Proxy Variable and Dynamic Nonlinear Estimation of Production Function Updates

Justin Doty

October 29, 2020

Introduction: Random Coefficient Model

- Consider the random coefficient gross output production function (in logs)

$$y_{it} = \beta_0(\eta_{it}) + \beta_k(\eta_{it})k_{it} + \beta_l(\eta_{it})l_{it} + \beta_m(\eta_{it})m_{it} + \omega_{it} \qquad (1)$$

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

- If the production function is strictly increasing in η_{it} then the conditional quantile of (1) are

$$Q_\tau(y_{it}|\mathcal{I}_{it}) = \beta_0(\tau) + \beta_k(\tau)k_{it} + \beta_l(\tau)l_{it} + \beta_m(\tau)m_{it} + \omega_{it} \qquad (2)$$

  where $\mathcal{I}_{it}$ is the firm's information set at time t which is independent of η_{it}.

- In this specification, ω_{it} is a location-shifter of the conditional output distribution

Random Coefficient Model

- Doty and Song (2020) show how to identify and estimate this simple linear random coefficient model by extending the control variable approach of Levinsohn and Petrin, 2003 of a value-added production function
- I will briefly review our approach here. They assume the following **Assumption 1**
  1. The production function $y_{it} = f_t(k_{it}, l_{it}, \omega_{it}, \eta_{it})$ is strictly increasing in η_{it}
  2. 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 $\mathcal{I}_{it}$
  3. Firm's productivity shocks evolve according to a first-order Markov process

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it} \tag{3}$$

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

Random Coefficient Model

Assumption 1 continued

1. Firms accumulate capital according to

$$K_{it} = \kappa_t(I_{it-1}, K_{it-1}). \tag{4}$$

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

2. Firm's intermediate input demand function is given by $m_{it} = m_t(k_{it}, \omega_{it})$
3. The intermediate input demand function $m_t(k_{it}, \omega_{it})$ is strictly increasing in ω_{it}

- 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:

$$y_{it} = \beta_k(\eta_{it})k_{it} + \beta_l(\eta_{it})l_{it} + m_t^{-1}(k_{it}, m_{it}) = \beta_l(\eta_{it})l_{it} + \Phi(k_{it}, m_{it}, \eta_{it}) \tag{5}$$

Random Coefficient Model

- Using Assumption 1 we have the following identification condition for the first stage:

$$P\left(y_{it} \leq \beta_l(\tau)l_{it} + \Phi(k_{it}, m_{it}; \tau)\middle|\mathcal{I}_{it}\right) = \tau \tag{6}$$

- 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 write a second stage identification condition in only the ξ_{it} component, similar to Ackerberg, Caves, and Frazer, 2015 by concentrating out the constant $\beta_0(\tau)$ and ρ. For a hypothetical guess of $\beta_k(\tau)$ we can write

$$\tilde{\omega}_{it} = \beta_0(\tau) + \omega_{it} = y_{it} - \hat{\beta}_l(\tau)l_{it} - \beta_k(\tau)k_{it} = \hat{\Phi}(k_{it}, m_{it}; \tau) - \beta_k(\tau)k_{it} \tag{7}$$

Random Coefficient Model

- We estimate the AR(1) model with a constant so that

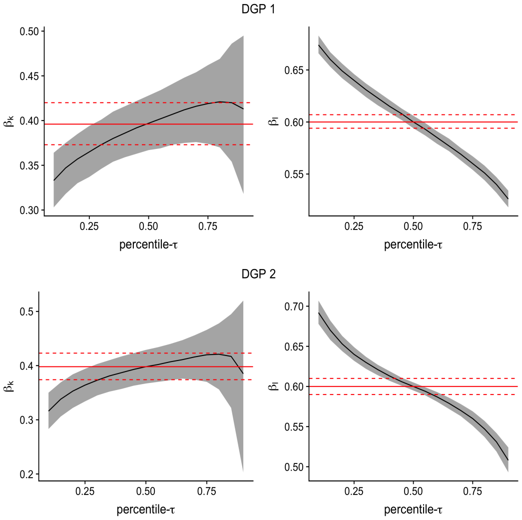
$$Q_\tau(\tilde{\omega}_{it}|\mathcal{I}_{it-1}) = \alpha + \rho(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1}) + F_\xi^{-1}(\tau) \quad (8)$$

