Weak Identification Robust Methods for Production Function Estimation

Jorge de la Cal Medina

University of Amsterdam, Tinbergen Institute

August 27, 2025

Introduction

- **Production function (PF) estimation:**

$$y = x'\beta + \omega + \varepsilon$$

  Firms choose inputs x after observing state ω (Marschak and Andrews, 1944).
- **Control function approach (CFA):** Proxying ω with a control function allows to estimate β with GMM (Ackerberg et al., 2015).
- **Weak identification:** When proxies have low explanatory power (measurement error) Jacobian near-singular, leading to biased, non-normal estimates and invalid standard inference.

This Paper

Contributions

1. Identification with noisy proxy feasible (under restrictions)
2. Weak proxies (small signal/noise) can cause biased estimates with non-normal distributions.
3. Adapt weak identification-robust methods: bootstrap tests for proxy strength (Angelini et al., 2024) and robust confidence sets (Stock and Wright, 2000).

Related Literature

Production-function identification

- Olley and Pakes (1996) (OP), Levinsohn and Petrin (2003) (LP), Ackerberg et al. (2015) (ACF) discuss identification using (strong) proxies; Gandhi et al. (2020) (GNR) discuss non-identification (in nonparametric models)
- *We consider sources for weak identification*

Weak identification in structural models

- DSGE (Canova & Sala 2009), NK Phillips curve (Mavroeidis, Plagborg-Møller & Stock 2013), BLP demand (Armstrong 2016)
- *We extend this list with structural production-function estimation*

Empirical work

- Empirical applications using control function estimator
  - Trade: Pavcnik (2002); Topalova (2010); Fernandes (2007)
  - Market power: De Loecker et al. (2020); Autor et al. (2020)
  - Finance: Gopinath et al. (2017); Gourinchas et al. (2020)
- *We give guidance on model specification and add to toolbox*

Structure of Presentation

- Identification Analysis
  - Present identification strategy of CFA
  - Weak proxies
- Monte Carlo simulation
- Empirical Analysis
- Conclusion

Outline

The Control Function Approach

- General PF specification for firm i at period t:

$$y_{it} = x_{it}'\beta + \omega_{it} + \varepsilon_{it}, \quad \mathbb{E}[\varepsilon_{it}|\mathcal{I}_{it}] = 0 \tag{1}$$

- Markov process for state variable

$$\omega_{it} = \rho\,\omega_{it-1} + \xi_{it}, \quad \mathbb{E}[\xi_{it} \mid \mathcal{I}_{it-1}] = 0 \tag{2}$$

- Demand function h of intermediate inputs m

$$m_{it} = h(\omega_{it}, x_{it}) + \nu_{it} \tag{3}$$

  - strictly monotonic in ω_{it}
  - allow for unobservable ν_{it} (unlike OP, LP, ACF)
- Control function

$$\begin{aligned} \omega_{it} &= h^{-1}(m_{it} - \nu_{it}, x_{it}) \\ &= \Psi(m_{it}, x_{it}) + \eta_{it}, \quad \mathbb{E}[\eta_{it}|x_{it}, m_{it}] = 0 \end{aligned}$$

  - $\Psi(m, x) = E[\omega_{it} \mid m_{it} = m,\ x_{it} = x]$

Estimation

- Estimation equation

$$y_{it} = x'_{it}\beta + \rho\left(y_{it-1} - x'_{it-1}\beta\right) - \rho\varepsilon_{it-1} + \xi_{it} + \varepsilon_{it} \tag{4}$$

- GMM Moment conditions

$$\mathbb{E}\Big[z_{it}\left(y_{it} - x'_{it}\beta - \rho\left(y_{it-1} - x'_{it-1}\beta\right)\right)\Big] = 0$$

- Instrument vector z_{it} contains m_{it-1} to instrument for y_{it-1} (through control function)
- Straightforward to get CUE (Hansen et al., 1996)

