A Dynamic Panel Data Framework for Identification and Estimation of Nonlinear Production Functions

Justin Doty[*]and Suyong Song[†]

1   Introduction

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

$$Y_{it} = \sum_{k=1}^{K_1} \beta_k(\eta_{it}) g_k(L_{it}, M_{it}, K_{it}, \omega_{it}) \tag{1}$$

where Y_{it} is firm i's' output at time t and L_{it}, M_{it}, K_{it} denotes optimal input choices for labor, materials, and capital respectively. The unobserved productivity is denoted by ω_{it} which is correlated to input choices of the firm at time t. We let the output elasticities β to be functionally dependent on unobserved production shocks $\eta_{i1}, \ldots, \eta_{iT}$ that are uncorrelated to input choices and productivity at time t.

Without loss of generality we normalize η_{it} to be uniformly distributed on the interval $[0, 1]$. This model corresponds to a nonlinear random coefficient model where the outcome y_{it} is monotonic in η_{it}. In practice we can allow for nonlinear interactions between inputs and unobserved productivity at different quantiles so that marginal effects can be modeled as non-Hick's neutral. For empirical simplicity we can model separability in the unobserved productivity to calculate total factor productivity (TFP). The function g is an unknown nonlinear function.

In this model, heterogeneity in production technology across firms is driven by the rank of the unobserved production shocks η_{it}. We specify the productivity process as

$$\omega_{it} = \sum_{k=1}^{K_2} \rho_k(\xi_{it}) h_k(\omega_{it-1}) \tag{2}$$

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} We let h be another unknown nonlinear function that allows the persistence in productivity in firms to be nonlinear across different quantiles.

2   Econometric Model

We introduce a dynamic model of firm investment that is a slight modification of Ericson and Pakes (1995). We begin by providing the evolution process for $\omega_t \in \mathbb{R}$ which is given by:

$$\omega_t = g(\omega_{t-1}, \xi_t), \tag{3}$$

---
[*]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

where the function $g(\cdot, \xi_t)$ is strictly increasing in the iid innovation shock, $\xi_t \in \mathbb{R}$. 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).

Capital accumulates according to the following process:

$$K_t = \kappa_t(K_{t-1}, I_{t-1}, \upsilon_t), \tag{4}$$

where the function κ is strictly increasing in its last argument and υ_t denotes an iid shock independent of the other arguments. A special case of equation (4) is the usual capital accumulation law $K_t = (1-\delta)K_{t-1} + I_{t-1} + \upsilon_t$. Here υ_t are other factors that affect the capital accumulation process. Breaking this deterministic relationship in the capital accumulation process allows for variation in investment I_t conditional on K_{t+1} and K_t which will be important in the identification results in the next section.

The investment decision problem is an extension of the Ericson and Pakes (1995) framework which can also be found in Hu and Shum (2013) and Ackerberg *et al.* (2007). In each period, a firm chooses investment to maximize its discounted future profits:

$$I_t = I^*(K_t, \omega_t, \zeta_t) = \underset{I_t \geq 0}{\operatorname{argmax}} \left[\Pi_t(K_t, \omega_t, \zeta_{it}) - c(I_t) + \beta \mathbb{E}\left[V_{t+1}(K_{t+1}, \omega_{t+1}, \zeta_{t+1}) | \mathcal{I}_t \right] \right], \tag{5}$$

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. Current costs to investment are given by $c(I_t)$ and β is the firm's discount factor. Pakes (1994) provides specific conditions for which the investment policy function is strictly increasing in its scalar unobservable component. However, the difference here is that we require the condition to be on the unobservable demand shock, as opposed to the unobserved productivity shock, ω_{it}. We require that the firm's individual demand function is strictly increasing in ζ_{it} which is required in the quantile regression literature for demand functions, see for example (Blundell *et al.*, 2017) regarding non-separable demand functions.We can interpret conditions for strict monotonicity here and leave a detailed explanation in the appendix. Without loss of generality, we normalize $\zeta_t \sim U[0,1]$. In a later section we show how to modify identification and estimation strategies to cases when investment is censored or there is selection bias due to endogenous entry/exit which can be the case in certain industries.

