Nonparametric Identification of Production Function, Total Factor Productivity, and Markup from Revenue Data

Hiroyuki Kasahara[1] Yoichi Sugita[2]

[1]Vancouver School of Economics, University of British Columbia

[2]Faculty of Economics, Hitotsubashi University

Issues

- Production function estimation is a core empirical tool for analyzing market outcomes (TFP and Markup).
- Firm-level dataset typically contains only firm's revenues but not output price/quantity.
- Can we identify markups and TFP separately from firm's revenue data without output price data?

Markups are estimated using revenue elasticity in practice

- Markup estimation via production function (Hall, 1988; De Loecker and Warzynski, 2012) is widely applied in various topics.
- Many published papers use revenue data to estimate output elasticity for computing markup (Javorcik and Poelhekke, 2017; Garcia-Marin and Voigtlander, 2019; Brandt et al., 2017; Varela, 2018):

$$\widehat{\frac{P_{it}}{MC_{it}}} = \frac{\partial \log \widehat{R}(M_{it}, K_{it}, L_{it}, \omega_{it}, z_{it}) / \partial \log M_{it}}{P_{it}^{M} M_{it} / R_{it}},$$

where

$$R_{it} = P(Y_{it}, z_{it}) Y_{it} \quad \text{with} \quad Y_{it} = F(M_{it}, K_{it}, L_{it}, \omega_{it}).$$

Bond, Hashemi, Kaplan, and Zoch (2020)

$$\log R_{it} = \log Y_{it} + \log P_{it}$$

$$\Rightarrow$$

$$\frac{\partial \log R_{it}/\partial \log M_{it}}{P_{it}^M M_{it}/R_{it}} = \frac{\partial \log Y_{it}/\partial \log M_{it}}{P_{it}^M M_{it}/R_{it}} + \frac{\partial \log P(Y_{it}, z_{it})/\partial \log M_{it}}{P_{it}^M M_{it}/R_{it}}$$

$$= \frac{P_{it}}{MC_{it}} + \text{Downward Bias}$$

$$= \mathbf{1}$$

Little is known about identification of markups from revenue data.

Contributions of this paper

- New constructive nonparametric identification of TFP, production function, and markup from revenue data
- “Standard” assumptions and data found in typical applications (e.g., Ackerberg et al. 2015)
- Various firm-level objects are identified
  - Gross production function, TFP, markup, price and quantity
  - Demand system and representative consumer’s utility

Related Literature

- Production function estimation:
    - Simultaneity bias: Marschak and Andrews (1944)
    - Control function approach: Olley and Pakes (1996); Levinsohn and Petrin (2003); Ackerberg, Caves, and Frazer (2015)
    - First Order Conditions: Doraszelski and Jaumandreu (2018); Gandhi, Navarro, and Rivers (2020)
- Markup: Klette and Griliches (1996); De Loecker (2011); De Loecker et al. (2016); Lu and Yu (2015); Nishioka and Tanaka (2019)

Related Literature

- Production function estimation from revenue data
  - Klette and Griliches (1996); De Loecker (2011): a constant and identical elastic demand
  - Katayama, Lu, and Tybout (2009): data on firm-level marginal costs are required
  - De Loecker, Eeckhout, and Unger (2020): an exogenous variable removing a price variation from revenue data
- Nonparametric identification of transformation model: Chiappori, Komunjer, and Kristensen (2015), Ekeland, Heckman, and Nesheim (2004).

$$\Lambda(Y) = f(X) + \epsilon, \ \Lambda'(Y) > 0.$$

Setup

- Demand elasticity depends on y and z:

$$p_{it} = \psi_t(y_{it}, z_{it})$$

- Revenue:

$$r_{it} = p_{it} + y_{it} = \underbrace{\psi_t(y_{it}, z_{it}) + y_{it}}_{:=\varphi_t(y_{it}, z_{it})},$$

- Profit maximization implies the monotonicity of φ_t:

$$0 < \frac{\partial \varphi_t(y_{it}, z_{it})}{\partial y_{it}} = \frac{MC}{P} \leq 1.$$

$$r_{it} = \varphi_t(y_{it}, z_{it}) \quad \Rightarrow \quad \varphi_t^{-1}(r_{it}, z_{it}) = y_{it}.$$

Setup

- Production function

$$y_{it} = f(x_{it}) + \omega_{it} \quad \text{with} \quad x_{it} = (m_{it}, k_{it}, \ell_{it})'$$

