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

Justin Doty[*] and Suyong Song[†]

November 19, 2020

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 identified by conditional quantile restrictions and estimated using the implied unconditional sample moment restrictions. We show that this method allows us to capture heterogeneity in output elasticities along the firm-size distribution that would not be estimated in conditional mean models. 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, Nonlinear 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) or an intermediate material input using Levinsohn and Petrin (2003) or Ackerberg *et al.* (2015). These proxies are a function of a state variable such as capital and the unobserved productivty components. Under certain assumptions, this demand function is strictly increasing in its scalar unobserved productivity component. Inverting this

[*]Department of Economics, University of Iowa, S321 Pappajohn Business Building, 21 E Market St, Iowa City, IA 52242. Email: `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`

demand function controls for unobserved 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 firm's heterogeneous behavior and decision.[1] 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 firm heterogeneity in production technology beyond Hick's neutral productivity shock to be driven by the rank of the unobserved production shock, η_{it}. We simultaneously extend the proxy variable approach to this framework in order to control for the part of production unobservables that are correlated with inputs. Since applying the quantile regression requires non-smooth criterion function, it is not straightforward to estimate the production functions by allowing for endogenous inputs and their heterogeneous coefficients. We are not aware of any published paper which takes into account for 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 most current control function approaches of Levinsohn and Petrin (2003) and Ackerberg *et al.* (2015) and is successful in capturing heterogeneous output elasticities along the conditional distribution of firm's output. In our empirical application, we consider several popular firm and plant-level manufacturing datasets and compare our estimator to control function approaches. We show that heterogeneity in these estimates implies differences in capital intensity and TFP growth over time.

The rest of the paper is organized as follows. Section 2 reviews prior approaches for production function estimation and the literature on panel data quantile regression. Section 3 introduces the econometric model and the proposed estimator. Section **??** 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.

[1]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 by construction.

2 Literature Review

2.1 Production Function Estimation

We briefly review the LP (2003) procedure for estimating a *value-added* production function (in logs) [2,3].

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it}. \tag{1}$$

where y_{it} denotes value-added output, l_{it} denotes labor input for firm i at time t, k_{it} denotes capital input, ω_{it} is unobserved productivity and η_{it} denotes an iid shock to production.

To control for the correlation between ω_{it} and inputs k_{it} and l_{it}. LP introduce an intermediate input demand defined as [4]

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

where the function f is strictly increasing in ω_{it} for all k_{it}. Productivity can then be expressed as

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

Substituting into the production function

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

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

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

where $\mathcal{I}_{it}$ denotes the firm's information at time t. A linear approximation can be used, in which case estimates can be obtained from a simple linear regression.

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

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

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

[2] We consider a value-added production function here to be consistent with the model we introduce in Section 3 for reasons which we will discuss in the corresponding section

[3] We drop the constant β_0 since it is not separately identified from ω_{it} without a location normalization

[4] 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

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

where $\tilde{y}_{it} = y_{it} - \hat{\beta}_l l_{it}$ 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 estimators complicates asymptotic inference.

2.2 Production Functions and Quantile Regression

Connecting variation in the random error, η_{it}, to differences in a firm's final output decisions is not straightforward in the standard production function model. We briefly review a subfield of production function estimation that facilitates a more natural interpretation; production frontier models. We discuss limitations of these applications and return to our interpretation in Section 3.

A (stochastic) frontier (SFA) model of production proposed by Aigner *et al.* (1977) introduces statistical error into a frontier model. Frontier models assume firms firms deviate from an optimal frontier of production. The SFA model is typically written as

$$y_i = f(x_i, \beta) + \varepsilon_i, \tag{8}$$

where $\varepsilon_i = \eta_i - u_i$, x_i are inputs to production and β are the parameters. The error term η_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 β are typically obtained using maximum likelihood which requires strong distributional assumptions on the error terms. Estimation of the efficient frontier is then a conditional mean estimator rather than a maximal value estimator as noted by Bernini *et al.* (2004). They suggest quantile regression could then be used to estimate the highest percentiles of the conditional output distribution as it relates to the stochastic frontier, however, a theoretical difficulty is then choosing which quantile corresponds to the frontier. A more detailed derivation of a quantile representation of the frontier was introduced by Aragon *et al.* (2005) in a nonparametric model which requires inversion of a conditional empirical CDF. Since the purpose of this paper is not to compare the advantages and disadvantages of production frontier models and ours we leave this discussion for future studies and acknowledge the theoretical challenges of quantile frontier models.

There are two main challenges of implementing a quantile regression framework to the standard production function model. Firstly, as we alluded to earlier in this section, if we maintain a structural interpretation of firm production, reduced-form linear quantile regression models are likely not sufficient in linking a firm's output choice as a function of the error term η_{it} as we elaborate in the next section. Secondly, addressing the endogeneity of ω_{it} using traditional panel data methods have challenges specific towards the production function literature and quantile models.

Regarding the second point, 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), Canay (2011) which acts as a pure location shifter of the conditional quantile function. This approach may have two main disadvantages. First, assuming the unobservable is time-invariant is restrictive and Griliches and Hausman (1986) has shown to leads to low estimates of β_k. Secondly, 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). Abrevaya and Dahl (2008) adopt this approach with a linear data generating process for birth outcomes and linking it to its quantile function to estimate the effect of birth inputs over the birth-weight distribution. This approach is further developed by Bache *et al.* (2012). One downside of this approach is that it is difficult to describe the behavior of the conditional quantile function as it 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^k_{it} and p^l_{it} 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 β_k and β_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 $y_{it} = Q(k_{it}, l_{it}, U_{it})$ where $\tau \in (0, 1]$ denotes the quantile index, the model is identified from a quantile type moment restriction

$$P[y_{it} \leq Q(k_{it}, l_{it}, \tau)|k_{it}, l_{it}, p^k_{it}, p^l_{it}] = \tau \tag{9}$$

while this identification argument is critical in our model, we do not use the estimation procedure for reasons explained in the next section. One of these reasons are that input prices may not have enough variation across firms and exogeneity can be violated if they capture input quality differences as argued by Griliches and Hausman (1986).

3 A Random Coefficient Production Function

We specify a *value-added* production function as a random coefficient model: [5]

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

The variables in equation (10) have the same interpretation as the ones we introduced in the LP model. The only difference here is that we allow the output elasticities to be functionally dependent on the production shock η_{it} while productivity still maintains its additive separability. [6]

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

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

Which implies that the τth conditional quantile of y_{it} is given by

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

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

The formulation of (11) is not new to the production function literature. The assumption 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. The most common application of those models are in 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 η_{it} may be smaller. 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.

We note that under quantile preferences a firm who maximizes the τ 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 opti-

[5] A value-added specification is not without loss of generality. Only specific value-added production functions such as the Leontief value-added model can be mapped to its gross-output counterpart while also avoiding the non-identification results of Gandhi *et al.* (2020). This becomes more difficult with this specification.

[6] More specifically, productivity is only a location shifter of the conditional output distribution. We cannot allow $\omega_{it} = \omega_{it}(\eta_{it})$ since this would violate the scalar unobservability assumption of our proxy variable

mal expenditure on inputs. Different managers may have different preferences for risk. Quantile utility maximization is not a new concept. 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. We illustrate a basic estimation procedure using static first order conditions for this type of model in the Appendix for the reader's interest.

Returning to our misspecification interpretation in Equation (10), we follow LP in the usual set of assumptions on timing of input choices and scalar unobservability.

Assumption 3.1

(a) The production function $y_{it} = f_t(k_{it}, l_{it}, \omega_{it}, \eta_{it})$ is strictly increasing in η_{it}

(b) The firm's information set at time t includes current and past productivity shocks $\{\omega_{it}\}_{t=0}^{t}$, but does not include past productivity shocks $\{\omega_{it}\}_{t=t+1}^{\infty}$. η_{it} is independent of $\mathcal{I}_{it}$

(c) Firm's productivity shocks evolve according to a first-order Markov process

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it} \tag{13}$$

where the iid productivity innovations ξ_{it} are independent of $\mathcal{I}_{it-1}$.
We require that $P[\xi_{it} \leq F_{\xi}^{-1}(\tau)|\mathcal{I}_{it-1}] = P[\xi_{it} \leq F_{\xi}^{-1}(\tau)] = \tau$

(d) Firms accumulate capital according to

$$K_{it} = \kappa_t(I_{it-1}, K_{it-1}). \tag{14}$$

where K_{it-1} and I_{it-1} denote previous period capital and investment

(e) Firm's intermediate input demand function is given by $m_{it} = m_t(k_{it}, \omega_{it})$

(f) The intermediate input demand function $m_t(k_{it}, \omega_{it})$ is strictly increasing in ω_{it}

Given these Assumption (3.1)(e, f), we invert intermediate input demand $\omega_{it} = m^{-1}(k_{it}, m_{it})$ and substitute into the production function. We treat m_t^{-1} as a nonparametric function (k_{it}, m_{it}). We then have:

$$y_{it} = \beta_k(\eta_{it})k_{it} + \beta_l(\eta_{it})l_{it} + m_t^{-1}(k_{it}, m_{it}) = \beta_l(\eta_{it})l_{it} + \Phi(k_{it}, m_{it}, \eta_{it}) \tag{15}$$

Using Assumption (3.1)(a, b) we have the following identification condition for the first stage:

$$P\big(y_{it} \le \beta_l(\tau)l_{it} + \Phi(k_{it}, m_{it}; \tau)\big|\mathcal{I}_{it}\big) = \tau \tag{16}$$

