Heterogenous Production Functions, Panel Data, and Productivity Dispersion

Jeremy T. Fox[1] Vitor Hadad[2] Stefan Hoderlein[3]
Amil Petrin[4] Robert P. Sherman[5]

[1]Rice University & NBER

[2]Stanford University

[3]Boston College

[4]University of Minnesota & NBER

[5]California Institute of Technology

UBC
March 2019

Panel Data

- **Online marketing** One popular website, single advertising campaign
  - Panel ID: browser cookies or required to log in
- **Productivity** firm or plant level data for a medium to large country
  - Panel ID: government statistical agency collects data
  - *Productivity dispersion in India today*
- **Nonparametric identification, estimation**
- Identification could motivate parametric estimation approaches

Some Uses of Panel Data

- *Some* **non-structural** machine learning
  - Predictive nonparametric regression
  - Polynomials, splines, trees, nearest neighbor, ...
- *Some* **structural demand estimation** with consumer panel data
  - Unobservable preference heterogeneity (random coefficients)
  - Dynamics of consumer decision making (learning, stockpiling)
- *Some* theoretical **econometricians**
  - **Correlation of unobservables and explanatory variables**
  - Few **conduct** assumptions in model for explanatory variables
  - Conduct: no first order condition for firm advertising choice, say
  - Common paradigm: **fixed effects**

Correlated Random Coefficients

- **Correlated random coefficients** combine ideas from
  - Heterogeneous parameters
  - Fixed effects
- Unobserved heterogeneity from, say, consumer panels or production data
  - **Random coefficients** in regression setting
  - *Treatment effect heterogeneity*
  - Random coefficients **not fixed over time** for individual, firm
- Random coefficients can be **correlated with explanatory variables**
  - Firms with more return to capital use more capital
  - Consumers more responsive to ads targeted by more ads
  - No explicit **conduct** model of capital or advertising FOCs
  - Certainly behavioral assumptions

Applications

- Consumer panels for demand estimation
  - Many endogeneity issues with non randomized marketing exposures
  - Fox, Hoderlein & Shashoua (2019): TV advertising on offline sales
- **Productivity in India**
  - Plant level productivity data on India
  - Debate about low, dispersed productivity in India
  - Random coefficients: **vector-valued productivity**

Correlated Random Coefficients

- Linear equation with random coefficients

$$Y_{i,t} = A_{i,t} + B_{i,t} X_{i,t}$$

- $Y_{i,t}$ outcome variable of interest
- $X_{i,t}$ explanatory variable of interest
- $A_{i,t}$ additive unobservable
- $B_{i,t}$ random slope
- Call *both* $A_{i,t}, B_{i,t}$ **random coefficients**
- Identify **moments of distribution**

$$F_t\left(A_{i,t}, B_{i,t} \mid X_{i,t}\right)$$

- Random coefficients $A_{i,t}, B_{i,t}$ can be **correlated** with $X_{i,t}$
- Single X with random coefficient for now, **more later**

Productivity Dispersion

- Cobb Douglas production function

$$Y_{i,t} = A_{i,t} + B^K K_{i,t} + B^L L_{i,t}$$

- **Productivity dispersion**: standard deviation of $A_{i,t}$ across plants or firms in same industry
- Syverson (2011) surveys productivity dispersion findings
- Typical finding: some plants produce more than **twice as much** output for same inputs
- *Assuming* $\left(B^K, B^L\right)$ *identical for plants in an industry too restrictive?*
- If $\left(B^K_{i,t}, B^L_{i,t}\right)$ vary by plant i & time, much of this heterogeneity enters $A_{i,t}$

Constant Returns to Scale

- Impose **constant returns to scale** (CRS)
- Skip simple algebra…

$$Y_{i,t} = A_{i,t} + B_{i,t} X_{i,t}$$

- $Y_{i,t}$ log output minus log labor under CRS
- $X_{i,t}$ log capital minus log labor
- $A_{i,t}$ total factor productivity
- $B_{i,t}$ input elasticity for capital, **heterogeneous** in this paper
- **Random coefficients** $A_{i,t}, B_{i,t}$ can be **correlated** with $X_{i,t}$
- Identify **moments of distribution**