The specifications for the static inputs, labor and material are much easier to state. We let optimal labor demand be given by:

$$l_t = \ell(k_t, \omega_t, \epsilon_t) \tag{6}$$

where the function, $\ell(\cdot, \cdot, \epsilon_t)$ is strictly increasing in ϵ_t which is assumed to be independent of the other arguments. We normalize this to be standard uniform each period. We follow a similar model for materials:

$$m_t = m_t(k_t, \omega_t, \varepsilon_t) \tag{7}$$

where the function, $m(\cdot, \cdot, \varepsilon_t)$ is strictly increasing in ε_t which is assumed to be independent of the other arguments. We normalize this to be standard uniform each period.

3   Identification

Our goal is identification of the Markov law of motion $f_{Y_t, L_t, M_t, K_t, I_t \omega_t | Y_{t-1}, L_{t-1}, M_{t-1}, K_{t-1}, I_{t-1} \omega_{t-1}}$ which we allow to be stationary. We assume the researcher observes a panel dataset consisting of i.i.d observations of firm output and input choices with the number of time periods $T \geq 4$ for a large number of firms. We introduce the following assumptions that simplify the expression for the law of motion.

Assumption 3.1. *(Production Dynamics)*

1. *Non-dynamic output and inputs:* $f_{Y_t,L_t,M_t,K_t,I_t,\omega_t|Y_{t-1},L_{t-1},M_{t-1},K_{t-1},I_{t-1}\omega_{t-1}} = f_{Y_t,L_t,M_t,K_t,I_t\omega_t|K_{t-1},I_{t-1}\omega_{t-1}}$

2. *First-order Markov:* $f_{Y_t,L_t,M_t,K_t,I_t\omega_t|K_{t-1},I_{t-1}\omega_{t-1},\mathcal{I}_{<t-1}} = f_{Y_t,L_t,M_t,K_t,I_t\omega_t|K_{t-1},I_{t-1}\omega_{t-1}}$, where $\mathcal{I}_{<t-1}$
   is the firm's information set up to time $t-1$

3. *Limited Feedback:* $f_{Y_t,L_t,M_t,K_t,I_t|K_{t-1},I_{t-1},\omega_t,\omega_{t-1}} = f_{Y_t,L_t,M_t,K_t,I_t|K_{t-1},I_{t-1},\omega_t}$

4. *Mutual Independence: $\eta_t, \epsilon_t, \varepsilon_t$ and ζ_t are mutually independent conditional on K_t and ω_t*

Assumption $(3.1)(1)$ states that input decisions for labor and material inputs maximize profits in the current period and do not have any dynamic implications. This can be relaxed to include cases where labor has dynamic effects due to hiring/firing costs. This assumption also does not allow for dynamic effects of output, which is standard in the production function literature. Assumption $(3.1)(2)$ is a standard condition in Markovian dynamic decision models. Decision and state variables any period before the previous period are irrelevant to the firm given the information contained in last periods decision and state variables. The investment model we introduce in the next few paragraphs satisfies this assumption and is trivially satisfied for any of the static decision variables.

Assumption $(3.1)(3)$ states that previous period's productivity does not directly affect the firms decision and state variables in the current period. Instead, they indirectly affect these variables through the previous periods decision and state variables. Assumption $(3.1)(4)$ are conditional independence assumptions for output and the flexible input decisions. Conditional on current period labor, materials, and capital, the previous period's capital as well as investment are not relevant in determining output in the current period. Also, capital summarizes the information about investment relevant in determining labor, but also assumes that labor decisions are mutually independent of material input decisions. The latter assumption is satisfied using timing assumptions that are common in the literature.

We will proceed in steps in factoring the Markov law of motion into the densities we are interested in identying. First, using Assumption (3.1) 1 and 2:

$$
\begin{aligned}
&f(Y_t, L_t, M_t, K_t, I_t, \omega_t | Y_{t-1}, L_{t-1}, M_{t-1}, K_{t-1}, I_{t-1}, \omega_{t-1}, \mathcal{I}_{<t-1}) \\
&= f(Y_t, L_t, M_t, K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1})
\end{aligned}
\tag{8}
$$

