

Production Function Estimation with Unobserved Input Price Dispersion*

Paul L. E. Grieco^a, Shengyu Li^a, and Hongsong Zhang^b

^aThe Pennsylvania State University

^bThe University of Hong Kong

November 4, 2013

Abstract

We propose a method to consistently estimate production functions in the presence of input price dispersion when intermediate input quantities are not observed. The traditional approach to dealing with unobserved input quantities—using deflated expenditure as a proxy—requires strong assumptions for consistency. Instead, we control for heterogeneous input prices by exploiting the first order conditions of the firm’s profit maximization problem. We show that the traditional approach tends to underestimate the elasticity of substitution and biases estimates of the distribution parameters. Our approach applies to a general class of production functions. It can accommodate both heterogeneity in input prices and a variety of heterogeneous intermediate input types. A Monte Carlo study illustrates that the omitted price bias is significant in the traditional approach, while our method consistently recovers the production function parameters. We apply our method to a firm-level data set from Colombian manufacturing industries. The empirical results are consistent with the predictions that the use of expenditure as a proxy for quantities biases the elasticity of substitution downward. Moreover, using our preferred method, we provide evidence of significant input price dispersion and even wider productivity dispersion than is estimated using proxy methods.

Keywords: *production functions, unobserved price bias, productivity dispersion*

*The authors would like to thank participants at the 22nd Annual Meeting of the Midwest Econometrics Group, the 11th Annual International Industrial Organization Conference, the 2013 North American Meetings of the Econometric Society, and the 2013 Conference of the European Association for Research in Industrial Economics for very helpful comments. In addition, we benefited from thoughtful comments provided by Robert Porter, Mark Roberts, David Rivers, James Tybout, and two anonymous referees. We are also grateful to Mark Roberts and James Tybout for providing the data used in the empirical application. All errors are the authors’ own responsibility. Correspondence: Grieco: paul.grieco@psu.edu, Li: sul223@psu.edu, Zhang: hszhang@hku.hk.

1 Introduction

In applications of production function estimation, many datasets do not contain a specific accounting of intermediate input prices and quantities, but instead provide only information on the total expenditure on material inputs. This presents a challenge for consistent estimation when input prices are not homogeneous across firms or when different firms have access to different types of inputs (for example, parts of varying quality). To address this issue, many previous studies assume a homogenous intermediate input is purchased from a single, perfectly competitive market. This assumption facilitates the use of input expenditures as a proxy for quantities (e.g., Levinsohn and Petrin, 2003). However, if this assumption does not hold—for example, if transport costs create price heterogeneity across geography—then the traditional proxy-based estimator is inconsistent. The logic of the inconsistency is straightforward: input price heterogeneity will be observed by firms who respond to price differences both by substituting across inputs and adjusting their total output, causing an endogeneity problem that cannot be controlled for using a Hicks-neutral structural error term. Even in a narrowly defined industry, perfect competition in input markets is not likely to hold, so the proxy approach is clearly not ideal. Fortunately, observed variation in labor input quantities, together with labor and materials expenditures, contains useful information on the intermediate input price variation across firms. By utilizing this variation within a structural model of firms maximizing profits, we introduce a method to consistently estimate firms’ production function in the presence of unobserved intermediate input price heterogeneity.

The omitted price problem for production function estimation was first recognized by Marschak and Andrews (1944). They proposed the use of expenditures and revenues as proxies for input and output quantities under the assumption that prices were homogeneous across firms. In practice, the literature has documented significant dispersions in both input and output prices across firms and over time (Dunne and Roberts, 1992; Roberts and Supina, 1996, 2000; Beaulieu and Matthey, 1999; Bils and Klenow, 2004; Ornaghi, 2006; Foster, Haltiwanger, and Syverson, 2008; Kugler and Verhoogen, 2012). Klette and Griliches (1996) show the consequence of ignoring the output price dispersion is a downward bias in the scale estimate of production function.¹ The effect of input

¹Klette and Griliches (1996) provide a structural approach for controlling for output price variation, we incorporate

price dispersion is slightly more complicated. Using a unique data set containing both inputs price and quantity data, Ornaghi (2006) documents input price bias under the Cobb-Douglas production function. In Section 2, we discuss how input price dispersion also biases both the output elasticity and substitution parameters in more general specifications.²

A typical data set for production function estimation contains firm-level revenue, intermediate (i.e., material) expenditure, total wage expenditure, capital stock, investment, and additional wage rate/labor quantity. However, quantities and prices for intermediate input are typically not available. The basic idea of our approach is to exploit the first order conditions of firms' profit maximization to impute the unobserved physical quantities of inputs from their expenditures.³ We then use this recovered physical quantity of intermediate inputs to replace the physical quantity of inputs in the estimation.

Due to these data restrictions, almost all approaches treat materials as a single, homogenous input. However, firms often purchase a wide variety of intermediate inputs at a variety of prices. For example, a clothing manufacturer may purchase high or low quality cloth, in addition to a variety of threads, buttons, etc. In the case where firms are selecting a vector of intermediate inputs, datasets typically contain only total input expenditure. In Section 4, We show how our method can be extended to account for a vector of unobserved heterogeneous inputs by imputing a quantity index as well as a quality-adjusted input price index for each firm. We confirm in Monte Carlo experiments that our method works well even when firms face a complicated input choice across several potential inputs.

In addition, our method recovers the underlying input price distribution, an important source of heterogeneity across firms. Accounting for input price heterogeneity can give rise to richer

their approach into our model which additionally controls for input price variation. Of course, because we assume profit maximization, it is important that our model include a demand function so that we can derive the firm's first order conditions.

²Ornaghi (2006) considers the Cobb-Douglas specification of the production function, where the elasticity of substitution is assumed to be fixed at one and leaves how price dispersion may bias the elasticity of substitution unexplored. Section 2 shows how input price dispersion also biases substitution parameters in more flexible production function specifications, which we confirm in our Monte Carlo study.

³To be precise, we recover a quality-adjusted index for the physical quantity of materials used by the firm. The associated materials price also represents a quality-adjusted price. In Section 4 we extend the model to consider the case where the firm chooses from several unobserved input types. Our procedure follows the common practice of assuming that observed inputs (labor and capital) are homogeneous to production. See Fox and Smeets (2011) for a study on the role of input heterogeneity in production function estimation.

explanations of firm policies. For example, if firms' exit decisions are modeled as a cutoff in firm productivity levels, input price heterogeneity would imply that less productive firms may remain in the market when they have access to lower input prices.

The idea of exploiting the first order conditions of profit maximization is also employed in many other studies. Taking materials inputs as given, Gandhi, Navarro, and Rivers (2013) use the transformed first-order conditions of the firm's profit maximization problem to estimate the elasticity of substitution and separate the non-structural errors as their first step in their production function estimation procedure. Doraszelski and Jaumandreu (forthcoming), also assuming labor and materials quantities are observed, use the first-order conditions of labor and material choices to impute the unobserved productivity. Together with a Markov assumption on productivity evolution, this identifies the production function parameters. Katayama, Lu, and Tybout (2009) use the first-order conditions for profit maximization to construct a welfare-based firm performance measure based on Bertrand-Nash equilibrium in a differentiated product industry when input and output prices are unobserved, instead of the traditional productivity measure. Epple, Gordon, and Sieg (2010) instead develop a procedure using the first order condition of the indirect profit function to estimate the housing supply function. Zhang (2012) uses first order conditions as constraints to directly control for structural errors to estimate a production function with biased technology change in Chinese manufacturing industries. De Loecker (2011), De Loecker and Warzynski (2012), and De Loecker, Goldberg, Khandelwal, and Pavcnik (2012) also use the first order condition of labor choice and/or material choice of profit maximization to estimate firm-level markup. The recovered markup is then used to analyze firm performance in international trade. Santos (2012) instead uses the first order condition of labor and material choices to recover demand shocks by adding a timing restriction on the sequence of input choices. Our work is also related to production estimation literature based on factor share regression (Klein, 1953; Solow, 1957; Walters, 1963), which also uses expenditure data to estimate production function using first order conditions.⁴

⁴Share regression can consistently (but may be biased as Walters (1963) point out) recover the production parameters when firms are price-takers in output market and technology shows constant return to scale. For many applications, Cobb-Douglas is a good approximation of production function. However, as is well known, it implies constant expenditure share for static inputs, even when firms face different input prices. This is not the case in the micro-level data, which usually suggests a large dispersion of expenditure shares among firms. So, we think it will be more realistic to recognize the dispersion of expenditure share, especially when the purpose is to consider firm

Our method is closest to Doraszelski and Jaumandreu (forthcoming) and Gandhi, Navarro, and Rivers (2013). These papers also assume that both material and labor choices are static and use the first order conditions of profit maximization as constraints to identify production parameters. Our method differs from these papers in both the data requirement and how we back out the unobserved productivity. Doraszelski and Jaumandreu (forthcoming) use both wage and material prices to directly back out the unobserved productivity using a timing restriction. Our method, without requiring the observation of material price (or quantity), uses the relationship between labor and materials expenditures and quantities to help back out productivity and the unobserved material quantity (and material price). Gandhi, Navarro, and Rivers (2013) show that, when materials quantities are directly observed, it is possible to use first order conditions to non-parametrically identify the production function. We rely on a parametric approach, but avoid the need to observe material quantities directly.

Our method applies to a very general set of production function parameterizations. To illustrate, we introduce our method with a general production function form. After employing our assumptions to control for productivity and unobserved price heterogeneity, the approach results in a revenue function which can be estimated by nonlinear least squares. A complication is that, in general, not all structural parameters of the production function will be identified from this revenue regression (e.g., some may only be identified up to ratios). Fortunately, we are able to provide two procedures to fully recover the structural parameters for the most common production function specifications: CES and translog. One approach relies on cross-sectional restrictions implied by the CES form, while the latter relies on panel assumptions, making use of the common assumption that productivity evolves according to a first-order Markov process to provide identifying moment restrictions. This later approach is available for extremely flexible production specifications, such as third or fourth order translog.

We demonstrate our approach using the CES production function specification, and evaluate it by carrying out a Monte Carlo study that compares its performance to the traditional estimator and an “oracle” estimator that observes input prices and quantities directly. The results show

behavior.

that our approach recovers the true parameters very well. In contrast, the traditional approach causes systematic biases in the parameter estimates. In particular, the elasticity of substitution is underestimated in the traditional approach as predicted, and the distribution parameters are also biased. This bias could mislead researchers attempting to make policy recommendations. For example, in a trade policy setting, this bias could result in erroneous counterfactual estimates of demand and supply changes of *all* inputs and outputs due to a proposed change to tariff rates on imported intermediate inputs.

We apply our approach to a plant-level data set from Colombian manufacturing industries and compare our results with those derived using the traditional estimator. The results are consistent with both our predictions and the results of the Monte Carlo experiments. That is, compared with our method, the elasticity of substitution from the traditional approach is consistently lower. Moreover, the distribution parameter estimates of the traditional method differ significantly from those of our method.

Our results indicate significant input price dispersion in all industries, providing further indication of the importance of controlling for unobserved price heterogeneity. The recovered distribution and evolution of intermediate prices are similar to that for studies in which input prices are directly observed (e.g., Atalay, 2012). We also find a positive correlation between intermediate input prices, wages and productivity, also corroborating earlier studies (Kugler and Verhoogen, 2012). Finally, the distribution of productivity estimated using our approach is even wider than using traditional approaches, suggesting that there is more productivity dispersion in Colombian manufacturing than previously thought.

The following section reviews the omitted input and output price biases in detail. Section 3 introduces a model with unobserved price heterogeneity and outlines our procedure to consistently estimate the model. Section 4 then extends the model to multiple heterogeneous inputs. Section 5 presents Monte Carlo experiments that evaluate the performance of our estimator and confirm the biases in traditional methods when unobserved price heterogeneity is present. We apply our method to a data set on Colombian manufacturing in Section 6, and conclude in Section 7.

2 Omitted Price Biases

In ideal cases where physical quantities of input and output are available, they can be used directly in the estimation of production functions (Eslava, Haltiwanger, Kugler, and Kugler, 2004; Ornaghi, 2006; Grieco and McDevitt, 2012). However, many datasets contain information on the total expenditure on intermediate inputs but not a specific accounting of their prices and quantities. In this case, the traditional approach is to use the deflated value (by industry-level price indices) of inputs and output (De Loecker, 2011) as proxy of physical quantities. This procedure implicitly requires that firms operate in perfectly competitive input markets so that all firms in the industry face the same prices. However, markets are more likely to be imperfectly competitive and are characterized by heterogenous features. For example, transportation costs may create input price differences between firms based on geography. Firm-level input and output prices vary across firms and over time, which impact the firm-level input choice. Consequently, the traditional approach will induce bias in the estimation (Klette and Griliches, 1996; Van Beveren, 2010).

Consider a production function in logarithm form $q_{jt} = f(x_{jt}, \theta_0)$, where q_{jt} is the log firm-level physical quantity of output produced by a vector of log physical input x_{jt} , and θ_0 is the parameter of interest. For commonly available firm-level production data sets, q_{jt} and x_{jt} are not available. Instead, we observe the (deflated) revenue $r_{jt} = q_{jt} + p_{jt}^q$ and (deflated) input expenditures $v_{jt} = x_{jt} + p_{jt}^x$, where p_{jt}^q and p_{jt}^x are the log firm-level output and input prices which are deflated by industry-level price indices. In the traditional estimation, q_{jt} and x_{jt} are often substituted by r_{jt} and v_{jt} . Thus, while the true model is,

$$r_{jt} = f(v_{jt} - p_{jt}^x, \theta_0) + p_{jt}^q + \epsilon_{jt},$$

the estimated model is,

$$r_{jt} = f(v_{jt}, \hat{\theta}) + u_{jt},$$

where ϵ_{jt} are i.i.d measurement errors, and u_{jt} contains both the measurement error ϵ_{jt} and the omitted terms p_{jt}^q and p_{jt}^x . Note that if input and output prices vary across firms, and input x_{jt} is chosen after these prices are observed, then the input expenditure v_{jt} is correlated with both input

and output prices. Ignoring this will result in an inconsistent estimator of θ_0 . We now provide two specific examples to illustrate the effect of unobserved price dispersion on the estimation of production functions.

Example 2.1:⁵ Consider the Cobb-Douglas production function,

$$q_{jt} = \beta_0 + \beta_l \ell_{jt} + \beta_m m_{jt} + \epsilon_{jt},$$

where lower case letters represent the logarithm value and for simplicity ϵ_{jt} is an i.i.d. measurement error.⁶ In commonly available data sets, where q_{jt} and m_{jt} are not available, revenue $r_{jt} = q_{jt} + p_{jt}^q$ and material expenditure $v_{jt}^m = m_{jt} + p_{jt}^m$ are used as proxy. Thus, the model is estimated via,

$$r_{jt} = \beta_0 + \beta_l \ell_{jt} + \beta_m v_{jt}^m + \underbrace{p_{jt}^q - \beta_m p_{jt}^m}_{u_{jt}} + \epsilon_{jt},$$

where $u_{jt} = p_{jt}^q - \beta_m p_{jt}^m + \epsilon_{jt}$ is the error term. The endogeneity problem arises since $E(v_{jt}^m u_{jt}) \neq 0$. For simplicity, suppose β_l is known, then the estimated coefficient for material is given by,⁷

$$\text{plim } \hat{\beta}_m = \beta_m + \frac{\text{cov}(v^m, p^q - \beta_m p^m)}{\text{var}(v^m)} = \beta_m + \frac{\text{cov}(v^m, p^q)}{\text{var}(v^m)} - \beta_m \frac{\text{cov}(v^m, p^m)}{\text{var}(v^m)}.$$

The estimate is inconsistent if the last two additional terms are non-zero. Note that input expenditures and output price are usually negatively correlated,⁸ we expect $\frac{\text{cov}(v^m, p^q)}{\text{var}(v^m)} < 0$. Similarly, if the correlation between materials price and materials expenditure is positive, then $\frac{\text{cov}(v^m, p^m)}{\text{var}(v^m)} > 0$. In this case, we have $\hat{\beta}_m < \beta_m$. So the material coefficients will be underestimated if quantities of material and output are substituted by deflated values. However, in the case that material expenditure and material price are negatively related, $\frac{\text{cov}(v^m, p^m)}{\text{var}(v^m)} < 0$, the direction of the bias is ambiguous. In both cases, input price bias is present as long as material choice depends on the omitted material price. Finally, it is easy to see that the same mechanism will induce inconsistency

⁵Ornaghi (2006) shows a similar example.

⁶Here we disregard the presence of unobserved productivity difference across firms for clarity. Our model in section 3 will account for the unobserved productivity.

⁷If β_l is not known and β_l and β_m are estimated jointly, then $\hat{\beta}_l$ will also be biased.

⁸Klette and Griliches (1996, 346) lists cases where this is true.

in the more flexible translog functional form, which also leads to a log-linear specification of the production function. ■

Example 2.2: For the CES production function, ignoring input price heterogeneity can bias the elasticity of substitution between inputs. To see this, consider the CES production function (without capital for simplicity),

$$Q_{jt} = F(M_{jt}, L_{jt}) = [\alpha M_{jt}^\gamma + (1 - \alpha)L_{jt}^\gamma]^\frac{1}{\gamma},$$

where $\gamma = \frac{\sigma-1}{\sigma}$, and σ is the constant elasticity of substitution which is defined as

$$\sigma = -\frac{\partial \ln(M/L)}{\partial \ln(F_M/F_L)},$$

where F_M and F_L are the marginal output of material and labor. If (deflated) expenditures (denoted as $E_M = P_M M$ and $E_L = P_L L$) are used instead of firm-level quantities, then the estimated model is

$$Q_{jt} = \hat{F}(E_{M_{jt}}, E_{L_{jt}}) = [\hat{\alpha} E_{M_{jt}}^{\hat{\gamma}} + (1 - \hat{\alpha}) E_{L_{jt}}^{\hat{\gamma}}]^\frac{1}{\hat{\gamma}},$$

where $\hat{\gamma} = \frac{\hat{\sigma}-1}{\hat{\sigma}}$. In particular, the elasticity of substitution $\hat{\sigma}$ is calculated using labor expenditure and material expenditure which is different from the original definition:

$$\begin{aligned} \hat{\sigma} &= -\frac{\partial \ln(E_M/E_L)}{\partial \ln(\hat{F}_{E_M}/\hat{F}_{E_L})} = -\frac{\partial \ln(E_M/E_L)}{\partial \ln(F_M/F_L)} \times \frac{\partial \ln(F_M/F_L)}{\partial \ln(\hat{F}_{E_M}/\hat{F}_{E_L})} \\ &= \left(-\frac{\partial \ln(M/L)}{\partial \ln(F_M/F_L)} - \frac{\partial \ln(P_M/P_L)}{\partial \ln(F_M/F_L)} \right) \times \frac{\partial \ln(F_M/F_L)}{\partial \ln(\hat{F}_{E_M}/\hat{F}_{E_L})} \\ &= (\sigma - 1) \times \frac{\partial \ln(F_M/F_L)}{\partial \ln(\hat{F}_{E_M}/\hat{F}_{E_L})} \end{aligned}$$

The last equation holds because $\sigma = -\frac{\partial \ln(M/L)}{\partial \ln(F_M/F_L)}$ by definition and $\frac{\partial \ln(P_M/P_L)}{\partial \ln(F_M/F_L)} = 1$ if the firm chooses M and L to minimize its variable cost. This equation shows that the bias comes from two sources. The first source is the use of (E_M, E_L) rather than (M, L) for the elasticity of substitution.

