A Dynamic Framework for Identification and Estimation of Nonlinear Production Functions

Justin Doty

September 3, 2021

Introduction

- This paper identifies and estimates a non-linear production function with unobserved heterogeneity
- Allows for interactions in the unobservable (productivity) with inputs, in a sense, allowing for non-Hicks neutral productivity shocks
- The conditional output distribution is allowed to vary with productivity
- Inputs can contain unobservable demand shocks which would violate identification arguments of OP, LP, ACF
- The identification argument does not rely on the control function approach so there are no identification issues between using value-added and gross-output production functions
- A more flexible model for productivity and non-linear persistence
- Productivity can be allowed to depend on the magnitude and sign of current and future innovation shocks

The Production Function

- Consider a nonlinear model for a firm's gross-output production function (in logs)

$$y_{it} = f_t(k_{it}, l_{it}, m_{it}, \omega_{it}, \eta_{it}) \tag{1}$$

- Allows for non-linear interactions between inputs and unobserved productivity
- Assume the following

Assumption 1 (Production Function)

1. *The unanticipated production shocks η_{it} are iid over firms and time.*
2. *The unanticipated production shock η_{it} follows a standard uniform distribution independent of $(\omega_{it}, k_{it}, l_{it}, m_{it})$.*
3. $\tau \to Q_t(y_{it}|k_{it}, l_{it}, m_{it}, \omega_{it}, \tau)$ *is strictly increasing on* $(0, 1)$.

Productivity

- Productivity evolves according to an exogenous first-order Markov process given by
$$\omega_{it} = h_t(\omega_{it-1}, \xi_{it}) \tag{2}$$
where $\xi_{i1}, \ldots, \xi_{iT}$ are independent uniform random variables which represent innovation shocks to productivity
- The function h is a nonlinear function that allows the persistence in productivity in firms to be nonlinear across different quantiles
- The process is exogenous, but this can relaxed to allow for productivity enhancing activities such as R&D or exporting status

Assumption 2 (Productivity)

1. *The productivity innovation shocks* ξ_{it} *are iid across firms and time.*
2. ξ_{it} *follows a standard uniform distribution independent of previous period productivity* ω_{it-1}*.*
3. $\tau \to Q_t(\omega_{it}|\omega_{it-1}, \tau)$ *is strictly increasing on* $(0,1)$*.*

Flexible Inputs

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

$$l_{it} = \ell_t(k_{it}, \omega_{it}, \epsilon_{\ell,it}) \tag{3}$$

$$m_{it} = \mu_t(k_{it}, \omega_{it}, \epsilon_{m,it}) \tag{4}$$

- $\epsilon_{l,it}$ and $\epsilon_{m,it}$ are iid unobservable input demand shocks that are assumed to be independent of current period state variables
- In the control function approach, with material inputs as a proxy, this function could not be inverted as an expression of productivity only
- This framework is similar to that of Hu, Huang, and Sasaki, 2019
- This can also be extended to the case where labor is chosen prior to material inputs or if labor has adjustment frictions

Flexible Inputs

Assumption 3 (Flexible Inputs)

1. *The unobserved input demand shocks* $\epsilon_{l,it}$ *and* $\epsilon_{m,it}$ *are iid across firms and time*
2. $\epsilon_{l,it}$ *and* $\epsilon_{m,it}$ *follow a standard uniform distribution independent of* (k_{it}, ω_{it})
3. $\tau \to Q_t(l_{it}|k_{it}, \omega_{it}, \tau)$ *and* $\tau \to Q_t(m_{it}|k_{it}, \omega_{it}, \tau)$ *are strictly increasing on* $(0, 1)$

- The monotonicity assumption is non-trivial
- Similar to conditions in Levinsohn and Petrin, 2003 this requires that that for fixed capital and productivity, the marginal productivity of labor and materials increases with higher demand shocks

Capital and Investment

- Capital accumulates to the following generalized law of motion

$$K_{it} = \kappa_t(K_{it-1}, I_{it-1}, \upsilon_{it-1}) \tag{5}$$

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

- Eliminates the deterministic relationship of capital with respect to previous period state and choice variables
- Assume this error term is independent of the arguments in the capital accumulation law
- 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:

$$I_{it} = I^*(K_{it}, \omega_{it}, \zeta_{it}) = \underset{I_t \geq 0}{\text{argmax}} \left[\Pi_t(K_{it}, \omega_{it}) - c(I_{it}, \zeta_{it}) + \beta \mathbb{E}\left[V_{t+1}(K_{it+1}, \omega_{it+1}, \zeta_{it+1}) | \mathcal{I}_t \right] \right], \tag{6}$$

Capital and Investment

- $\pi_t(\cdot)$ is current period profits as a function of the state variables ζ_{it} can be investment cost reduction innovation.

