

# Heterogeneity in Firms: A Proxy Variable Approach for Quantile Production Functions

Justin Doty\* and Suyong Song†

March 2, 2021

## Abstract

We propose a new approach to estimate firm-level production functions in which output elasticities are heterogeneous across the firm size distribution. This paper extends the proxy variable approach for estimating production functions to the conditional quantiles of firm production. Production function parameters are estimated in a simple two-stage approach which relies a location-shift assumption on unobserved productivity. We show that this method allows us to capture heterogeneity in output elasticities along the firm size distribution that may not be found in conditional mean estimates of [Olley and Pakes \(1996\)](#) and [Levinsohn and Petrin \(2003\)](#). We provide small-sample evidence in a Monte Carlo study to show that this approach is robust compared to other production function estimators. The method is applied to firm and plant-level manufacturing data from the US, Chile, and Colombia.

*Keywords:* Production functions, Heterogeneous elasticity, Quantile regression

*JEL Classification:* C14, C36, D24

## 1 Introduction

Production function estimation is an ongoing and historical empirical research topic that links firm's input to output decisions. Identification of the output elasticities and consequently the distribution of firm-level productivity is constrained by endogeneity issues. This is because productivity is unobserved by the econometrician, but observed by the firm when making input decisions.

A popular approach to address this issue is to introduce a proxy variable such as investment, made popular by [Olley and Pakes \(1996\)](#) (OP) or an intermediate input using [Levinsohn and Petrin \(2003\)](#) (LP) or [Akerberg et al. \(2015\)](#) (ACF). These proxies are functions of a state variable such as capital and unobserved productivity. Under certain assumptions, this function is strictly increasing in its scalar unobserved productivity component. Inverting this function controls for unobserved

---

\*Department of Economics, University of Iowa, S321 Pappajohn Business Building, 21 E Market St, Iowa City, IA 52242. Email: [justin-doty@uiowa.edu](mailto:justin-doty@uiowa.edu)

†Department of Economics and Finance, University of Iowa, W360 Pappajohn Business Building, 21 E Market St, Iowa City, IA 52242. Email: [suyong-song@uiowa.edu](mailto:suyong-song@uiowa.edu)

productivity and the production function parameters can be estimated with a simple two-stage estimator.

While these methods have been useful in identifying the production function parameters and recovering consistent estimates of total factor productivity (TFP), resulting estimates may be biased if there is additional heterogeneity in production technology across firms. Thus, allowing for heterogeneous coefficients is one possible way to capture these differences. The literature on heterogeneous production functions is small relative to the empirical research using the homogeneous coefficient model, even though many empirical studies have found heterogeneity in firm behaviors and decisions.<sup>1</sup> This is because estimating the homogeneous coefficient model by itself is very difficult due to the issue of unobserved productivity.

In our approach we allow for firm heterogeneity in production technology beyond a Hick’s neutral productivity shock to be driven by the rank of the unobserved ex-post production shock. We use the proxy variable approach in this framework in order to control for the part of production unobservables that are correlated with input decisions. The literature on control function approaches for quantile regression models is still a developing area therefore it is not straightforward to estimate production functions by allowing for endogenous inputs and their heterogeneous coefficients.<sup>2</sup> We are not aware of any published paper which takes into account the endogeneity issue of production functions in the conventional quantile regression framework. We fill the gap in this paper by proposing an easy-to-implement estimator.

We show through simulation, that our proposed two-step estimator performs relatively well to the popular control function approach of [Levinsohn and Petrin \(2003\)](#) and is successful in both capturing heterogeneous output elasticities and controlling for unobserved productivity. In our empirical application, we consider several popular firm and plant-level manufacturing datasets and compare our estimator to the LP estimator. We show that heterogeneity in these estimates implies differences in other features of firm production, such as returns to scale, capital intensity and productivity.

The rest of the paper is organized as follows. Section 2 reviews prior approaches for production function estimation and the interpretations of quantile production functions. Section 3 introduces the econometric model and the proposed estimator. Section 4 presents finite-sample behaviors of the estimator via Monte Carlo experiments and Section 5 applies this estimator to US, Chilean, and Colombian manufacturing datasets. Section 6 concludes with directions for future research.

---

<sup>1</sup>Some notable examples are [Kasahara, Schrimpf and Suzuki \(2017\)](#), [Balat, Brambilla and Sasaki \(2018\)](#), [Li and Sasaki \(2017\)](#) and [Dermirer \(2020\)](#) to name of few. Also [Gandhi \*et al.\* \(2020\)](#) who estimate a nonparametric production function and obtain heterogeneous estimates.

<sup>2</sup>See for example [Chesher \(2003\)](#), [Ma and Koenker \(2006\)](#), [Lee \(2007\)](#) and [Imbens and Newey \(2009\)](#)

## 2 Literature Review

### 2.1 Production Function Estimation

We briefly review the LP (2003) procedure for estimating a *value-added* production function (in logs)<sup>3</sup>.

$$y_{it} = (\beta_0 + \omega_{it}) + \beta_k k_{it} + \beta_l l_{it} + \eta_{it}, \quad (1)$$

where  $y_{it}$  denotes value-added output for firm  $i$  at time  $t$ ,  $l_{it}$  denotes labor input,  $k_{it}$  denotes capital input,  $\omega_{it}$  is unobserved productivity and  $\eta_{it}$  denotes an iid shock to production.<sup>4</sup>

To control for the correlation between  $\omega_{it}$  and inputs  $k_{it}$  and  $l_{it}$ , LP introduces an intermediate input demand defined as<sup>5</sup>

$$m_{it} = m_t(k_{it}, \omega_{it}), \quad (2)$$

where the function  $m$  is strictly increasing in  $\omega_{it}$  for all  $k_{it}$ . Productivity can then be expressed as

$$\omega_{it} = m_t^{-1}(k_{it}, m_{it}). \quad (3)$$

Substituting this equation into the production function they obtain

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + m_t^{-1}(k_{it}, m_{it}) + \eta_{it} = \beta_l l_{it} + \Phi(k_{it}, m_{it}) + \eta_{it}. \quad (4)$$

An estimate for  $\beta_l$  and  $\Phi_t(k_{it}, m_{it})$  can be obtained by the following first stage moment restriction

$$\mathbb{E}[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, m_{it}) | \mathcal{I}_{it}] = 0, \quad (5)$$

where  $\mathcal{I}_{it}$  denotes the firm's information at time  $t$ . First stage estimates of  $\beta_l$  and  $\Phi$  can be obtained by a local linear regression or a polynomial regression in  $(k_{it}, m_{it})$ .

A second stage moment restriction identifies the coefficient on capital. Assume that productivity follows an auto-regressive process

$$\omega_{it} = \mathbb{E}[\omega_{it} | \omega_{it-1}] + \xi_{it} = g(\omega_{it-1}) + \xi_{it} \quad (6)$$

where  $\xi_{it}$  denotes an innovation to productivity and satisfies  $\mathbb{E}[\xi_{it} | \mathcal{I}_{it-1}] = 0$ .

---

<sup>3</sup>We consider a value-added production function here to be consistent with the model we introduce in Section 3

<sup>4</sup>We present the production function this way (with the constant and productivity in parenthesis) since the constant is not separately identified from the mean of productivity. This notation is consistent with our model we introduce later.

<sup>5</sup>In the original paper of Levinsohn and Petrin (2003) they consider multiple intermediate inputs such as energy, fuels, and materials as potential proxies. We focus on material inputs as the proxy.

Then, the production function parameters can be estimated from the moment restrictions

$$\begin{aligned} \mathbb{E}[\xi_{it} + \eta_{it} | \mathcal{I}_{it-1}] = \\ \mathbb{E}[\hat{y}_{it} - \beta_k k_{it} \\ - g(\hat{\Phi}_{t-1}(k_{it-1}, m_{it-1}) - \beta_k k_{it-1}) | \mathcal{I}_{it-1}] = 0, \end{aligned} \quad (7)$$

where  $\hat{y}_{it} = y_{it} - \hat{\beta}_l l_{it}$  and  $\hat{\Phi}$  denotes estimates from the first stage. LP proceed by using instruments from  $\mathcal{I}_{it-1}$  and minimize a Generalized Method of Moments (GMM) criterion function. Standard errors are obtained using a bootstrap procedure since the two-step nature of this estimator complicates asymptotic inference.

## 2.2 Production Functions and Quantile Regression

A quantile regression framework for production functions may facilitate many different interpretations. Since this framework has been applied in the production frontier literature, we briefly review how this may be a natural interpretation and the limitations of such a model. A (stochastic) frontier (SFA) model of production proposed by [Aigner \*et al.\* \(1977\)](#) introduces statistical error into a frontier model. Frontier models assume firms deviate from an optimal frontier of production. The SFA model is typically written as

$$y_i = f(x_i, \beta) + \varepsilon_i, \quad (8)$$

where  $\varepsilon_i = \eta_i - u_i$ ,  $x_i$  are inputs to production and  $\beta$  are the parameters. The error term  $\eta_i$  denotes the statistical noise in the model such as measurement error and  $u_{it}$  represents one-sided deviations from the production frontier. Estimates of  $\beta$  are typically obtained using maximum likelihood which requires strong distributional assumptions on the error terms. [Bernini \*et al.\* \(2004\)](#) suggests quantile regression could be used to estimate the highest percentiles of the conditional output distribution as an estimate of the production frontier. [Aragon \*et al.\* \(2005\)](#) uses a characterization of a deterministic frontier that interprets production functions as being from a continuous interval  $\tau \in [0, 1]$  where  $\tau = 1$  corresponds to the efficient frontier. The difficulty with these approaches is choosing which quantile to estimate as the frontier and performing inference on these extremal quantiles.<sup>6</sup> Predicting differences between the frontier and a production process for a given firm is also complicated since the predicted error is a composite of technical inefficiency and noise. Other econometric issues, such as endogeneity of input choices with respect to inefficiency are difficult to incorporate in this framework. The purpose of this paper is not to debate the advantages and disadvantages of production frontier models so we leave this discussion for future studies and acknowledge the theoretical challenges of quantile frontiers and instead focus on the standard

---

<sup>6</sup>See [Chernozhukov \(2005\)](#) on how to extend inference in quantile regression to extreme tails in the conditional distribution

production function model.