We use a linear approximation $\Phi(k_{it}, m_{it}, \eta_{it}) = \beta_k(\eta_{it})k_{it} + \beta_m(\eta_{it})m_{it}$ so we can estimate $\beta_l(\tau)$ and $\Phi(k_{it}, m_{it}, \tau)$ using linear quantile regression. We can then plug these estimates back into the production function and apply Assumption (3.1)(c) to obtain:

$$y_{it} = \beta_k(\eta_{it})k_{it} + \hat{\beta}_l(\tau)l_{it} + \rho(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\eta_{it})k_{it-1}) + \xi_{it} \tag{17}$$

The main challenge of the identification result in the second stage is that it is not true in general that

$$P\big(y_{it} \le \beta_k(\tau)k_{it} + \hat{\beta}_l(\tau)l_{it} + \rho(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1}) + \xi_{it}|\mathcal{I}_{it-1}\big) = \tau \tag{18}$$

The lack of identification in this model is often lamented in the quantile panel data models using correlated random effects. For example, Canay (2011) notes that the conditional behavior of (18) depends on the joint distribution of η_{it} and ξ_{it} which makes identification of $(\rho, \beta_k(\tau))$ in the second stage a difficult task. We circumvent this by framing an identification condition in only the ξ_{it} component, similar to Ackerberg *et al.* (2015) by concentrating out the constant $\beta_0(\tau)$ and ρ. For a hypothetical guess of $\beta_k(\tau)$ we can write

$$\beta_0(\tau) + \omega_{it} = y_{it} - \hat{\beta}_l(\tau)l_{it} - \beta_k(\tau)k_{it} = \hat{\Phi}(k_{it}, m_{it}; \tau) - \beta_k(\tau)k_{it} \tag{19}$$

We can rewrite the AR(1) productivity process as

$$\hat{\Phi}(k_{it}, m_{it}; \tau) - \beta_k(\tau)k_{it} = \beta_0(\tau) + \rho(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1}) + \xi_{it} \tag{20}$$

and note that

$$Q_\tau(\omega_{it}|\mathcal{I}_{it-1}) = \beta_0(\tau) + \rho(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1}) + F_\xi^{-1}(\tau) \tag{21}$$

We estimate $(\beta_0(\tau) + F_\xi^{-1}(\tau), \rho)$ using the procedure of He (1997)

1. First, a median regression of $\hat{\Phi}(k_{it}, m_{it}; \tau) - \beta_k(\tau)k_{it}$ on $\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1}$ to obtain estimates of $(\beta_0(\tau), \rho)$

2. Second, let $\hat{\xi}_{it} = \hat{\Phi}(k_{it}, m_{it}; \tau) - \beta_k(\tau)k_{it} - \hat{\beta}_0(\tau) - \hat{\rho}(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1})$. Then take the τth quantile of $\hat{\xi}_{it}$ as an estimate for $\hat{F}_\xi^{-1}(\tau)$

Therefore, we have the following identification condition

$$P[\hat{\xi}_{it}(\beta_k(\tau)) \le \hat{F}_\xi^{-1}(\tau)|\mathcal{I}_{it-1}] = P[\hat{\xi}_{it}(\beta_k(\tau)) \le \hat{F}_\xi^{-1}(\tau)] = \tau \tag{22}$$

This can be represented by conditional moment restrictions

$$\mathbb{E}[\mathbb{1}\{\hat{\xi}_{it}(\beta_k(\tau)) - \hat{F}_{\xi}^{-1}(\tau) \leq 0\} - \tau | \mathcal{I}_{it-1}] = 0 \tag{23}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. To estimate the production function parameters we use the unconditional moments implied by (23)

$$\mathbb{E}[Z_{it-1}(\mathbb{1}\{\hat{\xi}_{it}(\beta_k(\tau)) - \hat{F}_{\xi}^{-1}(\tau) \leq 0\} - \tau)] = 0 \tag{24}$$

where Z_{it-1} includes the instruments used in the firm's information set at time $t-1$.

The presence of the indicator function in equation (16) makes estimation of the production function parameters intractable. We propose smoothing the indicator function using the methodology proposed by de Castro, Galvao, Kaplan and Liu (2018) for nonlinear conditional quantile models. These models can be exactly identified or over-identified by using more components in $\mathcal{I}_{it-1}$. To fix notation, let $z_{it} \subseteq \mathcal{I}_{it-1}$ denote the instruments used for estimation. For our application, $z_{it} = k_{it}$. Let $x_{it} = (k_{it}, k_{it-1}, m_{it-1})$ and let $\Lambda(\cdot)$ denote the residual function that defines the conditional quantile restriction in (15) that is known up to the finite-dimensional parameters $\beta(\tau) = (\beta_k(\tau), \rho(\tau))$. The sample analog of (16) can then be written as

$$\hat{M}_n(\beta, \tau) = \frac{1}{n} \sum_{i=1}^{n} \left[\tilde{I}\left(\frac{\Lambda(\tilde{y}_{it}, x_{it}, \beta(\tau))}{h_n} \right) - \tau \right] \tag{25}$$

where h_n is a bandwidth (sequence) and $\tilde{I}(\cdot)$ is a smoothed version of the indicator function $\mathbb{1}\{\cdot \leq 0\}$ used by Horowitz (1998), Whang (2006), and Kaplan and Sun (2016):

$$\tilde{I}(u) = \mathbb{1}\{-1 \leq u \leq 1\}\big[0.5 + \frac{105}{64}\big(u - \frac{5}{3}u^3 + \frac{7}{5}u^5 - \frac{3}{7}u^7\big)\big] + \mathbb{1}\{u > 1\}. \tag{26}$$

Despite the model defined in (19) being exactly identified we choose the smoothed GMM estimator proposed by de Castro *et al.* (2018) for over-identification. Let $\hat{W}$ be a symmetric, positive definite weighting matrix. The GMM estimator minimizes the following criterion function

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \hat{M}_n(\beta, \tau)^\top \hat{W} \hat{M}_n(\beta, \tau). \tag{27}$$

For the weighting matrix, we use the estimator of the inverse long run variance of the sample moments (17). Consistency of an estimator of the long run variance for quantile models is not well established in the literature and is only further complicated by the presence of the nonparametric control function estimated in the first stage. A discussion about the asymptotics of our semiparametric estimator will be provided in a later version of this paper.

The objective function in (22) may be non-convex. The global minimum is found using a simulated annealing algorithm from the GenSA package in R (Xiang *et al.*, 2013). A reasonable initial value for the optimization algorithm can be found by estimating an exactly-identified version (Method of Moments) of (19) using a root-finding algorithm such as the Newton-Raphson method (newtonsys in the Pracma package in R) or from a one-step estimator such as the one proposed by Newey and McFadden (1994).

4 Monte Carlo Experiments

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

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + (0.7k_{it} - 0.6l_{it})\eta_{it} \tag{28}$$

with $\beta_0 = 0$, $\beta_k = 0.4$ and $\beta_l = 0.6$. For each simulation we simulate two DGPs with $\eta_{it} \sim N(0, 0.1)$ and $\eta_{it} \sim Laplace(0, 0.1)$.

To produce consistent estimates of the labor coefficient in the first stage, we do not allow for any wage variation across firms and labor is chosen at time t with perfect information about ω_{it}. However, we add optimization error in labor.

An AR(1) process is specified for productivity $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$ where $\rho = 0.7$. The variance of ξ_{it} and initial value ω_{i0} is set so that the standard deviation of ω_{it} is constant over time and equal to 0.3

We compare the LP estimation procedure with the two-step procedure using Quantile GMM (QGMM) under the two different sets of experiments specified earlier. We estimate the model for $\tau \in \{0.1, 0.15, \dots, 0.85, 0.9\}$ fixing the smoothing bandwidth as $h = 0.1$ and using instruments current period capital, k_{it} as our instrument so that our model is exactly identified. For the weighting matrix, we use an estimate of the long-run variance covariance matrix using a truncated kernel with bandwidth choice of Andrews (1991). We use a continuously updated GMM procedure such that estimates of $\beta_k(\tau)$ in the second stage are estimated simultaneously with the weighting matrix. We initialize the algorithm at the true value of $\beta_k(\tau)$ however we find that the estimation is robust to reasonable initial values as well as bandwidth choices.