We can simplify the density on the last line of the equation (8) as

$$
\begin{aligned}
&f(Y_t, L_t, M_t, K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1}) \\
&= f(Y_t | L_t, M_t, K_t, K_{t-1}, I_t, I_{t-1}, \omega_t, \omega_{t-1}) f(L_t | M_t, K_t, K_{t-1}, I_t, I_{t-1}, \omega_t, \omega_{t-1}) \\
&\times f(M_t | K_t, K_{t-1}, I_t, I_{t-1}, \omega_t, \omega_{t-1}) f(K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1})
\end{aligned}
\tag{9}
$$

Using Assumption (3.1) 3 and 4, equation (9) becomes

$$
\begin{aligned}
&f(Y_t, L_t, M_t, K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1}) \\
&= f(Y_t | L_t, M_t, K_t, \omega_t) f(L_t | K_t, \omega_t) f(M_t | K_t, \omega_t) f(K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1})
\end{aligned}
\tag{10}
$$

The last density on the second line of equation (10) is the Markov law of motion for capital, investment, and productivity. It can be factored as:

$$
\begin{aligned}
&f(K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1}) \\
&= f(I_t | K_t, K_{t-1}, I_{t-1}, \omega_t, \omega_{t-1}) f(K_t | I_{t-1}, K_{t-1}, \omega_{t-1}) f(\omega_t | K_t, K_{t-1}, \omega_{t-1}) \\
&= f(I_t | K_t, \omega_t) f(K_t | K_{t-1}, I_{t-1}) f(\omega_t | \omega_{t-1})
\end{aligned}
\tag{11}
$$

The last line of equation (11) uses a combination of Assumption (3.1) 3 and the way investment, capital, and the productivity process in the previous section.

Combining all of the above assumptions, the density we are interested in identifying is:

$$f(Y_t, L_t, M_t, K_t, I_t, \omega_t | K_{t-1}, I_{t-1}, \omega_{t-1}) = f(Y_t | L_t, M_t, K_t, \omega_t) f(L_t | K_t, \omega_t) f(M_t | K_t, \omega_t) f(I_t | K_t, \omega_t) f(K_t | K_{t-1}, I_{t-1}) f(\omega_t | \omega_{t-1}) \tag{12}$$

3.1 Hu and Shum (2012) Identification

STILL A WORK IN PROGRESS, THESE DO NOT REFLECT MOST RECENT CHANGES ABOVE

We outline the identification procedure similar to Hu and Shum (2012). First, let $V_t = \{X_t, W_t\}$. Under assumption (3.1), we can write:

$$f_{X_{t+1}, W_{t+1}, X_t, W_t, X_{t-1}, W_{t-1}, X_{t-2}, W_{t-2}} = \int f_{X_{t+1}, W_{t+1} | W_t, \omega_t} f_{X_t, W_t | W_{t-1}, \omega_t} f_{X_{t-1}, W_{t-1}, X_{t-2}, W_{t-2}, \omega_t} d\omega_t \tag{13}$$

Should this be factored further?

The linear operator $L_{V_{t-2}, \bar{x}_{t-1}, \bar{w}_{t-1}, \bar{x}_t, \bar{w}_t, V_{t+1}}$ as a mapping from the $\mathcal{L}^p$ space of functions of V_{t+1} to the $\mathcal{L}^p$ space of functions of V_{t-2} is defined as

$$\left(L_{V_{t-2}, \bar{x}_{t-1}, \bar{w}_{t-1}, \bar{x}_t, \bar{w}_t, V_{t+1}} h \right)(v_{t-2})$$
$$= \int f_{V_{t-2}, X_{t-1}, W_{t-1}, X_t, W_t, V_{t+1}} (v_{t-2}, \bar{x}_{t-1}, \bar{w}_{t-1}, \bar{x}_t, \bar{w}_t, v_{t+1}) h(v_{t+1}) dv_{t+1} : \tag{14}$$
$$h \in \mathcal{L}^p(\mathcal{V}_{t+1}), \bar{x}_{t-1} \in \mathcal{X}_{t-1}, \bar{x}_t \in \mathcal{X}_t, \bar{w}_{t-1} \in \mathcal{W}_{t-1}, \bar{w}_t \in \mathcal{W}_t,$$

where $\mathcal{V}_t, \mathcal{X}_{t-1}, \mathcal{X}_t, \mathcal{W}_{t-1}$ and $\mathcal{W}_t$ are the supports of $V_t, X_{t-1}, X_t, W_{t-1}$ and W_t respectively. We also define the diagonal operator as:

