Identification and Estimation of Production Function with Unobserved Heterogeneity

Hiroyuki Kasahara
University of British Columbia

Paul Schrimpf
University of British Columbia

Michio Suzuki
Tohoku University

July 2020

Motivation

- Estimation of production function and TFP is important in empirical applications (Empirical IO, Trade, Macro).

- In most empirical applications, the coefficients of Cobb-Douglas production function are assumed to be common across firms

$$y_{it} = \beta_0 + \beta_m m_{it} + \beta_\ell \ell_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}$$

$\rightarrow$ $(\beta_m, \beta_\ell, \beta_k)$ is assumed to be common across i's within "narrowly defined" industry

Is $(\beta_m, \beta_\ell, \beta_k)$ really common across firms?

Implications of Cobb-Douglas Production Function:

$$\frac{P_{M,t}M_{it}}{P_{Y,t}Y_{it}} \approx \beta_m$$

$$\frac{P_{M,t}M_{it}}{P_{M,t}M_{it} + W_t L_{it}} \approx \frac{\beta_m}{\beta_m + \beta_\ell}$$

$\overline{\left(\frac{P_{M,t} M_{it}}{P_{Y,t} Y_{it}}\right)}_i$ over 25 yrs in Japanese Concrete Product and Electric Audio Industries, 30+ Workers

Figure: Concrete Product (2223)

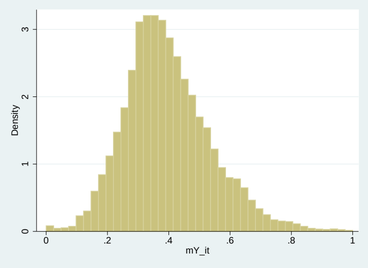
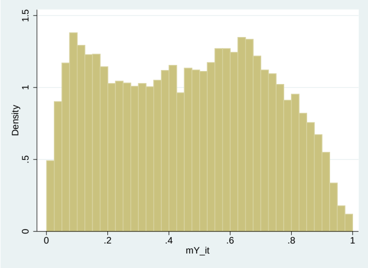


Figure: Electric Audio (2814)



The share of materials is **heterogenous** across firms and **persistent** over time within firm.

$\overline{\left(\frac{P_{M,t}M_{it}}{P_{M,t}M_{it}+W_tL_{it}}\right)_i}$ over 25 years

Figure: Concrete Product (2223)

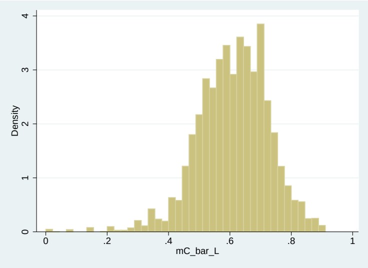
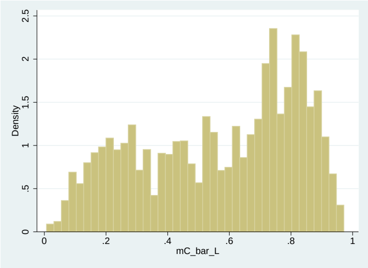


Figure: Electric Audio (2814)



$\Rightarrow$ The difference in markups is not the main reason for the heterogeneity in $\frac{PM_{it}}{PY_{it}}$.

$\overline{\left(\frac{P_{M,t}M_{it}}{P_{Y,t}Y_{it}}\right)_i}$: 2-digit vs. 3-digit vs. 4-digit

Figure: 2-digit Electric Parts, Device, Circuit (28)

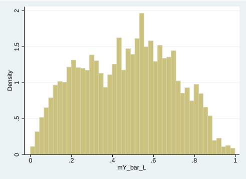


Figure: 3-digit Electric Device (281)

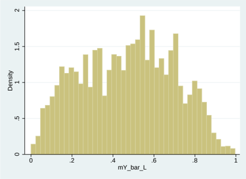
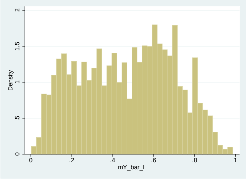


Figure: 4-digit Electric Audio (2814)



Industry Code : Name	No. of Obs.	90-10 diff in $\left(\frac{PM_{it}}{PQ_{it}}\right)_i$	90-10 diff in $\left(\frac{PM_{it}}{PM_{it}+WL_{it}}\right)_i$
28: Electric Parts/Devise/Circuit	33,467	0.62	0.66
281: Electric Device	21,712	0.62	0.67
2814: Electric Audio	12,441	0.63	0.67

