competition structure. Our suggested two-step estimator is simple in concept and implementation, requiring only common regression techniques and information available in most data settings.

JEL: E2, C33, D24, D4, L11
Keywords: *markups, omitted price bias, production functions, productivity*

*Kirov: Analysis Group, ikirov@uchicago. Mengano: University of Zurich, paolo.mengano @econ.uzh.ch. Traina: Federal Reserve Bank of San Francisco, traina@uchicago.edu. We thank Erik Hurst, Chad Syverson, Ali Hortaçsu, Brent Neiman, and Luigi Zingales for their support. We also thank working group participants at the UChicago IO Lunch, as well as Anhua Chen, Zach Flynn, Mike Gibbs, Arshia Hashemi, and Uyen Tran for helpful comments and discussions. Corresponding email address: traina@uchicago.edu.

The composition of the economy is changing with firms growing bigger and more profitable, industries becoming more concentrated, and business dynamism declining.[1] Several explanations have been proposed for these secular trends. One set of hypotheses argues that these trends are the outcome of technological change or intangible capital, typically with positive welfare implications for consumers (Autor et al. 2020, Bessen 2020). Another set of hypotheses argues that these trends are the outcome of rising market power or entry barriers, typically with negative welfare implications for consumers (De Loecker, Eeckhout, and Unger 2020, Gutiérrez and Philippon 2019). Although both sets of hypotheses may be partly true, their policy responses differ dramatically. A key indicator to analyze industry competition and market power is firm-level markup, the ratio of output price to marginal costs. The scale of the potential problem calls for economists to accurately measure firm-level markups in a way that generalizes across industries and over long periods of time.

Recent macroeconomic research responds to this call by adopting the production approach to measure firm-level markups on large datasets representative of entire industries or economies. This approach consists in applying standard methods to estimate production functions and then using these to recover markups (De Loecker, Eeckhout, and Unger 2020, Traina 2018, Diez, Leigh, and Tambunlertchai 2018, Calligaris, Criscuolo, and Marcolin 2018). However, econometricians originally designed these estimators to measure productivity in competitive environments, where variation in prices can only reflect variation in quality. In imperfectly competitive environments, firms internalize their price-setting ability, and this confounds the link between revenues and quantities (Klette and Griliches, 1996; Doraszelski and Jaumandreu, 2019; Bond et al., 2021). Therefore, these estimators require additional detailed information on output prices to remain unbiased and consistent.[2] However, most economy-wide or industry-wide datasets do not include such information. For instance, the commonly used US Census of Manufactures, Compustat, and the Colombian, Chilean, and Slovenian manufacturing surveys all only have revenue data. Absent additional assumptions, it is not possible to use existing methods to disentangle whether firms have higher revenues because of higher productivity or higher markups. Researchers and policymakers interested in studying competition

[1] See for example Grullon, Larkin, and Michaely (2019) and Decker et al. (2016).

[2] Note that using an industry price index to recover quantities does not solve this problem because it ignores the importance of firm specific prices, which are essential to understanding firms' market power (Bond et al., 2021).

thus face a sizable gap between models and data.

To bridge this gap, we propose an approach that offers unbiased and consistent estimates using only revenue data. Our approach relies on two main assumptions. First, markups can be spanned by a flexibly specified function of observables and fixed effects. Such a function hinges on markups being endogenous objects and absorbs markup variation due to changes in demand or conduct. This is in line with the traditional approach of production function estimators that uses a control function to proxy for physical productivity. In the absence of price data, we show that controlling for markups instead solves the omitted price bias generated using only revenues. Second, we assume that revenue productivity follows a Markov process. Also, in this case, an assumption equivalent to existing work on production function estimation (Klette and Griliches, 1996; De Loecker, 2011; De Loecker and Warzynski, 2012), and consistent with the persistence results of Foster, Haltiwanger, and Syverson (2008). Given this minimal setup, we propose a two-step estimator requiring only standard regression techniques that is unbiased, consistent, and simple in concept and implementation. And that can be used on easily accessible data sources.

