Econometrics of Organizational Economics

John Van Reenen

Organizational Economics, Spring 2020

ORGANIZATION

- Syllabus and Reading List
- Recitations

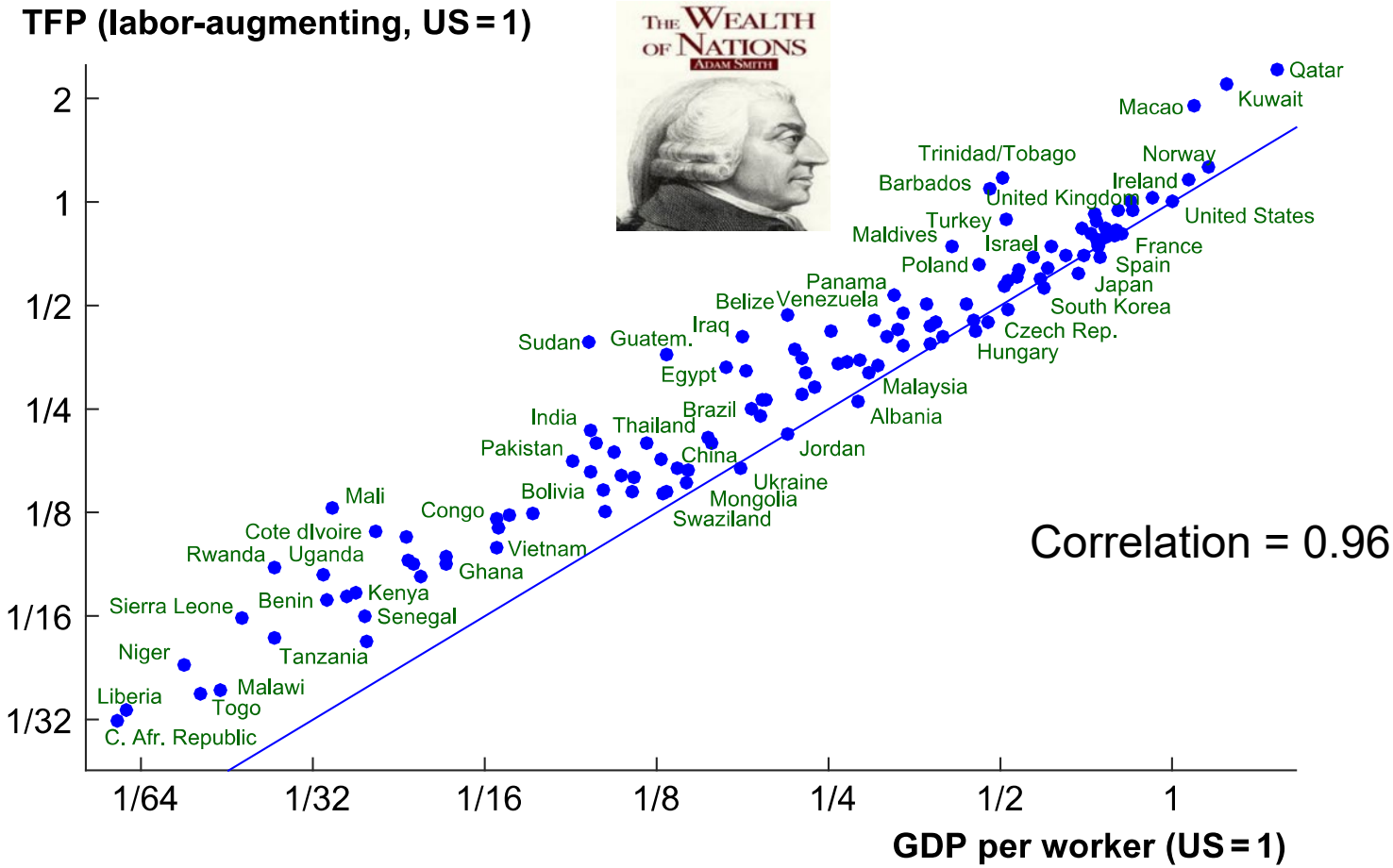
SOME KEY QUESTIONS IN ECONOMETRICS OF ORGANIZATIONAL ECONOMICS

1. What is causal impact of different forms of organizational practices on performance?
   - **Example:** What is effect of incentive pay on productivity (Lazear, 2000)
   - Identification Revolution in econometrics (Angrist & Pischke, 2015): RCTs; RDD, Diff in Diffs; IV, etc.
2. What determines the distribution of organizational practices in an economy?
   - e.g. Why some firms adopt incentive pay & others don't
   - More a "between firm" question (rather than "within firm")
3. Can the answers to 1. & 2. explain heterogeneity of performance across firms and countries?
   - Both an accounting & an equilibrium question
4. "Reduced form" & "Structural" approaches complements

AGENDA FOR MY LECTURES

1. Facts: Focus on Persistent Productivity Differences ("PPDs") across countries & firms
2. Managerial explanations of heterogeneity
   - Top managers vs. management practices
   - Measurement of management practices
   - Impact of management practices
   - Determinants of management practices
3. Explaining Macro with micro: Productivity & management
4. Decentralization and Power
5. Knowledge hierarchies
6. Organizational Complementarities
7. Inequalities
   - Top pay: CEOs
   - Between labor and capital
   - Between workers

Big spread of productivity between countries



Notes: 2010 data; Total Factor Productivity (α=1/3);
Source: Penn World Tables 8.0; Jones (2015)

DEVELOPMENT ACCOUNTING

Tries to decompose GDP per worker differences across countries into different factor inputs

- Caselli (2005, 2016)
  - Physical, natural and human capital (including health capital and school quality via test scores).
  - Alternative production functions; substitution assumptions, etc.
  - Large (most?) unexplained TFP residual between countries

- Gennaioli et al (2013)
  - Uses sub-national data. Important role for highly skilled entrepreneurs/managers

Micro Facts on Firms

Firm heterogeneity

- **General issues**
- Firm Size Distribution
- Productivity Dispersion

Reallocation & aggregate productivity

FIRM HETEROGENEITY HAS LONG BEEN RECOGNIZED

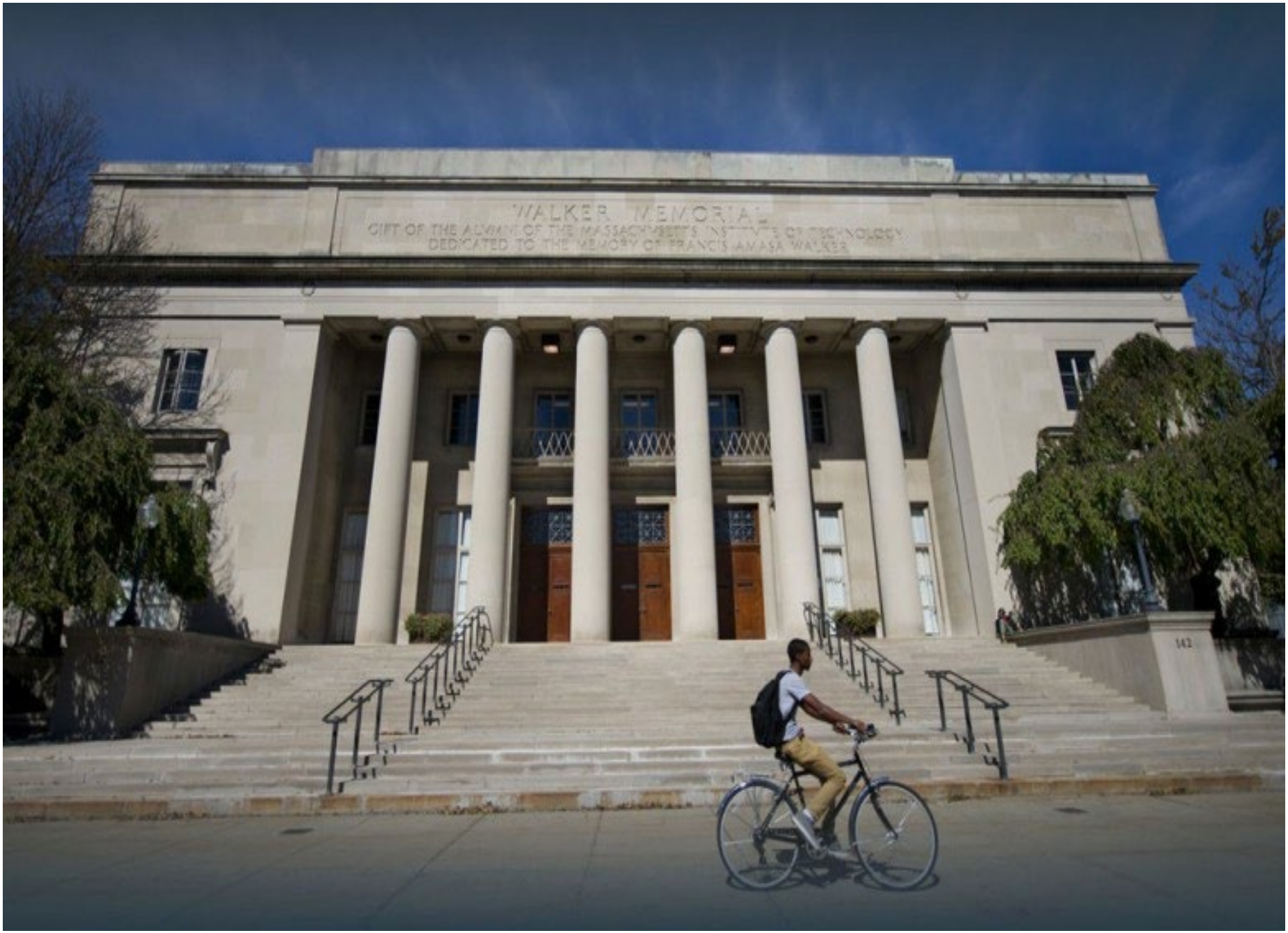
"…we have the phenomenon in every community and in every trade, in whatever state of the market, of some employers realizing no profits at all, while others are making fair profits; others, again, large profits; others, still, colossal profits."