Nonidentification & Weak Proxies

Identification

- requires full-rank Jacobian

$$\mathcal{J} := \Big(\mathbb{E}\big[z_{it}(x_{it}' - \rho_0 x_{it-1}')\big] \vdots \mathbb{E}\big[z_{it}(y_{it-1} - x_{it-1}'\beta_0)\big] \Big)$$

Nonidentification

- Assume linear control function $\Psi(m, x) = \gamma_m m + x'\gamma_x$
- $\mathcal{J}$ is rank-deficient, if
  - x_{it} and x_{it-1} collinear
  - $\gamma_m \to 0$

Weak proxies

- Small γ_m (weak proxies) lead to near rank-deficiency $\Rightarrow$ Weak identification
- Example: classical measurement error in proxies attenuates γ_m

Outline

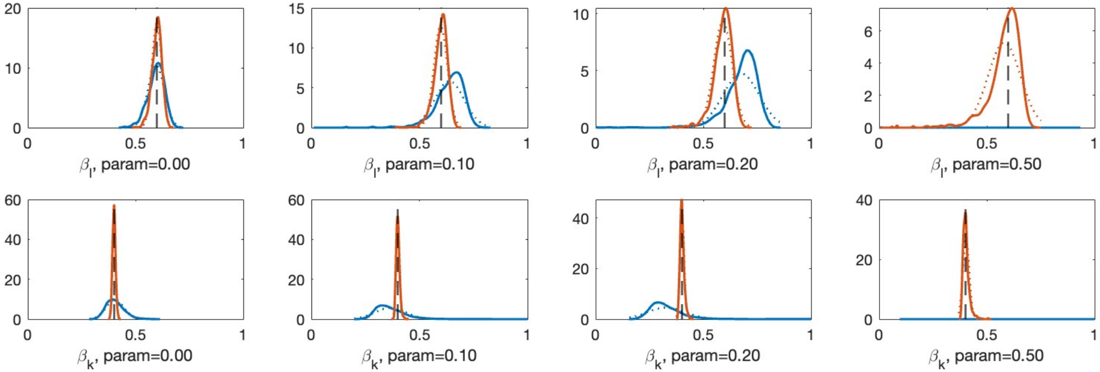
Monte Carlo

- Dynamic model of firm investment and production (ACF)
  - Two inputs capital k and labor l
  - Intermediate inputs, m, subject to measurement error
  - Tweak: reduce optimization error in l from 10% to 5%
- Estimators
  - ACF estimator
  - CUE

Table 1: Bias of simulated estimates (with standard errors). Measurement error in terms of additional % variance of m.

	ACF		CUE	
Meas.	β_l	β_k	β_l	β_k
0.0	−0.005 (0.038)	+0.005 (0.041)	−0.001 (0.023)	0.000 (0.007)
0.1	−0.500 (0.300)	−0.301 (0.299)	+0.038 (0.068)	−0.035 (0.070)
0.2	−0.445 (0.985)	−0.493 (0.983)	−0.461 (0.346)	−0.300 (0.300)
0.5	—	—	−0.532 (1.014)	−0.436 (0.968)

Monte Carlo Estimates of PF Parameters



1000 replications. ACF in blue. CUE in red. Empirical densities: solid. Normal pdf: dotted. Panels ordered left to right for larger measurement-error magnitudes.

- Measurement error ↑: stronger bias and non-normality

Monte Carlo: Inference on PF Parameters

Table 2: Rejection frequencies at 5 % nominal level

Meas.	Joint test (non-robust)[1]	Joint test (robust)[2]	BS joint normality[3]
0.0	0.064	0.060	0.057
0.1	0.059	0.048	0.068
0.2	0.075	0.062	0.167
0.5	0.128	0.050	0.431

Based on 1000 replications. [1]Wald. [2]Stock and Wright (2000). [3]Doornik and Hansen (2008).