$$\left(D_{\bar{x}_t, \bar{w}_t | \bar{w}_{t-1}, \omega_t} h \right)(\omega_t) = f_{X_t, W_t |, W_{t-1}, \omega_t}(\bar{x}_t, \bar{w}_t | \bar{w}_{t-1}, \omega_t) h(\omega_t)$$
$$h \in \mathcal{L}^p(\Omega_t), \bar{x}_t \in \mathcal{X}_t, \bar{w}_t \in \mathcal{W}_t, \bar{w}_{t-1} \in \mathcal{W}_{t-1} \tag{15}$$

We provide a few lemmas that allow us to represent the observed density $f_{X_{t+1}, W_{t+1}, X_t, W_t, X_{t-1}, W_{t-1}, X_{t-2}, W_{t-2}}$ and the Markov law of motion $f_{X_t, W_t, \omega_t | X_{t-1}, W_{t-1}, \omega_{t-1}}$ using operator notation.

Lemma 3.1. *(Representation of the Observed Density)   For any $t \in \{3, \ldots, T-1\}$. Assumption (3.1) implies that, for any $(x_t, w_t, x_{t-1}, w_{t-1}) \in \mathcal{X}_t \times \mathcal{W}_t \times \mathcal{X}_{t-1} \times \mathcal{W}_{t-1}$,*

$$L_{V_{t+1}, x_t, w_t, x_{t-1}, w_{t-1}} = L_{V_{t+1} | w_t, \omega_t} D_{x_t, w_t | w_{t-1}, \omega_t} L_{x_{t-1}, w_{t-1}, V_{t-2}, \omega_t} \tag{16}$$

Lemma 3.2. *(Representation of the Markov Law of Motion)   For any $t \in \{3, \ldots, T-1\}$ and under certain invertibility conditions provided below give, for any $(x_t, w_t, x_{t-1}, w_{t-1}) \in \mathcal{X}_t \times \mathcal{W}_t \times \mathcal{X}_{t-1} \times \mathcal{W}_{t-1}$,*

$$L_{x_t, w_t, \omega_t | w_{t-1}, \omega_{t-1}} = L_{V_{t+1} | w_t, \omega_t}^{-1} L_{V_{t+1}, x_t, w_t, x_{t-1}, w_{t-1}, V_{t-2}} L_{V_t, x_{t-1}, w_{t-1}, V_{t-2}}^{-1} L_{V_t | w_{t-1}, \omega_t} \tag{17}$$

The identification argument then depends on the existence and uniqueness a spectral decomposition of the linear operator $L_{V_{t+1}, x_t, w_t, x_{t-1}, w_{t-1}}$ which corresponds to the observed density $f_{X_{t+1}, W_{t+1}, X_t, W_t, X_{t-1}, W_{t-1}, X_{t-2}, W_{t-2}}$. We state these assumptions and their interpretation in the production function literature and in context of our model.

Assumption 3.2. *(Invertibility)   There exists variable(s) V_t such that*

a) *For any $w_t \in \mathcal{W}_t$, $L_{V_{t+1} | w_t, \omega_t}$ is one-to-one*

b) *For any $(x_{t-1}, w_{t-1}) \in \mathcal{X}_{t-1} \times \mathcal{W}_{t-1}$, $L_{V_t, x_{t-1}, w_{t-1}, V_{t-2}}$ is one-to-one*

c) *For any $(x_t, w_t) \in \mathcal{X}_t \times \mathcal{W}_t$, there exists a $(x_{t-1}, w_{t-1}) \in \mathcal{X}_{t-1} \times \mathcal{W}_{t-1}$ and a neighborhood $\mathcal{N}^r$ around $(x_t, w_t, x_{t-1}, w_{t-1})$ such that, for any $(\bar{x}_t, \bar{w}_t, \bar{x}_{t-1}, \bar{w}_{t-1}) \in \mathcal{N}^r$, $L_{V_{t-2}, \bar{x}_t, \bar{w}_t, \bar{x}_{t-1}, \bar{w}_{t-1}, V_{t-2}}$ is one-to-one*

4   Estimation Strategy

Our empirical specification for the Markovian transitions of productivity, output, and capital evolution closely resemble Arellano *et al.* (2017). Let $z_{it} = (l_{it}, m_{it})$ denote the flexible input choices for notational simplicity.

