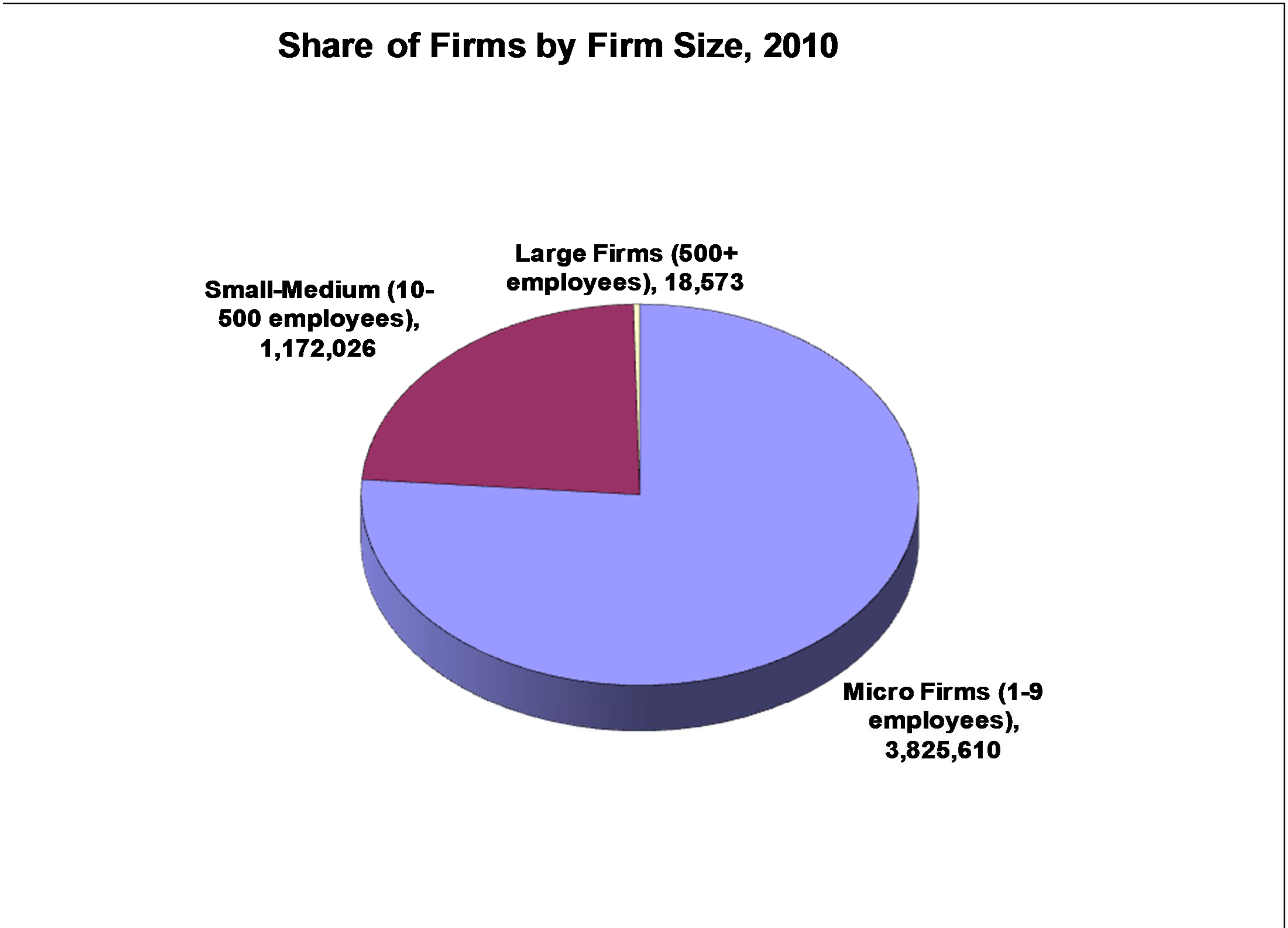
Title: Productivity, Job Creation and Entrepreneurship

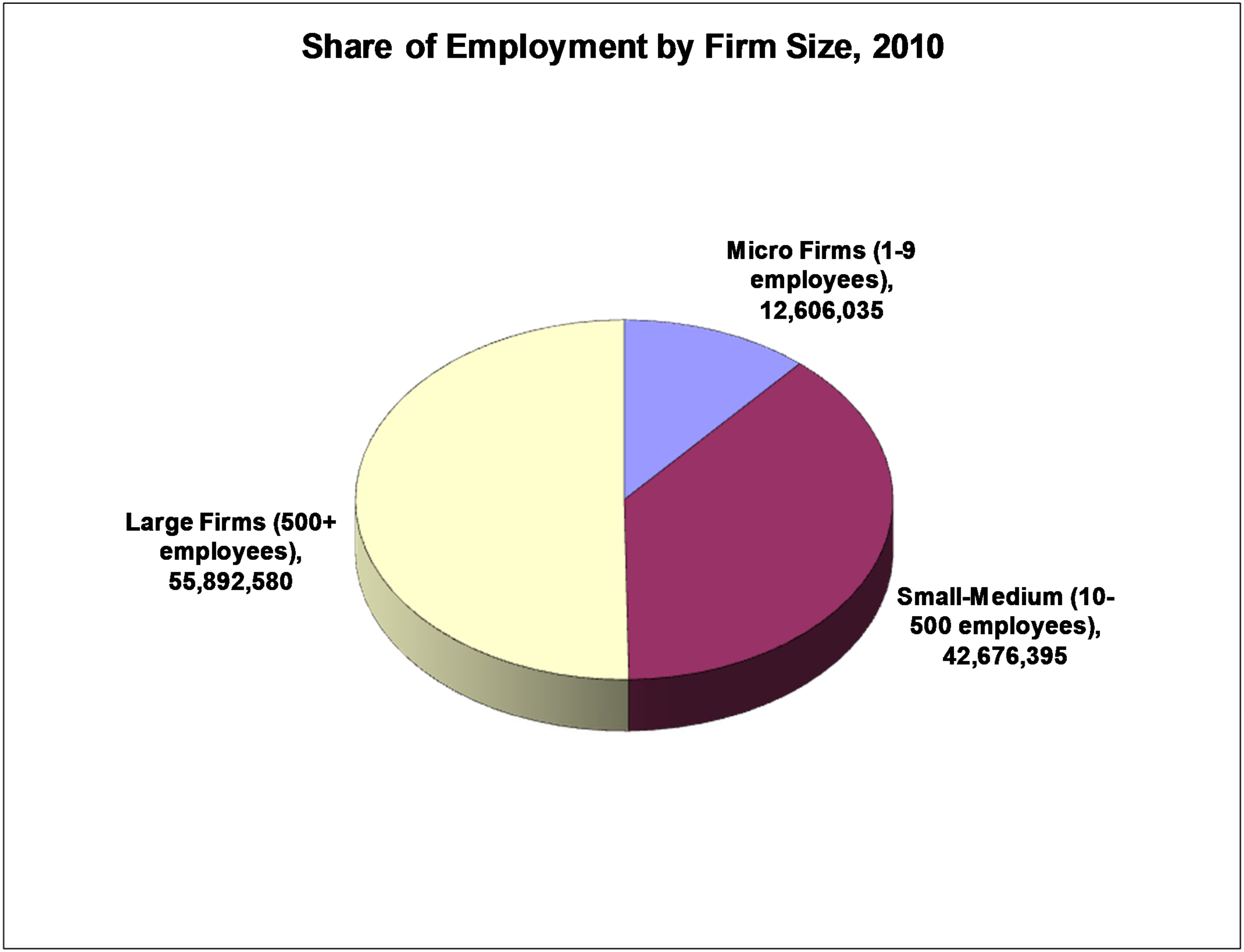
Lecture Notes for NBER Entrepreneurship Bootcamp

July 2015

By John Haltiwanger

University of Maryland and NBER


Share of Firms by Firm Size, 2010
Large Firms (500+ employees), 18,573
Small-Medium (10-500 employees), 1,172,026
Micro Firms (1-9 employees), 3,825,610



Share of Employment by Firm Size, 2010
Micro Firms (1-9 employees), 12,606,035
Small-Medium (10-500 employees), 42,676,395
Large Firms (500+ employees), 55,892,580