- We estimate $(\alpha + F_\xi^{-1}(\tau), \rho)$ using the procedure of He, 1997 called Restricted Regression Quantiles (RQQ)
  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 ρ so that it does not vary over τ
  2. Second, let $\hat{\xi}_{it} = \hat{\Phi}(k_{it}, m_{it}; \tau) - \beta_k(\tau)k_{it} - \hat{\rho}(\hat{\Phi}(k_{it-1}, m_{it-1}; \tau) - \beta_k(\tau)k_{it-1})$. Then take the τth quantile of $\hat{\xi}_{it}$ as an estimate for $\hat{\alpha} + \hat{F}_\xi^{-1}(\tau)$
- The idea is to eliminate the Quantile Crossing problem by estimating ρ separately
- Since they do not vary over τ, slope estimates are often poor if using quantile regression, especially at the tails

Quantile Crossing Solution

- Other than to avoid the problem of the writing a moment condition that corresponds to the joint distribution of η_{it} and ξ_{it} I have no theoretical justification for using the RQQ for concentrating out moment conditions in ξ_{it}
- Generally, quantile crossing is a finite sample problem that goes away in the limit.
- Typically this problem makes charts hard to interpret as crossing can occur in certain areas of the sample space
- For example a predicted 95th percentile of the response could be smaller than the 90th percentile which is impossible
- The RQQ method enforces monotonicity, but is also inefficient compared to estimating quantiles simultaneously
- This method also looked good for Monte Carlo simulations (next slide) in this paper. However, applying this methodology to the data has been a pain

Proxy Variable Monte Carlo Results



Random Coefficient Model

- Therefore, we have the following identification condition

$$P[\hat{\xi}_{it}(\beta_k(\tau)) \leq \hat{F}_{\xi}^{-1}(\tau)|\mathcal{I}_{it-1}] = P[\hat{\xi}_{it}(\beta_k(\tau)) \leq \hat{F}_{\xi}^{-1}(\tau)] = \tau \quad (9)$$

  where $\mathcal{I}_{t-1}$ should include a constant and other instruments such as inputs determined in the prior period
- This can be represented by conditional moment restrictions

$$\mathbb{E}[\mathbb{1}\{\hat{\xi}_{it}(\beta_k(\tau)) - \hat{F}_{\xi}^{-1}(\tau) \leq 0\} - \tau|\mathcal{I}_{it-1}] = 0 \quad (10)$$

  where $\mathbb{1}\{\cdot\}$ is the indicator function. To estimate the production function parameters we use the unconditional moments implied by (10)

$$\mathbb{E}[Z_{it-1}(\mathbb{1}\{\hat{\xi}_{it}(\beta_k(\tau)) - \hat{F}_{\xi}^{-1}(\tau) \leq 0\} - \tau)] = 0 \quad (11)$$

- We propose smoothing the indicator function using the methodology proposed by Kaplan and Sun, 2016 and Castro, Galvao, Kaplan, and Liu, 2018 for nonlinear conditional quantile models.

Limitations and Extensions

- Proxy variable approach only works for linear production function models
- Nonlinear production functions would allow for non-Hicks neutral technology shocks
- A more general model could allow for location-scale effects of productivity or even more general distributional effects (e.g $\omega_{it} = \omega_{it}(\tau)$)
- The productivity process could be generalized to a quantile auto-regressive process, allowing for rich distributional effects of innovation shocks on future productivity
- As a consequence, we could document the different effects an innovation shock has on the firm-size distribution which we cannot do in the linear model.
- In order to do this, we apply Arellano and Bonhomme, 2016 and use nonlinear IV identification results from Hu and Schennach, 2008 and Hu and Shum, 2012 for dynamic models

Dynamic Panel Data Framework

- Consider a nonlinear model for a firm's gross-output production function

$$Y_{it} = F_t(K_{it}, L_{it}, M_{it}, \omega_{it}, \eta_{it}) \tag{12}$$

- Without loss of generality we normalize η_{it} to be uniformly distributed on the interval $[0, 1]$
- We assume F is strictly increasing in η_{it}
- Labor inputs are chosen to maximize current period profits and therefore are a function of current period state variables

$$L_{it} = \ell_t(K_{it}, \omega_{it}, \epsilon_{it}) \tag{13}$$

  where ϵ_{it} is iid and independent of current period state variables.
- We assume the labor demand function ℓ is strictly increasing in ϵ_{it} which is normalized to be uniformly distributed on the interval $U[0, 1]$

