Estimating Multi-Product Production Functions: What Can We Learn Without Demand Assumptions?

Nicole Scholz

University of Warwick

Job Market Talk

Outline

The Problem

Estimating Multi-Product Production Functions Without Demand Assumptions?

Multi-product firms make up the vast majority of production

- 72.5% in Finland (Kuosmanen & Valmari, 2023)
- 87% in US (Bernard et al., 2010)

Production function estimation is a key tool to estimate

- input-output elasticities → market power (Autor et al., 2020; CMA, 2024; Montag, 2024)
- returns to scale & productivity → firm dynamics (Syverson, 2004; Gao and Kehrig, 2025)

But datasets only contain aggregated inputs at the firm level ⟹ traditional methods fail

Revenue vs Quantity

Traditionally, researchers have used revenue production functions, **but** revenues conflate prices and quantities

$\Longrightarrow$ misleading conclusions about productivity dynamics, e.g.

- overstating role of productivity in firm survival (Foster et al., 2008)
- changing sign of correlations (Hsieh and Klenow, 2009)

$\Longrightarrow$ markups are not identified (Bond et al., 2021; De Ridder et al., 2024)

Solution: Use quantity data, **but** need to solve the input allocation problem

The Current Approach

Current estimators

- make unrealistic assumption on productivity (De Loecker et al., 2016; Gong and Sickles, 2021),
- or impute input allocations using assumptions on market structure, firms' price setting behaviour, and the specific form of demand (Orr, 2022; Valmari, 2023)

However, unclear for many industries how to correctly specify the demand side (e.g. US brewing industry)

- Miller et al. (2021): Model of price leadership → strongly reject Bertrand pricing
- De Loecker and Scott (2024): Model of brewer-supermarket links → fail to reject Bertrand pricing

This Paper

How important are the demand assumptions?

- Identification if and only if some firms are expected to
  - grow arbitrarily fast
  - get arbitrarily close to shutting down
- Therefore, **in finite sample**, any production function can be rationalised with the right demand assumptions

This Paper

How important are the demand assumptions?

- Identification if and only if some firms are expected to
  - grow arbitrarily fast
  - get arbitrarily close to shutting down
- Therefore, **in finite sample**, any production function can be rationalised with the right demand assumptions

A novel estimator for mature industries

- If the **economy-wide** distribution of firm-product productivity pairs is constant over time (stationarity), the production function is **set-identified**

This Paper (Cont'd)

The identified set

- becomes more informative with greater time-series variation in the input distribution
- can be used to test firms' competitive conduct
- can be estimated without instrumental variables and numeric minimizers

This Paper (Cont'd)

The identified set

- becomes more informative with greater time-series variation in the input distribution
- can be used to test firms' competitive conduct
- can be estimated without instrumental variables and numeric minimizers

Using simulations, I show that

- the estimator is robust to mild non-stationarity of the productivity distribution
- near the true parameter vector, rejection becomes increasingly harder

Literature

- Multi-product production function estimation
  - De Loecker et al. (2016) - no within-firm productivity differences
  - Gong and Sickles (2021) - productivity can be decomposed into firm-product fixed effect and time fixed effect
  - Orr (2022) - demand estimation to recover input allocation
  - Valmari (2023) - simultaneous estimation under monopolistic competition
  - Chen and Liao (2022) - simultaneous estimation under flexible demand structure

    I provide a complementary approach to recover identification
- Production frontier estimation
  - Dhyne et al. (2022); Caselli et al. (2025); Cairncross et al. (2023)

    I provide an alternative way to estimate production without needing to allocate inputs
- Partial identification
  - Romano et al. (2014); Kaido et al. (2019)
  - Pakes et al. (2025)

    I further show the usefulness of partial identification in IO

Outline

Motivation

Example: Market power in the beer/cider industry

Goal

- Estimate production function to recover markups and markdowns

Data

- Price ($P_{i,j,t}$) and quantity ($Y_{i,j,t}$) for each beer/cider (i) produced by a firm (j)
- Aggregate inputs ($M_{j,t}$,$L_{j,t}$,$K_{j,t}$) for each firm
- Material prices ($P^M_{j,t}$) for each firm