That is,

$$\frac{\partial \ln(E_M/E_L)}{\partial \ln(F_M/F_L)} \neq \frac{\partial \ln(M/L)}{\partial \ln(F_M/F_L)}.$$

The second source of bias is due to inconsistency between the estimated production function and the true production function due to the use of (E_M, E_L) instead of (M, L) in estimation, i.e.,

$$\frac{\partial \ln(F_M/F_L)}{\partial \ln(\hat{F}_{E_M}/\hat{F}_{E_L})} \neq 1.$$

If $\frac{\partial \ln(F_M/F_L)}{\partial \ln(\hat{F}_{E_M}/\hat{F}_{E_L})}$ is less than 1, then the elasticity of substitution is underestimated ($\hat{\sigma} < \sigma$). The intuition is that because of cost minimization, the physical input ratio will change in a direction against the change in input price ratio. As a result, the change in the input expenditure ratio E_M/E_L is offset partially by the change in P_M/P_L . For example, suppose there is an increase in material price, so P_M/P_L rises. With cost minimization, the physical input M will be partially substituted by labor, thus M/L drops. However, the percentage drop in the expenditure ratio E_M/E_L is less than that in M/L , because $E_M/E_L = (P_M/P_L) \cdot (M/L)$ and P_M/P_L rose. As a result, the estimated elasticity of substitution measured using expenditure proxies will be biased. Note that this also implies the distribution parameter (α) is also biased. ■

These examples show that when firms face heterogeneous input prices, traditional production function estimates that rely on materials expenditure to proxy for the quantity of materials are inconsistent. The bias affect estimates of both output elasticity and the elasticity of substitution between inputs. It is easy to see how this inconsistency could lead to misleading counterfactual analysis to policy questions of interest. For example, consider a proposed tariff increase on some intermediate input: researchers who estimate the production function while failing to account for input price heterogeneity would mis-predict both the change in firm output and the degree of substitution from the imported input to other inputs resulting from the tariff. In the rest of the paper, we propose and demonstrate a structural approach which uses information on the relative expenditure on inputs to control for unobserved input price and consistently estimate the production function.

3 Estimation with Unobserved Price Dispersion

In this section, we introduce a model of firms' decision-making in a standard monopolistically competitive output market. The goal is to find an approach to estimating the production function when we have the commonly available data on output value, all inputs values, wage rate and labor quantity. Instead of substituting quantities with deflated values, our approach exploits the first order conditions implied by profit maximization to impute unavailable physical quantities of intermediate inputs from expenditures.

3.1 The General Model

To fix ideas, we first present our approach for general forms of production and demand functions and outline our approach for consistent estimation. In the next subsections, we will show how our approach can be used to estimate two commonly used specifications of the production function: CES and translog.

Suppose at each period t , each firm j produces a single product using labor (L_{jt}), intermediate material (M_{jt}), and capital (K_{jt}) using the production function,

$$Q_{jt} = \exp(\omega_{jt})F(L_{jt}, M_{jt}, K_{jt}; \theta),$$

where Q_{jt} is the output, ω_{jt} is a Hicks-neutral productivity shock observed by the firm (but not by researchers), and θ is the set of parameters in the production function. For notational clarity, we assume that materials is a scalar input, however we show that this can be generalized in Section 4.

The inverse demand function is,

$$P_{jt} = P_t(Q_{jt}; \eta),$$

where P_{jt} is the output price and η is the set of parameters in the demand function. The inverse demand function $P_t(\cdot, \cdot)$ is continuous and decreasing in its first coordinate. We allow the demand to be different over time. We make the following assumptions:

Assumption 1 (Smooth Production Function) *Production function $F(\cdot)$ is known up to a*

finite dimensional parameter θ , strictly increasing in inputs, and continuously differentiable up to second order.

Assumption 2 (Exogenous Input Prices) *Firms are price takers in input markets. Suppliers use linear pricing, but input prices are allowed to be different across firms and over time.*

Assumption 3 (Profit Maximization) *After observing their productivity draw, ω_{jt} , firms optimally choose labor and material inputs to maximize the profit in each period. The firm's capital stock for period t is chosen prior to the revelation of ω_{jt} .*

Several points are worth highlighting. As in Olley and Pakes (1996) and Levinsohn and Petrin (2003), capital is pre-fixed in the short run while both labor and material inputs are flexibly chosen at the beginning of each period.⁹ However, in contrast to the previous literature, these choices depend on idiosyncratic input prices. This is an additional source of firm heterogeneity along with the well-known Hicks-neutral technology shifter, ω_{jt} . The assumption that firms are price takers does not preclude firms being offered different prices on the basis of their size (i.e., capital stock), productivity, or negotiating ability, but does assume that firms do not receive “quantity discounts,” which would endogenously affect purchasing decisions. As shown in the previous section, ignoring the variation in input prices and using the deflated input expenditures as proxies of physical inputs will introduce a bias into the estimation.

To control for price and productivity differences across firms, we use an explicit model of profit maximization to construct estimation equations involving only commonly available expenditure data. Using first order conditions implies stronger assumptions in terms of static profit maximization/cost minimization by choosing material and labor. In particular, we exploit the structure of the first order conditions to impute materials prices and quantities from expenditure data.

While, relative to Olley and Pakes (1996), we strengthen some assumptions by requiring profit maximization, we are able to relax others. Because we use the first order conditions to recover the

⁹It is likely that labor may not be easily adjusted especially when a very large adjustment is needed and this will make this assumption less appealing. However, in practice firms can use temporary workers and overtime to meet production fluctuations. This implies that labor and material are still flexible if the fluctuation is not too large. Ultimately, the validity of this assumption will depend on the industry under study.

unobserved productivity, ω_{jt} , we will not need to use a “proxy” (such as investment) to recover it. Indeed, investment will not be used in our procedure at all, so there is no need for an invertability condition on the investment function. Instead, materials quantities and productivity will be jointly recovered from the two first order conditions.

3.2 Recovering the Unobserved Input Prices

We assume the econometrician observes revenue $R_{jt} = P_{jt}Q_{jt}$, inputs expenditure $E_{M_{jt}} = P_{M_{jt}}M_{jt}$, wage rate $P_{L_{jt}}$, number of workers or number of working hours L_{jt} , and capital stock K_{jt} . But she does not observe the material inputs prices or quantities ($P_{M_{jt}}$ and M_{jt}) or output prices and quantities (P_{jt} and Q_{jt}).

At the beginning of each period, after observing capital stock, K_{jt} , productivity shock, ω_{jt} and idiosyncratic input prices $P_{L_{jt}}$ and $P_{M_{jt}}$, firm j chooses its own labor and materials inputs to maximize its period profit. The firm’s static decision problem is:

$$\begin{aligned} \max_{L_{jt}, M_{jt}} & P_t(Q_{jt})Q_{jt} - P_{L_{jt}}L_{jt} - P_{M_{jt}}M_{jt} \\ \text{s.t.} & Q_{jt} = \exp(\omega_{jt})F(L_{jt}, M_{jt}, K_{jt}; \theta). \end{aligned}$$

The corresponding first order conditions are,

$$\begin{aligned} \exp(\omega_{jt})F_{L_{jt}} \left[P_t(Q_{jt}) + Q_{jt} \frac{\partial P_t(Q_{jt})}{\partial Q_{jt}} \right] &= P_{L_{jt}}, \\ \exp(\omega_{jt})F_{M_{jt}} \left[P_t(Q_{jt}) + Q_{jt} \frac{\partial P_t(Q_{jt})}{\partial Q_{jt}} \right] &= P_{M_{jt}}. \end{aligned} \tag{1}$$

Dividing the two first order conditions, multiplying both sides by $\frac{L_{jt}}{M_{jt}}$, and rearranging yield,

$$\frac{F_{L_{jt}}L_{jt}}{F_{M_{jt}}M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} = 0, \tag{2}$$

where $F_{L_{jt}}$ and $F_{M_{jt}}$ are the partial derivatives of F with respect to labor and material, and $E_{L_{jt}} = P_{L_{jt}}L_{jt}$ and $E_{M_{jt}} = P_{M_{jt}}M_{jt}$ are expenditures on labor and material, which are observed in the data.

Equation (2) is the key to our approach. It relates labor and the intermediate input, given that they are optimally chosen to maximize profits. The following proposition gives conditions under which we are able to impute M_{jt} from the observed data.

Proposition 1 *Define,*

$$z(M_{jt}, L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta) = \frac{F_{L_{jt}} L_{jt}}{F_{M_{jt}} M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}}.$$

Suppose assumptions 1-3 are satisfied and we observe a random sample of $(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}})$ generated by model parameter $\theta \in \Theta \in \mathbb{R}^h$ ¹⁰, which implies that $Z(M_{jt}, L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta) = 0$ $\frac{\partial z}{\partial M_{jt}} \neq 0$. Then there exists a neighborhood $V_\delta(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta) \in \mathbb{R}^{h+4}$, an open set $W \in \mathbb{R}$ containing M_{jt} , and a real valued function $M^ : V \rightarrow W$, such that,*

$$z(M^*(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta), L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta) = 0.$$

The proof for this proposition is an application of the implicit function theorem and is provided in Appendix A. The imputation problem is essentially about whether we can separate M_{jt} and $P_{M_{jt}}$ from what we can observe, i.e., $E_{M_{jt}} = P_{M_{jt}} M_{jt}$. The condition that $\frac{\partial z}{\partial M_{jt}} \neq 0$ has intuitive economic meaning. It implies that we observe firms optimally reacting to changes in input prices ($P_{M_{jt}}$ and $P_{L_{jt}}$) by adjusting their expenditures on intermediate inputs (M_{jt}) and labor (L_{jt}) and the intermediate-labor output elasticity ratio. In other words, $\frac{\partial z}{\partial M_{jt}} \neq 0$ implies that the elasticity of substitution between materials and labor (i.e. σ) is not always one. To see why, suppose $\sigma = 1$. By definition, $\sigma = -\frac{\partial \ln(L_{jt}/M_{jt})}{\partial \ln(F_{L_{jt}}/F_{M_{jt}})} = 1$. This implies that $\frac{F_{L_{jt}} L_{jt}}{F_{M_{jt}} M_{jt}}$ is a constant and $\frac{\partial z}{\partial M_{jt}} = 0$.

While Proposition 1 guarantees that there is at least one material price which satisfies the first order conditions, it does not guarantee that it is unique. Fortunately, for most common specifications of the production function, monotonicity and concavity assumptions ensure a unique solution. As we will see below, in the CES specification, it is possible to derive a closed form solution for M^* . For the remainder of the paper, we will assume that M^* is unique.

Once we recover $M_{jt}^* = M^*(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta)$, we can replace the unobserved intermediate

¹⁰In this proposition, h is the dimension of parameter set θ ; δ is the radius of an open ball

inputs M_{jt} in (1) to back out the productivity shock as $\omega_{jt}^* = \omega^*(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta)$. In turn, we can substitute M_{jt}^* and ω_{jt}^* into an estimation equation, where the only remaining unknown is the exogenous error term, so the model can be estimated via nonlinear least squares.

The following three examples illustrate Proposition 1. In the first example, we show that the condition is *not* satisfied for the Cobb-Douglas specification. This is because, as is well-known, the Cobb-Douglas production function assumes that expenditure shares are constant within the data, eliminating the source of variation we need to separate input prices and quantities. In Example 3.2, we show that for the more general CES production function, we can impute materials quantities as long as the elasticity of substitution is not 1 (the Cobb-Douglas case). We are able to test for this restriction by observing whether the expenditure shares are constant across firms in the data. Finally, we show that under the translog specification, M^* can be recovered under a simple invertibility condition.

Example 3.1 (Cobb-Douglas Production Function)

For Cobb-Douglas production function

$$Q_{jt} = \exp(\omega_{jt})F(L_{jt}, M_{jt}, K_{jt}; \theta) = \exp(\omega_{jt})L_{jt}^{\alpha_L} M_{jt}^{\alpha_M} K_{jt}^{\alpha_K},$$

where $\theta = (\alpha_L, \alpha_M, \alpha_K)$, it is straightforward to show that,

$$\begin{aligned} z(M_{jt}, L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta) &= \frac{F_{L_{jt}} L_{jt}}{F_{M_{jt}} M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \\ &= \frac{\alpha_L K_{jt}^{\alpha_K} L_{jt}^{\alpha_L - 1} M_{jt}^{\alpha_M} L_{jt}}{\alpha_M K_{jt}^{\alpha_K} L_{jt}^{\alpha_L} M_{jt}^{\alpha_M - 1} M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \\ &= \frac{\alpha_L}{\alpha_M} - \frac{E_{L_{jt}}}{E_{M_{jt}}}. \end{aligned}$$

We cannot recover the unobserved material from equation (2), since $\frac{\partial z}{\partial M_{jt}} = 0$. The intuition is that, because the elasticity of substitution is fixed at one, when the relative inputs price ($\frac{P_L}{P_M}$) changes firms always choose labor and material such that the percentage increase (or decrease) of the labor-material ratio ($\frac{L}{M}$) equals the percentage decrease (or increase) of relative price ($\frac{P_L}{P_M}$). As

a result, the expenditure ratio $\frac{E_{L_{jt}}}{E_{M_{jt}}}$ remains constant ($\frac{\alpha_L}{\alpha_M}$). In this case, we cannot separate the price and quantity of materials from the information on expenditure ratio $\frac{E_{L_{jt}}}{E_{M_{jt}}}$. ■

Of course, because we observe the expenditure ratio between materials and labor in the data, it is easy to verify that it is not constant across all firms. As long as there is variation, it is reasonable to specify a production function that allows the ratio to vary and use this variation to impute materials prices and quantities.

Example 3.2 (CES Production Function)

Consider CES production function,¹¹

$$Q_{jt} = \exp(\omega_{jt})F(L_{jt}, M_{jt}, K_{jt}; \theta) = \exp(\omega_{jt})[\alpha_L L_{jt}^\gamma + \alpha_M M_{jt}^\gamma + \alpha_K K_{jt}^\gamma]^{\frac{1}{\gamma}}, \quad (3)$$

where $\gamma = \frac{\sigma-1}{\sigma}$ (σ is the elasticity of substitution), and $\theta = (\alpha_L, \alpha_M, \alpha_K, \sigma)$. We can show that,

$$\begin{aligned} z(M_{jt}, L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}, \theta) &= \frac{F_{L_{jt}} L_{jt}}{F_{M_{jt}} M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \\ &= \frac{\exp(\omega_{jt})[\alpha_L L_{jt}^\gamma + \alpha_M M_{jt}^\gamma + \alpha_K K_{jt}^\gamma]^{\frac{1}{\gamma}-1} \alpha_L L_{jt}^{\gamma-1} L_{jt}}{\exp(\omega_{jt})[\alpha_L L_{jt}^\gamma + \alpha_M M_{jt}^\gamma + \alpha_K K_{jt}^\gamma]^{\frac{1}{\gamma}-1} \alpha_M M_{jt}^{\gamma-1} M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \\ &= \frac{\alpha_L L_{jt}^\gamma}{\alpha_M M_{jt}^\gamma} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \end{aligned}$$

When $\gamma \neq 0$ (i.e., $\sigma \neq 1$), we can show that,¹²

$$\frac{\partial z}{\partial M_{jt}} = -\gamma \frac{\alpha_L L_{jt}^\gamma}{\alpha_M M_{jt}^{\gamma+1}} \neq 0,$$

and we can recover the unobserved material from equation (2). In this case, equation (2) has a closed-form solution,

$$M^*(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta) = \left(\frac{\alpha_L}{\alpha_M} \frac{E_{M_{jt}}}{E_{L_{jt}}} \right)^{\frac{1}{\gamma}} L_{jt}.$$

¹¹Below we will work with a normalized form of the CES production function, but we use an unnormalized form here for expositional simplicity.

¹²When $\gamma = 0$ (i.e., $\sigma = 1$), the CES function is equivalent to the Cobb-Douglas case (Example 3.1) and we cannot recover the unobserved materials from equation (2). This case must be excluded from the parameter set Θ .

Intuitively, it is possible to infer information about material quantity M_{jt} from the inputs expenditure ratio $\frac{E_{M_{jt}}}{E_{L_{jt}}}$. This feature, together with the definition $E_{M_{jt}} = P_{M_{jt}}M_{jt}$ help us separate the quantity M_{jt} and material price $P_{M_{jt}}$ from each other. ■

Example 3.3 (Translog Production Function)

Finally, we consider the translog production function specification,

$$\begin{aligned} Q_{jt} &= \exp(\omega_{jt})F(L_{jt}, M_{jt}, K_{jt}; \theta) \\ &= \exp(\omega_{jt}) \exp \left\{ \alpha_k \ln K_{jt} + \alpha_l \ln L_{jt} + \alpha_m \ln M_{jt} + \frac{1}{2}\alpha_{kk} (\ln K_{jt})^2 + \frac{1}{2}\alpha_{ll} (\ln L_{jt})^2 + \frac{1}{2}\alpha_{mm} (\ln M_{jt})^2 \right. \\ &\quad \left. + \alpha_{kl} (\ln K_{jt}) (\ln L_{jt}) + \alpha_{km} (\ln K_{jt}) (\ln M_{jt}) + \alpha_{lm} (\ln L_{jt}) (\ln M_{jt}) \right\} \end{aligned} \quad (4)$$

Where $\theta = (\alpha_k, \alpha_l, \alpha_m, \alpha_{kk}, \alpha_{ll}, \alpha_{mm}, \alpha_{kl}, \alpha_{km}, \alpha_{lm})$ are the structural parameters. The translog is a more flexible generalization of the Cobb-Douglass production function which allows for the elasticity of substitution to be a function of the inputs. Under this specification,

$$\begin{aligned} z(M_{jt}, L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}, \theta) &= \frac{F_{L_{jt}}L_{jt}}{F_{M_{jt}}M_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \\ &= \frac{\alpha_l + \alpha_{ll} \ln L_{jt} + \alpha_{kl} \ln K_{jt} + \alpha_{ml} \ln M_{jt}}{\alpha_m + \alpha_{mm} \ln M_{jt} + \alpha_{km} \ln K_{jt} + \alpha_{ml} \ln L_{jt}} - \frac{E_{L_{jt}}}{E_{M_{jt}}} \end{aligned}$$

It is now easy to show that $\frac{\partial z}{\partial M_{jt}} \neq 0$ as long as $E_{M_{jt}} > 0$ and $\alpha_{ml} - \frac{E_{L_{jt}}}{E_{M_{jt}}}\alpha_{mm} \neq 0$, which will hold generically in the data as long as there is variation in the expenditure ratio and $\alpha_{mm} \neq 0$.¹³

Like in the CES case, this restriction relates to our inability to recover M_{jt} in the Cobb-Douglass specification, which is a special case of the translog in which $\alpha_{ml} = \alpha_{mm} = 0$. When this invertibility condition does hold, we can derive a closed form solution for materials,

¹³A sufficient condition for $\alpha_{ml} - \frac{E_{L_{jt}}}{E_{M_{jt}}}\alpha_{mm} \neq 0$ to always hold is if $\alpha_{mm}\alpha_{ml} < 0$.

$$M^*(L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta) = \exp \left(\frac{\frac{E_{Ljt}}{E_{Mjt}} (\alpha_m + \alpha_{km} \ln K_{jt} + \alpha_{ml} \ln L_{jt}) - (\alpha_l + \alpha_{ll} \ln L_{jt} + \alpha_{kl} \ln K_{jt})}{(\alpha_{ml} - \alpha_{mm} \frac{E_{Ljt}}{E_{Mjt}})} \right). \quad (5)$$

Finally, we note that for higher order translog specifications, a similar procedure may be available, but it will necessitate finding the roots of a polynomial (rather than linear) equation in M , introducing the possibility that $z(\cdot)$ may have multiple solutions.¹⁴ ■

3.3 Estimation Equation

We now turn to estimation of the parameters of the production function, θ , and the inverse demand function, η , using commonly available data on output value, all input values, and the wage rate. If our assumptions are satisfied, the intermediate input quantity can be uniquely imputed as,

$$M_{jt}^* = M^*(L_{jt}, M_{jt}, E_{Mjt}, E_{Ljt}; \theta), \quad (6)$$

Once the quantity of material M_{jt} is recovered, we can plug it back into either of the first order conditions and recover the unobserved productivity,

$$\omega_{jt}^* = \omega^*(L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}, M_{jt}^*; \theta). \quad (7)$$

Different from Olley and Pakes (1996), here we recover the unobserved productivity parametrically from the firm's first order conditions.¹⁵ There are several advantages to this method. First, the estimation does not require the investment data. Second, there is no need to rely on invertibility of the investment policy function, which may be problematic when adjustment costs generate lumpiness in the optimal investment policy. Moreover, our method of controlling for endogeneity does

¹⁴It would be ideal to be able to impute M^* for a nonparametric production function. Unfortunately, our method requires a parametric approach to recovering M^* , as is illustrated by these examples. Nonetheless, we can in principle accommodate an arbitrarily flexible parametric specification.