4.1   Persistent Productivity

Let age_{it} denote the age of firm i at time t, we specify productivity to transition according to:

$$Q_t(\omega_{it-1}, \tau) = Q(\omega_{it-1}, age_{it}, \tau) = \sum_{k=1}^{K} \rho_k(\tau) g_k(\omega_{it-1}, age_{it}) \tag{18}$$

The quantile function for ω_{i1} is specified in a similar way

$$Q_{\omega_1}(age_{i1}, \tau) = \sum_{k=0}^{K} \rho_k^1(\tau) g_k(age_{i1}) \tag{19}$$

4.2   Output

We specify the output equation as follows:

$$\begin{aligned} Q_t(z_{it}, k_{it}\omega_{it}, \tau) &= Q(z_{it}, k_{it}, age_{it}, \omega_{it}, \tau) \\ &= \sum_{k=1}^{K} \beta_k g_k(z_{it}, k_{it}, age_{it}, \omega_{it}) + \beta_0(\tau) \end{aligned} \tag{20}$$

4.3   Investment Demand

We specify the investment demand equation as:

$$\begin{aligned} \iota_t(\omega_{it}, k_{it}, \tau) &= \iota(\omega_{it}, k_{it}, age_{it}, \tau) \\ &= \sum_{k=1}^{K} \iota_k g_k(\omega_{it}, k_{it}, age_{it}) + \iota_0(\tau) \end{aligned} \tag{21}$$

4.4   Capital

We specify the distribution of initial capital as:

$$Q_\kappa(\omega_{i1}, age_{i1}, \tau) = \sum_{k=0}^{K} \kappa_k^1 g_k(\omega_{i1}, age_{i1}) \tag{22}$$

The evolution of capital is specified as:

$$\begin{aligned} \kappa_t(\omega_{it}, K_{it-1}, I_{it-1}, \tau) &= \kappa_t(\omega_{it}, K_{it-1}, I_{it-1}, age_{it}, \tau) \\ &= \sum_{k=1}^{K} \kappa_k g_k(\omega_{it}, K_{it-1}, I_{it-1}, age_{it}) + \kappa_0(\tau) \end{aligned} \tag{23}$$

4.5   Static Inputs

Finally, we specify a reduced-form demand function for the static inputs of $z = (m, l, u)$ (in logs) as:

$$\begin{aligned} \mu_t^z(k_{it}, \omega_{it}, \tau) &= \mu^z(k_{it}, \omega_{it}, age_{it}, \tau) \\ &= \sum_{k=1}^{K} \mu_{zk} g_k(\omega_{it}, k_{it}, age_{it}) + \mu_{z0}(\tau) \end{aligned} \tag{24}$$

5 Implementation

To ease notation, we let $\rho_k, \rho_k, \beta_0, \iota_0, \kappa_0, \mu_{z0}$ be indexed by a finite dimensional parameter vector θ that also contains $\beta_1, \ldots, \beta_K, \iota_1, \ldots, \iota_K, \kappa_1, \ldots, \kappa_K$ and $\mu_{z1}, \ldots, \mu_{zK}$. We model the functional parameters using Wei and Carroll (2009) and Arellano and Bonhomme (2016). For example, the function ρ_ω is modeled as a piecewise-polynomial interpolating splines on a grid $[\tau_1, \tau_2], [\tau_3, \tau_4], \ldots, [\tau_{J-1}, \tau_J]$, contained in the unit interval and is constant on $[0, \tau_1]$ and $[\tau_J, 1)$ The intercept coefficient ρ_0 is specified as the quantile of an exponential distribution on $(0, \tau_1]$ (indexed by λ_ω^-) and $[\tau_{J-1}, 1)$ (indexed by λ_ω^+).

The function ρ_ω is modeled as piecewise linear on $[\tau_1, \tau_J]$. The remaining functional parameters are modeled similarly. We take $J = 11$ and $\tau_j = \frac{j}{J+1}$. Following Arellano *et al.* (2017) we parameterize the distribution of η to be log-normal so we set, for example, $\beta_0(\tau) = \alpha_0 + \sigma \Phi^{-1}(\tau)$. Similar parameterizations for made for ι_0, κ_0 and μ_{z0}. In the following section we outline the model's restrictions and a feasible estimation strategy.