Assumptions

Firms

- report quantity ($Y_{i,j,t}$) without error,
- produce beer and cider using the same **Cobb Douglas** production function with different **firm-product-time-specific total factor productivities** ($\omega_{i,j,t}$):

$$\ln(Y_{i,j,t}) = \underbrace{\beta_m \ln(M_{i,j,t}) + \beta_l \ln(L_{i,j,t}) + \beta_k \ln(K_{i,j,t})}_{F(\cdot)} + \omega_{i,j,t}$$

- can costlessly re-allocate inputs between products,
- choose total materials ($M_{j,t}$) and input shares ($S_{i,j,t}$) to minimise costs.[1]

$\Longrightarrow$ Firms choose same share for each input

[1] Equivalently, firms maximise profit, but demand is fully flexible.

Outline

Review

How can we estimate production functions?

1. Estimate revenue production functions → incorrect markups
2. Use demand assumptions → unclear which assumptions to use

Table: Markups differ wildly by conduct (De Loecker and Scott, 2024)

Brewer competition	(1)	(2)	(3)
Nash Bertrand	2.29	2.15	1.80
Nation-wide monopoly	14.91	14.88	14.68
Product-by-product	1.57	1.52	1.39
Retail cost correction	Yes	No	No
Distribution cost correction	Yes	Yes	No

AR(1) Productivity

Are these assumptions sufficient to identify the production function?

Make identification easier by imposing an AR(1) productivity process:

$$\omega_{i,j,t} = \rho\omega_{i,j,t-1} + \zeta_{i,j,t}$$

with **known** persistence, $\rho \in (0, 1)$, and innovation,$\zeta_{i,j,t}$, independent across firms and time.

What condition are we trying to satisfy?

- Cost minimisation $\implies$ can always explain input choices
- Only restrictions come from productivity process!

Productivity Residual

Denote the vector of production function parameters as

$$\vec{\theta} = (\beta_m, \beta_l, \beta_k)$$

Can re-write firm-product inputs as sum of firm inputs and firm-product input shares:

$$\ln(M_{i,j,t}) = \ln(M_{j,t}) + \ln(S_{i,j,t})$$

Productivity is residual from an accounting identity:

$$\omega(\vec{\theta}, S, Y, M, L, K) = \ln(Y) - \underbrace{\beta_m \ln(M) - \beta_l \ln(L) + \beta_k \ln(K)}_{F(\vec{\theta}, M, L, K)} - \underbrace{(\beta_m + \beta_l + \beta_k)}_{\text{Returns to Scale}} \ln(S)$$

Recovering Innovation

Can now use the AR(1) productivity process to write

$$\zeta_{i,j,t} = \omega_{i,j,t} - \rho\omega_{i,j,t-1}$$

The innovation in productivity is a difference of two residuals[2]

$$\zeta(\vec{\theta}, \vec{S}_{t:t-1}, \vec{Y}_{t:t-1}, \vec{M}_{t:t-1}, \vec{L}_{t:t-1}, \vec{K}_{t:t-1})$$

$\Longrightarrow$ **Statistical restrictions**: $\zeta(\cdot)$ should be independent of, e.g.

$$\left.\begin{array}{l} \text{Fixed inputs } (L_{j,t}, K_{j,t}) \\ \text{Past inputs } (M_{j,t-1}, L_{j,t-1}, K_{j,t-1}, \dots) \\ \text{Past output } (Y_{i,j,t-1}, Y_{i,j,t-2}, \dots) \end{array}\right\} \vec{Z}_t$$

[2] Let $\vec{S}_{t:t-1} = (S_t, S_{t-1})$.

Moment Equalities

Denote the average innovation by

$$\bar{\zeta} = E[\zeta(\vec{\theta}, \vec{S}_{t:t-1}, \vec{Y}_{t:t-1}, \vec{M}_{t:t-1}, \vec{L}_{t:t-1}, \vec{K}_{t:t-1})]$$

The GMM-style moments are

$$E[\zeta(\vec{\theta}, \vec{S}_{t:t-1}, \vec{Y}_{t:t-1}, \vec{M}_{t:t-1}, \vec{L}_{t:t-1}, \vec{K}_{t:t-1}) \mid \vec{Z}_t] - \bar{\zeta} = 0$$

Why Does Identification Fail?

