Nonparametric Identification and Estimation of Production Function, Total Factor Productivity, and Consumer Demand Function from Firm-Level Revenue Data

Chun Pang Chow[1] Hiroyuki Kasahara[1] Yoichi Sugita[2]

[1]Department of Economics, University of British Columbia

[2]Faculty of Business and Commerce, Keio University

Introduction

- Production function estimation is widely used for analyzing market outcomes
  - Total factor productivity (TFP), technological change
  - Price markup, wage markdown (labor monopsony)
- Firm-level datasets typically contain only revenue but not output quantity.
- Many studies use revenue (deflated by an industry price index) as a proxy for output $\Rightarrow$ Bias in general!
- Can we still identify production function, TFP, markups, and consumer demand system from firm's revenue data without output quantity data?

What This Paper Does

- We identify:
  - Gross production function, TFP, markup, output quantity
  - Nonparametric Homothetic Single Aggregator (HSA) demand system by Matsuyama and Ushchev (2017) and consumer utility function
  - Counterfactual analysis of market outcomes and welfare
- Semi-parametric estimator and simulation
- Application to Chilean manufacturing plant data
  - Welfare analysis of firm's market power

Model Setup

Setup

- Price p depends on output y and temporary demand shock ϵ:

$$p_{it} = \tilde{D}_t(y_{it}, \epsilon_{it}) = D_t(y_{it}, u_{it}), \quad u_{it} := F_\epsilon(\epsilon_{it}) \overset{iid}{\sim} U[0,1]$$

where $u_{it} \perp\!\!\!\perp \mathbb{I}_{it-2}$ (temporary).

- Revenue:

$$r_{it} = D_t(y_{it}, u_{it}) + y_{it} =: \varphi_t(y_{it}, u_{it}),$$

- Markup:

$$\text{Markup}_{it} := \left(\frac{\partial \varphi_t(y_{it}, u_{it})}{\partial y_{it}} \right)^{-1} \geq 1.$$

Setup

- Output y_{it}, Production Function $f_t(.)$, and TFP ω_{it}:

$$y_{it} = f_t\left(x_{it}\right) + \omega_{it}, \text{ with } \quad x_{it} = (m_{it}, k_{it}, \ell_{it})'$$
$$\omega_{it} = h\left(\omega_{it-1}\right) + \eta_{it}, \quad \eta_{it} \overset{iid}{\sim} G_\eta$$

- Flexible: m_{it}
- Fixed: $k_{it} = \mathbb{K}(\mathbb{I}_{it-1})$ and $l_{it} = \mathbb{L}(\mathbb{I}_{it-1})$ (Timing assumption)
- Independence
  - $\eta_{it} \perp\!\!\!\perp k_{it}, l_{it}, m_{it-1}$
  - $\eta_{it} \perp\!\!\!\perp u_{it}$

Structure

- Profit maximization:

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

- Model:

$$\underbrace{\mathbb{M}_t^{-1}\left(x_{it}, u_{it}\right)}_{\omega_{it}} = h\left(\underbrace{\mathbb{M}_{t-1}^{-1}\left(x_{it-1}, u_{it-1}\right)}_{\omega_{it-1}}\right) + \eta_{it}, \quad \eta_{it} \overset{iid}{\sim} G_\eta,$$

$$r_{it} = \varphi_t\left(\overbrace{f_t\left(x_{it}\right) + \mathbb{M}_t^{-1}\left(x_{it}, u_{it}\right)}^{y_{it}}, u_{it}\right) \qquad u_{it} \overset{iid}{\sim} U[0,1]$$

- Structure to be identified:

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

Identification

Simpler Setting for Today's Presentation

- AR(1) TFP process $\omega_{it} = \rho\omega_{it-1} + \eta_{it}$
- Cobb-Douglas Production Function

$$y_{it} = \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + \omega_{it}$$

$\Rightarrow$ Separability on control function

$$\omega_{it} = \mathbb{M}_t^{-1}(x_{it}, u_{it}) = \lambda_t\left(m_{it}, u_{it}\right) - \theta_k k_{it} - \theta_l l_{it}$$

simplifies

$$\begin{aligned} r_{it} &= \varphi_t\left(\overbrace{\theta_m m_{it} + \lambda_t\left(m_{it}, u_{it}\right)}^{y_{it}}, u_{it}\right) \\ &= \phi_t\left(m_{it}, u_{it}\right) \end{aligned}$$

- MA(1) demand shock $\epsilon_{it} \Rightarrow u_{it} \perp\!\!\!\perp m_{it-2}$

Step 1: Identification of the rank of demand shock u_{it}

- Revenue inversion by monotonicity

$$r_{it} = \phi_t\left(m_{it}, u_{it}\right) \Rightarrow u_{it} = \phi_t^{-1}\left(m_{it}, r_{it}\right) \overset{iid}{\sim} U[0,1]$$