5.1 Model Restrictions

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 in section 2. Therefore, we estimate the parameters of interest from the following conditional moment restrictions. For $t \geq 2$

$$\mathbb{E}\left[\Psi_{\tau_j}(\xi_{it})\Big|\omega_{it-1}, age_{it}\right] = 0 \tag{25}$$

For initial productivity,

$$\mathbb{E}\left[\Psi_{\tau_j}(\xi_{i1})\Big|\omega_{i1}, age_{i1}\right] = 0 \tag{26}$$

and for initial capital

$$\mathbb{E}\left[\Psi_{\tau_j}(\upsilon_{i1})\Big|\omega_{i1}, age_{i1}\right] = 0 \tag{27}$$

For the nonlinear specifications we have

$$\mathbb{E}\left[\eta_{it}\Big|\omega_{it}, z_{it}, k_{it}, age_{it}\right] = 0 \tag{28}$$

$$\mathbb{E}\left[\zeta_{it}\Big|\omega_{it}, k_{it}, age_{it}\right] = 0 \tag{29}$$

$$\mathbb{E}\left[\upsilon_{it}\Big|\omega_{it}, K_{it-1}, I_{it-1}, age_{it}\right] = 0 \tag{30}$$

$$\mathbb{E}\left[\varepsilon_{it}^Z\Big|\omega_{it}, k_{it}, age_{it}\right] = 0 \tag{31}$$

To fix ideas, we first specify how to estimate the conditional moments for persistent productivity and production.

For $t \geq 2$

$$\mathbb{E}\left[\Psi_{\tau_j}(\xi_{it})\Big|\omega_{it-1}, age_{it}\right] = $$
$$\mathbb{E}\left[\Psi_{\tau_j}(\omega_{it} - \sum_{k=1}^{K} \bar{\rho}_k(\tau_j)g_k(\omega_{it-1}, age_{it}))\Big|\omega_{it-1}, age_{it}\right] = 0 \tag{32}$$

$$\mathbb{E}\left[\eta_{it}\Big|\omega_{it}, z_{it}, k_{it}, age_{it}\right] =$$

$$\mathbb{E}\left[(y_{it} - \bar{\alpha} - \sum_{k=1}^{K}\bar{\beta}_k g_k(z_{it}, k_{it}, \omega_{it}, age_{it}))\Big|\omega_{it}, z_{it}, k_{it}, age_{it}\right] = 0 \tag{33}$$

Here $\bar{\rho}_k(\tau_j), \bar{\beta}_k$ and $\bar{\alpha}$ denote the true values of $\rho_k(\tau_j), \beta_k$ and α for $j \in \{1, \ldots, J\}$ and $k \in \{1, \ldots, K\}$. 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:

$$\mathbb{E}\left[\int\left(\Psi_{\tau_j}(\omega_{it} - \sum_{k=1}^{K}\bar{\rho}_k(\tau_j)g_k(\omega_{it-1}, age_{it})) \otimes \begin{pmatrix} 1 \\ \omega_{it-1} \\ age_{it} \end{pmatrix}\right)g_i(\omega_{it}; \bar{\theta})d\omega_{it}\right] = 0 \tag{34}$$

and

$$\mathbb{E}\left[\int\left((y_{it} - \bar{\alpha} - \sum_{k=1}^{K}\bar{\beta}_k g_k(z_{it}, k_{it}, \omega_{it}, age_{it})) \otimes \begin{pmatrix} 1 \\ \omega_{it} \\ z_{it} \\ k_{it} \\ age_{it} \end{pmatrix}\right)g_i(\omega_{it}; , \bar{\theta})d\omega_{it}\right] = 0, \tag{35}$$

where $\bar{\theta}$ denotes the true values of θ. The posterior distribution is specified as (age omitted for ease of notation):

$$f_i(\omega_{it}; \bar{\theta}) = f(\omega_{it}|y_{it}, z_{it}, k_{it}, I_{it}; \bar{\theta}) \propto$$

$$\prod_{t=1}^{T} f(y_{it}|\omega_{it}, z_{it}, k_{it}; \bar{\theta}) \times \prod_{t=1}^{T} f(z_{it}|\omega_{it}, k_{it}; \bar{\theta}) \times$$

$$\prod_{t=2}^{T} f(K_{it}|\omega_{it}, K_{it-1}, I_{it-1}; \bar{\theta}) \times \prod_{t=1}^{T} f(I_{it}|\omega_{it}, k_{it}; \bar{\theta}) \times \tag{36}$$

$$\prod_{t=2}^{T} f(\omega_{it}|\omega_{it-1}; \bar{\theta}) \times f(\omega_{i1}; \bar{\theta}) \times f(K_{i1}|\omega_{i1}; \bar{\theta})$$