There are two main challenges with implementing a quantile regression framework to the standard production function model. First, addressing the endogeneity of  $\omega_{it}$  using traditional panel data methods or control function approaches for quantile models is still a developing field that faces many challenges such as incidental parameters or non-smooth criterion functions. Second, it is not straightforward to link firms output decisions to variations in the ex-post shock  $\eta_{it}$  without imposing some strict structural model of firm production.

Addressing the first point, we will argue that recent quantile panel data and quantile IV models should not be used for the same reasons panel data and IV models are not used for the conditional mean model. Quantile panel data models allow for flexible interactions between unobserved heterogeneity and the quantiles of the conditional response function. Some well known approaches assume a time-invariant fixed effect such as [Koenker \(2004\)](#), [Lamarche \(2010\)](#), or [Canay \(2011\)](#) which assumes the fixed effect shifts only the location of the conditional quantile function. This approach has two disadvantages. First, assuming the unobservable is time-invariant is restrictive. [Griliches and Hausman \(1986\)](#) show that a fixed effect for productivity,  $\omega_i$ , leads to low estimates of the capital elasticity. Second, with the exception of the [Canay \(2011\)](#) estimator, the fixed effects of these models are incidental parameters so as the sample size grows, so does the number of parameters that need to be estimated which makes it computationally costly. An alternative to fixed effect estimation is to model the unobserved heterogeneity as a projection onto the observables plus a disturbance in the spirit of [Chamberlain \(1984\)](#). This correlated random effect (CRE) approach was used by [Abrevaya and Dahl \(2008\)](#) and variations of this type of estimator have been developed by other researchers. One issue with the CRE approach is that identification of the conditional quantile function is difficult because it now depends on the joint distribution of unobservables in the response function and the random effect.

Another alternative is to make use of valid instruments if they are available. The conventional argument for using input prices  $p_{it}^k$  and  $p_{it}^l$  as instruments is that they must be uncorrelated with the error term  $\omega_{it} + \eta_{it}$  and correlated with input choices for capital and labor. Then one could use two-stage least squares to obtain consistent estimates of  $\beta_k$  and  $\beta_l$ . This idea can be extended to quantile-IV models such as [Chernozhukov and Hansen \(2005\)](#). In their identification arguments, one would need to strengthen assumptions to conditional independence as well as monotonicity of a quantile structural function (QSF) in  $U_{it} = \omega_{it} + \eta_{it}$ . Then if one writes the QSF for the production function as  $y_{it} = Q(k_{it}, l_{it}, U_{it})$ , then the model is identified using

$$P[y_{it} \leq Q(k_{it}, l_{it}, \tau) | k_{it}, l_{it}, p_{it}^k, p_{it}^l] = \tau, \quad \tau \in (0, 1]. \quad (9)$$

We do not use this procedure since input prices may not have enough variation across firms and exogeneity can be violated if they capture input quality differences as shown by [Griliches and Hausman \(1986\)](#). Therefore, it is not straightforward to develop an identification strategy that

also leads to a computationally simple estimator. This is a common challenge in the quantile IV literature. However, we will show that a simple adaption of the [Canay \(2011\)](#) estimator works well in practice.

### 3 A Random Coefficient Production Function

We specify a *value-added* production function as a random coefficient model:

$$y_{it} = (\beta_0 + \omega_{it}) + \beta_k(\eta_{it})k_{it} + \beta_l(\eta_{it})l_{it} \quad (10)$$

The variables in equation (10) have the same interpretation as the ones we introduced in the LP model. In this specification we allow the output elasticities to be functionally dependent on the production shock  $\eta_{it}$  while productivity  $\omega_{it}$  maintains its additive separability. The model presented here assumes that  $(\beta_0 + \omega_{it})$  is only a location-shift of the conditional output distribution. We assume the constant in the production function does not vary over  $\tau$  since the constant is not separately identified from the mean of productivity. The location-shift specification is crucial in our approach since  $(\beta_0 + \omega_{it})$  also appears in the conditional mean counterpart of equation (10) so that if this object can be consistently estimated from popular proxy variable approaches such as OP, LP or ACF, then with additional restrictions from the quantile IV literature, the production function parameters  $\beta_k(\tau)$  and  $\beta_l(\tau)$  can also be consistently estimated.

We should also mention that a value-added specification in equation (10) is non-trivial. Value-added production functions are common in the empirical literature, however the objects recovered from a value-added model such as the output elasticities and TFP can only be mapped to its gross-output counterpart under special structural production functions such as Leontief value-added. Since the production shock  $\eta_{it}$  enters (10) non-separably, it is difficult to recover gross-output objects from value-added since the latter is composed of the additional error term. One consequence of a value-added model is that estimates of productivity may be more disperse than gross-output estimates which controls for material inputs. Therefore, in our empirical application we interpret the elasticities and productivity with some caution.

One advantage of the value-added model is that the rank of  $\eta_{it}$  can be interpreted as the rank of firm size measured by output net of material inputs. This way, we are able to measure firm size by the value of the firm's contribution to output rather than the value of the units they sell. The value-added approach also avoids the non-identification results of [Gandhi \*et al.\* \(2020\)](#). We leave the connection between this value-added production function and its possibly underlying gross-output production function as future research agenda.

An interesting question that follows from the firm size interpretation is then what are the determinants of firm size? Controlling for productivity also controls for firm size determinants such as managerial ability so variation in firm size might be limited in our model. There is a large

and historical literature on the determinants of firm size. [Kumar \*et al.\* \(1999\)](#) survey some of the theories and test them using data. We offer a possible explanation below, but we focus on whether there is heterogeneity in firm size and its consequences.

A special case of (10) is the location-scale model,

$$y_{it} = (\beta_0 + \omega_{it}) + \beta_k k_{it} + \beta_l l_{it} + (\mu_k k_{it} + \mu_l l_{it}) \eta_{it} \quad (11)$$

Which implies that the  $\tau$ -th conditional quantile of  $y_{it}$  is given by

$$Q_{y_{it}}(\tau | \mathcal{I}_{it}) = (\beta_0 + \omega_{it}) + \beta_k k_{it} + \beta_l l_{it} + (\mu_k k_{it} + \mu_l l_{it}) F^{-1}(\tau) \quad (12)$$

where  $F^{-1}(\tau)$  is the quantile function of production shocks  $\eta_{it}$ .

The idea that input choices can impact firm's production beyond the conditional mean has important consequences for firm's attitude towards production risk. A volume of literature that originated in the late 1970's challenged the standard stochastic specifications of production functions ([Just and Pope, 1978, 1979](#)) by considering a specification that allowed firm's inputs to both increase or decrease the marginal variability of final output. These models are commonly applied to studying the agricultural industry where the variance on the yield of harvested crops could be increased by adverse weather or decreased by pesticide usage. Since manufacturing businesses tend to operate in a more controlled environment, risk is less prevalent in these industries so the conditional variance of  $\eta_{it}$  may be smaller.

The connection between production risk and heterogeneity in quantiles is interesting so we briefly mention that under quantile preferences, a firm who maximizes the  $\tau$  level of utility of profits could explain heterogeneity in the output distribution. Unlike risk-neutral firms, firms could have a utility function that is represented by preferences of the firm manager(s) who decides the optimal expenditure on inputs. Different managers may have different preferences for risk. A short list of papers have considered quantile utility maximization such as [Manski \(1988\)](#), [Rostek \(2009\)](#), [Chambers \(2007\)](#), and [Bhattacharya \(2009\)](#). Dynamic input choices such as investment are much more difficult to solve using the quantile utility framework and the reader can refer to [de Castro and Galvao \(2017\)](#) for a treatment of dynamic quantile utility models. As far as we know, the quantile utility framework has not been applied to firm decision problems and a more thorough treatment of such is outside the scope of this paper.

A general quantile model such as the one specified in (10) can be seen as an extension of the higher-order moment estimation of risk initiated by [Antle \(1983\)](#). However, it can also be seen purely as an econometric specification issue as we are unaware of any tests that could distinguish between higher order moment production risk and misspecification. We choose the latter interpretation for our model.

### 3.1 A Two-Step Estimator for Quantile Production Functions

Let  $\varepsilon_{it} = k_{it}[\beta_k(\eta_{it}) - \beta_k^\mu] + l_{it}[\beta_l(\eta_{it}) - \beta_l^\mu]$  where  $(\beta_k^\mu, \beta_l^\mu) = \beta^\mu = \mathbb{E}[\beta(\eta_{it})]$  is the mean of the random coefficient. The conditional mean version of the random coefficient production function can then be written as

$$y_{it} = (\beta_0 + \omega_{it}) + \beta_k^\mu k_{it} + \beta_l^\mu l_{it} + \varepsilon_{it}, \quad (13)$$

where  $\mathbb{E}[\varepsilon_{it} | \mathcal{I}_{it}] = 0$ . This identity is useful since we can now use a consistent estimator of  $(\beta_0 + \omega_{it})$  which can be plugged into the random coefficient production function in equation (10). For example, if one is willing to make the appropriate assumptions regarding the timing of labor decisions and the structure of the proxy variable mentioned earlier, then  $(\beta_0 + \omega_{it}) = \hat{\Phi}_t(k_{it}, m_{it}) - \hat{\beta}_k^\mu k_{it}$  in the LP procedure. The ACF estimator can likewise be implemented by using  $(\beta_0 + \omega_{it}) = \hat{\Phi}_t(k_{it}, l_{it}, m_{it}) - \hat{\beta}_k^\mu k_{it} - \hat{\beta}_l^\mu l_{it}$  as the labor elasticity is estimated in the second stage of their procedure.