Figure 1: QLP estimated coefficients of $\beta_k(\tau)$ and $\beta_l(\tau)$. Dotted line is LP estimator

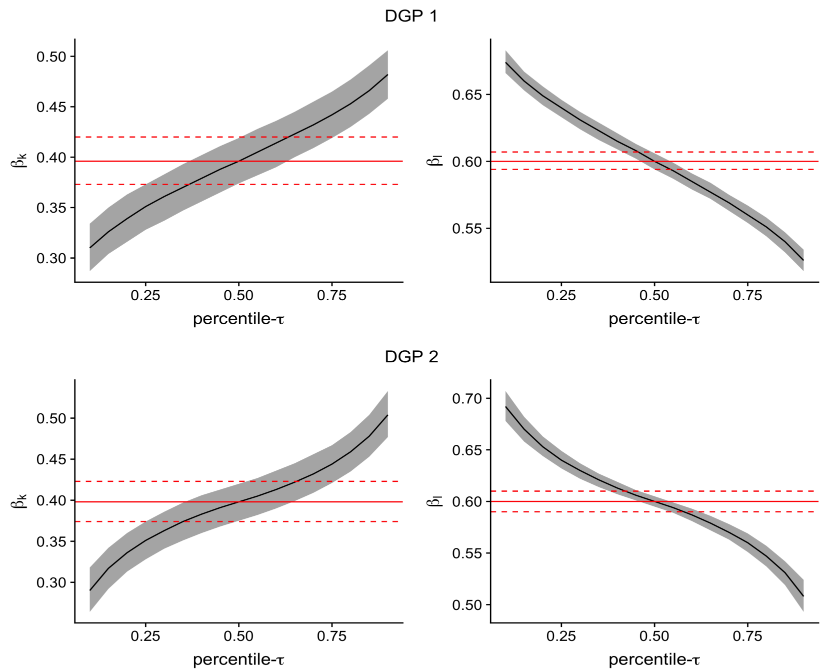
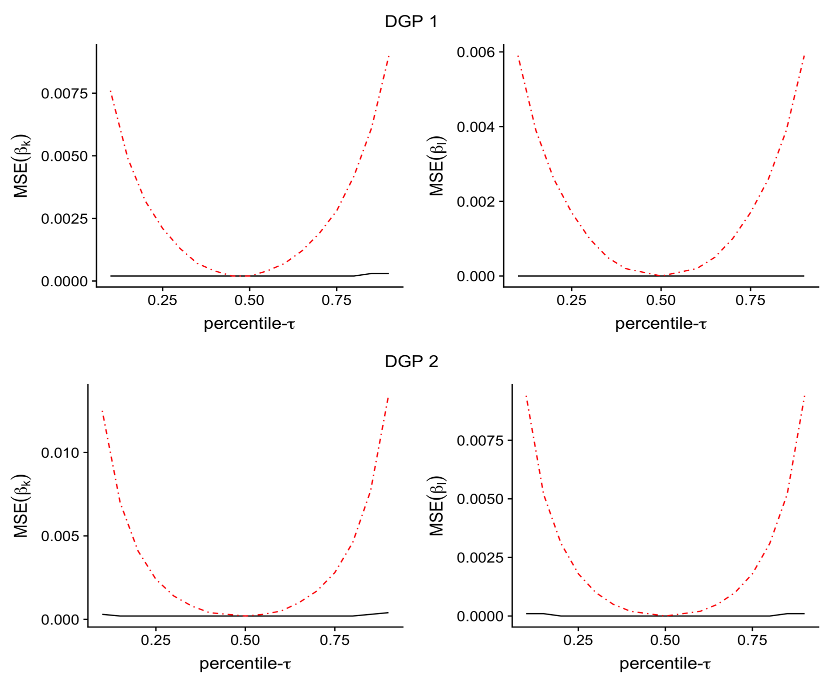
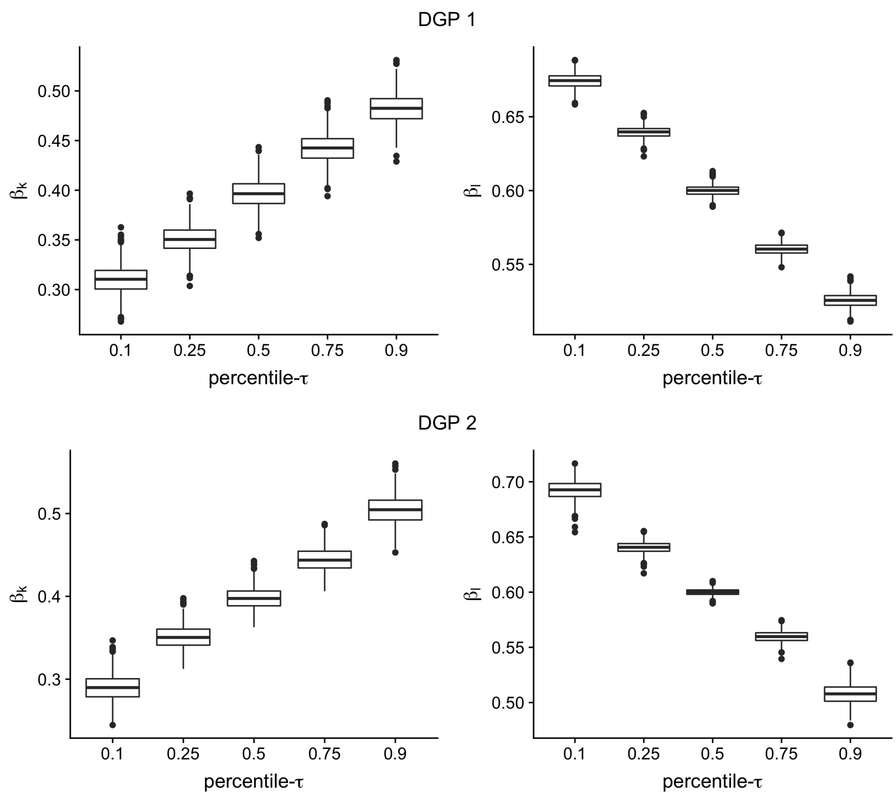


Figure 2: Simulated precision of QLP estimators of $\beta_k(\tau)$ and $\beta_l(\tau)$s. Dotted line is LP estimator.



The next figure summarizes distributional information from our Monte Carlo experiment. These simulation results show that our estimator has good finite-sample performance and does reasonably well at capturing firm heterogeneity across quantiles.

Figure 3: LP Box Plots for estimated coefficients from 1000 replications of Monte Carlo



5 Application

5.1 US Compustat

Table 1: 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

Table 2: Coefficient Estimates and Standard Errors for US Manufacturing Firms

		Capital		Labor		Returns to Scale	
Industry (NAICS code)	τ	Coef.	s.e.	Coef.	s.e.	Coef.	s.e
31	0.10	0.107	0.0261	0.631	0.0304	0.738	0.0331
	0.25	0.146	0.0273	0.558	0.0353	0.704	0.0362
	0.50	0.183	0.0205	0.486	0.0378	0.669	0.0359
	0.75	0.224	0.0415	0.442	0.0355	0.666	0.0471
32	0.10	0.092	0.0656	0.694	0.0457	0.786	0.0760
	0.25	0.103	0.0216	0.641	0.0323	0.744	0.0325
	0.50	0.137	0.0170	0.594	0.0246	0.731	0.0284
	0.75	0.166	0.0159	0.548	0.0237	0.714	0.0250
33	0.10	0.046	0.0266	0.215	0.0664	0.261	0.0595
	0.25	0.075	0.0182	0.358	0.0388	0.433	0.0350
	0.50	0.095	0.0138	0.449	0.0275	0.545	0.0253
	0.75	0.116	0.0132	0.469	0.0205	0.584	0.0194
All	0.10	0.099	0.0144	0.380	0.0368	0.479	0.0359
	0.25	0.120	0.0119	0.421	0.0253	0.541	0.0232
	0.50	0.128	0.0098	0.478	0.0203	0.606	0.0184
	0.75	0.150	0.0095	0.485	0.0158	0.635	0.0149