- Productivity process

$$\omega_{it} = h(\omega_{it-1}) + \eta_{it}, \quad \eta_{it} \overset{iid}{\sim} G_{\eta}$$

- Demand function for m_{it}

$$\begin{aligned} m_{it} &= \mathbb{M}_t(\omega_{it}, k_{it}, \ell_{it}, z_{it}) && \text{(demand function for } m_{it}\text{)} \\ \Rightarrow \quad \omega_{it} &= \mathbb{M}_t^{-1}(m_{it}, k_{it}, \ell_{it}, z_{it}) && \text{(by monotonicity)} \end{aligned}$$

Model and Structure

- Model

$$\varphi_t^{-1}(r_{it}, z_{it}) = f_t(x_{it}) + \omega_{it}, \quad \text{with} \quad x_{it} := (m_{it}, k_{it}, l_{it}),$$

$$\omega_{it} = h(\omega_{it-1}) + \eta_{it}, \quad \eta_{it} \overset{iid}{\sim} G_\eta$$

$$\omega_{it} = \mathbb{M}_t^{-1}(x_{it}, z_{it})$$

- Structure to be identified:

$$\{\mathbb{M}_t^{-1}(\cdot),\ \varphi_t^{-1}(\cdot),\ f_t(\cdot),\ h(\cdot),\ G_\eta\}$$

- Identification issues
    - Correlation of m_{it} and ω_{it}
    - Nonlinearity of $\varphi_t(\cdot)$

Location and Scale Normalization

- We aim to identify $\{\varphi_t^{-1}(\cdot), f_t(\cdot), \mathbb{M}_t^{-1}(\cdot)\}$ from constraint:

$$\varphi_t^{-1}(r_{it}, z_{it}) = f_t(x_{it}) + \mathbb{M}_t^{-1}(x_{it}, z_{it})$$
$$\mathbb{M}_t^{-1}(x_{it}, z_{it}) = h\left(\mathbb{M}_{t-1}^{-1}(x_{it-1}, z_{it-1})\right) + \eta_{it}$$

- For any constant $(a_1, a_2, b) \in \mathbb{R}^2 \times \mathbb{R}_{++}$,

$$\tilde{\varphi}_t^{-1} = a_1 + a_2 + b\varphi_t^{-1},\ \tilde{f}_t = a_1 + bf_t,\ \tilde{\mathbb{M}}_t^{-1} = a_2 + b\mathbb{M}_t^{-1}$$

also satisfies the above restriction.

Step 1: Identification of Control Function M_t^{-1}

$$\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it}) = \underbrace{\bar{h}_t(x_{it-1}, z_{it-1})}_{=h\left(\mathbb{M}_{t-1}^{-1}(x_{it-1}, z_{it-1})\right)} + \eta_{it},$$

where

$$v_{it} := (k_{it}, \ell_{it}, x_{it-1}, z_{it-1}) \perp\!\!\!\perp \eta_{it},$$

$$m_{it} \not\perp\!\!\!\perp \eta_{it},$$

$$\frac{\partial \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})}{\partial m_t} > 0.$$

We apply the identification argument for transformation models by Chiappori et al. (2015).

Step 1: Identification of Control Function M_t^{-1}

Assumption (Normalization)

For some $(m_0^*, m_1^*, k^*, l^*, z^*)$,

$$f_t\left(m_0^*, k^*, l^*\right) = 0,$$
$$\mathbb{M}_t^{-1}\left(m_0^*, k^*, l^*, z^*\right) = 0, \text{ and } \mathbb{M}_t^{-1}\left(m_1^*, k^*, l^*, z^*\right) = 1.$$

Proposition

Under regularity conditions, we may identify $\mathbb{M}_t^{-1}\left(m_t, k_t, l_t, , z_t\right)$ *up to scale and location from the conditional distribution of* m_t *given* $v_t := (k_t, l_t, z_t, x_{t-1}, z_{t-1})$*, denoted by* $G_{m_t|v_t}(m|v)$.

