Weak Identification Robust Methods for Production Function Estimation

Jorge de la Cal Medina

University of Amsterdam, Tinbergen Institute

January 10, 2026

Introduction

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

  $$y = \beta_k k + \beta_\ell \ell + \omega + \varepsilon.$$

  Firms choose inputs (k, ℓ) after observing productivity ω (Marschak and Andrews, 1944)

- **Control function approach (CFA):** Use a proxy input m to control for ω and estimate (β_k, β_ℓ) with GMM (Ackerberg et al., 2015)

- **Weak identification:** *Multicollinearity* between k and ℓ and *noisy proxies* $m \rightarrow$ near-singular Jacobian, biased / non-normal estimates and invalid standard inference

This Paper

Contributions

1. Discuss potential identification issues structural PF estimation and show consequences of weak identification (Stock and Wright, 2000)
2. Illustrate weak identification issues using Monte Carlo of dynamic firm optimization model
3. Adapt weak-ID-robust methods to CFA: CUE (Hansen et al., 1996a), bootstrap pre-tests for proxy strength (Angelini et al., 2024) and AR–GMM / constraint (subset) inference (Stock and Wright, 2000; Kleibergen and Mavroeidis, 2009)

Related Literature

Production-function identification

- Olley and Pakes (1996) (OP), Levinsohn and Petrin (2003) (LP), Ackerberg et al. (2015) (ACF): identification with (strong) proxies
- Gandhi et al. (2020) (GNR): non-identification in nonparametric CFA
- *We discuss weak identification in semi-/parametric CFA and weak proxies*

Weak identification in structural models

- DSGE (Canova & Sala 2009), NK Phillips curve (Mavroeidis, Plagborg-Møller & Stock 2013), BLP demand (Armstrong 2016)
- *We add structural PF estimation to this list*

Empirical applications

- CF estimators widely used in:
  - 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)
- *Applying methods to empirical data and adding to the empiricist's toolbox*

Structure of Presentation

Identification Analysis

Methods

Monte Carlo

Empirical application

Conclusion

Outline

The Control Function Approach

- PF for firm $i \in \{1, \ldots, n\}$ at period $t \in \{1, \ldots, T\}$:

$$y_{it} = \beta_k k_{it} + \beta_\ell \ell_{it} + \omega_{it} + \varepsilon_{it} \tag{1}$$

- k is predetermined, ℓ is flexible
- Productivity evolves as

$$\omega_{it} = \rho \, \omega_{it-1} + \xi_{it} \tag{2}$$

- Additional variable m_{it}, e.g. intermediate inputs to *proxy* ω_{it}
- Control function representation:

$$\omega_{it} = \delta_k \, k_{it} + \delta_m \, m_{it} + \eta_{it} \tag{3}$$

- Timing Assumptions:
  - Inputs chosen by firm i in period t: $x_{it} = (k_{it+1}, \ell_{it}, m_{it})'$
  - Sequential exogeneity conditions:

$$\mathbb{E}[\varepsilon_{it} \mid x_{i\tau}] = 0, \quad \mathbb{E}[\xi_{it} \mid x_{i\tau-1}] = 0, \quad \mathbb{E}[\eta_{it} \mid x_{i\tau}] = 0, \quad \tau \leq t \tag{4}$$

Identification and Estimation

- Analogy with IV: m_{it-1} plays the role of an instrument for y_{t-1}
- Combining (1) and (2) with the lagged production function

$$y_{it} = \beta_k k_{it} + \beta_\ell \ell_{it} + \rho\big(y_{it-1} - \beta_k k_{it-1} - \beta_\ell \ell_{it-1}\big) + \underbrace{\xi_{it} + \varepsilon_{it} - \rho\,\varepsilon_{it-1}}_{u_{it}} \qquad (5)$$

- Combining (1) and (3) one period lagged gives

$$y_{it-1} = (\beta_k + \delta_k)k_{it-1} + \beta_\ell \ell_{it-1} + \delta_m m_{it-1} + \underbrace{\varepsilon_{it-1} + \eta_{it-1}}_{v_{it}}$$