- Key assumption: $u_{it} \perp\!\!\!\perp m_{it-2}$ (demand shock is temporary)
- IV Quantile Regression (Chernozhukov and Hansen, 2005)

$$E\left[u - 1\left\{\underbrace{\phi_t^{-1}\left(m_{it}, r_{it}\right)}_{=u_{it}} \leq u\right\} |m_{it-2}\right] = 0 \text{ for all } u \in [0,1]$$

$$\Longrightarrow \phi_t(\cdot) \text{ and } u_{it} = \phi_t^{-1}(x_{it}, r_{it}) \text{ are identified}$$

- Next step: Identification of Control Function $\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, u_{it})$

Step 2: Identification of Control Function $\mathbb{M}_t^{-1}(.)$

Assumption
(Normalization) for some $(m_0^*, m_1^*, k^*, l^*, u^*)$,

$$\mathbb{M}_t^{-1}(m_0^*, k^*, l^*, u^*) = 0 \text{ and } \mathbb{M}_t^{-1}(m_1^*, k^*, l^*, u^*) = 1$$

Proposition
Under regularity conditions, we can identify $\omega_t = \mathbb{M}_t^{-1}(m_t, k_t, l_t, u_t)$ *up to scale and location from conditional CDF* $G_{m_t|v_t}(m_t|v_t)$ *for given* $v_t \equiv (k_t, l_t, u_t, x_{t-1}, u_{t-1})$.

Proof

$$\overbrace{\mathbb{M}_t^{-1}(x_t, u_t)}^{\omega_{it}} = \rho \overbrace{\mathbb{M}_{t-1}^{-1}(x_{t-1}, u_{t-1})}^{\omega_{it-1}} + \eta_t, \qquad \text{(AR(1) TFP process)}$$

$$\mathbb{M}_t^{-1}(x_t, u_t) = \lambda_t(m_t, u_t) - \theta_k k_t - \theta_\ell \ell_t \qquad \text{(Cobb-Douglas Separability)}$$

$\Rightarrow$

$$\eta_t = \Delta_\rho \lambda_t(m_t, u_t) - \theta_k \Delta_\rho k_t - \theta_\ell \Delta_\rho \ell_t,$$

where

$$\Delta_\rho k_t = k_t - \rho k_{t-1} \quad \text{etc.},$$

Proof

$$\eta_t = \overbrace{\mathbb{M}_t^{-1}\left(m_t, k_t, l_t, u_t\right)}^{\omega_t} - \rho \overbrace{\mathbb{M}_{t-1}^{-1}\left(x_{t-1}, u_{t-1}\right)}^{\omega_{t-1}} \quad \text{(AR(1) TFP process)}$$

Because of

(1) Monotonicty between m_t and ω_t

(2) $\eta \perp\!\!\!\perp v = (k_t, \ell_t, u_t, x'_{t-1}, u_{t-1})'$,

$$\begin{aligned}
&\underbrace{G_{m_t|v_t}\left(m_t|v_t\right)}_{\text{Observed}} \\
&= G_{\eta_t}\left(\mathbb{M}_t^{-1}\left(m_t, k_t, l_t, u_t\right) - \rho \mathbb{M}_{t-1}^{-1}\left(x_{t-1}, u_{t-1}\right) | v_t\right) \\
&= G_{\eta_t}\left(\Delta_\rho \lambda_t(m_t, u_t) - \theta_k \Delta_\rho k_t - \theta_\ell \Delta_\rho \ell_t\right)
\end{aligned}$$

Proof

Key idea: Differentiate and Divide! (Chiappori, Komunjer, and Kristensen, 2015)

$$\underbrace{G_{m_t|v_t}(m_t|v_t)}_{\text{Observed}} = G_{\eta_t}\left(\Delta_\rho \lambda_t(m_t, u_t) - \theta_k k_t + \rho\theta_k k_{t-1} - \theta_\ell \Delta_\rho \ell_t\right)$$

- Differentiation w.r.t. k_t and k_{t-1}:

$$\frac{\partial G_{m|v}(m_t|v_t)}{\partial k_t} = -g_\eta(\eta_t)\theta_k$$
$$\frac{\partial G_{m|v}(m_t|v_t)}{\partial k_{t-1}} = g_\eta(\eta_t)\rho\theta_k$$

$\Longrightarrow$

$$\rho = -\frac{\partial G_{m|v}(m_t|v_t)/\partial k_t}{\partial G_{m|v}(m_t|v_t)/\partial k_{t-1}} \text{ is identified.}$$

Proof

- Differentiation w.r.t. m_t and k_t:

$$\frac{\partial \lambda_t (m_t, u_t)}{\partial m_t} = \frac{\partial G_{m|v}(m_t|v_t)/\partial m_t}{\partial G_{m|v}(m_t|v_t)/\partial k_t}\, \theta_k$$

- From the location and scale normalization,

$$\begin{aligned} 1 &= \mathbb{M}_t^{-1}(m_1^*, k^*, l^*, u^*) - \mathbb{M}_t^{-1}(m_0^*, k^*, l^*, u^*) \\ &= \lambda_t (m_1^*, u^*) - \lambda_t (m_0^*, u^*) \\ &= \int_{m_0^*}^{m_1^*} \frac{\partial \lambda_t (m, u^*)}{\partial m_t} dm \\ &= \theta_k \underbrace{\int_{m_0^*}^{m_1^*} \frac{\partial G_{m|v}(m|v_t^*)/\partial m_t}{\partial G_{m|v}(m|v_t^*)/\partial k_t} dm}_{\equiv S_t} \end{aligned}$$

$\Longrightarrow \theta_k = 1/S_t$ and $\frac{\partial \lambda_t (m_t, u_t)}{\partial m_t}$ are identified up to scale.