The posterior density in equation (36) is a closed-form expression when using piece-wise linear splines for $\theta(\cdot)$. The estimation is an Expectation Maximization (EM) algorithm. In Arellano and Bonhomme (2016) and Arellano *et al.* (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

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

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

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

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

2. *Maximization Step*: For $j = 1, \ldots, J$, solve

$$\hat{\boldsymbol{\rho}}(\tau_j)^{(s+1)} = \underset{\boldsymbol{\rho}(\tau_j)}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{t=2}^{T} \sum_{m=1}^{M} \Psi_{\tau_j}(\omega_{it}^{(m)} - \sum_{k=1}^{K} \rho_k g_k(\omega_{it-1}^{(m)}, age_{it}))$$

$$\hat{\boldsymbol{\rho}^1}(\tau_j)^{(s+1)} = \underset{\boldsymbol{\rho}^1(\tau_j)}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{m=1}^{M} \Psi_{\tau_j}(\omega_{i1}^{(m)} - \sum_{k=1}^{K} \rho_k^1 g_k(age_{i1}))$$

$$\hat{\boldsymbol{\kappa}^1}(\tau_j)^{(s+1)} = \underset{\boldsymbol{\kappa}^1(\tau_j)}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{m=1}^{M} \Psi_{\tau_j}(K_{i1} - \sum_{k=1}^{K} \kappa_k^1 g_k(\omega_{i1}^{(m)}, age_{i1}))$$

The parameters of the production function equation in (35) can be estimated using a nonlinear regression for a given draw of $\omega_{it}^{(m)}$. Then, the variance of the shock η can be estimated using [1]

$$\hat{\sigma}^2 = \frac{1}{NTM} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{m=1}^{M} \left[\left(y_{it} - \hat{\alpha} - \sum_{k=1}^{K} \hat{\beta}_k g_k(\omega_{it}^{(m)}, z_{it}, k_{it}, age_{it}) \right)^2 \right] \tag{37}$$

---

[1]The parameters of the investment equation, capital accumulation process, and static inputs can be estimated similarily, however are omitted for clarity. See appendix for details.

References

ACKERBERG, D., BENKARD, C. L., BERRY, S. and PAKES, A. (2007). Chapter 63 econometric tools for analyzing market outcomes. In *Handbook of Econometrics*, Elsevier, pp. 4171–4276.

ARELLANO, M., BLUNDELL, R. and BONHOMME, S. (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica*, **85** (3), 693–734.

— and BONHOMME, S. (2016). Nonlinear panel data estimation via quantile regressions. *The Econometrics Journal*, **19** (3), C61–C94.

BLUNDELL, R., HOROWITZ, J. and PAREY, M. (2017). Nonparametric estimation of a nonseparable demand function under the slutsky inequality restriction. *The Review of Economics and Statistics*, **99** (2), 291–304.

DORASZELSKI, U. and JAUMANDREU, J. (2013). R&d and productivity: Estimating endogenous productivity. *The Review of Economic Studies*, **80** (4), 1338–1383.

ERICSON, R. and PAKES, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, **62** (1), 53.

HU, Y. and SHUM, M. (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics*, **171** (1), 32–44.

— and — (2013). Identifying dynamic games with serially correlated unobservables. In *Structural Econometric Models*, Emerald Group Publishing Limited, pp. 97–113.

PAKES, A. (1994). Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interactions. In C. A. Sims (ed.), *Advances in Econometrics*, Cambridge University Press, pp. 171–274.

WEI, Y. and CARROLL, R. J. (2009). Quantile regression with measurement error. *Journal of the American Statistical Association*, **104** (487), 1129–1143.