$$F_t\left(A_{i,t}, B_{i,t} \mid X_{i,t}\right)$$

Moments & Quantiles to Identify

- **Cobb Douglas production**, constant returns to scale

$$Y_{i,t} = A_{i,t} + B_{i,t} X_{i,t}$$

- $E[A_{i,t} \mid X_{i,t}]$ firms with **more inputs** / capital intensity have higher TFP? **Allocation of inputs**
- $E[A_{i,t}]$ **unconditional mean of log TFP**, could see how log TFP varies over time, across countries
- $\mathrm{SD}[A_{i,t} \mid X_{i,t}]$ dispersion of log TFP for firms using **same capital intensity**
- $\mathrm{SD}[A_{i,t}]$ **dispersion is log TFP** within, say, a country? Huge literature
- $E[B_{i,t} \mid X_{i,t}]$ firms with **more capital intensity** have higher capital elasticities?
- $E[B_{i,t}]$ **unconditional mean capital elasticity**: percentage increase in output for percentage increase in capital
- $\mathrm{SD}[B_{i,t}]$ how **dispersed is technology for capital** across firms?

Endogeneity Concerns

- **Cobb Douglas production**, constant returns to scale

$$Y_{i,t} = A_{i,t} + B_{i,t} X_{i,t}$$

- Firms with higher capital elasticity $B_{i,t}$ might use more capital, be more capital intensive
- Firms with more log TFP $A_{i,t}$ might use more inputs

Two Time Periods

- Leave productivity example
- Simplify: periods $t = 1$ & $t = 2$ (fixed $T = 2$)
- Regression with random coefficients

$$Y_{i,1} = A_{i,1} + B_{i,1} X_{i,1}$$
$$Y_{i,2} = A_{i,2} + B_{i,2} X_{i,2}$$

- Identify **moments of distribution**

$$F\left(A_{i,1}, A_{i,2}, B_{i,1}, B_{i,2} \mid X_{i,1}, X_{i,2}\right)$$

- Random coefficients $(A_{i,t}, B_{i,t})$ **correlated** with $X_{i,1}$, $X_{i,2}$
- $F_t\left(A_{i,t}, B_{i,t} \mid X_{i,t}\right)$ can shift over time

Innovations to Random Coefficients

- Add time series process for **innovations to random coefficients** (shocks)

$$Y_{i,1} = A_{i,1} + B_{i,1}X_{i,1}$$
$$Y_{i,2} = A_{i,2} + B_{i,2}X_{i,2}$$
$$A_{i,2} = A_{i,1} + U_{i,2}$$
$$B_{i,2} = B_{i,1} + V_{i,2}$$

- Drop i subscripts in what follows
  - Almost everything depends on i in this paper!
- A_t & B_t: **random walks with drift**
  - AR(1) results in paper
- U_2 & V_2: innovations to production functions
- Innovations not necessarily mean 0, allow drift in random coefficients distribution!
- Seasonality, learning, productivity growth, etc.

Identification Assumptions

- Recall dropping i subscript

$$
\begin{aligned}
Y_1 &= A_1 + B_1 X_1 \\
Y_2 &= A_2 + B_2 X_2 \\
A_2 &= A_1 + U_2 \\
B_2 &= B_1 + V_2
\end{aligned}
$$

- **Most important identification assumptions**

$$
\begin{aligned}
&(U_2, V_2) \perp (A_1, B_1) \mid X_1, X_2 \\
&(U_2, V_2) \perp (X_1, X_2)
\end{aligned}
$$

Productivity: Identification Assumptions

- **Most important identification assumptions**

$$(U_2, V_2) \perp (A_1, B_1) \mid X_1, X_2$$
$$(U_2, V_2) \perp (X_1, X_2)$$

- Productivity: **timing assumption**
- Firms choose inputs X_2 in previous period 1 before (U_2, V_2) seen
- Story has economic content: could verify by talking to firms about when they choose inputs

Productivity: Timing Assumption vs Proxy Variable Literature

- **Most important identification assumptions**

$$(U_2, V_2) \perp (A_1, B_1) \mid X_1, X_2$$
$$(U_2, V_2) \perp (X_1, X_2)$$

- Related timing assumption to literature **without random coefficient** on capital
- Olley & Pakes (1996), Levinsohn & Petrin (2003), Wooldridge (2009), Ackerberg et al (2015), Gandhi et al (2018)
- Some inputs chosen in period $t-1$ with knowledge of only $t-1$ production function
- A_t often nonparametric Markov, distribution $\Pr(A_t \mid A_{t-1})$
- We have random walk, AR(1) extension
- We do not allow i.i.d. output error not in A_t (proxy variable methods require scalar unobservables)
- Related: Arellano & Bond (1991) literature