Our chief contribution is to offer an approach that identifies physical output elasticities to infer markups without requiring information on output prices: a long-standing critical issue in this literature. The first stage of our suggested estimator is a share regression in the spirit of Gandhi, Navarro, and Rivers (2020), which directly estimates the firm's cost minimization first-order condition to recover its revenue elasticity. The second stage then uses information and timing assumptions to separate the effects of inputs on prices and output, thus recovering its output elasticity. We can then measure unbiased and consistent firm-level markups with these estimates. In addition, we also suggest ways to modify other production function estimators that impose more structure on the productivity process, as in Blundell and Bond (2000). The approach applies to a wide range of data and market settings, including ones where existing methods might not work. For example, even if we had information on output prices, we would need to quality-adjust physical quantities to ensure that outputs are comparable across firms, yet quality data are very rare in practice. Many firms also produce multiple products, which complicates the interpretation of quantity data. Our approach alleviates both problems and is thus attractive even with price data.

The literature on production function estimation has long noted the omitted price bias in measuring markups with revenue data. Klette and Griliches (1996) shows that markup estimators are biased and inconsistent without information on output prices. This bias arises because unobserved prices enter the residual of the estimating equation but also affect firm input choices (e.g., higher prices induce the firm to increase input use to increase output). Consequently, omitted prices cause a classic simultaneity problem. Doraszelski and Jaumandreu (2019) extends these results to modern production function estimators, while Bond et al. (2021) further shows that the bias should lead to unitary markups for profit-maximizing firms. Our paper generalizes the solution in Klette and Griliches (1996), which relies on the restrictive assumptions of a constant elasticity demand system and monopolistic competition. Beyond typical parametric critiques, these assumptions narrow the very concept of competition and imply zero markup variation within industries. More recently, De Loecker et al. (2016) uses observed output prices to control for unobserved input price biases, but the authors operate in a setting where output price data are available. Foster, Haltiwanger, and Syverson (2008) suggests that the omitted price bias is so significant that it changes the observed correlation between physical productivity and revenue productivity. More recently, De Ridder, Grassi, and Morzenti (2021) shows again that it is not possible to recover precise markup levels from revenue data but that we can learn something about the dispersion both in the cross-section variation and over time. In sum, the literature has shown that the omitted price bias is important but has not yet offered a general solution for resolving it. We see the contribution of this paper as a promising avenue to study market power by tailoring the productivity process and proxy variable assumptions.

Besides addressing the omitted price bias, our paper also relates to the literature on transmission bias: firms optimally choose their inputs as a function of their productivity, which means the productivity error simultaneously determines both output and inputs (Marschak and Andrews 1944). A regression of output on inputs does not identify a firm's production function, as the productivity error term is "transmitted" into the input decision, which creates a classic endogeneity problem, as the error term is correlated with the explanatory variable. The literature overcomes this bias with assumptions about a firm's production or productivity process. Broadly, proxy variable estimators (Olley and Pakes 1996, Levinsohn and Petrin 2003, Ackerberg, Caves, and Frazer 2015, Gandhi, Navarro, and Rivers 2020) assume that observable inputs can control for unobserved productivity, while dynamic panel estimators (Blundell and Bond 2000) parametrically impose linearity on the productivity process. Our suggested method is in the tradition of proxy variable estimators, but it allows researchers to relax the critical assumption of competitive output markets. Our work also relates to earlier attempts at forming control functions for revenue productivity directly, as in Flynn, Traina, and Gandhi (2019). One appealing feature is that in general models of competition, higher planned markups induce lower chosen flexible inputs; this suggests a straightforward way to microfound the markup control function in an input demand equation, as in Olley and Pakes 1996. We also derive results showing how to extend methods in the tradition of dynamic panel estimators.

Our main estimator builds on this line of work by addressing omitted price and transmission bias simultaneously. We do so by specifying a control function for markups, as opposed to the typical control function for physical productivity. We then directly estimate the firm's first-order condition by regressing its flexible input's log cost share of revenue on log inputs and fixed effects. Though similar in spirit to the first stage of Gandhi, Navarro, and Rivers (2020), we do not identify the flexible input's output elasticity from this share regression. Instead, we use it to partially identify the elasticity by identifying a combination of the elasticity and markup. We then substitute this composite into the revenue production equation to control for unobserved revenue productivity. Together, this method allows us to identify physical output elasticities relying only on revenue data.

The rest of the paper proceeds as follows. In Section 1, we summarize a standard production model. In Section 2, we describe how to adapt the production model for imperfectly competitive environments, and how to use our proposed share regression to estimate a control function and recover markups. We compare our estimator to existing production function estimation methods in Section 3, and provide specific examples of markup controls in Section 4. Section 5 concludes.

1 A Standard Production Model

In this section, we outline the typical structural model and data that researchers use to estimate production functions and markups. We suppose throughout that the researcher wants to estimate firm-level markups, and only has data on firm revenues (that is, not quantities).