Dynamic Panel Data Framework

- Material inputs are chosen to maximize current period profits and therefore are a function of current period state variables

$$M_{it} = \mu_t(K_{it}, \omega_{it}, \varepsilon_{it}) \tag{14}$$

  where ε_{it} is iid and independent of current period state variables.
- We assume the material demand function μ is strictly increasing in ε_{it} which is normalized to be uniformly distributed on the interval $U[0,1]$.
- We can extend this to the case where labor is chosen prior to choosing material inputs in which case we would include L_{it} as a state variable in equation (14).
- Capital accumulates to the following generalized law of motion

$$K_{it} = \kappa_t(K_{it-1}, I_{it-1}, \upsilon_{it-1}) \tag{15}$$

  where I_{it-1} denotes firm investment in the prior period.

Dynamic Panel Data Framework

- We introduce a random error term υ_{it-1} which eliminates the deterministic relationship of capital with respect to previous period state and choice variables.
- We assume this error term is independent of the arguments in the capital accumulation law and that the function $\kappa$is strictly increasing in this term and is normalized to be uniformly distributed on the interval $U[0, 1]$.
- Productivity evolves according to the exogenous first order Markov process:

$$\omega_{it} = g(\omega_{it-1}, \xi_{it}) \tag{16}$$

- where $\xi_{i1}, \ldots, \xi_{iT}$ are independent uniform random variables which represent innovation shocks to productivity.
- We assume ω_{it} is monotonic in ξ_{it}
- The exogeneity of the productivity process can be relaxed when we consider productivity enhancing activities such as R&D similar to Doraszelski and Jaumandreu, 2013.

Dynamic Panel Data Framework

- We introduce a dynamic model of firm investment that is a slight modification of Ericson and Pakes, 1995
- In each period, a firm chooses investment to maximize its discounted future profits:

$$I_{it} = I^*(K_{it}, \omega_{it}, \zeta_{it}) = \underset{I_t \geq 0}{\operatorname{argmax}} \left[\Pi_t(K_{it}, \omega_{it}, \zeta_{it}) - c(I_{it}) + \beta \mathbb{E}\left[V_{t+1}(K_{it+1}, \omega_{it+1}, \zeta_{it+1}) | \mathcal{I}_t \right] \right], \tag{17}$$

  where $\pi_t(\cdot)$ is current period profits as a function of the state variables and an unobservable demand shock ζ_{it}.
- These are shocks to a firm's product demand which are privately observed by each firm and i.i.d across i and t.
- We assume these shocks are independent from the firm's state variables.
- Pakes, 1994 provides specific conditions for which the investment policy function is strictly increasing in its unobservable components.
- Without loss of generality, we normalize $\zeta_t \sim U[0, 1]$

Dynamic Panel Data Framework: Estimation

- Our empirical specification for the Markovian transitions of productivity, output, and capital evolution closely resemble Arellano, Blundell, and Bonhomme, 2017. We let $(y_{it}, k_{it}, l_{it}, m_{it}, i_{it})$ denote the logarithms of $(Y_{it}, K_{it}, L_{it}, M_{it}, I_{it})$ respectively.
  Output: Let age_{it} denote the age of firm i at time t. We specify the output equation as follows:

$$Q_t(y_{it}|k_{it}, l_{it}, m_{it}, \omega_{it}, \tau) = Q(y_{it}|k_{it}, l_{it}, m_{it}, \omega_{it}, \tau) = \sum_{j=1}^{J} \beta_j(\tau)\psi_j(k_{it}, l_{it}, m_{it}, \omega_{it}) \tag{18}$$

  Labor Input: We specify the labor input demand equation as follows:

$$Q_t(l_{it}|k_{it}, \omega_{it}, \tau) = Q(l_{it}|k_{it}, \omega_{it}, \tau) = \sum_{j=1}^{J} \gamma_j(\tau)\psi_j(k_{it}, \omega_{it}) \tag{19}$$

Dynamic Panel Data Framework: Estimation

Material Input: We specify the material input demand equation as follows

$$\begin{aligned} Q_t(m_{it}|k_{it},\omega_{it},\tau) &= Q(m_{it}|k_{it},\omega_{it},\tau) \\ &= \sum_{j=1}^{J}\delta_j(\tau)\psi_j(k_{it},\omega_{it}) \end{aligned} \tag{20}$$