Key insight: Without additional assumptions, quantity differences are not indicative of differences in input shares

$$Y_{Beer} >> Y_{Cider} \quad \not\Rightarrow \quad S_{Beer} > S_{Cider}$$

But, to rationalise wrong production function parameters ($\vec{\theta} \neq \vec{\theta}^*$)

- if log input differences are large ($|E[\ln(M_t) - \rho \ln(M_{t-1}) \mid \vec{Z}_t]| >> 0$)
- need extreme input shares ($E[\ln(S_t) \mid \vec{Z}_t] << 0 \to \infty$)

Therefore,

- as the log input difference goes to infinity ($|E[\ln(M_t) - \rho \ln(M_{t-1}) \mid \vec{Z}_t]| \to \infty$) we can recover identification in the limit
- but finite samples provide no restrictions on $\vec{\theta}$ without further assumptions

Identification at Infinity

Theorem - Identification of β_M

Under the above assumptions and two technical assumptions on the distribution of the true input shares ($S^*_{i,j,t}$) and productivities ($\omega^*_{i,j,t}$), β_M is identified if and only if there is enough independent variation in total inputs, $(M_{j,t}, L_{j,t}, K_{j,t})$, and

$$\exists\{\vec{Z}_r\}_{r=1}^{\infty} \text{ s.t. } \lim_{r\to\infty} |E[\ln(M_t) - \rho \ln(M_{t-1})|\vec{Z}_t = \vec{Z}_r]| = \infty.$$

Implications

Condition

$$\exists\{\vec{Z}_r\}_{r=1}^{\infty} \text{ s.t. } \lim_{r\to\infty} |E[\ln(M_t) - \rho\ln(M_{t-1})|\vec{Z}_t = \vec{Z}_r]| = \infty$$

Intuitively, need to expect some firms to either

- get arbitrarily close to shutting down ($\ln(M_t) - \rho\ln(M_{t-1}) \to -\infty$)
- grow arbitrarily fast ($\ln(M_t) - \rho\ln(M_{t-1}) \to \infty$).

The proof takes any vector of production function parameters, $\vec{\theta} = (\beta_M, \beta_L, \beta_K)$, and constructs input allocations that rationalise them.

$\implies$ **Finite samples**, contain no information without further assumptions

Outline

Moving to Identification

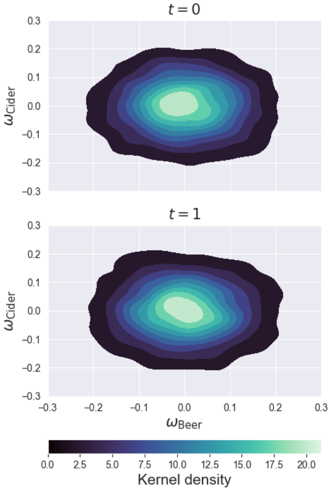
Problem: Unrestricted input shares/productivity
$\rightarrow$ non-identification

Potential solution: Industry is in long-run equilibrium (productivity dynamics relatively stable)

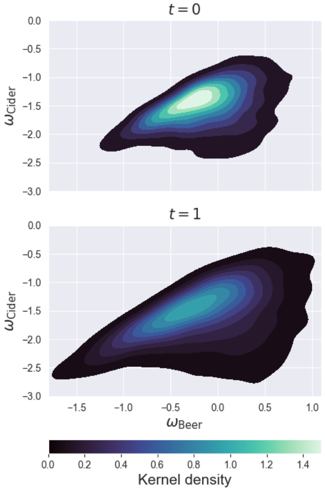
Assumption: Stationarity

The data-generating joint probability density of productivity, denoted by $\Omega^*(\cdot)$, is stationary; that is,