Assumption 4 (Capital Accumulation and Investment)

1. *The unobserved investment demand shocks ζ_{it} is iid across firms and time*
2. *ζ_{it} follows a standard uniform distribution independent of (k_{it}, ω_{it})*
3. *The production shock η_{it} and ζ_{it} are independent conditional on $(k_{it}, l_{it}, m_{it}, \omega_{it})$. In addition, υ_{it} is independent of η_{it} conditional on $(k_{it}, l_{it}, m_{it}, \omega_{it})$*
4. *$\tau \to Q_t(i_{it}|K_{it}, \omega_{it}, \tau)$ is strictly increasing on $(0, 1)$*

- Pakes, 1994 provides specific conditions for which the investment policy function is strictly increasing in its unobservable components.
- Satisfied if (6) is supermodular in I_t and ζ_{it} for all (K_{it}, ω_{it})

Nonparametric Identification

- I show that the conditional densities corresponding to production, inputs, and productivity are nonparametrically identified
- Identification is similar to Hu, Huang, and Sasaki, 2019
- Let $Z_t = (l_t, k_t, m_t, k_{t+1})$ denote conditioning variables

Assumption 5 (Conditional Independence)

1. $f(y_t|y_{t+1}, l_t, \omega_t, Z_t) = f(y_t|\omega_t, Z_t)$
2. $f(l_t|y_{t+1}, \omega_t, Z_t) = f(l_t|\omega_t, Z_t)$

- First equality states that conditional on ω_t and Z_t, y_{t+1} and l_t do not provide any additional information about y_t
- Second equality states that conditional on ω_t and Z_t, y_{t+1} does not provide any additional information about l_t
- Satisfied by mutual independence assumptions on η_t and ζ_t conditional on $(\omega_t, k_t, l_t, m_t)$

Nonparametric Identification

- Begin by relating a conditional density as a function of observable to densities containing unobserved productivity
- Using the conditional independence assumption, I can write

$$f_{y_t, I_t | y_{t+1}, Z_t} = \int f_{y_t | Z_t, \omega_t} f_{I_t | Z_t, \omega_t} f_{\omega_t | y_{t+1}, Z_t} d\omega_t \tag{7}$$

- The identification strategy follows HS by using a eigenvalue-eigenfunction decomposition of integral operators of (7)

Definition 1

(Integral Operator) Let a and b denote random variables with supports $\mathcal{A}$ and $\mathcal{B}$. Given two corresponding spaces $\mathcal{G}(\mathcal{A})$ and $\mathcal{G}(\mathcal{B})$ of functions with domains $\mathcal{A}$ and $\mathcal{B}$, let $L_{b|a}$ denote the operator mapping $g \in \mathcal{G}(\mathcal{A})$ to $L_{b|a} g \in L_{b|a} \mathcal{G}(\mathcal{B})$ defined by

$$[L_{b|a} g](b) \equiv \int_{\mathcal{A}} f_{b|a}(b|a) g(a) da,$$

where $f_{b|a}$ denotes the conditional density of b given a.

Nonparametric Identification

- The observed density in (7) can be written in operator notation

$$L_{y_t,I_t|y_{t+1},Z_t} = L_{y_t|Z_t,\omega_t}\Delta_{I_t|Z_t,\omega_t}L_{\omega_t|y_{t+1},Z_t} \tag{8}$$

- Will show that under a set of assumptions, the conditional density is identified from an eigenvalue-eigenfunction decomposition of (8)

Assumption 6 (Injectivity)

The operators $L_{y_t|Z_t,\omega_t}$ and $L_{y_{t+1}|Z_t,\omega_t}$ are injective

- The above assumption allows us to take inverses of the operators.
- Consider the operator $L_{y_t|Z_t,\omega_t}$, injectivity of this operator can be interpreted as its corresponding density $f_{y_t|Z_t,\omega_t}(I_t|K_t,\omega_t)$ having sufficient variation in ω_t given Z_t.
- This assumption is often phrased as completeness condition in the nonparametric IV literature on the density $f_{y_t|Z_t,\omega_t}(y_t|Z_t,\omega_t)$.

Nonparametric Identification

Assumption 7 (Uniqueness)

For any $\bar{\omega}_t, \tilde{\omega}_t \in \Omega$, the set $\{f_{I|\omega,Z}(I_t|\bar{\omega}_t, Z_t) \neq f_{I|\omega,Z}(I_t|\tilde{\omega}_t, Z_t)\}$ has positive probability whenever $\bar{\omega}_t \neq \tilde{\omega}_t$

- This assumption is relatively weak
- Satisfied if there is conditional heteroskedasticity in $f_{I|\omega,Z}$
- Satisfied if any functional of its distribution is strictly increasing in ω_t
- I assume $E[I_t|\omega_t, Z_t]$ is strictly increasing in ω_t
- Similar to the invertibility condition in Olley and Pakes, 1996
- Censoring is not an issue in the data used in the application, but is a common problem in some datasets