No Conduct Assumptions Like Static FOCs

- Estimate B as expenditure share on capital input
- Based on **static profit maximizing** FOC for Cobb-Douglas
- Treats input as having no adjustment costs, opposite of our approach
- Based on strong **conduct assumption**: static cost minimization
- See Doraszelski & Jaumandreu (2013, 2016), Gandhi et al (2015)
- We do not impose profit maximization or cost minimization, **important for studying unproductive firms**
- We allow unobservables outside production function to affect input choice
  - Firm-specific input prices
  - Adjustment costs
  - Product demand

Consumer Panels: Analogous Issues

- **Most important identification assumptions**

$$(U_2, V_2) \perp (A_1, B_1) \mid X_1, X_2$$
$$(U_2, V_2) \perp (X_1, X_2)$$

- **Targeting** based on previous behavior, not current period
- Targeting can be complex function of past X's
- No explicit targeting FOC / supply side model
- Higher frequency data?
- Consumer preferences evolve over time

Correlated Random Coefficients Literature

- Literature on **correlated random coefficients** in panel data
  - Chamberlain (1992), Arellano & Bonhomme (2012), Graham & Powell (2012), Evdokimov (2011), Laage (2018), etc.
- Models have **time invariant** random coefficients (A_i, B_i)

$$Y_{i,t} = A_i + \bar{A}_{i,t} + B_i X_{i,t}$$

- Correlation with $X_{i,t}$ usually only arises for time invariant components (A_i, B_i)
- Sometimes, $\bar{A}_{i,t}$ assumed independent of $\bar{A}_{i,t-1}$
- Time invariance assumption impede investigating firm or industry growth
- Our results only for scalar $X_{i,t}$

Recent Papers on Heterogeneous Production Functions

- Kasahara, Shrimpf & Suzuki (2017)
  - Kasahara and Shimotsu (2009) finite mixture approach
- Balat, Brambilla & Sazaki (2018)
  - Extend proxy variable approaches but require variables to be chosen in period $t-1$ with only static considerations, as in FOCs
- Navarro & Rivers (2018)
  - Extend Gandhi, Navarro & Rivers (2018)
- Li & Sasaki (2018)
  - Scalar underlying random variable drives multiple Cobb-Douglas parameters
- Ackerberg & Hahn (2016)
  - Scalar unobservable enters into nonparametric production function

All Identification Assumptions

- Model, dropping *i* subscript

$$
\begin{aligned}
Y_1 &= A_1 + B_1 X_1 \\
Y_2 &= A_2 + B_2 X_2 \\
A_2 &= A_1 + U_2 \\
B_2 &= B_1 + V_2
\end{aligned}
$$

- **Most important identification assumptions**

$$
\begin{aligned}
&(U_2, V_2) \perp (A_1, B_1) \mid X_1, X_2 \\
&(U_2, V_2) \perp (X_1, X_2)
\end{aligned}
$$

- **Other identification assumptions**
  - X_1 & X_2 have continuous, common support
  - X_1 & X_2 not linearly dependent

Identify Conditional Moments

- **Conditional distribution of random coefficients**

$$F(A_1, A_2, B_1, B_2 \mid X_1, X_2)$$

- **Means** $E[A_1 \mid X_1, X_2]$, $E[B_1 \mid X_1, X_2]$, $E[A_2 \mid X_1, X_2]$, $E[B_2 \mid X_1, X_2]$
- **Standard deviations** $\mathrm{SD}[A_1 \mid X_1, X_2]$, $\mathrm{SD}[B_1 \mid X_1, X_2]$, $\mathrm{SD}[A_2 \mid X_1, X_2]$, $\mathrm{SD}[B_2 \mid X_1, X_2]$
- Can then identify **unconditional moments** and other estimands

$$E[A_1] = E_{X_1,X_2}(E[A_1 \mid X_1, X_2])$$
$$\mathrm{median}_{X_1,X_2}(E[A_1 \mid X_1, X_2])$$

- Also identify entire joint distribution $F(A_1, A_2, B_1, B_2 \mid X_1, X_2)$
  - Characteristic functions, in paper