$$\Omega^*(\cdot|t) = \Omega^*(\cdot|t') \quad \forall t, t'.$$



Testing Stationarity



Key question

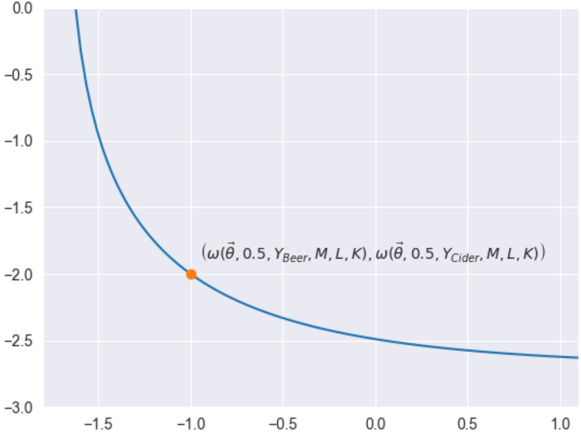
When does a (conditional) distribution of input shares, $\mathcal{S}$ exist so $\Omega(\cdot;\vec{\theta},\mathcal{S})$ is stationary?

OR

Given $\Omega(\cdot;\vec{\theta},\mathcal{S})$, when can we reallocate inputs to make the distribution stationary?

$\Omega(\cdot;\vec{\theta},\mathcal{S})$

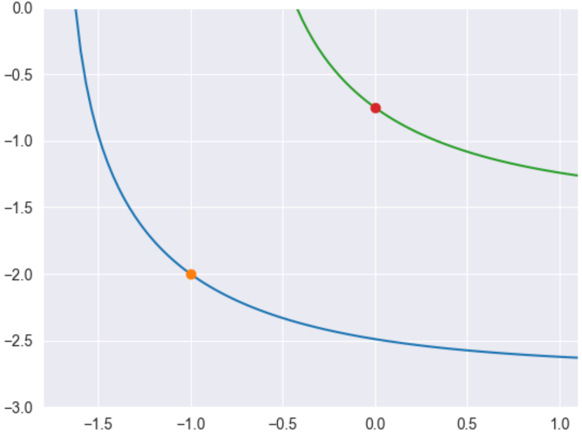
Potential Productivities



Zoom in on one firm - split its inputs equally:

- Can move its productivity along the blue line

Potential Productivities



Zoom in on one firm - split its inputs equally:

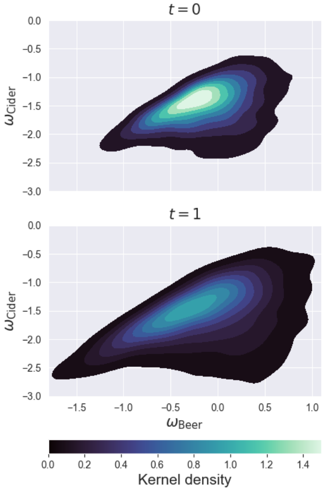
- Can move its productivity along the blue line

Grab a second, strictly more productive, firm:

- The green and blue lines never cross

$\implies$ can use this to construct an 'only if' condition

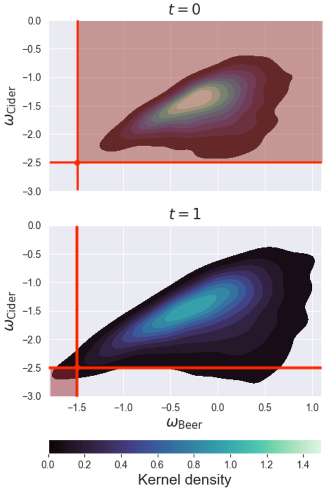
Bounding the Identified Set



Key insight: Cannot make a firm more (or less) productive at producing all its outputs

1. Pick a **reference** point, $\vec{\omega}_r = (\omega_{r1}, \omega_{r2})$
2. Calculate % of firms strictly **less efficient** than $\vec{\omega}_r$ in period t
3. Calculate % of firms strictly **more efficient** than $\vec{\omega}_r$ in period t'
4. $Pr(\vec{\omega} > \vec{\omega}_r | t) + Pr(\vec{\omega} < \vec{\omega}_r | t') > 100\% \implies \vec{\theta}$ cannot rationalise the data