CHRYSLER

Ford

GM

kmart

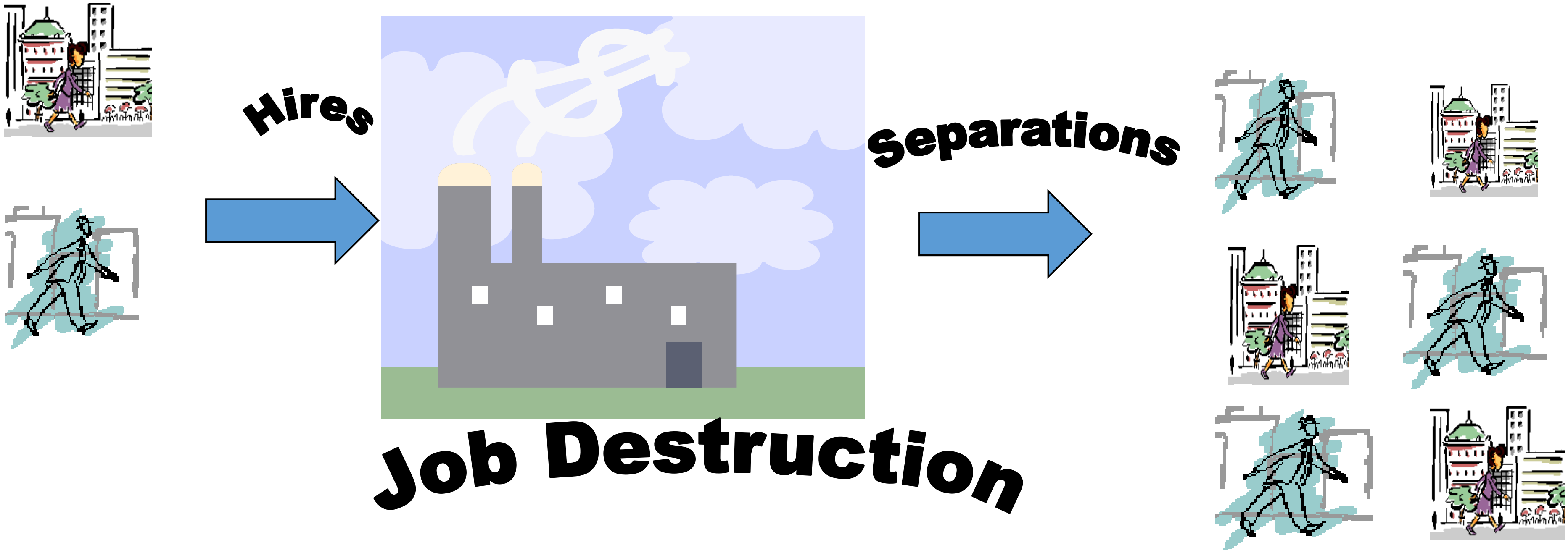
USS

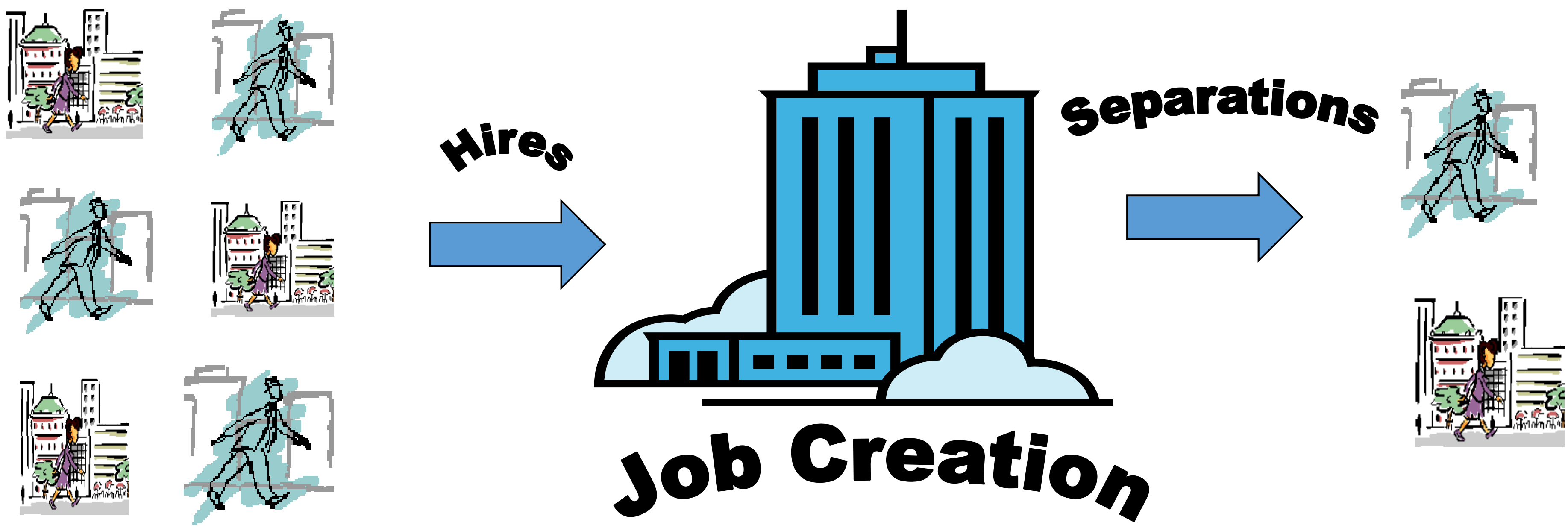
TOYOTA

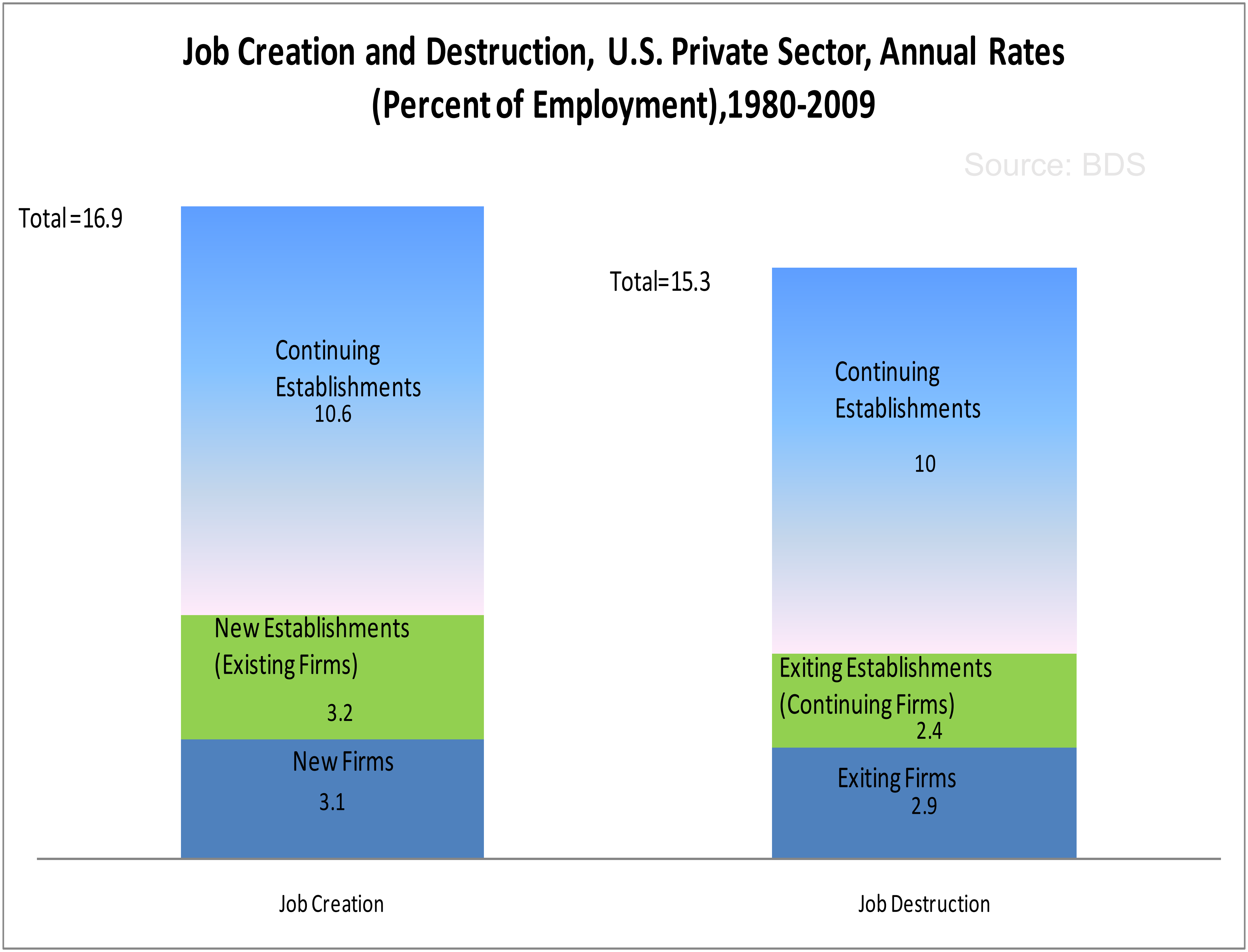
Google

HONDA

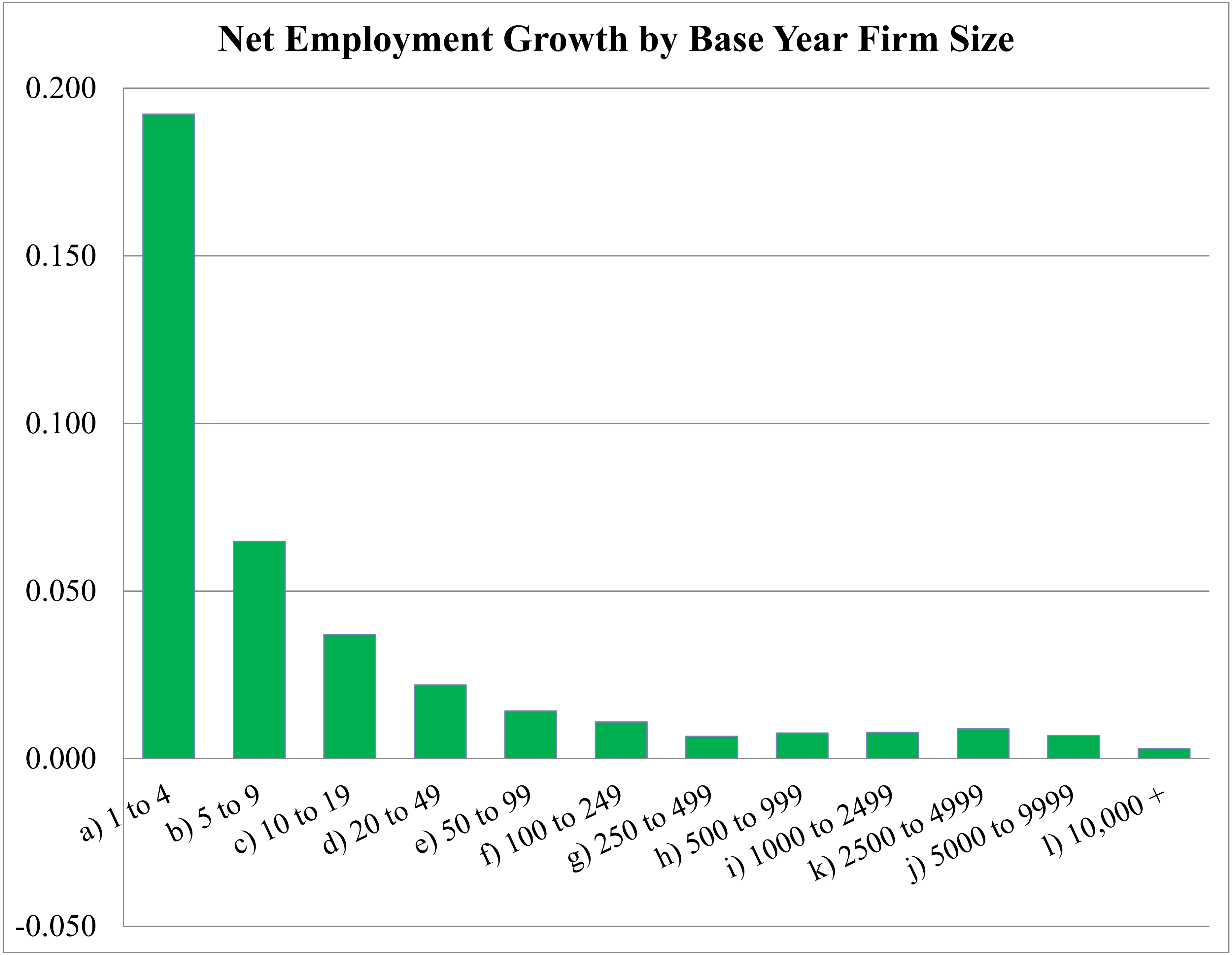
STARBUCKS COFFEE

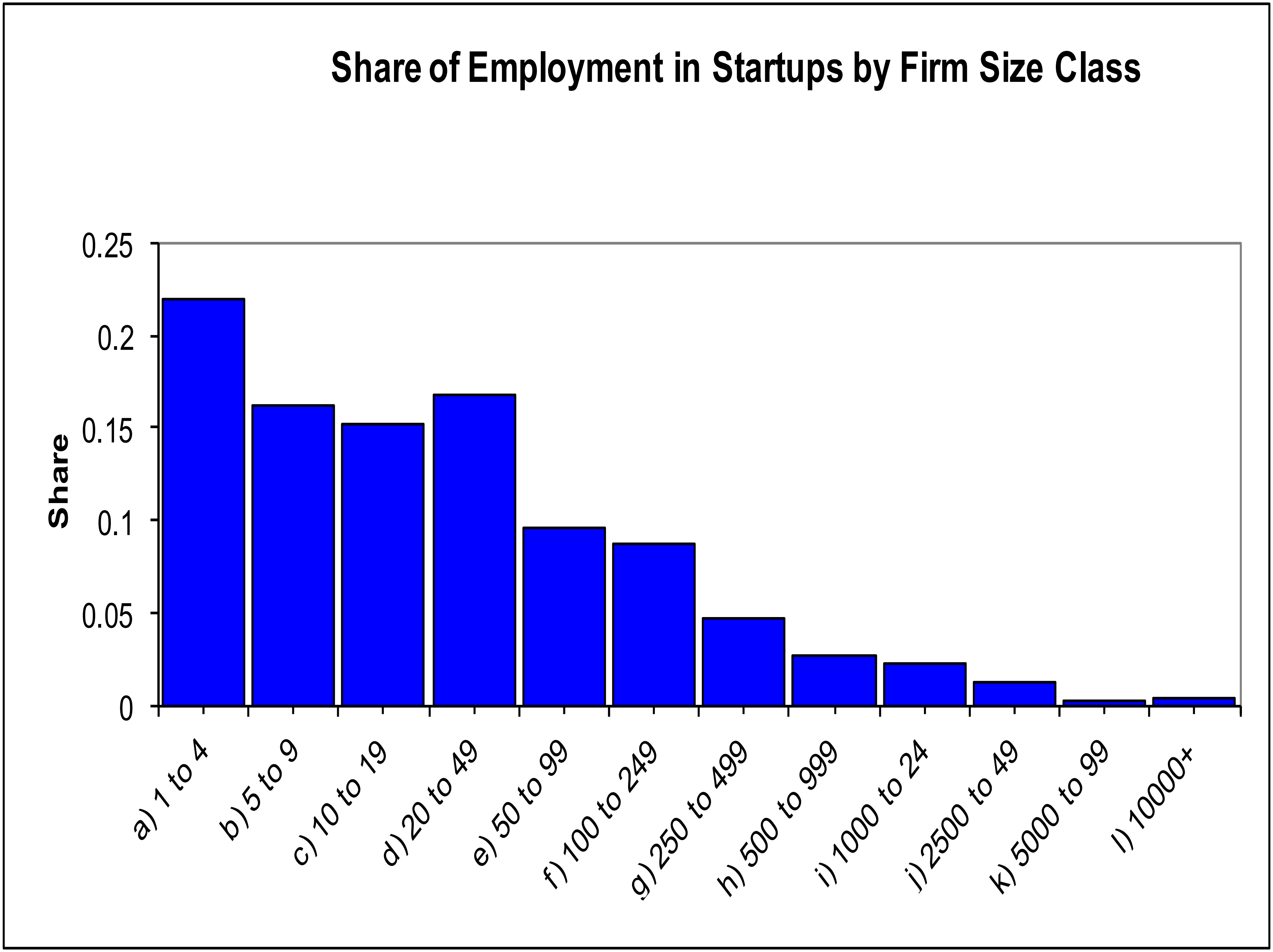

Hires
Separations
Job Destruction


Hires
Separations