Proof (Chiappori et al., 2015)

Because $\eta \perp\!\!\!\perp v = (k_t, \ell_t, z_t, x'_{t-1}, z_{t-1})'$ and
$\eta_t = \mathbb{M}_t^{-1}(x_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})$,

$$\underbrace{G_{m_t|v_t}(m|v_t)}_{\text{Data}} = G_\eta\left(\mathbb{M}_t^{-1}(m, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right)$$

Differentiate by $q_t \in \{x_t, z_t\}$ and $q_{t-1} \in \{x_{t-1}, z_{t-1}\}$ at some point $(\tilde{x}_{t-1}, \tilde{z}_{t-1})$:

$$\frac{\partial G_{m_t|v_t}(m_t|\tilde{v}_t)}{\partial q_t} = \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial q_t} g_{\eta_t}\left(\mathbb{M}_t^{-1}(x_t, z_t) - \bar{h}_t(\tilde{x}_{t-1}, \tilde{z}_{t-1})\right)$$

$$\frac{\partial G_{m_t|v_t}(m_t|\tilde{v}_t)}{\partial q_{t-1}} = \frac{\partial \bar{h}_t(\tilde{x}_{t-1}, \tilde{z}_{t-1})}{\partial q_{t-1}} g_{\eta_t}\left(\mathbb{M}_t^{-1}(x_t, z_t) - \bar{h}_t(\tilde{x}_{t-1}, \tilde{z}_{t-1})\right)$$

$\tilde{v}_t := (k_t, l_t, z_t, \tilde{x}_{t-1}, \tilde{z}_{t-1})$

Proof

$$\Rightarrow \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial m_t} = -\underbrace{\frac{\partial G_{m_t|v_t}(m|\tilde{v}_t)/\partial m_t}{\partial G_{m_t|v_t}(m|\tilde{v}_t)/\partial q_{t-1}}}_{\text{Data}} \frac{\partial \bar{h}_t(\tilde{x}_{t-1}, \tilde{z}_{t-1})}{\partial q_{t-1}}$$

From the location and scale normalization,

$$\begin{aligned}
1 &= \mathbb{M}_t^{-1}(m_1^*, k^*, l^*, z^*) \\
&= \underbrace{\mathbb{M}_t^{-1}(m_0^*, k^*, l^*, z^*)}_{=0} + \int_{m_0^*}^{m_1^*} \frac{\partial \mathbb{M}_t^{-1}(u, k^*, l^*, z^*)}{\partial m_t} du \\
&= -\frac{\partial \bar{h}_t(x_{t-1}^*, z_{t-1}^*)}{\partial q_{t-1}} \underbrace{\int_{m_0^*}^{m_1^*} \frac{\partial G_{m|v}(u|k^*, l^*, z^* \tilde{x}_{t-1}, \tilde{z}_{t-1})/\partial m_t}{\partial G_{m|v}(u|k^*, l^*, z^*, \tilde{x}_{t-1}, \tilde{z}_{t-1})/\partial q_{t-1}} du}_{\equiv S_t}
\end{aligned}$$

$$\Rightarrow \frac{\partial \bar{h}_t(x_{t-1}^*, z_{t-1}^*)}{\partial q_{t-1}} = -1/S_t.$$

Proof

Therefore, we identify

$$\frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial q_t} = \frac{1}{S_t} \frac{\partial G_{m|v}(m_t|\tilde{v}_t)/\partial q_t}{\partial G_{m|v}(m_t|\tilde{v}_t)/\partial q_{t-1}} \quad \text{for } q_t \in \{m_t, k_t, l_t, z_t\}.$$

$\Rightarrow$

$$\mathbb{M}_t^{-1}(x_t, z_t) = \underbrace{\mathbb{M}_t^{-1}(m_0^*, k^*, l^*, z^*)}_{=0} + \int_{m_0^*}^{m_t} \frac{\partial \mathbb{M}_t^{-1}(s, k_t, l_t, z_t)}{\partial m_t} ds$$

$$+ \int_{k^*}^{k_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*, s, l_t, z_t)}{\partial k_t} ds + \int_{l^*}^{l_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, s, z_t)}{\partial l_t} ds$$

$$+ \int_{z^*}^{z_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, l_t^*, s)}{\partial z_t} ds.$$

□

Step 3: Identification of Markup and Production Function

Assumption (F.O.C. with respect to *m*)

$$\frac{\partial f(x_{it})}{\partial m_{it}} = \frac{\partial \varphi_t^{-1}(r_{it}, z_{it})}{\partial r_{it}} \frac{P_{it}^M M_{it}}{R_{it}}$$

for all observations.