Bounding the Identified Set



Key insight: Cannot make a firm more (/less) productive for all outputs

1. Pick a **reference** point, $\vec{\omega}_r = (\omega_{r1}, \omega_{r2})$
2. Calculate % of firms strictly **less efficient** than $\vec{\omega}_r$ in period t
3. Calculate % of firms strictly **more efficient** than $\vec{\omega}_r$ in period t'
4. $Pr(\vec{\omega} > \vec{\omega}_r | t) + Pr(\vec{\omega} < \vec{\omega}_r | t') > 100\% \implies \vec{\theta}$ cannot rationalise the data

Theorem

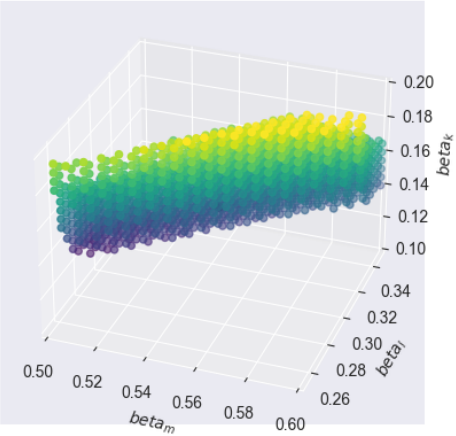
If the marginal distribution of at least one input differs between any two periods, e.g. $g_{M|t}(\cdot) \neq g_{M|t'}(\cdot)$, then there are at least some parameter values $\vec{\theta} \neq \vec{\theta}^*$ that cannot rationalise the distribution of observables, i.e.

$$\Theta_I^C := \left\{ \vec{\theta} : \Omega(\cdot; \vec{\theta}, \mathcal{S}|t) \neq \Omega(\cdot; \vec{\theta}, \mathcal{S}|t') \text{ for all } \mathcal{S} \right\} \neq \emptyset.$$

Additionally, if **any** distribution of inputs can be observed with positive probability, then the production function is identified at the limit, i.e. $\Theta_I \to \{\vec{\theta}^*\}$ as $T \to \infty$.

Outline

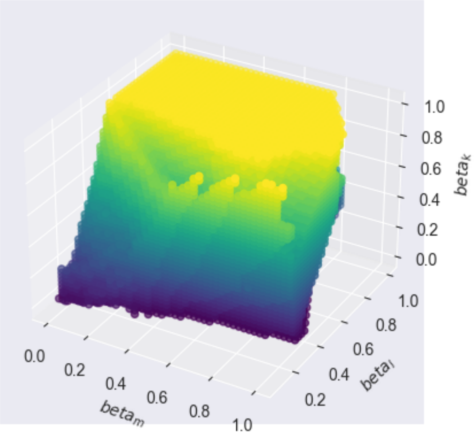
Misspecification



Increase variance of initial productivity draws relative to stationary equilibrium (7.84x σ_ω^*)

- Even with highly informative sets, the true vector (θ^*), was never rejected over 100 datasets
- However, it seems that the identified set shrinks

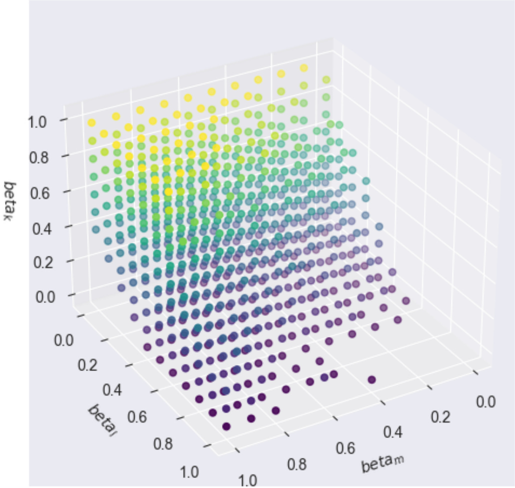
Reference Productivities and Moment Inequalities



There exists a set of moment inequalities for every $\vec{\omega}_r \in \mathbb{R}^2$