- Valid instruments

$$z_{it} = (k_{it}, \ell_{it-1}, m_{it-1}, \dots)'$$

- Moment conditions for $\theta = (\beta_k, \beta_\ell, \rho)'$,

$$\mathbb{E}[f_{it}(\theta_0)] = \mathbb{E}\Big[z_{it}\Big(y_{it} - \beta_k k_{it} - \beta_\ell \ell_{it} - \rho\big(y_{it-1} - \beta_k k_{it-1} - \beta_\ell \ell_{it-1}\big)\Big)\Big] = 0$$

- Straightforward to implement CUE (Hansen et al., 1996b), important for ID-robust methods

Nonidentification

Rank conditions

- Identification requires full-rank Jacobian

$$\mathbb{E}[\partial f_{it}(\theta)/\partial\theta'] =$$
$$\left(\mathbb{E}\left[z_{it-1}(k_{it} - \rho\, k_{it-1})\right] \vdots\, \mathbb{E}\left[z_{it-1}(\ell_{it} - \rho\, \ell_{it-1})\right] \vdots\, \mathbb{E}\left[z_{it-1}\left(y_{it-1} - \beta_k k_{it-1} - \beta_\ell \ell_{it-1}\right)\right]\right)$$

Sources of nonidentification:

- *Linear labor policy.* If labor is linear in k and ω,

$$\ell_{it} = a_{\ell k} k_{it} + a_{\ell\omega}\omega_{it}$$

- *Static capital and uninformative proxy.* Capital is fixed over time, $k_t = k_{t-1}$, and the proxy is uninformative, that is, $\delta_m = 0$ in (3)

Proxy strength and measurement error

- Subsume model to linear IV
- Linear policy for materials with classical measurement error ν_{it}:

$$m_{it} = \gamma_\omega \omega_{it} + \gamma_k k_{it} + \nu_{it},$$

- Under joint normality, concentration parameter at cross-section t:

$$\mu_t^2 = \frac{n\, \gamma_\omega^2\, \sigma_{\omega|k}^2}{\sigma_\varepsilon^2 \gamma_\omega^2 + \left(\frac{\sigma_\varepsilon^2 + \sigma_{\omega|k}^2}{\sigma_{\omega|k}^2}\right) \sigma_\nu^2}$$

- μ_t^2 decreases monotonically in σ_ν^2: more proxy noise $\to$ weaker proxy $\to$ weaker ID
- Analogue of classical measurement error weakening instruments in linear IV

Strong ID: Standard Asymptotics

Standard Setup (Hansen, 1982)

- Parameter vector $\theta = (\beta_k, \beta_\ell, \rho)^\top$
- Sample moments

$$\bar{f}_n(\theta) = \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=2}^{T} f_{it}(\theta)$$

- Estimator and objective

$$\hat{\theta}_n = \arg\min_{\theta \in \Theta} Q_n(\theta), \quad Q_n(\theta) = n\, \bar{f}_n(\theta)^\top W_n \bar{f}_n(\theta)$$

- Under standard assumptions, if Jacobian has full column rank:
  - Consistency: $\hat{\theta}_n \xrightarrow{p} \theta_0$
  - Asymptotic normality: $\sqrt{n}\,(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, V_\theta)$

Weak ID: Nonstandard Asymptotics

Partial weak-ID setup (Stock and Wright, 2000)

- Parameter split: $\theta = (\beta', \rho)'$, with $\beta = (\beta_k, \beta_\ell)'$
- (Partial local-to-zero Jacobian)

$$\mathbb{E}\left[\partial \bar{f}_n(\beta, \rho)/\partial \beta'\right] = \frac{1}{\sqrt{n}} C(\rho), \quad \mathbb{E}\left[\partial \bar{f}_n(\beta_0, \rho)/\partial \rho\right] = r, \quad [C(\rho)\ r] \text{ full rank}$$

- (Weak convergence): $\sqrt{n}\left\{\bar{f}_n(\theta) - \mathbb{E}\left[\bar{f}_n(\theta)\right]\right\} \Rightarrow \psi_0^f$,[1]
- (Pos. definiteness) $W_n \xrightarrow{p} W_0$