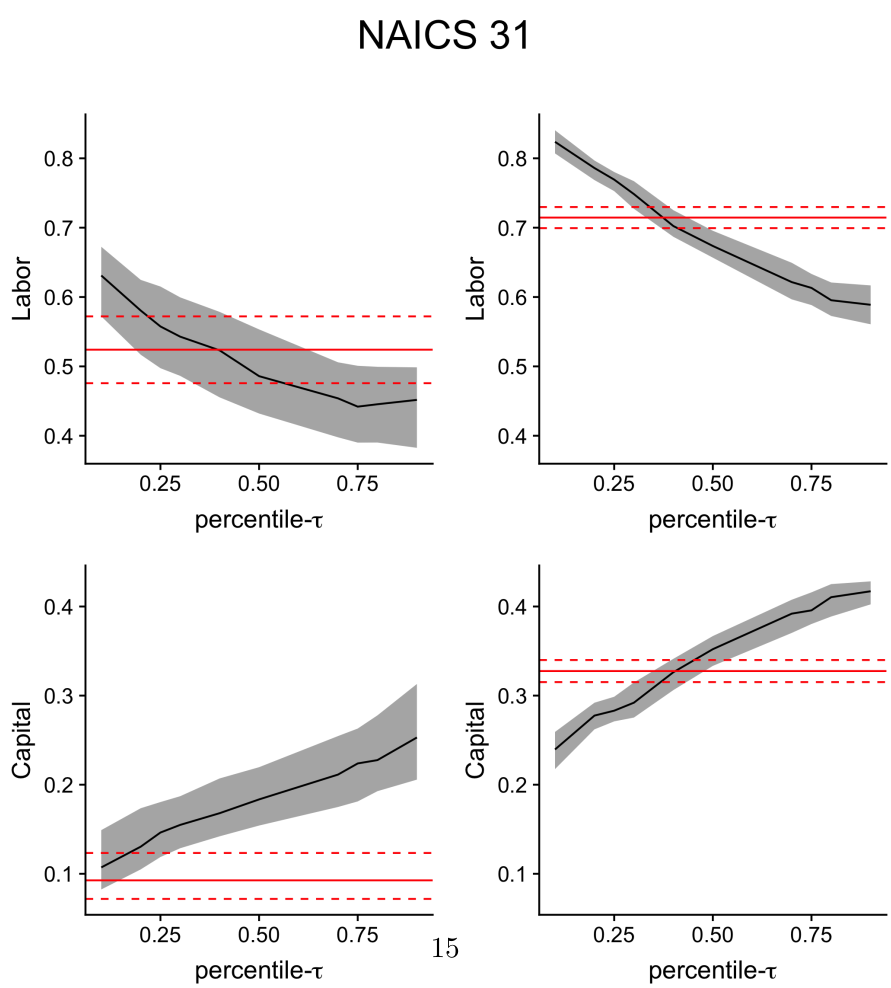


Figure 4: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression

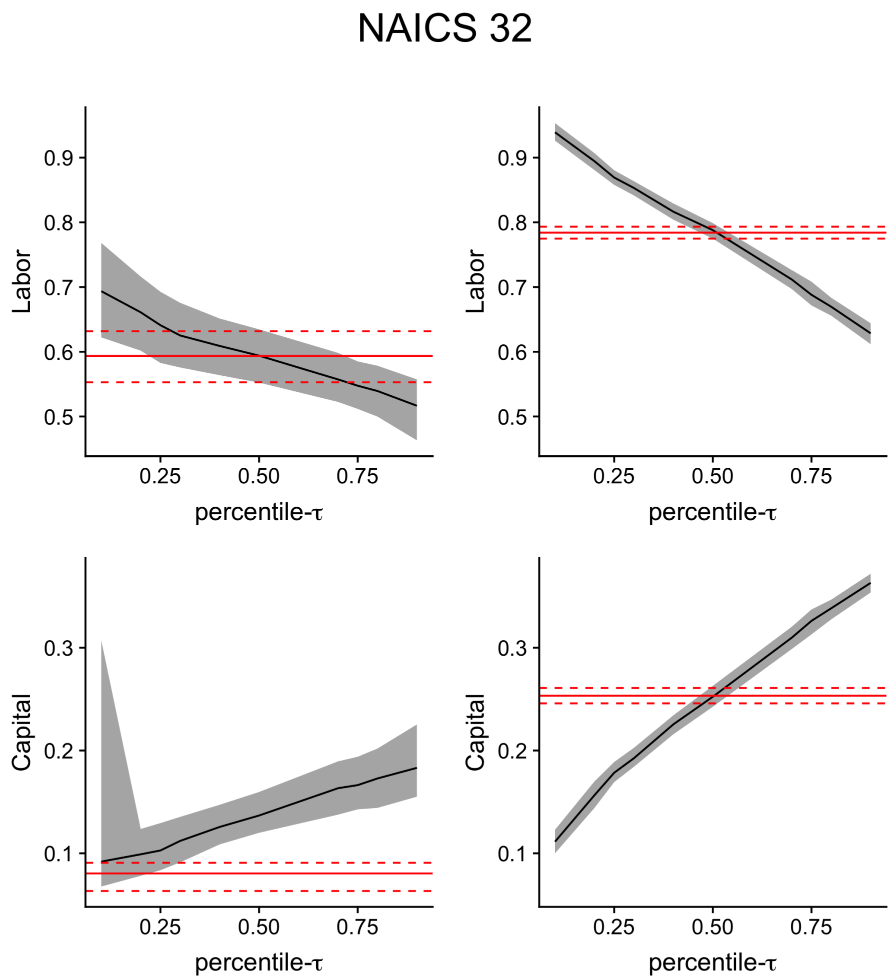


Figure 5: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

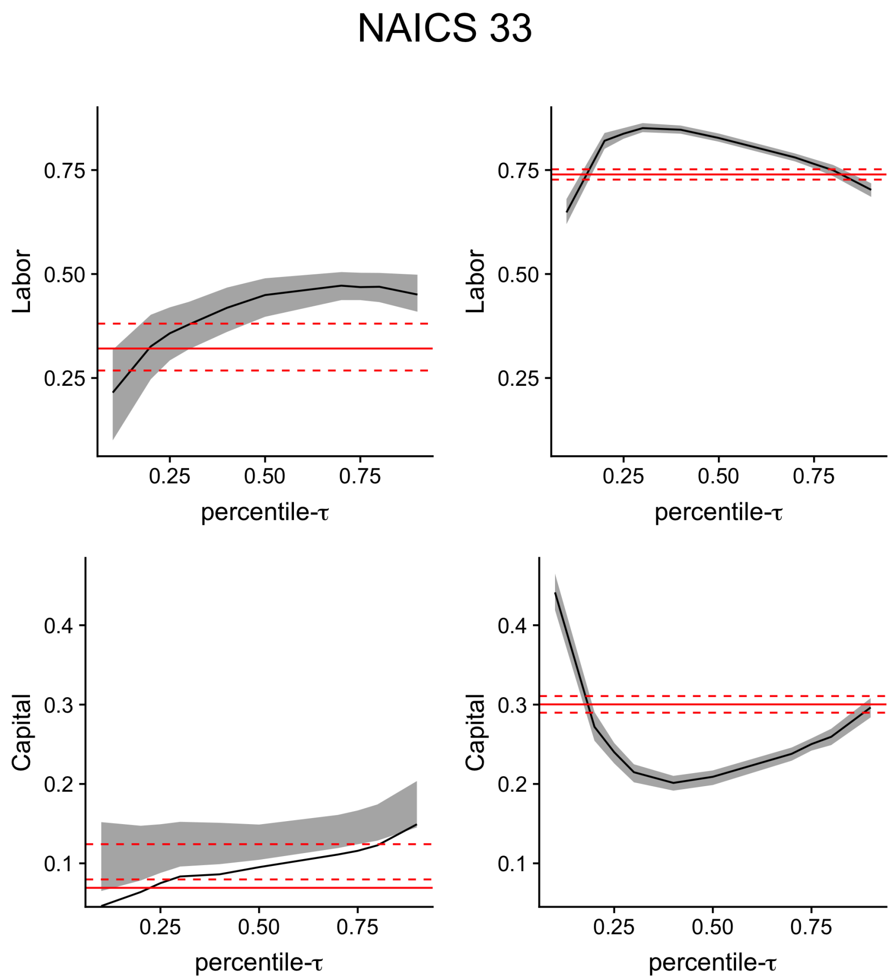


Figure 6: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

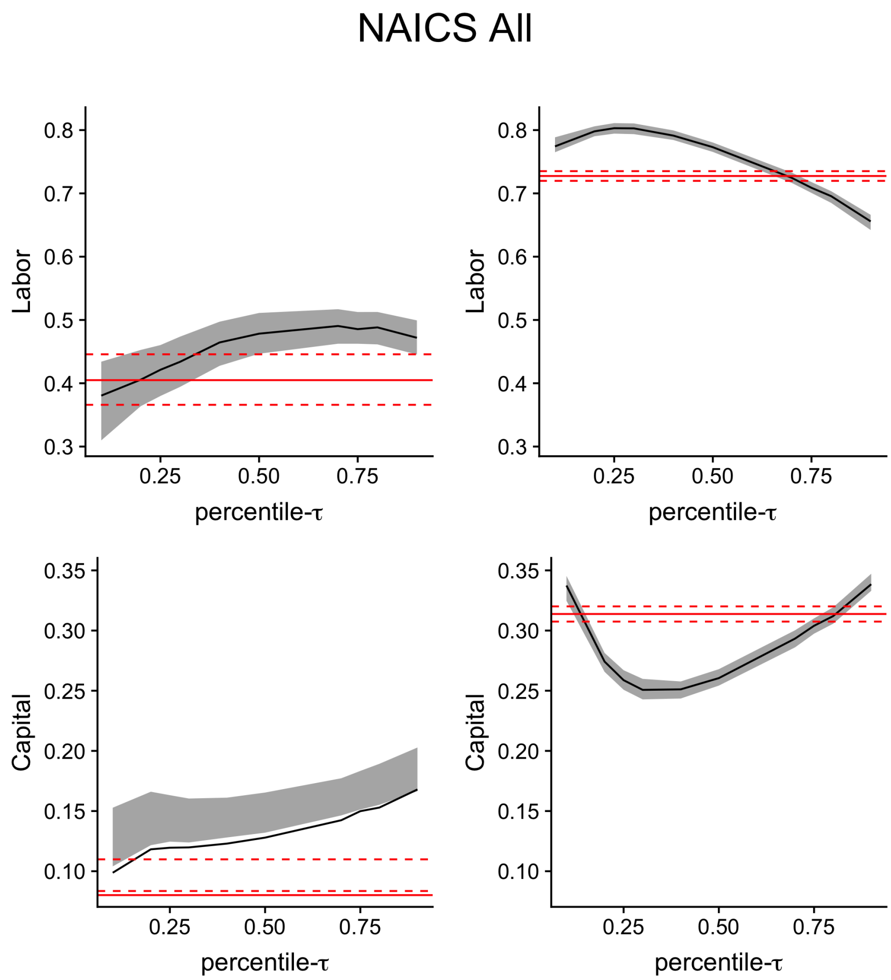


Figure 7: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

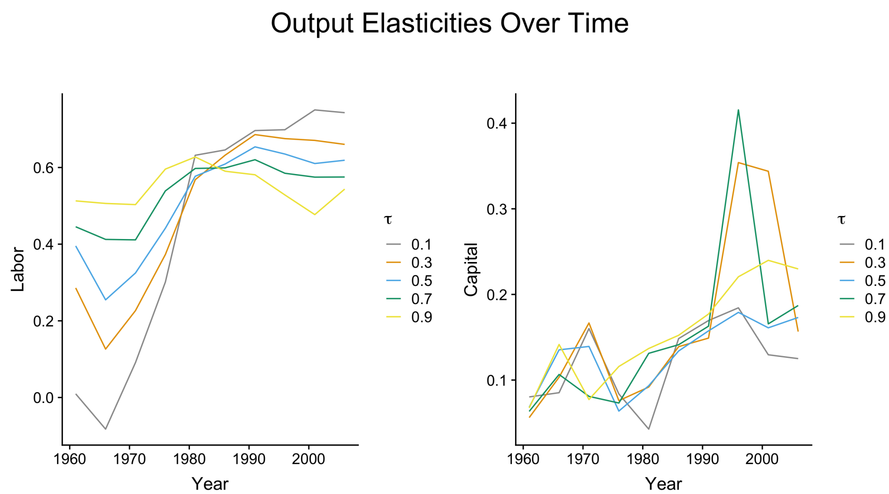


Figure 8: Estimated values of production function coefficients over time estimated at 5 year intervals