- Even with extreme underlying variation, choosing the wrong equalities $\rightarrow$ less informative sets
- Choosing moments optimally is important

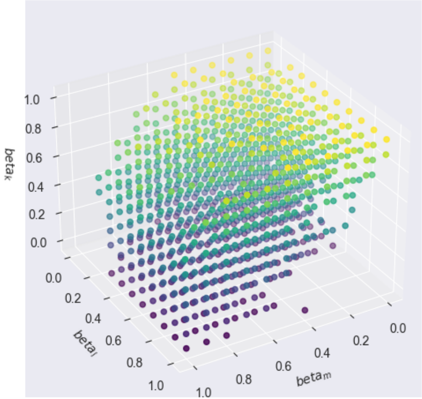
Aggregate Input Shocks

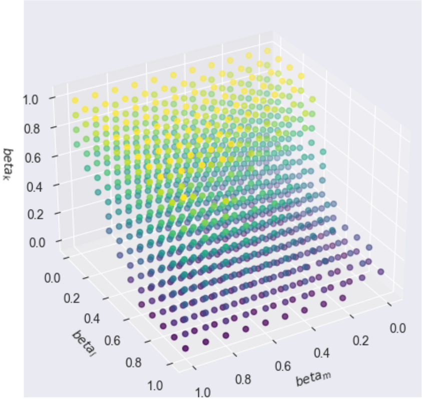


Exogenously increase materials by 20% (ceteris paribus):

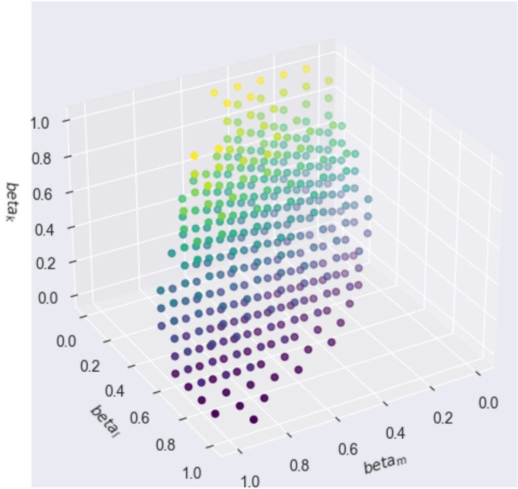
- Bounds are uninformative
- But the set is very informative

Aggregate Input Shocks





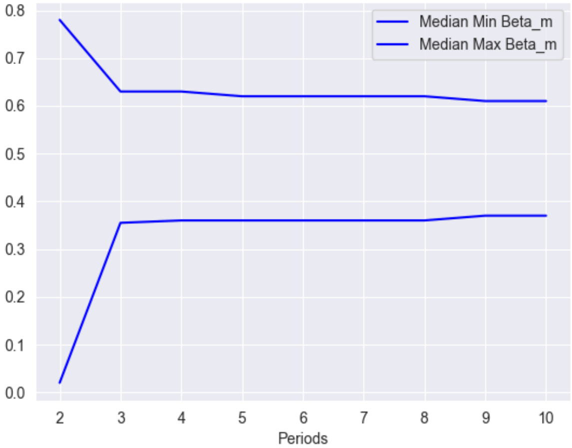
The Sum Is Greater Than Its Parts



Combining the datasets gives a lot more information

- How informative depends on the question we want to answer
- E.g. for testing firms' conduct the absolute size of the set is unimportant

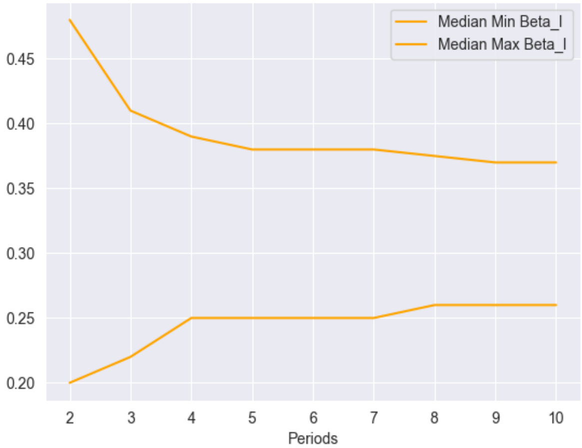
Identified Set Shrinks Non-Monotonically