Francis Walker (*Quarterly Journal of Economics*, '87)

FIRM HETEROGENEITY HAS LONG BEEN RECOGNIZED

enon in every community and in
r state of the market, of some
profits at all, while others are
ers, again, large profits; others,

***arterly Journal of Economics*, 1887)**


WALKER MEMORIAL
142

Some dimensions of firm heterogeneity

- **Profits & markups**
  - Schmalensee (1985); de Loecker et al (2019)

- **Firm Size** (Gibrat, 1932; this lecture)

- **Productivity** (this lecture)

- **Wages** (this lecture & next)

- **Management Practices** more widely (next lecture)

Firm heterogeneity

- General issues
- **Firm Size Distribution**
- Productivity Dispersion

Turnover & Lumpiness

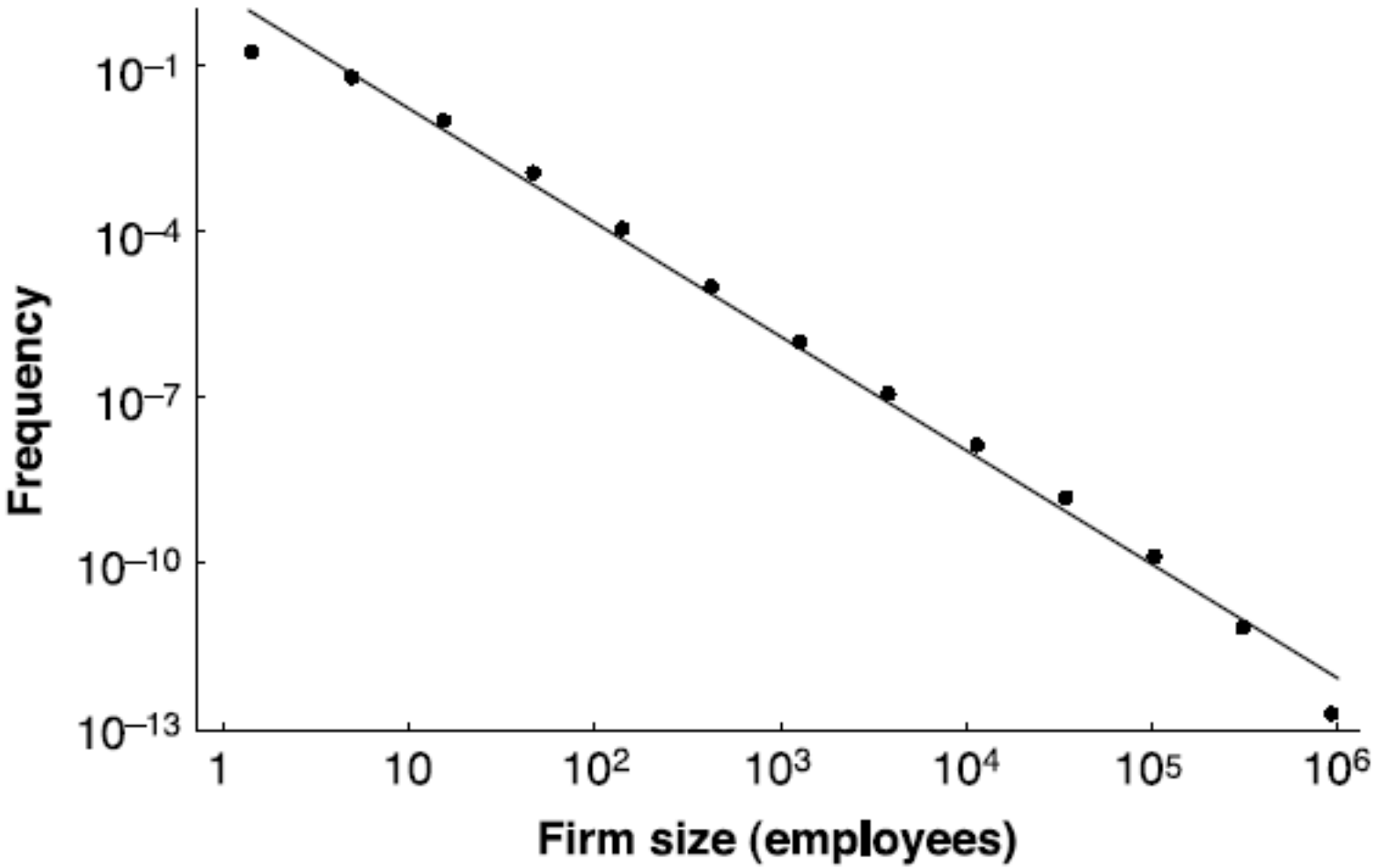
Reallocation & aggregate productivity

Big firms account for large fraction of activity (e.g. over a third of all US employees in ~2k biggest firms)

Size class	Share firms	Share jobs	#Firms	employees
<10 workers	78.91%	10.33%	4,597,138	12,511,024
10 to 4,999 workers	21.06%	55.7%	1,226,334	67,430,297
5,000+ workers	0.03%	33.97%	1,986	41,128,623
Total	100.00	100.00	5,825,458	121,069,944

Source: US SBA, 2014 data
https://www.sba.gov/advocacy/firm-size-data

Firm Size distribution from Census data (mainly) follows a power law



Source: Robert L. Axtell, *Science* (2001), US Economic Census, 1997

Firm heterogeneity

- General issues
- Firm Size Distribution
- **Productivity Dispersion**

Turnover & Lumpiness

Reallocation & aggregate productivity

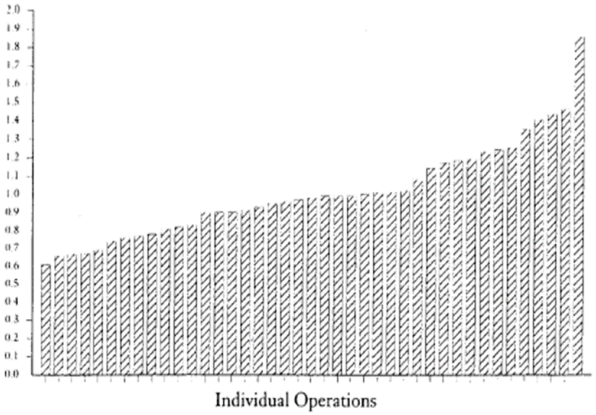
Productivity Heterogeneity a classic issue

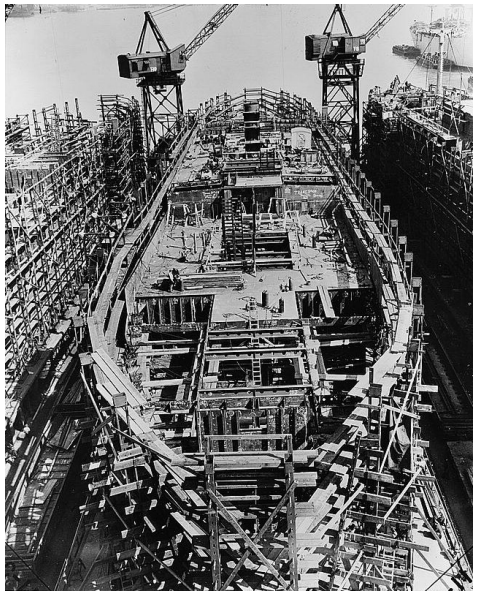
- **Salter (1960)** 1911-26 UK pig-iron industry: best factory twice as much output per worker as average factory

- **Argote et al (1990)** US Liberty ships in WW2 across 16 separate shipyards. Focuses on learning by doing but also big yard-specific effects

- **Chew et al (1990)** 40 operating units in commercial food division of large US corporation. Top ranked unit twice as productive as bottom ranked (within firm)

Figure 5-1 Multifactor Productivity Index





MEASURING PRODUCTIVITY

1. **Residual "accounting approach"**
2. Estimating Production Function
   - IV Approach with "External instruments"
   - Fixed effects (FE)
   - IV+FE (Arellano & Bond, 1992; Blundell & Bond, 2000)
   - Olley Pakes (1996) + extensions (e.g. Levinsohn & Petrin, 2003; Ackerberg, Caves & Frazer, 2015; de Loecker et al, 2016, 2019)

Measuring firm or establishment productivity

Simplest way is Labor Productivity (LP): value added (Q) of establishment *i* at time *t* normalized on labor inputs

$$\ln(LP_{it}) = \ln(Q_{it}) - \ln(L_{it})$$

Ideally L = quality-adjusted hours; although often only number of workers available (or wage bill)

Value Added = Sales – intermediate inputs (Y – M). Micro equivalent of GDP. Deflated by (usually industry level) price index

Illustrate for Cobb-Douglas, but issues generalize to CES, Translog, etc,

$$Y = AF(K, L_t) = AL^{\alpha_L} K^{\alpha_K}$$

$$\ln A = \ln Y - \alpha_L \ln L - \alpha_K \ln K$$

$$\Pi = PY - wL - rK$$

Assuming perfect competition in all output & factor markets

$$\alpha_L = WL / PY \equiv s_L ; \alpha_K = rK / PY \equiv s_K$$

Technological parameters are replaced by factor shares in revenue

Measuring productivity

"Two factor" TFP:

$$\ln A_{it} = \ln Y_{it} - s_L \ln L_{it} - s_K \ln K_{it}$$

Easily generalizable to more factors, e.g. "3 factor" TFP:

$$\ln A_{it} = \ln Y_{it} - s_L \ln L_{it} - s_K \ln K_{it} - s_M \ln M_{it}$$

where M = intermediate inputs

If constant returns to scale (CRTS) 2 factor TFP is:

$$\ln A = \ln Y - \boldsymbol{s}_L lnL - (1 - \boldsymbol{s}_L)\ln K$$

Problems with accounting approach

- Many problems with this way of measuring productivity

- Strong assumptions over the functioning of markets (e.g. perfect competition in output & input markets) and no adjustment costs
  - Some assumptions can be partially relaxed (e.g. if imperfect product market competition use <u>cost shares</u> as revenue shares partially reflects markup)
  - Main alternative is to econometrically estimate the parameters of production function

- But a transparent & simple approach (the broad facts I'll describe also emerge in more complex estimation method). Very popular so lots of papers to draw on

Productivity Heterogeneity: basic facts

- Typical gap between 10th and 90th percentiles within same US four digit industry (Syverson, 2004, 2011)
  - Labor Productivity (output per worker) 4:1 ratio
  - Total Factor Productivity 2:1 ratio

- Important to look within industry for TFP as “apples & oranges” problem across industries with different production functions

- Productivity dispersion generally larger in other countries
  - Hsieh & Klenow (2009) China and India 5:1 ratio
  - Bartelsman, Haltiwanger & Scarpetta (2013) 9 OECD countries.
  - And now replications in very many countries: OECD Multiprod data initiative (Criscuolo et al, 2016); ECB CompuNet; World Bank, etc.