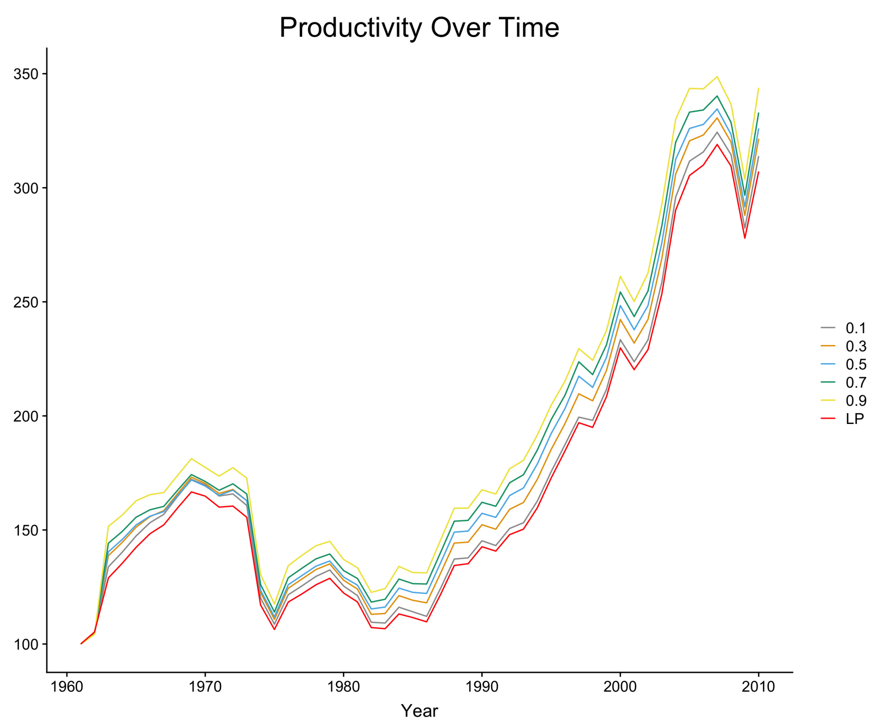


Figure 9: Estimated average TFP over time for the US. Base productivity in 1961 is set to 100.

5.2 Chilean Manufacturing

Table 3: 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

Table 4: Coefficient Estimates and Standard Errors for Chile Manufacturing Firms

Industry (ISIC code)	τ	Capital Coef.	Capital s.e.	Labor Coef.	Labor s.e.	Returns to Scale Coef.	Returns to Scale s.e
311	0.10	0.310	0.0218	0.405	0.0380	0.715	0.0348
	0.25	0.324	0.0206	0.411	0.0306	0.735	0.0274
	0.50	0.331	0.0175	0.398	0.0231	0.729	0.0234
	0.75	0.310	0.0168	0.411	0.0218	0.721	0.0233
381	0.10	0.355	0.0770	0.743	0.0631	1.098	0.0895
	0.25	0.356	0.0227	0.630	0.0439	0.986	0.0380
	0.50	0.357	0.0201	0.567	0.0365	0.924	0.0303
	0.75	0.351	0.0227	0.522	0.0399	0.873	0.0322
321	0.10	0.326	0.0718	0.705	0.0710	1.031	0.0878
	0.25	0.324	0.0447	0.640	0.0529	0.964	0.0557
	0.50	0.319	0.0267	0.585	0.0432	0.904	0.0342
	0.75	0.314	0.0240	0.556	0.0378	0.869	0.0305
All	0.10	0.416	0.0115	0.525	0.0208	0.941	0.0172
	0.25	0.432	0.0095	0.503	0.0170	0.934	0.0141
	0.50	0.432	0.0085	0.463	0.0141	0.894	0.0116
	0.75	0.426	0.0083	0.434	0.0156	0.859	0.0127