Job Creation

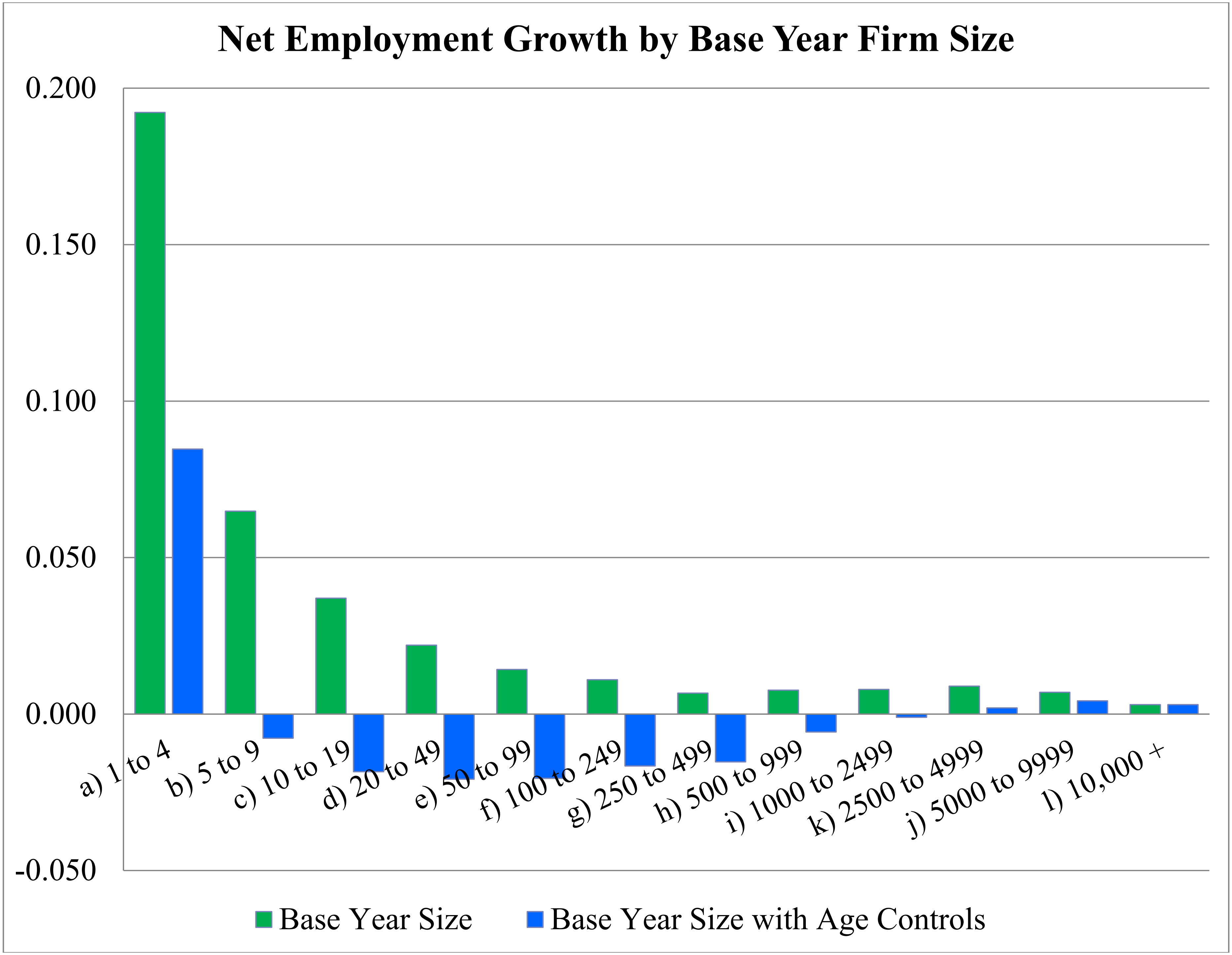


Job Creation and Destruction, U.S. Private Sector, Annual Rates
(Percent of Employment),1980-2009
Source: BDS
Total =16.9
Continuing Establishments
10.6
New Establishments (Existing Firms)
3.2
New Firms
3.1
Total=15.3
Continuing Establishments
10
Exiting Establishments (Continuing Firms)