Intuitively, the appropriate choice for estimation of  $(\beta_0 + \omega_{it})$  places an identification restriction on the location of the conditional distribution of  $S_t = \beta_k(\eta_{it})k_{it} + \beta_l(\eta_{it})l_{it}$ . Canay (2011) approaches identification more formally. He shows using a deconvolution type argument that the distribution of the random coefficient component of the response variable is identified up to a location. The location of the conditional distribution in his model is identified from regularity conditions in a panel data model with first-differences of an unobserved individual effect. It is not straightforward how to extend this idea to our model where the unobserved effect, productivity, varies over time and exhibits serial correlation. We are currently working on a paper that proposes a flexible identification strategy for  $S_t$  and where the location of the conditional distribution can be identified under alternative production function approaches. For example, one may be interested in utilizing the methodology of Gandhi *et al.* (2020) to estimate productivity using an average gross-output production function.

Once estimates of  $(\beta_0 + \omega_{it})$  are obtained, estimates of  $\beta_k(\tau)$  and  $\beta_l(\tau)$  can be obtained from a quantile regression of  $\hat{y}_{it} = y_{it} - (\beta_0 + \omega_{it})$  with a constant omitted in estimation. To summarize these steps:

1. Let  $\hat{\beta}_k^\mu$  and  $\hat{\beta}_l^\mu$  be consistent estimators of  $\beta_k^\mu$  and  $\beta_l^\mu$  from a value-added production function. For example, using the OP, LP, or ACF estimation strategy. Construct the estimator,  $(\beta_0 + \omega_{it})$  using these estimates
2. Let  $\hat{y}_{it} = y_{it} - (\beta_0 + \omega_{it})$  and define the two-step estimator of  $\beta_k(\tau)$  and  $\beta_l(\tau)$  as:

$$\hat{\beta}(\tau) = \underset{\beta \in \mathcal{B}}{\operatorname{argmin}} \mathbb{E}[\rho_\tau(\hat{y}_{it} - \beta_k k_{it} - \beta_l l_{it})],$$

where  $\rho_\tau(u) = u[\tau - \mathbb{I}\{u < 0\}]$  and  $\mathbb{I}\{\cdot\}$  denotes the indicator function.



## 4 Monte Carlo Experiments

We use a location-scale version of [Levinsohn and Petrin \(2003\)](#) and replicate [Akerberg \*et al.\* \(2015\)](#) simulations by sampling 500 datasets consisting of 1000 firms. We simulate optimal input choices for 100 time periods, using the last 10 periods for estimation.

$$y_{it} = (\beta_0 + \omega_{it}) + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + (\gamma_k k_{it} + \gamma_l l_{it}) \eta_{it} \quad (14)$$

with  $\beta_0 = 0$ ,  $\beta_k = 0.4$  and  $\beta_l = 0.6$ . The location scale parameters are set as  $\gamma_k = 0.7$  and  $\gamma_l = -0.6$ . For each simulation we simulate two DGPs with  $\eta_{it} \sim N(0, 0.1)$  and  $\eta_{it} \sim Laplace(0, 0.1)$ .

Following the data generating process in [Akerberg \*et al.\* \(2015\)](#), we do not allow for any wage variation across firms and labor is chosen at time  $t$  with perfect information about  $\omega_{it}$ . We add optimization error to labor decisions to break the functional dependence problem. An AR(1) process is specified for productivity  $\omega_{it} = \rho \omega_{it-1} + \xi_{it}$  where  $\rho = 0.7$ . The variance of  $\xi_{it}$  and initial value  $\omega_{i0}$  is set so that the standard deviation of  $\omega_{it}$  is constant over time and equal to 0.3. We compare the LP estimation procedure with the QLP two-step procedure under the two different sets of experiments specified earlier. We estimate the model for  $\tau \in \{0.05, 0.1, \dots, 0.9, 0.95\}$

Table 1 provides estimates of the bias and MSE for DGP 1 and DGP 2. They are both very small and decrease as the estimates approach  $\tau = 0.5$  in DGP 2. We plot the MSE of our estimator as well as the MSE of the LP estimator compared to the true values of  $\beta_k(\tau)$  and  $\beta_l(\tau)$  in Figure 1. The MSE for both estimators is plotted over  $\tau \in \{0.05, 0.1, \dots, 0.9, 0.95\}$  with the black line denoting the QLP estimates and the dotted red line denoting the LP estimate. It is clear from this plot that our estimator does relatively well at capturing the heterogeneity in our model. Lastly we test whether our estimates control for unobserved productivity by comparing our estimates to standard quantile regression estimates without controlling for productivity. Figure 2 plots estimates of these differences along with their 95% confidence intervals for each  $\tau$ . The plots show significant differences in these estimates with capital being underestimated and labor being overestimated in the presence of the simultaneity bias.

Table 1: Bias and MSE

DGP	$\tau$	Capital		Labor	
		Bias	MSE	Bias	MSE
1	0.10	-0.0007	0.0002	0.0029	0.0000
	0.25	0.0008	0.0002	0.0005	0.0000
	0.50	0.0010	0.0002	0.0000	0.0000
	0.75	0.0022	0.0002	-0.0005	0.0000
	0.90	0.0027	0.0002	-0.0019	0.0000
2	0.10	-0.0037	0.0002	0.0046	0.0001
	0.25	-0.0005	0.0002	0.0006	0.0000
	0.50	0.0000	0.0002	0.0000	0.0000
	0.75	0.0005	0.0002	-0.0016	0.0000
	0.90	0.0037	0.0002	-0.0036	0.0000

Figure 1: Simulated precision of QLP estimators of  $\beta_k(\tau)$  and  $\beta_l(\tau)$ s. Dotted line is LP estimator.

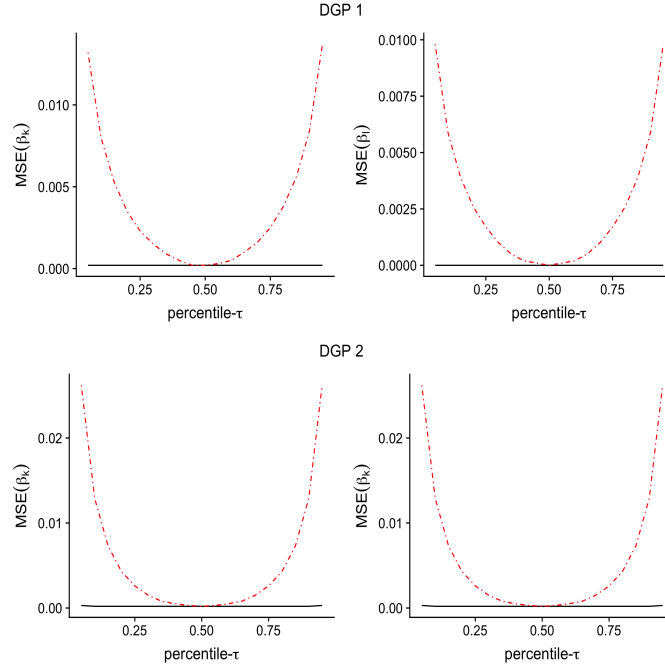
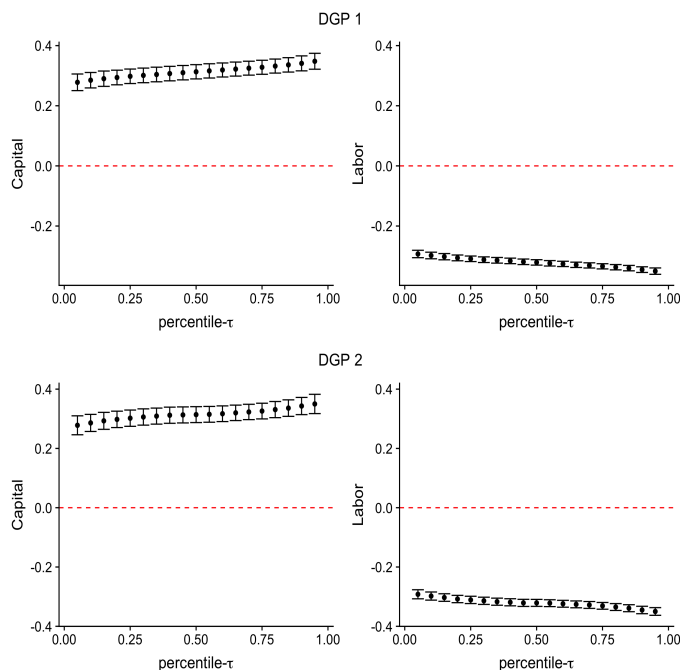


Figure 2: Difference between QLP estimators and QR estimates of  $\beta_k(\tau)$  and  $\beta_l(\tau)$ s.



## 5 Application

We apply our estimator to popular firm and plant level manufacturing datasets from the US, Chile, and Colombia to examine heterogeneity in the output distribution <sup>7</sup>. For each country we examine estimates across different manufacturing industries as well as how these estimates have changed over time. We use the QLP estimator presented in this paper and compare it to the LP estimates. We also compare our estimates to the quantile regression estimates without controlling for productivity. In the LP estimation procedure, we estimate the labor coefficient and  $\Phi$  with a 3rd degree polynomial with interactions in capital and materials. To estimate capital, we use the LP criterion function mentioned earlier. Since we only need a consistent estimator of the production function parameters, we do not consider any over-identification conditions in this step. We use bootstrap to estimate standard errors of  $\beta_k(\tau)$  and  $\beta_l(\tau)$  with the number of iterations set to 500.

### 5.1 US Compustat

The source for the US manufacturing data is from Compustat which covers publicly traded firms and contains data from their financial statements. We collect a sample between 1961 and 2010 on sales, capital expenditures, number of workers, and other expenses to construct measures of output, capital, labor, and material inputs using 3-digit deflators from [Bartelsman and Gray \(1996\)](#). Data

<sup>7</sup>We thank Mert Demirer for providing the datasets from Chile and Colombia

preparation follows [Keller and Yeaple \(2009\)](#) and [Dermirer \(2020\)](#). Some issues regarding the Compustat dataset is that since the data is reported in the firm’s financial statements, deflated output and input measures may not completely capture firm’s actual usage. Also, since this sample only contains publicly traded firms, it is only a fraction of all manufacturing firms in the US. Summary statistics for these deflated values are provided in Table 2. We present a series of output elasticity estimates in Table 3 which are illustrated graphically in Figures 3, 4, 5, and 6.