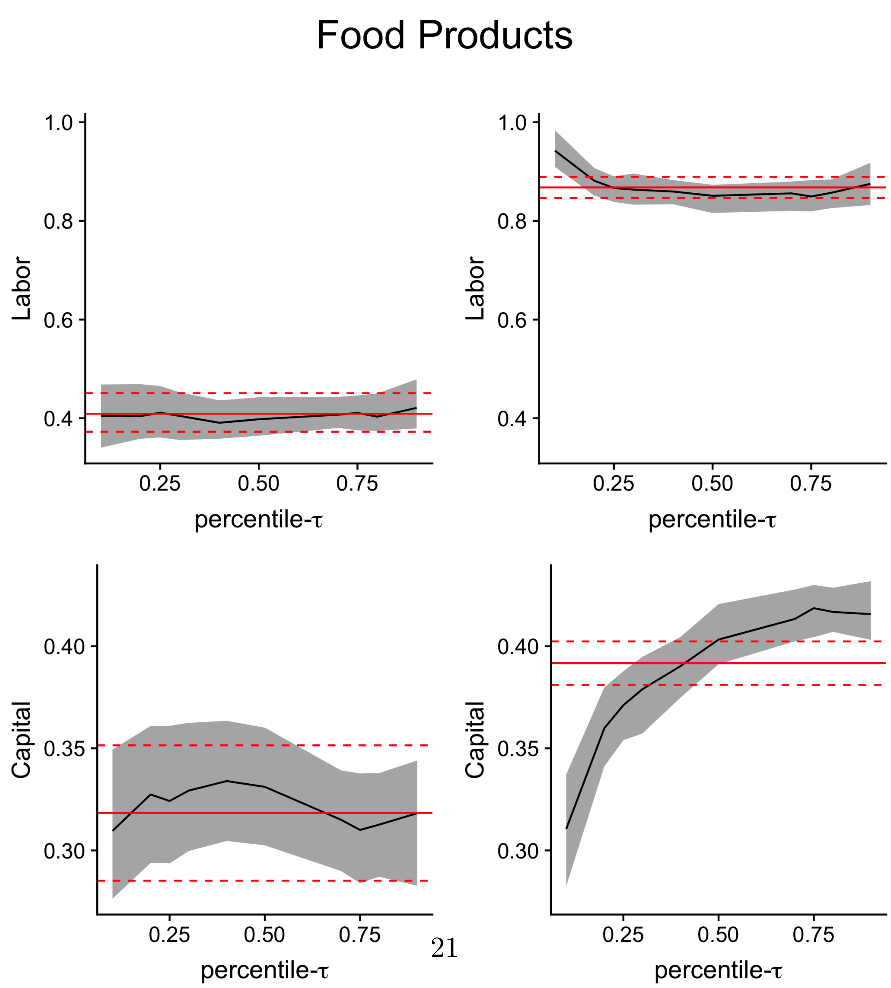


Figure 10: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression

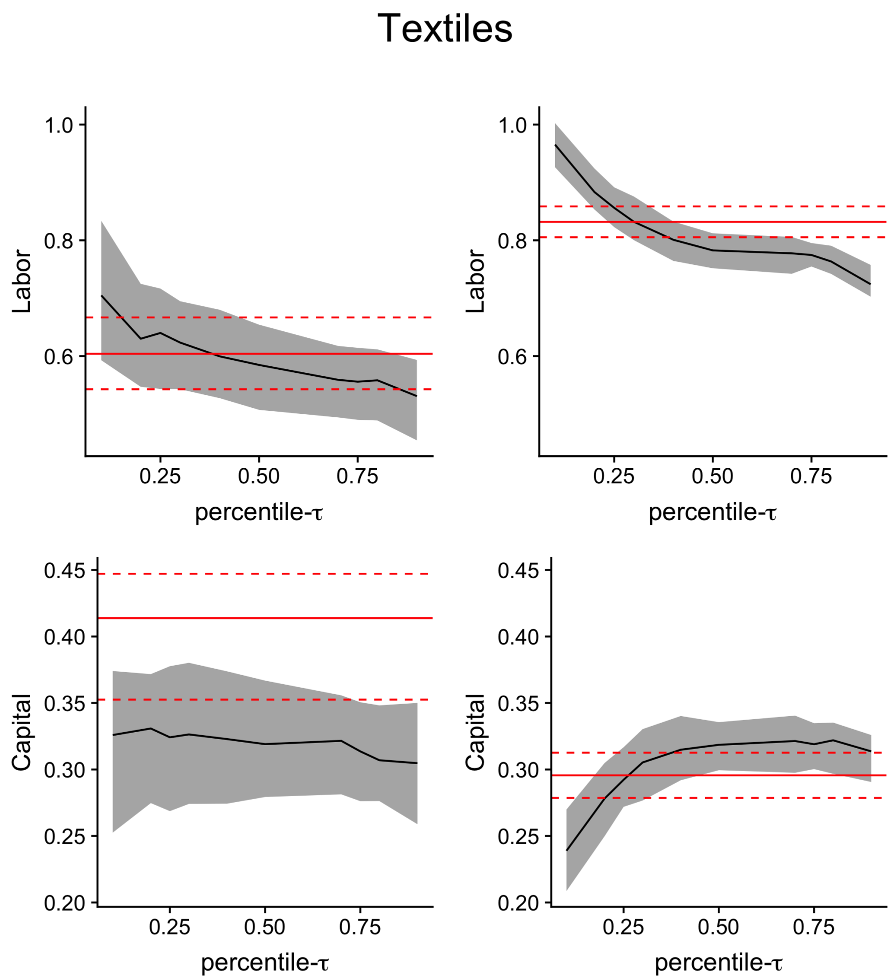


Figure 11: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

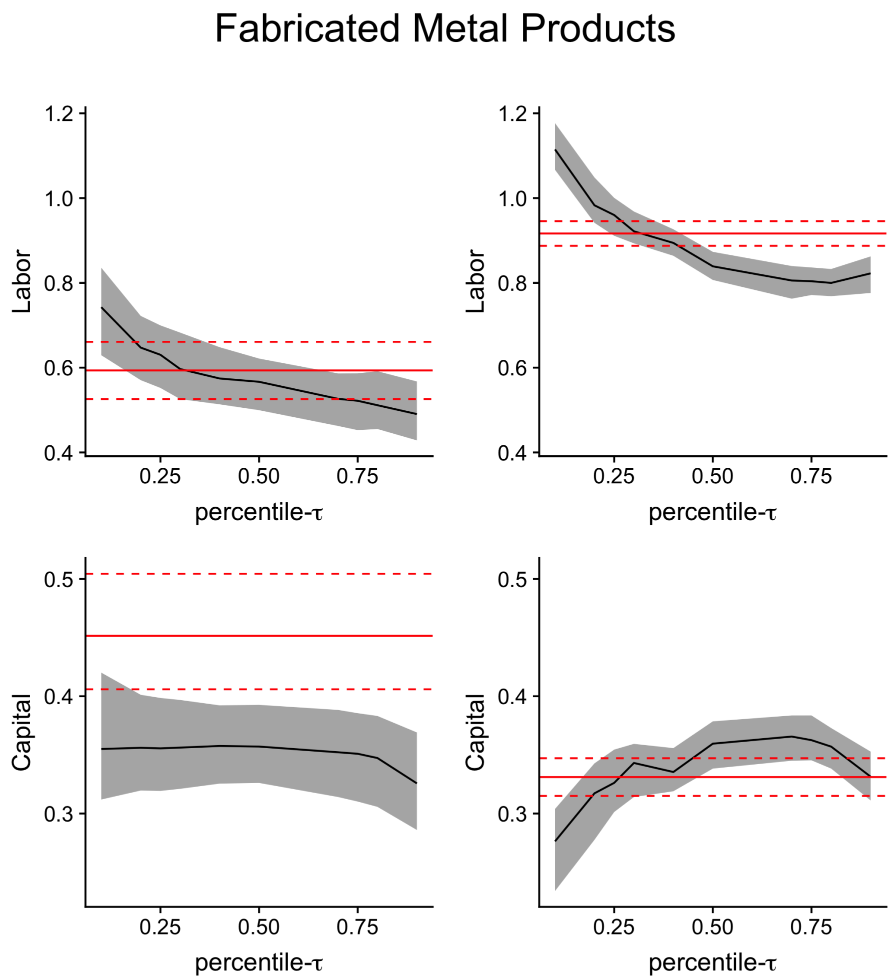


Figure 12: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

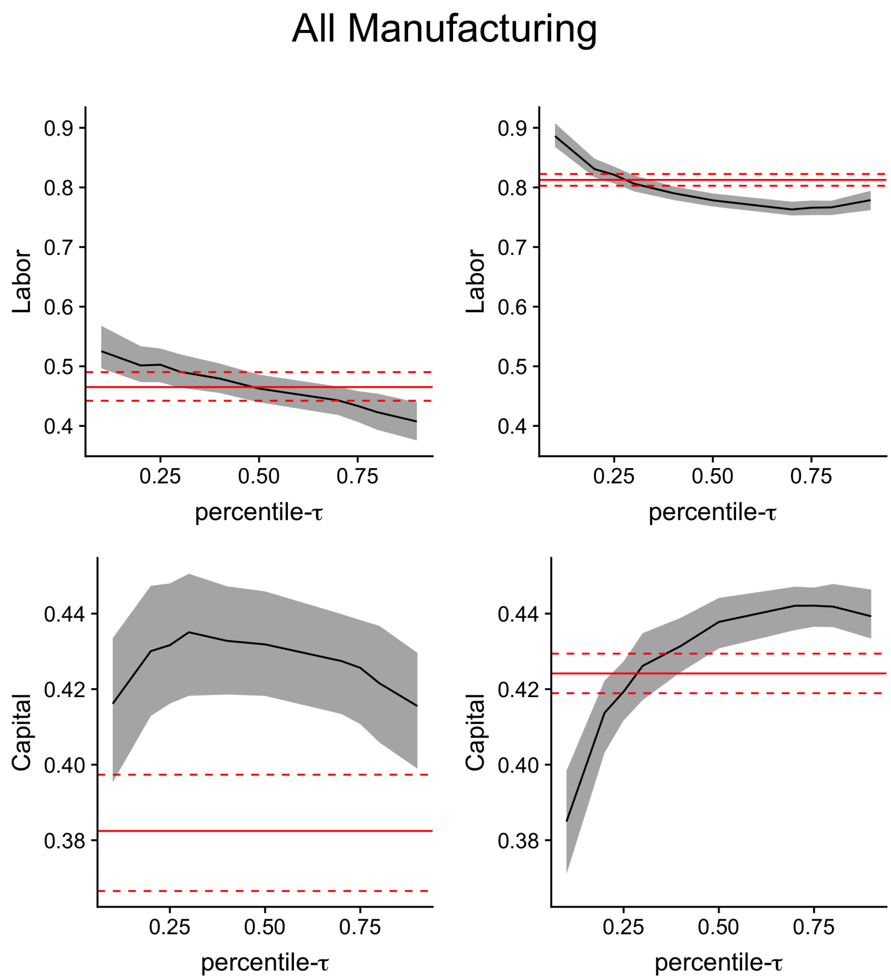


Figure 13: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

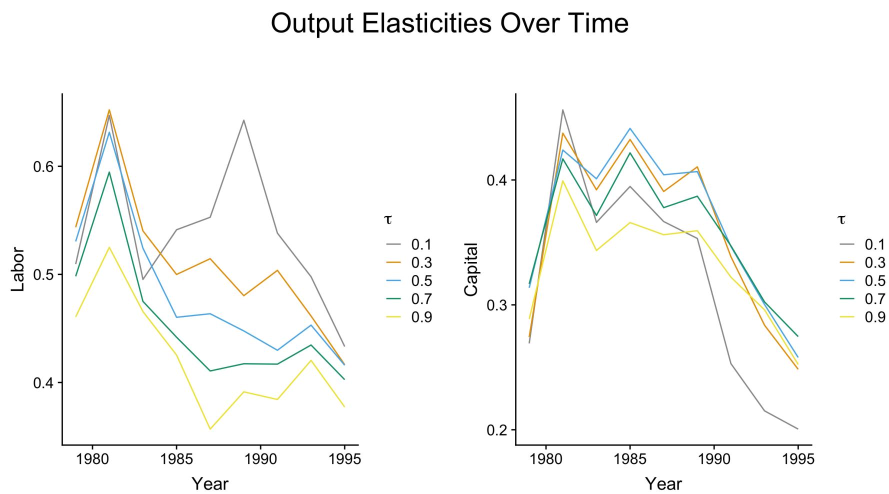


Figure 14: Estimated values of production function coefficients over time estimated at 2 year intervals

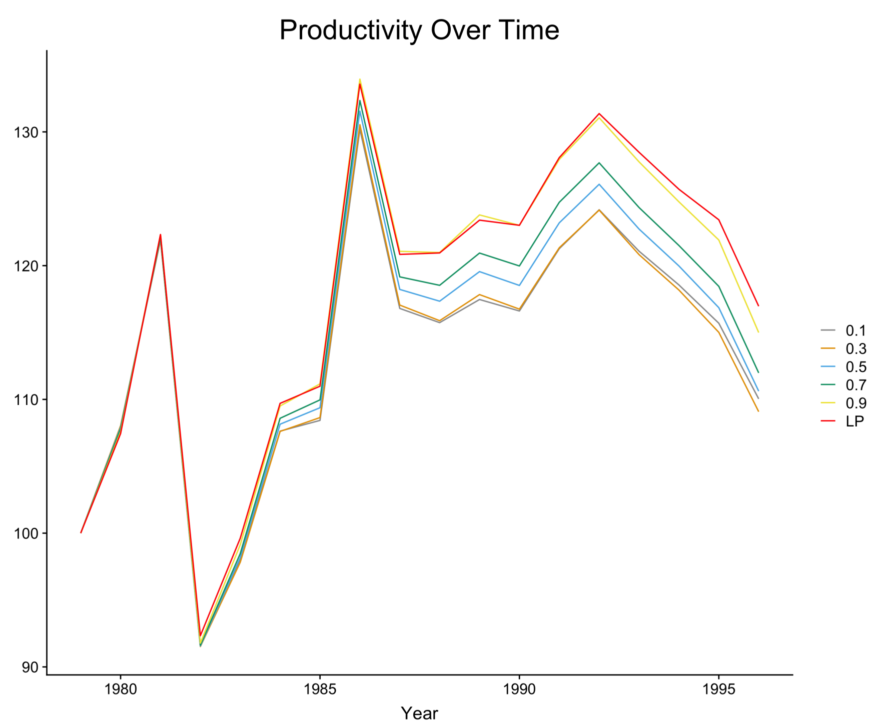


Figure 15: Estimated average TFP over time for Chile. Base productivity in 1979 is set to 100.