2.4
Exiting Firms
2.9
Job Creation
Job Destruction

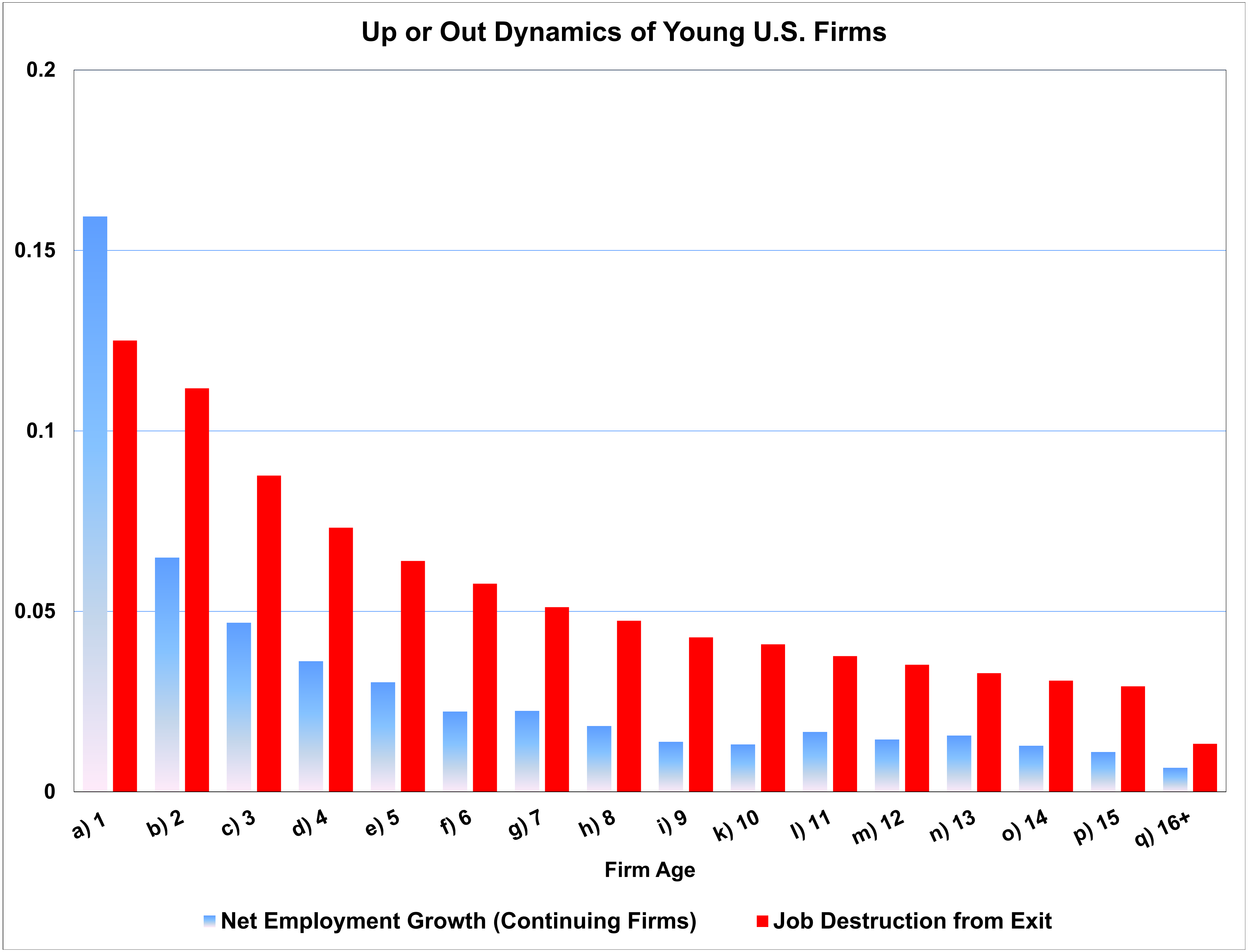

Net Employment Growth by Base Year Firm Size
0.200
0.150
0.100
0.050
0.000
-0.050
a) 1 to 4
b) 5 to 9
c) 10 to 19
d) 20 to 49
e) 50 to 99
f) 100 to 249
g) 250 to 499
h) 500 to 999
i) 1000 to 2499
k) 2500 to 4999
j) 5000 to 9999
l) 10,000 +