Performance Differentials are Persistent

- If productivity differences completely transitory, may be of less interest

- Not completely persistent as there is turnover. But this does not happen immediately (competition works at a distance)
  - Transition matrix of TFP over a 5 year period (Bailey et al, 1993): 2/3 of plants in top TFP quintile were in top quintile 5 years later

- Basis of idea of firm “capabilities” (fundamental to Strategy). Penrose (1959); Winter (1988); Nelson (1991); Teece et al, (1997); Henderson & Clark (2007)

- Classical measurement error means we exaggerate transitory component (but also exaggerate cross sectional spread)

Are TFP differences "just measurement error?"

- **Parallels old debate in macro growth accounting**
  - Solow (1957) vs Griliches & Jorgensen (1967)
- **It's not all measurement error:** as measured TFP correlated with future survival & growth as dynamic models suggest
- **Some types of measurement error**
  1. *Missing input quantities:* e.g. intangible capital (like IP)
  2. *Missing output prices* (focus of literature)
  3. *Heterogeneous* inputs: (i) factor "prices" could reflect quality input shifts (e.g. labor); (ii) incorrect aggregation
  4. *Heterogeneous* outputs: multi-product plants
  5. *Other issues*
     - Recording errors (exaggerate spread). ~50% of US Census TFP measurement error (Collard-Wexler, 2003)
     - Imputation implies we underestimate spread (e.g. White, Reiter & Petrin, 2017, US Census in 2007 73% of obs have at least one TFP element imputed)

Mis-measured Inputs

- **Labor Services**: Not simply #hours, but also in efficiency units. Take into account education, demographics, etc.
  - Use wage bill instead of #workers, weight each worker by their wage (simple but assumes perfect substitutability)

- **Capital services**
  - Capital stock tricky to measure: dynamic not static
  - As with labor have to aggregate over heterogeneous types of capital (structures, buildings, etc.) & weight using imperfect substitutability
  - Historical accounting values vs. replacement value (usually via Perpetual Inventory Method)

- **Intermediate inputs**
  - These are often disguised “Intangible Capital” (R&D, software, IP, advertising, management, etc.)

Measurement of output prices

- Revenue based TFP ("TFPR") vs quantity based TFP ("TFPQ"). Problem is that conventional TFP mixes up firm-specific price-cost margins with efficiency
- In almost all micro datasets we do not observe quantity, but revenue (= output price x quantity)
- Researchers deflate revenues by industry output price deflator (P_I), equivalently express in within-industry deviations
- But if prices vary within industry ($P_i \neq P_I$), a high price firm (e.g. with market power) looks like it is high TFP
- What we measure in data is real revenue/sales (R) not quantity (Q). $\ln R_i - \ln P_I \equiv (\ln Q_i + \ln P_i) - \ln P_I$
- Changes to environment that raise TFPR by increasing market power & so reduce consumer welfare might look like they are raising efficiency (TFPQ), giving opposite policy conclusions!

PRODUCTIVITY OR MARK-UPS?

- Q = output, but actually revenues deflated with industry prices: firm prices unobserved (use PPIs) .
  - Estimated coefficients mix tech parameters & markups
- **Solutions:**
  - Get better data on firm prices (Foster, Haltiwanger and Syverson, 2008)
  - Make explicit demand side & jointly estimate mark-ups.
- **Example:** Klette & Griliches (1996) monopolistic competition: firm specific demand function:

$$\mathbf{q_i = -\eta(p_i - p_I) + d_I}$$

- $\mathbf{p_i}$ firm-specific log price (unobservable), $\mathbf{p_I}$ is industry log price (observable), $\mathbf{d_I}$ = industry demand shifters, $\boldsymbol{\eta}$ = the elasticity of demand.

PRODUCTIVITY OR MARK-UPS?

- Solving for firm prices $p_i = p_I - (q_i - d_I)/\eta$
- What we measure in data is real sales (r) not quantity (q)
- $r_i - p_I \equiv (q_i + p_i) - p_I$
- Cobb-Douglas: $q_i = a_i + \alpha_L l_i + \alpha_K k_i$, real sales equation is:

$$\boldsymbol{r_i = \left(1 - \frac{1}{\eta}\right)(a_i + \alpha_L l_i + \alpha_K k_i) + \frac{1}{\eta} d_I}$$

- $1/\eta$ is the mark-up of price over marginal cost (as η gets large mark-up = 0). Factor input coefficients mix mark-ups with technological coefficients
- If we have demand shifters (e.g. Industry output) d_I then can get elasticity (η) & separate from productivity
- De Loecker (2011, ECMA) generalizes this to OP set-up (multiproduct firms creates some extra variation)

Examples in Foster et al (2008)

Result:

- TFPQ spread is as big in these sectors as in the TFPR spread in other sectors
- Positive TFPQ,TFPR correlation. But underestimate entrant TFPQ

TFP Spread in hospitals (case-mix adjusted survival rates From AMI) <u>and</u> concrete plants

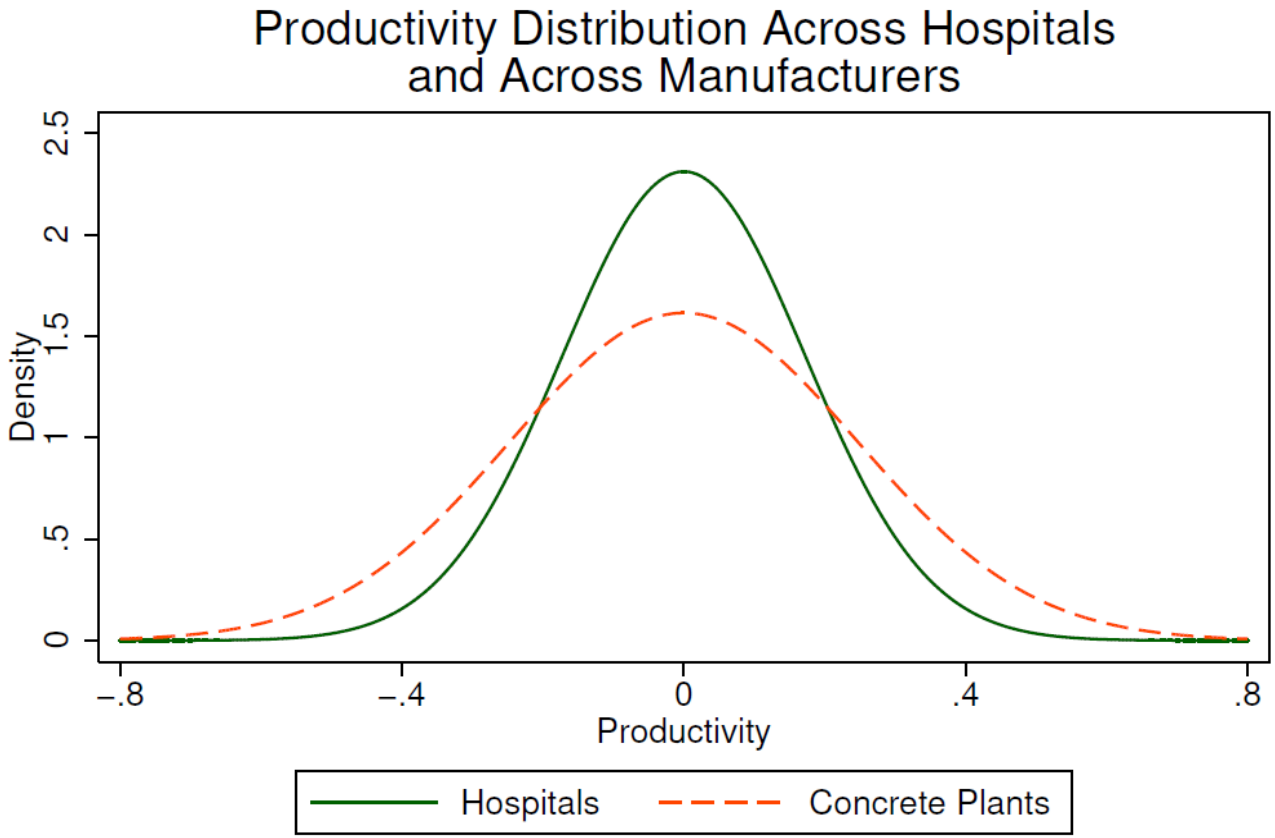
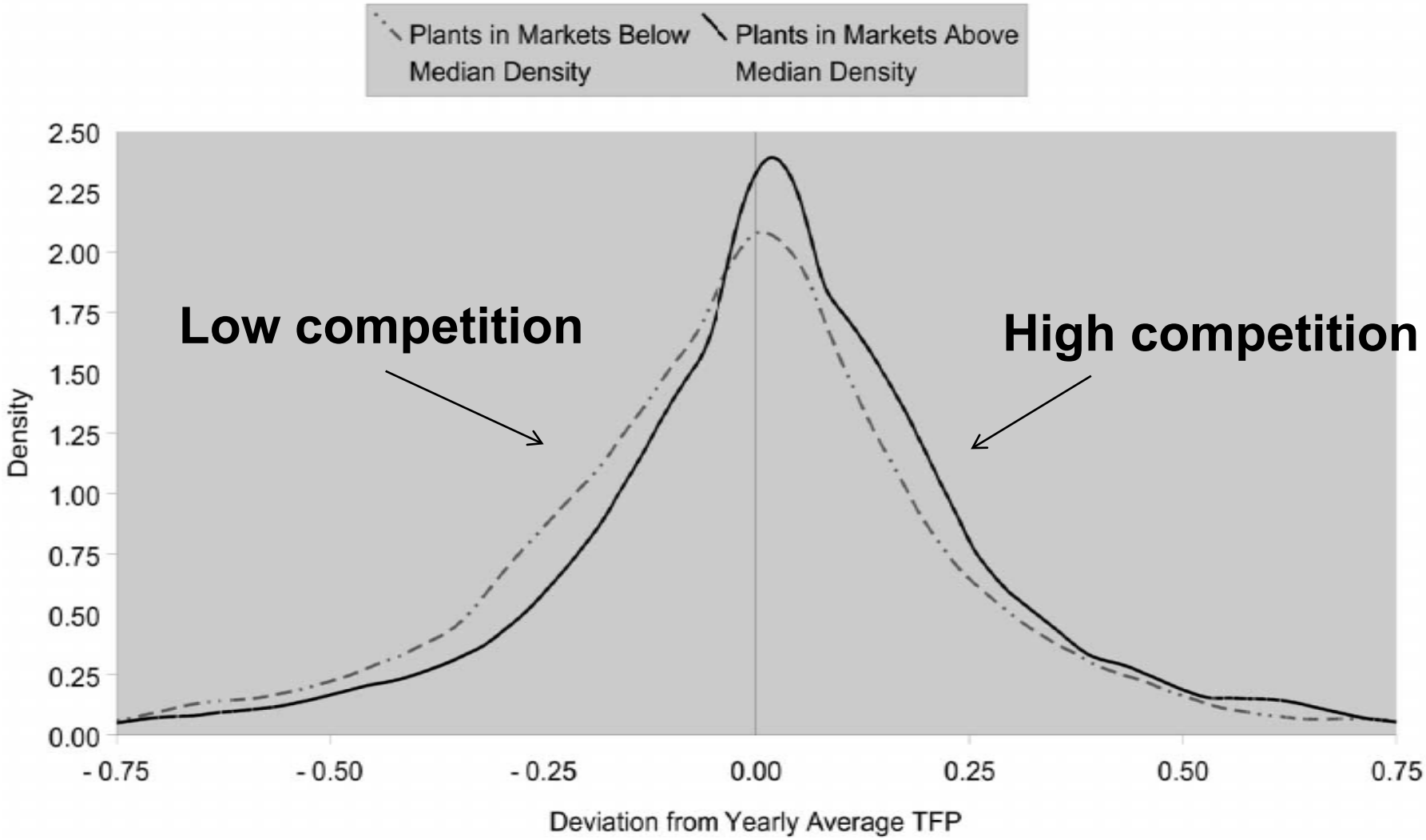