5.3 Colombia Manufacturing

Table 5: 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

Table 6: Coefficient Estimates and Standard Errors for Colombia Manufacturing Firms

		Capital		Labor		Returns to Scale	
Industry (ISIC code)	τ	Coef.	s.e.	Coef.	s.e.	Coef.	s.e
311	0.10	0.217	0.0983	0.747	0.0284	0.964	0.0999
	0.25	0.253	0.0484	0.629	0.0175	0.882	0.0490
	0.50	0.275	0.0173	0.519	0.0181	0.795	0.0203
	0.75	0.315	0.0198	0.418	0.0216	0.733	0.0211
322	0.10	0.109	0.0188	0.844	0.0270	0.952	0.0228
	0.25	0.107	0.0171	0.773	0.0238	0.881	0.0195
	0.50	0.144	0.0187	0.670	0.0255	0.814	0.0210
	0.75	0.213	0.0230	0.540	0.0307	0.753	0.0250
381	0.10	0.219	0.0343	1.180	0.0580	1.400	0.0438
	0.25	0.263	0.0771	0.938	0.0305	1.201	0.0793
	0.50	0.292	0.1299	0.760	0.0312	1.053	0.1323
	0.75	0.316	0.1846	0.655	0.0295	0.970	0.1876
All	0.10	0.222	0.0846	0.871	0.0151	1.093	0.0854
	0.25	0.252	0.0139	0.747	0.0101	1.000	0.0146
	0.50	0.274	0.0065	0.647	0.0088	0.921	0.0075
	0.75	0.295	0.0073	0.572	0.0104	0.867	0.0088