¹⁵Zhang (2012) uses a similar approach to allow for biased technical change in the production function.

not require the Markov assumption on the productivity evolution process. Of course, assumptions on the productivity evolution process may still be needed in identifying the production function, as we will discuss below. Finally, as we fully exploit the structural assumptions of the parametric production function and corresponding first order conditions to recover unobserved productivity and material quantity, as in Doraszelski and Jaumandreu (forthcoming), we do not have to rely on nonparametric methods to estimate these functions. As a result, even when both expenditures may be functions of the same set of variables in equations (8) (implicit in ω_{jt}^*) and (12) below, we still have identification as long as labor and materials expenditures are not perfectly correlated.¹⁶ Akerberg, Caves, and Frazer (2006) also discussed this possibility (page 16, version of December 28, 2006).

Since output quantities are not directly observed, we follow Klette and Griliches (1996) and use the revenue function as the estimating equation. The revenue function is,

$$R_{jt} = \exp(u_{jt})P_t(Q_{jt}; \eta) Q_{jt}.$$

Where R_{jt} is the observed revenue of the firm, $Q_{jt} = e^{\omega_{jt}^*} F(L_{jt}, M_{jt}^*, K_{jt}; \theta)$ is the predicted quantity of physical output based on observed inputs and the model parameters (θ, η) , and u_{jt} is a revenue error term which incorporates measurement error as well as demand and productivity shocks that are unanticipated by the firm and are assumed to be i.i.d. across firms and over time. Taking the logarithm of the revenue function yields,

$$\ln R_{jt} = \ln P_t \left(e^{\omega_{jt}^*} F(L_{jt}, M_{jt}^*, K_{jt}; \theta); \eta \right) + \ln \left[e^{\omega_{jt}^*} F(L_{jt}, M_{jt}^*, K_{jt}; \theta) \right] + u_{jt} \quad (8)$$

In this equation, the unobserved productivity and material quantity, ω_{jt}^* and M_{jt}^* , are recovered as functions of observed variables as in equations (6) and (7). The only remaining unobservable, u_{jt} , is unknown to the firm and is uncorrelated with the observed inputs. Therefore, we can consistently estimate (8) via non-linear least squares. However, in general, all parameters of the production function, θ , may not be identified by the revenue function. Indeed, this is the case in both of

¹⁶The data in our application shows that the labor-material expenditure ratio has large variation across firms (as required by our empirical model), supporting the idea that these two expenditures are not perfectly correlated.

our primary examples, the CES and translog specifications. Fortunately, in both cases, additional restrictions are available to identify all of the structural parameters. For the CES, these restrictions are related to aggregate measures and do not involve any additional assumptions. For the translog, we rely on the additional assumption that productivity moves according to a Markov process to provide moment restrictions with which to identify the remaining parameters. This second approach is quite general and can provide identifying restrictions for many functional forms (including the CES, if it were necessary). The following two subsections illustrate the implementation of these two specifications.

3.4 Implementation with CES specification

As presented in the examples above, we consider a CES production function with an elasticity of substitution σ .¹⁷ It has been commonly recognized that the CES production function needs to be normalized to give meaningful identification of its parameters. A branch of the literature has analyzed the importance and the method of normalization (de La Grandville, 1989; Klump and de La Grandville, 2000; Klump and Preissler, 2000; de La Grandville and Solow, 2006; Leon-Ledesma, McAdam, and Willman, 2010). We follow this literature and normalize the CES production function according to the geometric mean.¹⁸ Specifically, let the baseline point for our normalization be the geometric mean of $(Q_{jt}, L_{jt}, M_{jt}, K_{jt})$, denoted as $\bar{Z} = (\bar{Q}, \bar{L}, \bar{M}, \bar{K})$ where $\bar{X} = \sqrt[n]{X_1 X_2 \cdots X_n}$.¹⁹ Then the normalized CES production function can be written as,

$$Q_{jt} = e^{\omega_{jt}} \bar{Q} \left[\alpha_L \left(\frac{L_{jt}}{\bar{L}} \right)^\gamma + \alpha_M \left(\frac{M_{jt}}{\bar{M}} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{\bar{K}} \right)^\gamma \right]^{\frac{1}{\gamma}}, \quad (9)$$

¹⁷We follow the literature in assuming constant returns to scale in this specification. This assumption can be relaxed by adding a scale parameter. This does not affect the estimation procedure but make the scale parameter and demand elasticity not separately identified. However, if Markov process of productivity is assumed, then one can easily (using Equation (55) in Appendix D) separately identify these two parameters using the variation in industrial-level output and prices.

¹⁸For the details of this normalization and how we implement it in this paper, see Appendix B.

¹⁹In principle, any point $Z_0 = (Q_0, L_0, M_0, K_0)$ (which satisfies normalization conditions, i.e., Equation (43)-(45) in Appendix B) can be chosen as the baseline point, for example a default choice could be $(1, \dots, 1)$. The entire CES production function is identified up to the knowledge of the baseline point.

where $\gamma = \frac{\sigma-1}{\sigma}$ and $\alpha_L, \alpha_M, \alpha_K$ are the distribution parameters, which sum to 1,

$$\alpha_L + \alpha_M + \alpha_K = 1. \quad (10)$$

The normalization has three advantages for our purposes. First, it scales the level of inputs according to an industry average, eliminating the effect of units on the parameters. Second, the geometric mean (\bar{Z}) is computable using the observed data, and will be convenient to use in constructing an additional restriction to identify the distribution parameters. Third, this scaling gives the distribution parameters a precise interpretation. Specifically, they are the marginal return to inputs (in normalized units) for a firm with the geometric mean level of inputs, productivity, and input prices.²⁰

Since analysts typically observe revenues rather than output quantities, we follow Klette and Griliches (1996) and also specify a classic Dixit-Stiglitz demand function,

$$\frac{Q_{jt}}{Q_t} = \left(\frac{P_{jt}}{P_t} \right)^\eta, \quad (11)$$

where Q_t and P_t are industry-level output quantity and price in period t , and η is the demand elasticity. As discussed earlier, the CES production function satisfies the condition of Proposition 1 if $\sigma \neq 1$ (i.e., $\gamma \neq 0$).

Given our specification for the production function (9) and demand function (11), we follow our proposed procedure from the previous section to derive the estimating revenue equation,²¹

$$\ln R_{jt} = \ln \frac{\eta}{1 + \eta} + \ln \left[E_{M_{jt}} + E_{L_{jt}} \left(1 + \tau \left(\frac{K_{jt}/\bar{K}}{L_{jt}/\bar{L}} \right)^\gamma \right) \right] + u_{jt}, \quad (12)$$

where $\tau = \alpha_K/\alpha_L$.²² It is easy to see that while (12) provides identification of the elasticity of substitution and the slope of the demand curve, it does not identify the distribution parameters.

This is due to the substitution of our structural equation for the unobserved materials inputs.

²⁰Note that the normalized input of the baseline point is simply (1, 1, 1).

²¹See Appendix C for the complete derivation.

²²See Appendix C for a detailed economic interpretation of τ .

Fortunately, with the estimate of τ in hand, we can recover the distribution parameters by making use of two additional restrictions. The first is simply the adding up constraint (10), the second is implied by the first order conditions of profit maximization. To see this, note that the first order conditions for each firm j at period t are

$$\lambda \bar{\alpha}_L e^{\omega_{jt}} \bar{Q} \left[\alpha_L \left(\frac{L_{jt}}{\bar{L}} \right)^\gamma + \alpha_M \left(\frac{M_{jt}}{\bar{M}} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{\bar{K}} \right)^\gamma \right]^{\frac{1}{\gamma}} \left(\frac{L_{jt}}{\bar{L}} \right)^{\gamma-1} \frac{1}{\bar{L}} = P_{L_{jt}},$$

and

$$\lambda \bar{\alpha}_M e^{\omega_{jt}} \bar{Q} \left[\alpha_L \left(\frac{L_{jt}}{\bar{L}} \right)^\gamma + \alpha_M \left(\frac{M_{jt}}{\bar{M}} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{\bar{K}} \right)^\gamma \right]^{\frac{1}{\gamma}} \left(\frac{M_{jt}}{\bar{M}} \right)^{\gamma-1} \frac{1}{\bar{M}} = P_{M_{jt}},$$

where λ is the lagrangian multiplier. The ratio of the two equations yields,

$$\frac{\alpha_L (L_{jt}/\bar{L})^\gamma}{\alpha_M (M_{jt}/\bar{M})^\gamma} = \frac{P_{L_{jt}} L_{jt}}{P_{M_{jt}} M_{jt}} \equiv \frac{E_{L_{jt}}}{E_{M_{jt}}}. \quad (13)$$

This equation holds for each firm j at each period t . Taking the geometric mean of (13) across all observations implies,²³

$$\frac{\alpha_M}{\alpha_L} = \frac{\bar{E}_M}{\bar{E}_L}, \quad (14)$$

where \bar{E}_M and \bar{E}_L are the geometric mean of $E_{M_{jt}}$ and $E_{L_{jt}}$ respectively. Because expenditures on materials and labor are observed in the data for all observations, the right hand side of this restriction can be directly computed.

Then, we combine (10), (14) and the estimate of τ to derive explicit formulas for the distribution parameter estimates in terms of observables and $\hat{\tau}$,

$$\begin{aligned} \hat{\alpha}_L &= \frac{\bar{E}_L}{\bar{E}_L + \bar{E}_M + \hat{\tau} \bar{E}_L} \\ \hat{\alpha}_M &= \frac{\bar{E}_M}{\bar{E}_L + \bar{E}_M + \hat{\tau} \bar{E}_L} \\ \hat{\alpha}_K &= 1 - \hat{\alpha}_L - \hat{\alpha}_M \end{aligned} \quad (15)$$

That is, once $\hat{\tau}$ is estimated using (12), the distribution parameters estimates $\hat{\alpha}_L$, $\hat{\alpha}_M$ and $\hat{\alpha}_K$

²³Recall that the geometric mean of a ratio is the ratio of geometric means.

can be recovered by (15).

With all of the production function parameters estimated we can solve for each firm's productivity level from (7), which is reduced to (55) in Appendix D.

3.5 Implementation with Translog specification

We now turn to the translog specification, which was introduced in (4). As with the CES implementation, we assume demand follows a Dixit-Stiglitz specification (11). Again, not all parameters of the production function are identified by the revenue equation. However we can use time series assumptions as additional restrictions to identify the remaining parameters. In particular, we follow Olley and Pakes (1996) to assume that firm productivity follows first order Markov process in order to generate the needed moment restrictions.

Example 3.3 shows how to recover materials. We can insert equation (5) into one of the first order conditions to uniquely recover the unobserved productivity ω_{jt}^* ,

$$\omega_{jt}^*(\theta, \eta) = \frac{1}{1 + 1/\eta} [\ln P_{Ljt} - \ln F_{Ljt} - \ln(1 + 1/\eta) - 1/\eta \ln F_{jt}]. \quad (16)$$

Substituting (16) into the log revenue function yields our estimating equation,

$$\begin{aligned} \ln R_{jt} &= \ln P_{Ljt} - \ln F_{Ljt} - \ln(1 + 1/\eta) + u_{jt} \\ &= \ln P_{Ljt} - \ln(1 + 1/\eta) - \ln \left\{ \alpha_l + \alpha_{ll} \ln L_{jt} + \alpha_{kl} \ln K_{jt} \right. \\ &\quad \left. + \alpha_{ml} \left(\frac{\frac{E_{Ljt}}{E_{Mjt}} (\alpha_m + \alpha_{km} \ln K_{jt} + \alpha_{ml} \ln L_{jt}) - (\alpha_l + \alpha_{ll} \ln L_{jt} + \alpha_{kl} \ln K_{jt})}{(\alpha_{ml} - \alpha_{mm} \frac{E_{Ljt}}{E_{Mjt}})} \right) \right\} + u_{jt}. \end{aligned} \quad (17)$$

It is now clear that all production and demand parameters, except a_k and α_{kk} , are identified from the revenue equation. The remaining parameters, α_k and α_{kk} , are canceled when computing $\ln F_{Ljt}$.

Therefore, we need additional restrictions to help identify a_k and a_{kk} , either from the cross section restriction or the time series restriction. With the translog specification, cross sections restriction are not easy to find. Instead we follow Olley and Pakes (1996) and Doraszelski and Jaumandreu (forthcoming) in using time series restrictions on productivity to help identify remaining

parameters. In particular, we assume that productivity follows a first order Markov process,

$$\omega_{jt} = g(\omega_{jt-1}) + \epsilon_{jt}.$$

With this additional restriction, we can identify all parameters. The estimation can be done in two stages. In the first stage, we use (17) to estimate $\hat{\alpha}_l, \hat{\alpha}_m, \hat{\alpha}_{kl}, \hat{\alpha}_{km}, \hat{\alpha}_{lm}, \hat{\alpha}_{ll}, \hat{\alpha}_{mm}$, and $\hat{\eta}$. Then define,

$$\begin{aligned} \hat{\phi}_{jt}(\alpha_k, \alpha_{kk}) &= \ln F_{jt} - \alpha_k \ln K_{jt} - \frac{1}{2} \alpha_{kk} (\ln K_{jt})^2 \\ &= \hat{\alpha}_l \ln L_{jt} + \hat{\alpha}_m \ln M_{jt}^* + \frac{1}{2} \hat{\alpha}_{ll} (\ln L_{jt})^2 + \frac{1}{2} \hat{\alpha}_{mm} (\ln M_{jt}^*)^2 \\ &\quad + \hat{\alpha}_{kl} (\ln K_{jt}) (\ln L_{jt}) + \hat{\alpha}_{km} (\ln K_{jt}) (\ln M_{jt}^*) + \hat{\alpha}_{lm} (\ln L_{jt}) (\ln M_{jt}^*) \end{aligned}$$

with $\ln M_{jt}^*$ replaced by equation (5).²⁴ Then according to equation (16) the productivity for firm j at time t for a given value of the remaining parameters is,

$$\omega_{jt}^*(\alpha_k, \alpha_{kk}) = \frac{\hat{\eta}}{1 + \hat{\eta}} [\ln \hat{R}_{jt} - \frac{1}{\hat{\eta}} (\hat{\phi}_{jt} + \alpha_k \ln K_{jt} + \frac{1}{2} \alpha_{kk} (\ln K_{jt})^2)], \quad (18)$$

where $\ln \hat{R}_{jt}$ represents the fitted value from the first stage estimation of (17).

In the second stage, we estimate the remaining parameters α_k and α_{kk} using the additional Markov productivity assumption. By replacing ω_{jt} and ω_{jt-1} in the Markov productivity process using equation (18), we derive the second stage estimation equation,

$$\begin{aligned} \hat{\eta} \ln \hat{R}_{jt} - \hat{\phi}_{jt} &= \alpha_k \ln K_{jt} + \frac{1}{2} \alpha_{kk} (\ln K_{jt})^2 \\ &\quad + (1 + \hat{\eta})g \left(\frac{1}{1 + \hat{\eta}} (\hat{\eta} \ln \hat{R}_{jt-1} - \hat{\phi}_{jt-1} - \alpha_k \ln K_{jt-1} + \frac{1}{2} \alpha_{kk} (\ln K_{jt-1})^2) \right) + \epsilon_{jt}. \end{aligned} \quad (19)$$

Recall that ϵ_{jt} , the productivity innovation, is independent of capital by the Markov assumption.

²⁴Note that (5) does not involve α_k or α_{kk} .

So, we can now estimate α_k and α_{kk} using the moment conditions,

$$E \left(\epsilon_{jt} \begin{pmatrix} \ln K_{jt} \\ (\ln K_{jt})^2 \\ \ln K_{jt-1} \\ (\ln K_{jt-1})^2 \\ \hat{\eta} \ln \hat{R}_{jt-1} - \hat{\phi}_{jt-1} \end{pmatrix} \right) = 0. \quad (20)$$

Note that this procedure actually over-identifies the parameters of interest, allowing for a test of over-identifying restrictions.

While we use the translog as a motivating example, this procedure can be directly applied to more general (parametric) production functions as long as the productivity shock is additively separable from F , and can be assumed to follow a Markov process.

4 Multiple Materials Inputs

So far, we have followed the literature in assuming that firms purchase a single homogeneous intermediate input. Indeed, the ability to treat the imputed firm-specific price and quantity choices as quality-adjusted scalars representing a single homogenous input is critical since our demand specification assumes that outputs are horizontally differentiated.²⁵ In reality, intermediate input expenditures are an aggregate of a wide variety of different input goods. Ideally, an analyst would be able to account for each of these goods separately in the production function. Unfortunately, datasets typically contain only total input expenditure, not information on the various types used, much less prices and quantities for each. With such limited data, it is clearly not possible to learn the impact of individual inputs. However, if the effect of inputs on production can be summarized through a homogeneous materials index function, we show that the remaining production function parameters can consistently be recovered using only total expenditure information.

To be specific, suppose the firm may use up to D different types of materials. Denote the vector of materials quantities used in production as $M_{jt} = (M_{1jt}, M_{2jt}, \dots, M_{Djt})$. These input types

²⁵We thank an anonymous referee for making this point.

may be entirely different input goods (thread versus fabric) or the same input good of different quality (cotton versus polyester fabric). However, only the total expenditure on all components $E_{M_{jt}} = \sum_{d=1}^D P_{M_{djt}} M_{djt}$, rather than each specific component M_{djt} , is known to the econometrician. Assume inputs enter into the production function as,

$$Q_{jt} = e^{\omega_{jt}} F(L_{jt}, \mu(M_{jt}), K_{jt}; \theta), \quad (21)$$

where $\mu : \mathbb{R}^D \rightarrow \mathbb{R}$ is a homogeneous index function which summarizes the contribution of all materials inputs to production.²⁶ As part of the production function, we assume that μ is known to the firm. While this structure allows materials to substitute for each other in an unknown manner, it does restrict the substitution patterns between materials and other production inputs, namely labor and capital. As a result, we can allow for vertically or horizontally differentiated materials and treat them as different elements of the materials vector in M_{jt} . The corresponding idiosyncratic material prices for each component is summarized in price vector $P_{M_{jt}} = (P_{M_{1jt}}, P_{M_{2jt}}, \dots, P_{M_{Djt}})$, which is observed by firms but not by researchers.

The firm's static optimization problem is now to choose L_{jt} and the vector M_{jt} to maximize the profit given productivity and capital stock:

$$\begin{aligned} \max_{L_{jt}, M_{jt}} \quad & P_t(Q_{jt})Q_{jt} - P_{L_{jt}}L_{jt} - P_{M_{jt}}M_{jt} \\ \text{s.t.} \quad & Q_{jt} = \exp(\omega_{jt})F(L_{jt}, \mu(M_{jt}), K_{jt}; \theta). \end{aligned}$$

The first order conditions for L_{jt} and all components of vector M_{jt} are:

$$\exp(\omega_{jt})F_{L_{jt}} \left[P_t(Q_{jt}) + Q_{jt} \frac{\partial P_t(Q_{jt})}{\partial Q_{jt}} \right] = P_{L_{jt}}, \quad (22)$$

$$\exp(\omega_{jt})F_{\mu_{jt}} \left[P_t(Q_{jt}) + Q_{jt} \frac{\partial P_t(Q_{jt})}{\partial Q_{jt}} \right] \mu_d(M_{jt}) = P_{M_{djt}}, \quad \forall d = 1, 2, \dots, D \quad (23)$$

where $\mu_d(M_{jt}) = \frac{\partial \mu(M_{jt})}{\partial M_{djt}}$.

²⁶In order for Assumption 1 to continue to hold, we assume that μ is differentiable in M_{jt} . However we are able to allow for some non-differentiability in μ , as we will see in the Monte Carlo experiment in Section 5.6.

Denote the optimal choice of the firm as L_{jt}^* and vector M_{jt}^* . Thus the total expenditure on materials, which is observed by the researcher, is $E_{M_{jt}}^* = \sum_{d=1}^D P_{M_{djt}} M_{djt}^*$. Define the material price index as $P_{\mu_{jt}} = \frac{E_{M_{jt}}^*}{\psi(M_{jt}^*)}$, where $\psi(M_{jt}^*) = \sum_{d=1}^D M_{djt}^* \mu_d(M_{jt}^*)$. Using this price index, the information in (23) can be summarized into a single equation by multiplying (23) by M_{djt}^* , summing them up across d , and dividing it by $\psi(M_{jt}^*)$,

$$\exp(\omega_{jt}) F_{\mu_{jt}} \left[P_t(Q_{jt}) + Q_{jt} \frac{\partial P_t(Q_{jt})}{\partial Q_{jt}} \right] = P_{\mu_{jt}}. \quad (24)$$

This equation together with (22) can be viewed as the first order conditions of the firm's optimization problem if it faced labor price $P_{L_{jt}}$ and a materials price $P_{\mu_{jt}}$ for *single* material μ . The following proposition shows how our method can be adapted to an unknown vector of inputs.