Figure shows estimated productivity dispersion across hospitals for heart attack treatments and across concrete plants for the production of ready–mixed concrete. We show the average within–year fitted normal density for each. Hospital productivity estimates (which reflect the hospital's ability to produce patient survival given a fixed set of inputs), are from our baseline specification (Table 2, column 1); concrete productivity estimates are from Table A7. See text for more details on the construction of these estimates.

Source: Chandra, Finkelstein, Sacarny & Syverson (2016, AER)

Competition as an explanation for heterogeneity: Ready mix concrete



Source: Syverson (2004)

Firm heterogeneity

- General issues
- Firm Size Distribution
- Productivity Dispersion

Reallocation & aggregate productivity

Defining industry (or aggregate) productivity

Define a simple industry productivity index: P_t

$$PROD_t = \sum s_{it}\omega_{it}$$

Where:

ω_{it} is the ln(productivity) of establishment *i* in period *t* (e.g. ln(labor productivity) or ln(TFP))

s_{it} is the size share of establishment *i* in the industry in period *t* (e.g. the share of establishment value added in industry value added)

Industry productivity can increase through two channels

- **Within Firms (Traditional view)**
  - The same firms become more productive (e.g. new technology spreads quickly to all firms, like Internet)

- **Between Firms (Schumpeterian view)**
  - Low TFP firms exit and resources are reallocated to high TFP firms
    - **Intensive margin**: High TFP firms expand (e.g. more jobs) & low TFP firms contract (e.g. less jobs)
    - **Extensive margin**: Exit/entry

These two effects are well known to cricket fans

Within batsman (each batsman improves)

Between batsman (more time for your best batsman)

Shift-Share Decompositions of productivity

Productivity growth for a balanced panel of establishments can be broken down into three terms:

$$
\begin{aligned}
PROD_t - PROD_{t-1} &= \sum s_{i,t}\omega_{i,t} - \sum s_{i,t-1}\omega_{i,t-1} \\
&= \sum s_{i,t-1}(\omega_{i,t} - \omega_{i,t-1}) \text{ Within term} \\
&\quad + \sum (s_{i,t} - s_{i,t-1})\omega_{i,t} \text{ Between term}
\end{aligned}
$$

"Reallocation"

Within term is focus of representative agent models

Decomposing productivity

Allowing for <u>entry and exit</u> requires two more terms:

$$
\begin{aligned}
PROD_t - PROD_{t-1} = & \sum_{i \in Survivors} s_{i,t-1}(\omega_{i,t} - \omega_{i,t-1}) && \text{Within incumbent term} \\
& + \sum_{i \in Survivors} (s_{i,t} - s_{i,t-1})\omega_{i,t} && \text{Between incumbent term} \\
& + \sum_{i \in Entry} s_{i,t}^{Entry} \omega_{i,t}^{Entry} && \text{Entry term} \\
& - \sum_{i \in Exit} s_{i,t-1}^{Exit} \omega_{i,t-1}^{Exit} && \text{Exit term}
\end{aligned}
$$

This is the original Bailey, Hulten and Campbell (1992) decomposition. Used Census data 1972-1987

Decomposing productivity

Foster, Haltiwanger & Krizan (2001) similar, but allow cross term & use average productivity as reference group

$$PROD_t - PROD_{t-1} = \sum_{i \in Survivors} s_{i,t-1}(\tilde{\omega}_{i,t} - \tilde{\omega}_{i,t-1}) \quad \text{Within term}$$

"reallocation"

$$+ \sum_{i \in Survivors} (s_{i,t} - s_{i,t-1})\tilde{\omega}_{i,t-1} \quad \text{Between term}$$

$$+ \sum_{i \in Survivors} (s_{i,t} - s_{i,t-1})(\tilde{\omega}_{i,t} - \tilde{\omega}_{i,t-1}) \quad \text{Cross term}$$

"net entry"

$$+ \sum_{i \in Entry} s_{i,t}^{Entry} \tilde{\omega}_{i,t}^{Entry} \quad \text{Entry term}$$

$$- \sum_{i \in Exit} s_{i,t-1}^{Exit} \tilde{\omega}_{i,t-1}^{Exit} \quad \text{Exit term}$$

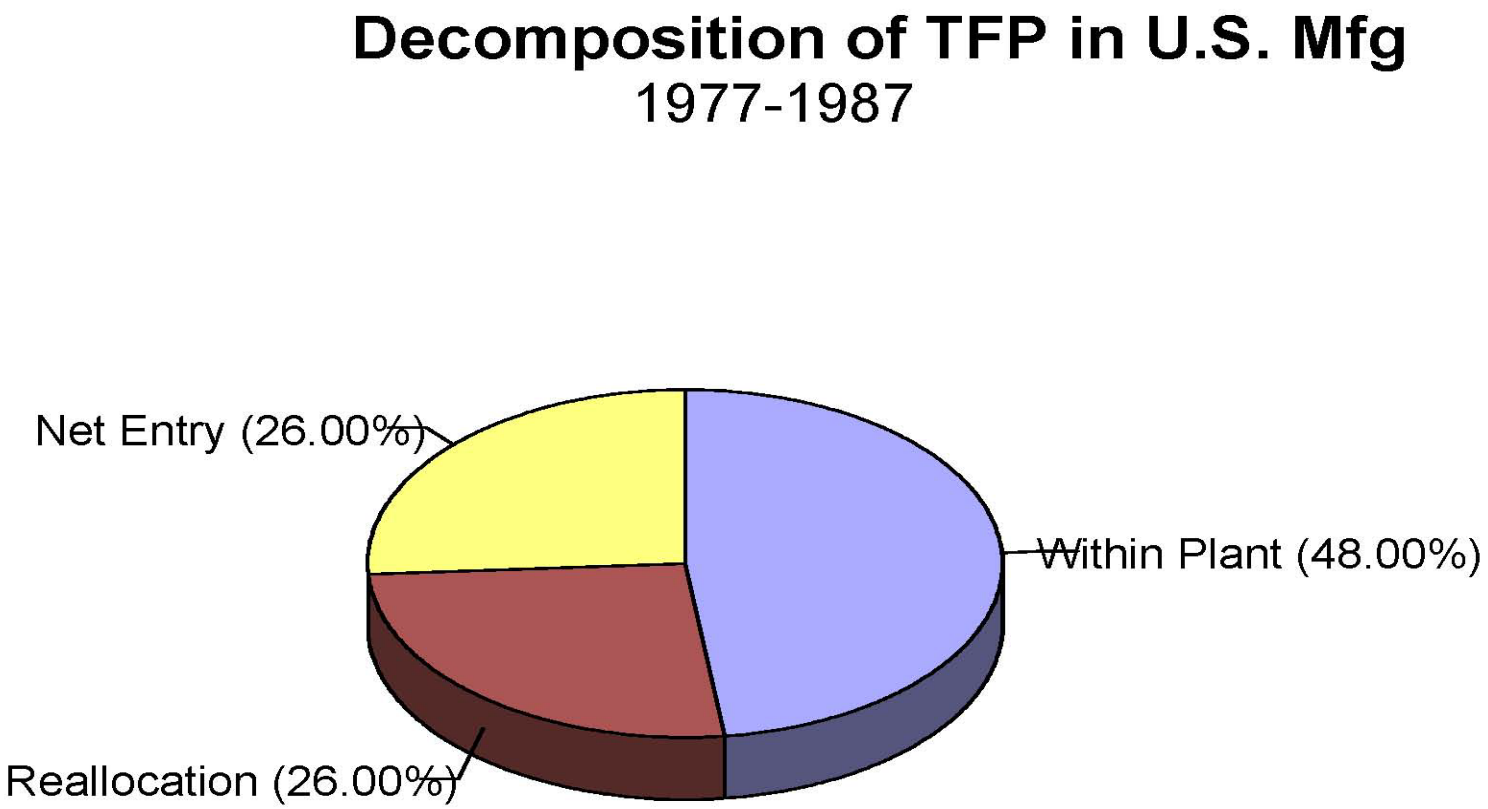
$$\tilde{\omega}_{i,t} = \omega_{i,t} - \bar{\omega}; \bar{\omega} = (PROD_t + PROD_{t-1}) / 2$$