Investment Demand: We specify the investment demand equation as

$$\begin{aligned} i_t = \iota_t(k_{it},\omega_{it},\zeta_{it}) &= \iota(k_{it},\omega_{it},\zeta_{it}) \\ &= \iota_0 + \sum_{j=1}^{J}\iota_j\psi_j(k_{it},\omega_{it}) + \zeta_{it}, \end{aligned} \tag{21}$$

where $E[\zeta_{it}|k_{it},\omega_{it}] = 0$

- The above specification is a nonlinear regression model.
- The conditional mean zero assumption provides the normalization assumption required

Dynamic Panel Data Framework: Estimation

Persistent Productivity: We specify productivity to transition according to:

$$Q_t(\omega_{it-1}, \tau) = Q(\omega_{it-1}, \tau) = \sum_{j=1}^{J} \rho_j(\tau)\psi_j(\omega_{it-1}) \tag{22}$$

- I decided to remove age dependence in the above specifications for two reasons
  1. Most results show age has negligent effects on production such as Olley and Pakes, 1996
  2. I cannot find age in the Compustat database. Not sure about using IPO date as initial age because most companies are private before they go public. Other datasets have similar concern. Is date appearing in the panel equal to initial age? Maybe not.

I specify a parametric specification for initial productivity

$$\omega_{i1} \sim N(\mu, \sigma^2) \tag{23}$$

These parameters are easy to estimate in the M-step of the algorithm

Dynamic Panel Data Framework: Estimation

- To ease notation, we let the finite and functional parameters be indexed by a finite dimensional parameter vector θ.
- We model the functional parameters using Wei and Carroll, 2009 and Arellano and Bonhomme, 2016. For example, the function $\beta_j(\tau_q)$ is modeled as a piecewise-polynomial interpolating splines on a grid $[\tau_1, \tau_2], [\tau_3, \tau_4], \ldots, [\tau_{Q-1}, \tau_Q]$, contained in the unit interval and is constant on $[0, \tau_1]$ and $[\tau_Q, 1)$
- The intercept coefficient β_0 is specified as the quantile of an exponential distribution on $(0, \tau_1]$ (indexed by λ^-) and $[\tau_{Q-1}, 1)$ (indexed by λ^+).
- The remaining functional parameters are modeled similarly. We take $Q = 11$ and $\tau_q = \frac{q}{Q+1}$.
- Following Arellano, Blundell, and Bonhomme, 2017 we parameterize the distribution of ζ_{it} to be log-normal so we set, for example, $\iota_0(\tau_q) = \iota_0 + \sigma_\zeta \Phi^{-1}(\tau_q)$. In the following section we outline the model's restrictions and a feasible estimation strategy.

Dynamic Panel Data Framework: Estimation

- Let $\Psi_\tau(u) = \tau - \mathbb{1}\{u < 0\}$ denote the first derivative of the quantile check function $\psi_\tau(u) = (\tau - \mathbb{1}\{u < 0\})u$.
- The following conditional moment restrictions hold as an implication of the conditional independence restrictions. Therefore, we estimate the parameters of interest from the following conditional moment restrictions.
- To fix ideas, we focus on how to estimate the production function and investment equation

$$\mathbb{E}\left[\Psi_{\tau_q}(\eta_{it})\middle| k_{it}, l_{it}, m_{it}\right] = 0 \tag{24}$$

$$\mathbb{E}\left[\zeta_{it}\middle|\omega_{it}, k_{it}\right] = 0 \tag{25}$$

Dynamic Panel Data Framework: Estimation

Rewriting these moment conditions as:

$$\mathbb{E}\left[\Psi_{\tau_q}(\eta_{it})\Big| k_{it}, l_{it}, m_{it}, \omega_{it}\right] = \mathbb{E}\left[\Psi_{\tau_q}(y_{it} - \sum_{j=1}^{J} \bar{\beta}_j(\tau_q)\psi_j(k_{it}, l_{it}, m_{it}, \omega_{it}))\Big| k_{it}, l_{it}, m_{it}, \omega_{it}\right] = 0 \tag{26}$$

and

$$\mathbb{E}\left[\zeta_{it}\Big| k_{it}, \omega_{it}\right] = \mathbb{E}\left[i_{it} - \bar{\iota}_0 - \sum_{j=1}^{J} \bar{\iota}_j\psi_j(k_{it}, \omega_{it})\Big| k_{it}, \omega_{it}\right] = 0 \tag{27}$$