Identification & Estimation Steps

1. Identify unconditional moments of innovations

$$E[U_2], E[V_2]$$

2. Identify conditional moments of outcomes like

$$E[Y_1 \mid X_1, X_2]$$

   - Directly from data
   - Nonparametric regression

3. Identify first period conditional moments of random coefficients like

$$E[A_1 \mid X_1, X_2]$$

4. Identify second period conditional moments like

$$E[A_2 \mid X_1, X_2]$$

5. Identify unconditional moments, other estimands

$$E[A_1] = E_{X_1,X_2}(E[A_1 \mid X_1, X_2])$$

$$\mathrm{median}_{X_1,X_2}(E[A_1 \mid X_1, X_2])$$

Step 1: Means of Innovations

- Identify means of innovations $E[U_2]$, $E[V_2]$
- Focus on (X_1, X_2) such that $X_1 = X_2 = x$: input choices remain constant
- Difference time periods using panel data, giving

$$Y_2 - Y_1 = (A_2 - A_1) + (B_2 - B_1)x = U_2 + V_2 x$$

- Take expectations

$$E[Y_2 - Y_1 \mid X_1 = X_2 = x] = E[U_2] + E[V_2]x$$

- Uses mean independence relaxation of

$$(U_2, V_2) \perp (X_1, X_2)$$

- Local kernel regression, local to $X_1 = X_2 = x$, identify $E[U_2]$, $E[V_2]$
- Estimation: with optimal kernel bandwidth, rate 2/5 (less than 1/2), normality

Step 1: Intuition for Identifying Innovation Moments

- Consider plants that do not change input X
- Mean of change in output only from innovations

Step 2: Outcome Moments

- Data directly identify conditional moments of outcomes

$$E\left[Y_1 \mid X_1, X_2\right], E\left[Y_1^2 \mid X_1, X_2\right], E\left[Y_1 \cdot Y_2 \mid X_1, X_2\right]$$

- Estimation: nonparametric regression
- Kernel with optimal bandwidth gives rate 1/6, asymptotic normality
- Have also simulated splines, polynomials with Lasso

Step 3: Two Linear Equations for Means

- Production functions

$$Y_1 = A_1 + B_1 X_1$$
$$Y_2 = A_2 + B_2 X_2$$

- Then

$$E[Y_1 \mid X_1, X_2] = E[A_1 \mid X_1, X_2] + E[B_1 \mid X_1, X_2] x_1$$
$$E[Y_2 \mid X_1, X_2] = E[A_2 \mid X_1, X_2] + E[B_2 \mid X_1, X_2] x_2$$

- Rewrite period 2 production function at $X_1 = x_1$ and $X_2 = x_2$

$$E[Y_2 \mid X_1, X_2] =$$
$$E[A_1 + U_2 \mid X_1, X_2] + E[B_1 + V_2 \mid X_1, X_2] x_2 =$$
$$E[A_1 \mid X_1, X_2] + E[B_1 \mid X_1, X_2] x_2 + E[U_2] + E[V_2] x_2$$

Step 3: Means of Random Coefficients

- Fix $X_1 = x_1$ and $X_2 = x_2$
- Two linear equations

$$E[Y_1 \mid X_1, X_2] = E[A_1 \mid X_1, X_2] + E[B_1 \mid X_1, X_2]x_1$$
$$E[Y_2 \mid X_1, X_2] = E[A_1 \mid X_1, X_2] + E[B_1 \mid X_1, X_2]x_2 + E[U_2] + E[V_2]x_2$$

- $E[U_2]$, $E[V_2]$, $E[Y_1 \mid X_1, X_2]$, $E[Y_2 \mid X_1, X_2]$ previously identified
- Two unknowns: $E[A_1 \mid X_1, X_2]$ and $E[B_1 \mid X_1, X_2]$
- Linear equations have unique solution if $x_1 \neq x_2$
- Identify **conditional means of random coefficients** $E[A_1 \mid X_1, X_2]$ and $E[B_1 \mid X_1, X_2]$ for $x_1 \neq x_2$
- **Step 3 avoids variation in X_1, X_2 because of panel data solution to endogeneity**

Step 4: Second Period Means of Random Coefficients

- Identified for $X_1 = x_1$ and $X_2 = x_2$

$$E[A_1 \mid X_1, X_2], E[B_1 \mid X_1, X_2], E[U_2], E[V_2]$$

- Now identify conditional means of second period random coefficients for $X_1 = x_1$ and $X_2 = x_2$

$$E[A_2 \mid X_1, X_2] = E[A_1 + U_2 \mid X_1, X_2] = E[A_1 \mid X_1, X_2] + E[U_2]$$
$$E[B_2 \mid X_1, X_2] = E[B_1 + V_2 \mid X_1, X_2] = E[B_1 \mid X_1, X_2] + E[V_2]$$