Estimates of the capital elasticity are slightly increasing in the firm size distribution in every industry as well as the combined sample. The estimates for labor elasticity for each industry and the combined sample are decreasing. For NAICS 33, the labor elasticity increases at very small  $\tau$ , but then decreases. In each industry, there is evidence that our model captures some heterogeneity compared to the LP model. In NAICS 31, the capital elasticity is smaller and the labor elasticity is larger than the LP estimates for very small  $\tau$ . In NAICS 32, there is more heterogeneity in the tails of the firm size distribution for capital estimates, but only the labor estimates are lower than the LP estimate for very large  $\tau$ . The same relationship in these estimates is true in NAICS 33 and in the combined sample.

Table 2: Summary Statistics (in logs) for the US Manufacturing Data

Industry (NAICS code)		1st Qu.	Median	3rd Qu.	Mean	sd
31 (Total=3271)	Output	19.05	20.24	21.57	20.3	1.77
	Capital	18.66	20.37	21.76	20.19	2.12
	Labor	17.42	19.08	20.61	19.02	2.21
	Materials	17.96	19.59	21.15	19.54	2.21
32 (Total=7207)	Output	15.67	17.04	18.51	17.01	2.05
	Capital	15.65	17.51	19.13	17.31	2.41
	Labor	14.44	16.01	17.57	16.01	2.29
	Materials	14.89	16.53	18.25	16.52	2.37
33 (Total=13978)	Output	7.38	8.58	9.8	8.5	1.67
	Capital	6.67	8.29	9.74	8.15	1.95
	Labor	6.01	7.42	8.91	7.48	1.93
	Materials	6.33	7.82	9.29	7.82	1.95
All (Total=24456)	Output	18.58	19.78	21.23	19.85	1.79
	Capital	18.14	19.86	21.26	19.67	2.16
	Labor	16.98	18.59	20.13	18.56	2.17
	Materials	17.49	19.12	20.66	19.06	2.2

### NAICS 31

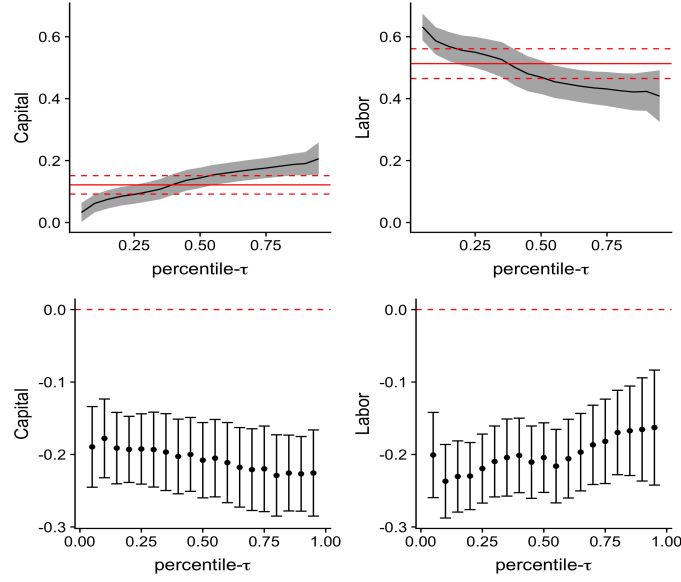


Figure 3: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

### NAICS 32

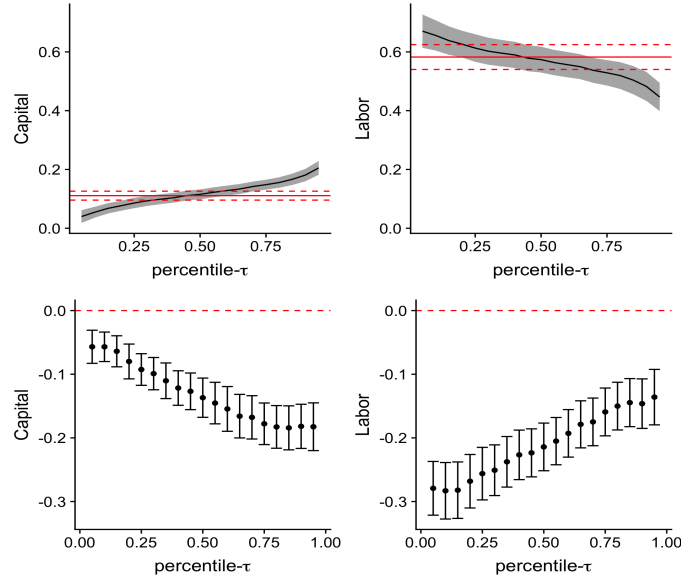


Figure 4: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

# NAICS 33

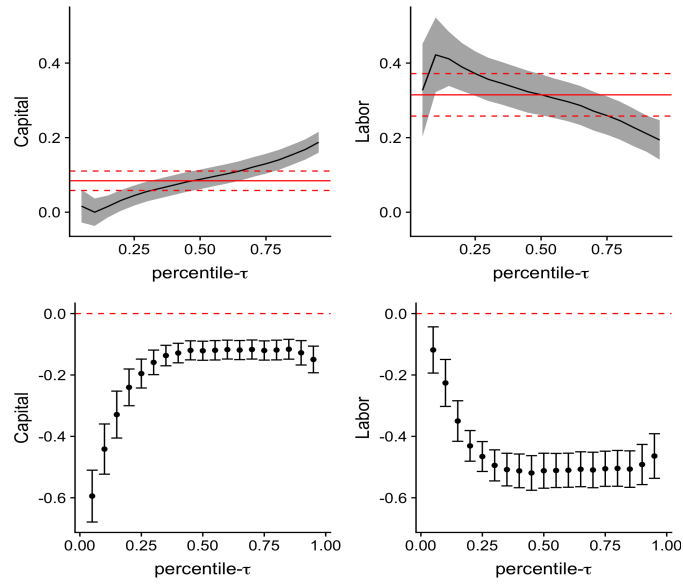


Figure 5: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

## NAICS All

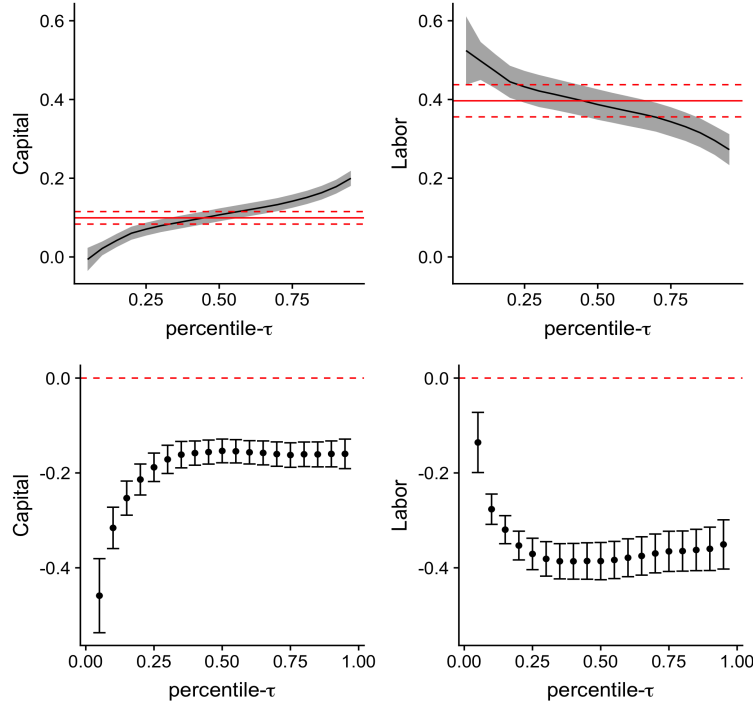


Figure 6: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

In each industry we compare the differences between QLP and QR estimates to test whether our model corrects for endogeneity from unobserved productivity. Bootstrap is used to construct confidence intervals of the difference between the two estimates. We find that there are significant differences between these estimates in all of the samples. This suggests that even after controlling for productivity differences across firms, there is still unaccounted heterogeneity between firms.

We use the estimates from the output elasticities to construct measures of returns to scale and capital intensity in Table 3. The results for returns to scale are puzzling as they are all significantly different from constant returns to scale. Each sample and each quantile of firm size exhibits large decreasing returns to scale and these returns decrease when  $\tau$  increases. Previous papers that estimate returns to scale using the Compustat dataset such as [Keller and Yeaple \(2009\)](#) and [Dermirer \(2020\)](#) show constant returns to scale using a gross-output production function. Therefore it is possible that the empirical value-added (deflated sales minus intermediate input expenditure) is a poor proxy for value-added in our model and that value-added biases the returns to scale estimates. Differences in returns to scale in value-added and gross-output production functions are explored by [Basu and Fernald \(1997\)](#). We also report estimates of capital intensity measured by

the ratio of capital to labor elasticity for each quantile. Each industry has estimates of capital intensity that increases with  $\tau$ . This result is consistent with previous findings such as [Holmes and Mitchell \(2008\)](#), [Kumar \*et al.\* \(1999\)](#) and [Dermirer \(2020\)](#).