Proof

- Similarly, we can also identify

$$\frac{\partial \lambda_t\left(m_t, u_t\right)}{\partial u_t} = \frac{\partial G_{m|v}(m_t|v_t)/\partial u_t}{\partial G_{m|v}(m_t|v_t)/\partial k_t}\theta_k$$

$$\theta_l = \frac{\partial G_{m|v}(m_t|v_t)/\partial l_t}{\partial G_{m|v}(m_t|v_t)/\partial k_t}\,\theta_k$$

  up to scale.

- $\lambda_t(m_t, u_t)$ is identified up to scale and location from $\partial \lambda_t\left(m_t, u_t\right)/\partial m_t$ and $\partial \lambda_t\left(m_t, u_t\right)/\partial u_t$ by integration.

- $\mathbb{M}_t^{-1}\left(x_t, u_t\right)$ is identified up to scale and location as

$$\mathbb{M}_t^{-1}(x_t, u_t) = \lambda_t\left(m_t, u_t\right) - \theta_k k_t - \theta_l l_t$$

Step 3: Identification of Markup and Production Function

Differentiate $r_t = \phi_t(m_t, u_t) = \varphi_t(f(x_t) + \mathbb{M}_t^{-1}(x_t, u_t), u_t)$ by m_t

$$\underbrace{\frac{\partial \phi_t(m_t, u_t)}{\partial m_t}}_{\text{known}} = \frac{\partial \varphi_t(y_t, u_t)}{\partial y_t} \left(\theta_m + \underbrace{\frac{\partial \mathbb{M}_t^{-1}(x_t, u_t)}{\partial m_t}}_{\text{known up to scale}} \right) \quad (1)$$

From F.O.C. (De Loecker and Warzynski, 2012),

$$\frac{\partial \varphi_t(y_t, u_t)}{\partial y_t} \theta_m = \underbrace{\frac{P_t^M M_t}{R_t}}_{\text{observed}}. \quad (2)$$

$\Rightarrow$

$\dfrac{\partial \varphi_t(y_t, u_t)}{\partial y_t} = \dfrac{1}{\text{Markup}_t}$ and θ_m are identified up to scale from (1)-(2).

Step 3: Identification of Output

- Output is identified up to scale and location by

$$y_t = \theta_m m_t + \theta_l l_t + \theta_k k_t + \mathbb{M}_t^{-1}(.)$$

- So far some objects are just identified up to scale and/or location.

Step 3: Recovering the Scale Parameter

- Local returns to scale: for some $c > 0$,

$$\underbrace{\theta_m}_{\text{primitive}} + \underbrace{\theta_l}_{\text{primitive}} + \underbrace{\theta_k}_{\text{primitive}} = c.$$

$$\Leftrightarrow$$

$$b_t \cdot \underbrace{\tilde{\theta}_m}_{\text{identified}} + b_t \cdot \underbrace{\tilde{\theta}_l}_{\text{identified}} + b_t \cdot \underbrace{\tilde{\theta}_k}_{\text{identified}} = c.$$

- Constant Returns to Scale: $c = 1 \Rightarrow b_t$ is identified.

Step 3: Recovering the Location Parameter on TFP

$$\underbrace{\tilde{\mathbb{M}}_t^{-1}}_{\text{identified}} = \underbrace{b_t}_{\text{identified}} \cdot \underbrace{\mathbb{M}_t^{-1}}_{\text{primitive}} + a_t$$

- TFP following an AR(1) process:

$$E[\omega_t] = E[\mathbb{M}_t^{-1}] = 0$$

$\Rightarrow E[\tilde{\mathbb{M}}_t^{-1}] = a_t$ is identified.

Step 3: Identification of the Revenue Function $\varphi_t(.)$

- Output is fully identified as

$$y_t = \theta_m m_t + \theta_l l_t + \theta_k k_t + \mathbb{M}_t^{-1}(.)$$

- Revenue function $r_t = \varphi_t(y_t, u_t)$ is fully identified from $\{r_t, y_t, u_t\}$.

All the concerning model structures so far are fully identified!

Step 4: Identification of HSA Demand System

- Homothetic Single Aggregator (HSA) demand system (Matsuyama and Ushchev, 2017)

$$
\begin{aligned}
r_{it} &= s_t^*\left(y_{it} - q_t\left(\mathbf{y_t}, \mathbf{u_t}\right), u_t\right) + \Phi_t && \text{(primitive)} \\
&= \varphi_t\left(y_{it}, u_t\right) && \text{(reduced-form)}
\end{aligned}
$$

  - $s_t^*(\cdot)$: Primitive budget/market share
  - Φ_t: Log total income/total industry revenue
  - $q_t\left(\mathbf{y_t}, \mathbf{u_t}\right)$: Aggregate quantity index
- Extension of the CES demand system (CES: s_t^* linear in y_{it})
- Counterfactual? The identified reduced-form $\varphi_t(.)$ is not invariant to changes in the aggregate state $\left(\mathbf{y_t}, \mathbf{u_t}\right)$ due to change in $q_t\left(\mathbf{y_t}, \mathbf{u_t}\right)$!