Decomposing productivity

- Many other ways of decomposing productivity
- Foster, Haltiwanger & Krizan (2001) similar, but allow cross term & use average productivity as reference group
- Melitz & Polanec (2015, RAND) has a good survey of these different methods with pros & cons
  - They also introduce a method that extends Olley & Pakes (1996) decomposition method to allow for entry/exit

- Qualitative conclusions below do not depend on details of the decompositions

Total reallocation (between, entry and exit) accounts for about ½ of manufacturing TFP growth



Source: John Haltiwanger; same conclusion as Bailey et al (1993)

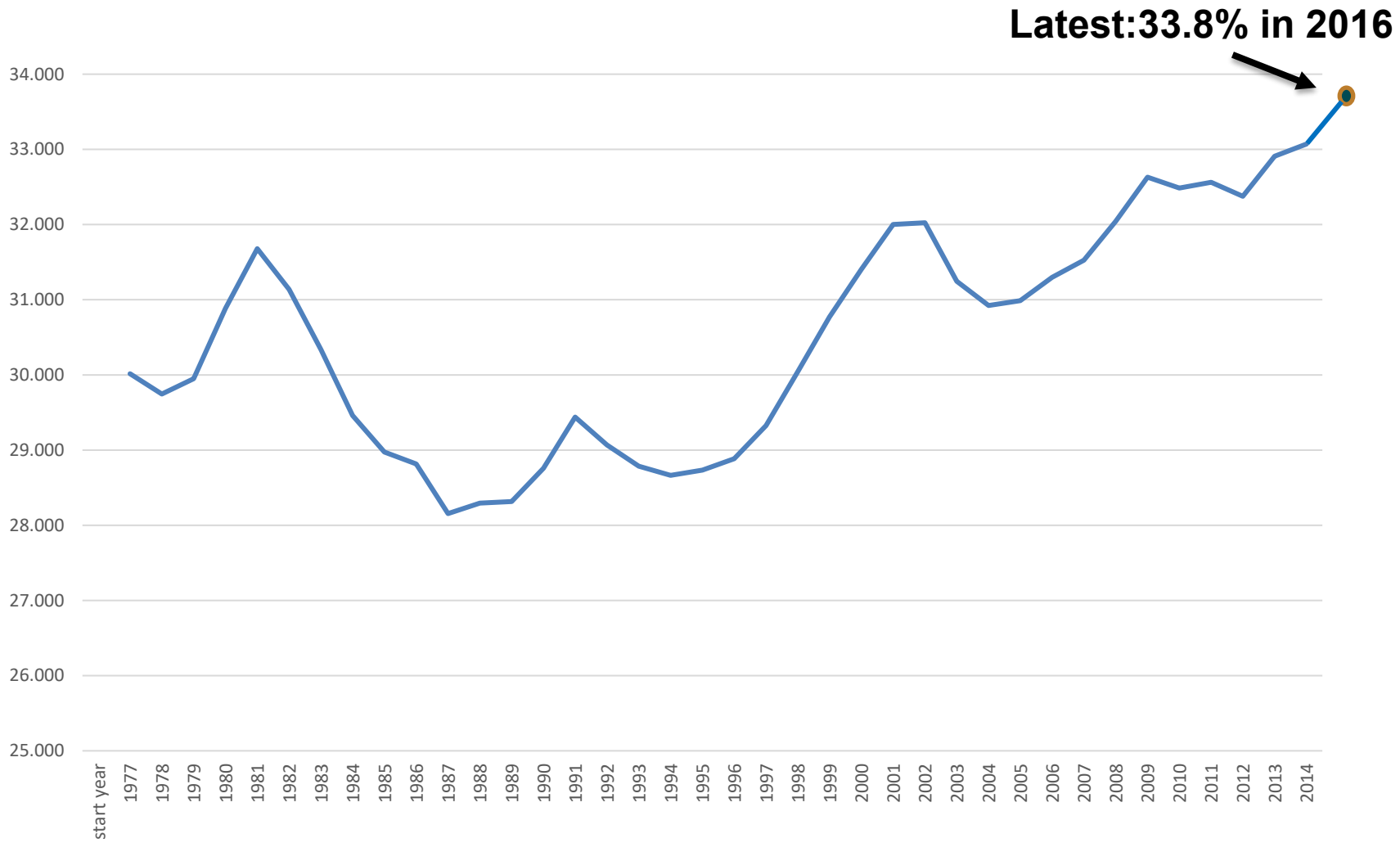
Recent Trends

- Evidence that firms appear increasingly different on a number of dimensions in US & overseas
  - Size (Autor et al, 2017, 2019 on growth of “superstar” firms: return to this in last lecture)
  - Productivity (OECD: Andrews et al, 2014)
  - Wages (Song et al, 2019)

- Greater heterogeneity between firms mirrors increase inequality between individuals

- Many indicators suggest declining US business dynamism
  - Decker et al (2016)

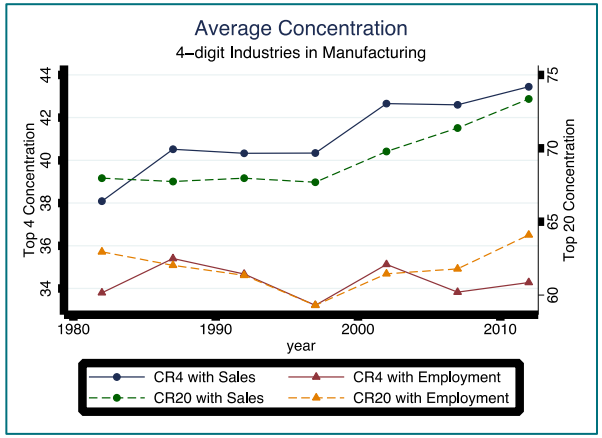
Since mid '80s Big Firms getting bigger: % jobs in firms with 5,000+ workers rose from 28% in 1987 to 34% in 2016



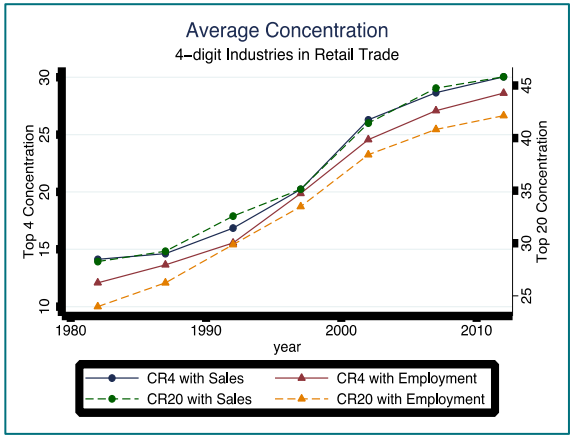
Source: SBA, https://www.sba.gov/advocacy/firm-size-data#susb