Nonparametric Identification

Assumption 8 (Normalization)

There exists a functional Γ such that $\Gamma[f_{y|\omega,Z}(y_t|\omega_t, Z_t)] = \omega_t$

- This functional does not need to be known
- Sufficient to consider a known function of the data distribution as shown by Arellano and Bonhomme, 2016
- In my empirical application, I consider a nonseparable translog production function
- The assumption can be satisfied by the normalization $E[y_t|\omega_t, 0] = \omega_t$
- For more generalized production functions, if $E[y_t|\omega_t, Z_t]$ is strictly increasing in ω_{it}, then one could normalize $\omega_t = E[y_t|\omega_t, Z_t]$
- These restrictions are easily adaptable in estimation as it amounts to centering the coefficients in the model

Nonparametric Identification

Theorem 1 (Identification)

Under Assumptions 5, 6, 7, and 8, given the observed density $f_{y_t, I_t | y_{t+1}, Z_t}$, the equation

$$f_{y_t, I_t | y_{t+1}, Z_t} = \int f_{y_t | \omega_t, Z_t} f_{I_t | \omega_t, Z_t} f_{\omega_t | y_{t+1}, Z_t} d\omega_t \tag{9}$$

admits a unique solution for $f_{y_t | \omega_t, Z_t}$, $f_{I_t | \omega_t, Z_t}$ and $f_{\omega_t | y_{t+1}, Z_t}$

- The proof follows using Hu and Schennach, 2008
- However it does not directly identify the Markov transition function for productivity $f_{\omega_{it+1} | \omega_{it}}(\omega_{it+1} | \omega_{it})$

Nonparametric Identification

Corollary 1 (Identification of Markov Process: Stationarity Case)

Suppose that the production function is stationary, $f_{y_t|\omega_t,Z_t} = f_{y_1|\omega_1,Z_1} \forall t \in \{1, \cdots, T\}$. Then, under Assumptions 5, 6, 7, and 8, the observed density $f_{y_t,I_t|y_{t+1},Z_t}$ uniquely determines the density $f_{\omega_{t+1}|\omega_t} \forall t \in \{1, \ldots, T-1\}$

Corollary 2 (Identification of Markov Process: Non-Stationarity Case)

Under Assumptions 5, 6, 7, and 8, the observed density $f_{y_{t+1},I_{t+1}|y_{t+2},Z_{t+1}}$ uniquely determines the density $f_{\omega_{t+1}|\omega_t} \forall t \in \{1, \ldots, T-2\}$

Econometric Procedure: Production

- The production function is specified as Translog with non-Hicks neutral effects

$$
\begin{aligned}
&Q_t(y_{it}|k_{it}, l_{it}, m_{it}, \omega_{it}; \beta(\tau)) = \\
&\quad \gamma_0(\tau) + (\gamma_k(\tau) + \sigma_k(\tau)\omega_{it})k_{it} + (\gamma_l(\tau) + \sigma_l(\tau)\omega_{it})l_{it} + (\gamma_m(\tau) + \sigma_m(\tau)\omega_{it})m_{it} \\
&\quad + (\gamma_{kl}(\tau) + \sigma_{kl}(\tau)\omega_{it})k_{it}l_{it} + (\gamma_{lm}(\tau) + \sigma_{lm}(\tau)\omega_{it})l_{it}m_{it} + (\gamma_{km}(\tau) + \sigma_{km}(\tau)\omega_{it})k_{it}m_{it} \\
&\quad + (\gamma_{kk}(\tau) + \sigma_{kk}(\tau)\omega_{it})k_{it}^2 + (\gamma_{ll}(\tau) + \sigma_{ll}(\tau)\omega_{it})l_{it}^2 + (\gamma_{mm}(\tau) + \sigma_{mm}(\tau)\omega_{it})m_{it}^2 + \sigma_\omega(\tau)\omega_{it}
\end{aligned}
\tag{10}
$$

- Similar model was estimated by Ackerberg and Chen (2015)
- In my approach I can simulate productivity from estimated intital conditions and Markov process to compute average derivative effects
- Provides a better picture of heterogeneity instead of reporting individual coefficients

Econometric Procedure: Productivity

- I specify productivity using 3rd order polynomial

$$Q_t(\omega_{it}|\omega_{it-1},\tau) = \rho_0(\tau)+\rho_1(\tau)\omega_{it-1}+\rho_2(\tau)\omega_{it-1}^2+\rho_3(\tau)\omega_{it-1}^3. \quad (11)$$

- Initial productivity, $\omega_{i1} \sim N(\mu_{\omega_1}, \sigma_{\omega_1}^2)$
- I can also consider the case where productivity may evolve endogenously
- Consider a productivity enhancing activity such as R&D (in logs) as

$$r_{it} = R_t(k_{it}, \omega_{it}, \varrho_{it}), \quad (12)$$

  where ϱ_{it} is a shock to R&D that is independent from the state variables and normalized to be $U(0,1)$.
- In this case, it can be included in the process for productivity as

$$Q_t(\omega_{it}|\omega_{it-1}, r_{it-1}, \tau) = \sum_{j=1}^{J} \rho_j(\tau)\phi_{\omega,j}(\omega_{it-1}, r_{it-1}) \quad (13)$$

  where $\phi_{\omega,j}$ can capture interaction effects between productivity and R&D expenditure.