Proposition

Under regularity conditions, we can identify $\varphi_t^{-1}(\cdot)$ and $f_t(\cdot)$ up to scale and location and each firm's markup $\partial \varphi_t^{-1}(r_{it}, z_{it})/\partial r_t$ up to scale.

Proof: Identification of Markup

Differentiate $E[r_{it}|x_{it}, z_{it}] = \varphi_t\left(f_t(x_{it}) + \mathbb{M}_t^{-1}(x_{it}, z_{it}), z_{it}\right)$ by m_t

$$\underbrace{\frac{\partial E[r_{it}|x_{it}, z_{it}]}{\partial m_t}}_{\text{known}} = \frac{\partial \varphi_t(y_{it}, z_{it})}{\partial y_t}\left(\frac{\partial f_t(x_{it})}{\partial m_t} + \underbrace{\frac{\partial \mathbb{M}_t^{-1}(x_{it}, z_{it})}{\partial m_t}}_{\text{known}}\right)$$

From F.O.C. for m_{it},

$$\frac{\partial \varphi_t(y_{it}, z_{it})}{\partial y_t}\frac{\partial f_t(x_{it})}{\partial m_t} = \underbrace{\frac{P_{it}^M M_{it}}{R_{it}}}_{\text{known}},$$

we can identify the markup

$$\mu_{it} = \frac{1}{\partial \varphi_t(y_{it}, z_{it})/\partial y_t} = \frac{\partial \mathbb{M}_t^{-1}(x_{it}, z_{it})}{\partial m_t}\left(\frac{\partial E[r_{it}|x_{it}, z_{it}]}{\partial m_t} - \frac{P_{it}^M M_{it}}{R_{it}}\right)^{-1}.$$

Identification of Elasticities

- $\partial f_t(x_{it}) / \partial q_t$ for $q_t \in \{m_t, k_t, l_t\}$ are identified from:

$$\underbrace{\frac{\partial E[r_{it}|x_{it}, z_{it}]}{\partial q_t}}_{\text{known}} = \underbrace{\frac{\partial \varphi_t(y_{it}, z_{it})}{\partial y_{it}}}_{\text{known}} \left(\frac{\partial f_t(x_{it})}{\partial q_t} + \underbrace{\frac{\partial \mathbb{M}_t^{-1}(x_{it}, z_{it})}{\partial q_t}}_{\text{known}} \right).$$

- $\partial \varphi_t(y_{it}, z_{it}) / \partial z_{it}$ is identified from:

$$\underbrace{\frac{\partial E[r_{it}|x_{it}, z_{it}]}{\partial z_t}}_{\text{known}} = \frac{\partial \varphi_t(y_{it}, z_{it})}{\partial z_{it}} + \underbrace{\frac{\partial \varphi_t(y_{it}, z_{it})}{\partial y_{it}} \frac{\partial \mathbb{M}_t^{-1}(x_{it}, z_{it})}{\partial z_t}}_{\text{known}}.$$

Identification of Production Function

$$f_t(x_t) = \underbrace{f_t(m_0^*, k^*, l^*)}_{=0} + \int_{m_0^*}^{m_t} \frac{\partial f_t(s, k_t, l_t)}{\partial m_t} ds + \int_{k^*}^{k_t} \frac{\partial f_t(m_0^*, s, l_t)}{\partial k_t} ds + \int_{l^*}^{l_t} \frac{\partial f_t(m_0^*, k^*, s)}{\partial l_t} ds.$$

Identification of Price and Quantity

- Price and quantity are identified by

$$y_{it} = f_t(x_{it}) + \mathbb{M}_t^{-1}\left(x_{it}, z_{it}\right)$$
$$p_{it} = r_{it} - y_{it}.$$

- $\varphi_t(y, z)$ can be identified by integrating $\partial \varphi_t\left(y_{it}, z_{it}\right) / \partial y_{it}$ and $\partial \varphi_t\left(y_{it}, z_{it}\right) / \partial z_{it}$ as for $f_t(\cdot)$.