Step 5: Unconditional Moments

- Form unconditional moments such as

$$E[A_1] = E_{X_1,X_2}[E[A_1 \mid X_1, X_2]]$$

- Identified if points $X_1 = X_2$ have measure zero
- Consistency, estimation discussed later: need censoring
- Censoring requires new consistency theorem, below
- Also identify quantiles of conditional moments

$$\text{median}_{X_1,X_2}(E[A_1 \mid X_1, X_2])$$

Use of Means of Random Coefficients

- Identified

$$E[A_1 \mid X_1, X_2], E[A_1], E[B_1 \mid X_1, X_2], E[B_1], E[U_2], E[V_2]$$

- Explore growth in mean TFP & mean input elasticity (say capital elasticity) over time
- Compare *unconditional means* of random coefficients across industries, countries
- *Conditional means*: inputs **allocated to most productive firms?**
- Later: **aggregate productivity growth**

Second Moments: Productivity Dispersion

- Our empirical focus is **productivity dispersion**
- Productivity dispersion refers to **second moment** of productivity
- In our model, productivity refers to both TFP & capital input elasticity (random coefficients)
- How firms vary in TFP / capital input elasticities?
- Identify conditional **standard deviation** and **correlation** of TFP and input elasticity…

$$\mathrm{SD}(A_1 \mid X_1, X_2), \mathrm{SD}(B_1 \mid X_1, X_2), \\ \mathrm{Corr}(A_1, B_1 \mid X_1, X_2), \mathrm{Corr}(A_1, A_2 \mid X_1, X_2), \ldots$$

- Also unconditional moments such as $\mathrm{SD}(A_1)$ if points $X_1 = X_2$ have probability zero

Step 1: Variances, Covariances of Shocks (U_2, V_2)

- Like with first moments, focus on $X_1 = X_2 = x$: input choices remain constant
- First differences

$$Y_2 - Y_1 = (A_2 - A_1) + (B_2 - B_1)x = U_2 + V_2 x$$

- Take conditional variance & exploit independence from X of second moments of innovations

$$\mathrm{Var}[Y_2 - Y_1 \mid X_1 = X_2 = x] = \mathrm{Var}[U_2] + \mathrm{Var}[V_2]x^2 + 2\mathrm{Cov}(U_2, V_2)x$$

- Local variation in x identifies $\mathrm{Var}[U_2]$, $\mathrm{Var}[V_2]$ & $\mathrm{Cov}(U_2, V_2)$

Step 3: Variances, Covariances of Random Coefficients

- Skip steps 2, 4, 5 for conciseness
- Fix $X_1 = x_1$ and $X_2 = x_2$
- Algebra and our main independence assumption show

$$\begin{aligned}\mathrm{Var}(Y_1 \mid X_1, X_2) = \mathrm{Var}(A_1 \mid X_1, X_2) + \mathrm{Var}(B_1 \mid X_1, X_2)\, x_1^2 + \\ 2\mathrm{Cov}(A_1, B_1 \mid X_1, X_2)\, x_1\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(Y_2 \mid X_1, X_2) = \mathrm{Var}(A_1 \mid X_1, X_2) + \mathrm{Var}(B_1 \mid X_1, X_2)\, x_2^2 + \\ 2\mathrm{Cov}(A_1, B_1 \mid X_1, X_2)\, x_2 + \mathrm{Var}(U_2 + V_2 x_2)\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}(Y_1, Y_2 \mid X_1, X_2) = \mathrm{Var}(A_1 \mid X_1, X_2) + \mathrm{Var}(B_1 \mid X_1, X_2)\, x_1 x_2 + \\ \mathrm{Cov}(A_1, B_1 \mid X_1, X_2)(x_1 + x_2)\end{aligned}$$

- Three linear equations, three unknowns
- Identify

$$\begin{aligned}\mathrm{Var}(A_1 \mid X_1 = x_1, X_2 = x_2),\ \mathrm{Var}(B_1 \mid X_1 = x_1, X_2 = x_2), \\ \mathrm{Cov}(A_1, B_1 \mid X_1 = x_1, X_2 = x_2)\end{aligned}$$

- Again, no identification when $x_1 = x_2$

Use of Second Moments

- Literature on **productivity dispersion within an industry** all about **second moments**
- Here, identify **dispersion in production functions** instead of just dispersion in total factor productivity, A_t
- Generalize output level & output growth decompositions for higher dimensional notion of productivity
- Decompositions involve productivity & allocation of inputs to firms
- We identify second moments conditional on $X_1 = x_1, X_2 = x_2$
- Important for questions on allocation of inputs