1.1 Definitions

We observe data for a panel of firms over periods $t = 1, 2, \ldots, T$. We omit panel subscripts and let the data take a short panel form: the number of firms grows large for a fixed T. For each firm, we observe revenue $R = PQ$, expenditures on a competitively supplied flexible input X with cost C, and a vector of nonflexible inputs K. X is flexible in the sense that it is both variable and static: firms may adjust it in each period after observing the realization of state variables such as productivity, and its choice has no dynamic implications. K is fixed, dynamic, or both. Throughout the rest of the paper, we use uppercase to refer to levels of variables and lowercase to refer to logs of variables. For exposition, we refer to X as intermediates, K as capital, and assume capital is fixed so that it may not respond to current period state variables. In practice, the choice of X might be energy, materials, labor, or some combination thereof. We recommend choosing the input that is most likely to satisfy the competitively supplied and flexible assumptions based on the empirical setting.

1.2 Pricing, Technology, and Productivity

Firms may have market power in the output market, so that each firm's residual demand curve may not be perfectly elastic.[3] Markups can come from a mix of demand, cost, conduct, or other sources. Here we are agnostic; markups may, for instance, arise because of product differentiation, consumer tastes, technological advantages, or concentrated markets. Cost-minimizing firms internalize inframarginal price reductions and choose a price and output pair $\{P, Q\}$ along their residual demand curves subject to technological constraints. Firms may choose to price above marginal cost, and we define the markup $\mathcal{M} = \frac{P}{\Lambda}$, where the denominator represents the shadow price of producing a unit of output, namely the marginal cost.

[3]Here we refer to the firm's residual demand curve net of potential strategic interaction concerns.

Inputs generate output according to a constant returns to scale production function with Hicks-neutral productivity: $Q = AF(X, K)$.[4] A represents physical productivity (TFPQ in the parlance of Foster, Haltiwanger, and Syverson (2008)), and $F(X, K)$ represents the firm's production function. Finally, the log productivity term a is additively separable into a part known to the firm when making input decisions ω and an i.i.d. error term ε. Firm production is thus

$$q = f + \omega + \varepsilon$$

We note that $f = f(x, k)$.

1.3 Optimization, Information, and Timing

We assume that firms minimize costs. Each firm uses expected output in its minimization problem because it knows that it must account for an as-yet-unknown portion of productivity ε.

The timing of the problem is as follows. First, a firm uses its expectation about its productivity a conditional on the part known to the firm ω, along with other information observable to the firm about its residual demand curve, to plan a markup μ. Second, the firm chooses the corresponding intermediate inputs X to implement this plan given its residual demand curve, technology, capital stock, and expected productivity. Finally, productivity is fully realized, production occurs, and markups are realized. Markups are generally not orthogonal to productivity, since the firm uses its expectation about productivity to plan a markup.

Given this setup, the firm's cost minimization problem with time t information $\mathcal{I}$ is

$$\begin{aligned} \min_{X} \quad & CX \\ s.t. \quad & Q = \mathbb{E}[A|\mathcal{I}]F(X, K) \end{aligned}$$

with Lagrangian

$$CX + \Lambda(Q - \mathbb{E}[A|\mathcal{I}]F(X, K)$$

[4] Any returns-to-scale parameter may be substituted for CRS, but some assumption is needed because markups are not identified separately from the scale parameter (Flynn, Traina, and Gandhi 2019).

and first-order condition for X

$$[X] \qquad C = \Lambda \mathbb{E}[A|\mathcal{I}] F_X$$

with $F_X = \frac{\partial F}{\partial X}$. Denote $\mathcal{M}$ as the markup, i.e. the ratio of output price P over marginal cost Λ. We can manipulate the first-order condition to yield

$$\begin{aligned}
\mathcal{M} &= \frac{P}{\Lambda} \\
&= \mathbb{E}[A|\mathcal{I}] F_X \frac{P}{C} \\
&= \frac{\mathbb{E}[A|\mathcal{I}]}{A} F_X \frac{AX}{Q} \frac{PQ}{CX} \\
&= \frac{\mathbb{E}[A|\mathcal{I}]}{A} F_X \frac{X}{F} \frac{PQ}{CX} \\
&= \frac{\mathbb{E}[\mathcal{E}]}{\mathcal{E}} f_x \frac{R}{CX}
\end{aligned}$$