Proposition 2 *If the index function $\mu_{jt} = \mu(M_{jt}) : \mathbb{R}^D \rightarrow \mathbb{R}$ is homogeneous of degree $k > 0$, then,*

$$z(\mu_{jt}, L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta, k) \equiv \frac{F_{L_{jt}} L_{jt}}{F_{\mu_{jt}} \mu_{jt}} - \frac{E_{L_{jt}}}{\frac{1}{k} E_{M_{jt}}} = 0, \quad (25)$$

and it is possible to recover $\mu^(L_{jt}, K_{jt}, E_{M_{jt}}, E_{L_{jt}}; \theta, k)$ when the conditions of Proposition 1 hold for $z(\cdot)$ as defined in (25).*

The proof of this proposition is provided in Appendix A and directly follows the Euler's Theorem for homogeneous functions and Proposition 1. We can now substitute μ^* into the revenue function as with M^* in Section 3, albeit with k as an additional scale parameter. In some specifications (e.g., translog), k may not be separately identified from the production function parameters. In this case k can be scale normalized to 1 without loss of generality, since it is absorbed in the primary parameters of the production function. In other cases (e.g., CES), k can be identified through the revenue function, where it represents returns to scale of the materials aggregator index μ .²⁷ In this case, estimation still follows our method from Section 3, except that now we substitute the material expenditure with $\frac{1}{k} E_{M_{jt}}$ and estimate k as an additional parameter of the production function. For

²⁷Whether or not k is identified depends on whether or not the production function already accounts for scale effects on materials independent of other inputs. That is, suppose the production function is $F(L, M, K; \theta)$ and consider the alternative $\tilde{F}(L, M, K; k, \theta) = F(L, M^k, K; \theta)$. In $\tilde{F}(\cdot)$, k may or may not be identified depending on the form of F . If k is not identified, normalizing $k = 1$ simply returns the researcher to the original specification. For theoretical reasons, some researchers may still want to impose that $k = 1$ even when it is formally identified. This would be essentially equivalent to assuming constant returns to scale in the materials aggregator μ .

example, in the CES specification of Section 3.4, we can employ the following revenue equation:²⁸

$$\ln R_{jt} = \ln \frac{\eta}{1 + \eta} + \ln \left[\frac{1}{k} E_{M_{jt}} + E_{L_{jt}} \left(1 + \tau \left(\frac{K_{jt}/\bar{K}}{L_{jt}/\bar{L}} \right)^\gamma \right) \right] + u_{jt}, \quad (26)$$

where $\tau = \alpha_K/\alpha_L$.

In sum, our method still works if the effect of material inputs on production can be summarized by a homogeneous materials index function.²⁹ As we would expect, the functional form of $\mu(\cdot)$ is not identified without more information, but its functional form (indeed, even its dimension) is not needed to recover the other production parameters, θ .

Although we assume that $\mu(\cdot)$ is homogeneous to make our method work, this still allows a vast set of flexible functional forms that may incorporate both vertically and horizontally differentiated materials inputs. We verify the validity of this approach through a Monte Carlo experiment in Section 5.6.

5 Monte Carlo Experiment

This section presents Monte Carlo experiments that evaluate the performance of our method, and show how it corrects for input price heterogeneity. For brevity, we concentrate on the CES version of the estimator. We first describe the data generation process, then estimate the model in three different ways based on assumed data availability. At the end of this section, we examine the performance of our method when the firm chooses from a vector of inputs.

5.1 Data Generation

Using the CES specification of the production function (9) and a Dixit-Siglitz demand system (11), we generate N replications of simulated data sets, given a set of true parameters of interest (η , σ ,

²⁸We conducted a Monte Carlo experiment to verify this result with $\mu(M_{jt}) = M_{jt}^k$, and our method works very well in recovering the primary parameter θ as well as k . Results are available upon request.

²⁹If $\mu(\cdot)$ is not homogenous, then the “total material expenditure” implied by Equation (24) together with Equation (22) will be $\mu(M_{jt}^*)P_{\mu_{jt}} = \frac{\mu(M_{jt}^*)}{\psi(M_{jt}^*)}E_{M_{jt}^*}$ which is not observable (since generally $\frac{\mu(M_{jt}^*)}{\psi(M_{jt}^*)}$ is not a constant). In this case, information on expenditure alone is insufficient to control for variation in materials inputs even if prices are homogenous across firms.

α_L , α_M and α_K). In each replication, there are J firms in production for T periods. We simulate a sequence of productivity for each firm (ω_{jt}) from an AR(1) process. We allow the firm's capital stock (K_{jt}) to evolve based on an investment rule (investment is denoted as I_{jt}) that depends on its productivity and capital stock,

$$\log(I_{jt}) = \xi\omega_{jt} + (1 - \xi)\log(K_{jt}).$$

Finally, we allow input prices ($P_{L_{jt}}$ and $P_{M_{jt}}$) to vary across firms and over time. Table 1 lists the underlying parameters used to generate the data set. A full description of the data generating process is provided in Appendix D.

Given these variables and industrial-level outputs and prices (Q_t and P_t), we derive a sequence of optimal choices of labor and material inputs (L_{jt} and M_{jt} with corresponding input expenditures $E_{L_{jt}}$, $E_{M_{jt}}$), the optimal output quantity (Q_{jt}), price (P_{jt}) and revenue (R_{jt}) for firm j in each period t . In this way, we generate a data set of $\{\omega_{jt}, K_{jt}, I_{jt}, L_{jt}, M_{jt}, E_{L_{jt}}, E_{M_{jt}}, Q_{jt}, R_{jt}, Q_t, P_t\}$ for each firm j and period t . All these variables are observable to firms, however, usually only a subset of them are available to researchers.

5.2 Our Method

We first estimate the model with our method. In this case, we assume a researcher observes $\{K_{jt}, L_{jt}, E_{L_{jt}}, E_{M_{jt}}, R_{jt}, P_t, Q_t\}$ for each firm and each period. The researcher is not required to observe firm's investment, material input quantity, physical outputs quantity or, of course, productivity. As described in the previous section, we exploit the first order conditions to impute firm-level material quantities from labor quantities and expenditures. This approach allows us to estimate the production function while controlling for unobserved price dispersion. We will evaluate our method by comparing our estimates with the true parameters, as well as with those derived from two alternative estimation methods that require additional data.

5.3 Traditional Method with Direct Proxy

For our first point of comparison, we estimate the model using a direct proxy method that substitutes $E_{M_{jt}}$ for M_{jt} . The method follows Olley and Pakes (1996) in using a control function approach to utilize investment data to control for unobserved productivity. Traditionally, researchers have used deflated expenditure on materials inputs to proxy for intermediate input quantities when applying this and similar methods (e.g., Levinsohn and Petrin, 2003), and we follow that practice here. We will refer to this method as the ‘‘OP’’ procedure, although we should emphasize that it is the direct proxy, rather than the OP procedure, that is introducing the bias. In contrast with our own method, the OP procedure takes output quantities as observable. Hence there will be no output price bias and any resultant bias is caused by the substitution of physical material input by its deflated cost.³⁰

Specifically, researchers using this method observe $\{K_{jt}, L_{jt}, E_{L_{jt}}, E_{M_{jt}}, Q_{jt}, I_{jt}, P_t, Q_t\}$ and estimate parameters via the (logarithm) production function:

$$\ln\left(\frac{Q_{jt}}{Q}\right) = \left\{ \omega_{jt} + \frac{1}{\gamma} \ln \left[\alpha_L \left(\frac{L_{jt}}{L}\right)^\gamma + \alpha_M \left(\frac{E_{M_{jt}}}{E_M}\right)^\gamma + \alpha_K \left(\frac{K_{jt}}{K}\right)^\gamma \right] \right\} + u_{jt}, \quad (27)$$

where the error term u_{jt} accounts for measurement error of revenue and productivity shocks that are unanticipated by the firm.

To control for endogeneity bias, Olley and Pakes (1996) assume that productivity follows a first order Markov process. Following our data generating process, we are more specific and assume that productivity follows an AR(1) specification,

$$\omega_{jt+1} = g_0 + g_1\omega_{jt} + \epsilon_{jt+1}.$$

Since the true data generating process is in fact AR(1), this rules out specification error associated with the productivity evolution process, so that the Monte Carlo focuses on the bias caused by dispersion in input prices. Within our data generating process, the investment decision is a function

³⁰We could easily incorporate a revenue function into this procedure. We do not in order to emphasize that the direct proxy is the cause of the resulting bias.

of current capital stock and the unobservable heterogenous productivity and therefore, the OP method can approximate the productivity by a control function of investment and capital stock: $\omega_{jt} = \omega_t(I_{jt}, K_{jt})$. Substituting it into equation (27) yields,³¹

$$\ln\left(\frac{Q_{jt}}{Q}\right) = \phi(L_{jt}, E_{M_{jt}}, K_{jt}, I_{jt}, \Phi_t) + u_{jt}, \quad (28)$$

where Φ_t represents time dummies to capture aggregate investment shifters. This equation can be estimated non-parametrically. This estimation is consistent since the right-hand-side variables are all uncorrelated with u_{jt} . We estimate ϕ using the method of sieves.³² Denote $\hat{\phi}_{jt}$ as the fitted value of $\phi(L_{jt}, E_{M_{jt}}, K_{jt}, I_{jt}, \Phi_t)$. Then productivity can be expressed as,

$$\omega_{jt} = \hat{\phi}_{jt} - \frac{1}{\gamma} \ln \left[\alpha_L \left(\frac{L_{jt}}{L} \right)^\gamma + \alpha_M \left(\frac{E_{M_{jt}}}{E_M} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{K} \right)^\gamma \right]. \quad (29)$$

Substituting ω_{jt+1} and ω_{jt} into the evolution process of productivity, we obtain,

$$\epsilon_{jt+1} = \omega_{jt+1} - (g_0 + g_1\omega_{jt}).$$

Note that ϵ_{jt+1} is uncorrelated with variables up to period t , so we can construct the set of moment conditions with which we estimate the model's parameters,

$$E(\epsilon_{jt+1}x_{jt+1}) = 0, \quad (30)$$

where $\epsilon_{jt+1}(\eta, \tau, \gamma, g_0, g_1) = \omega_{jt+1} - (g_0 + g_1\omega_{jt})$, and x_{jt+1} is a combination of variables that are uncorrelated with the innovation term in period $t + 1$, e.g., $L_{jt}, E_{L_{jt}}, E_{M_{jt}}, K_{jt}, K_{jt+1}$.³³

³¹In contrast to the original OP paper, we follow Akerberg, Caves, and Frazer (2006) in recovering the labor and materials parameter out of the second stage of the OP estimation to avoid collinearity issues in the first stage.

³²In practice, we model $\phi(\cdot)$ with a cubic function with interactions.

³³In the Monte Carlo experiment, we choose

$$x_{jt+1} = \left(\frac{L_{jt}}{L}, \frac{E_{L_{jt}}}{E_L}, \frac{E_{M_{jt}}}{E_M}, \frac{K_{jt}}{K}, \frac{K_{jt+1}}{K}, \left(\frac{L_{jt}}{L} \right)^2, \left(\frac{E_{L_{jt}}}{E_L} \right)^2, \left(\frac{E_{M_{jt}}}{E_M} \right)^2, \left(\frac{K_{jt}}{K} \right)^2, \left(\frac{K_{jt+1}}{K} \right)^2 \right)$$

to serve as instruments.

5.4 Oracle-OP Procedure

Finally, we compare our method to a first-best case when input quantities are actually observed. We refer to this as the “Oracle-OP” case as it uses the Olley and Pakes (1996) inversion to recover productivity but uses the actual materials input quantities instead of a proxy. That is, we observe $\{K_{jt}, L_{jt}, M_{jt}, E_{L_{jt}}, E_{M_{jt}}, Q_{jt}, I_{jt}, P_t, Q_t\}$ for each firm and each period. This enables us to estimate the production function in (27) without using expenditure as a proxy. The only difference between the oracle case and previous OP’s procedure is that material quantity is not substituted by its proxy, since the true quantity is observable in this case. In comparison to our own method, this method requires that the researcher observes investment, output quantity, and materials input quantities.

5.5 Results

The results of the Monte Carlo experiments for three separate elasticities of substitution are presented in Table 2. For each method, the listed parameter represents the median estimate of the 1000 Monte Carlo replications with standard errors in parenthesis. The square brackets contain the root mean squared error of the estimates. Across all parameterizations, our method recovers the parameters well. In contrast, the elasticity of substitution (σ) and α_K are severely underestimated by OP. This corresponds to the biases due to input price dispersion as documented in Section 2. The results for the oracle-OP method confirm that when input price heterogeneity is observed, the bias is eliminated. Interestingly, it appears there is little loss in efficiency between the oracle-OP method and the method we propose, despite the fact that we do not use investment, output quantity, or input quantities. Of course, our method makes use of the additional structure implied by the firm’s first order conditions, which is not used within the OP framework.

To further investigate the performance of the estimators, Figure 1 plots the density of $\hat{\sigma}$ for the three cases. The dashed line represents the true value of σ . Clearly, our method generates estimates that are concentrated around the true elasticity of substitution. However, the proxy-OP method produces biased estimates of σ . This bias is economically significant, implying an elasticity of substitution up to 20% lower than the true value. As expected, when we allow the researcher to observe input quantities directly, the oracle-OP method performs well. The validation of the expenditure

proxy requires no heterogeneity in input prices across firms. When heterogeneity is present, the unobserved input price dispersion will bias the estimation. The advantage of our method is that it provides a consistent way of imputing unobservable input quantities from observable expenditures of both materials and labor, which is necessary to control for the input price heterogeneity.

In addition to controlling for input price dispersion, our method allows the researcher to recover estimates of the unobserved input prices. In short, material quantities and prices can be imputed from Equation (54) (in Appendix C). Figure 2 presents the kernel density estimation of the imputed material prices from our method and compares it to the true density of material prices in the Monte Carlo.³⁴ It shows that the imputed material price density matches the true density quite well.

5.6 Multiple Materials Inputs

The earlier Monte Carlo experiments assumed a scalar materials input. Here, we allow firms to endogenously select a vector of inputs, as in Section 4, to provide an illustration of how our method works under these conditions.

For consistency, we again use the basic CES formulation, and only emphasize the introduction of multiple materials here. Specifically, in addition to labor and capital stock, firms may choose to use any combination of three components (M_1 , M_2 , M_3) of materials in production. They enter into the production function through the index function μ ,³⁵

$$\mu(M_{jt}) = \max \left(\left[(\beta M_{1jt})^{\gamma_1} + M_{3jt}^{\gamma_1} \right]^{1/\gamma_1}, \left[M_{2jt}^{\gamma_2} + M_{3jt}^{\gamma_2} \right]^{1/\gamma_2} \right), \quad (31)$$

where $\gamma_1 = \frac{\delta_1 - 1}{\delta_1}$ and $\gamma_2 = \frac{\delta_2 - 1}{\delta_2}$. The functional form of μ is observable to the firms but not to the econometrician. The experiment's basic structure is inspired by the following scenario:³⁶ M_1 and M_2 are vertically differentiated versions of the same type of input (e.g., two versions of the same part). They differ in their quality such that the efficiency for M_1 in the production process is $\beta < 1$, while the efficiency for M_2 is normalized to be one. They also differ in regard to their substitutability with the third component (δ_1 and δ_2 may not be equal). They are produced within

³⁴We present the case for true $\sigma = 2.5$, but the results from other cases are very similar.

³⁵Note that this function is homogenous of degree 1, as seems reasonable given our scenario.

³⁶We thank a referee for suggesting a version of this Monte Carlo design.

an imperfectly competitive industry and their prices are P_{M_1} and P_{M_2} offered to all firms.³⁷ The functional form of $\mu(M_{jt})$ implies that the firm will optimally use either M_1 or M_2 since using both can provide no benefit over only purchasing one. The third component is a homogenous material M_3 , with idiosyncratic price $P_{M_{3jt}}$. The production function is,

$$Q_{jt} = e^{\omega_{jt}} \bar{Q} \left[\alpha_L \left(\frac{L_{jt}}{L} \right)^\gamma + \alpha_M \left(\frac{\mu(M_{jt})}{M} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{K} \right)^\gamma \right]^{\frac{1}{\gamma}}. \quad (32)$$

The firm observes the price vector, $(P_{M_1}, P_{M_2}, P_{M_{3jt}})$, and optimally chooses its vector of inputs, $(M_{1jt}, M_{2jt}, M_{3jt})$. However, only total materials expenditure $E_M = \sum_{d=1}^3 P_{M_{djt}} M_{djt}$ is observed by the researcher, who is attempting to recover the production function parameters $(\alpha_L, \alpha_M, \alpha_K, \gamma)$, as well as the distribution of $\mu(M_{jt})$, its price index $P_{\mu jt}$, and productivity ω_{jt} .

Importantly, the price of M_3 will affect the optimal decision to employ M_1 or M_2 in production (see Figure 3). Each firm has a cutoff point $\tilde{P}_{M_{3jt}}$, and the choice of quality levels depends on whether it faces a price for M_3 above or below this cutoff. To generate our data, we solve the optimization problem for each firm j and period t , and obtain the input demand $(M_{1jt}^*, M_{2jt}^*, M_{3jt}^*, L_{jt}^*)$ ³⁸, which is then substituted into the demand function and the production function to calculate the other endogenous variables. Therefore, we have generated the entire data set of firm-level variables for each firm j and period t : $\{\omega_{jt}, K_{jt}, I_{jt}, L_{jt}, M_{jt}, E_{L_{jt}}, E_{M_{jt}}, Q_{jt}, R_{jt}, Q_t, P_t\}$, where $M_{jt} = (M_{1jt}^*, M_{2jt}^*, M_{3jt}^*)$. Then we estimate the model with our method only using data set on $\{K_{jt}, L_{jt}, E_{L_{jt}}, E_{M_{jt}}, R_{jt}, Q_t, P_t\}$.

Table 3 presents the result for N replications. As discussed earlier, the form of $\mu(M)$ is not identified, but we find that our method recovers the primary parameters (i.e., all α 's, σ and η) very well. Also, the material quantity index and price index can be recovered. In Figure 4, we compare the recovered material quantity index $\hat{\mu}(M)$ and price index \hat{P}_μ with the true indexes $\mu(M)$ and P_μ . Both show that our method recovers the underlining material quantity index and price index accurately in this example.

³⁷We have also experimented with allowing these prices to be heterogeneous and have also been able to successfully recover the production function parameters.

³⁸Either M_{1jt}^* or M_{2jt}^* should be zero because of the substitution between the two.

6 Application: Colombian Data

To evaluate the performance of our estimator using real data, we apply our method to a dataset of Colombian manufacturing firms from 1981 to 1989, which was collected by the Departamento Administrativo Nacional de Estadística (DANE).³⁹ This application serves two purposes. The first is to compare our results with those found using the traditional proxy method to account for unobserved materials inputs. The second is to illustrate additional information which can be recovered using our method, including the distribution of input prices and their relationship to productivity.

This dataset contains detailed information about firm-level revenue (R), labor and material input expenditure (E_L and E_M), capital stock (K), employment (L), the wage bill (E_L), and investment (I). However, firm-level price information about input and output is not available. Moreover, only total expenditure on “raw materials, materials and packaging” is available (E_M) rather than total quantities (M). This includes expenditure on raw materials such as cloth and gasoline, but does not include consumption of electrical energy, “general expenses” such as professional services and advertising, or “industrial expenses” such as spare or replacement parts, all of which are reported separately. It is extremely common in the literature to treat materials as a homogenous input (e.g. Levinsohn and Petrin, 2003) and our approach can be interpreted as following this tradition. However, as shown in Section 4 it can also be employed if firms are optimally choosing from a vector of heterogenous inputs. In this case, the imputed input price represents the shadow price of increasing the use of inputs in production. This is important since material expenditure represents the sum of several different input types that may vary across firms even within an industry.⁴⁰

First, we estimate the model using our method using the CES specification of the production function normalized at the geometric mean as illustrated in Section 3.4. As a primary basis of comparison, we also estimate the production function using materials expenditure as a proxy for materials inputs as in Olley and Pakes (1996). To focus on the impact of input price heterogeneity,

³⁹For a detailed introduction to the data set, see Roberts and Tybout (1997).