Table 3: Coefficient Estimates and Standard Errors for US Manufacturing Firms

NAICS	$\tau$	Capital		Labor		Returns to Scale		Capital Intensity	
		Coef.	s.e	Coef.	s.e	Coef.	s.e	Coef.	s.e
31	0.10	0.062	0.0176	0.587	0.0269	0.649	0.0318	0.105	0.0305
	0.25	0.091	0.0184	0.550	0.0303	0.641	0.0329	0.165	0.0366
	0.50	0.144	0.0201	0.469	0.0329	0.614	0.0333	0.307	0.0522
	0.90	0.190	0.0226	0.423	0.0382	0.614	0.0346	0.450	0.0879
32	0.10	0.054	0.0113	0.656	0.0313	0.710	0.0280	0.083	0.0190
	0.25	0.086	0.0107	0.613	0.0285	0.699	0.0266	0.140	0.0203
	0.50	0.115	0.0101	0.574	0.0267	0.689	0.0259	0.201	0.0217
	0.90	0.181	0.0129	0.482	0.0300	0.663	0.0265	0.376	0.0428
33	0.10	0.000	0.0224	0.422	0.0609	0.422	0.0459	0.000	0.0709
	0.25	0.045	0.0163	0.372	0.0363	0.417	0.0323	0.120	0.0591
	0.50	0.088	0.0160	0.315	0.0332	0.403	0.0305	0.279	0.0816
	0.90	0.169	0.0165	0.211	0.0296	0.379	0.0278	0.800	0.1931
All	0.10	0.022	0.0107	0.498	0.0293	0.519	0.0259	0.043	0.0245
	0.25	0.071	0.0100	0.432	0.0246	0.503	0.0234	0.163	0.0300
	0.50	0.107	0.0100	0.387	0.0235	0.494	0.0228	0.276	0.0379
	0.90	0.179	0.0107	0.296	0.0222	0.474	0.0214	0.604	0.0734

We also use our quantile production function estimates to construct measures of firm level productivity which we define as

$$\hat{w}_{it,\tau} = \exp(y_{it} - \hat{\beta}_k(\tau)k_{it} - \hat{\beta}_l(\tau)l_{it}) \quad (15)$$

We use these measures to compare productivity growth over time to LP estimates over the distribution of firm size. Figure 7 reports average productivity for all US firms in the sample with the base year of the sample period set to 100. We can see that productivity growth was rapid in the beginning of the sample period but then declined after 1970, but then increases after 1980. Growth trends for each percentile of firm size were similar although larger firms in this sample were more productive than smaller ones. The LP estimates are close to the productivity estimates for medium-sized firms at  $\tau = 0.5$ .

We are also interested in examining the firm size distribution over time and whether there are differences in within-firm and across-firm technology heterogeneity. Figure 8 plots estimates of the output elasticities over 5 year intervals. For labor elasticity, there is heterogeneity across the firm size distribution in the beginning of the sample period. Larger firms had greater estimates of labor elasticity than small firms. This heterogeneity decreases up until 1980 when the relationship



between firm size and labor elasticity reverses. At the end of our sample period, very large firms have smaller estimates of labor elasticity than very small firms. Large firms have higher estimates of capital elasticity than small and medium-sized firms. These estimates are increasing until the year 2000, but then decreases. The relationship between firm size and capital estimates stays the same throughout the sample period.

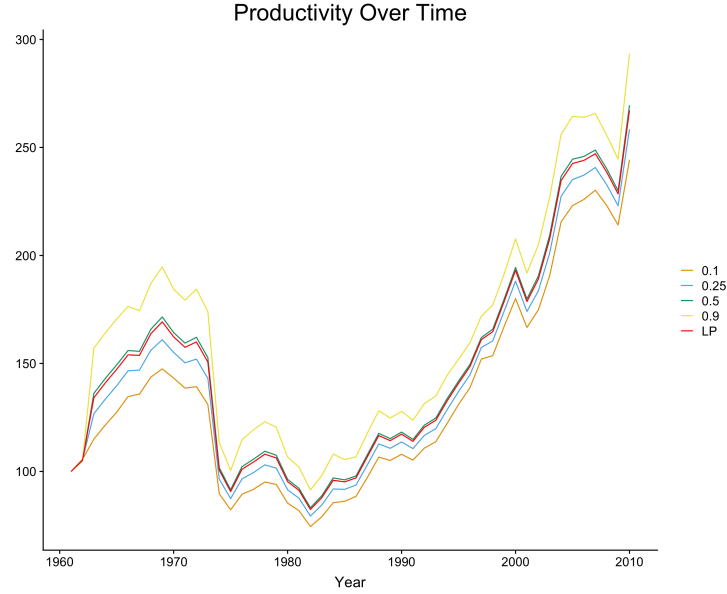


Figure 7: Estimated average productivity over time for the US. Base productivity in 1961 is set to 100.

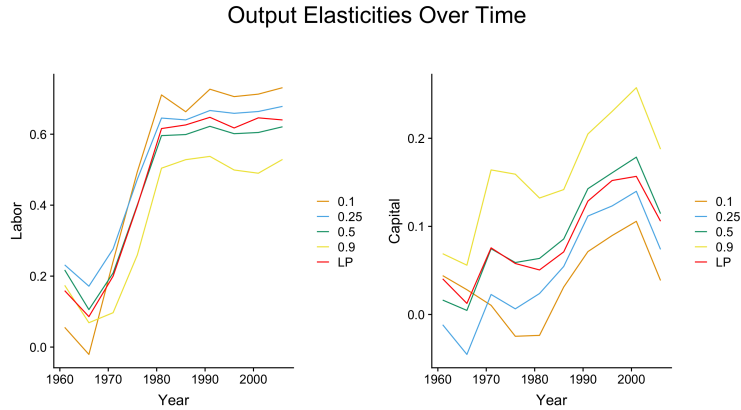


Figure 8: Estimated values of production function coefficients over time estimated at 5 year intervals

## 5.2 Chilean Manufacturing

This data comes from the census of Chilean manufacturing plants conducted by the Instituto Nacional de Estadística (INE). The sample is collected between 1979 and 1996 for firms with more than 10 employees. We divide our estimates into the three largest manufacturing industries: Food (ISIC 311), Fabricated Metals (ISIC 381), and Textiles (ISIC 321). We also aggregate the three industries with the other smaller industries to obtain estimates from the entire sample. Summary statistics for the data we use are provided in Table 4.

Figures 9, 10, and 11 illustrate our estimates from our model compared to LP estimates as well as their differences to QR estimates. In each industry and the combined sample, estimates of capital elasticity are increasing in firm size. In ISIC 311 and ISIC 381, our model estimates lower capital elasticities than the LP estimate for small  $\tau$ . In ISIC 321 and the combined sample, there are differences between these estimates for both small and large  $\tau$ . The labor estimates in ISIC 311 are increasing and different than the LP estimate for large  $\tau$ . In ISIC 321, labor estimates are decreasing and not different from the LP estimate. In ISIC 381, labor estimates are mostly constant and not different from the LP estimate. Finally, in the combined sample, our estimates exhibit a U-shaped relationship although overall these are not different from the LP estimate. In ISIC 321 our capital estimates do not differ much from QR estimates. From Table 5, our estimates of returns to scale are more reasonable compared to the US estimates except for small firms who experience decreasing returns to scale. In each industry, returns to scale are increasing with firm size. Capital intensity is increasing as well except for large firms at  $\tau = 0.9$  in ISIC 311.

Figure 13 reports average productivity for all Chilean plants in the sample with base period set to 100. Productivity decreases in the beginning of the 1980s but then increases for the rest of the sample period. These estimates suggest that small firms had higher productivity than large firms. The LP estimates are similar to productivity of medium-sized firms at  $\tau = 0.5$ . Figure 14 shows the time trends in output elasticities. The estimate of labor elasticity are high for each quantile of firm size and decreases steadily with the exception of small firms ( $\tau = 0.1$ ). The estimates for capital elasticities follow similar trends for each quantile. Capital elasticity estimates are higher for large firms than small firms.

Table 4: Summary Statistics (in logs) for Chile Manufacturing Data

Industry (ISIC code)		1st Qu.	Median	3rd Qu.	Mean	sd
311 (Total=13838)	Output	10.21	10.84	12.22	11.36	1.58
	Capital	10.56	11.4	12.4	11.52	1.37
	Labor	10.49	11.4	12.54	11.53	1.43
	Materials	10.38	11.28	12.53	11.56	1.6
381 (Total=4311)	Output	6.69	7.66	9.06	8.02	1.98
	Capital	7.52	8.51	9.7	8.65	1.68
	Labor	7.21	8.34	9.56	8.4	1.72
	Materials	7.22	8.35	9.72	8.54	1.92
321 (Total=4302)	Output	2.77	3.22	3.91	3.49	0.99
	Capital	2.89	3.47	4.22	3.71	1.08
	Labor	2.94	3.48	4.37	3.69	0.95
	Materials	2.89	3.43	4.28	3.67	1.02
All (Total=51567)	Output	9.84	10.46	11.81	10.94	1.56
	Capital	9.91	10.75	11.79	10.86	1.41
	Labor	9.68	10.62	11.75	10.73	1.48
	Materials	9.81	10.68	11.89	10.93	1.62