Econometric Procedure: Productivity

- Olley and Pakes, 1996 derive an optimal investment demand function and exit rule from a Markov perfect Nash equilibrium
- Firm's exit at some threshold productivity determined by their capital stock
- Firms with larger capital stocks can expect larger future returns for any given level of current productivity
- These firms can survive at lower productivity realizations
- This leads to a negative bias in the capital coefficients
- I follow OP and Dermirer, 2020 in deriving a correction to selection
- The exit rule is written as

$$\chi_{it} = \begin{cases} 1 & \text{if} \quad \omega_{it} \geq \underline{\omega}_t(k_{it}) \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where $\underline{\omega}_t(k_{it})$ is the threshold productivity level

Econometric Procedure: Productivity

- The exit rule can be written as

$$\begin{aligned} h_t(\omega_{it-1},\xi_{it}) &\geq \underline{\omega}_t(k_{it}), \\ \xi_{it} &\geq h_t^{-1}(\omega_{it-1}, \underline{\omega}_t(k_{it})), \\ \xi_{it} &\geq h_t^{-1}(\omega_{it-1}, k_{it}), \\ \xi_{it} &\geq \underline{\omega}_t(\omega_{it-1}, k_{it}) \end{aligned} \tag{15}$$

- Then if ξ_{it} is independent of (k_{it}, ω_{it-1}), $\xi_{it}|(k_{it}, \omega_{it-1}) \sim U(0,1)$
- The cutoff for which firms stay in operation can be estimated from

$$\underline{\omega}_t(\omega_{it-1}, k_{it}) = \text{Prob}(\chi_{it} = 1|\omega_{it-1}, k_{it}) \equiv p(\omega_{it-1}, k_{it}) \tag{16}$$

- Firms that receive an innovation shock greater than $p(\omega_{it-1}, k_{it})$ continue to operate
- Then, the distribution of $\xi_{it}|(k_{it}, \omega_{it-1}, \chi_{it} = 1)$ is

$$\xi_{it}|(k_{it}, \omega_{it-1}, \chi_{it} = 1) \sim U(p(\omega_{it-1}, k_{it}), 1) \tag{17}$$

Econometric Procedure: Productivity

- Consider a simple process for productivity $\omega_{it} = \rho(\xi_{it})\omega_{it-1}$

$$\begin{aligned} \text{Prob}(\omega_{it} \leq \rho(\tau)\omega_{it-1} | \omega_{it-1}, k_{it}, \chi_{it} = 1) \\ = \text{Prob}(\xi_{it} \leq \tau | \omega_{it-1}, k_{it}, \chi_{it} = 1) \\ = \frac{\tau - p(\omega_{it-1}, k_{it})}{1 - p(\omega_{it-1}, k_{it})} \equiv G(\tau, p) \end{aligned} \tag{18}$$

- If ω_{it} is known, then $\rho(\tau)$ is the solution to a rotated quantile regression problem

$$\hat{\rho}(\tau) = \underset{\rho}{\text{argmin}} \sum_{i=1}^{N} \sum_{t=2}^{T} \chi_{it} \Big[G(\tau, \hat{p})(\omega_{it} - \rho\omega_{it-1})^{+} + (1 - G(\tau, \hat{p}))(\omega_{it} - \rho\omega_{it-1})^{-} \Big] \tag{19}$$

where $a^{+} = max(a, 0)$ and $a^{-} = max(-a, 0)$

- $p(\omega_{it-1}, k_{it})$ can be estimated using a probit model
- This approach is similar to that of Arellano and Bonhomme, 2017

Econometric Procedure: Flexible Inputs

- I specify the labor input demand equation as follows:

$$Q_t(l_{it}|k_{it}, \omega_{it}, \tau) = \sum_{j=1}^{J} \alpha_{l,j}(\tau)\phi_{l,j}(k_{it}, \omega_{it}) \tag{20}$$

  where $\phi_{l,j}$ can be another non-linear function

- Material inputs are specified as

$$Q_t(m_{it}|k_{it}, \omega_{it}, \tau) = \sum_{j=1}^{J} \alpha_{m,j}(\tau)\phi_{m,j}(k_{it}, \omega_{it}) \tag{21}$$

- Again, $\phi_{m,j}$ can be a non-linear function where l_{it} could be added as an additional state variable

Econometric Procedure: Investment

- The investment demand function (in logs) is specified as

$$i_{it}^* = Q_t(i_{it}|k_{it}, \omega_{it}, \tau) = \sum_{j=1}^{J} \delta_j(\tau)\phi_{\iota,j}(k_{it}, \omega_{it}) \tag{22}$$

  where $\phi_{\iota,j}$ is specified similarly as the labor and material input decision rule.