⁴⁰For simplicity, we normalize k (the degree of homogeneity of μ) to be 1 in the application.

we control for output price bias by incorporating a demand function in this approach, as suggested in Klette and Griliches (1996). Thus, the only difference between the two estimators is in their treatment of input quantities. We refer to the second method as OP-KG in the text and tables. Of course, there are many other approaches that may be used to estimate production functions. In Appendix F, we compare our method to several alternative approaches, including employing first order conditions to recover productivity while using the proxy approach for materials and following well-known panel data methods (Arellano and Bond, 1991).

Estimates for four large industries are displayed in Table 4: clothing, bakery products, printing and publishing, and metal furniture.⁴¹ In all these industries, the estimate of the elasticity of substitution is significantly lower using the OP-KG method compared with the results from our method. This is consistent with both our intuition about the bias generated by unobserved price dispersion and the pattern shown in the Monte Carlo experiments. Moreover, the elasticity of substitution estimates are significantly greater than one in all industries when using our method. This implies that production function is not likely Cobb-Douglas in these industries. The results support the conclusion that ignoring input price dispersion would lead to inconsistent estimates of elasticities of substitution, and that our method is capable of controlling for unobserved price dispersion.

Biased estimates of elasticity of substitution (σ) using the OP-KG method will contaminate estimates of the distribution parameters. However, the direction of the bias is unclear. We find that our method produces estimates of α_K that are at least 30 percent larger, and sometimes more than twice as large, as the estimates of α_K using the OP-KG method. These results mirror the findings from the Monte Carlo study, where α_K is also underestimated by the traditional method. It appears that ignoring price dispersion is likely to lead researchers to underestimate the degree of capital intensity in production.

A key output from production function estimation is the implied productivity distribution of firms within an industry. We find that there are substantial differences in the estimates of this distribution between the two methods. Figure 5 shows the productivity distributions estimated

⁴¹We have estimated the model for a wide variety of industries and found these results to be representative with respect to the performance of the estimators. Additional results are available by request.

using our method and the OP-KG method for each of the four industries.⁴² For all industries, the productivity distribution in OP-KG is more concentrated than using our method to control for price dispersion. The result is most stark for the bakery products industry, where our implied distribution has an inter-quartile range that is 3.4 times as wide as that using the OP-KG method. But even in the clothing industry, where the two productivity distributions are most similar, our distribution has an inter-quartile range more than 60 percent larger than is found using OP-KG. This suggests that omitting the unobserved input price dispersion tends to underestimate the firm heterogeneity in productivity. One possible reason might be a positive correlation between input prices and productivity, which we report below. Intuitively, positive correlation between the productivity and input prices could bias productivity estimates since a firm with low productivity tends to use low-price materials. In the OP-KG method, where all firms are assumed to have the same material price, the total material quantity used by low-productivity firms is underestimated, resulting in overestimates of their productivity. Similarly, OP-KG would underestimate the productivity for high-productivity firms facing high prices. As a result, OP-KG, by not controlling for the unobserved input prices, would underestimate the degree of productivity dispersion within the industry. A large literature, recently reviewed by Syverson (2011), is devoted to understanding and explaining heterogeneity of productivity among firms.⁴³ Our finding indicates that the “true” productivity heterogeneity may be even larger than is indicated by estimations that fail to control for unobserved input price dispersion.

In addition to results on the production function and the distribution of productivity, our method also provides estimates of the unobserved input prices and quantities across firms. Because these prices and quantities are recovered from the first order condition, they reflect quality-adjusted quantity indices and the imputed prices are purged of the effect of quality differences. In Figure 6, we present the kernel density estimations of imputed material prices (in logarithm) from our method for each of the four industries pooled across all years. In all industries, the distributions

⁴²Figure 5 follows Olley and Pakes (1996) in defining productivity as the sum of ω_{it} , which is known to the firm when it chooses labor and materials, and u_{it} which is unanticipated productivity and measurement error. In Appendix E, we compare the distributions of $\omega_{it} + u_{it}$ with only anticipated productivity, ω_{it} . We find that for both methods the distributions are fairly similar, implying that the bulk of productivity dispersion is anticipated by firms.

⁴³An earlier review of this literature is provided by Bartelsman and Doms (2000).

of input prices are quite spread out, indicating that price dispersion is substantial. Our findings are partially corroborated by studies, such as Ornaghi (2006) and Atalay (2012), which observe input prices directly and also find significant dispersion. Since our input prices are quality adjusted they suggest that even after considering quality differences across inputs, which could account for significant dispersion when inputs are directly observed, there is still large dispersion in the input prices.

We are also able to use our method to analyze the dynamics of input price dispersion. While it is not assumed in our estimation, we would expect a significant amount of persistence in firms' input prices over time. To check this, we fit the input prices to a simple AR(1) process to analyze their persistence. Table 5 shows the estimated persistence with standard error. In all four industries, there is quite high persistence with mean around 0.75, which is close to the persistence reported in Atalay (2012) in which firm-level input prices and quantities are available. Thus, firms that are able to secure low prices today are likely to be able to secure them again in the future. This gives us some confidence that our imputed prices do not simply reflect estimation error, but are a persistent feature of firms.

Finally, we examine the joint relationship between input prices and productivity in our sample of firms. As shown in Table 5, the imputed input price is positively correlated with the recovered productivity. That is, higher productivity firms tend to pay higher input prices. As mentioned above, this correlation is one reason why our method indicates a higher degree of productivity dispersion than we see in traditional methods that assume input prices are homogeneous. Table 5 also reports the correlation between input prices and observed wages, and again finds a positive correlation—high productivity firms pay more for both labor and materials. These results are consistent with Kugler and Verhoogen (2012), who directly examine data on input prices and compares them with productivity estimates. In explaining their result, Kugler and Verhoogen (2012) emphasize the quality complementarity hypothesis—input quality and plant productivity are complementary in generating output. However, because we recover the input prices using the marginal contribution of inputs in production, our recovered input price is quality-adjusted, ruling out the quality-complementarity explanation. Even so, we find a positive correlation between

input prices and productivity. This indicates that alternative factors, such as plant-specific demand shocks or market power in input sectors, as discussed in Kugler and Verhoogen (2012), may also contribute to the dispersion of input prices within industries.

7 Conclusion

We analyze the problem of unobserved input prices and quantities in the estimation of production functions. Simply using expenditures as a proxy for quantities is likely to bias production function estimates in the presence of input price heterogeneity. To account for unobserved price dispersion, we introduce a method which exploits the first order conditions of profit maximization and imputes unobservable firm-level quantities of outputs and inputs from observable data on revenue and expenditures. We show how to apply our method using the commonly used CES production function specification.

To validate our method, we conduct Monte Carlo experiments to evaluate the performance of our estimation method. The results confirm that ignoring unobserved price dispersion biases the estimation when deflated values are used as proxies of quantities. In contrast, our method recovers the true parameters very well.

We further show that these differences matter in real data by applying the methods to a dataset on the Colombian manufacturing sector. The results are in line with theory and the Monte Carlo study. Specifically, the elasticity of substitution is significantly lower compared with our method when using the expenditure proxy. In addition, our results confirm the presence of unobserved price dispersion, and indicate that input prices and firm productivity are positively correlated. As a result, we find significantly more productivity dispersion in the industries we study than would be uncovered using a traditional estimator. The results suggest that our method of imputing unobservable firm-level input quantities from observable expenditures is important to effectively control for input price dispersion and consistently estimate the production function and the degree of dispersion in productivity.

References

- ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2006): “Structural Identification of Production Function,” UCLA, Deloitte and Touche, Rotman School of Management.
- ARELLANO, M., AND S. BOND (1991): “Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations,” *The Review of Economic Studies*, 58(2), 277–297.
- ATALAY, E. (2012): “Materials Prices and Productivity,” University of Chicago.
- BARTELSMAN, E., AND M. DOMS (2000): “Understanding Productivity: Lessons from Longitudinal Microdata,” *Journal of Economic Literature*, 38(3), 569–594.
- BEAULIEU, J., AND J. MATTEY (1999): “The effects of General Inflation and Idiosyncratic Cost Shocks on Within-Commodity Price Dispersion: Evidence from Microdata,” *Review of Economics and Statistics*, 81(2), 205–216.
- BILS, M., AND P. KLENOW (2004): “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy*, 112, 947–985.
- BLUNDELL, R., AND S. BOND (2000): “GMM estimation with persistent panel data: an application to production functions,” *Econometric Reviews*, 19, 321–340.
- DE LA GRANDVILLE, O. (1989): “In Quest of the Slutsky Diamond,” *American Economic Review*, 79(3), 468–481.
- DE LA GRANDVILLE, O., AND R. M. SOLOW (2006): “A Conjecture on General Means,” *Journal of Inequalities in Pure and Applied Mathematics*, 7(1).
- DE LOECKER, J. (2011): “Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity,” *Econometrica*, Vol. 79, No. 5, pp. 1407–1451.
- DE LOECKER, J., P. K. GOLDBERG, A. K. KHANDELWAL, AND N. PAVCNİK (2012): “Prices, Markups and Trade Reform,” *NBER Working Paper 17925*.
- DE LOECKER, J., AND F. WARZYŃSKI (2012): “Markups and Firm-Level Export Status,” *American Economic Review*, 102(6), 2437–71.
- DORASZELSKI, U., AND J. JAUMANDREU (forthcoming): “R&D and Productivity: Estimating Endogenous Productivity,” *Review of Economic Studies*.
- DUNNE, T., AND M. J. ROBERTS (1992): “Costs, Demand, and Imperfect Competition as Determinants of Plant-Level Output Prices,” in *Empirical Studies in Industrial Organization: Essays in Honor of Leonard W. Weiss*, ed. by D. Audretsch, and J. Sigfried. Kluwer Academic Publishers.
- EPPLE, D., B. GORDON, AND H. SIEG (2010): “A New Approach to Estimating the Production Function for Housing,” *American Economic Review*, 100(3), 905–924.
- ESLAVA, M., J. HALTIWANGER, A. KUGLER, AND M. KUGLER (2004): “The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia,” *Journal of Development Economics*, 75, 333–371.

- 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.
- FOX, J., AND V. SMEETS (2011): “Does Input Quality Drive Measured Productivity Differences in Firm Productivity,” *International Economic Review*, 52(4), 961–989.
- GANDHI, A., S. NAVARRO, AND D. RIVERS (2013): “On the Identification of Production Functions: How Heterogeneous is Productivity?,” University of Wisconsin-Madison and University of Western Ontario.
- GRIECO, P., AND R. MCDEVITT (2012): “Productivity and Quality in Health Care: Evidence from the Dialysis Industry,” Penn State Univ.
- KATAYAMA, H., S. LU, AND J. TYBOUT (2009): “Firm-Level Productivity Studies: Illusions and a Solution,” *International Journal of Industrial Organization*, 27(3), 403–413.
- KLEIN, L. R. (1953): “A Textbook of Econometrics,” *Evanston, Illinois, and White Plains, New York: Row, Peterson and Company*, pp. pp. ix, 355.
- 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.
- KLUMP, R., AND O. DE LA GRANDVILLE (2000): “Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions,” *American Economic Review*, 90(1), 282–291.
- KLUMP, R., AND H. PREISLER (2000): “CES Production Functions and Economic Growth,” *Scandinavian Journal of Economics*, 102(1), 41–56.
- KUGLER, M., AND E. VERHOOGEN (2012): “Prices, Plant Size, and Product Quality,” *Review of Economic Studies*, 79(1), 307–339.
- LEON-LEDESMA, M. A., P. MCADAM, AND A. WILLMAN (2010): “Identifying the Elasticity of Substitution with Biased Technical Change,” *American Economic Review*, 100(4), 1330–1357.
- LEVINSOHN, J., AND A. PETRIN (2003): “Estimating Production Functions Using Inputs to Control for Unobservables,” *The Review of Economic Studies*, 70(2), pp. 317–341.
- MARSCHAK, J., AND W. H. ANDREWS (1944): “Random Simultaneous Equations and the Theory of Production,” *Econometrica*, 12(3/4), 143–205.
- OLLEY, G. S., AND A. PAKES (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64(6), pp. 1263–1297.
- ORNAGHI, C. (2006): “Assessing the Effects of Measurement Errors on the Estimation of Production Functions,” *Journal of Applied Econometrics*, 21(6), 879–891.
- ROBERTS, M., AND J. TYBOUT (1997): “The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs,” *American Economic Review*, 87, 545–564.
- ROBERTS, M. J., AND D. SUPINA (1996): “Output price, markups, and producer size,” *European Economic Review*, 40, 909–921.

- (2000): “Output Price and Markup Dispersion in Micro Data: The Roles of Producer Heterogeneity and Noise,” in *Advances in Applied Microeconomics*, ed. by M. Baye, vol. 9. JAI Press.
- SANTOS, C. D. (2012): “Identifying demand shocks from production data,” Tillburg University.
- SOLOW, R. M. (1957): “Technical Change and the Aggregate Production Function,” *Review of Economics and Statistics*, Vol. 39, pp. 312–320.
- SYVERSON, C. (2011): “What Determines Productivity?,” *Journal of Economic Literature*, 49(2), 326–65.
- VAN BEVEREN, I. (2010): “Total Factor Productivity Estimation: A Practical Review,” *Journal of Economic Surveys*, 26, 98–128.
- WALTERS, A. A. (1963): “Production and Cost Functions: An Econometric Survey,” *Econometrica*, 31(1/2), pp. 1–66.
- ZHANG, H. (2012): “Biased Technology and Contribution of Technological Change to Economic Growth: Firm-Level Evidence from China,” working paper, Pennsylvania State University.

Appendices

Appendix A Proofs

Proposition 1 Define,

$$z(M_{jt}, L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta) = \frac{F_{Ljt} L_{jt}}{F_{Mjt} M_{jt}} - \frac{E_{Ljt}}{E_{Mjt}}.$$

Suppose assumptions 1-3 are satisfied and we observe a random sample of $(L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt})$ generated by model parameter $\theta \in \Theta \in \mathbb{R}^h$, which implies that $Z(M_{jt}, L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta) = 0$ $\frac{\partial z}{\partial M_{jt}} \neq 0$. Then there exists a neighborhood $V_\delta(L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta) \in \mathbb{R}^{h+4}$, an open set $W \in \mathbb{R}$ containing M_{jt} , and a real valued function $M^* : V \rightarrow W$, such that,

$$z(M^*(L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta), L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta) = 0.$$

Proof: Under assumption 1, F is second order continuously differentiable and $F_{M_{jt}} > 0$. This implies that $z(M_{jt}, K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}, \theta)$ is first order continuously differentiable. According to the implicit function theorem, $z(M_{jt}, K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}, \theta) = 0$ and $\frac{\partial z}{\partial M_{jt}} \neq 0$ for all $\theta \in \Theta$ implies that there exist an open ball B_x of $(\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt})$ centered at $(K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt})$ and an open ball B_M of \widetilde{M}_{jt} centered at M_{jt} , such that

$$\begin{aligned} & \left\{ \left((\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt}), \widetilde{M}_{jt} \right) \in B_x \times B_M : z \left((\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt}), \widetilde{M}_{jt}; \theta \right) = 0 \right\} \\ = & \left\{ \left((\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt}), M_{jt}^* \left((\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt}); \theta \right) \right) : (\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt}) \in B_x \right\} \end{aligned}$$

Since the observed sample $(K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}) \in B_M$, let

$$(K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}) = \left(\widetilde{K}_{jt}, \widetilde{L}_{jt}, \widetilde{E}_{Mjt}, \widetilde{E}_{Ljt} \right),$$

we have

$$z(M_{jt}, K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}, \theta) = 0 \Leftrightarrow M_{jt} = M_{jt}^*(K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}, \theta)$$

which implies that $z(M_{jt}^*(K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}, \theta), K_{jt}, L_{jt}, E_{Mjt}, E_{Ljt}, \theta) = 0$. ■

Proposition 2 (Multiple Inputs) If the index function $\mu_{jt} = \mu(M_{jt}) : \mathbb{R}^D \rightarrow \mathbb{R}$ is homogeneous of degree $k > 0$, then,

$$z(\mu_{jt}, L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta, k) \equiv \frac{F_{Ljt} L_{jt}}{F_{\mu_{jt}} M_{jt}} - \frac{E_{Ljt}}{\frac{1}{k} E_{Mjt}} = 0. \quad (33)$$

and it is possible to recover $\mu^*(L_{jt}, K_{jt}, E_{Mjt}, E_{Ljt}; \theta, k)$ when the conditions of Proposition 1 hold for $z(\cdot)$ as defined in (33).

Proof: This proposition directly follows from the Euler's Theorem for homogeneous functions, which implies that $\sum_{d=1}^D M_{djt}^* \mu_d(M_{jt}^*) = k\mu(M_{jt}^*)$ if $\mu(\cdot)$ is homogenous of degree k . Therefore, $\psi(M_{jt}^*) = k\mu(M_{jt}^*)$, and in turn, $\frac{\mu(M_{jt}^*)}{\psi(M_{jt}^*)} E_{M_{jt}^*} = \frac{1}{k} E_{M_{jt}^*}^*$. Taking the ratio of (22) and (24) and using this substitution yields (33). Thus, the firm facing price vector $(P_{L_{jt}}, P_{\mu_{jt}})$, will optimally choose $(L_{jt}^*, \mu(M_{jt}^*))$, with associated labor and material expenditure $E_{L_{jt}}^*$ and $\frac{1}{k} E_{M_{jt}}^*$. Recovering the value of μ^* is a straightforward generalization of Proposition 1. Note that in the single input case, μ is just the identity function, which is homogeneous of degree 1. Thus in this special case we know $k = 1$. ■

Appendix B The Standard CES Normalization

The purpose of this appendix is to show how the CES function is normalized in the literature, which can be compared to our normalization in the following appendix.

B.1 Motivation of Normalization

It has been commonly recognized that the CES production function need to be normalized to give meaningful identification of its parameters. There is a branch of literature analyzing the importance and the method of normalization, which includes de La Grandville (1989), Klump and de La Grandville (2000), Klump and Preissler (2000), de La Grandville and Solow (2006), and Leon-Ledesma, McAdam and Willman (2010).

The current literature has illustrated the key motivation of the normalization in details for two-factor-input production function (see Brown and de Cani (1963), Klump and Preissler (2000) and Leon-Ledesma, McAdam and Willman (2010)). However, we will work with three-factor-input production function, $Q = F(L, M, K)$. It is defined as a linear homogeneous function in which the elasticity of substitution between any two factors is a constant. The idea and motivation of the standard normalization procedure can be easily extended to our case. To see this, let us follow the literature by stating the definition of elasticity of substitution σ :

$$\begin{cases} \frac{\partial \ln(M/L)}{\partial \ln(F_L/F_M)} = \sigma \\ \frac{\partial \ln(K/L)}{\partial \ln(F_L/F_K)} = \sigma \end{cases} \quad (34)$$

This definition provides us with a second-order partial differential equation system. Given the assumption of the linear homogenous function, the general solution of the equation system is given by,

$$Q = F(L, M, K) = \lambda_1 [L^\gamma + \lambda_2 M^\gamma + \lambda_3 K^\gamma]^{\frac{1}{\gamma}}, \quad (35)$$

where $\gamma = \frac{\sigma-1}{\sigma}$, and λ_s are three arbitrary constants of integration emerging in the process of solving the differential equation system. One particular functional form used in the literature is obtained by taking $\tilde{\alpha}_L = \frac{1}{1+\lambda_2+\lambda_3}$, $\tilde{\alpha}_M = \frac{\lambda_2}{1+\lambda_2+\lambda_3}$, $\tilde{\alpha}_K = 1 - \tilde{\alpha}_L - \tilde{\alpha}_M$ and $C = \lambda_1(1 + \lambda_2 + \lambda_3)^{\frac{1}{\gamma}}$, thus

$$Q = F(L, M, K) = C[\tilde{\alpha}_L L^\gamma + \tilde{\alpha}_M M^\gamma + \tilde{\alpha}_K K^\gamma]^{\frac{1}{\gamma}}. \quad (36)$$

Here $\tilde{\alpha}_L$, $\tilde{\alpha}_M$ and $\tilde{\alpha}_K$ are referred as distribution parameters.⁴⁴ However, one can obtain different function forms by taking different specifications for λ s. Each of these forms is called a family of CES functions. Examples of different families include ones used in Pitchford (1960), Arrow et al. (1961), and David and van de Klundert (1965). Therefore, as shown in the literature, a common baseline point is needed to compare different families of CES functions whose members are distinguished only by different elasticities of substitution. To this end, one needs to fix baseline point for the level of production (Q_0), factor inputs (L_0, M_0, K_0), and the marginal rates of substitution (μ_{ML0}, μ_{KL0}), which are equal to the price ratios ($P_{M_0}/P_{L_0}, P_{K_0}/P_{L_0}$) because of the cost minimization.⁴⁵ For detailed motivation of normalization, refer to La Grandville (1989) and Leon-Ledesma, McAdam, and Willman (2010).