Location and scale normalization across different periods

- $\{\varphi_t^{*-1}, f_t^*, \mathbb{M}_t^{*-1}\}$: true structure
- $\{\varphi_t^{-1}, f_t, \mathbb{M}_t^{-1}\}$: identified structure.
- Location and scale normalization, (a_{1t}, a_{2t}, b_t):

$$\varphi_t^{-1} = a_{1t} + b_t \varphi_t^{*-1},\ f_t = a_{2t} + b_t f_t^*,\ \mathbb{M}_t^{-1} = a_t + b_t \mathbb{M}_t^{-1*}.$$

- Location and scale normalization differs period-by-period:

$$(a_{1t}, a_{2t}, b_t) \neq (a_{1,t+1}, a_{2,t+1}, b_{t+1}).$$

- The relationship between b_t and b_{t+1} needs to be identified to know how productivities and markups change over periods.

Location and scale normalization across different periods

b_t and b_{t+1} can be linked by assuming the stability of some function:

1. Variance $Var(\eta_t) = Var(\eta_{t+1})$
2. Elasticity $\partial f_t(x)/\partial q = \partial f_t(x)/\partial q$ for some x and some factor $q \in \{m, k, l\}$
3. Local returns to scale: for some $c > 0$ and x_t,

$$\frac{\partial f_t(x_t)}{\partial m_t} + \frac{\partial f_t(x_t)}{\partial l_t} + \frac{\partial f_t(x_t)}{\partial k_t} = c \qquad \text{for } t \text{ and } t+1.$$

Identification in Alternative Settings

- Ex-post iid demand elasticity shock ε_{it}:

$$r_{it} = \varphi(y_{it} e^{\varepsilon_{it}}, z_{it}),$$

  where ε_{it} is not known when m_{it} is chosen.

- Additively separable iid measurement error ϵ_{it}:

$$r_{it} = \varphi(y_{it}, z_{it}) + \epsilon_{it}.$$

- Ex-post iid shock e_t to TFP:

$$r_{it} = \varphi_t \left(f(x_t) + \omega_t + e_t \right).$$

Identification in Alternative Settings

- Endogenous l_t and z_t
  - l_{t-1} and z_{t-1} can be used IVs as in Ackerberg et al. (2007)
- Discrete z_t
- Unobserved demand shifter (quality) ξ_{it}:
  - Identify quality-adjusted variables $\hat{\omega}_t := \omega_t + \xi_t$, $\hat{y}_t := y_t + \xi_t$, and $\hat{p}_t := p_t - \xi_t$

Semi-Parametric Estimator (in Progress)

- Cobb-Douglas production function

$$f(x) = \theta_m m_t + \theta_k k_t + \theta_l l_t.$$

$$\Rightarrow \quad \mathbb{M}_t^{-1}(x_t, z_t) = \lambda_t(m_t, z_t) - \theta_k k_t - \theta_l l_t.$$

- A semi-para transformation model for $\omega_t = h(\omega_{t-1}) + \eta_t$:

$$\lambda_t(m_t, z_t; \alpha^m) = \theta_k k_t + \theta_l l_t + \bar{h}_t(x_{t-1}, z_{t-1}; \alpha^h) + \eta_t$$

where $\lambda_t(m_t, z_t; \alpha^m)$ and $\bar{h}_t(x_{t-1}, z_{t-1}; \alpha^h)$ are B-splines.

Semi-Parametric Estimator (in Progress)

The profile likelihood estimator by Linton, Sperlich, and Van Keilegom (2008)

$$\alpha_{PL}^m \in \arg\min_{\alpha^m} \sum_{i=1}^{N} \ln g_{\eta_t}(\eta_{it}(\alpha^m)) + \ln \frac{\partial \lambda_t \left(m_{it}, z_{it}; \alpha^m\right)}{\partial m_t}$$

- $g_{\eta_t}(\eta_{it})$ is estimated by kernel-density
- the scale and location normalization of $\lambda_t \left(m_{it}, z_{it}; \alpha^m\right)$
- monotonicity and positivity constraints

Semi-Parametric Estimator (in Progress)

- First Order Condition

$$\frac{\partial \hat{f}_t(x_{it})}{\partial m_{it}} = \frac{\partial \hat{\mathbb{M}}_t^{-1}(x_{it}, z_{it})}{\partial m_t} \left(\frac{\partial \widehat{E[r_{it}|x_{it}, z_{it}]}}{\partial m_t} - \frac{P_{it}^M M_{it}}{R_{it}} \right)^{-1} \frac{P_{it}^M M_{it}}{R_{it}}.$$

- Material elasticity

$$\hat{\theta}_m = median \left\{ \frac{\partial \hat{f}_t(x_{it})}{\partial m_t} \right\}$$