The average difference between the 90th and the 10th percentiles at 2-, 3-, and 4-digit industry classifications

Industry Classifications	No. of Industries	Ave. 90-10 diff in $\left(\frac{PM_{it}}{PQ_{it}}\right)_i$	Ave. 90-10 diff in $\left(\frac{PM_{it}}{PM_{it}+WL_{it}}\right)_i$	Ave. No. of Obs.
2-digit	24	0.46	0.44	53,205
3-digit	149	0.42	0.41	8,570
4-digit	279	0.38	0.37	2,666

More general than Cobb-Douglas? (Electric Audio)

$$\frac{P_{M,t} M_{it}}{P_{Y,t} Y_{it}} = \text{2nd order polynomials of } (l_{it}, k_{it}, m_{it}) + e_{it}$$

$$\hat{\xi}_i := T^{-1} \sum_{t=1}^{T} \hat{e}_{it}$$

- 90-10 diff of $\hat{\xi}_i = 0.52$
- 90-10 diff of $\left(\frac{PM_{it}}{PQ_{it}} \right)_i = 0.63$
- Evidence for **persistent heterogeneity** under more general functions than Cobb-Douglas.

Issues

- Can we identify production function in the presence of unobserved heterogeneity?
- Can we estimate production function with random coefficients using a typical firm/plant-level panel data?

Issues in Estimation of Firm-Level Production Function

Simultaneity Bias

- Corr. between productivity & input factors (Marschak and Andrews, 1944)
- Control function/"Proxy variable" approach: OP (Olley and Pakes, 1996), LP (Levinsohn and Petrin, 2003), Wooldrige (2009), ACF (Ackerberg, Caves and Frazer, 2015)
- Dynamic Panel: Arellano and Bond (1991), Blundell and Bond (1998)
- "Exogenous" Input Price as IV: Doraszelski and Jaumandreu (2018)

→ OP/LP is widely used in empirical analysis.

Identification of Firm-Level Production Function

Nonparametric Identification Problem

- Bond and Söderbom (2005) and ACF (2015): Colineality of material and labor.
- GNR (Gandhi, Navarro, and Rivers, 2020): No exogenous variation to identify flexible input's elasticity after conditioning on quasi-fixed input factors.
- GNR exploit the first order condition w.r.t. flexible inputs for identification.

Production Function with Random Coefficients

- Mairesse and Griliches (1990), Van Biesebroeck (2003), Doraszelski and Jaumandreu (2018)
- Identification without input price instruments?

  $\rightarrow$ Unresolved issue.

Our Paper's contribution

- Non-parametric identification of production function with unobserved heterogeneity
- Estimation of production function with random coefficients
- Empirical application: Japanese Census of Manufacture

Nonparametric Identification

Setup

- Panel Data: $\{\{Y_{it}, \tilde{L}_{it}, K_{it}, M_{it}, B_{it}\}_{t=1}^{T}\}_{i=1}^{n}$
  - Y_{it} Output; $\tilde{L}_{it}$ # of workers; K_{it} Capital stock; M_{it} Intermediate input; B_{it} Total wage bills
- K_{it}: Predetermined
- $\tilde{L}_{it}$, M_{it}: Flexibly chosen
- $\mathcal{I}_{it}$: Information available for choosing L_{it} & M_{it}

Production function and Capital

- J unobserved types with $\pi^j := \Pr(\text{type}=j)$

$$Y_{it} = e^{\omega_{it} + \epsilon_{it}} F_t^j(L_{it}, K_{it}, M_{it}), \quad \epsilon_{it}|\mathcal{I}_{it} \sim_{iid} g_{\epsilon,t}^j(\epsilon_{it})$$

$$\omega_{it} = h^j(\omega_{it-1}) + \eta_{it}, \quad \eta_{it}|\mathcal{I}_{it-1} \sim_{iid} g_\eta^j(\eta_{it})$$

- K_{it} predetermined with

$$P_t(K_{it}|\mathcal{I}_{t-1}, D_i = j) = P_t^j(K_{it}|K_{it-1}, \omega_{it-1})$$

Labor Input and Wage Bills

- L_{it} not directly observed but related to # of workers $\tilde{L}_{it}$ as

$$L_{it} = e^{\psi_t^j} \tilde{L}_{it}$$

- Total Wage Bills:

$$B_{it} = e^{v_{it} + \zeta_{it}} P_{L,t} L_{it}, \quad v_{it} | \mathcal{I}_{it-1} \sim_{iid} g_v^j(v_{it})$$