B.2 Standard Normalization Procedure

We follow de La Grandville (1989) and Leon-Ledesma, McAdam, and Willman (2010) to illustrate the normalization of the three-factor-input CES function. Given the elasticity of substitution σ , for any baseline point $Z_0 = (L_0, M_0, K_0, Q_0, \mu_{ML0}, \mu_{KL0})$, there are four equations about four parameters that characterize one particular family of CES functions:

$$\tilde{\alpha}_L + \tilde{\alpha}_M + \tilde{\alpha}_K = 1, \quad (37)$$

$$\left(\frac{F_M}{F_L}\right)_0 = \frac{\tilde{\alpha}_M}{\tilde{\alpha}_L} \left(\frac{L_0}{M_0}\right)^{1-\gamma} = \mu_{ML0} \equiv \frac{P_{M_0}}{P_{L_0}}, \quad (38)$$

$$\left(\frac{F_K}{F_L}\right)_0 = \frac{\tilde{\alpha}_K}{\tilde{\alpha}_L} \left(\frac{L_0}{K_0}\right)^{1-\gamma} = \mu_{KL0} \equiv \frac{P_{K_0}}{P_{L_0}}, \quad (39)$$

$$Q_0 = C[\tilde{\alpha}_L L_0^\gamma + \tilde{\alpha}_M M_0^\gamma + \tilde{\alpha}_K K_0^\gamma]^{\frac{1}{\gamma}}. \quad (40)$$

The Equation (38) and (39) are implied by cost minimization. Note that the validation of Equation (39) implicitly assumes the optimal choice of capital stock in the short run. The last equation holds since Q_0 is the physical output produced by its corresponding factor inputs. De La Grandville (1989) provides a graphical representation of the normalization. He shows that, after normalization all CES functions in the same family share the common baseline point of tangency, although their elasticities of substitution are different. Therefore, the purpose of normalization is to compare different CES functions in a meaningful way: on the one hand, different families of CES functions can be characterized by different baseline points, on the other hand, the members of each family sharing common baseline point are distinguished only by different elasticities of substitution.

These four equations imply a solution of four parameters:

$$\tilde{\alpha}_L(\sigma, Z_0) = \frac{P_{L_0} L_0^{\frac{1}{\sigma}}}{P_{M_0} M_0^{\frac{1}{\sigma}} + P_{L_0} L_0^{\frac{1}{\sigma}} + P_{K_0} K_0^{\frac{1}{\sigma}}},$$

⁴⁴We use $\tilde{\alpha}$'s to denote the un-normalized (or "original") distribution parameters, while α 's are reserved for the normalized distribution parameters, unless otherwise noticed.

⁴⁵Note that P_{K_0} is the user price of capital, which usually is not accurately measured. To this end, we will extend the normalization to cases where P_{K_0} is not available.

$$\begin{aligned}\tilde{\alpha}_M(\sigma, Z_0) &= \frac{P_{M_0}M_0^{\frac{1}{\sigma}}}{P_{M_0}M_0^{\frac{1}{\sigma}} + P_{L_0}L_0^{\frac{1}{\sigma}} + P_{K_0}K_0^{\frac{1}{\sigma}}}, \\ \tilde{\alpha}_K(\sigma, Z_0) &= \frac{P_{K_0}K_0^{\frac{1}{\sigma}}}{P_{M_0}M_0^{\frac{1}{\sigma}} + P_{L_0}L_0^{\frac{1}{\sigma}} + P_{K_0}K_0^{\frac{1}{\sigma}}}, \\ C(\sigma, Z_0) &= Q_0 \left[\frac{P_{L_0}L_0^{\frac{1}{\sigma}} + P_{M_0}M_0^{\frac{1}{\sigma}} + P_{K_0}K_0^{\frac{1}{\sigma}}}{P_{L_0}L_0 + P_{M_0}M_0 + P_{K_0}K_0} \right]^{\frac{\sigma}{\sigma-1}}.\end{aligned}$$

Note that given the elasticity of substitution, the value of parameters depend on the choice of baseline point Z_0 . Hence, comparing any two CES functions is not informative unless they are specified with the same baseline point.

Substituting the value of these parameters into the original function, we obtain:

$$Q = C(\sigma, Z_0)[\tilde{\alpha}_L(\sigma, Z_0)L^\gamma + \tilde{\alpha}_M(\sigma, Z_0)M^\gamma + \tilde{\alpha}_K(\sigma, Z_0)K^\gamma]^{\frac{1}{\gamma}}.$$

After re-parameterizations, one particular family of CES production function with corresponding normalized parameters is given by

$$Q = Q_0 \left[\alpha_{L_0} \left(\frac{L}{L_0} \right)^\gamma + \alpha_{M_0} \left(\frac{M}{M_0} \right)^\gamma + \alpha_{K_0} \left(\frac{K}{K_0} \right)^\gamma \right]^{\frac{1}{\gamma}},$$

where:

$$\begin{cases} \alpha_{L_0} = \frac{E_{L_0}}{E_{L_0} + E_{M_0} + E_{K_0}} \\ \alpha_{M_0} = \frac{E_{M_0}}{E_{L_0} + E_{M_0} + E_{K_0}} \\ \alpha_{K_0} = 1 - \alpha_{L_0} - \alpha_{M_0} \end{cases}$$

and $E_{L_0} = P_{L_0}L_0$, $E_{M_0} = P_{M_0}M_0$ and $E_{K_0} = P_{K_0}K_0$ are expenditures of labor, material and capital respectively⁴⁶. Hence a normalized CES function is characterized by the baseline point Z_0 and elasticity of substitution σ : while each baseline point specifies a family of CES production functions, the members of each family sharing a common baseline values are distinguished only by different elasticities of substitution. The normalized distribution parameters now solely depend on the baseline point. Thus they can be prefixed before the estimation if normalization equations (38)-(39) hold (thus the normalization is valid).

B.3 Our CES Normalization

In the standard normalization literature, capital is assumed to be a static input which is chosen optimally in each period. However, in practice, capital may be chosen dynamically. For this reason, we extend the standard normalization approach to allow that capital is not running at the cost-minimizing level in the short run.

⁴⁶Note that the expenditure on capital E_{K_0} is different from the capital stock K_0 . But they are related by $E_{K_0} = P_{K_0}K_0$, where P_{K_0} is the user price of capital stock.

Specifically, although capital could be optimally chosen in the long run, the user price of capital (P_K , if available) may not reflect the marginal cost of capital in the short run. To this end, we assume the choice of capital can deviate from the short-run optimal value by certain magnitude of τ which is treated as a parameter to be estimated. This extension also allows for additional flexibility to deal with situations when the user cost of capital service ($E_K = P_K K$) is not available.

We start from the original production function

$$Q = \exp(\tilde{\omega})F(L, M, K) = \exp(\tilde{\omega})[\tilde{\alpha}_L L^\gamma + \tilde{\alpha}_M M^\gamma + \tilde{\alpha}_K K^\gamma]^{\frac{1}{\gamma}}, \quad (41)$$

where $\tilde{\omega}$ is the firm-level productivity.

As suggested by Leon-Ledesma, McAdam, and Willman (2010), the baseline point is chosen as the geometric sample mean:

$$\bar{Z} = (\bar{L}, \bar{M}, \bar{K}, \bar{Q}, \bar{\mu}_{ML}),$$

where $\bar{\mu}_{ML}$ is the average marginal rate of substitution between material and labor (i.e., \bar{P}_M/\bar{P}_L).

Note that the choice of the baseline value specifies a family of CES functions. Given the baseline value, the equations that characterize this family are:

$$\tilde{\alpha}_L + \tilde{\alpha}_M + \tilde{\alpha}_K = 1, \quad (42)$$

$$\left(\frac{F_M}{F_L}\right)_{\bar{Z}} = \frac{\tilde{\alpha}_M}{\tilde{\alpha}_L} \left(\frac{\bar{L}}{\bar{M}}\right)^{1-\gamma} = \bar{\mu}_{ML}, \quad (43)$$

$$\left(\frac{F_K}{F_L}\right)_{\bar{Z}} = \frac{\tilde{\alpha}_K}{\tilde{\alpha}_L} \left(\frac{\bar{L}}{\bar{K}}\right)^{1-\gamma} = \left(\tau \frac{\bar{E}_L}{\bar{E}_K}\right) \bar{\mu}_{KL} = \tau \frac{\bar{L}}{\bar{K}}, \quad (44)$$

$$\bar{Q} = e^{\bar{\omega}}[\tilde{\alpha}_L \bar{L}^\gamma + \tilde{\alpha}_M \bar{M}^\gamma + \tilde{\alpha}_K \bar{K}^\gamma]^{\frac{1}{\gamma}}, \quad (45)$$

where $\bar{\omega}$ is the ‘‘average’’ productivity associated with producing \bar{Q} by $(\bar{L}, \bar{M}, \bar{K})$.

Here τ in equation (44) is introduced as an inefficiency parameter to measure the mean deviation of capital stock from its optimal level. This extension is important for multiple reasons compared with the standard normalization procedure. First, by introducing such an additional flexible parameter, we allow for the case when the capital stock is not optimally chosen in the short run (although it could be optimal in the long run). Specifically, when $\tau = \frac{\bar{E}_K}{\bar{E}_L}$, the marginal rate of substitution of labor and capital at the baseline point is equal to the price ratio, which implies the capital stock is indeed optimally chosen; when $\tau \neq \frac{\bar{E}_K}{\bar{E}_L}$, the actual capital deviates from the optimal amount. We will not specify the value of τ but leave it to be revealed by data as a parameter to estimate. Second, in our empirical application, such a flexible parameter enables us to deal with situations where the average ‘‘price’’ (or the user cost) of capital stock \bar{P}_K (or \bar{E}_K) is not available or accurately measured. In other words, instead of assuming that \bar{P}_K or \bar{E}_K is known, we let it be absorbed in the parameter τ which can be estimated from data.

Given γ and τ , the distribution parameters implied by the equations (42), (43) and (44) are

given by:

$$\begin{aligned}
\tilde{\alpha}_L(\gamma, \tau) &= \frac{\frac{\bar{E}_L}{L^\gamma}}{\frac{\bar{E}_L}{L^\gamma} + \frac{\bar{E}_M}{M^\gamma} + \tau \frac{\bar{E}_L}{K^\gamma}} \\
\tilde{\alpha}_M(\gamma, \tau) &= \frac{\frac{\bar{E}_M}{M^\gamma}}{\frac{\bar{E}_L}{L^\gamma} + \frac{\bar{E}_M}{M^\gamma} + \tau \frac{\bar{E}_L}{K^\gamma}} \\
\tilde{\alpha}_K(\gamma, \tau) &= 1 - \tilde{\alpha}_L(\gamma, \tau) - \tilde{\alpha}_M(\gamma, \tau)
\end{aligned} \tag{46}$$

As in the standard normalization procedure, we plug the distribution parameters into the original CES function to obtain the normalized CES function after re-parametrization:

$$Q = e^{\omega \bar{Q}} \left[\alpha_L \left(\frac{L}{\bar{L}} \right)^\gamma + \alpha_M \left(\frac{M}{\bar{M}} \right)^\gamma + \alpha_K \left(\frac{K}{\bar{K}} \right)^\gamma \right]^{\frac{1}{\gamma}},$$

where

$$\begin{aligned}
\alpha_L &= \frac{\bar{E}_L}{\bar{E}_L + \bar{E}_M + \tau \bar{E}_L} \\
\alpha_M &= \frac{\bar{E}_M}{\bar{E}_L + \bar{E}_M + \tau \bar{E}_L}, \\
\alpha_K &= 1 - \alpha_L - \alpha_M
\end{aligned} \tag{47}$$

and

$$\omega = \tilde{\omega} - \bar{\omega}. \tag{48}$$

Note that, these equations imply $\frac{\alpha_K}{\alpha_L} = \tau$, which is why we define the ratio of α_K and α_L as τ in Equation (12).

Appendix C CES Estimation Equation

In this appendix, we develop the estimation equation for normalized CES production function. Each firm j chooses labor and material quantities to maximize the profit in each period t , given its capital stock and productivity. The firm's static problem is:

$$\max_{L_{jt}, M_{jt}} P_t(Q_{jt})Q_{jt} - P_{L_{jt}}L_{jt} - P_{M_{jt}}M_{jt}, \tag{49}$$

where

$$Q_{jt} = e^{\omega_{jt} \bar{Q}} \left[\alpha_L \left(\frac{L_{jt}}{\bar{L}} \right)^\gamma + \alpha_M \left(\frac{M_{jt}}{\bar{M}} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{\bar{K}} \right)^\gamma \right]^{\frac{1}{\gamma}},$$

and

$$\frac{Q_{jt}}{Q_t} = \left(\frac{P_{jt}}{P_t} \right)^\eta.$$

Note that L_{jt} , M_{jt} and Q_{jt} are physical quantities of labor and material input and output respectively. The first order conditions with respect to labor and material are

$$\frac{1 + \eta}{\eta} \frac{\partial Q_{jt}}{\partial L_{jt}} \frac{Q_{jt}^{1/\eta} P_t}{Q_t^{1/\eta}} = P_{L_{jt}}, \quad (50)$$

$$\frac{1 + \eta}{\eta} \frac{\partial Q_{jt}}{\partial M_{jt}} \frac{Q_{jt}^{1/\eta} P_t}{Q_t^{1/\eta}} = P_{M_{jt}}. \quad (51)$$

Note that $E_{L_{jt}} = P_{L_{jt}} L_{jt}$ and $E_{M_{jt}} = P_{M_{jt}} M_{jt}$, and plug the demand function into above equations we obtain:

$$\frac{1 + \eta}{\eta} \frac{\partial Q_{jt}}{\partial L_{jt}} \frac{L_{jt}}{Q_{jt}} = \frac{E_{L_{jt}}}{R_{jt}}, \quad (52)$$

$$\frac{1 + \eta}{\eta} \frac{\partial Q_{jt}}{\partial M_{jt}} \frac{M_{jt}}{Q_{jt}} = \frac{E_{M_{jt}}}{R_{jt}}, \quad (53)$$

where $R_{jt} = P_{jt} Q_{jt}$ is the revenue for firm j at period t .

Take the ratio with respect to both sides of the equations, and we can solve for material quantity:

$$\frac{M_{jt}}{L_{jt}} = \left[\frac{\alpha_L E_{M_{jt}}}{\alpha_M E_{L_{jt}}} \right]^{\frac{1}{\gamma}} \frac{L_{jt}}{L}. \quad (54)$$

This implies that material quantity can be imputed from observables ($E_{L_{jt}}$, $E_{M_{jt}}$, and L_{jt}) up to unknown parameters. Substitute this M_{jt} in the first order condition for labor and solve for ω_{jt} , we have

$$e^{-\frac{1+\eta}{\eta} \omega_{jt}} = \alpha_L \frac{1 + \eta}{\eta} \frac{P_t}{Q_t^{1/\eta}} \left(\frac{L_{jt}}{L} \right)^\gamma \frac{Q_t^{1/\eta}}{E_{L_{jt}}} \left[\alpha_L \left(\frac{E_{L_{jt}} + E_{M_{jt}}}{E_{L_{jt}}} \right) \left(\frac{L_{jt}}{L} \right)^\gamma + \alpha_K \left(\frac{K_{jt}}{K} \right)^\gamma \right]^{\frac{1}{\gamma} + \frac{1}{\gamma} - 1} \quad (55)$$

Note that the imputed ω_{jt} is also a function of observables.

Plug the imputed M_{jt} and ω_{jt} into the revenue equation:

$$R_{jt} = \exp(u_{jt}) P_t(Q_{jt}) Q_{jt}, \quad (56)$$

after some algebra we have,

$$\ln R_{jt} = \ln \frac{\eta}{1 + \eta} + \ln \left[E_{M_{jt}} + E_{L_{jt}} \left(1 + \tau \left(\frac{K_{jt}/K}{L_{jt}/L} \right)^\gamma \right) \right] + u_{jt}, \quad (57)$$

where $\tau = \frac{\alpha_K}{\alpha_L}$ is the flexible parameter in the normalization of the CES family.

Appendix D Monte Carlo Description

In this appendix, we outline the data generating process for the Monte Carlo experiments. Specifically, the Monte Carlo experiments consist of N replications of simulated data sets, given a set of

true parameters of interest (η , σ , α_L , α_M and α_K).⁴⁷ In each replication, we simulate a sequence of productivity (ω_{jt}), idiosyncratic input prices ($P_{L_{jt}}$ and $P_{M_{jt}}$), and capital stock (K_{jt}) for each firm j over time. Given these variables and random shocks, we derive a sequence of optimal choices of labor and material inputs (L_{jt} and M_{jt}), the optimal output quantity (Q_{jt}) and price (P_{jt}) for firm j in each period t .

In each replication, there are J firms in production for T periods. The evolution process of productivity for each firm is assumed to be a first order Markov process:

$$\omega_{jt+1} = g_0 + g_1\omega_{jt} + \varepsilon_{jt+1}^\omega, \quad (58)$$

where $\varepsilon_{jt+1}^\omega$ is the innovation shock realized in period $t + 1$, which is assumed to be a normally distributed i.i.d. error term with zero mean and standard deviation $se(\varepsilon^\omega)$. The initial productivity of each firm (ω_{j0}) is drawn from a normal distribution of mean ω_0 and standard deviation $se(\omega_0)$.

The investment rule and the capital evolution process are set as,

$$\log(I_{jt}) = \xi\omega_{jt} + (1 - \xi)\log(K_{jt}), \quad (59)$$

$$K_{jt+1} = K_{jt} + I_{jt}, \quad (60)$$

where $\xi \in (0, 1)$ is an arbitrary weight. The initial capital stock of each firm (K_{j0}) is drawn from a normal distribution of mean K_0 and standard variance $se(K_0)$.

The idiosyncratic labor and material input prices ($P_{L_{jt}}$ and $P_{M_{jt}}$) are generated as follows:

$$P_{L_{jt}} = \bar{P}_{L_t} e^{\varepsilon_{jt}^{PL}}, \quad (61)$$

$$P_{M_{jt}} = \bar{P}_{M_t} e^{\varepsilon_{jt}^{PM}}, \quad (62)$$

where \bar{P}_{L_t} and \bar{P}_{M_t} are the industrial-level labor and material prices in period t , which can be drawn from $N(\bar{P}_L, se(\bar{P}_L))$ and $N(\bar{P}_M, se(\bar{P}_M))$ independently, or set to 1 for simplicity as in our implementation. ε_{jt}^{PL} and ε_{jt}^{PM} are deviations from the industrial-level input prices, which are independently drawn from $N(0, se(\varepsilon^{PL}))$ and $N(0, se(\varepsilon^{PM}))$ respectively.

For the demand side, we assume the industrial-level output quantities are generated by,

$$Q_t = r^t Q_0 e^{\varepsilon_t^Q}, \quad (63)$$

where r is the growth rate of the industrial-level output quantity, Q_0 is the initial industrial-level output quantity, and $\varepsilon_t^Q \sim N(0, se(\varepsilon^Q))$ is an independent random shock. The industrial-level output prices are generated by,

$$P_t = Q_t^{1/\eta}, \quad (64)$$

where η is the demand elasticity.