Three or More Random Coefficients

- Say three random coefficients: two slopes, one intercept
- Not surprisingly need three time periods

$$Y_1 = A_1 + B_1 X_1 + C_1 Z_1$$
$$Y_2 = A_2 + B_2 X_2 + C_2 Z_2$$
$$Y_3 = A_3 + B_3 X_3 + C_3 Z_3$$

- **New timing assumption**

$$(U_2, V_2, W_2, U_3, V_3, W_3) \perp (A_1, B_1) \mid X_1, X_2, X_3$$
$$(U_2, V_2, W_2, U_3, V_3, W_3) \perp (X_1, X_2, X_3)$$

- Firm makes decisions in period 1 (or before) about inputs in periods 2,3
- Can ask firms when they make decisions, as before
- Relates to decision period of firm & frequency of data
- If previous timing assumption valid for annual data, **new one valid for data every six months**
  - Ackerberg (2018): same point to lower standard errors

Estimation Steps

1. Estimate moments of innovations U_2, V_2
   - Kernel local to $x_1 = x_2$, local polynomial regression
2. Estimate conditional moments of outcomes like $E[Y_1 \mid X_1, X_2]$ using Nadarya-Watson kernels
   - Nonparametric regression: could use other methods (machine learning...)
3. Estimate first period conditional moments of random coefficients like $E[A_1 \mid X_1, X_2]$
   - Solve for conditional moments given estimates like $\hat{E}[Y_1 \mid X_1, X_2]$
4. Estimate second period conditional moments like $E[A_2 \mid X_1, X_2]$
5. Estimate unconditional moments like $E[A_1] = E_{X_1, X_2}[E[A_1 \mid X_1, X_2]]$
   - **Requires censoring because of non-identification at $X_1 = X_2$**

Estimator of Unconditional First Moments

- Estimand $E[A_1]$
- Estimator $\hat{\alpha}$ is $\hat{\alpha} = \arg\min_{\alpha \in \mathbb{R}} \hat{G}_n(\alpha)$ where

$$\hat{G}_n(\alpha) = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{E}[A_1 \mid X_{1,i}, X_{2,i}] - \alpha \right)^2 \{|X_{2,i} - X_{1,i}| > h_n\}$$

- $\hat{E}[A_1 \mid X_{1,i}, X_{2,i}]$ is kernel estimator from step 3
- h_n scalar defining **censoring rule**
- Estimator $\hat{\alpha} = \arg\min_{\alpha \in \mathbb{R}} \hat{G}_n(\alpha)$ censored sample average

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} \hat{E}[A_1 \mid X_{1,i}, X_{2,i}] \{|X_{2,i} - X_{1,i}| > h_n\}}{\sum_{i=1}^{n} \{|X_{2,i} - X_{1,i}| > h_n\}}$$

Consistency for Unconditional First Moments

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} \hat{E}\left[A_1 \mid X_{1,i}, X_{2,i}\right]\{|X_{2,i} - X_{1,i}| > h_n\}}{\sum_{i=1}^{n}\{|X_{2,i} - X_{1,i}| > h_n\}}$$

- $E[A_1]$ identified if points $X_1 = X_2$ have probability zero
- Say cross-sectional observations (X_1, X_2, Y_1, Y_2) are i.i.d.

$\hat{\alpha}$ **consistent estimator of** $E[A_1]$ **if** $h_n \propto n^{-a}$ **when** $a < 1/4$**.**

- Uniform convergence under **censoring** main technical step
- Related to Graham & Powell (2012) with weaker regularity conditions

Monte Carlo: Productivity Inspired DGP

- Value added production function
- Constant returns to scale (subtract log labor from log output, log capital)
- Everything multivariate normal
- Autocorrelation in capital intensity X_1 and X_2
- Capital coefficient B_1 correlated with capital intensity X_1
- Correlation between log TFP A_1 and capital elasticity B_1
- High productivity dispersion, $\mathrm{SD}(A_1)$