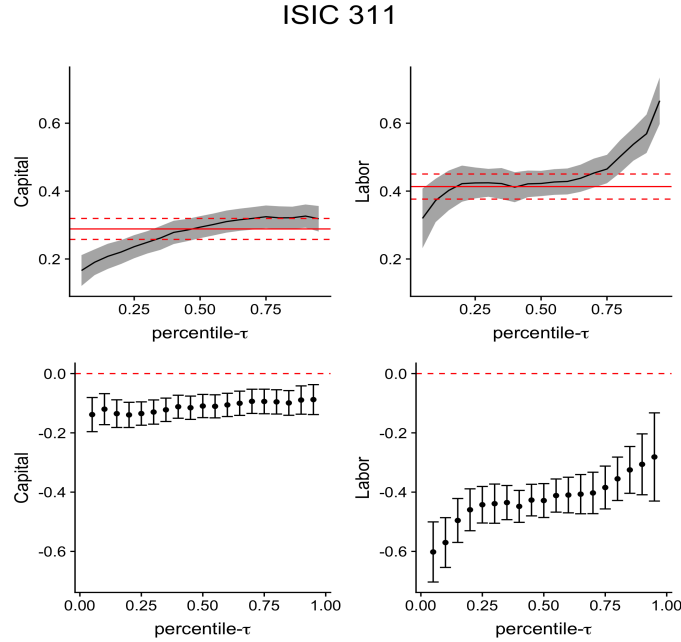


Figure 9: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

### ISIC 321

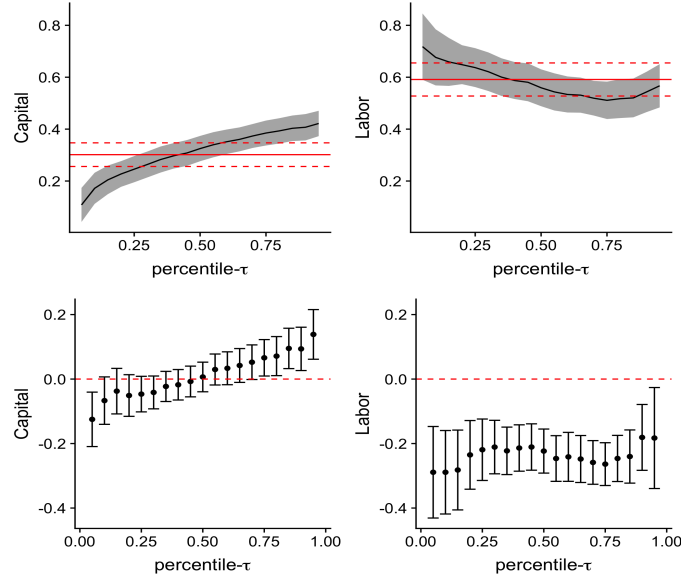


Figure 10: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

### ISIC 381

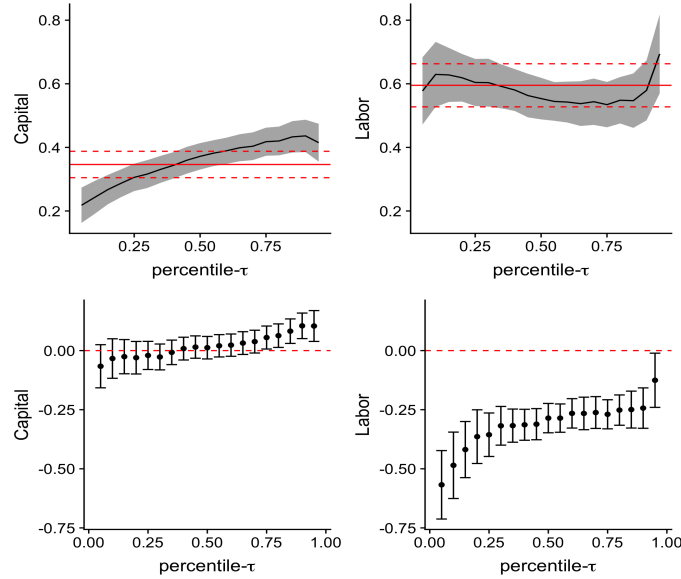


Figure 11: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

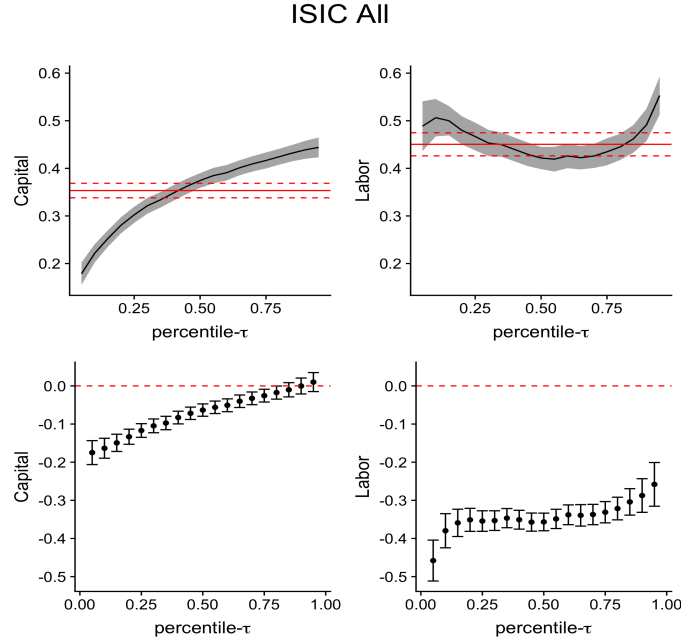


Figure 12: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

Table 5: Coefficient Estimates and Standard Errors for Chilean Manufacturing Firms

ISIC	$\tau$	Capital		Labor		Returns to Scale		Capital Intensity	
		Coef.	s.e	Coef.	s.e	Coef.	s.e	Coef.	s.e
311	0.10	0.190	0.0230	0.373	0.0388	0.563	0.0314	0.511	0.1097
	0.25	0.236	0.0207	0.424	0.0277	0.660	0.0266	0.557	0.0747
	0.50	0.294	0.0197	0.422	0.0227	0.716	0.0245	0.695	0.0714
	0.90	0.326	0.0208	0.569	0.0344	0.895	0.0305	0.574	0.0622
381	0.10	0.243	0.0315	0.630	0.0621	0.872	0.0457	0.386	0.0848
	0.25	0.306	0.0262	0.604	0.0450	0.910	0.0380	0.506	0.0725
	0.50	0.372	0.0253	0.553	0.0398	0.925	0.0350	0.672	0.0826
	0.90	0.436	0.0309	0.579	0.0573	1.015	0.0422	0.753	0.1175
321	0.10	0.172	0.0364	0.676	0.0657	0.848	0.0470	0.254	0.0749
	0.25	0.246	0.0306	0.637	0.0456	0.883	0.0381	0.386	0.0694
	0.50	0.325	0.0296	0.559	0.0432	0.884	0.0371	0.581	0.0874
	0.90	0.407	0.0308	0.543	0.0472	0.950	0.0390	0.750	0.1168
All	0.10	0.222	0.0117	0.506	0.0240	0.728	0.0176	0.437	0.0407
	0.25	0.302	0.0103	0.467	0.0177	0.769	0.0141	0.648	0.0415
	0.50	0.374	0.0094	0.422	0.0142	0.796	0.0123	0.888	0.0449
	0.90	0.439	0.0115	0.491	0.0206	0.930	0.0148	0.893	0.0572

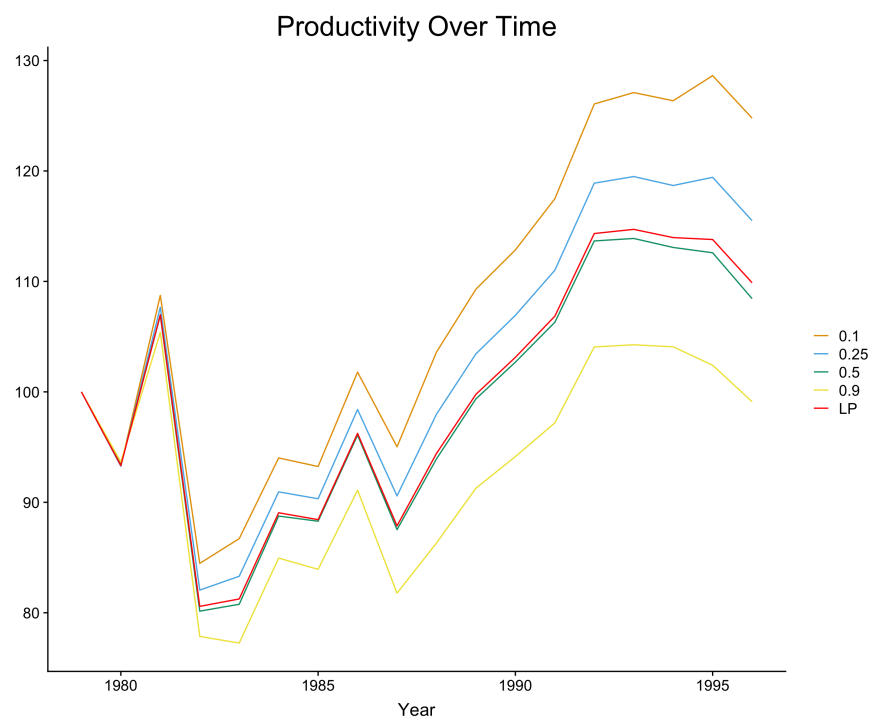


Figure 13: Estimated average productivity over time for Chile. Base productivity in 1979 is set to 100.

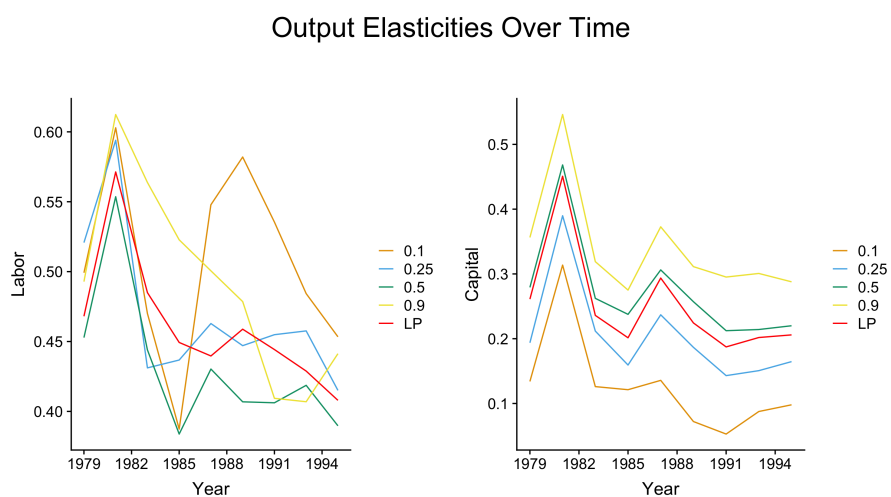


Figure 14: Estimated values of production function coefficients over time estimated at 2 year intervals

### 5.3 Colombian Manufacturing

This data comes from the Colombian manufacturing census conducted by the Departamento Administrativo Nacional de Estadística. The sample is collected between 1977 and 1991. We divide our estimates into the three largest manufacturing industries: Food (ISIC 311), Apparel (ISIC 322), and Fabricated Metals (ISIC 381). As we did with the Chilean sample, we also aggregate the three industries with other smaller industries to obtain estimates from the entire sample of manufacturing plants. Summary statistics for this data is provided in Table 6.