Adding additional periods shows:

- As the variance in the data increases bounds become more informative at a **decreasing rate**
- Bounds are not symmetric around the true parameter ($\beta_m^* = 0.55$)

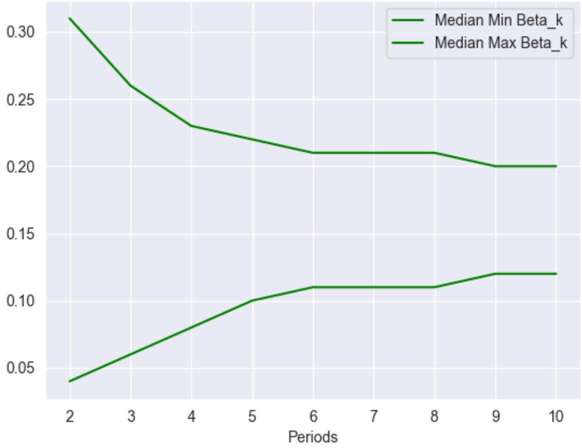
Identified Set Shrinks Non-Monotonically



Adding additional periods shows:

- As the variance in the data increases bounds become more informative at a **decreasing rate**
- Bounds are not symmetric around the true parameter ($\beta_l^* = 0.3$)

Identified Set Shrinks Non-Monotonically



Adding additional periods shows:

- As the variance in the data increases bounds become more informative at a **decreasing rate**
- Bounds are not symmetric around the true parameter ($\beta_k^* = 0.15$)

Summary

Key take-aways

- Without demand/additional assumptions input shares are unconstrained $\rightarrow$ any set of production parameters, $\vec{\theta}$, can be rationalised
- When the productivity distribution is stationary, the production function is set-identified
- Can estimate the identified set by adding the percentage of firms better/worse than a reference level of productivity

Continued work

- How to optimally choose the moment inequalities
- How does the estimator perform on real-world datasets

Research Agenda

Providing tools for policy-makers and studying firms, market power, and innovation

- When is multi-good production inherently interdependent? (Work in Progress)
  - Returns to scope have received large amount of attention
  - Current production frontier methods limited to assuming interdependence
  - → Show under what conditions a production frontier can be derived from independent production functions
- Are demand or supply factors driving firms to diversify production?
  - Combine demand and production function estimation tools
  - Quantify impact of return to scope, productivity dynamics, demand drivers
- Can production processes predict innovation?
  - Use data on firms' collateral for detailed capital information
  - Using machine learning, map out production processes and calculate closeness
  - Understand how multi-product production can improve innovation

Thank you!

Appendix

Assumption - Stationarity

Let $\mathcal{S}$ be the distribution of input shares conditional on the firm's observed inputs and outputs

$$\mathcal{S} : [0,1]^2 \times \mathbb{R}^5 \to \mathbb{R}_+$$

Remember, that productivity is a residual

$$\omega(\vec{\theta}, s, \beta_0, y, m, l, k) = y - F(\vec{\theta}, \beta_0, m, l, k) - (\beta_m + \beta_l + \beta_k)s$$

Then the counterfactual productivity distribution $\Omega(\vec{\theta}, \mathcal{S})$ is the distribution of ω induced by the empirical distribution and $\mathcal{S}$

$$\Omega(\vec{\omega}; \vec{\theta}, \mathcal{S}|t) \;=\; \int_{\mathbb{Y}\times\mathbb{X}} \int_{[0,1]^n} \delta\Big(\vec{\omega} - \vec{\omega}(\vec{\theta}, \vec{s}, \vec{y}, \vec{x})\Big)\, \mathcal{S}_t(\vec{y}, \vec{x}, \vec{s})\, g_{\vec{Y},\vec{X}|t}(\vec{y}, \vec{x})\, d\vec{s}\, d\vec{y}\, d\vec{x},$$

where $\delta(\cdot)$ is the Dirac delta function. Intuitively, $\Omega_t(\vec{\omega}; \vec{\theta}, \mathcal{S})$ aggregates the probability mass of all triples $(\vec{y}, \vec{x}, \vec{s})$ that map into the productivity vector $\vec{\omega}$.