- Here $\bar{\beta}_j(\tau_q), \bar{\iota_0}$ and $\bar{\iota_j}$ denote the true values of $\beta_j(\tau_q), \iota_0$ and ι_j for $j \in \{1, \ldots, J\}$ and $q \in \{1, \ldots, Q\}$.
- Clearly, estimating the above conditional moment restrictions are infeasible due to the unobserved productivity component.
- Therefore, we use the following unconditional moment restrictions and posterior distributions for ω_{it} to integrate out the unobserved productivity.
- Due to the law of iterated expectations we now have the following integrated moment conditions:

$$\mathbb{E}\left[\int \left(\Psi_{\tau_q}(y_{it} - \sum_{j=1}^{J} \bar{\beta}_j(\tau_q)\psi_j(k_{it}, l_{it}, m_{it}, \omega_{it})) \otimes \begin{pmatrix} k_{it} \\ l_{it} \\ m_{it} \\ \omega_{it} \end{pmatrix}\right) f_i(\omega_{it}; \bar{\theta}) d\omega_{it}\right] = 0 \tag{28}$$

Dynamic Panel Data Framework: Estimation

and for investment

$$\mathbb{E}\left[\int \left(\left(i_{it} - \bar{\iota}_0 - \sum_{j=1}^{J} \bar{\iota}_j \psi_j(k_{it}, \omega_{it})\right) \otimes \begin{pmatrix} 1 \\ k_{it} \\ \omega_{it} \end{pmatrix}\right) f_i(\omega_{it}; , \bar{\theta}) d\omega_{it}\right] = 0, \tag{29}$$

where $\bar{\theta}$ denotes the true values of θ. The posterior distribution is specified as:

$$\begin{aligned} f_i(\omega_{it}; \bar{\theta}) &= f(\omega_{it} | y_{it}, k_{it}, l_{it}, m_{it}, i_{it}; \bar{\theta}) \propto \\ &\prod_{t=1}^{T} f(y_{it} | k_{it}, l_{it}, m_{it}, \omega_{it}; \bar{\theta}) f(l_{it} | k_{it}, \omega_{it}; \bar{\theta}) f(m_{it} | k_{it}, \omega_{it}; \bar{\theta}) \\ &\times f(i_{it} | k_{it}, \omega_{it}; \bar{\theta}) \prod_{t=2}^{T} f(\omega_{it} | \omega_{it-1}; \bar{\theta}) f(\omega_{i1}; \bar{\theta}) \end{aligned} \tag{30}$$

Dynamic Panel Data Framework: Estimation

- The posterior density in equation (30) is a closed-form expression when using piecewise linear splines for $\theta(\cdot)$.
- The estimation is an Expectation Maximization (EM) algorithm. In Arellano and Bonhomme, 2016 and Arellano, Blundell, and Bonhomme, 2017, the "M-step" is performed using quantile regression.
- Given an initial parameter value $\hat{\theta}^0$. Iterate on $s = 0, 1, 2, \ldots$ in the following two-step procedure until converge to a stationary distribution:

1. *Stochastic E-Step*: Draw M values $\omega_i^{(m)} = (\omega_{i1}^{(m)}, \omega_{i2}^{(m)}, \ldots, \omega_{iT}^{(m)})$ from

$$f_i(\omega_{it}; \hat{\theta}^{(s)}) = f(\omega_{it} | y_{it}, k_{it}, l_{it}, m_{it}, i_t, ; \hat{\theta}^{(s)}) \propto$$

$$\prod_{t=1}^{T} f(y_{it} | k_{it}, l_{it}, m_{it}, \omega_{it}; \hat{\theta}^{(s)}) f(l_{it} | k_{it}, \omega_{it}; \hat{\theta}^{(s)}) f(m_{it} | k_{it}, \omega_{it}; \hat{\theta}^{(s)})$$

$$\times f(i_{it} | k_{it}, \omega_{it}; \hat{\theta}^{(s)}) \prod_{t=2}^{T} f(\omega_{it} | \omega_{it-1}; \hat{\theta}^{(s)}) f(\omega_{i1}; \hat{\theta}^{(s)})$$

Dynamic Panel Data Framework: Estimation

2. *Maximization Step*: For $q = 1, \ldots, Q$, solve

$$\hat{\boldsymbol{\beta}}(\tau_q)^{(s+1)} = \underset{\boldsymbol{\beta}(\tau_q)}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} \Psi_{\tau_q}\left(y_{it} - \sum_{j=1}^{J} \beta_j(\tau_q)\psi_j(k_{it}, l_{it}, m_{it}, \omega_{it}^{(m)}) \right)$$

$$\hat{\boldsymbol{\iota}}^{(s+1)} = \underset{\boldsymbol{\iota}}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} \left(i_{it} - \iota_0 - \sum_{j=1}^{J} \iota_j \psi_j(k_{it}, \omega_{it}^{(m)}) \right)^2$$