Table 6: Summary Statistics (in logs) for Colombia Manufacturing Data

Industry (ISIC code)		1st Qu.	Median	3rd Qu.	Mean	sd
311 (Total=13215)	Output	9.03	10.21	11.59	10.42	1.8
	Capital	8.69	9.37	10.22	9.49	1.18
	Labor	8.52	9.3	10.33	9.54	1.43
	Materials	8.7	9.62	10.88	9.92	1.67
322 (Total=12182)	Output	6.02	7.07	8.35	7.24	1.78
	Capital	5.47	6.14	6.93	6.23	1.21
	Labor	5.89	6.75	7.81	6.93	1.55
	Materials	5.9	6.89	8.16	7.12	1.77
381 (Total=7411)	Output	2.56	3.09	3.97	3.36	1.1
	Capital	2.77	3.3	3.95	3.42	0.92
	Labor	2.64	3.18	3.91	3.37	0.98
	Materials	2.71	3.3	4.11	3.5	1.09
All (Total=87783)	Output	8.39	9.73	11.26	9.87	2
	Capital	7.62	8.53	9.46	8.48	1.51
	Labor	7.77	8.65	9.72	8.8	1.58
	Materials	7.89	8.93	10.26	9.15	1.88

Figures 15, 16, and 17 illustrate estimates from our model compared to the LP estimates as well as their differences from QR estimates. The capital estimates in each industry are increasing a significantly different than the LP estimates. The labor estimates in each industry are decreasing. In ISIC 311, 322, and the combined sample, these estimates are significantly different from LP estimates. In ISIC 381, there are only differences in the tails of the firm size distribution. In this industry, there is also not much difference between our estimates and the QR estimates.

### ISIC 311

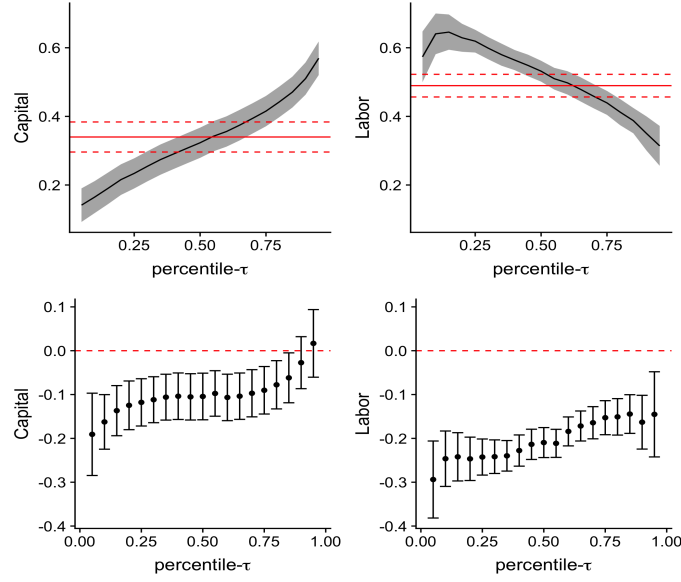


Figure 15: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

### ISIC 322

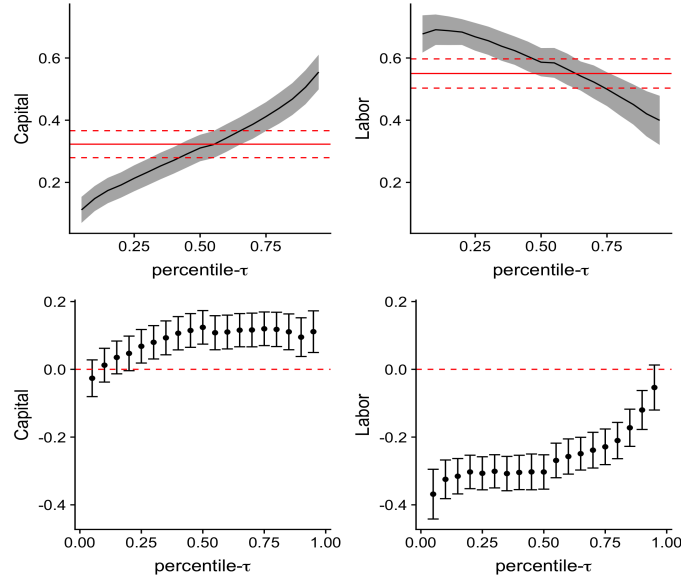


Figure 16: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.



### ISIC 381

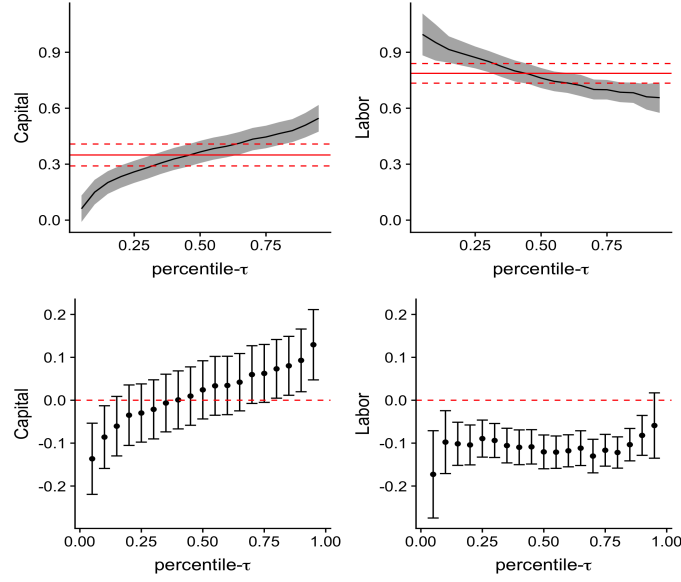


Figure 17: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

### ISIC All

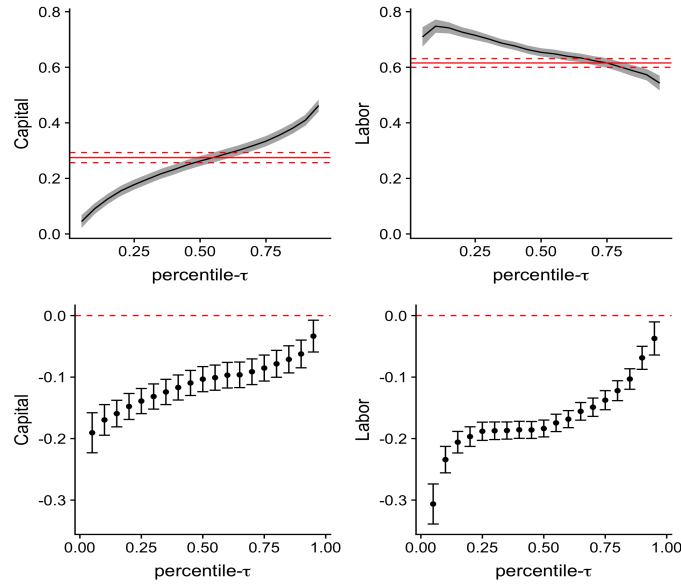


Figure 18: Top row: Estimated values of production function coefficients and their point-wise 90% confidence interval. Bottom row: Difference between QLP and quantile regression estimates and their 95% confidence intervals.

Using these estimates we construct measures of returns to scale and capital intensity for each industry in Table 7. Most firms experience constant or slightly decreasing returns to scale. Returns to scale and capital intensity are increasing in firm size.

Table 7: Coefficient Estimates and Standard Errors for Colombian Manufacturing Firms

ISIC	$\tau$	Capital		Labor		Returns to Scale		Capital Intensity	
		Coef.	s.e	Coef.	s.e	Coef.	s.e	Coef.	s.e
311	0.10	0.165	0.0283	0.641	0.0361	0.805	0.0361	0.257	0.0530
	0.25	0.234	0.0267	0.619	0.0202	0.853	0.0300	0.378	0.0476
	0.50	0.323	0.0268	0.532	0.0185	0.855	0.0293	0.608	0.0582
	0.90	0.510	0.0286	0.350	0.0302	0.860	0.0326	1.457	0.1759
322	0.10	0.148	0.0246	0.691	0.0298	0.839	0.0310	0.214	0.0394
	0.25	0.213	0.0253	0.668	0.0272	0.881	0.0290	0.319	0.0434
	0.50	0.311	0.0259	0.586	0.0277	0.897	0.0287	0.530	0.0562
	0.90	0.506	0.0332	0.421	0.0442	0.927	0.0320	1.203	0.1780
381	0.10	0.150	0.0402	0.952	0.0588	1.102	0.0529	0.158	0.0471
	0.25	0.259	0.0360	0.872	0.0356	1.131	0.0450	0.297	0.0446
	0.50	0.366	0.0362	0.762	0.0327	1.127	0.0434	0.480	0.0538
	0.90	0.509	0.0398	0.661	0.0417	1.171	0.0436	0.770	0.0861
All	0.10	0.091	0.0118	0.748	0.0146	0.839	0.0136	0.122	0.0173
	0.25	0.178	0.0110	0.715	0.0104	0.893	0.0121	0.248	0.0172
	0.50	0.263	0.0109	0.654	0.0090	0.916	0.0116	0.402	0.0194
	0.90	0.410	0.0120	0.573	0.0125	0.983	0.0122	0.715	0.0321

Figure 19 reports average productivity for all Colombian plants in the sample with base period set to 100. Productivity decreases in the beginning of the sample period but then increases for the rest of the sample period after 1980 with some sharp periods of productivity decline and incline. Each percentile of firm size has similar productivity levels at the beginning of the sample period, but slightly diverges after 1984. The LP estimates show just smaller than productivity of other estimates after this time period. Figure 20 shows the time trends in output elasticities. The estimates of labor elasticity are about 0.6 for each quantile of firm size and increases steadily until about 1981 then starts to decrease. At the end of the sample period there is more heterogeneity in these estimates. Capital estimates exhibit a clear distinction across quantiles with the larger firms having higher estimates than smaller firms. These estimates are mostly increasing until 1985 when they begin to decrease.