Rising Concentration in US since 1982

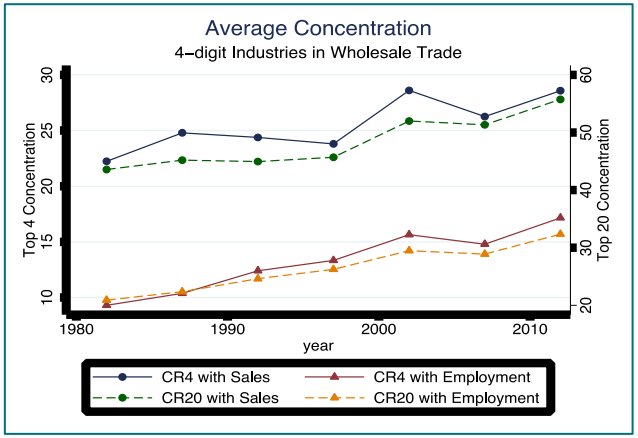
Manufacturing



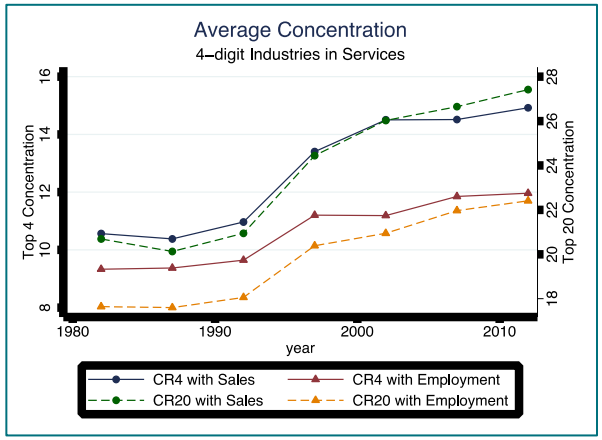
Retail Trade



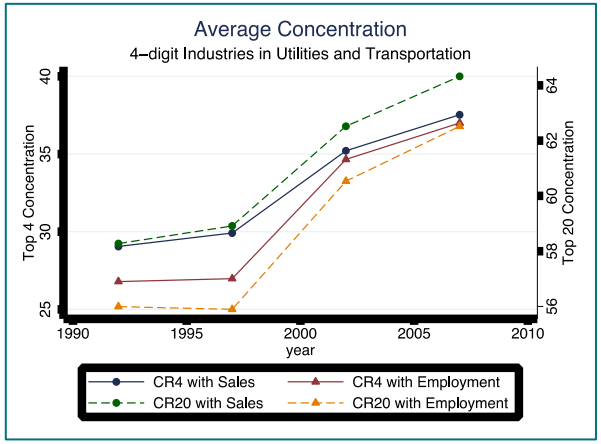
Wholesale Trade



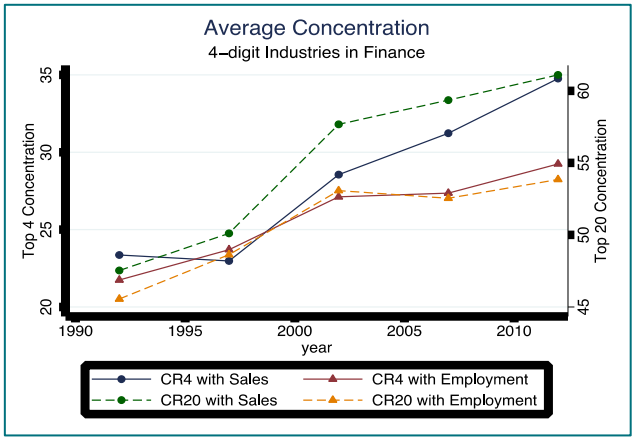
Services



Utilities + Transportation

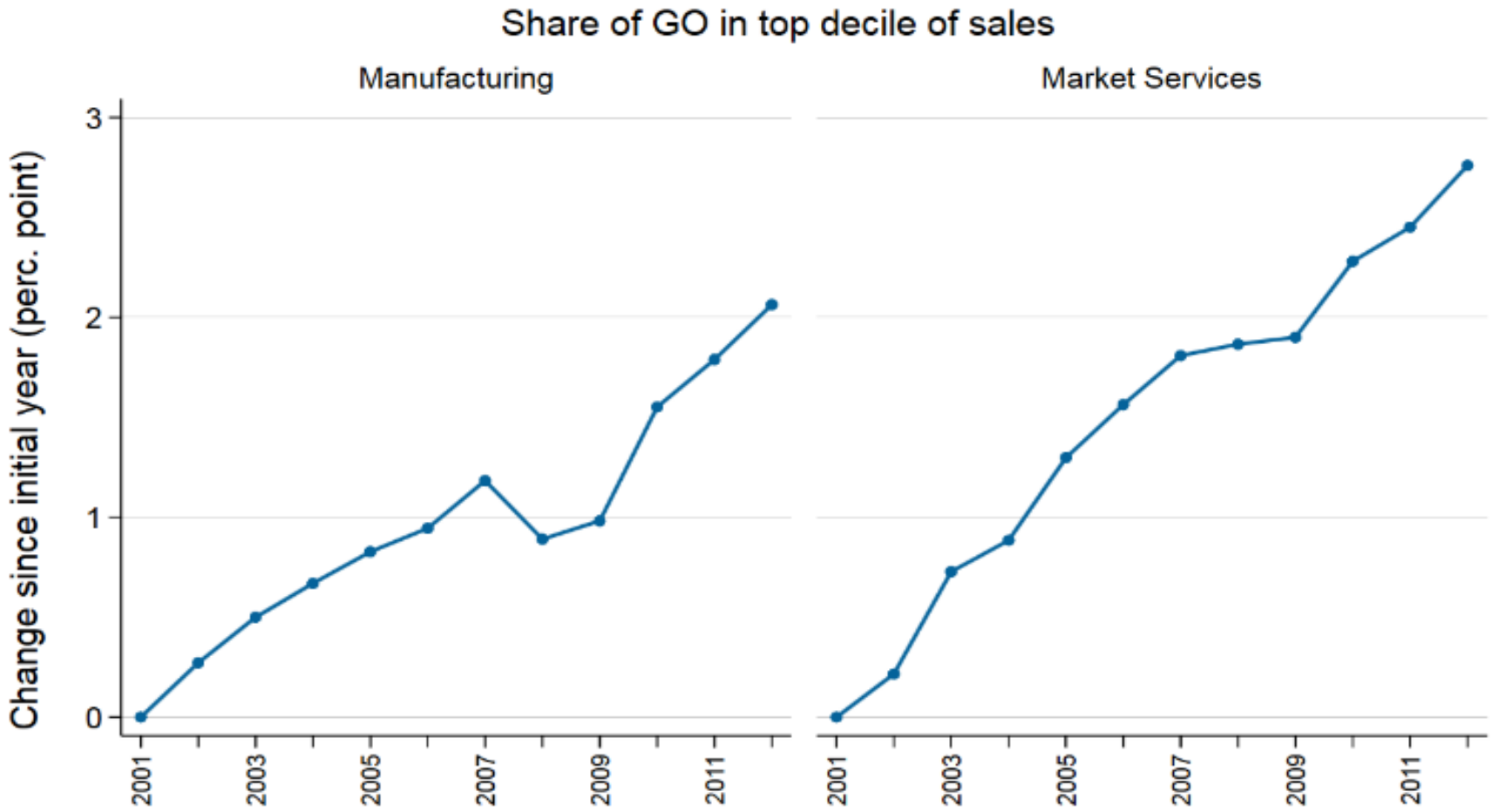


Finance



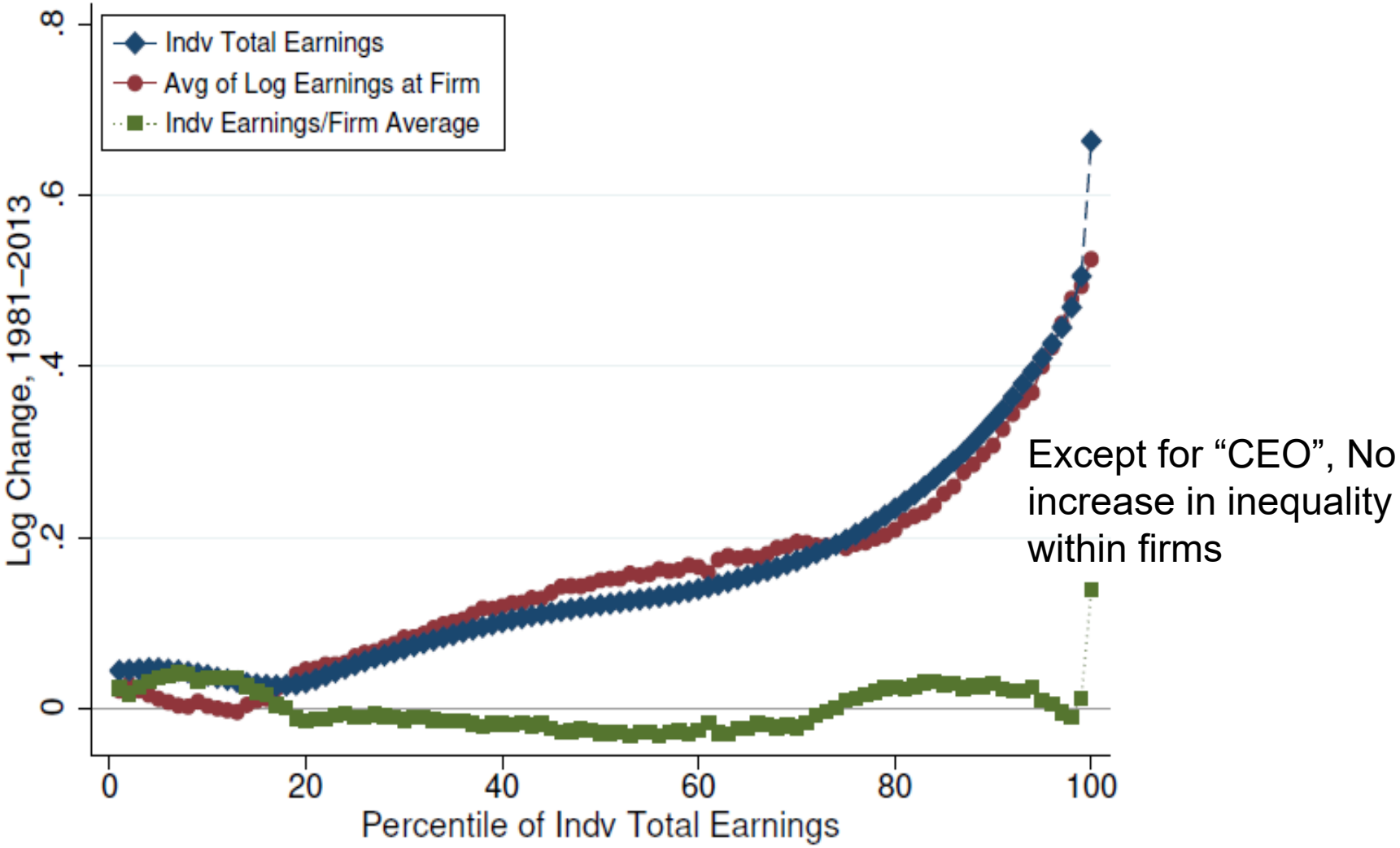
Notes: Autor et al (2019) from Economic Census; Weighted av. of concentration across the SIC-4's within each sector. Manufacturing:388 inds; Retail:58; Services:95; Utilities/Transportation:48; Wholesale:56; Finance:31