- Output and price

$$\hat{y}_{it} = \hat{\theta}_m m_{it} + \hat{\theta}_k k_{it} + \hat{\theta}_l l_{it} + \hat{\mathbb{M}}_t^{-1}(x_t, z_t)$$

$$\hat{p}_{it} = r_{it} - \hat{y}_{it}$$

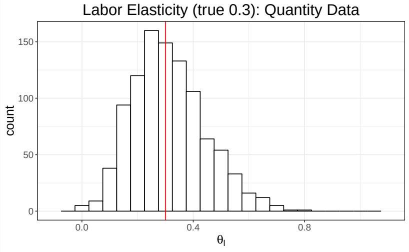
Simulation

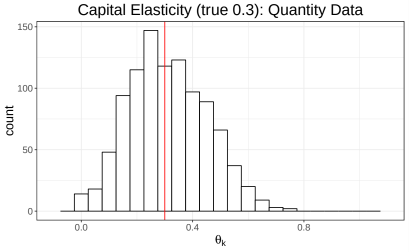
- Cobb-Douglas production, constant response demand (Mrázová and Neary, 2019), AR1 TFP process:

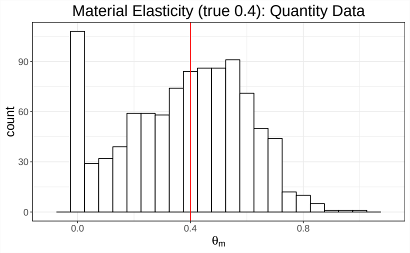
$$
\begin{aligned}
y_{it} &= 0.4m_{it} + 0.3k_{it} + 0.3l_{it} + \omega_{it} \\
r_{it} &= 2.3 + 2.03 \ln\left(\exp\left(0.35 y_{it}\right) + 0.6\right) + \epsilon_{it}, \;\; \epsilon_{it} \sim N(0,4) \\
\omega_{it} &= 0.7\omega_{it-1} + \eta_{it}, \;\; \eta_{it} \sim N(0,2)
\end{aligned}
$$

- 1000 samples of 300 firms
- Markup distribution: (P25, P50, P75)=(1.7, 2.2, 3.7)
- Estimation method: CRS-imposed Ackerberg et al. (2015) by Flynn, Gandhi, and Traina (2019)