where the last line comes from the fact that the firm knows part of its productivity before its input choice $\frac{\mathbb{E}[A|\mathcal{I}]}{A} = \frac{\Omega \mathbb{E}[\mathcal{E}]}{\Omega \mathcal{E}} = \frac{\mathbb{E}[\mathcal{E}]}{\mathcal{E}}$, the definition of the output elasticity $f_x = F_X \frac{X}{F}$ and revenue $R = PQ$. Defining $b = \log \mathbb{E}[\mathcal{E}]$ and writing this result in logs gives us

$$\mu = \log f_x + r - cx + b - \varepsilon \tag{1}$$

That is, the log markup μ is the sum of the intermediate input's log output elasticity f_x, the intermediate input's log inverse cost share of revenue $r - cx$, and the unplanned productivity term $b - \varepsilon$. Recovering markups from this equation requires an estimate of f_x. However, we cannot simply regress revenue on inputs to get this estimate for two critical reasons. The first reason is the omitted price bias emphasized in Klette and Griliches (1996): higher markups induce the firm to decrease input use, which increases prices and thus (all else equal) revenues. The second reason is the transmission bias emphasized in Marschak and Andrews (1944): higher physical productivity induces the firm to increase input use, which increases output and thus (all else equal) revenues.

One approach in the literature is to ignore the distinction between revenues and quantities, treating the revenue production function estimates as though they are physical production function estimates. However, this empirical strategy confounds demand with productivity (Foster, Haltiwanger, and Syverson, 2008) and results in markup estimates devoid of empirical content (Bond et al., 2021). We discuss our solutions to these issues next.

2 Measuring Markups with Revenue Data

In this section, we discuss the assumptions needed to estimate markups and production function parameters with revenue data. We first interpret the first-order condition 1 as a markup function and use it to identify revenue elasticities. We then combine it with information and timing assumptions on revenue productivity to identify physical output elasticities.

2.1 Assumption 1: The Markup Control Function

Let markups be a function of inputs, firm and time fixed effects ι and τ, and a vector of other firm-time varying observables relevant in determining markups $\mathbf{D}$:

$$\mu = h(x, k, \iota, \tau, \mathbf{D}) \tag{2}$$

We return to examples of possible observables in Section 4. One appealing feature is that in general models of competition, higher planned markups induce lower chosen intermediates; this suggests a straightforward way to microfound the markup control function in an input demand equation, as in Olley and Pakes (1996). Rewrite the first-order condition 1 as

$$cx - r = \log f_x - \mu + b - \varepsilon$$

The left-hand side is the intermediates log cost share of revenue. The term $\log f_x - \mu$ on the right-hand side is the log revenue elasticity with respect to input x, a mix of supply and demand parameters.[5] As productivity is Hicks-neutral, the elasticity

[5] $\log f_x - \mu = \log f_x - \log(\frac{1}{1+p_q}) = \log f_x + \log(f_x p_q) = \log q_x + \log p_x$.

term f_x is a function of inputs only: $f_x = f_x(x, k)$. Combining the revenue elasticity terms into a single function $s(x, k, \iota, \tau, \mathbf{D}) = \log f_x(x, k) - h(x, k, \iota, \tau, \mathbf{D})$, our first stage estimating equation becomes:

$$cx - r = s(x, k, \iota, \tau, \mathbf{D}) + b - \varepsilon$$

To operationalize this equation, a researcher may (nonparametrically) regress the intermediates share of revenues on inputs, fixed effects, and a vector of markup determinants $\mathbf{D}$ to get an estimate of the revenue elasticity $\hat{s} = \widehat{\log f_x - \mu}$. This share regression estimates the specified markup control function: it describes the determinants of wedges between prices and marginal costs, and is similar to the share regression in Gandhi, Navarro, and Rivers (2020), but adapted for cases of imperfect competition with unobserved prices.

Importantly, it also recovers an estimate of the error $\hat{\varepsilon}$, and therefore $\hat{b}$. Estimating $\hat{\varepsilon}$ is the primary function of the first stage of proxy variable estimators (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves, and Frazer, 2015). Estimating it here allows us to replace the physical productivity control function assumption of these models with a markup control function assumption. However, the share regression alone cannot separate the impact of markups from output elasticities, since it still contains the unknown f_x. We now turn to identifying the physical elasticity f_x by adding information and timing assumptions to the revenue productivity process.