Like US, Sales Concentration seems to have has also increased in the EU



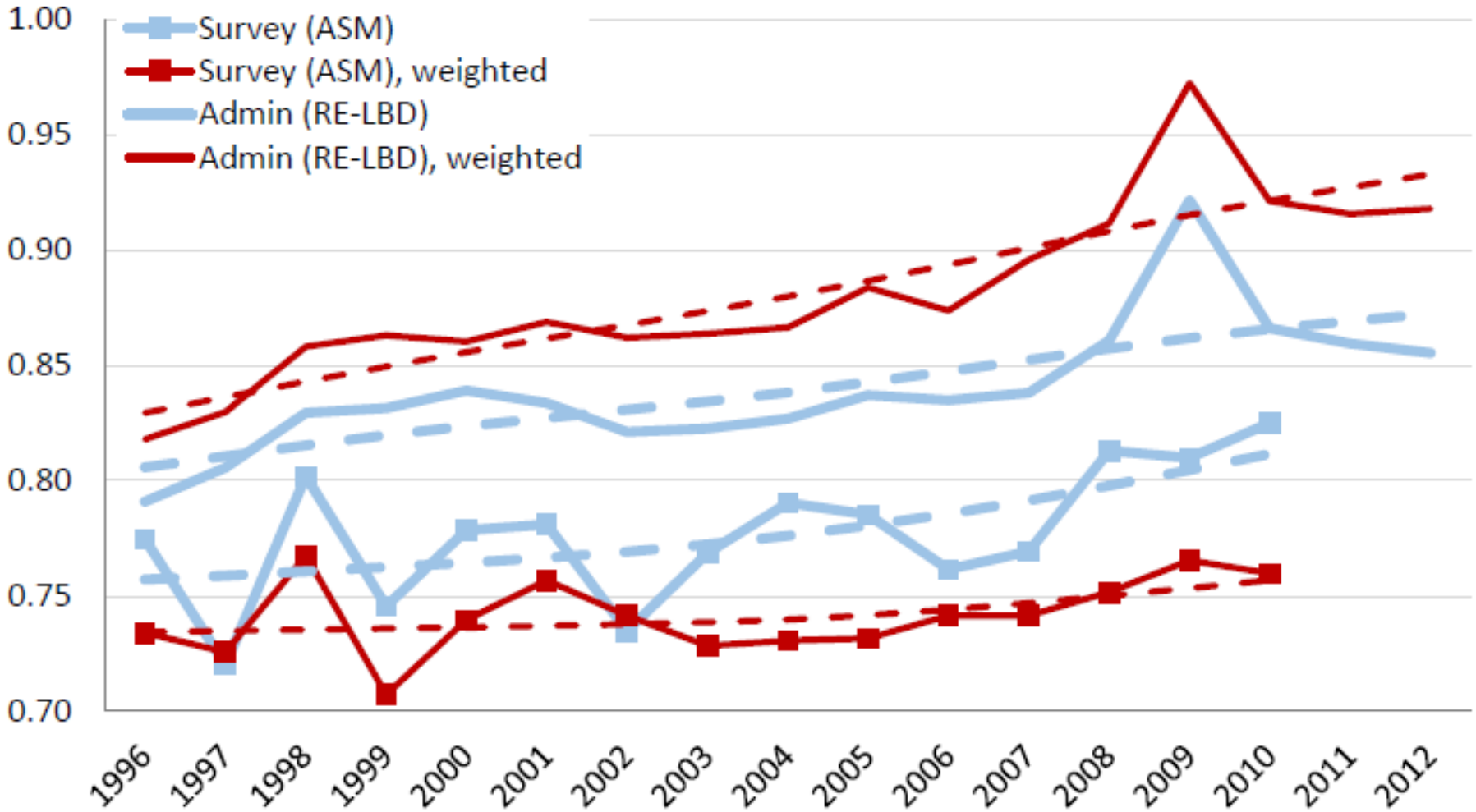
Source: OECD Multiprod, https://www.oecd.org/sti/ind/multiprod.htm; Criscuolo (2018)
Notes: Year effects from regressions with country-industry dummies and year dummies (BEL, DEU, DNK, FIN, FRA, HUN,NOR, PRT, SWE)

Change in individual US earnings inequality is a between firm (rather than within firm) phenomena, 1981-2013



Source: Song et al (2019), SSA data

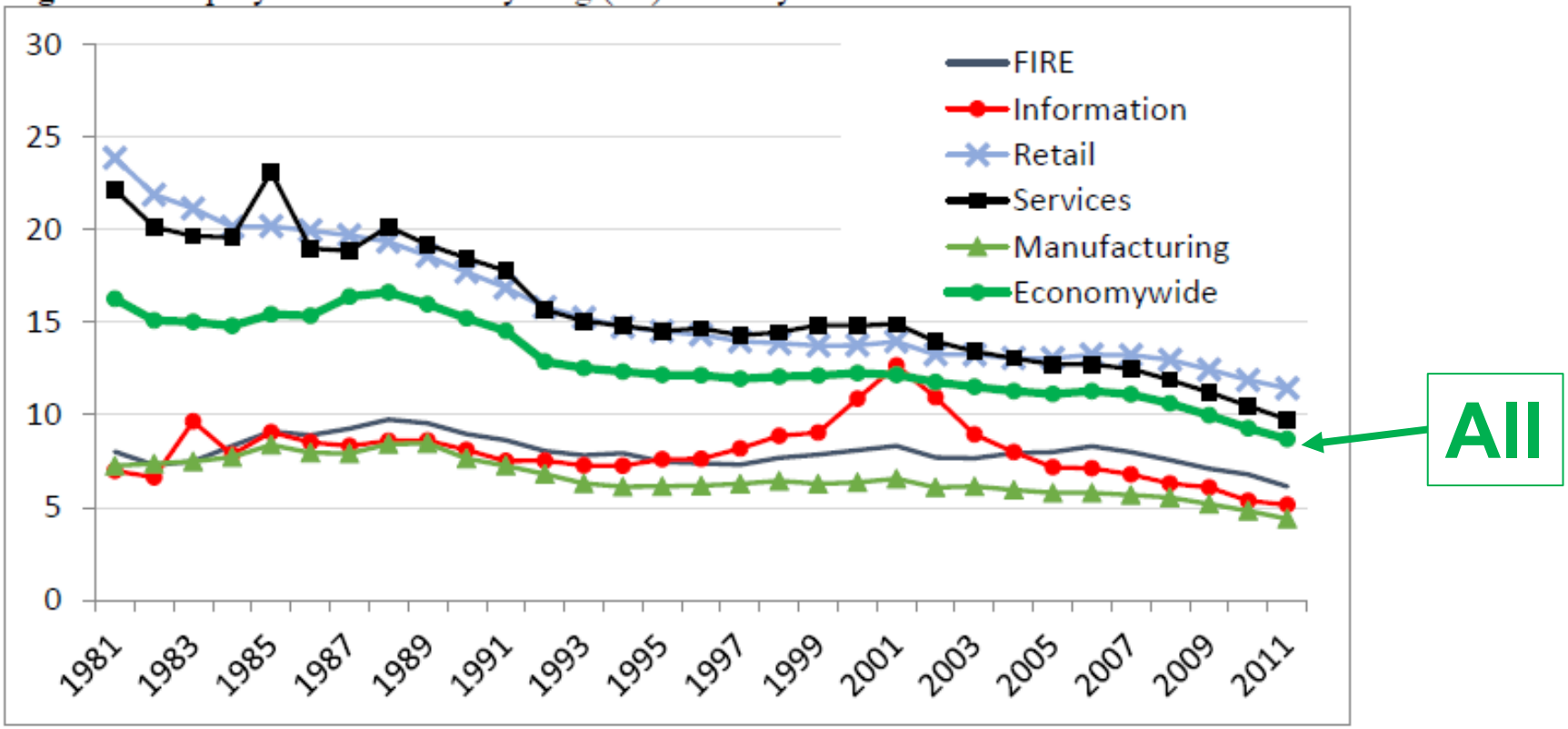
Rising US productivity dispersion (manufacturing)



Source: Decker, Haltiwanger, Jarmin & Miranda (2018, Figure A6)
Notes: Standard Deviation of log(real sales/employment) normalized in a NAICS 6 digit industry-year. HP filtered series in dashed lines. LBD is population whereas ASM is corrected for sample selection. Weights are employment weights.

Declining US Business Dynamism?

Figure 2: Employment shares for young (<5) firms by broad sector



Note: Young firms have age less than 5. Industries are defined on a consistent NAICS basis. Data include all firms (new entrants, exiters, and continuers). Author calculations from the Longitudinal Business Database.

Source: Decker et al (2017)

Conclusions on micro facts

- Large degree of persistent performance differentials ("PPD's") across firms/plants even within detailed private sector industries, even in advanced economies like US
  - Not all measurement error

- Reallocation accounts for large fraction of aggregate TFP growth over time and cross-country differences

- How to map these "moments" into theoretical models (later)

- What causes these large differences?

Back Up

Be careful generalizing from US publicly listed firms

Size class	COMPUSTAT	Census
0	2,576	719,978
1 to 4	123	2,638,070
5 to 9	149	1,006,897
10 to 19	251	593,696
20 to 99	1,287	487,491
100 to 499	2,123	79,707
500+	4,267	16,079
Total	10,776	5,541,918

Note: 1997 data from Census, Axtell (2001)

Power law in firm size

Discrete Pareto distributed random variable, *S*, where s_0 (*scale* parameter) is minimum size; α (*shape* parameter determines skewness).

$$\Pr[S \geq s_i] = \left(\frac{s_0}{s_i}\right)^{\alpha}, \; s_i \geq s_0, \; \alpha > 0$$

$$CDF : F(s) = 1 - \left(\frac{s_0}{s_i}\right)^{\alpha}$$

"Zipf's Law" is when α = 1 (*shape* parameter).
Firm size distribution close to this (& many other things – e.g. city size).