- In the case where investment is censored, I can write

$$Q_\tau(i_{it}|k_{it}, \omega_{it}) = \max\{0, \sum_{j=1}^{J} \delta_j(\tau)\phi_{\iota,j}(k_{it}, \omega_{it})\}, \tag{23}$$

  due to the equivariance properties of quantiles

- The censored quantile regression model avoids distributional assumptions at the cost of computational complexity
- Censored investment levels are not an issue in Compustat

Econometric Restrictions

Note that the following conditional moment restrictions hold as an implication of Assumptions 1-4

$$\mathbb{E}\left[\Psi_\tau\left(y_{it} - Q_t(y_{it}|k_{it}, l_{it}, m_{it}, \omega_{it}; \beta(\tau))\right)\middle| k_{it}, l_{it}, m_{it}\right] = 0 \tag{24}$$

$$\mathbb{E}\left[\Psi_\tau\left(l_{it} - \sum_{j=1}^{J} \alpha_{l,j}(\tau)\phi_{l,j}(k_{it}, \omega_{it})\right)\middle| k_{it}, \omega_{it}\right] = 0 \tag{25}$$

$$\mathbb{E}\left[\Psi_\tau\left(m_{it} - \sum_{j=1}^{J} \alpha_{m,j}(\tau)\phi_{m,j}(k_{it}, \omega_{it})\right)\middle| k_{it}, \omega_{it}\right] = 0 \tag{26}$$

$$\mathbb{E}\left[\Psi_\tau\left(i_{it} - \sum_{j=1}^{J} \delta_j(\tau)\phi_{\iota,j}(k_{it}, \omega_{it})\right)\middle| k_{it}, \omega_{it}\right] = 0 \tag{27}$$

For $t \geq 2$,

$$\mathbb{E}\left[\Psi_\tau\left(\omega_{it} - \rho_0(\tau) - \rho_1(\tau)\omega_{it-1} - \rho_2(\tau)\omega_{it-1}^2 - \rho_3(\tau)\omega_{it-1}^3\right)\middle| \omega_{it-1}\right] = 0, \tag{28}$$

Econometric Restrictions

- The function $\Psi_\tau(u) = \tau - \mathbb{1}\{u < 0\}$
- Estimating the above conditional moment restrictions is infeasible due to the unobserved productivity component
- Use the unconditional moment restriction and integrate out productivity
- Let the finite and functional parameters be indexed by a finite dimensional parameter vector $\theta(\cdot)$.
- To fix ideas, consider the unconditional moment restriction corresponding to the production function

$$\mathbb{E}\left[\int_\Omega \Psi_\tau\left(y_{it} - Q_t(y_{it}|k_{it}, l_{it}, m_{it}, \omega_{it}; \beta(\tau))\right) \otimes \begin{pmatrix} 1 \\ k_{it} \\ l_{it} \\ m_{it} \\ \omega_{it} \end{pmatrix} g_i(\omega_{it}; \theta(\cdot)) d\omega_{it}\right] = 0 \qquad (29)$$

- The posterior density $g_i(\omega_{it}; \theta(\cdot)) = f(\omega_{it}|y_{it}, k_{it}, l_{it}, m_{it}, I_t, \theta(\cdot))$ involves the entire set of model parameters

Implementation

- Therefore, it is impossible to estimate the model parameters in a τ-by-τ procedure
- To eliminate the intractability of this problem, the continuous model parameters are approximated by piece-wise linear splines
- θ is 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.
- With piece-wise linear splines, the posterior density has a closed form expression without relying on strong distributional assumptions for estimation

Implementation

- In order to estimate the model, the integral inside the expectation of Equation (29) needs to be approximated
- This can be done using quadrature methods or Monte Carlo integration by converting the problem into a weighted quantile regression
- Due to the high-dimensionality of my application I choose to use a random-walk Metropolis Hastings algorithm to compute the integral
- This is known as a Monte Carlo Expectation Maximization (MCEM) procedure where the maximization step is performed using quantile regression
- This type of estimator is used by Arellano and Bonhomme, 2016 and Arellano, Blundell, and Bonhomme, 2017