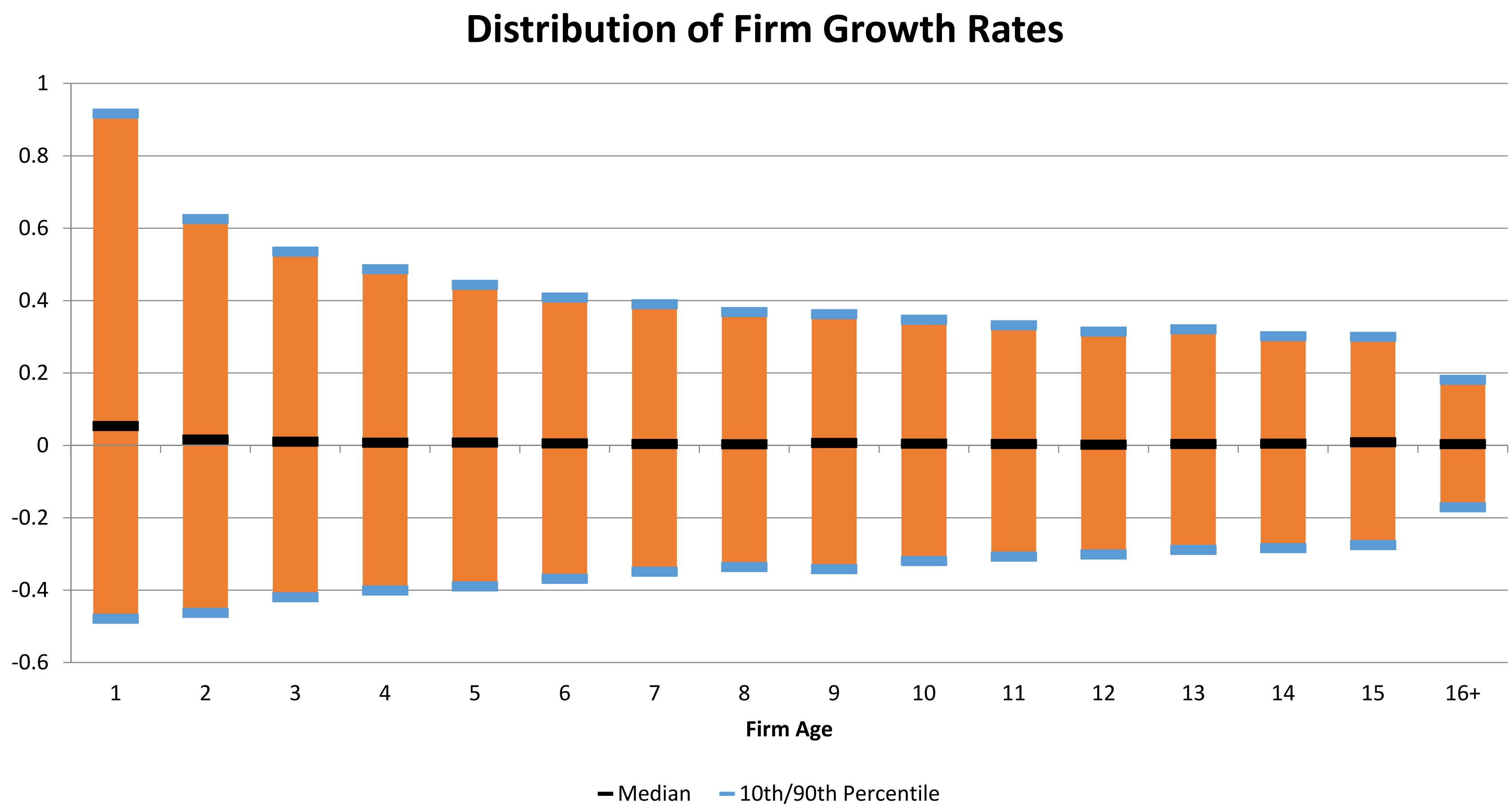
Share of Employment in Startups by Firm Size Class
Share
0.25
0.2
0.15
0.1
0.05
0
a) 1 to 4
b) 5 to 9
c) 10 to 19
d) 20 to 49
e) 50 to 99
f) 100 to 249
g) 250 to 499
h) 500 to 999
i) 1000 to 24
j) 2500 to 49
k) 5000 to 99
l) 10000+

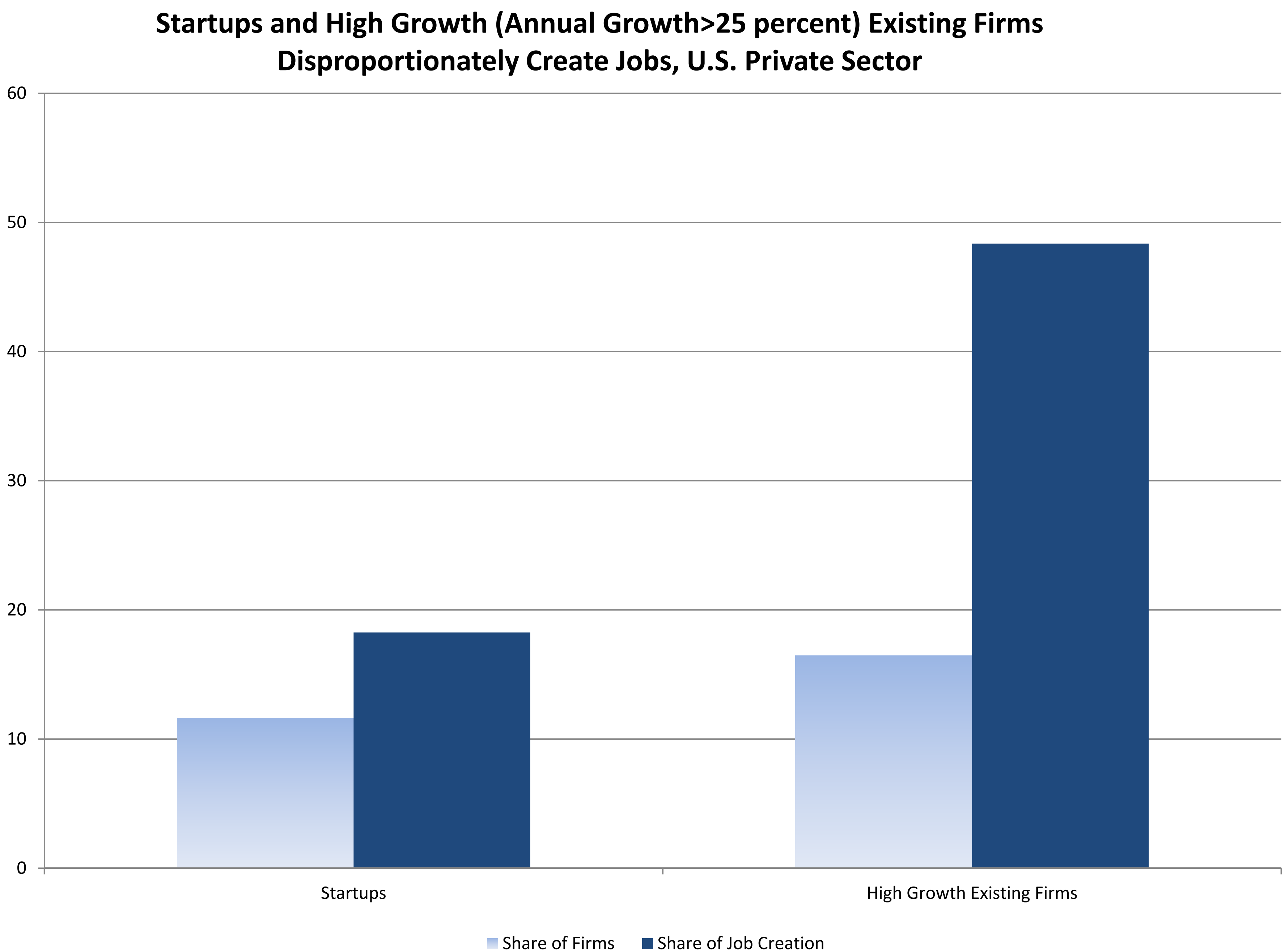


Net Employment Growth by Base Year Firm Size
0.200
0.150
0.100
0.050
0.000
-0.050
a) 1 to 4
b) 5 to 9
c) 10 to 19
d) 20 to 49
e) 50 to 99
f) 100 to 249
g) 250 to 499
h) 500 to 999
i) 1000 to 2499
k) 2500 to 4999
j) 5000 to 9999
l) 10,000 +
Base Year Size
Base Year Size with Age Controls


Up or Out Dynamics of Young U.S. Firms
0.2
0.15
0.1
0.05
0
a) 1
b) 2
c) 3
d) 4
e) 5
f) 6
g) 7
h) 8
i) 9
k) 10
l) 11
m) 12
n) 13
o) 14
p) 15
q) 16+
Firm Age
Net Employment Growth (Continuing Firms)
Job Destruction from Exit