Back

Assumptions

Bounded input shares

All moments of the conditional log input differences exist and are bounded. For every $\rho \in (0,1)$

$$\sup_{i,t} \operatorname*{ess\,sup}_{\vec{Z}_t \in \mathbb{Z}_t} E\left[(s^*_{i,t} - \rho s^*_{i,t-1})^d | \vec{Z}_t\right] < \infty, \ \forall d \in \mathbb{N},$$

where $\mathbb{Z}_t = \{\vec{Z}_t : \vec{Z}_t \in \text{supp}(g_{\vec{Z}_t}(\cdot))\}$ is the set of observed controls.

Bounded productivity

All the moments of last period's productivity exist and are bounded:

$$\sup_{i,t} \operatorname*{ess\,sup}_{\vec{Z}_t \in \mathbb{Z}_t} E[|\omega^*_{i,t-1}|^d | \vec{Z}_t] < \infty \ \forall d \in \mathbb{N}$$

Back

Theorem

Under the above assumptions, β_{X_m} is identified if and only if

$$\exists\{\vec{Z}_r\}_{r=1}^{\infty} \text{ s.t. } \lim_{r\to\infty} |E[x_{m',t} - \rho x_{m',t-1}|\vec{Z}_t = \vec{Z}_r]| = \infty$$

and, for all $m' \neq m$, one of the following holds:

(a) $\exists\{\vec{Z}_r\}_{r=1}^{\infty}$ such that

$$\lim_{r\to\infty} \frac{E[x_{m',t} - \rho x_{m',t-1}|\vec{Z}_t = \vec{Z}_r]}{E[x_{m,t} - \rho x_{m,t-1}|\vec{Z}_t = \vec{Z}_r]} = c,$$

for $c \in \{-\infty, 0, \infty\}$

(b) $\exists\{\vec{Z}_r\}_{r=1}^{\infty}$, $\{\vec{Z'}_r\}_{r=1}^{\infty}$ such that

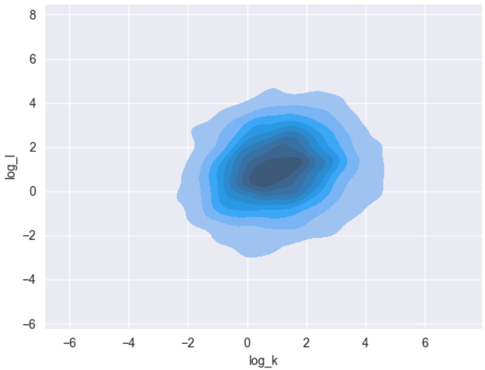
$$\lim_{r\to\infty} \frac{E[x_{m',t} - \rho x_{m',t-1}|\vec{Z}_t = \vec{Z}_r]}{E[x_{m,t} - \rho x_{m,t-1}|\vec{Z}_t = \vec{Z}_r]} = c,$$

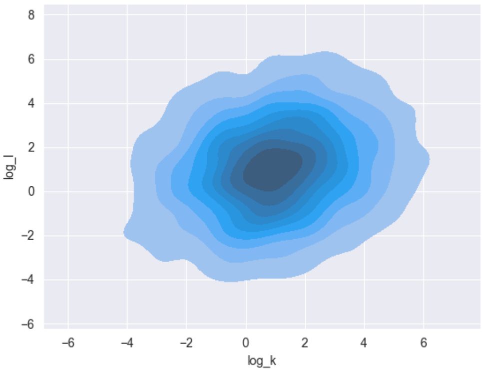
$$\lim_{r\to\infty} \frac{E[x_{m',t'} - \rho' x_{m',t'-1}|\vec{Z}_t = Z'_r]}{E[x_{m,t'} - \rho x_{m,t'-1}|\vec{Z}_t = Z'_r]} = c',$$

for $c \neq c'$.

Back

Underlying distributions





Underlying distributions