- The parameters of the production function equation in (29) can be estimated using a nonlinear regression for a given draw of $\omega_{it}^{(m)}$.
- Then, the variance of the shock ζ_{it} can be estimated using

$$\hat{\sigma_\zeta}^2 = \frac{1}{NTM} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} \left[\left(i_{it} - \hat{\iota_0} - \sum_{j=1}^{J} \hat{\iota}_j \psi_j(k_{it}, \omega_{it}^{(m)}) \right)^2 \right] \qquad (31)$$

so that $\hat{\iota_0}(\tau_q) = \hat{\iota_0} + \hat{\sigma_\zeta} \Phi^{-1}(\tau_q)$

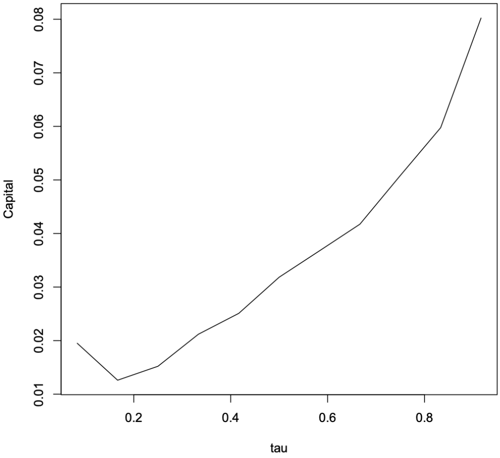
Empirical Implementation

- I used a tensor product of Hermite polynomials for ψ_j
- For example in the output equation, I specify $(1, 2, 1, 2)$ for capital, labor, materials, and productivity respectively (36 parameters in total)
- In their paper, Arellano and Bonhomme, 2016 and Arellano, Blundell, and Bonhomme, 2017 take $M = 1$ which is makes this estimation algorithm a Stochastic EM Algorithm (stEM) as opposed to a Monte Carlo EM Algorithm (MCEM)
- A stEM algorithm means that with $M = 1$ a single draw of the unobservables is used to estimate the model which can be done using basic quantile regression
- A large M is computationally costly. Even though a large M could be used to "smooth" the moment conditions
- I tried variations of stochastic gradient descent algorithms that although had decent convergence results, were too costly to use in practice.

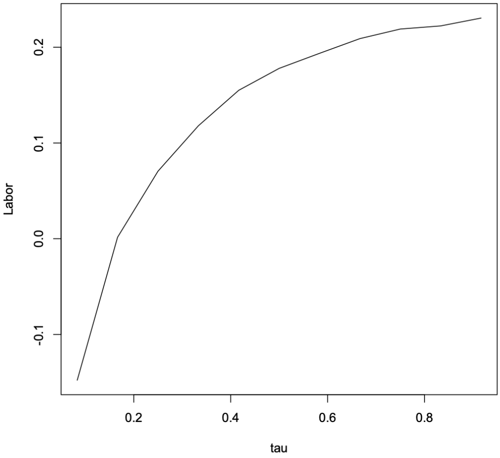
The Stochastic EM Algorithm

- The sequence of maximizers $\hat{\theta}^{(s)}$ is a time-homogeneous Markov chain
- Nielsen, 2000 shows that under certain conditions (e.g aperiodic and irreducible) than this Markov chain is ergodic and will converge to the stationary distribution of the Markov chain
- They further show that this distribution is asymptotically normal with a variance term that decreases in M.
- The asymptotic distribution is much harder to characterize in this quantile based approach although Arellano and Bonhomme, 2016 make some effort
- The paper by Nielsen, 2000 provides many possible improvements of the stochastic EM algorithm such as multiple starting points or multiple maximizations
- In the next few slides ill present a few preliminary results from a linear random coefficient model for US NAICS 33