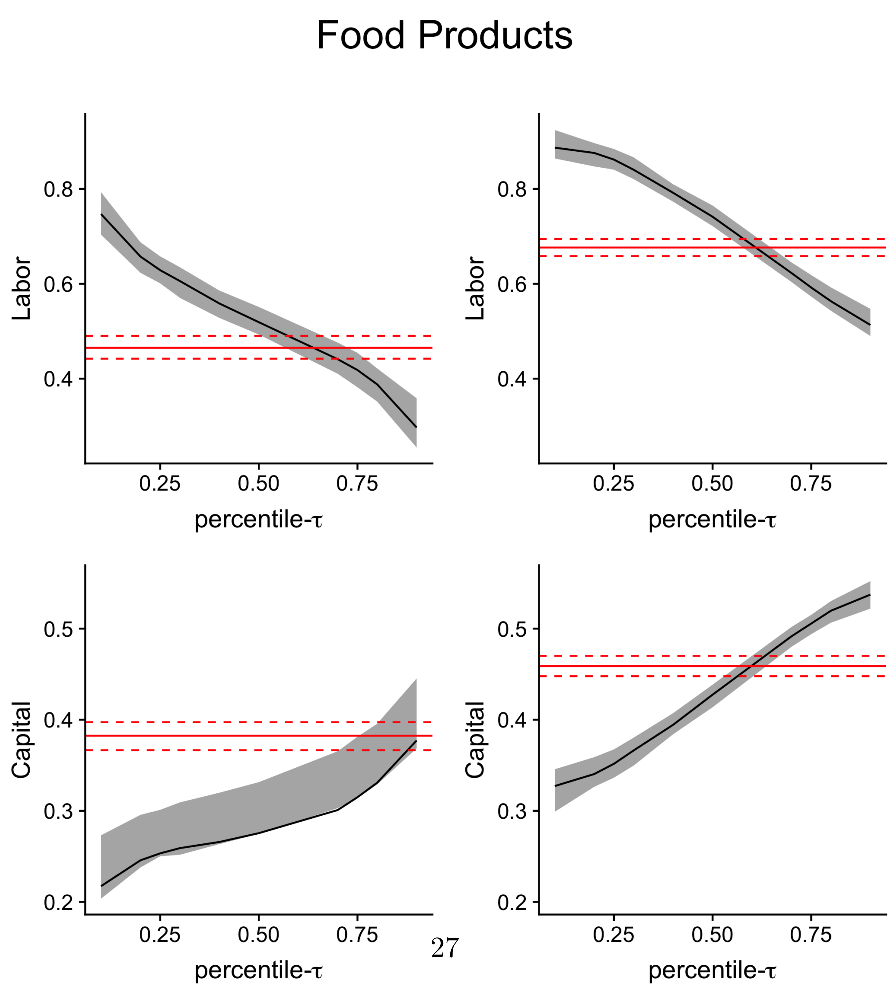


Figure 16: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression

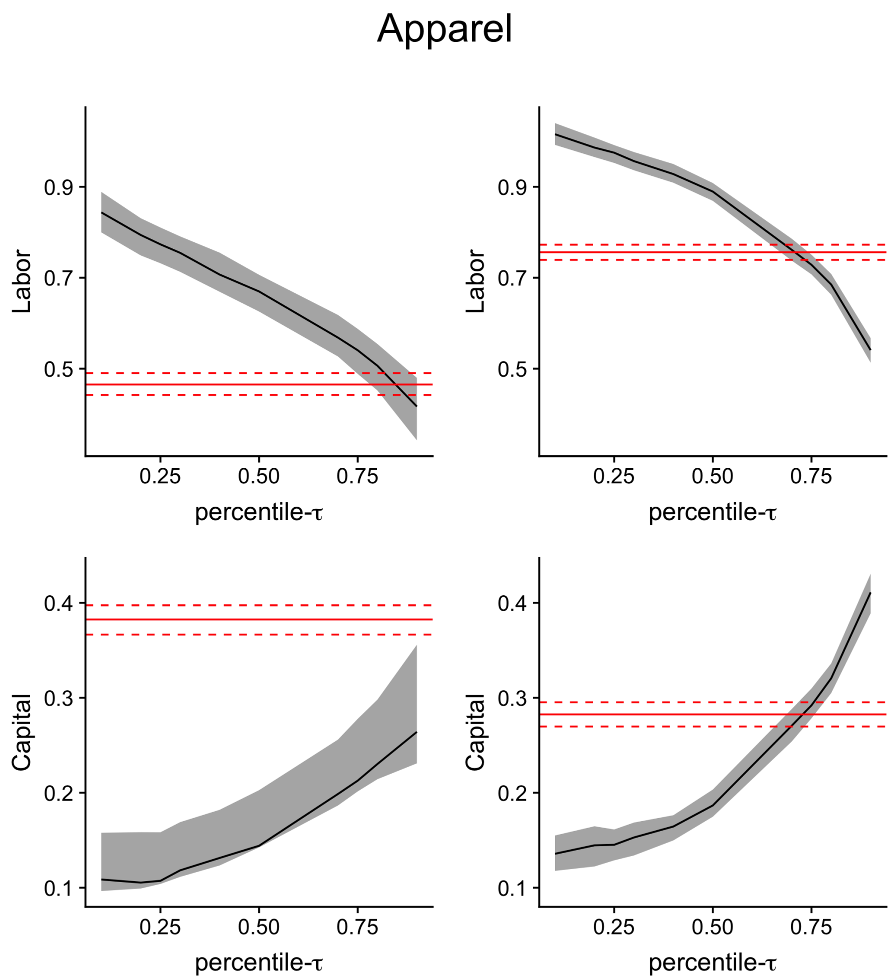


Figure 17: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

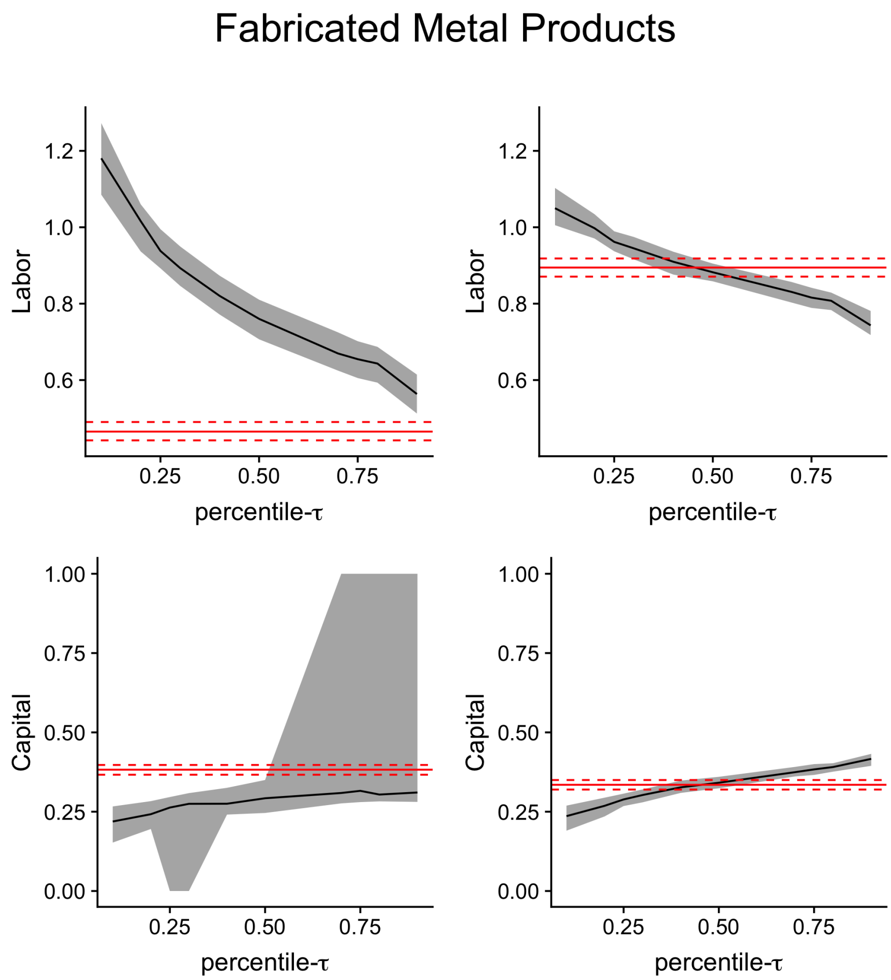


Figure 18: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

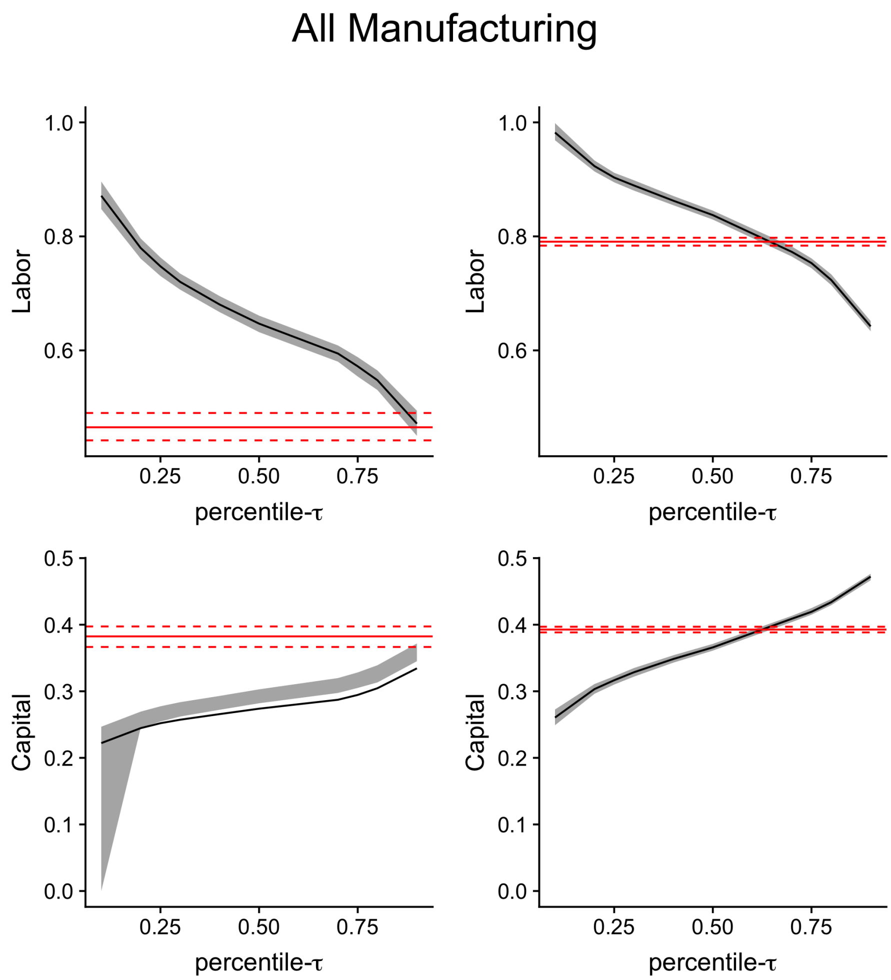


Figure 19: Estimated values of production function coefficients and their 90% confidence interval. The plots on the LHS are the QLP and LP estimates. The plots on the RHS are quantile regression and OLS estimates.

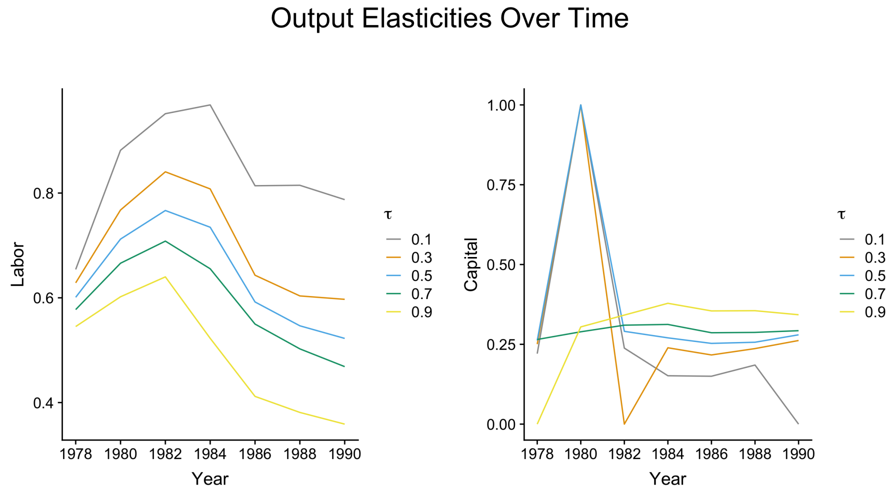


Figure 20: Estimated values of production function coefficients over time estimated at 2 year intervals

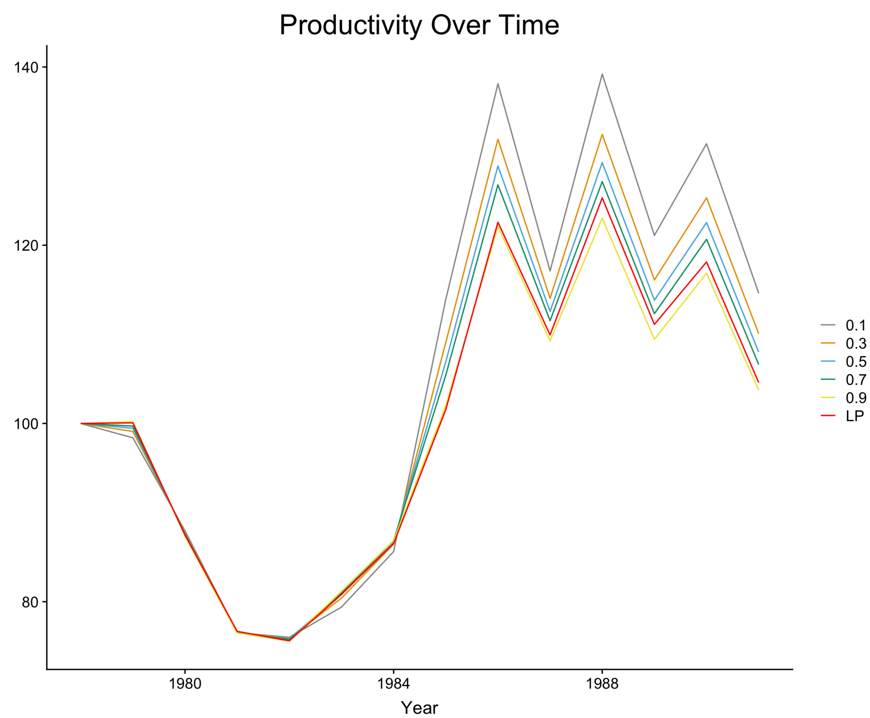


Figure 21: Estimated average TFP over time for Colombia. Base productivity in 1978 is set to 100.

6 Conclusions

We proposed a method that extends the intermediate input proxy variable approach to estimating quantiles of the conditional quantiles of firm production. The method is computationally attractive as it resembles the two-stage estimator introduced in the control function literature with conditional quantile restrictions at each stage. As a result, practitioners are able to easily apply the proposed estimator to production function models where the data reveal significant heterogeneous output elasticities along the conditional distribution of firm's output. We showed that this estimator works well in finite samples by replicating the experiment of ACF (2015) and showing that it captures heterogeneity in firm-size under different data generating processes.

Econometric issues with this estimator are currently being explored. A method to consistently estimate the long-run variance of the sample moment conditions to achieve the semi-parametric efficiency bound under general conditions is desired and extending the asymptotic results of de Castro, Galvao, Kaplan and Liu (2018) might not be straightforward. We also seek to estimate and interpret the resulting estimates of total factor productivity. Once these are addressed, an application using data such as the Chilean firm-level data may reveal heterogeneity in production technology along the distribution of firm output. We leave them as future research agenda.

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.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59** (3), 817.

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).

Bache, S. H. M., Dahl, C. M. and Kristensen, J. T. (2012). Headlights on tobacco road to low birthweight outcomes. *Empirical Economics*, **44** (3), 1593–1633.

Balat, J., Brambilla, I. and Sasaki, Y. (2018). Heterogeneous firms: skilled-labor productivity and the destination of exports, Working paper.

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.

Borchers, H. W. (2015). Pracma: practical numerical math functions. R package version 2.25.

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. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, **73** (1), 245–261.

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.

He, X. (1997). Quantile curves without crossing. *The American Statistician*, **51** (2), 186.

Horowitz, J. L. (1998). Bootstrap methods for median regression models. *Econometrica*, **66** (6), 1327.

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.

Kaplan, D. M. and Sun, Y. (2016). Smoothed estimating equations for instrumental variables quantile regression. *Econometric Theory*, **33** (1), 105–157.

Kasahara, H., Schrimpf, P. and Suzuki, M. (2017). Identification and estimation of production function with unobserved heterogeneity, Working paper.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, **91** (1), 74–89.

Lamarche, C. (2010). Robust penalized quantile regression estimation for panel data. *Journal of Econometrics*, **157** (2), 396–408.

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.

Manski, C. F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision*, **25** (1), 79–104.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In: R. F. Engle and D. L. McFadden (eds) Handbook of econometrics. vol. 4, *36*, Elsevier, pp. 2111–2245.

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.

Whang, Y.-J. (2006). Smoothed empirical likelihood methods for quantile regression models. *Econometric Theory*, **22** (02).

Xiang, Y., Gubian, S., Suomela, B. and Hoeng, J. (2013). Generalized simulated annealing for global optimization: The GenSA package. *The R Journal*, **5** (1), 13.