  $\Rightarrow$ The primitive $s_t^*(\cdot)$ is required for counterfactual analysis.

Step 4: Identification of HSA Demand System

Assumption

(a) Monopolistic competition (without free entry); specifically, each firm takes aggregate quantity index $q_t\left(\mathbf{y}_t, \mathbf{u}_t\right)$ *and total industry revenue* Φ_t *as given.*

(b) The aggregate quantity index at the baseline aggregate state is normalized as $q_t(\mathbf{y}_t, \mathbf{u}_t) = 0$*.*

Proposition

(a) The primitive budget share function $s_t^*(\cdot)$ *is identified.*

(b) There exists an unique and homothetic preference $\succsim$ *that generates* $\{s_t^*(\cdot), q_t(\cdot)\}$ *such that the utility function* $\ln U_t(\mathbf{y}_t, \mathbf{u}_t)$ *is identified.*

Counterfactual Analysis: Market Power

- Short-run partial equilibrium counterfactual
  - From Monopolistic competition equilibrium (MCE) to Marginal cost pricing equilibrium (MCPE)
  - Firms adjust only m_{it} while $(k_{it}, l_{it}; u_{it}, \omega_{it})$ and income Φ_t are fixed across both states; normalize $p_{mt} = 0$.

Counterfactual Analysis: Monopolistic competition

- Monopolistic competition equilibrium (MCE):
  (y_t^m, q_t^m) solves

$$\underbrace{\exp\left(s_t^*\left(y_{it}^m - q_t^m, u_{it}\right) + \Phi_t\right) \frac{\partial s_t^*\left(y_{it}^m - q_t^m, u_{it}\right)}{\partial y_{it}}}_{\text{Marginal Revenue}}$$

$$=$$

$$\underbrace{\exp\left(\chi_{it}\left(y_{it}^m\right)\right) \frac{\partial \chi_{it}(y_{it}^m)}{\partial y_{it}}}_{\text{Marginal Cost}},$$

  subject to

$$\sum_{i=1}^{N_t} \exp\left(s_t^*\left(y_{it}^m - q_t^m, u_{it}\right)\right) = 1.$$

  where χ_{it} is the inverse production function w.r.t. m_{it}.

Counterfactual Analysis: Marginal Cost Pricing Equilibrium

- (Counterfactual) Marginal cost pricing equilibrium (MCPE): (y_t^c, q_t^c) solves

$$\underbrace{\exp\left(s_t^*\left(y_{it}^c - q_t^c, u_{it}\right) + \Phi_t\right)}_{\text{Price}}$$

$$=$$

$$\underbrace{\exp\left(\chi_{it}(y_{it}^c)\right)\frac{\partial \chi_{it}(y_{it}^c)}{\partial y_{it}}}_{\text{Marginal Cost}},$$

subject to

$$\sum_{i=1}^{N_t} \exp\left(s_t^*\left(y_{it}^c - q_t^c, u_{it}\right)\right) = 1.$$

where χ_{it} is the inverse production function w.r.t. m_{it}.

Welfare Cost of Firm's Market Power

- Compensation Variation CV_t such that

$$\ln U(\mathbf{y}_t^c(\Phi_t + CV_t), \mathbf{u}_t) = \ln U(\mathbf{y}_t^m, \mathbf{u}_t)$$

- Short run profit change

$$\Pi^c - \Pi^m = -\sum_{i=1}^{N_t} \left[\exp\left(m_{it}^c\right) - \exp\left(m_{it}^m\right)\right]$$

Estimation

Semiparametric Estimator: Overview

- Step 1: Estimate $r_{it} = \phi_t(m_{it}, u_{it})$ via B-spline and GMM Quantile Regression with Kernel Smoothing (Firpo, Galvao, Pinto, Poirier, and Sanroman (2022))

- Step 2: Profile Likelihood Estimation for $\lambda_t(\cdot)$, θ_k, and θ_ℓ (Linton, Sperlich, and Van Keilegom (2008)) with $g_{\eta_t}(\cdot)$ estimated via kernel density.

- Step 3: Recovering Material Elasticities θ_m, Markup μ_{it}, TFP ω_{it}, and quantity y_{it} with normalization from CRS and zero-mean AR(1) TFP process.