- Measurement error $\uparrow$:
  - Non-robust test: size $> 5\%$
  - Robust test: size $\approx 5\%$
  - BS normality test: rejections $\uparrow$

Outline

Data & empirical setting (Chile & U.S.)

- Replicate Raval (2023) and apply CUE estimator and identification-robust methods
- Chile (ENIA, plant-level)
  - Fabricated metal products (ISIC 381)
  - years 1979–1996
  - plant-year obs. $\approx$ 4,000
  - Materials: plant-reported intermediate consumption (*high precision*)
- U.S. (Compustat, firm-level)
  - Manufacturing (mostly durables incl. metal products; NAICS 33)
  - years 1970–2010
  - firm-year obs. $\approx$ 8,000
  - Materials: proxy via COGS $-$ XLR; mixes labor/overhead (*lower precision*)
- Model & estimators.
  - Cobb–Douglas in k, l, m
  - Control function: third-order polynomial in k, l, m
  - Estimation: ACF baseline + CUE

Elasticities & bootstrap normality — Chile (ISIC 381)

Table 3: Production function estimates (Chile, ISIC 381)

	Estimates		95% CI		BS Normality[3]
Parameter	ACF	CUE	Nonrobust[1]	Robust[2]	p
β_k	0.064	0.047	[0.014, 0.080]	[0.003, 0.087]	0.0
β_l	0.122	0.053	[−0.090, 0.195]	[−0.083, 0.185]	0.0
β_m	0.875	0.956	[0.834, 1.078]	[0.848, 1.072]	0.0
Returns to scale	1.060	1.060	—		—

[1] Wald. [2] Subset S Stock and Wright (2000). [3] Shapiro–Wilk (not reported).

- ACF and CUE estimates in ballpark
- Non-robust and robust CI similar and rel. tight
- Normality is rejected still

Elasticities & bootstrap normality — U.S. (NAICS 33)

Table 4: Production function estimates (U.S., NAICS 33)

	Estimates		95% CI		BS Normality[3]
Parameter	ACF	CUE	Nonrobust[1]	Robust[2]	p
β_k	0.422	0.187	$[-0.660, 1.034]$	$[-0.153, 0.507]$	0.0
β_l	0.411	0.333	$[-1.135, 1.801]$	$(-\infty, 0.787]$	0.0
β_m	0.237	0.219	$[0.106, 0.332]$	$[0.052, 0.332]$	0.0
Returns to scale	1.070	0.739	—		—

[1] Wald. [2] Subset S Stock and Wright (2000). [3] Shapiro–Wilk.

- CUE aligns less with ACF
- robust CIs are large, unbounded for β_l (weak ID)
- Normality again rejected

Conclusion

Findings

- Proxies with poor explanatory power (weak proxies) are source of weak identification
- Monte Carlo shows potential large imprecision and non-normality in case of weak proxies
- Robust methods ensure valid inference

Outlook

- Pretest that does not rely on BS (Andrews, 2018)
- More powerful robust inference (Kleibergen, 2005; Moreira, 2003)
- Other empirical applications

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Andrews, I. (2018). Valid two-step identification-robust confidence sets for gmm. *The Review of Economics and Statistics*, 100(2):337–348.

Angelini, G., Cavaliere, G., and Fanelli, L. (2024). An identification and testing strategy for proxy-svars with weak proxies. *Journal of Econometrics*, 238(2):105604.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The Fall of the Labor Share and the Rise of Superstar Firms. *The Quarterly Journal of Economics*, 135(2):645–709.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2).

Doornik, J. A. and Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, 70(5):927–939.

Gandhi, A., Navarro, S., and Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8). Electronically published.

Gopinath, G., Kalemli-Özcan, , Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe. *The Quarterly Journal of Economics*, 132(4):1915–1967.

Hansen, L. P., Heaton, J., and Yaron, A. (1996). Finite-sample properties of some alternative