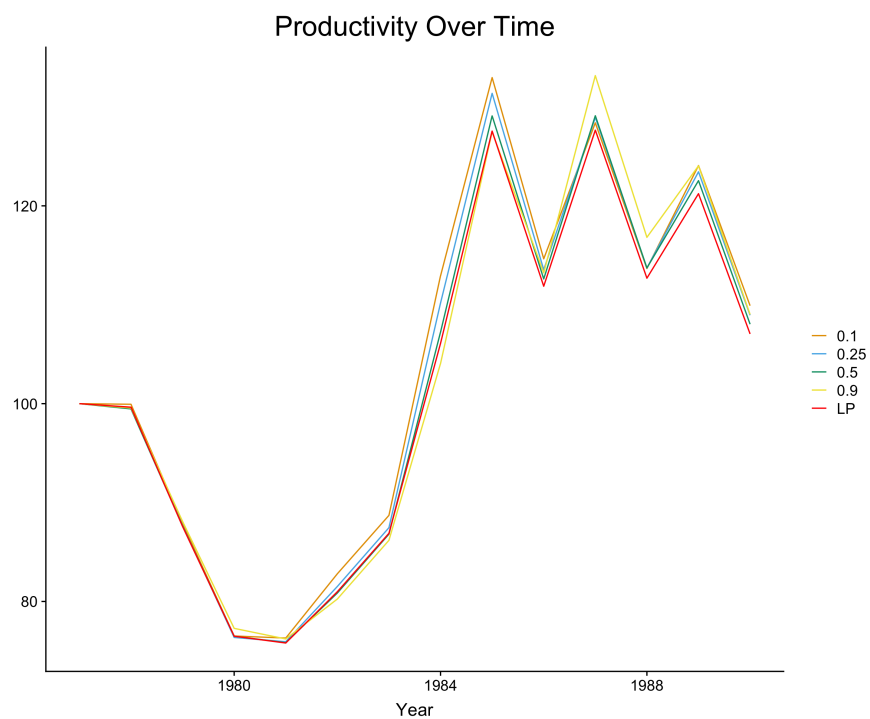


Figure 19: Estimated average productivity over time for Colombia. Base productivity in 1978 is set to 100.

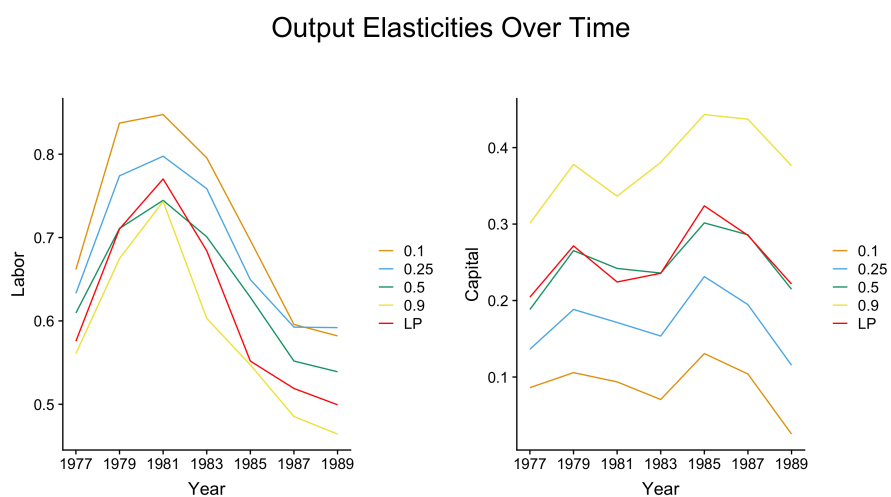


Figure 20: Estimated values of production function coefficients over time estimated at 2 year intervals

## 6 Conclusions

We propose a method that extends the proxy variable approach to estimating the conditional quantiles of firm size. The method is computationally attractive as it resembles the two-stage estimator proposed by [Canay \(2011\)](#). As a result, practitioners are able to easily apply the proposed estimator to production function models where the data may reveal significant heterogeneous output elasticities along the conditional firm size distribution. We showed that this estimator works well in finite samples and showed that it captures heterogeneity in firm size under different data generating processes. An application to widely used datasets from the US, Chile, and Colombia showed that in some industries, our estimator captures unobserved heterogeneity that the LP estimator does not.

Improvements and extensions of this estimator are currently being explored. For example, using a value-added production function may show more heterogeneity in estimates of elasticities and productivity than a gross-output production function. However, using a gross-output production function with an intermediate input proxy variable suffers from non-identification. We are currently working on an identification strategy that allows for different types of production function estimators such as the one proposed by [Gandhi \*et al.\* \(2020\)](#). This paper also makes an interesting but brief connection between the literature on production risk and quantile utility maximization. Currently, quantile utility maximization problems and estimation of these models are being studied by [de Castro and Galvao \(2017\)](#) and [de Castro \*et al.\* \(2018\)](#) in the context of dynamic consumption problems. It would be interesting to explore a model for a firm who maximizes quantile utility of profits which could provide an alternative interpretation for unobserved heterogeneity that are obtained from quantile regression estimates.

This paper contributes to the growing literature on production functions with unobserved heterogeneity. We show that differences in firm size correspond to the the rank of the ex-post shock. The location-shift model for productivity we propose here restricts us from examining other dimensions of firm heterogeneity. Therefore, allowing richer distributional effects of productivity would be an interesting extension. This approach also restricts us from examining non-Hicks neutral productivity shocks such as factor-augmenting productivity. We are currently working on an extension of this paper to a non-separable model to address these last two points, but the estimator we propose here is computationally attractive and easy to implement in empirical research.

## References

- ABREVAYA, J. and DAHL, C. M. (2008). The effects of birth inputs on birthweight. *Journal of Business & Economic Statistics*, **26** (4), 379–397.
- ACKERBERG, D., CAVES, K. and FRAZER, G. (2015). Identification properties of recent production function estimators. *Econometrica*, **83** (6), 2411–2451.
- AIGNER, D., LOVELL, C. and SCHMIDT, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, **6** (1), 21–37.
- ANTLE, J. M. (1983). Testing the stochastic structure of production: A flexible moment-based approach. *Journal of Business & Economic Statistics*, **1** (3), 192–201.
- ARAGON, Y., DAOUIA, A. and THOMAS-AGNAN, C. (2005). Nonparametric frontier estimation: A conditional quantile based approach. *Econometric Theory*, **21** (02).
- BALAT, J., BRAMBILLA, I. and SASAKI, Y. (2018). Heterogeneous firms: skilled-labor productivity and the destination of exports, Working paper.
- BARTELSMAN and GRAY (1996). *The NBER Manufacturing Productivity Database*. Tech. Rep. 205, National Bureau of Economic Research.
- BASU, S. and FERNALD, J. G. (1997). Returns to scale in u.s. production: Estimates and implications. *Journal of Political Economy*, **105** (2), 249–283.
- BERNINI, C., FREO, M. and GARDINI, A. (2004). Quantile estimation of frontier production function. *Empirical Economics*, **29** (2), 373–381.
- BHATTACHARYA, D. (2009). Inferring optimal peer assignment from experimental data. *Journal of the American Statistical Association*, **104** (486), 486–500.
- CANAY, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, **14** (3), 368–386.
- CHAMBERLAIN, G. (1984). Panel data. In: Griliches Z, Intriligator MD (eds) *Handbook of econometrics*. vol. 2, Elsevier Science B.V., New York, pp. 1247–1318.
- CHAMBERS, C. P. (2007). Ordinal aggregation and quantiles. *Journal of Economic Theory*, **137** (1), 416–431.
- CHERNOZHUKOV, V. (2005). Extremal quantile regression. *The Annals of Statistics*, **33** (2), 806–839.
- and HANSEN, C. (2005). An IV model of quantile treatment effects. *Econometrica*, **73** (1), 245–261.

- CHESHER, A. (2003). Identification in nonseparable models. *Econometrica*, **71** (5), 1405–1441.
- DE CASTRO, L., GALVAO, A. F., KAPLAN, D. M. and LIU, X. (2018). Smoothed GMM for quantile models. *Journal of Econometrics*, forthcoming.
- DE CASTRO, L. I. and GALVAO, A. F. (2017). Dynamic quantile models of rational behavior. *SSRN Electronic Journal*.
- DERMIRER, M. (2020). Production function estimation with factor augmenting technology: An application to markups.
- GANDHI, A., NAVARRO, S. and RIVERS, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, pp. 000–000.
- GRILICHES, Z. and HAUSMAN, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, **31** (1), 93–118.
- HOLMES, T. J. and MITCHELL, M. F. (2008). A theory of factor allocation and plant size. *The RAND Journal of Economics*, **39** (2), 329–351.
- IMBENS and NEWEY (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica*, **77** (5), 1481–1512.
- JUST, R. E. and POPE, R. D. (1978). Stochastic specification of production functions and economic implications. *Journal of Econometrics*, **7** (1), 67–86.
- and — (1979). Production function estimation and related risk considerations. *American Journal of Agricultural Economics*, **61** (2), 276–284.
- KASAHARA, H., SCHRIMPF, P. and SUZUKI, M. (2017). Identification and estimation of production function with unobserved heterogeneity, Working paper.
- KELLER, W. and YEAPLE, S. R. (2009). Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the united states. *Review of Economics and Statistics*, **91** (4), 821–831.
- KOENKER, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, **91** (1), 74–89.
- KUMAR, K., RAJAN, R. and ZINGALES, L. (1999). *What Determines Firm Size?* Tech. rep.
- LAMARCHE, C. (2010). Robust penalized quantile regression estimation for panel data. *Journal of Econometrics*, **157** (2), 396–408.
- LEE, S. (2007). Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics*, **141** (2), 1131–1158.

- LEVINSOHN, J. and PETRIN, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, **70** (2), 317–341.
- LI, T. and SASAKI, Y. (2017). Constructive identification of heterogeneous elasticities in the cobb-douglas production function, Working paper.
- MA, L. and KOENKER, R. (2006). Quantile regression methods for recursive structural equation models. *Journal of Econometrics*, **134** (2), 471–506.
- MANSKI, C. F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision*, **25** (1), 79–104.
- OLLEY, G. S. and PAKES, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, **64** (6), 1263.
- ROSTEK, M. (2009). Quantile maximization in decision theory. *The Review of Economic Studies*, **77** (1), 339–371.