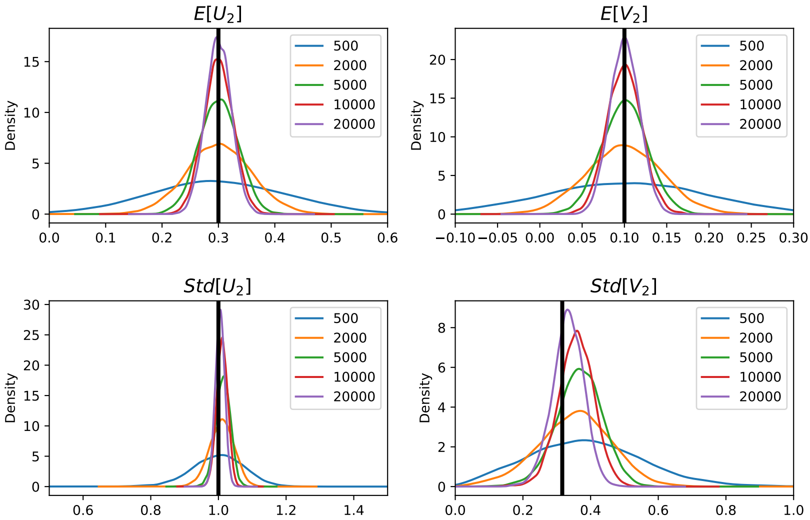
Monte Carlo: Correlation Matrix in DGP

- Correlation matrix (for multivariate normal) for

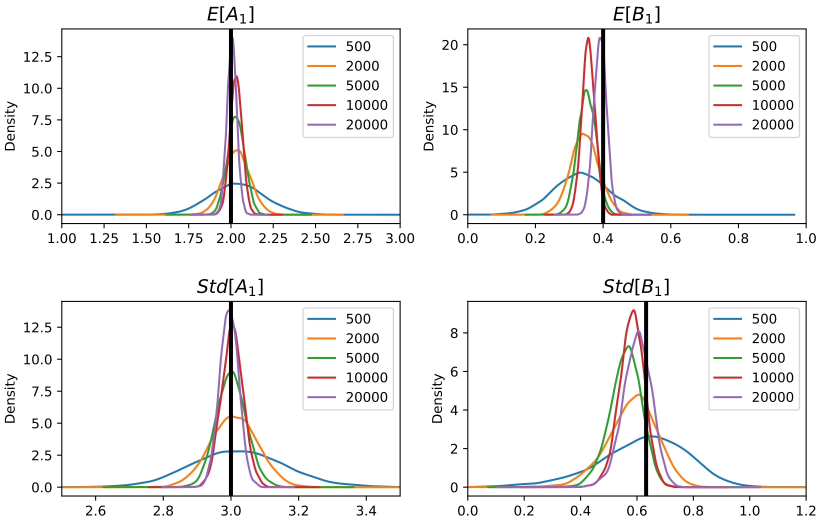
$$(A_1, B_1, X_1, X_2, U_2, V_2)$$

$$\begin{bmatrix} 1 & 0.5 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0,5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 1 \end{bmatrix}$$

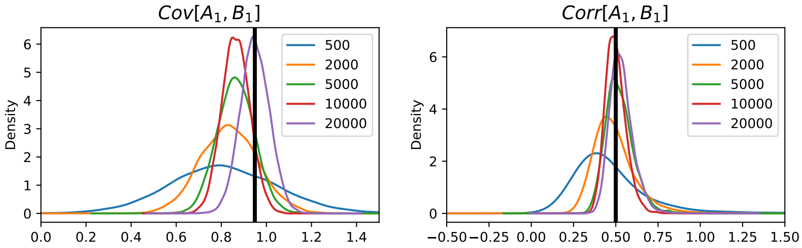
Moments of Innovations U_2, V_2



Unconditional Moments of Random Coefficients



Unconditional Covariance, Correlation of Random Coefficients



Bootstrap Confidence Region Coverage

	$E[A_1]$	$E[B_1]$	$\text{Std}[A_1]$	$\text{Std}[B_1]$	$\text{Cov}[A_1, B_1]$	$\text{Corr}[A_1, B_1]$
0	0.945	0.958	0.951	0.928	0.948	0.945

	$E[A_2]$	$E[B_2]$	$\text{Std}[A_2]$	$\text{Std}[B_2]$	$\text{Cov}[A_2, B_2]$	$\text{Corr}[A_2, B_2]$
0	0.935	0.954	0.951	0.932	0.945	0.941

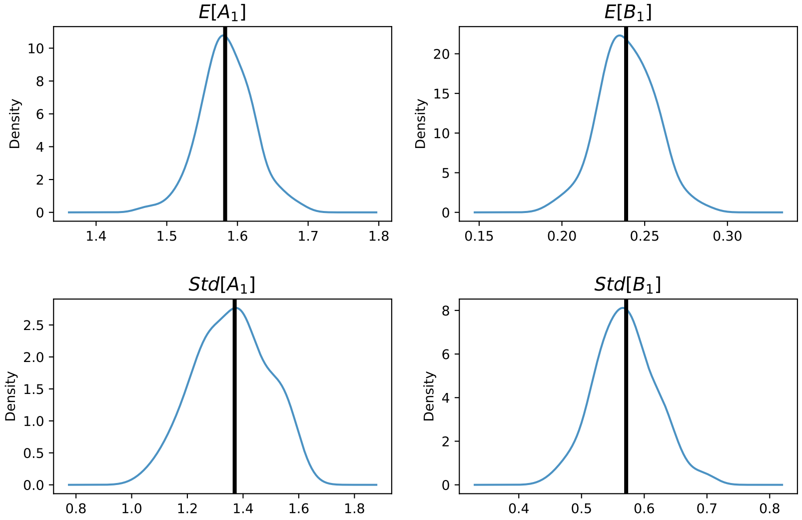
- Nominal size 95%
- 20,000 observations
- No theoretical results on bootstrap validity for unconditional moments (censoring)