[1] $\Rightarrow$ denotes weak convergence of random functions on Θ with respect to the supremum norm (see van der Vaart, 2000, Sec. 18)

Weak ID: Nonstandard Asymptotics

Proposition 1: $\hat{\rho}$ is $\sqrt{n}$–consistent for ρ_0

Proposition 2: Limit representation

$$Q_n(\beta, \hat{\rho}) \Rightarrow \left(\psi_0^f(\beta, \rho_0) + C(\rho_0)(\beta - \beta_0) + \lambda r\right)' W_0 \left(\psi_0^f(\beta, \rho_0) + C(\rho_0)(\beta - \beta_0) + \lambda r\right)$$

Implications

- $Q_n(\beta, \hat{\rho}) = O_p(1)$ for any $\beta \neq \beta_0$
- $\hat{\beta}$ is only $O_p(1)$, inconsistent and asymptotically non-normal
- $\hat{\rho}$ is consistent but has a non-normal limit (because it depends on the random limit of $\hat{\beta}$)

Outline

Bootstrap pre-test for proxy relevance (Angelini et al., 2024)

- H_0: Proxies (and instruments) are strong
- Let full-sample CUE be $\hat{\theta}_n$ with variance V_θ, and BS estimate $\hat{\theta}_n^*$
- Statistic:

$$\Gamma_n^* = \sqrt{n}\,\hat{V}_\theta^{-1/2}(\hat{\theta}_n^* - \hat{\theta}_n)$$

- Asymptotically Gaussian *only* under strong proxies:

$$\Gamma_n^* \xrightarrow{d}_{H_0} \mathcal{N}(0, I)$$

- Testing procedure
  - Draw cluster bootstrap samples at firm level and compute $\Gamma_{n,b}^*$ for each b
  - Test (multivariate) normality of $\{\Gamma_{n,b}^*\}_{b=1}^B$ (e.g. Doornik–Hansen)
  - If reject normality: treat proxies/instruments as weak and move to ID-robust inference
- Key features
  - Post-test inference unaffected (pretest asymptotically independent of estimator)
  - Simple: only needs $\hat{\theta}_n$, $\hat{V}_\theta$ and bootstrap re-estimates

AR-GMM (Stock and Wright, 2000)

- $H_0 : \theta_0 = \theta^*$
- AR–GMM (S-statistic):

$$S(\theta^*) = n\bar{f}_n(\theta^*)'\hat{V}_{ff}(\theta^*)^{-1}\bar{f}_n(\theta^*)$$

- Asymptotically chi-squared:

$$S(\theta^*) \xrightarrow{d}_{H_0} \chi^2_{d_f},$$

  where d_f is the number of moment conditions
- Confidence set by inversion:

$$CS(\alpha) = \{\theta^* : S(\theta^*) \leq \chi^2_{d_f}(1-\alpha)\} \tag{6}$$

- Key features
  - Confidence sets remain valid even under weak identification
  - Easily implementable once we have CUE
  - More powerful robust tests exist, e.g. Kleibergen (2002), Moreira (2003)

Constraint Inference and Subset Inference

- Often we care about *linear combinations* of θ:

$$\psi := R\theta \in \mathbb{R}^{d_\psi}, \qquad \theta = (\beta_k, \beta_\ell, \rho)'$$

  e.g. returns to scale, or a particular elasticity
- Constraint null:

$$H_0 : \ \psi = \psi^*$$

- Define restricted CUE:

$$\hat{\theta}(\psi^*) := \arg \min_{\theta:\, R\theta = \psi^*} S(\theta),$$

- Test statistic:

$$S\left(\psi^*\right) := S\left(\hat{\theta}(\psi^*)\right)$$

- Under H_0, limiting distribution is *stochastically dominated* by $\chi^2_{d_f - d_\theta - d_\psi}$ (Kleibergen and Mavroeidis, 2009)