Implementation

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

$$
\begin{aligned}
g_i(\omega_{it}; \hat{\theta}^{(s)}) = f(\omega_{it} | y_{it}, k_{it}, l_{it}, m_{it}, i_t, ; \hat{\theta}^{(s)}) \propto \\
\prod_{t=1}^{T} f(y_{it} | k_{it}, l_{it}, m_{it}, \omega_{it}; \hat{\beta}^{(s)}) f(l_{it} | k_{it}, \omega_{it}; \hat{\alpha}_l{}^{(s)}) f(m_{it} | k_{it}, \omega_{it}; \hat{\alpha}_m^{(s)}) \\
\times f(i_{it} | k_{it}, \omega_{it}; \hat{\delta}^{(s)}) \prod_{t=2}^{T} f(\omega_{it} | \omega_{it-1}; \hat{\rho}^{(s)}) f(\omega_{i1}; \hat{\mu}_{\omega_1}, \hat{\sigma}^2_{\omega_1})
\end{aligned}
$$

2. *Maximization Step*: For $q = 1, \ldots, Q$, solve (e.g for production function)

$$
\hat{\beta}(\tau_q)^{(s+1)} = \underset{\beta(\tau_q)}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} \psi_{\tau_q} \left(y_{it} - Q_t(y_{it} | k_{it}, l_{it}, m_{it}, \omega_{it}^{(m)}; \beta(\tau_q)) \right)
$$

Empirical Implementation

- $\psi_\tau(u) = (\tau - \mathbb{1}\{u < 0\})u$ is the "check" function from quantile regression
- Repeat Step 2 for estimating the productivity process, input decision rules and investment
- Take $M = 1$ in the MCEM algorithm and the report estimates as the average of the last $\tilde{S} = S/2$ draws
- This is known as the stochastic EM algorithm (stEM) of Celeux and Diebolt, 1985
- The sequence of maximizers $\hat{\theta}^{(s)}$ is a time-homogeneous Markov chain which, if ergodic, will converge to its stationary distribution
- Nielsen, 2000 provides sufficient conditions for ergodicity and asymptotic properties of the estimator when the "M-step" is solved using maximum likelihood
- Arellano and Bonhomme, 2016 discuss the asymptotic properties of the estimator when the M-step is solved using quantile regression

Empirical Implementation

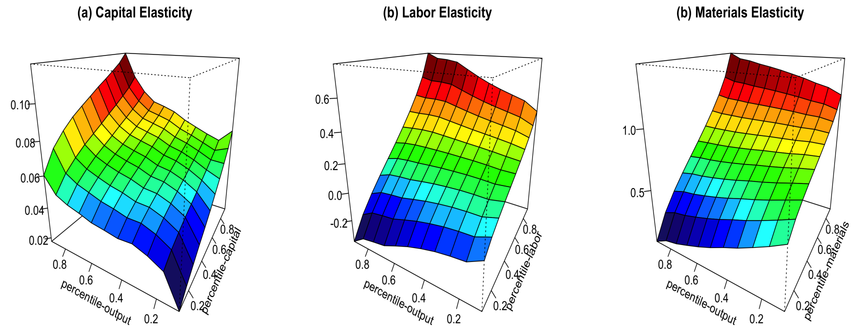
- I run the estimation procedure with 500 random walk Metropolis-Hastings steps keeping the last draw for estimation
- 200 EM steps where the average is taken over half the draws
- I take $Q = 11$ as the grid size for the interpolating spline
- Experimented with many different proposal distributions and initial values
- I use a uniform distribution centered at the current draw of productivity with step size equal to 0.5
- Initial values for productivity are simulated from TFP estimated from the LP model
- I specify 3rd order polynomials for labor, materials, and investment
- Replication code is available on author's Github

Application

- I apply this estimator to the same US data from Doty and Song, 2021
- I focus on more recent data past 1997
- To report my estimates, I simulate productivity data from its estimated parameters and use that to construct investment, inputs, and output using their estimated parameters
- Capital is simulated from a linear accumulation process with constant depreciation rate 0.1
- Results are not too different from reasonable specifications for the capital accumulation process
- I am interested in a variety of average and individual marginal quantile effects
- Using these estimates, I can analyze how firms react to latest shocks to production, inputs, and productivity
- How long does it take for firms to recover from bad shocks to productivity?

Production Elasticities

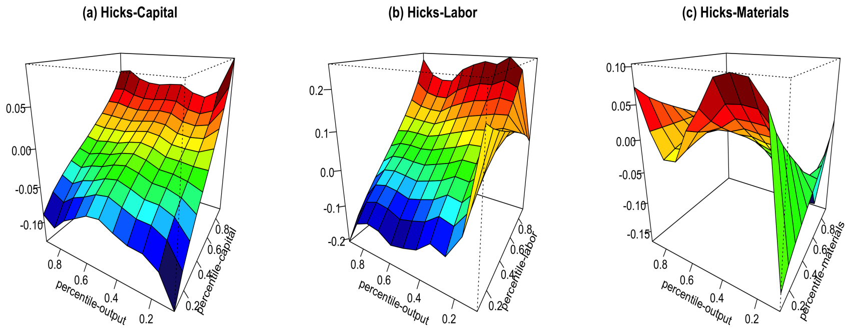
Figure 1: Individual Output Elasticities



*Panel (a): Capital elasticity evaluated at τ_η and percentiles of capital τ_k averaged over values of (l_{it}, m_{it}) that correspond to τ_k. Panel (b): Labor elasticity evaluated at τ_η and percentiles of labor τ_l averaged over values of (k_{it}, m_{it}). Panel (c): Materials elasticity evaluated at τ_η and percentiles of materials τ_m averaged over values of (k_{it}, l_{it}).