Skewness of young continuing firms underlies high mean net growth of young firms



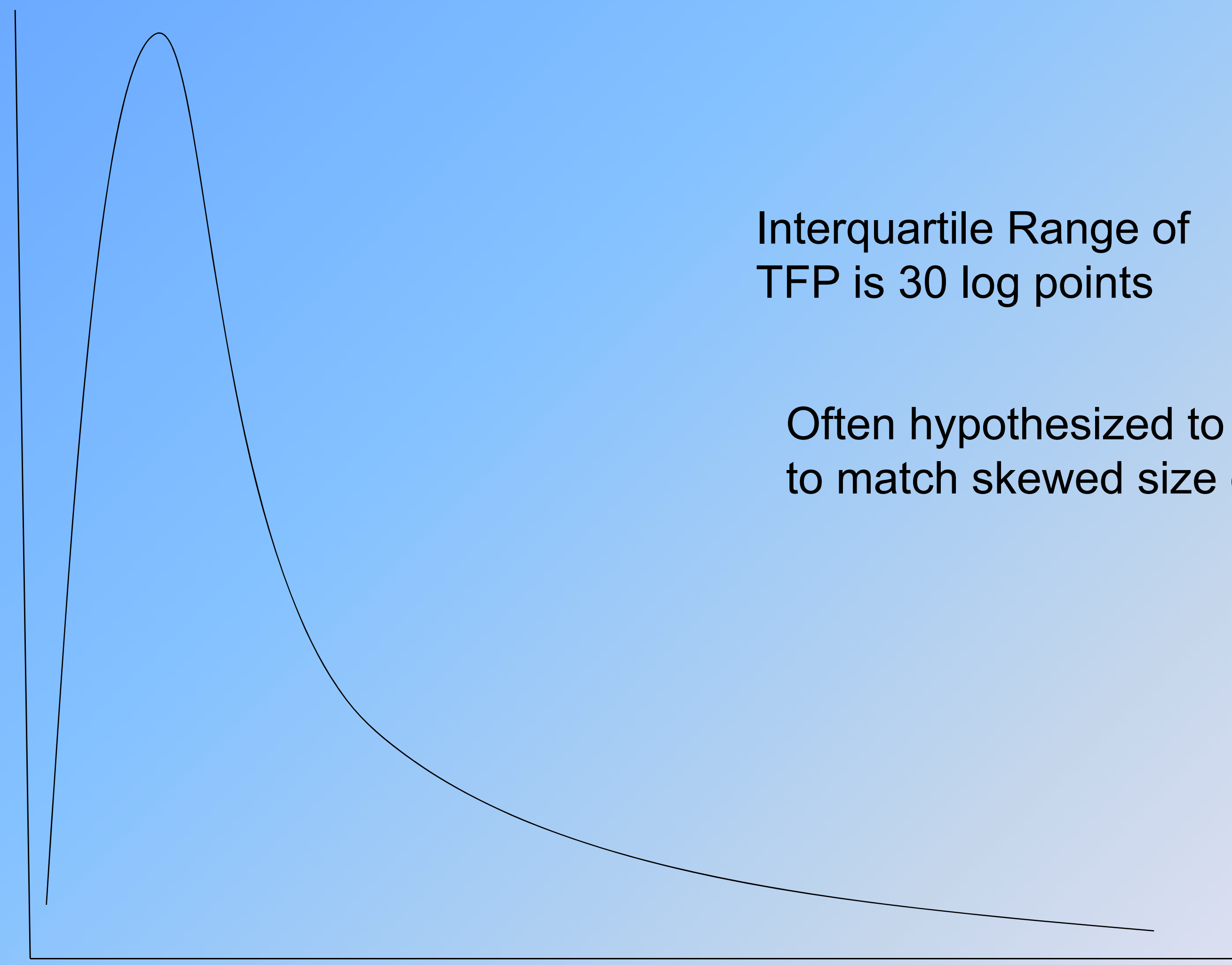


Source: Tabulations from Firm-Level Data Used in Haltiwanger, Jarmin and Miranda (

What accounts for cross sectional and dynamic patterns?

- Very skewed size distribution (often attributed to skewed productivity distribution)
- Very skewed distribution of *growth rates* of young firms
- These two findings presumably linked but complex:
  - One view is that entrants don't know type but draw from same distribution (hence skewed distribution of growth rates).
  - A complementary view is that young businesses engage in experimentation/innovation activity. Some are successful while others are not (hence skewed distribution of growth rates as probability of success is low).
  - In either view, what is happening to older incumbents?
- Part of the open debate:
  - Is this ex ante heterogeneity (Hurst and Pugsley, 2012, 2014) or ex post heterogeneity (draws from a pareto distribution of productivity or success at innovations)?
  - Some evidence we will see below suggests that Hurst and Pugsley holds for some industries while skewness of firm growth rates holds in others.
  - Important to take into account ex ante heterogeneity.

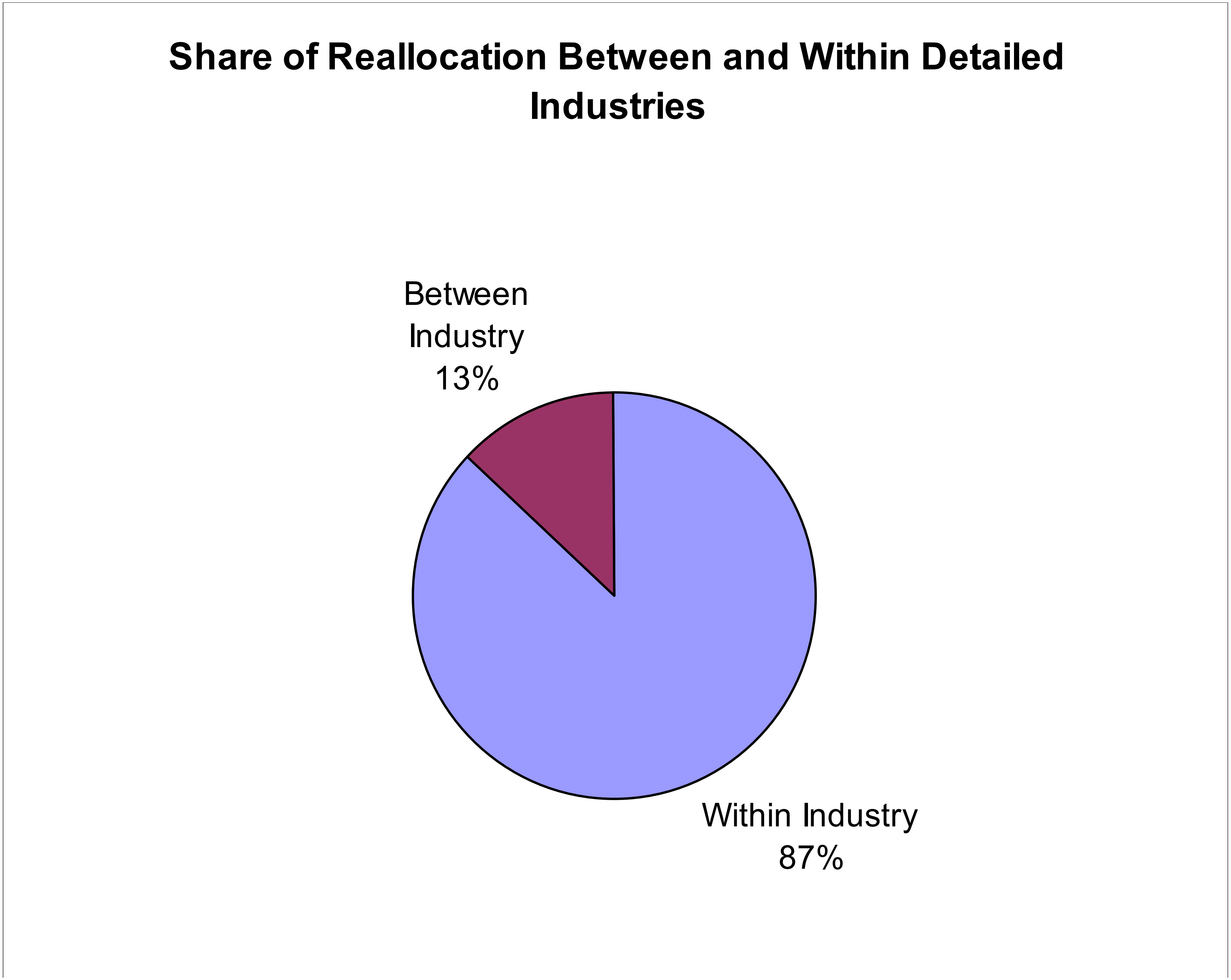
Productivity Distribution Within Narrowly Defined Industries