2.2 Assumption 2: Markovian Revenue Productivity

To separately identify markups and physical elasticities, we combine the share regression with structure on the revenue productivity process. Specifically, we assume revenue productivity $\nu = p + \omega$ (the sum of prices p and physical productivity ω) follows a Markov process with additively separable mean-zero shocks η. Relative to the existing literature, this assumption is equivalent to the existing assumptions of Klette and Griliches (1996), De Loecker (2011), and De Loecker and Warzynski (2012), and is consistent with the persistence results of Foster, Haltiwanger, and Syverson (2008).

Writing $\mathbb{L}$ as the lag operator, we have

$$\nu = g(\mathbb{L}[\nu]) + \eta \tag{3}$$

Now write firm revenues as

$$\begin{aligned} r &= p + q \\ &= p + f + \omega + \varepsilon \\ &= f + \nu + \varepsilon \end{aligned}$$

Where the second line comes from the definition of f, the third line comes from the definition of ν. Using our Markovian assumption leads to our second stage estimating equation:

$$r = f + g(\mathbb{L}[\nu]) + \eta + \varepsilon$$

Now rewrite the output of our first stage to isolate revenue productivity:

$$\begin{aligned} cx - r &= \hat{s} + \hat{b} - \hat{\varepsilon} \\ cx - f - \nu - \hat{\varepsilon} &= \hat{s} + \hat{b} - \hat{\varepsilon} \\ \nu &= cx - f - \hat{s} - \hat{b} \end{aligned}$$

Combining these equations yields our main estimating equation:

$$r = f + g(\mathbb{L}[cx - f - \hat{s} - \hat{b}]) + \eta + \hat{\varepsilon}$$

Two notes are in order. First, as shown in Flynn, Traina, and Gandhi (2019), we require a scale elasticity assumption. Second, there must be independent variation in μ which does not enter f. This variation can come from the fixed effects ι and τ, or the $\mathbf{D}$ vector of other firm-time observables (we discuss specific examples in Section 4). With this variation, the model identifies physical quantity elasticities and markups. Without this variation, the nonparametric underidentification arguments of Gandhi, Navarro, and Rivers (2020) will apply. However, we view this requirement as minimal – the existing revenue productivity literature suggests persistent dispersion across firms that is independent of other latent variables such as physical productivity (Foster, Haltiwanger, and Syverson, 2008). Moreover, it is consistent with empirically successful models of firm dynamics, such as those studied in Sutton (1991).

2.3 Estimation

Here we describe the steps to implement our suggested estimator.

1. Regress the intermediates log cost share of revenues ($cx - r$) on log inputs and firm and time fixed effects, along with other observed markup determinants **D**. Use the predicted residual $\hat{\varepsilon}$ to form $\hat{b} = \log \hat{\mathbb{E}}[\exp(\hat{\varepsilon})]$, and therefore $\hat{s} = \log \widehat{f_x - }\mu = cx - r - \hat{b} + \hat{\varepsilon}$.

2. Specify functional forms for f and g, such as Cobb-Douglas technology with AR(1) productivity, or translog technology with quadratic Markov productivity. Note that f must satisfy a scale elasticity assumption, such as constant returns to scale. In practice, the main concern is how to model the intermediates elasticity, because this term directly affects the markup as seen from the first-order condition 1.

3. Combine the estimates $\hat{s}$, $\hat{b}$, and $\hat{\varepsilon}$ with data r and cx and the specified functional forms for f and g to form the productivity shock $\hat{\eta} = r - f - g(\mathbb{L}[cx - f - \hat{s} - \hat{b}]) - \hat{\varepsilon}$. This shock will be a function of f and g parameters.

4. Estimate the parameters of f and g using the moment conditions formed by

$$\mathbb{E}[\hat{\eta}|k, \mathbb{L}[\hat{s}]] = 0$$

The specific moments depend upon the specifications of f and g. For instance, we may add moments for a translog production function f by including squares and interactions of the nonflexible inputs. We may also add lags and interactions of the revenue share vector $\hat{s}$ to control for more complicated specifications of the revenue productivity process g. For instance, we could use $\hat{s}^2$ to control for a second-order g process. By construction, these additional moments are orthogonal to revenue productivity innovations. In economic terms, costs, demand, and conduct are determinants of markups, which co-evolve with the TFPR process. Innovations to this process are therefore orthogonal to any functions of the costs, demand, or conduct embedded in the $\hat{s}$ vector. This two-step approach can also be implemented as a single step GMM problem by minimizing the residuals in the share regression together with the moment conditions from the second step.