Average wage $B_{it}/\tilde{L}_{it}$

$\Rightarrow$ Identification of unobserved quality/hours ψ_t^j

L and M are flexible inputs

- L_{it} & M_{it} are chosen flexibly before observing ϵ_{it} & ζ_{it}:

$$
\begin{aligned}
(M_{it}, L_{it}) &= (\mathbb{M}_t^j(K_{it}, \omega_{it}, v_{it}), \mathbb{L}_t^j(K_{it}, \omega_{it}, v_{it})) \\
&:= \underset{(M,L)\in\mathcal{M}\times\mathcal{L}}{\operatorname{argmax}} \; P_{Y,t} E^j[e^{\epsilon_{it}}|\mathcal{I}_{it}] e^{\omega_{it}} F_t^j(K_{it}, L, M) - P_{M,t} M \\
&\qquad - E^j[e^{\zeta_{it}}|\mathcal{I}_{it}] e^{v_{it}} P_{L,t} L,
\end{aligned}
$$

- $(\mathbb{M}_t^j(K_{it}, \omega_{it}, v_{it}), \mathbb{L}_t^j(K_{it}, \omega_{it}, v_{it}))$ is invertible w.r.t. (ω_{it}, v_{it}) with probability one

- Correlation between (L_{it}, M_{it}) and ω_{it} $\Rightarrow$ Endogeneity

First Order Conditions as Identifying Restrictions

- Material & Labor Share Equations:

$$\underbrace{\frac{P_{M,t}M_{it}}{P_{Y,t}Y_{it}}}_{:=S_{it}^m} = \underbrace{\frac{\partial F_t^j(L_{it},K_{it},M_{it})/\partial M}{F_t^j(L_{it},K_{it},M_{it})/M_{it}}}_{:=G_{M,t}^j(L_{it},K_{it},M_{it})} E[e^{\epsilon}|\text{type=j}]e^{-\epsilon_{it}}$$

$$\underbrace{\frac{B_{it}}{P_{Y,t}Y_{it}}}_{:=S_{it}^\ell} = \underbrace{\frac{\partial F_t^j(L_{it},K_{it},M_{it})/\partial L}{F_t^j(L_{it},K_{it},M_{it})/L_{it}}}_{:=G_{L,t}^j(L_{it},K_{it},M_{it})} \frac{E[e^{\epsilon}|\text{type=j}]}{E[e^{\zeta}|\text{type=j}]} e^{\zeta_{it}-\epsilon_{it}}$$

- Material/labor share equation $\Rightarrow$ Source of Identificaiton

Main Nonparametric Identification Results

Under regularity conditions, the model structure

$$\theta = \{g^j_{\epsilon\zeta,t}(\cdot), g^j_v(\cdot), g^j_\eta(\cdot), h^j(\cdot), \pi^j, \{G^j_{M,t}(\cdot), G^j_{L,t}(\cdot), F^j_t(\cdot)\}^T_{t=1}, P_{L,t}, \psi^j_t\}^J_{j=1},$$

is nonparametrically identified from the joint distribution of $\underline{T = 4}$ periods of the panel data,

$$\Pr(\{Y_t, B_t, M_t, \tilde{L}_t, K_t\}^4_{t=1})$$

Or equivalently, consider the following data

$$\Pr(\{S^m_t, S^\ell_t, M_t, \tilde{L}_t, K_t\}^4_{t=1}).$$

Cobb-Douglas Case

$$\begin{aligned}
y_t &= \beta_0^j + \beta_m^j m_t + \beta_\ell^j(\psi_t^j + \tilde{\ell}_t) + \beta_k^j k_t + \omega_t + \epsilon_t \\
s_t^m &= \ln \beta_m^j + \ln E[e^\epsilon|\text{type=j}] - \epsilon_t \\
s_t^\ell - s_t^m &= \ln(\beta_\ell^j/\beta_m^j) - \ln\left(E_t^j[e^\zeta]\right) + \zeta_t \\
m_t - \tilde{\ell}_t &= \underbrace{\ln(P_{L,t}/P_{M,t})}_{:=\alpha_t} + \ln\left(\beta_m^j/\beta_\ell^j\right) + \ln\left(E_t^j[e^\zeta]\right) + \psi_t^j + v_t
\end{aligned}$$

We may identify:

- $\{\beta_0^j, \beta_m^j, \beta_\ell^j, \beta_k^j, \psi_t^j, g_{\epsilon\zeta,t}^j(\cdot)\}_{j,t}$
- $\{\pi^j, h^j(\cdot), g_\eta^j(\cdot), g_v^j(\cdot)\}_j$, $\{\alpha_t\}_{t=1}^T$

from the distribution of $\{S_t, \tilde{X}_t\}_{t=1}^T$ for $T \geq 4$.