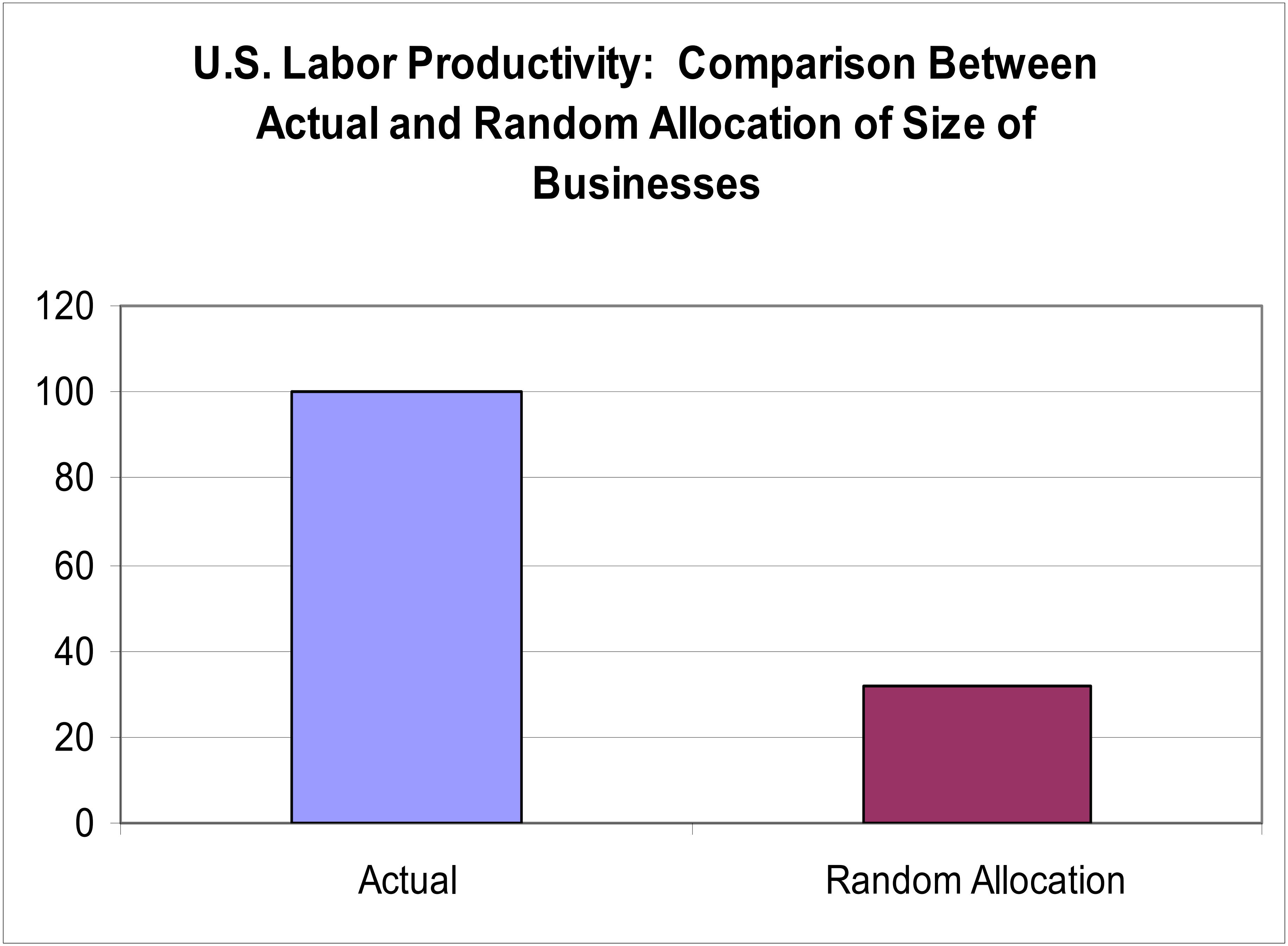


Interquartile Range of
TFP is 30 log points

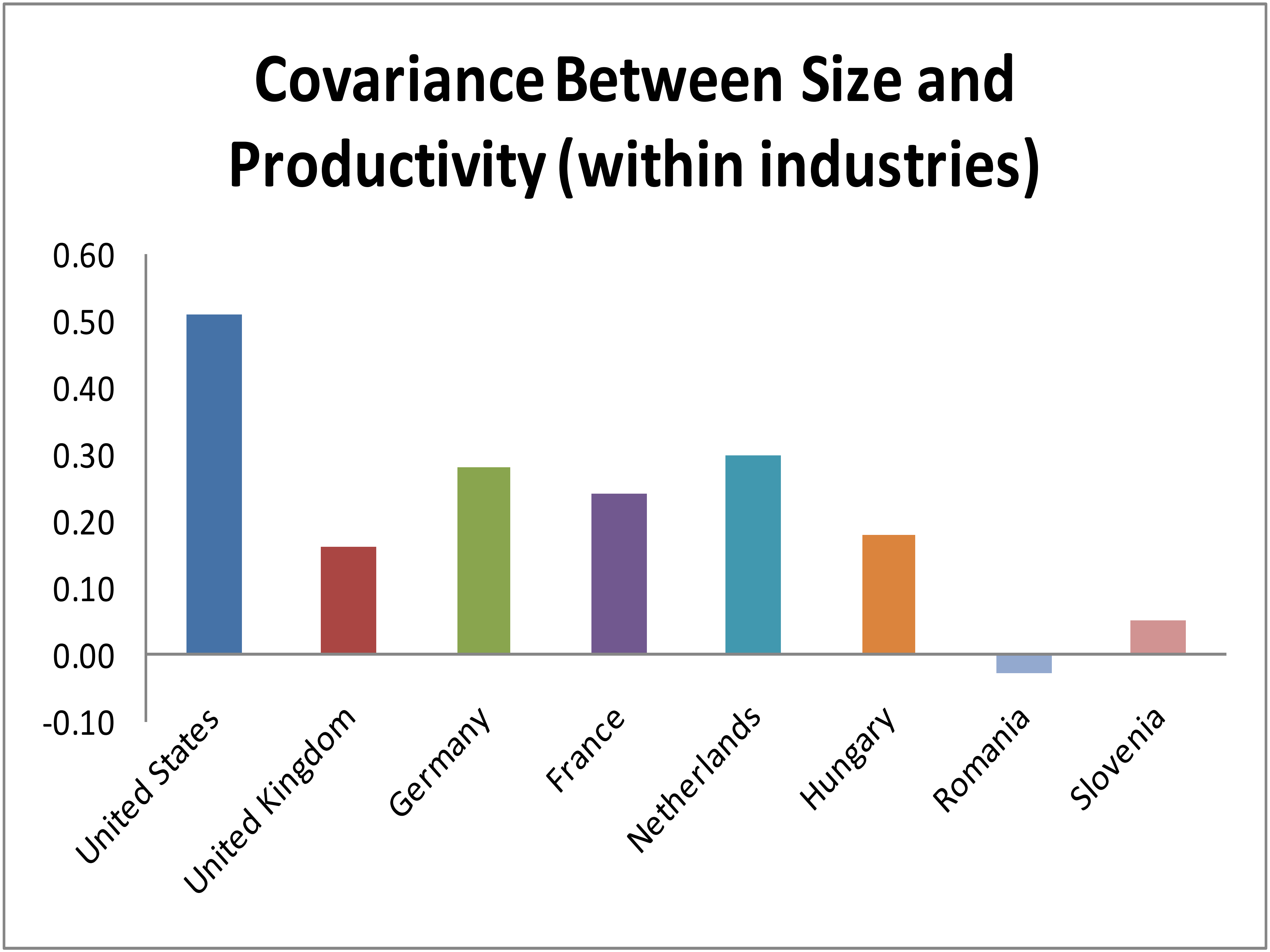
Often hypothesized to be pareto
to match skewed size distribution