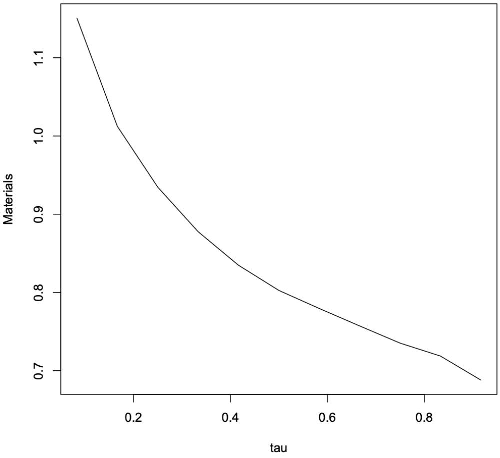
Preliminary Results: Capital Elasticity



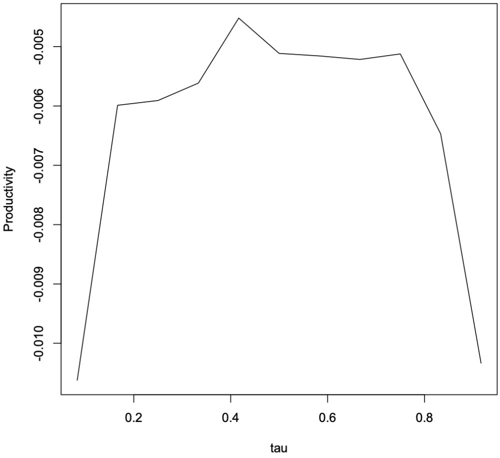
Preliminary Results: Labor Elasticity



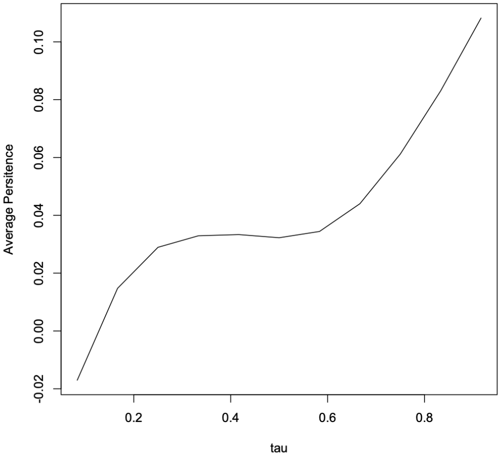
Preliminary Results: Materials Elasticity



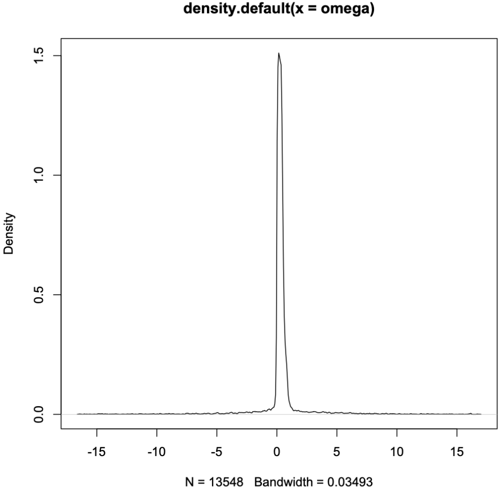
Preliminary Results: Productivity



Preliminary Results: Average Persistence



Preliminary Results: Productivity Distribution



Addressing a Sample Selection Problem

- Much of the time consumed writing this algorithm was taking into consideration unbalanced panel data
- Both Arellano and Bonhomme, 2016 and Arellano, Blundell, and Bonhomme, 2017 use balanced panel data in their empirical applications
- Olley and Pakes, 1996 address anomalies in production function coefficients estimated from balanced panels.
- They find that going from the balanced sample to the full sample more than doubles the capital coefficient and decreases the labor coefficient by about 20 percent
- Productivity is a major determinant of whether or not a plant exits the industry
- If firms exit decisions depend on beliefs of future productivity which are partially determined by current productivity then a balanced panel will be selected on the basis of the unobserved productivity realizations

Addressing a Sample Selection Problem

- The exit rule decision is given by

$$\chi_t = \begin{cases} 1 & \text{if } \omega_t \geq \bar{\omega}_t(a_t, k_t) \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

where a_t denotes another state variable of the firm (age)