Production Elasticities

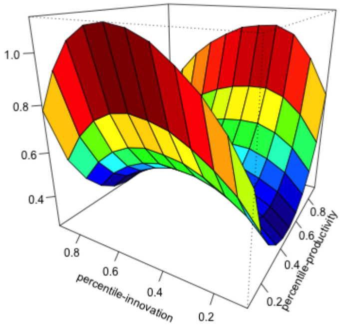
Figure 2: Individual Non-Hicks Neutral Elasticities



*Panel (a): Capital elasticity evaluated at τ_η and percentiles of capital τ_k averaged over values of (l_{it}, m_{it}) that correspond to τ_k. Panel (b): Labor elasticity evaluated at τ_η and percentiles of labor τ_l averaged over values of (k_{it}, m_{it}). Panel (c): Materials elasticity evaluated at τ_η and percentiles of materials τ_m averaged over values of (k_{it}, l_{it}).

Productivity

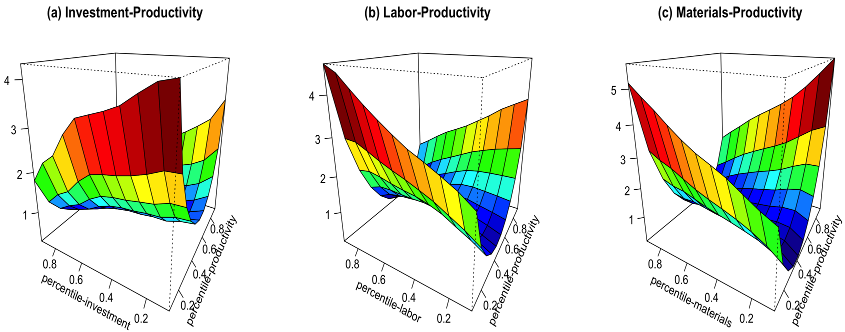
Figure 3: Productivity Persistence



Estimates of average productivity persistence evaluated at τ_{ξ} and percentiles of previous productivity.

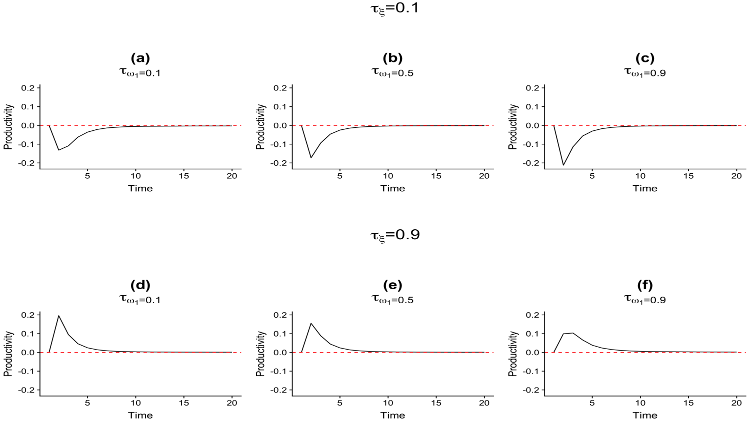
Marginal Productivities

Figure 4: Marginal Productivity of Inputs



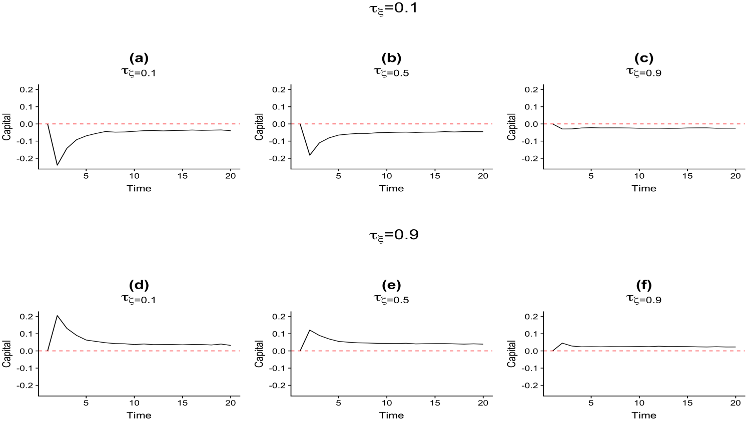
*Panel (a): Investment demand evaluated at τ_ζ and percentiles of productivity τ_ω averaged over values of k_{it}. Panel (b): Labor demand evaluated at τ_{ϵ_l} and percentiles of productivity τ_ω averaged over values of k_{it}. Panel (c): Material demand evaluated at τ_{ϵ_m} and percentiles of productivity τ_ω averaged over values of k_{it}