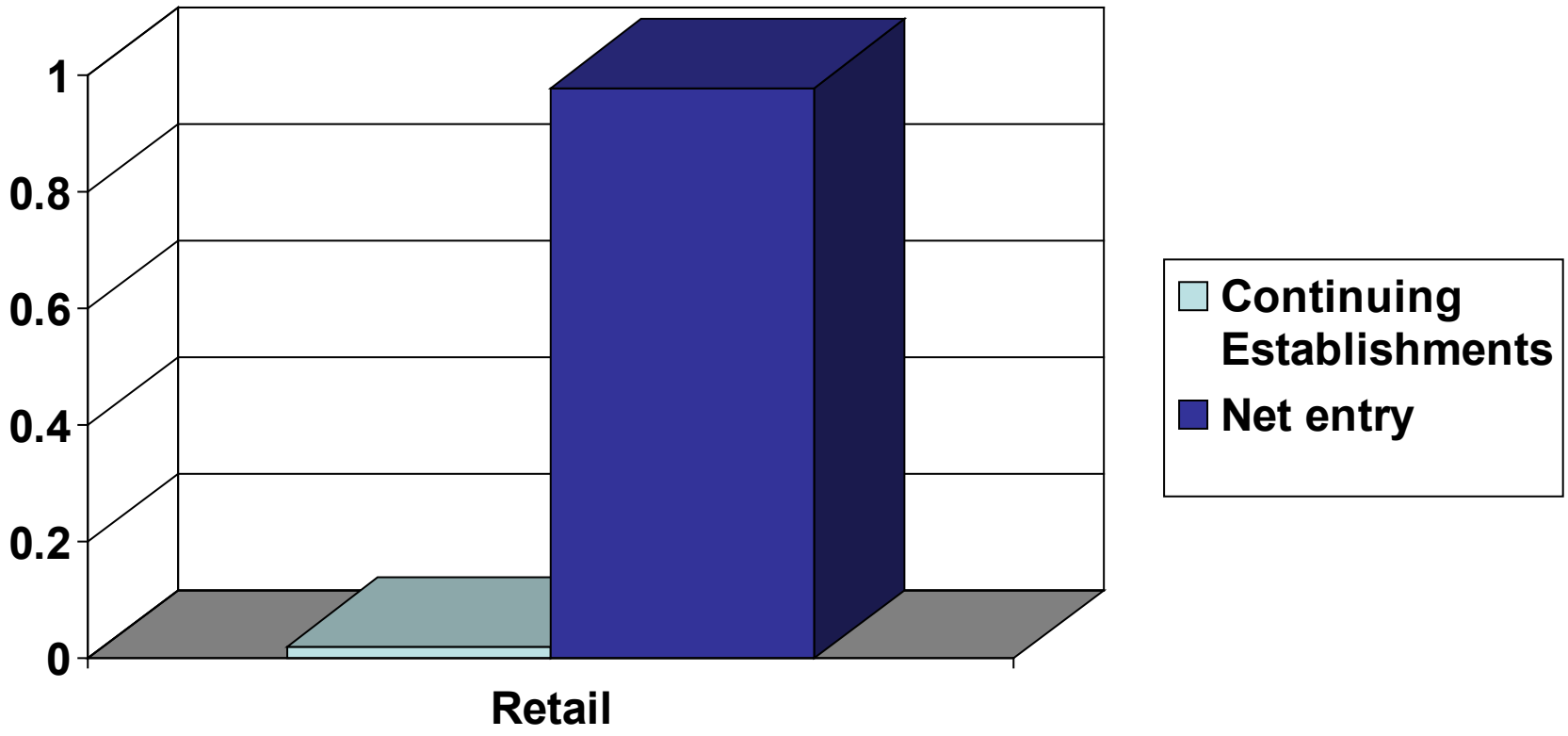
Firm size distribution is close to Zipf Law

Table 3. Theoretical power law exponents for U.S. firms over a 10-year period. Note that even though the number of firms and total employees each increased over this period, as did the average firm size, the value of α was approximately unchanged.

Year	Firms	Employees	Mean firm size	α, from (4)
1997	5,541,918	105,299,123	19.00	0.9966
1996	5,478,047	102,187,297	18.65	0.9986
1995	5,369,068	100,314,946	18.68	0.9983
1994	5,276,964	96,721,594	18.33	1.0004
1993	5,193,642	94,773,913	18.25	1.0008
1992	5,095,356	92,825,797	18.22	1.0009
1991	5,051,025	92,307,559	18.28	1.0004
1990	5,073,795	93,469,275	18.42	0.9995
1989	5,021,315	91,626,094	18.25	1.0006
1988	4,954,645	87,844,303	17.73	1.0039

Source: Axtell (2001)

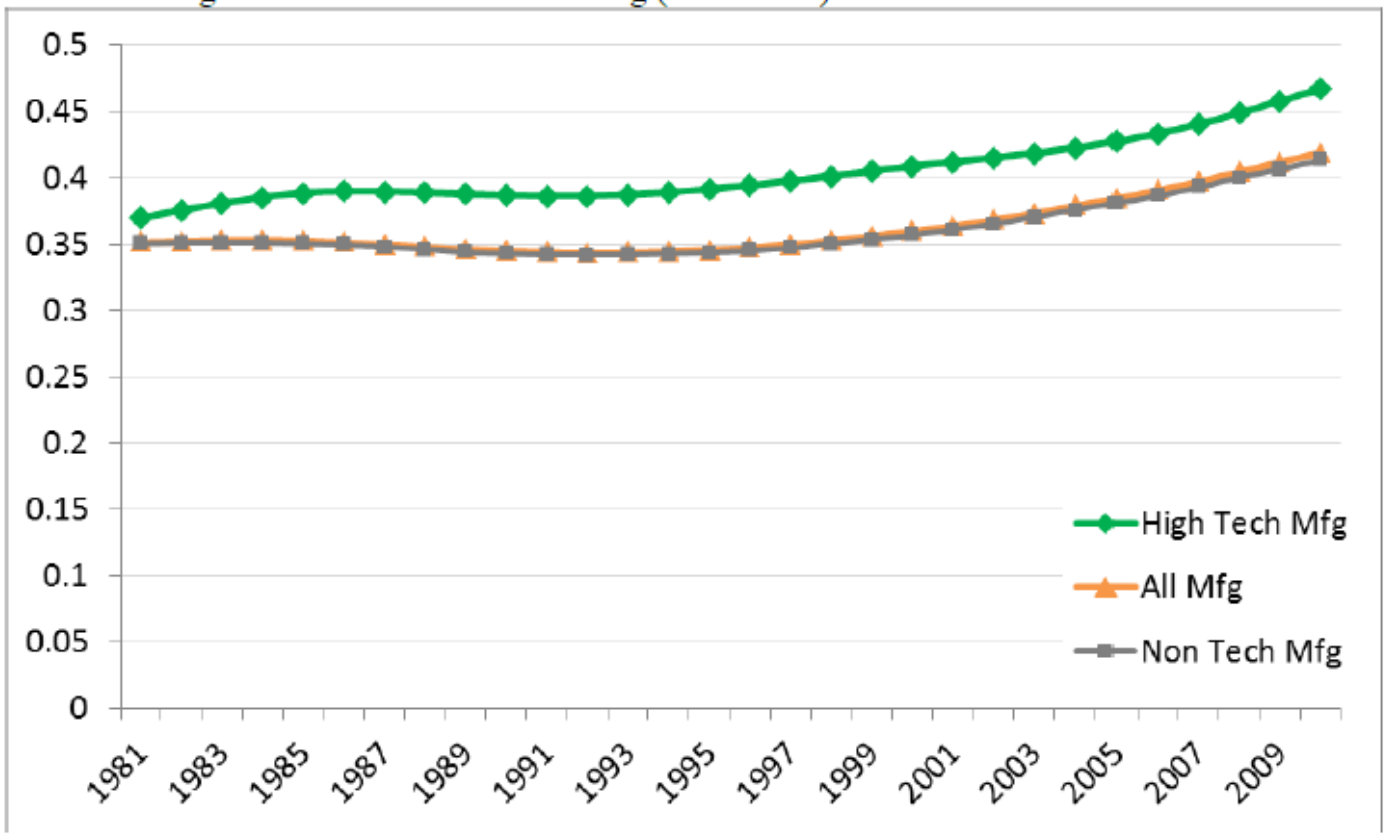
Reallocation (including entry) accounts for almost all Retail labor productivity growth



Source: Foster, Haltiwanger & Krizan (2006)

Declining US Business Dynamism?

Figure A1: Within-industry TFP dispersion (std deviation) in Manufacturing, High-Tech Manufacturing and Non-Tech Manufacturing (HP Trends)



Note: The standard deviation is the based on within-detailed industry log TFP. High-Tech is defined as in Hecker (2005). Manufacturing is defined on a consistent NAICS basis. Author calculations from the Longitudinal Business Database, the Annual Survey of Manufacturers, and the Census of Manufacturers. Hodrick Prescott Trends depicted.

Source: Decker et al (2017)

KEY QUESTIONS IN ECONOMETRICS OF ORGANIZATIONAL ECONOMICS

1. What is causal impact of different forms of organizational practices on performance?
   - **Example:** What is effect of incentive pay on productivity (Lazear, 2000)
   - Identification Revolution in econometrics (Angrist & Pischke, 2015): RCTs; RDD, Diff in Diffs; IV, etc.
2. What determines the distribution of organizational practices in an economy?
   - e.g. Why some firms adopt incentive pay & others don't
   - More a "between firm" question (rather than "within firm")
3. Can the answers to 1. & 2. explain heterogeneity of performance across firms and countries?
   - Both an accounting & an equilibrium question
4. "Reduced form" & "Structural" approaches complements