3 Comparison to Existing Competitive Models

In this section, we compare our estimator to two commonly used methods to estimate production functions. We first discuss the proxy variable class, then the dynamic panel class. Both methods form moments from information and timing assumptions about the productivity process, but differ in other assumptions to achieve identification.

3.1 Proxy Variable Estimators

The proxy model of production (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves, and Frazer, 2015; Gandhi, Navarro, and Rivers, 2020) identifies production functions using assumptions about the unobservable state. Proxy models assume that productivity is Markovian and that productivity can be written as a control function of observables. These two assumptions allow one to use lagged inputs to control for current productivity, solving the transmission bias. More formally, the assumptions of the proxy model are:

Proxy Variable Assumption 1: Physical productivity follows a first-order Markov process: $\omega = g(\mathbb{L}[\omega]) + \eta$.

Proxy Variable Assumption 2: Physical productivity is a control function of observables: $\omega = \omega(x, k)$.

Proxy Variable Assumption 2 ensures that past inputs can proxy for current productivity through the productivity process. Then substitution yields

$$\begin{aligned} q &= f + \omega + \varepsilon \\ &= f + g(\mathbb{L}[\omega]) + \eta + \varepsilon \\ &= f + g(\mathbb{L}[\omega(x,k)]) + \eta + \varepsilon \end{aligned}$$

Identification proceeds by forming moments with the composite error term $\eta + \varepsilon$.

With revenue data, this derivation becomes

$$\begin{aligned} r &= f + p + \omega + \varepsilon \\ &= f + p + g(\mathbb{L}[\omega]) + \eta + \varepsilon \\ &= f + p + g(\mathbb{L}[\omega(x,k)]) + \eta + \varepsilon \end{aligned}$$

The appearance of prices on the right-hand side is the origin of the omitted price bias terminology.

Our approach is a direct modification of these two assumptions. For the first, we assume revenue productivity (TFPR) follows a first-order Markov process, instead of physical productivity (TFPQ). This assumption is consistent with Foster, Haltiwanger, and Syverson (2008), which finds that revenue productivity exhibits similar levels of persistence as physical productivity. It is also implicit in existing work that estimates revenue production functions, or explicit in existing work that attempts to correct for the omitted price bias with richer data or stronger parametric structure (Klette and Griliches 1996, De Loecker 2011, De Loecker and Warzynski 2012). For the second, we assume markups are a control function of observables, instead of physical productivity. Although markups and productivity are both unobservable, researchers typically impose some markup-setting process or rule in modeling (for example, monopolistic competitors facing a constant price elasticity of substitution demand system). Therefore, we view our assumption on markups as less limiting. More generally, one can think of our model as a version of a proxy variable estimator, in which we are proxying for markups instead of proxying for productivity. It is built to estimate markups, and also allows us to relax some of the physical productivity assumptions.

3.2 Dynamic Panel Estimators

The dynamic panel approach (Blundell and Bond 2000) is a commonly used alternative to proxy variable approaches. Dynamic panel models maintain the same basic structure of production. However, they impose linearity on the productivity process.

Dynamic Panel Assumption: Physical productivity ω follows an AR(1) process: $\omega = \rho\mathbb{L}[\omega] + \eta$.

The unobserved term η is uncorrelated with all past and future input choices. Then differencing the production function yields

$$
\begin{aligned}
q - \rho\mathbb{L}[q] &= f - \rho\mathbb{L}[f] + \omega - \rho\mathbb{L}[\omega] + \varepsilon - \rho\mathbb{L}[\varepsilon] \\
q - \rho\mathbb{L}[q] &= f - \rho\mathbb{L}[f] + \eta + \varepsilon - \rho\mathbb{L}[\varepsilon] \\
q &= \rho\mathbb{L}[q] + f - \rho\mathbb{L}[f] + \eta + \varepsilon - \rho\mathbb{L}[\varepsilon]
\end{aligned}
$$

The appearance of lagged quantities on the right-hand side is the origin of the dynamic panel terminology. Identification proceeds by forming moments with the composite error term $\eta + \varepsilon - \rho\mathbb{L}[\varepsilon]$.