Semiparametric Estimator: Overview

- Step 4: Estimate parametric HSA Demand with composite NLS.

$$\min_{\beta_t,\delta_t,\gamma_t} \sum_i \left(r_{it} - \overbrace{\left(\Phi_t + \delta_t - \frac{1}{\beta_t} \ln\left(\frac{\exp(-\beta_t \hat{y}_{it} + \gamma_t) + \epsilon_{it}}{1+\epsilon_{it}}\right)\right)}^{\hat{r_{it}}} \right)^2$$
$$+ \sum_i \left(\hat{u}_{it} - \text{quantile}(\epsilon_{it})\right)^2$$

subject to the F.O.C.: $\epsilon_{it} = \frac{\hat{\mu}_{it}-1}{\exp(\beta_t \hat{y}_{it} - \gamma_t)}$ and the market share condition: $1 = \sum_i \exp\left(\delta_t - \frac{1}{\beta_t} \ln\left(\frac{\exp(-\beta_t \hat{y}_{it} + \gamma_t) + \epsilon_{it}}{1+\epsilon_{it}}\right)\right)$

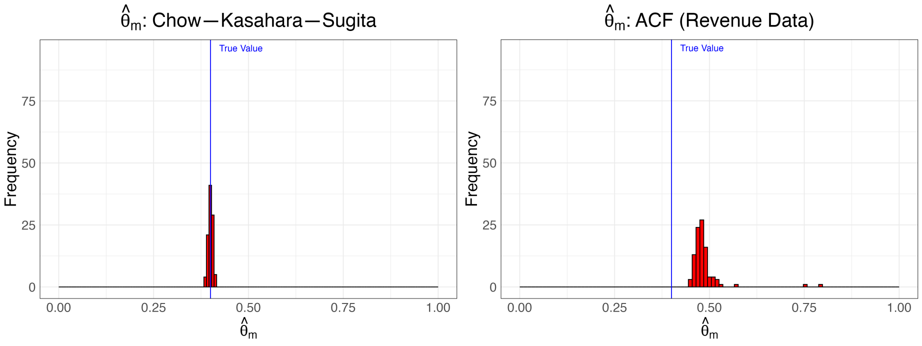
Simulation

- HSA Demand, Cobb-Dogulas Production, AR1 TFP, MA1 Demand Shock:

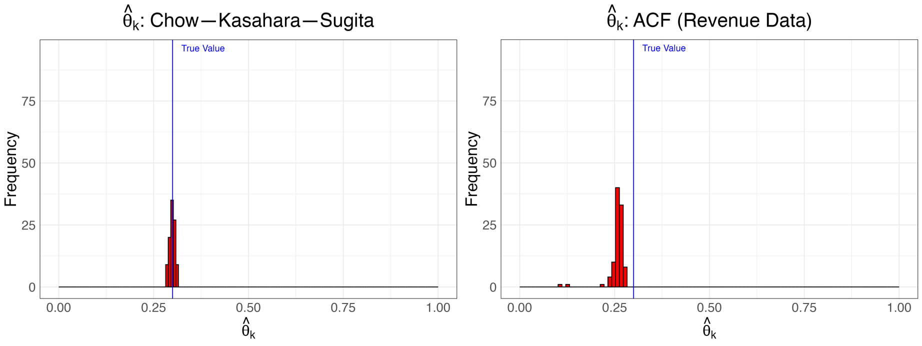
$$
\begin{aligned}
y_{it} &= 0.4m_{it} + 0.3k_{it} + 0.3\ell_{it} + \omega_{it}, \\
\omega_{it} &= 0.8\omega_{it-1} + \eta_{it},\ \ \eta_{it} \sim N(0, (0.05)^2), \\
\epsilon_{it} &= 0.8\xi_{it-1} + \xi_{it}, \quad \xi_{it} \perp\!\!\!\perp \xi_{it-1} \sim U[0, 0.3]. \\
s_t^* &:= -6.5 - \frac{1}{0.21} \ln\left(\frac{\exp(-0.21\,(y_{it} - q_t(\mathbf{y}_t, \boldsymbol{\epsilon}_t))) + \epsilon_{it}}{1 + \epsilon_{it}}\right) \\
\Phi_t &= 20, p_{mt} = 1.
\end{aligned}
$$

- 100 samples of 600 firms over 4 years
- We compare our method with Ackerberg, Caves, and Frazer (2015)'s method applied with revenue as output.