Estimation

Two-stage Parametric MLE: Additional Assumptions

- J unobserved types
- Cobb-Douglas Production Function
- Assume normal distribution:

$$
\begin{aligned}
&(\epsilon_{it}, \zeta_{it}, v_{it})' \overset{d}{\sim} N\left(0, \Sigma^j\right) \\
&\omega_{it} = \rho_\omega^j \omega_{it-1} + \eta_{it}, \quad \eta_i \overset{d}{\sim} N(0, \sigma_\eta^j) \\
&k_{it} | (k_{it-1}, \omega_{it-1}) \overset{d}{\sim} N(\rho_{k0}^j + \rho_{kk}^j k_{it-1} + \rho_{k\omega}^j \omega_{it-1}, (\sigma_k^j)^2) \\
&(k_{i1}, \omega_{i1}) \overset{d}{\sim} N(\mu_1^j, \Sigma_1^j).
\end{aligned}
$$

Two-stage Parametric MLE

Probability Density Function of $(s_{it}, \tilde{x}_{it})_{t=1}^{T}$ for type j

$$
\begin{aligned}
f_t^j(\{s_{it}, \tilde{x}_{it}\}_{t=1}^{T}) = & \underbrace{\prod_{t=1}^{T} f_t^j(s_{it}, \tilde{\ell}_{it} - m_{it}; \theta_1^j, \alpha_t)}_{=L_{1i}(\theta_1^j, \boldsymbol{\alpha})} \\
\times \quad & \underbrace{f_1^j(\tilde{x}_{i1} | \tilde{\ell}_{i1} - m_{i1}; \theta^j) \prod_{t=2}^{T} f_t^j(\tilde{x}_{it} | \tilde{\ell}_{it} - m_{it}, \tilde{x}_{it-1}; \theta^j)}_{=L_{2i}(\theta_1^j, \theta_2^j)},
\end{aligned}
$$

Two-stage Parametric MLE

- First Stage:

$$\hat{\boldsymbol{\theta}}_1 = \underset{(\pi^j,\theta_1^j)_j,\boldsymbol{\alpha}}{\operatorname{argmax}} \sum_{i=1}^{N} \log \left(\sum_{j=1}^{J} \pi^j L_{1i}(\theta_1^j, \boldsymbol{\alpha}) \right)$$

- Second Stage: Given $((\hat{\theta_1^j})_j, \hat{\boldsymbol{\alpha}})$,

$$\hat{\boldsymbol{\theta}}_2 = \underset{\boldsymbol{\pi},(\theta_2^j)_j}{\operatorname{argmax}} \sum_{i=1}^{N} \log \left(\sum_{j=1}^{J} \pi^j L_{1i}(\hat{\theta}_1^j, \hat{\boldsymbol{\alpha}}) L_{2i}(\hat{\theta}_1^j, \theta_2^j) \right)$$

Two-Step Parameteric MLE

- Advantage: Short T ($T \geq 4$), based on identification argument, type is identified from all data
- Disadvantage: Parametric assumptions, Computationally difficult due to local maxima.
- EM algorithm

Empical Application:

Japanese Census of Manufacute,

1986-2010

Data set

- Unbalanced panel data, Plants with 30+ workers
- Variable Construction
  - K_{it}: Fixed asset less land. Perpetual inventory method.
  - L_{it}: Number of employees
  - M_{it}: Material + Energy + Subcontracting expenses for consigned production
  - Y_{it}: Sales
- Electric Audio Equipment: 907 plants $\times$ 4$\sim$25 years

Specification

- Cobb-Douglas (3 Types):

$$Y_i = \beta_0^j + \beta_m^j m_{it} + \beta_k^j k_{it} + \beta_\ell^j \ell_{it} + \omega_{it} + \epsilon_{it}, \quad \text{for } j = 1, 2, 3$$