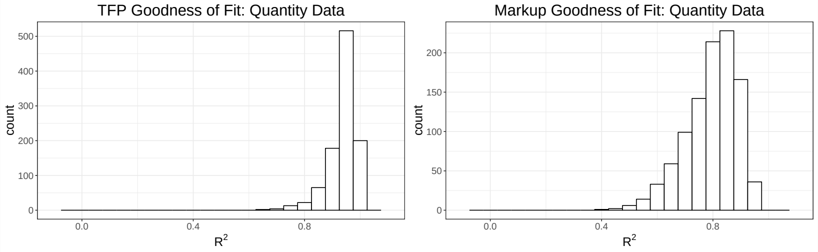
ACF with Quantity Data



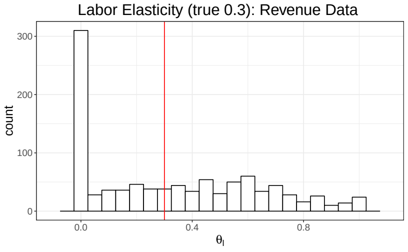


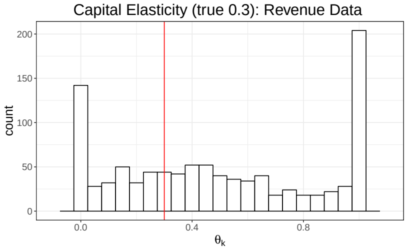


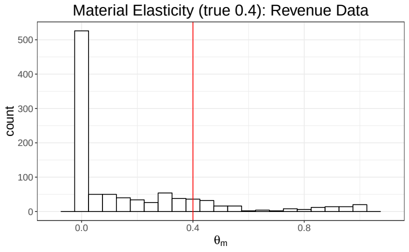
ACF with Quantity Data



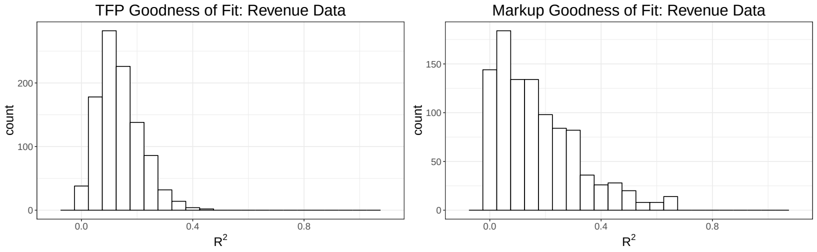
ACF with Revenue Data



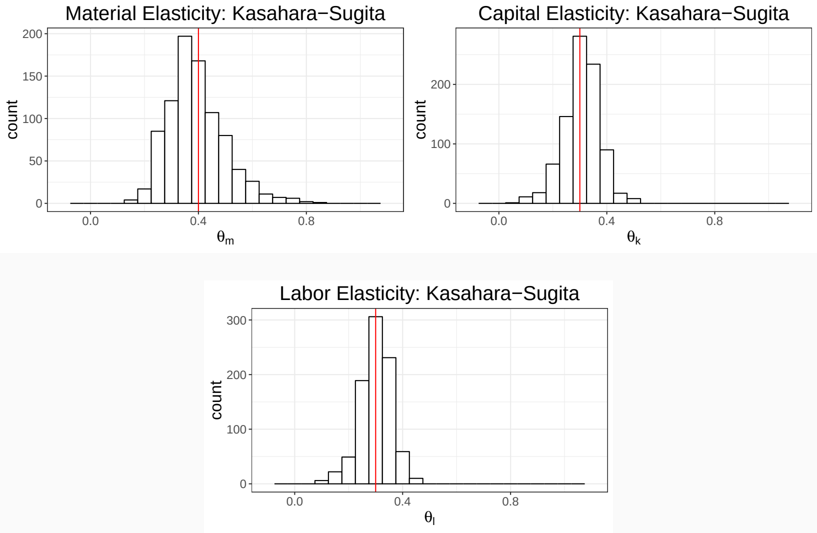




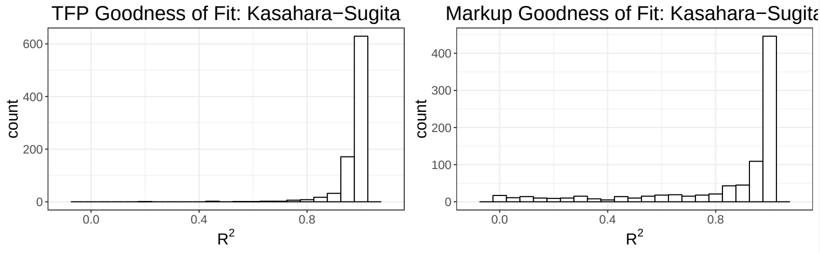
ACF with Revenue Data



Our Method with Revenue Data



Our Method with Revenue DataTFP and Markups



Conclusion

- Using revenue as output biases production function and markup estimation.
- This paper provides identification of production function and markup from revenue data.
- When revenue is modeled as a function of output (instead of a mere proxy of output), various economic objects can be identified from revenue data.
- We plain in an ongoing, followup research to estimate these objects from an actual dataset.

Identification of Demand System

- Homothetic utility with a single index (HSA)(Matsuyama and Ushchev, 2017)

$$\frac{P_{it}Y_{it}}{H_t} = S\left(\frac{Y_{it}}{A(\mathbf{Y},\mathbf{z})}, z_{it}\right)$$

  $H_t \equiv \sum_{i \in I} P_{it}Y_{it}$ is the total industry expenditure; $A(\mathbf{Y},\mathbf{z})$ an index of Y_{it}, z_{it}.

- HSA nests the CES demand, the translog demand and the constant response demand

Identification of Demand System

Proposition
(Matsuyama and Ushchev, 2017, Lemma1). Suppose $\varphi_t(y)$ is identified up to location (e.g., local CRS is assumed). Define

$$S(Y_{it}, z_{it}) \equiv \frac{\exp\left(\varphi_t\left(\ln Y_{it}, z_{it}\right)\right)}{H_t}$$

and identify $A\left(\mathbf{Y}', \mathbf{z}'\right)$ by solving

$$\sum_{i \in I_t} S\left(\frac{\mathbf{Y}'}{\mathbf{A}(\mathbf{Y}', \mathbf{z}')}, \mathbf{z}'\right) = 1.$$

Then, $\{S(\cdot), A(\cdot)\}$ is a HSA demand system.