Now we have simulated $\{\omega_{jt}, K_{jt}, I_{jt}, P_{L_{jt}}, P_{M_{jt}}, Q_t, P_t\}$ for each firm j and period t . Given these variables, we can derive the optimal labor and material input choices (L_{jt} and M_{jt}) and the corresponding output quantity (Q_{jt}) for each firm j and period t according to the first order conditions associated with the firm's static profit maximization problem. Specifically, the optimal

⁴⁷To simplify the notation, we drop the tilde over the distribution parameters in this section as well as the report in the associated tables. But please note that the data is generated from the "original" production function (41).

labor input is derived as,

$$L_{jt} = \left(\frac{\alpha_M P_{L_{jt}}}{\alpha_L P_{M_{jt}}} \right)^{\frac{1}{\gamma-1}} M_{jt}, \quad (65)$$

where the material input M_{jt} is given by,

$$M_{jt} = \left[\frac{(e^{-\omega_{jt}} Q_{jt})^\gamma - \alpha_K K^\gamma}{\alpha_M + \alpha_L \left(\frac{\alpha_M P_{L_{jt}}}{\alpha_L P_{M_{jt}}} \right)^{\frac{\gamma}{1-\gamma}}} \right]^{\frac{1}{\gamma}}, \quad (66)$$

and Q_{jt} is the solution of the following equation:

$$\frac{\eta + 1}{\eta} \left(\frac{P_t}{Q_t} \right) Q_{jt}^{\frac{1}{\eta}} = e^{-\omega_{jt}} \frac{\left[P_{M_{jt}} + P_{L_{jt}} \left(\frac{\alpha_M P_{L_{jt}}}{\alpha_L P_{M_{jt}}} \right)^{\frac{1}{\gamma-1}} \right]}{\left[\alpha_M + \alpha_L \left(\frac{\alpha_M P_{L_{jt}}}{\alpha_L P_{M_{jt}}} \right)^{\frac{\gamma}{1-\gamma}} \right]^{\frac{1}{\gamma}}} (1 - \alpha_K K^\gamma (e^{-\omega_{jt}} Q_{jt})^{-\gamma})^{\frac{1}{\gamma}-1}. \quad (67)$$

Given the derived variables and underlying true parameters, equation (67) is only about Q_{jt} . It is easy to verify that equation (67) implies a unique solution for Q_{jt} since given $\eta < -1$, the left hand side is decreasing in Q_{jt} while the right hand side is increasing in Q_{jt} . Denote the solution of the equation as Q_{jt}^* . Once we obtain Q_{jt}^* , we can derive the corresponding L_{jt} and M_{jt} from equation (65) and equation (66). Hence, the expenditures of input are given by $E_{L_{jt}} = P_{L_{jt}} L_{jt}$ and $E_{M_{jt}} = P_{M_{jt}} M_{jt}$. The observed output with a measurement error is given by

$$Q_{jt} = Q_{jt}^* e^{\varepsilon_{jt}^q}, \quad (68)$$

where $\varepsilon_{jt}^q \sim N(0, se(\varepsilon^q))$ is the measurement error. At last, firm level output price P_{jt} is derived by equation

$$\frac{Q_{jt}}{Q_t} = \left(\frac{P_{jt}}{P_t} \right)^\eta, \quad (69)$$

and the firm-level revenue is obtained by

$$R_{jt} = P_{jt} Q_{jt}. \quad (70)$$

Hence, we have generated a data set of $\{\omega_{jt}, K_{jt}, I_{jt}, L_{jt}, M_{jt}, E_{L_{jt}}, E_{M_{jt}}, Q_{jt}, R_{jt}, Q_t, P_t\}$ for each firm j and period t .

Appendix E Additional Application Results

E.1 Comparison to other estimation methods

While we use the OP-KG estimation method as our primary basis of comparison in the main body of the paper, there are many other approaches to estimating production functions. In this appendix, we compare our method to three additional approaches. First, we implement a simple nonlinear least squares estimator for the production function which proxies for materials with expenditure

and also ignores the presence of heterogeneity. Second we use an approach that uses the first order conditions to control for productivity, but continues to use a materials expenditure to proxy for materials quantities. Finally, we compare our estimator to a panel data estimator a la Arellano and Bond (1991), where the productivity term includes a fixed effect and an AR(1) process.

First, we estimate the model with naive nonlinear least square estimation, in which the material expenditure is used as a proxy of quantity and the productivity is lumped into the additive error term. Specifically, the following model is estimated:⁴⁸

$$\ln\left(\frac{R_{jt}}{\bar{R}}\right) = \ln\left(\frac{P_t}{\bar{P}_t}\right) - \frac{1}{\eta} \ln\left(\frac{Q_t}{\bar{Q}_t}\right) + \frac{1}{\gamma} \frac{1+\eta}{\eta} \ln\left[\alpha_L \left(\frac{L_{jt}}{\bar{L}}\right)^\gamma + \alpha_M \left(\frac{E_{M_{jt}}}{\bar{E}_M}\right)^\gamma + \alpha_K \left(\frac{K_{jt}}{\bar{K}}\right)^\gamma\right] + u_{jt}, \quad (71)$$

where $\gamma = \frac{\sigma-1}{\sigma}$ and σ is the elasticity of substitution. P_t and Q_t are industry-level output price and quantity. Note that u_{jt} contains both the productivity and measurement error.

The result is shown in the second column of Table A.1 under the title ‘NLLS’. For comparison purposes, the first column reproduces our estimates from Table 4, while the final column reproduces the OP-KG estimates. We can immediately see that controlling for productivity is essential to producing reasonable estimates of the demand parameter η , which has the wrong sign and an extremely high magnitude under the NLLS specification.

Secondly, we estimate the model with nonlinear least square estimation, with the proxy of material quantity (i.e., material expenditure) and the productivity imputed from the first order condition of labor input. To be specific, with the productivity imputed from the first order condition of labor input, the revenue equation can be derived similarly to Appendix C:

$$\ln R_{jt} = \ln \frac{\eta}{1+\eta} + \ln \left[E_{L_{jt}} \frac{\alpha_M}{\alpha_L} \left(\frac{M_{jt}/\bar{M}}{L_{jt}/\bar{L}} \right)^\gamma + E_{L_{jt}} \left(1 + \frac{\alpha_K}{\alpha_L} \left(\frac{K_{jt}/\bar{K}}{L_{jt}/\bar{L}} \right)^\gamma \right) \right], \quad (72)$$

where $\frac{\alpha_M}{\alpha_L} = \frac{\bar{E}_M}{\bar{E}_L}$ and $\frac{\alpha_K}{\alpha_L} = \tau$ (to be estimated).

Since M_{jt} is not observed, we use its proxy $E_{M_{jt}}$. Thus, the following empirical equation is estimated:

$$\ln R_{jt} = \ln \frac{\eta}{1+\eta} + \ln \left[E_{L_{jt}} \frac{\alpha_M}{\alpha_L} \left(\frac{E_{M_{jt}}/\bar{E}_M}{L_{jt}/\bar{L}} \right)^\gamma + E_{L_{jt}} \left(1 + \frac{\alpha_K}{\alpha_L} \left(\frac{K_{jt}/\bar{K}}{L_{jt}/\bar{L}} \right)^\gamma \right) \right] + e_{jt}, \quad (73)$$

where e_{jt} is the measurement error.

The result from this table is presented in column 3 of Table A.1 under the title ‘Prod.’. Controlling for productivity generates a reasonable demand parameter, as opposed to the earlier approach. However, we see that now, the elasticity of substitution parameter is substantially larger than in all other methods. This is intuitive. Note that the first order conditions for labor and material implies that $E_{L_{jt}} \frac{\alpha_M}{\alpha_L} \left(\frac{M_{jt}/\bar{M}}{L_{jt}/\bar{L}} \right)^\gamma = E_{M_{jt}}$. The difference between this estimation and our method is that we utilize this relationship rather than using a proxy of material quantity. As shown in the table, the elasticity of substitution is significantly larger than our estimates, because the variance of $E_{L_{jt}} \frac{\alpha_M}{\alpha_L} \left(\frac{E_{M_{jt}}/\bar{E}_M}{L_{jt}/\bar{L}} \right)^\gamma$ is larger (1.2 ~ 3 times) than the variance of $E_{M_{jt}}$.

Third, we estimate a CES version of persistent panel data method (Arellano and Bond, 1991;

⁴⁸To make results comparable, we estimate this normalized revenue equation instead of production function.

Blundell and Bond, 2000), with an AR(1) term and a fixed effect. In particular, consider the empirical equation (71), but now the error term is decomposed as

$$u_{jt} = \beta_j + \nu_{jt} + \epsilon_{jt}, \quad (74)$$

and

$$\nu_{jt} = \rho\nu_{jt-1} + v_{jt}, \quad (75)$$

where ϵ_{jt} is the i.i.d. measurement error and v_{jt} is the i.i.d. innovation term.

Following the persistent panel data method, we take the quasi-difference

$$D_t(\eta, \sigma, \tau, \rho) = u_{jt} - \rho u_{jt-1}, \quad (76)$$

$$D_{t-1}(\eta, \sigma, \tau, \rho) = u_{jt-1} - \rho u_{jt-2}, \quad (77)$$

where u_{jt} is given by⁴⁹

$$u_{jt} = \ln\left(\frac{R_{jt}}{\bar{R}}\right) - \left\{ \ln\left(\frac{P_t}{\bar{P}_t}\right) - \frac{1}{\eta} \ln\left(\frac{Q_t}{\bar{Q}_t}\right) + \frac{1}{\gamma} \frac{1+\eta}{\eta} \ln\left[\alpha_L \left(\frac{L_{jt}}{\bar{L}}\right)^\gamma + \alpha_M \left(\frac{E_{M_{jt}}}{\bar{E}_M}\right)^\gamma + \alpha_K \left(\frac{K_{jt}}{\bar{K}}\right)^\gamma \right] \right\}.$$

Then we construct the following moment condition for GMM estimation:

$$E\left[D_t(\eta, \sigma, \tau, \rho) - D_{t-1}(\eta, \sigma, \tau, \rho) | E_{L_{jt-2}}, E_{M_{jt-2}}, K_{jt-2}, K_{jt-2}^2\right] = 0. \quad (78)$$

The result is reported in the fourth column of A.1 under the title ‘AB’. Like the OP-KG estimator, we would expect this the elasticity of substitution to be biased downward using this approach, due to the way the expenditure proxy for materials is employed. In fact, we do see that this estimator, like OP-KG estimates a smaller $\hat{\sigma}$ relative to our method.

E.2 Comparing Productivity Measures

As with other structural approaches to production function estimation, there are two potential approaches to defining “productivity” in our model. In the body of the paper, we follow the most common approach, and report the distribution of $\widehat{\omega_{it} + u_{it}}$ which is the residual from the production function itself. This represents the sum of productivity anticipated by the firm as well as unanticipated productivity and potential measurement error in revenues. Alternatively, we could employ (7) to recover $\hat{\omega}_{it}$ alone from the system of first order conditions. This approach includes only an estimate of productivity anticipated by the firm when it makes its labor and materials decision. A similar approach could be recover $\hat{\omega}_{it}$ alone using the OP-KG procedure. However in this case, the anticipated productivity relates to the firm’s expectation of productivity when making the *investment* decision which occurs later than the hiring decision according to the timing assumptions.

It is interesting to see whether these different definitions of the productivity distribution yield substantially different results. We investigate this in Figure A.1, which plots the two distributions for the two different methods for the Clothing industry. (Results for other industries are similar and are available by request). We see that while there is a substantial difference across methods (as

⁴⁹Note that all α ’s are parameterized as in (47), as a function of τ .

is also visible in Figure 5), the difference across definitions for a given method is relatively small. It is particularly small when using our method. This implies that the bulk of the variance in the distribution of productivity is due to anticipated productivity differences, which further supports the importance of controlling for productivity differences when estimating production functions.

Finally, Figure A.2 presents the two distributions for our method only across all four industries. It shows that the result that the two distributions are quite similar is robust across the four industries we consider in the main body of the paper.

Table 1: Monte Carlo Parameter Values

Single material	Description	Value
η	Demand elasticity	-4
σ	Elasticity of substitution	0.8, 1.5, 2.5
α_L	Distribution parameter of labor	0.4
α_M	Distribution parameter of material	0.4
α_K	Distribution parameter of capital	0.2
g_0	Parameter in productivity evolution	0.2
g_1	Parameter in productivity evolution	0.95
$se(\varepsilon^\omega)$	Standard error of productivity innovation	0.01
$se(\varepsilon^{PL})$	Standard error of labor price shock	0.2
$se(\varepsilon^{PM})$	Standard error of material price shock	0.2
$se(\varepsilon^q)$	Standard error of output measurement error	0.01
T	Number of periods	10
J	Number of firms	100
N	Number of Monte Carlo replications	1000
Multiple materials ¹		
σ	Elasticity of substitution	1.5
P_{M1}	Price of $M1$ (constant)	0.1
P_{M2}	Price of $M2$ (constant)	0.18
$se(\varepsilon^{PM3})$	Standard error of price shock of $M3$	0.2
β	Effective factor of $M1$	0.65
δ_1	Elasticity of substitution between $M1$ and $M3$	2.1
δ_2	Elasticity of substitution between $M2$ and $M3$	2.0

¹ Only parameters different from the single material input setting are listed.

Table 2: Monte Carlo: Estimated results from three different methods¹

True	$\sigma = 0.8$			$\sigma = 1.5$			$\sigma = 2.5$		
	Us	OP	Oracle	Us	OP	Oracle	Us	OP	Oracle
$\hat{\eta}$	-4.00	-	-	-3.999 (0.001)	-	-	-3.999 (0.000)	-	-
				[0.146]			[0.015]		
$\hat{\sigma}$	0.800	0.669	0.802	1.501	1.208	1.502	2.499	1.988	2.495
				(0.011)	(0.174)	(0.030)	(0.027)	(0.039)	(0.068)
				[0.011]	[0.146]	[0.013]	[0.029]	[0.043]	[0.037]
$\hat{\alpha}_L$	0.40	0.483	0.400	0.400	0.478	0.400	0.400	0.459	0.401
				(0.007)	(0.007)	(0.007)	(0.004)	(0.002)	(0.003)
				[0.005]	[0.083]	[0.003]	[0.001]	[0.059]	[0.001]
$\hat{\alpha}_M$	0.40	0.484	0.400	0.400	0.479	0.400	0.400	0.460	0.400
				(0.007)	(0.007)	(0.007)	(0.004)	(0.005)	(0.003)
				[0.005]	[0.084]	[0.003]	[0.001]	[0.060]	[0.001]
$\hat{\alpha}_K$	0.20	0.033	0.199	0.200	0.043	0.201	0.200	0.082	0.199
				(0.014)	(0.015)	(0.013)	(0.009)	(0.021)	(0.005)
				[0.009]	[0.166]	[0.005]	[0.003]	[0.118]	[0.003]
$\hat{\tau}^2$		0.325	0.326	0.209	0.057	0.209	0.174	0.074	0.173
				(0.028)	(0.002)	(0.001)	(0.009)	(0.003)	(0.001)
				[0.028]	[0.296]	[0.004]	[0.009]	[0.099]	[0.003]

¹ The table reports the medians of N replications of each case. The standard errors are included in the parentheses and root mean squared errors are in the square brackets.

² $\hat{\tau}$ is endogenously implied by the choice of capital stock. True value of $\hat{\tau}$ for the three cases are 0.326, 0.209 and 0.173 respectively.

Figure 1: Monte Carlo experiments: true $\sigma = 0.8, 1.5, 2.5$

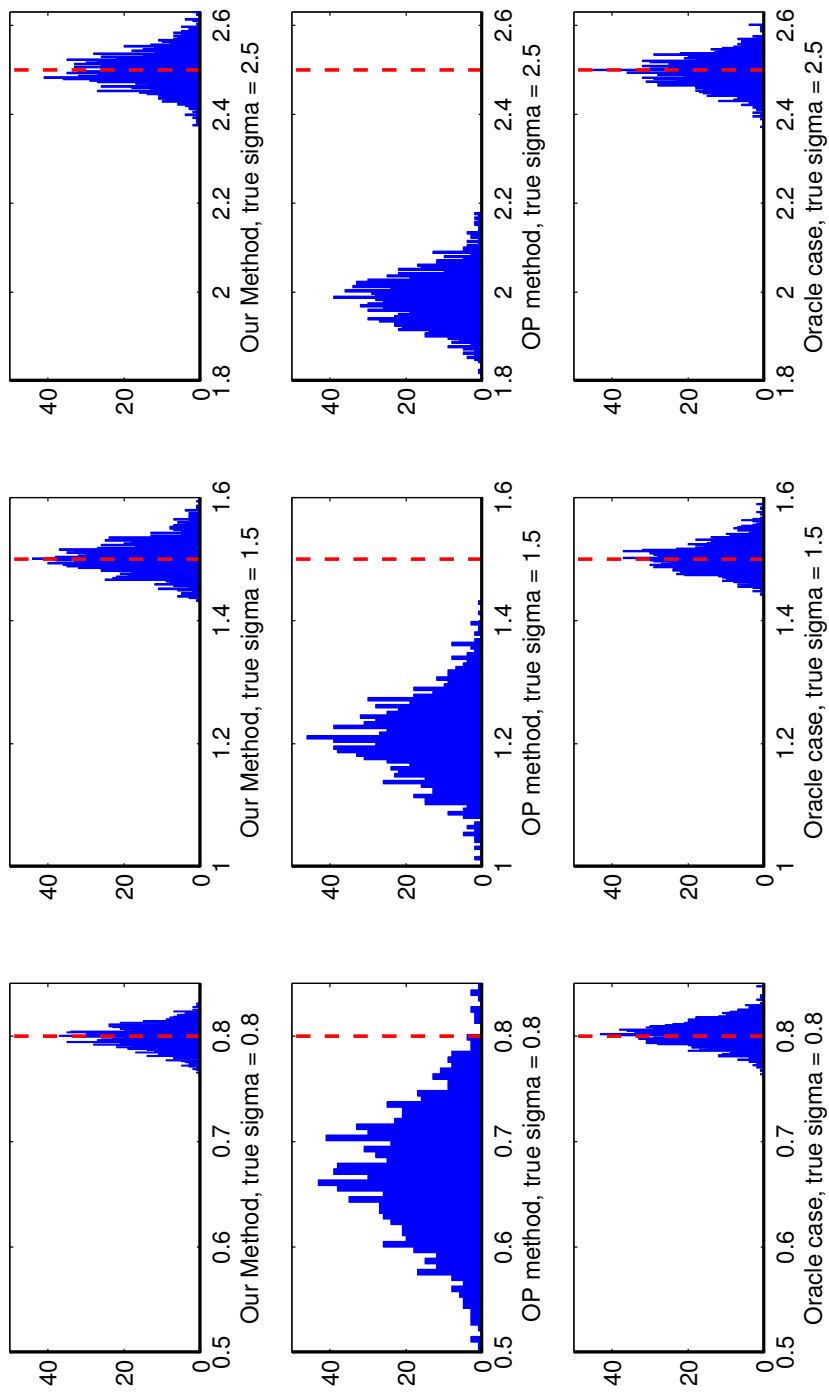


Figure 2: Kernel density estimation: imputed material prices v.s. true material prices