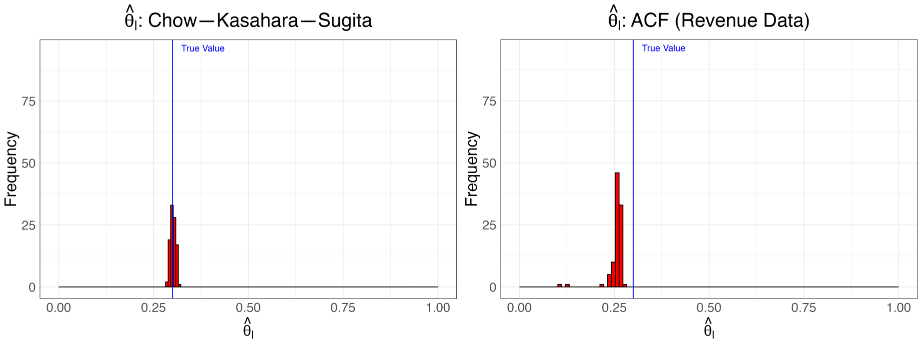
Simulation: Material Elasticity



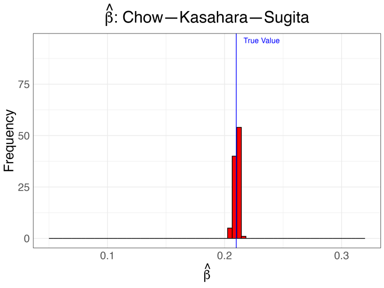
Simulation: Capital Elasticity

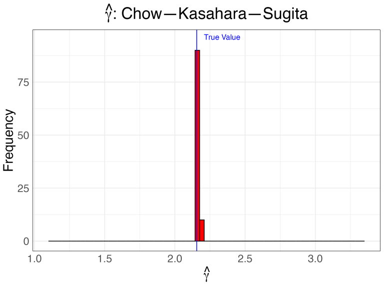


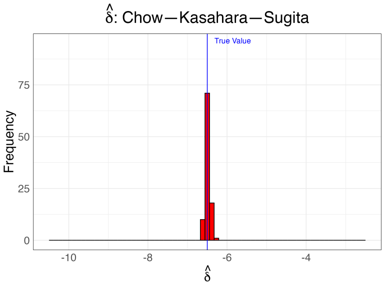
Simulation: Labor Elasticity



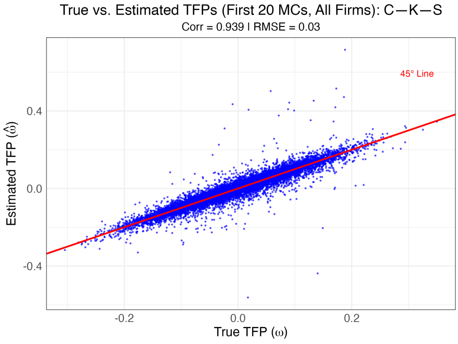
Simulation: HSA Demand System Parameters

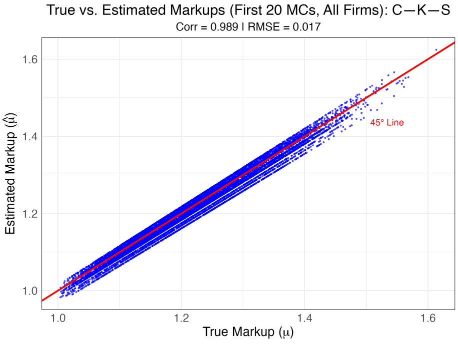






Simulation: TFP and Markup





Chilean Manufacturing Plant Data

- Three largest industries in 1996
  - 31 (Food, Beverage, and Tobacco)
  - 32 (Textiles, Apparel, and Leather Products)
  - 38 (Metal Products, Electric/Non-electric Machinery, Transport Equipment, and Professional Equipment)
  - We exclude firms with $K_{it} \leq 0$, $P_t^M M_{it}/R_{it} < 0$,$P_t^M M_{it}/R_{it} > 1$, the bottom and top 2 percentiles. In step 4 we also exclude firms with $\mu_{it} < 1$.
  - Computation time with 100 non-parametric bootstrap iterations: 6-13 min for each industry.

Chilean Manufacturing Plant Data: Step 1-3

Industry	n	$\hat{\theta}_m$	$\hat{\theta}_k$	$\hat{\theta}_l$	Avg. Markup
31	736	0.848	0.013	0.138	1.386
		(0.031)	(0.010)	(0.031)	(0.052)
32	463	0.756	0.079	0.164	1.503
		(0.049)	(0.032)	(0.045)	(0.099)
38	391	0.672	0.058	0.270	1.628
		(0.067)	(0.037)	(0.062)	(0.167)

Step 1, Step 2, and Step 3 (Industries 31, 32, and 38 in 1996). Standard errors in parentheses with 100 non-parametric bootstrap iterations.

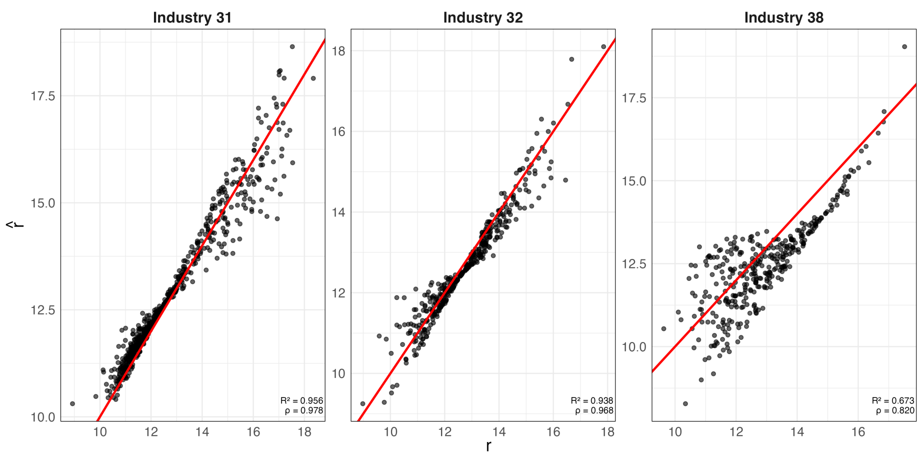
Chilean Manufacturing Plant Data: Step 4

$H_0 : \beta = 0$ implied by the CES demand is rejected.