India Dataset, Production Function

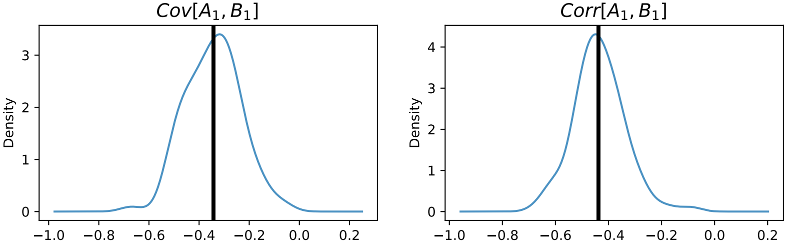
$$\log\left(\frac{Y-M}{L}\right) = A + B\left(\frac{K}{L}\right)$$

- Annual Survey of Industries, 2008 & 2009
- Used in Allcott, Collard-Wexler, & O'Connell (2016)
  - Use their deflators, etc
- Variable definitions
  - Output Y: sales
  - Capital K: value of assets with normal productive life more than a year
  - Labor L: wages
  - Materials M: value of raw materials, components, chemicals, packing material, ...
- After cleaning, 13,297 firms

Bootstrap Distribution, Point Estimate for India



Bootstrap Distribution, Point Estimate for India



Correlated Random Coefficients vs OLS, First Differences

$$\log\left(\frac{Y-M}{L}\right) = A + B\left(\frac{K}{L}\right)$$

	$E[A]$	$\mathrm{Std}[A]$	$E[B]$	$\mathrm{Std}[B]$	$\mathrm{Corr}[A,B]$
OLS, 2008	1.50	**0.99**	0.31	–	–
OLS, 2009	1.47	**1.00**	0.32	–	–
First Differences, OLS	–	–	0.28	–	–
Random Coefficients, 2008	1.58	**1.37**	0.24	0.57	-0.35
Random Coefficients, 2009	1.59	**1.54**	0.24	0.57	-0.31

- **Standard deviation of log TFP increases**
- Productivity dispersion with heterogeneous production function higher than OLS indicates

Aggregate Productivity Growth (APG)

- **Aggregate productivity growth** (APG) **defined** as
  - Change in aggregate final demand *minus*
  - Change in aggregate expenditure on inputs
- APG formula has, per plant
  - One term related to production function
  - One term related to allocation of inputs to heterogeneous firms
- APG allows for real-life features: varying input prices, market power in output markets,. . . .
- APG equal to change in consumer welfare under baseline assumptions
- Extend APG in Petrin & Levinsohn (2012) to **vector-valued productivity**

APG: Bayes Rule Density as Analog to Residuals

$$\log\left(\frac{Y-M}{L}\right) = A + B\left(\frac{K}{L}\right)$$

- APG definition does not incorporate econometrician's uncertainty
  - OLS, **residual** estimate $\hat{A}$ of log TFP A
  - Plug in $\hat{A}$ for A in APG definition for point estimate of APG
- **Vector-valued plant heterogeneity**
  - Straight line of A, B satisfy equation per firm
  - Bayes rule: **marginal density of** A over this line per firm
  - Still have density for each firm as number of firms in data grows large
- Extending APG to vector-valued heterogeneity now

Conclusions

- Random coefficients
  - Consumer panels with heterogeneous consumers
  - Production function parameters (A, B) vary across firms in same industry
- Random coefficients (A, B) can be **correlated with inputs** (X_1, X_2)
- Estimator allows **arbitrary correlations** of production function parameters (A_1, B_1) with inputs (X_1, X_2)
- **Timing assumptions** with **random coefficients** (A, B)
  - Input decisions made in period $t - 1$ with knowledge of $t - 1$ production function
- Estimation, consistency: censoring for unconditional means
- India: log TFP dispersion increases in random coefficients vs OLS