Outline

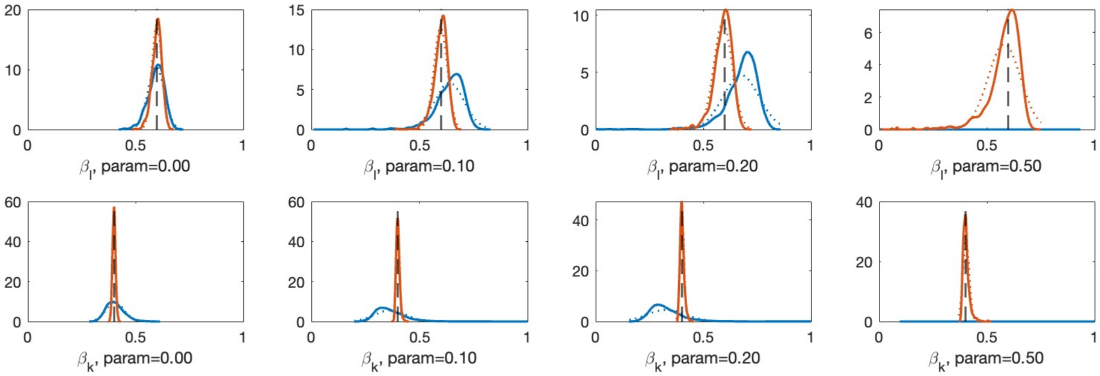
Monte Carlo

- Dynamic model of firm investment and production (ACF-type)
  - Two inputs capital k and labor ℓ
  - Intermediate inputs m subject to classical 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]AR-GMM. [3]Doornik and Hansen (2008).

- Measurement error $\uparrow$:
  - Wald over-rejects (non-robust to weak proxies)
  - AR-GMM keeps size
  - BS normality rejections $\uparrow$ (signals weak ID)

Outline

Identification Analysis

Methods

Monte Carlo

Empirical application

Conclusion

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

- Replicate Raval (2023) and apply CUE + ID-robust methods
- Chile (ENIA, plant-level)
  - Manufacture of food products (ISIC 311)
  - years 1979–1996
  - plant-year obs. $\approx$ 16,000
  - Materials: plant-reported intermediate consumption (*high precision*)
- U.S. (Compustat, firm-level)
  - Food Manufacturing (NAICS 31)
  - years 1970–2005
  - firm-year obs. $\approx$ 1,000
  - 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

ENIA, Chile (ISIC 311)

Table 3: Production function estimates (Chile, ISIC 311)

	Estimates		95% CI		BS normality
Parameter	ACF	CUE	Wald (Non-robust)[1]	AR-GMM (Robust)[2]	SW p-value[3]
β_k	0.040	0.019	[0.008, 0.030]	[0.003, 0.035]	0.896
β_ℓ	0.123	0.089	[0.060, 0.118]	[0.049, 0.129]	0.619
β_m	0.905	0.955	[0.935, 0.975]	[0.923, 0.987]	0.493
Returns to scale	1.067	1.063	[1.048, 1.078]	[1.039, 1.455]	—

Hansen J-test at CUE: $p = 0.776$.

- ACF and CUE coefficients in ballpark of eachother
- Robust confidence intervals are tight
- Hansen test does not reject the over-identifying restrictions.

US manufacturing (NAICS 31)

Table 4: Production function estimates (US, NAICS 31)

	Estimates		95% CI		BS normality
Parameter	ACF	CUE	Wald (Non-robust)	AR-GMM (Robust)	SW p-value
β_k	0.329	0.829	$[-4.241, 5.898]$	$[-0.271, 5.829]$	0.000
β_ℓ	0.132	-0.346	$[-5.363, 4.671]$	$[-4.046, > 5]$	0.000
β_m	0.561	0.290	$[-2.523, 3.103]$	$[< -4, > 5]$	0.000
Returns to scale	1.022	0.773	$[-1.950, 3.497]$	$[-3.127, > 5]$	—

Hansen J-test at CUE: $p = 0.793$

- ACF and CUE coefficients differ substantially
- Robust confidence intervals are very wide, indicating weak information about elasticities
- Hansen test does not reject the over-identifying restrictions

Conclusion

- Weak proxies are a key source of weak identification in CFA
- Robust methods are informative and deliver valid inference
- Monte Carlo: weak proxies $\rightarrow$ large imprecision and non-normality

References

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

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. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054.

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