With revenue data, this derivation becomes

$$
\begin{aligned}
r - \rho\mathbb{L}[r] &= f - \rho\mathbb{L}[f] + p - \rho\mathbb{L}[p] + \omega - \rho\mathbb{L}[\omega] + \varepsilon - \rho\mathbb{L}[\varepsilon] \\
r - \rho\mathbb{L}[r] &= f - \rho\mathbb{L}[f] + p - \rho\mathbb{L}[p] + \eta + \varepsilon - \rho\mathbb{L}[\varepsilon] \\
r &= \rho\mathbb{L}[r] + f - \rho\mathbb{L}[f] + p - \rho\mathbb{L}[p] + \eta + \varepsilon - \rho\mathbb{L}[\varepsilon]
\end{aligned}
$$

Absent additional assumptions, we cannot proceed further without price data. However, combining this equation with a version of our earlier Markovian revenue productivity assumption ($\nu = g(\mathbb{L}[\nu]) + \eta$), we can make progress. Suppose our earlier assumption of Markovian revenue productivity holds, and further suppose it is linear, so that revenue productivity follows an AR(1) process: $\nu = \rho\mathbb{L}[\nu] + \eta$. We can then proceed as follows:

$$\begin{aligned}
r - \rho\mathbb{L}[r] &= f - \rho\mathbb{L}[f] + p - \rho\mathbb{L}[p] + \omega - \rho\mathbb{L}[\omega] + \varepsilon - \rho\mathbb{L}[\varepsilon] \\
r - \rho\mathbb{L}[r] &= f - \rho\mathbb{L}[f] + \eta + \varepsilon - \rho\mathbb{L}[\varepsilon] \\
r &= \rho\mathbb{L}[r] + f - \rho\mathbb{L}[f] + \eta + \varepsilon - \rho\mathbb{L}[\varepsilon]
\end{aligned}$$

This derivation does not require a markup control function, and therefore suggests a similar tradeoff as in the competitive case: researchers may impose more structure (linear) on the productivity process to avoid assumptions that observables (inputs) and fixed effects span unobservables (markups).

4 Relation to Imperfect Competition Models

Our solution generalizes much of the existing literature. In our earlier setup, we showed that our markup function identifies markups so long as μ is determined partially independently from inputs. In this section, we offer several commonly used parametric examples of markup functions.

4.1 Monopolistic Competition with CES Demand

Suppose that firms are monopolistic competitors facing a constant price elasticity of substitution demand system. Suppose further that these firms compete in a number of industries j. In this environment, a firm i in industry j faces a demand curve given by $Q_i = Q_j(\frac{P_i}{P_j})^{-\sigma_j}$. Firm optimization implies that markups are constant within industries and given by $\mathcal{M}_j = \frac{\sigma_j}{\sigma_j - 1}$.

In the context of our markup function, assuming monopolistic competition with CES demand implies that $\mu = \mu_j$: markups are determined by a constant within industry. One may recover markups and elasticities by simply including an industry fixed effect in the share regression.

This setup is the setting in Klette and Griliches (1996). This paper shows that, in this case, one can estimate (industry-level) production functions by including controls for the industry quantity production. Klette and Griliches (1996) uses the estimating equation

$$r = \beta_0 + \frac{\sigma_j - 1}{\sigma_j}(\beta_x x + \beta_k k) - \frac{1}{\sigma_j} q_j + \nu$$

Here, q_j is an industry-level price index which comes from the monopolistically competitive environment. Estimation can then proceed using (observed) industry-level output.

Gandhi, Navarro, and Rivers (2020) extends the Klette and Griliches (1996) model to allow for time-varying price elasticities of demand. We can easily accommodate such an extension by including a time fixed effect in the share regression.

In sum, our approach is a generalization of the Klette and Griliches (1996) correction that allows for more conduct and demand structures than monopolistic competition with CES demand. This generalization is especially important as it allows credible scaling of production-based markup estimators across many industries and time periods, where these earlier assumptions might be considerably off.

4.2 Firm-Time Characteristics

If the researcher has data on firm-level characteristics which determine markups, then our model identifies market power and production elasticities by putting these into $\mathbf{D}$. For instance, advertising, managerial practices, research and development (Doraszelski and Jaumandreu 2013), export status (De Loecker and Warzynski 2012), or product mix (De Loecker 2011) all might be associated with markups, and therefore added to the share regression. Any other observable market characteristics that vary by firm-time, such as combining geographic variation with consumer income variation, may also be added, depending on the researcher's model.

If revenue market shares determine markups, then in our model we have $\mathbf{D} = \frac{R_i}{R_j}$, where R_j is industry revenues. Researchers may define industries j in whatever way appropriate, such as common industry codes, or broadly or narrowly defined product markets. Then the markup control function may again be used to identify markups and output elasticities by adding market shares to each. Unlike models such as nested CES, this approach does not impose a parametric relationship between market shares and markups. Rather, the data determine the relationship.