Productivity Innovation Shocks



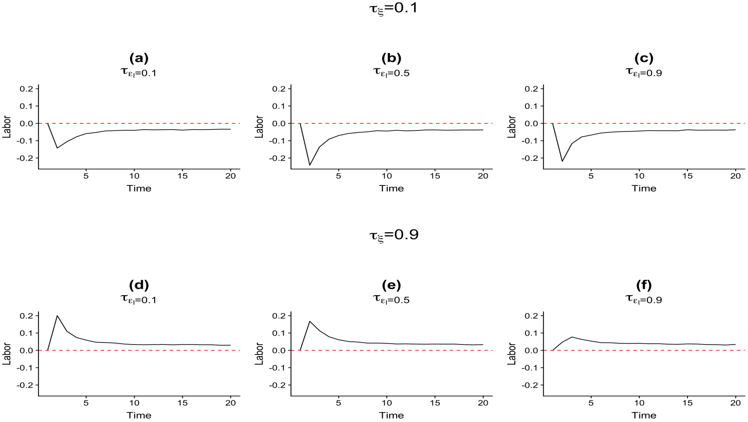
*Top row: Difference between firms hit with low productivity shock $\tau_\xi = 0.1$ and medium shock $\tau_\xi = 0.5$ at different levels of initial productivity. Bottom row: Difference between firms hit with high productivity shock $\tau_\xi = 0.9$ and medium shock $\tau_\xi = 0.5$ at different levels of initial productivity.

Productivity Innovation Shocks to Investment



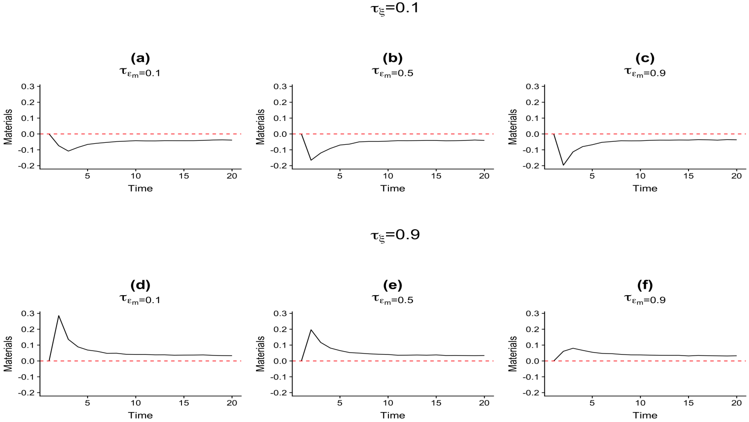
*Top row: Difference between firms hit with low productivity shock $\tau_\xi = 0.1$ and medium shock $\tau_\xi = 0.5$ at different levels of investment demand. Bottom row: Difference between firms hit with high productivity shock $\tau_\xi = 0.9$ and medium shock $\tau_\xi = 0.5$ at different levels of investment demand

Productivity Innovation Shocks to Labor



*Top row: Difference between firms hit with low productivity shock $\tau_\xi = 0.1$ and medium shock $\tau_\xi = 0.5$ at different levels of labor demand. Bottom row: Difference between firms hit with high productivity shock $\tau_\xi = 0.9$ and medium shock $\tau_\xi = 0.5$ at different levels of labor demand.

Productivity Innovation Shocks to Materials



*Top row: Difference between firms hit with low productivity shock $\tau_\xi = 0.1$ and medium shock $\tau_\xi = 0.5$ at different levels of materials demand. Bottom row: Difference between firms hit with high productivity shock $\tau_\xi = 0.9$ and medium shock $\tau_\xi = 0.5$ at different levels of materials demand.

Conclusion

- Extended the identification arguments of Hu, Huang, and Sasaki, 2019 under more general production functions and the entire distribution of productivity
- Proposed an estimation framework to document firm heterogeneity in production, input decisions, and productivity
- Firms which low productivity have higher probability of having low productivity for different sized innovation shocks. Same for high productivity firms
- Firms respond to productivity shocks by using more inputs, but this affect is asymmetric for different productivity levels and the rank of the conditional input distribution
- Asymmetric impact of innovation shocks to inputs and the length of recovery after bad/good shocks
- Extension to multi-dimensional unobservables: fixed effects, labor-augmenting productivity
- Implications for TFP estimates? Market power?