- Unobserved hours or labor quality ($2 \times 3 = 6$ Types):

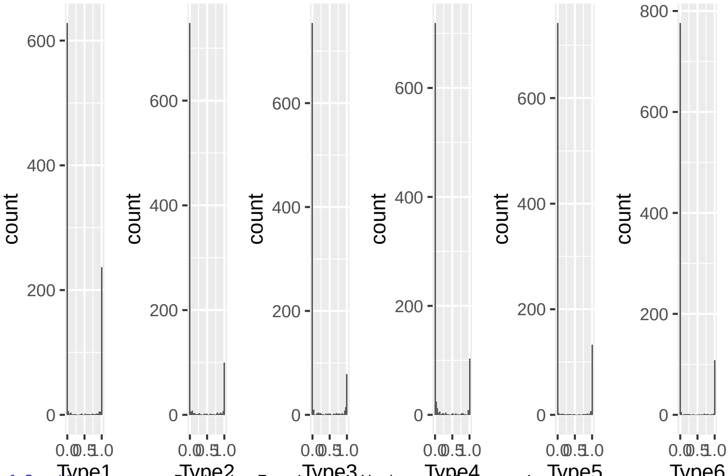
$$L_{it} = \psi^j \tilde{L}_{it} \quad \text{for } j = 1, 2, ..., 6$$

Estimates of Production Function (Electric Audio)

	J = 1	J = 6					
		Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
β_m^j	0.281	0.136		0.364		0.597	
	(0.008)	(0.007)		(0.022)		(0.010)	
β_ℓ^j	0.296	0.507		0.228		0.132	
	(0.006)	(0.012)		(0.012)		(0.006)	
β_k^j	0.076	0.122		0.185		0.202	
	(0.016)	(0.010)		(0.021)		(0.017)	
$\beta_m^j + \beta_\ell^j + \beta_k^j$	0.652	0.765		0.778		0.932	
β_k^j / β_ℓ^j	0.256	0.242		0.813		1.526	
ψ^j	0.000	-0.999	0.474	-0.258	0.201	-0.215	0.798
		(0.135)	(0.139)	(0.389)	(0.142)	(0.115)	(—)
π	1.000	0.278	0.137	0.134	0.158	0.164	0.128
		(0.023)	(0.017)	(0.163)	(0.149)	(0.015)	(0.014)

Posterior Probabilities for $J = 6$ Types (Electric Audio)

$$\hat{\pi}_i^j = \frac{\hat{\pi}^j L_i(\hat{\theta}^j)}{\sum_{k=1}^J \hat{\pi}^k L_i(\hat{\theta}^k)} \quad \text{for } j = 1, \ldots, J.$$



The Effect of ω_{it} on Investment (Electric Audio)

$$I_{it}/K_{it} = \alpha_0 + \alpha_\omega^j \omega_{it} + \text{quadratic of } k_{it}$$

	$J=1$	$J=6$					
		Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
α_ω^j	0.50	-0.19	-0.47	0.29	0.18	0.11	0.10
	(0.11)	(0.13)	(1.26)	(0.07)	(0.34)	(0.02)	(0.03)
β_m^j	0.28	0.14		0.36		0.60	
β_k^j/β_ℓ^j	0.26	0.24		0.81		1.53	

The effect of ω_{it} on I_{it}/K_{it} is larger among firms with

- high material share
- high capital-labor ratio

Quantile Regression (Electric Audio)

$$I_{it}/K_{it} = \alpha_0 + \alpha_\omega^j \omega_{it} + \text{quadratic of } k_{it}$$

	$J = 1$	$J = 6$					
		Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
$\alpha_\omega^j(0.10)$	0.00	0.00	0.00	0.01	0.00	-0.00	0.01
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
$\alpha_\omega^j(0.50)$	0.01	0.00	0.02	0.04	0.01	0.04	0.04
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
$\alpha_\omega^j(0.90)$	0.06	0.01	0.01	0.09	0.09	0.18	0.10
	(0.00)	(0.01)	(0.02)	(0.03)	(0.02)	(0.05)	(0.04)
β_m^j	0.28	0.14		0.36		0.60	
β_k^j/β_ℓ^j	0.26	0.24		0.81		1.53	

Conclusions

- Evidence for heterogeneity beyond Hick's neutral term
- Non-parametric identification
- Maximum likelihood estimation
- Empirical applications to Japanese Census Data
- Heterogeneous investment/export elasticities across unobserved types