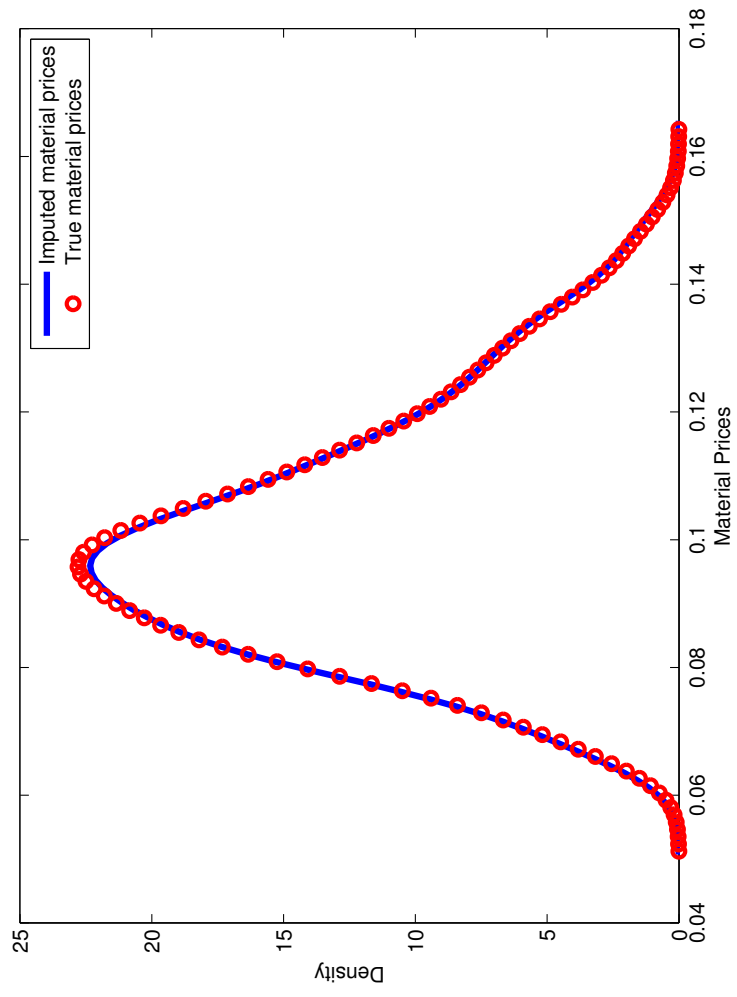


Figure 3: Profit difference in choosing different quality levels

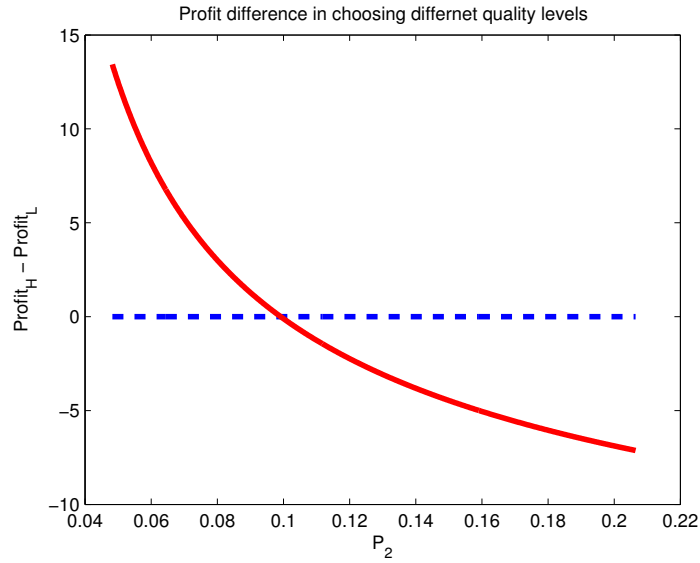


Table 3: Multiple Materials Monte Carlo: Parameter Estimates¹

Parameter	$\hat{\eta}$	$\hat{\sigma}$	$\hat{\alpha}_L$	$\hat{\alpha}_M$	$\hat{\alpha}_K$
True	-4.000	1.500	0.400	0.400	0.200
Estimation	-3.997	1.497	0.400	0.400	0.200
SE	(0.001)	(0.026)	(0.004)	(0.005)	(0.008)
RMSE	[0.052]	[0.027]	[0.003]	[0.001]	[0.003]

¹ The table reports the medians of $N = 1000$ replications of each case. The standard errors are included in the parentheses and root mean squared errors are in the square brackets.

Figure 4: Multiple Materials Monte Carlo: True and recovered material and material price index

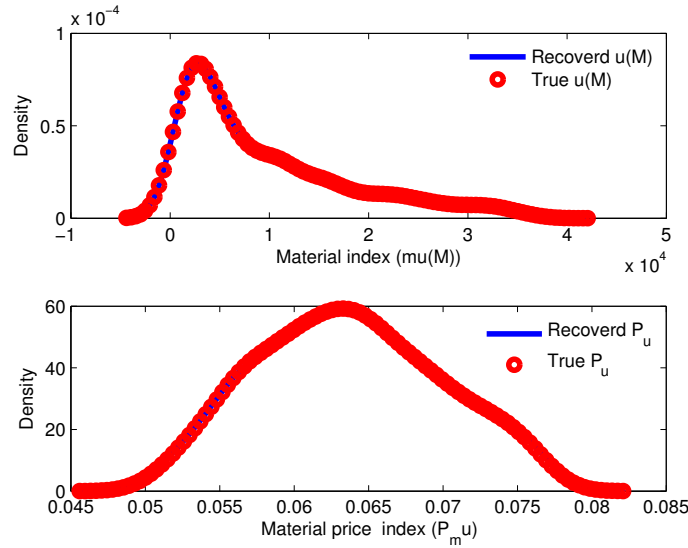


Table 4: Estimated results for Colombian industries

	Clothing		Bakery Products		Printing & Publishing		Metal Furniture	
	Us	OP-KG	Us	OP-KG	Us	OP-KG	Us	OP-KG
$\hat{\eta}$	-5.768 (0.121)	-8.465 (1.544)	-5.231 (0.188)	-5.253 (0.417)	-4.659 (0.236)	-12.161 (5.434)	-5.518 (0.433)	-6.947 (2.686)
$\hat{\sigma}$	1.948 (0.234)	0.361 (0.018)	1.443 (0.117)	0.401 (0.011)	2.555 (0.405)	0.593 (0.054)	1.772 (0.379)	0.393 (0.045)
$\hat{\alpha}_L$	0.361 (0.002)	0.371 (0.001)	0.244 (0.002)	0.251 (0.000)	0.372 (0.005)	0.381 (0.004)	0.300 (0.005)	0.304 (0.005)
$\hat{\alpha}_M$	0.601 (0.003)	0.618 (0.001)	0.705 (0.006)	0.725 (0.001)	0.537 (0.007)	0.549 (0.005)	0.637 (0.010)	0.647 (0.011)
$\hat{\alpha}_K$	0.038 (0.004)	0.011 (0.002)	0.050 (0.007)	0.025 (0.002)	0.091 (0.013)	0.070 (0.009)	0.064 (0.015)	0.049 (0.016)
$\hat{\tau}$	0.106 (0.012)	0.030 (0.005)	0.207 (0.032)	0.098 (0.008)	0.245 (0.037)	0.183 (0.025)	0.213 (0.055)	0.160 (0.054)
\hat{g}_0	0.008 (0.010)	0.101 (0.011)	0.039 (0.015)	0.148 (0.011)	-0.025 (0.015)	0.211 (0.039)	-0.033 (0.024)	0.219 (0.072)
\hat{g}_1	0.695 (0.014)	0.972 (0.010)	0.822 (0.012)	0.955 (0.005)	0.906 (0.019)	0.950 (0.014)	0.824 (0.026)	0.877 (0.026)
#Obs	5763		2269		2377		903	

Figure 5: Kernel density estimation of productivity ($\widehat{\omega_{jt} + u_{jt}}$)

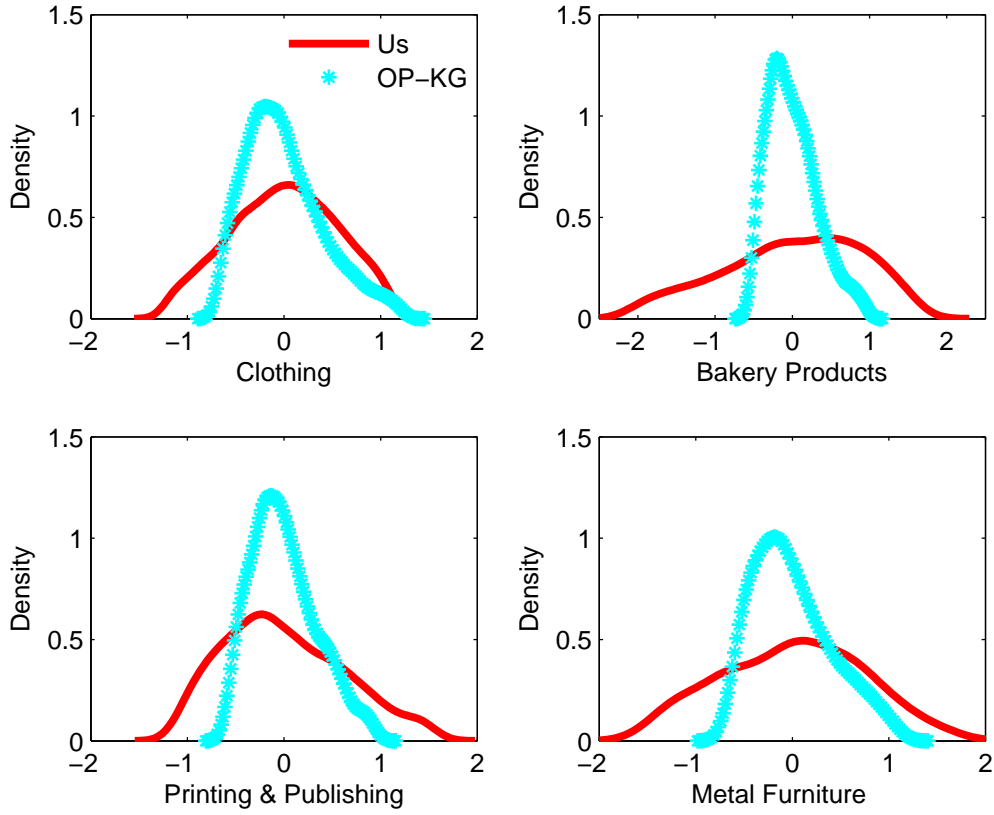


Table 5: Persistence of imputed material prices

	Persistence	Standard Error
Clothing	0.77	0.20
Bakery Products	0.77	0.21
Printing & Publishing	0.82	0.18
Metal Furniture	0.68	0.19

Table 6: Correlations between imputed productivity and input prices in logarithm

	$\text{corr}(\hat{\omega}, \log(\hat{P}_M))$	$\text{corr}(\hat{\omega}, \log(P_L))$	$\text{corr}(\log(\hat{P}_M), \log(P_L))$
Clothing	0.76	0.60	0.26
Bakery Products	0.93	0.58	0.40
Printing & Publishing	0.68	0.88	0.68
Metal Furniture	0.85	0.65	0.48

Figure 6: Kernel density estimation of imputed material prices in logarithm

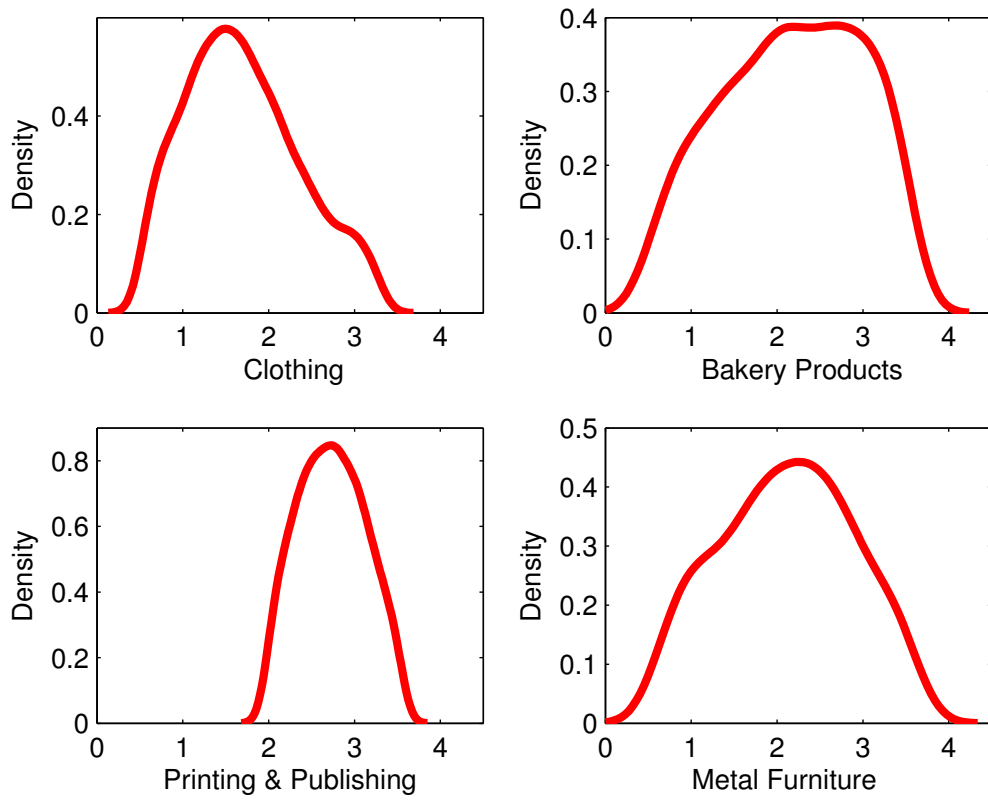


Table A.1: Estimated results by various methods for Colombian industries

	Clothing					Bakery Products				
	Us	NLLS	Prod.	AB	OP-KG	Us	NLLS	Prod.	AB	OP-KG
$\hat{\eta}$	-5.768 (0.121)	109.423 (39.317)	-5.020 (0.105)	-2.416 (0.373)	-8.465 (1.544)	-5.231 (0.188)	24.564 (1.927)	-4.650 (0.156)	-2.546 (0.484)	-5.253 (0.417)
$\hat{\sigma}$	1.948 (0.234)	1.880 (0.077)	6.826 (0.292)	0.569 (0.229)	0.361 (0.018)	1.443 (0.117)	1.411 (0.104)	4.369 (0.233)	0.363 (0.298)	0.401 (0.011)
$\hat{\alpha}_L$	0.361 (0.002)	0.372 (0.001)	0.370 (0.001)	0.358 (0.037)	0.371 (0.001)	0.244 (0.002)	0.253 (0.001)	0.254 (0.001)	0.244 (0.025)	0.251 (0.000)
$\hat{\alpha}_M$	0.601 (0.003)	0.620 (0.002)	0.617 (0.002)	0.597 (0.062)	0.618 (0.001)	0.705 (0.006)	0.731 (0.002)	0.734 (0.004)	0.704 (0.072)	0.725 (0.001)
$\hat{\alpha}_K$	0.038 (0.004)	0.008 (0.004)	0.013 (0.003)	0.045 (0.100)	0.011 (0.002)	0.050 (0.007)	0.016 (0.003)	0.012 (0.005)	0.053 (0.097)	0.025 (0.002)
$\hat{\tau}$	0.106 (0.012)	0.021 (0.010)	0.034 (0.009)	0.126 (0.278)	0.030 (0.005)	0.207 (0.032)	0.064 (0.013)	0.047 (0.021)	0.218 (0.419)	0.098 (0.008)
\hat{g}_0	0.008 (0.010)		0.048 (0.010)		0.101 (0.011)	0.039 (0.015)		0.029 (0.015)		0.148 (0.011)
\hat{g}_1	0.695 (0.014)		0.771 (0.018)		0.972 (0.010)	0.822 (0.012)		0.873 (0.022)		0.955 (0.005)
$\hat{\rho}$				0.547 (0.071)					0.264 (0.191)	
#Obs	5763					2269				
	Printing & Publishing					Metal Furniture				
	Us	NLLS	Prod.	AB	OP-KG	Us	NLLS	Prod.	AB	OP-KG
$\hat{\eta}$	-4.659 (0.236)	14.142 (0.986)	-3.857 (0.135)	-2.189 (1.739)	-12.161 (5.434)	-5.518 (0.433)	10.514 (0.918)	-4.186 (0.210)	-6.280 (6.713)	-6.947 (2.686)
$\hat{\sigma}$	2.555 (0.405)	2.218 (0.186)	3.326 (0.145)	0.313 (0.405)	0.593 (0.054)	1.772 (0.379)	1.395 (0.090)	5.600 (0.598)	0.429 (0.212)	0.393 (0.045)
$\hat{\alpha}_L$	0.372 (0.005)	0.379 (0.003)	0.389 (0.004)	0.298 (0.070)	0.381 (0.004)	0.300 (0.005)	0.295 (0.002)	0.319 (0.003)	0.233 (0.057)	0.304 (0.005)
$\hat{\alpha}_M$	0.537 (0.007)	0.546 (0.004)	0.560 (0.005)	0.429 (0.101)	0.549 (0.005)	0.637 (0.010)	0.627 (0.005)	0.677 (0.005)	0.494 (0.121)	0.647 (0.011)
$\hat{\alpha}_K$	0.091 (0.013)	0.075 (0.006)	0.051 (0.009)	0.273 (0.171)	0.070 (0.009)	0.064 (0.015)	0.078 (0.008)	0.004 (0.008)	0.273 (0.178)	0.049 (0.016)
$\hat{\tau}$	0.245 (0.037)	0.199 (0.018)	0.132 (0.024)	0.915 (0.722)	0.183 (0.025)	0.213 (0.055)	0.262 (0.028)	0.013 (0.024)	1.176 (0.841)	0.160 (0.054)
\hat{g}_0	-0.025 (0.015)		-0.018 (0.015)		0.211 (0.039)	-0.033 (0.024)		-0.003 (0.024)		0.219 (0.072)
\hat{g}_1	0.906 (0.019)		0.933 (0.017)		0.950 (0.014)	0.824 (0.026)		0.863 (0.033)		0.877 (0.026)
$\hat{\rho}$				0.358 (0.380)					-0.303 (0.462)	
#Obs	2377					903				

Figure A.1: Comparison: densities of $\hat{\omega}$ and $\hat{\omega} + \hat{u}$ – large demonstration

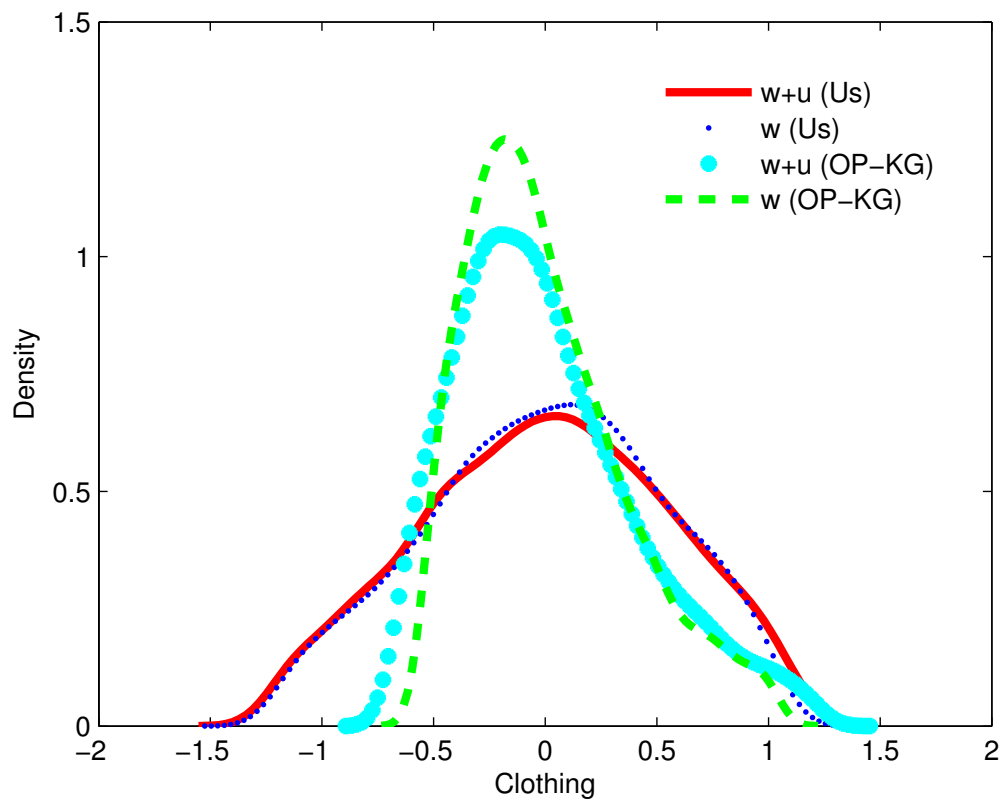


Figure A.2: Comparison: densities of $\hat{\omega}$ and $\hat{\omega} + \hat{u}$ from our method – small figures