Identification of Utility Function

Theorem

(Matsuyama and Ushchev, 2017, Lemma1).There exists a unique monotone, convex, continuous and homothetic rational preference that generates the HSA demand system. The associated utility function U is obtained as

$$\ln U(Y, z_t) = \ln A(Y, z_t) + \sum_{i \in I_t} \int_{c(z_t)}^{Y_i/A(Y, z_t)} \frac{S(\xi, z_{it})}{\xi} d\xi \quad (1)$$

where $c(z_t)$ is a constant for given z_t.

References

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007), "Econometric Tools for Analyzing Market Outcomes," *Handbook of econometrics*, 6, 4171–4276.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015), "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83, 2411–2451.

Reference ii

Bond, S., Hashemi, A., Kaplan, G., and Zoch, P. (2020), "Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation from Production Data," NBER Working Paper w27002.

Brandt, L., Biesebroeck, J. V., Wang, L., and Zhang, Y. (2017), "WTO accession and performance of Chinese manufacturing firms." *American Economic Review*, 107, 2784–2820.

Chiappori, P.-A., Komunjer, I., and Kristensen, D. (2015), "Nonparametric Identification and Estimation of Transformation Models," *Journal of Econometrics*, 188, 22–39.

Reference iii

De Loecker, J. (2011), "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79, 1407–1451.

De Loecker, J., Eeckhout, J., and Unger, G. (2020), "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 135, 561–644.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016), "Prices, Markups, and Trade Reform," *Econometrica*, 84, 445–510.

De Loecker, J. and Warzynski, F. (2012), "Markups and Firm-Level Export Status," *American Economic Review*, 102, 2437–71.

Reference iv

Doraszelski, U. and Jaumandreu, J. (2018), "Measuring the Bias of Technological Change," *Journal of Political Economy*, 126, 1027–1084.

Ekeland, I., Heckman, J. J., and Nesheim, L. (2004), "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 112, S60–S109.

Flynn, Z., Gandhi, A., and Traina, J. (2019), "Measuring Markups with Production Data," Unpublished.

Gandhi, A., Navarro, S., and Rivers, D. A. (2020), "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, forthcoming.

Reference v

Garcia-Marin, A. and Voigtlander, N. (2019), “Exporting and Plant-Level Efficiency Gains: It’s in the Measure,” *Journal of Political Economy*, 127, 1777–1825.

Hall, R. E. (1988), “The Relation between Price and Marginal Cost in US Industry,” *Journal of Political Economy*, 96, 921–947.

Javorcik, B. and Poelhekke, S. (2017), “Former foreign affiliates: Cast out and outperformed?” *Journal of the European Economic Association*, 15, 501–539.

Katayama, H., Lu, S., and Tybout, J. R. (2009), “Firm-Level Productivity Studies: Illusions and a Solution,” *International Journal of Industrial Organization*, 27, 403–413.

Klette, T. J. and Griliches, Z. (1996), “The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous,” *Journal of Applied Econometrics*, 11, 343–361.

Levinsohn, J. and Petrin, A. (2003), “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies*, 317–341.

Linton, O., Sperlich, S., and Van Keilegom, I. (2008), “Estimation of a semiparametric transformation model,” *The Annals of Statistics*, 36, 686–718.

Lu, Y. and Yu, L. (2015), "Trade liberalization and markup dispersion: evidence from China's WTO accession," *American Economic Journal: Applied Economics*, 7, 221–53.

Marschak, J. and Andrews, W. (1944), "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 143–205.

Matsuyama, K. and Ushchev, P. (2017), "Beyond CES: Three Alternative Cases of Flexible Homothetic Demand Systems," Buffett Institute Global Poverty Research Lab Working Paper 17-109.

Mrázová, M. and Neary, J. P. (2019), "IO for Exports(s)," *International Journal of Industrial Organization*, 102561.

Reference viii

Nishioka, S. and Tanaka, M. (2019), "Measuring Markups from Revenue and Total Cost: An Application to Japanese Plant-Product Matched Data," Rieti Discussion Paper Series 19-E-018.

Olley, G. S. and Pakes, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1263–1297.

Varela, L. (2018), "Reallocation, Competition, and Productivity: Evidence from a Financial Liberalization Episode," *Review of Economic Studies*, 85, 1279–1313.