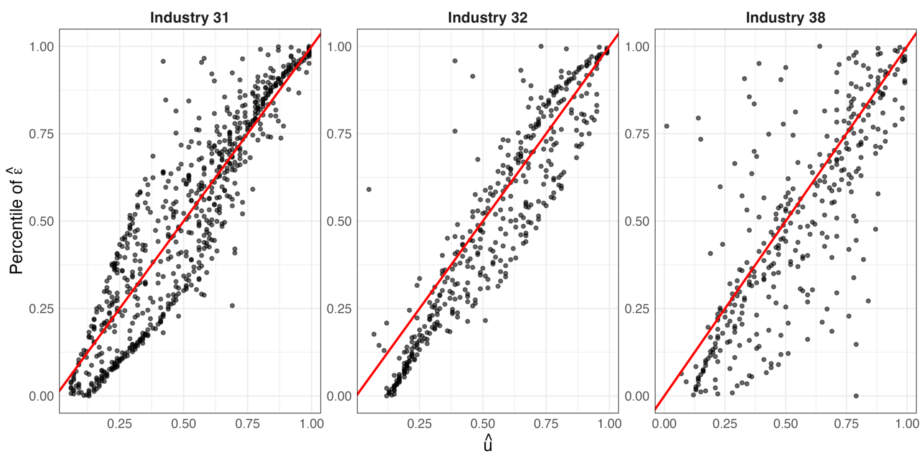
Industry	n	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$
31	698	0.154	1.770	-8.009
		(0.011)	(0.105)	(0.159)
32	409	0.085	0.951	-6.899
		(0.012)	(0.141)	(0.572)
38	347	0.103	1.367	-5.443
		(0.044)	(0.443)	(1.352)

Step 4 (Industries 31, 32, and 38 in 1996). Standard errors in parentheses with 100 non-parametric bootstrap iterations.

Goodness of Fit: Revenue



Goodness of Fit: Demand Shock (Step 1 vs Step 4)



Welfare Cost of Firm's Market Power

- % of industry revenue $\exp(\Phi_t)$

Industry	CV(%)	$\Pi^m - \Pi^c$(%)	Overall Welfare (%)
31	-14.1	11.0	-3.08
	(2.71)	(-2.39)	(0.58)
32	-16.1	9.89	-6.20
	(5.97)	(5.04)	(1.35)
38	-9.85	4.07	-5.78
	(6.75)	(4.55)	(2.92)

Standard errors in parentheses with 100 non-parametric bootstrap iterations.

Conclusion

- This paper provides identification of production function, TFP, and markup as well as the HSA demand system from revenue data.
- The identified HSA demand system enables counterfactual analysis of market outcomes and welfare.
- Our semi-parametric estimator performs well.
- Applying our method to Chilean manufacturing plant data, we reject the CES demand in favor of HSA.
- Counterfactual analysis indicates that market power reduces welfare by about 3-6% of industry revenue.

References

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

Chernozhukov, V. and Hansen, C. (2005), "An IV model of quantile treatment effects," *Econometrica*, 73, 245–261.

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

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

Firpo, S., Galvao, A. F., Pinto, C., Poirier, A., and Sanroman, G. (2022), "GMM quantile regression," *Journal of Econometrics*, 230, 432–452.

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

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, N/A.

— (2020), "Constant Pass-Through," CEPR Discussion Paper 15475, C.E.P.R. Discussion Papers.

Semiparametric Estimator

- Specify $\phi_t(m, u)$ with B-spline and use

$$E[u - \mathbb{I}(r_{it} - \phi_t(m_{it}, u) \leq 0) | m_{it-2}] = 0$$

with shape restriction

$$\frac{\partial \phi_t(m, u)}{\partial m} > 0, \quad \frac{\partial \phi_t(m, u)}{\partial u} > 0,$$

where

$$\phi_t(m, u) = \sum_{j_1=1}^{K_1} \sum_{j_2=1}^{K_2} B_{j_1}^m(m) B_{j_2}^u(u) \beta_{j_1 j_2} = \boldsymbol{B}(m, u)^\top \boldsymbol{\beta}.$$

Semiparametric Estimator: Step 1

- GMM quantile regression by Firpo, Galvao, Pinto, Poirier, and Sanroman (2022):
- Kernel smoothing
- GMM: for$u \in [0, 1]$,

$$E\left[\left(u - \mathcal{K}\left(\frac{\boldsymbol{B}(m_{it}, u)^{\top}\boldsymbol{\beta} - r_{it}}{b}\right)\right) B^{m}(m_{it-2})\right] = 0$$

$\Rightarrow$

$$\hat{\phi}_t\,(m, u) = \boldsymbol{B}(m, u)^{\top}\hat{\boldsymbol{\beta}}.$$

Inverting $r_{it} = \hat{\phi}_t(m_{it}, u_{it})$,

$$\hat{u}_{it} = \hat{\phi}_t^{-1}(m_{it}, r_{it}).$$

Semi-parametric Estimator: Step 2

$$\omega_t = \rho\omega_{t-1} + \eta_t$$

$$\underbrace{\mathbb{M}_t^{-1}(x_t, u_t)}_{\omega_t} = \lambda_t(m_t, u_t) - \theta_k k_t - \theta_\ell l_t$$