- Then the second stage of OP

$$\begin{aligned} &\mathbb{E}[y_{it} - \hat{\beta}_l l_{it} | \chi_t = 1, \mathcal{I}_{it}] \\ &\quad = \beta_0 + \beta_k k_{it} + \beta_a a_{it} + \mathbb{E}[\omega_{it} | \chi_t = 1, \mathcal{I}_{it}] \end{aligned} \tag{33}$$

where

$$\begin{aligned} \mathbb{E}[\omega_{it} | \chi_t = 1, \mathcal{I}_{it}] &= \mathbb{E}[\omega_{it} | \omega_t \geq \bar{\omega}_t(a_t, k_t), \mathcal{I}_{it}] \\ &\int_{\bar{\omega}_t(a_t,k_t)}^{\infty} \omega_{it} \frac{p(\omega_{it}|\omega_{it-1})}{\int_{\bar{\omega}_t(a_t,k_t)}^{\infty} p(\omega_{it}|\omega_{it-1}) d\omega_{it}} d\omega_{it} \\ &= g(\omega_{it-1}, \bar{\omega}_t(a_{it}, k_{it})) \end{aligned} \tag{34}$$

Addressing a Sample Selection Problem

- In OP, the first stage results conditional on parameters can be used as a proxy for ω_{it-1}
- $\bar{\omega}_t(a_{it}, k_{it})$ can be controlled using observed data on entry and exit

$$\begin{aligned} P(\chi_t = 1 | \mathcal{I}_{it-1}) &= P(\omega_{it} \geq \bar{\omega}_t(a_{it}, k_{it}) | \mathcal{I}_{it-1}) \\ &= \varphi(\omega_{it-1}, \bar{\omega}_t(a_{it}, k_{it})) \\ &= \varphi(\omega_{it-1}, a_{it}, k_{it}) \\ &= \varphi(i_{it-1}, k_{it-1}, a_{it-1}) = P_t \end{aligned} \tag{35}$$

  which can be nonparametrically estimated using probit with a flexible polynomial of $(i_{it-1}, k_{it-1}, a_{it-1})$ or by using kernel methods
- Lastly, we can invert the last line of the above equation to express $\bar{\omega}_t(a_{it}, k_{it})$ as a function of ω_{it-1} and P_t so that

$$g(\omega_{it-1}, \bar{\omega}_t(a_{it}, k_{it})) = \tilde{g}(\omega_{it-1}, P_t) \tag{36}$$

Addressing a Sample Selection Problem

- Can we apply a similar correction using Arellano and Bonhomme, 2017 for quantile models?

$$P(\omega_{it} \leq g(\omega_{it-1}, \xi_{it}) | \chi_t = 1, \mathcal{I}_{it-1}) = P(\xi_{it} \leq \tau | \bar{\omega}_t(a_{it}, k_{it}) \leq \omega_{it}, \mathcal{I}_{it-1}) \tag{37}$$

- They use a copula based method to estimate the dependence between the error in the selection equation and the error in the output equation
- Not sure if that method can be applied here or if I should try to see if the OP selection correction can be applied here
- This is something to consider in the far future

Conclusion

- Identification strategy is not new, most economists have avoided it due to injectivity assumptions and panel time length requirements
- These data requirements become even more stringent if other unobservables (firm fixed effects) are added
- Implementation in progress
- The estimation procedure is novel and allows us to document a variety of heterogeneous quantile marginal effects, for example, the average quantile marginal effect of labor would be

$$\phi_t(k_t, l_t, m_t, \tau) = \mathbb{E}\left[\frac{\partial Q(y_t | k_t, l_t, m_t, \omega_t, \tau)}{\partial l_t}\right] \tag{38}$$

- Distributional effects of productivity

$$\phi_t(k_t, l_t, m_t, \omega_t, \tau) = \frac{\partial Q(y_t | k_t, l_t, m_t, \omega_t, \tau)}{\partial \omega_t} \tag{39}$$

Conclusion

- Dynamic effects of innovation shocks on production

$$\frac{\partial}{\partial \xi_{it}} \left[\frac{\partial Q(y_t | k_t, l_t, m_t, Q_t(\omega_{t-1}, \xi_t), \tau)}{\partial \omega_t} \right] \Bigg|_{\xi = \tau_\xi} = \phi_t(k_t, l_t, m_t, Q_t(\omega_{t-1}, \tau_\xi), \tau) \frac{\partial Q_t(\omega_{t-1}, \tau_\xi)}{\partial \xi} \tag{40}$$

- "Impulse response" functions can be written as

$$\frac{\partial}{\partial \omega_{t-1}} \left[\frac{\partial Q_t(\omega_{t-1}, \tau_\xi)}{\partial \xi} \right] \tag{41}$$

  and can be approximated by finite differences across τ_ξ to study the asymmetric impacts of innovation shocks at different points on the output distribution over firm age

- Many more interesting effects can be formalized and calculated