In microfounding this control variable, typical models of competition such as homogenous product Cournot and differentiated product Bertrand result in a mapping from markups to quantity market shares, not revenue market shares. Of course, if we had quantity market shares, we would have quantity information, which would obviate the need for revenue data corrections.

4.3 Conduct Instruments

One readily constructed choice of μ instrument is competitors' input choices within industries j. Markups often depend upon industry conduct, and therefore, upon competitors' choices. We sketch here a straightforward application of this approach. Suppose that each firm operates under Markovian market conditions δ:

$$\delta = h(\mathbb{L}[\delta]) + u$$

Suppose these conditions evolve with error terms u orthogonal to revenue productivity innovations η. Firms may attempt to forecast the evolution of market conditions and use other firms' input choices as proxies, so that we can write $\delta = \delta(x_{-i}, k_{-i})$. Then we have $\mathbb{E}[\eta \times \mathbb{L}[\delta]] = 0$. In sum, firms use past competitor choices to form predictions about current market conditions, but these choices do not directly affect productivity in the current period. These conduct instruments are valid in many models of imperfect competition. Any model in which demand and supply conditions are persistent is consistent with using lagged d as instruments. For instance, a model of habit formation is consistent with our timing assumptions, as is a model with serially correlated input and output shocks. These instruments do not rule out perfect competition in intermediates markets or strategic interactions between firms (indeed, they rely on such interactions). They do, however, rule out dynamic strategic interactions, since such interactions would generate dynamic implications for the flexible input, violating our assumption that it is static. These conduct instruments are readily constructed in production datasets: all that is required is a choice of competitor market definition.

5 Concluding Remarks

This paper proposes a method to estimate markups with revenue data, without requiring information on prices. It combines firm cost minimization with a control function for markups and an assumption on the timing of firm decision-making to recover quantity elasticities and estimates of market power. In line with the literature on production function estimation, we show that it is possible to identify both physical production elasticities and markups, as long as a researcher has some variable which affects markups independently of inputs. This condition is less restrictive than it seems, for this independent variation could come from an industry-level constant (as in the monopolistic competition, CES demand case), fixed effects, market shares, or other inputs recoverable from production data. However, the researcher must spell out the determinants of markups in their model and relate them to observables in the data, albeit very generally. Our method is a variant of proxy variable production function estimators, but proxies explicitly for markups rather than for productivity. Contrary to traditional production function estimators, however, it directly addresses the pervasive omitted price bias arising from the lack of price information. This allows to estimate markups with minimal structural assumptions and data requirements. We believe it is a tool that applies to a wide variety of settings, and hope it can clarify and expand the competition literature.

References

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

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. Van Reenen (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics 135*(2), 645–709.

Bessen, J. (2020). Industry concentration and information technology. *Journal of Law and Economics 63*(3), 531–555.

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

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

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

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica 79*(5), 1407–1451.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics 135*(2), 561–644.

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

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

De Ridder, M., B. Grassi, and G. Morzenti (2021). The hitchhikerâs guide to markup estimation. Working paper.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2016). Declining business dynamism: What we know and the way forward. *American Economic Review 106*(5), 203–07.

Diez, F., D. Leigh, and S. Tambunlertchai (2018). Global market power and its macroeconomic implications. Working paper.

Doraszelski, U. and J. Jaumandreu (2013). R&D and productivity: Estimating endogenous productivity. *Review of Economic Studies 80*(4), 1338–1383.

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

Flynn, Z., J. Traina, and A. Gandhi (2019). Measuring markups with production data. Working paper.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review 98*(1), 394–425.

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

Grullon, G., Y. Larkin, and R. Michaely (2019). Are US industries becoming more concentrated? *Review of Finance 23*(4), 697–743.

Gutiérrez, G. and T. Philippon (2019). The failure of free entry. Working paper, National Bureau of Economic Research.

Klette, T. J. and Z. Griliches (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics 11*(4), 343–361.

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

Marschak, J. and W. Andrews (1944). Random simultaneous equations and the theory of production. *Econometrica*, 143–205.

Olley, S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Sutton, J. (1991). *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration.* MIT press.

Traina, J. (2018). Is aggregate market power increasing? Production trends using financial statements. Working paper, Stigler Center for the Study of the Economy and the State.