B-spline

$$\lambda_t(m_t, u_t; \alpha^m) = \sum_{j_1=1}^{K_1}\sum_{j_2=1}^{K_2} B_{j_1}^m(m) B_{j_2}^u(u)\alpha_{j_1 j_2}^m$$

$\Rightarrow$

$$\eta_{it}(\alpha^m, \theta) = \Delta_\rho \lambda(m_{it}, \hat{u}_{it}; \alpha^m) - \theta_\ell \Delta_\rho \ell_{it} - \theta_k \Delta_\rho k_{it}$$

Semi-parametric Estimator: Step 2

Linton, Sperlich, and Van Keilegom (2008)

$$(\hat{\alpha}^m, \hat{\theta}) \in \arg\max_{\alpha^m, \theta} \sum_{i=1}^{N} \ln g_{m|v}(m_{it}|v_{it}; \alpha^m, \theta)$$

where

$$\ln g_{m|v}(m_{it}|v_{it}; \alpha^m, \theta) := \ln g_{\eta_t}(\eta_{it}(\alpha^m, \theta)) + \ln \frac{\partial \lambda_t\left(m_{it}, \hat{u}_{it}; \alpha^m\right)}{\partial m_t}.$$

$g_{\eta_t}(\eta_{it})$: kernel-density

Semi-Parametric Estimator: Step 3

- Material elasticity:

$$\widehat{\theta}_m = \text{med}\left\{ \left(\frac{\partial \widehat{\phi}_t(m_{it}, u_{it})}{\partial m_{it}} - \frac{P^M_{it} M_{it}}{R_{it}} \right)^{-1} \frac{\partial \hat{\lambda}(m_{it}, u_{it})}{\partial m_{it}} \frac{P^M_{it} M_{it}}{R_{it}} \right\}$$

- TFP is estimated as

$$\hat{\omega}_{it} = \lambda_t(m_t, \hat{u}_t; \hat{\alpha}^m) - \hat{\theta}_k k_t - \hat{\theta}_\ell \ell_t$$

- Normalization
  - Scale: $\text{med}(\widehat{\frac{\partial f_t(x_{it})}{\partial m_{it}}}) + \hat{\theta}_k + \hat{\theta}_\ell = 1.$
  - Location: $\text{med}(\hat{\omega}_{it}) = 0.$

Semi-Parametric Estimator: Step 4

- So far the demand function is nonparametric.
- We estimate a parametric demand function for stable estimation.
- CoPaTh-HSA demand system by Matsuyama and Ushchev (2020)

$$\begin{aligned} & s_t^*(y_{it} - q_t(\mathbf{y}_t, \epsilon_t), \epsilon_{it}) = r_{it} - \Phi_t \\ := & \delta_t - \frac{1}{\beta_t} \ln \left(\frac{\exp(-\beta_t \left(y_{it} - q_t(\mathbf{y}_t, \boldsymbol{\epsilon}_t)\right) + \kappa_t) + \epsilon_{it}}{1 + \epsilon_{it}} \right) \end{aligned}$$

- Matsuyama and Ushchev (2020) allows (δ_t, κ_t) also to be firm-specific.
- We estimate its reduced form

$$\varphi_t(y_{it}, \epsilon_{it}) = \Phi_t + \delta_t - \frac{1}{\beta_t} \ln \left(\frac{\exp(-\beta_t y_{it} + \gamma_t) + \epsilon_{it}}{1 + \epsilon_{it}} \right)$$

where $\gamma_t \equiv \beta_t q_t(\mathbf{y}_t, \boldsymbol{\epsilon}_t) + \kappa_t$

Semi-Parametric Estimator: Step 4

- The CoPaTh-HSA demand predicts the markup as:

$$\frac{\mathsf{P}_{it}}{\mathsf{MC}_{it}} = \mu_{it} = 1 + \epsilon_{it} \exp(\beta_t y_{it} - \gamma_t).$$

- $\beta_t \to 0$: constant markup and complete pass-through, including the CES demand system

Semi-Parametric Estimator: Step 4

$$(\hat{\beta}_t, \hat{\delta}_t, \hat{\gamma}_t)'$$

$$\in \arg\min_{\beta_t, \delta_t, \gamma_t} \sum_i \left(r_{it} - \left(\Phi + \delta_t - \frac{1}{\beta_t} \ln \left(\frac{\exp(-\beta_t \hat{y}_{it} + \gamma_t) + \epsilon_{it}}{1 + \epsilon_{it}} \right) \right) \right)^2$$

$$+ \sum_i \left(\hat{u}_{it} - (\text{quantile}(\epsilon_{it})) \right)^2,$$

s.t.

$$\epsilon_{it} = \frac{\hat{\mu}_{it} - 1}{\exp(\beta_t \hat{y}_{it} - \gamma_t)},$$

$$1 = \sum_i \exp \left(\delta_t - \frac{1}{\beta_t} \ln \left(\frac{\exp(-\beta_t \hat{y}_{it} + \gamma_t) + \epsilon_{it}}{1 + \epsilon_{it}} \right) \right).$$