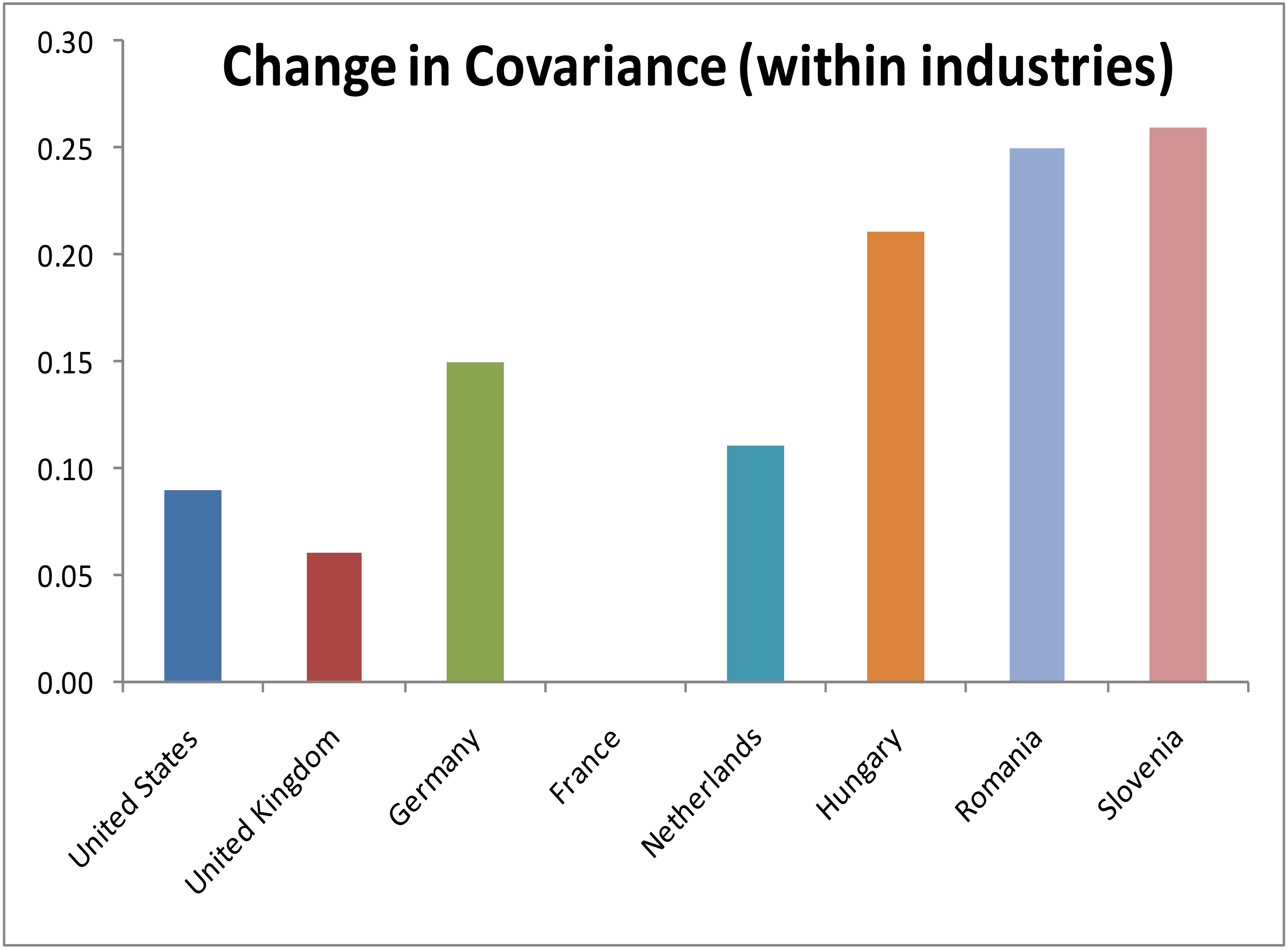
Productivity of Businesses

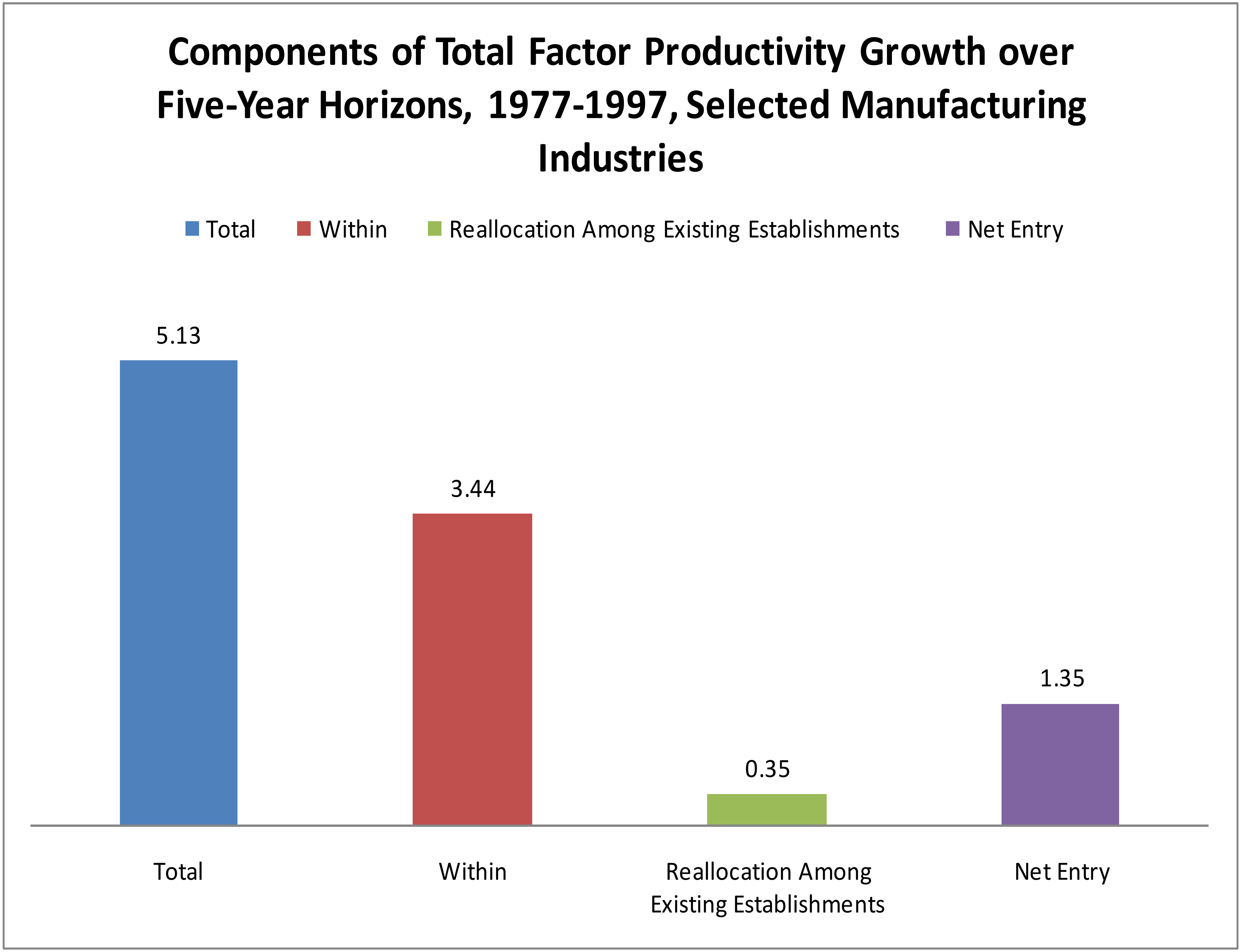

Share of Reallocation Between and Within Detailed Industries
Between Industry 13%
Within Industry 87%

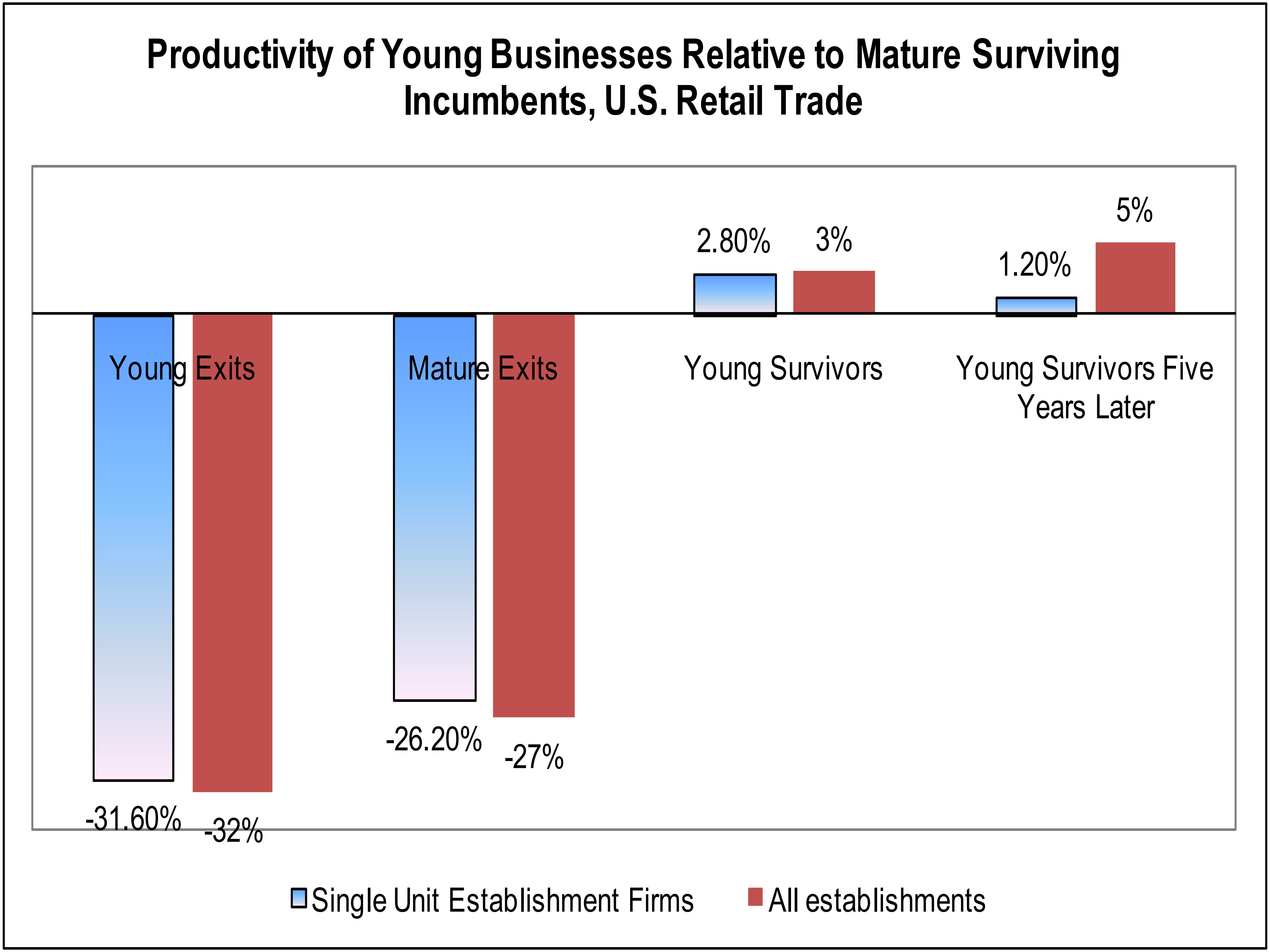


U.S. Labor Productivity: Comparison Between Actual and Random Allocation of Size of Businesses
120
100
80
60
40
20
0
Actual
Random Allocation


Covariance Between Size and Productivity (within industries)
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
United States
United Kingdom
Germany
France
Netherlands
Hungary
Romania
Slovenia

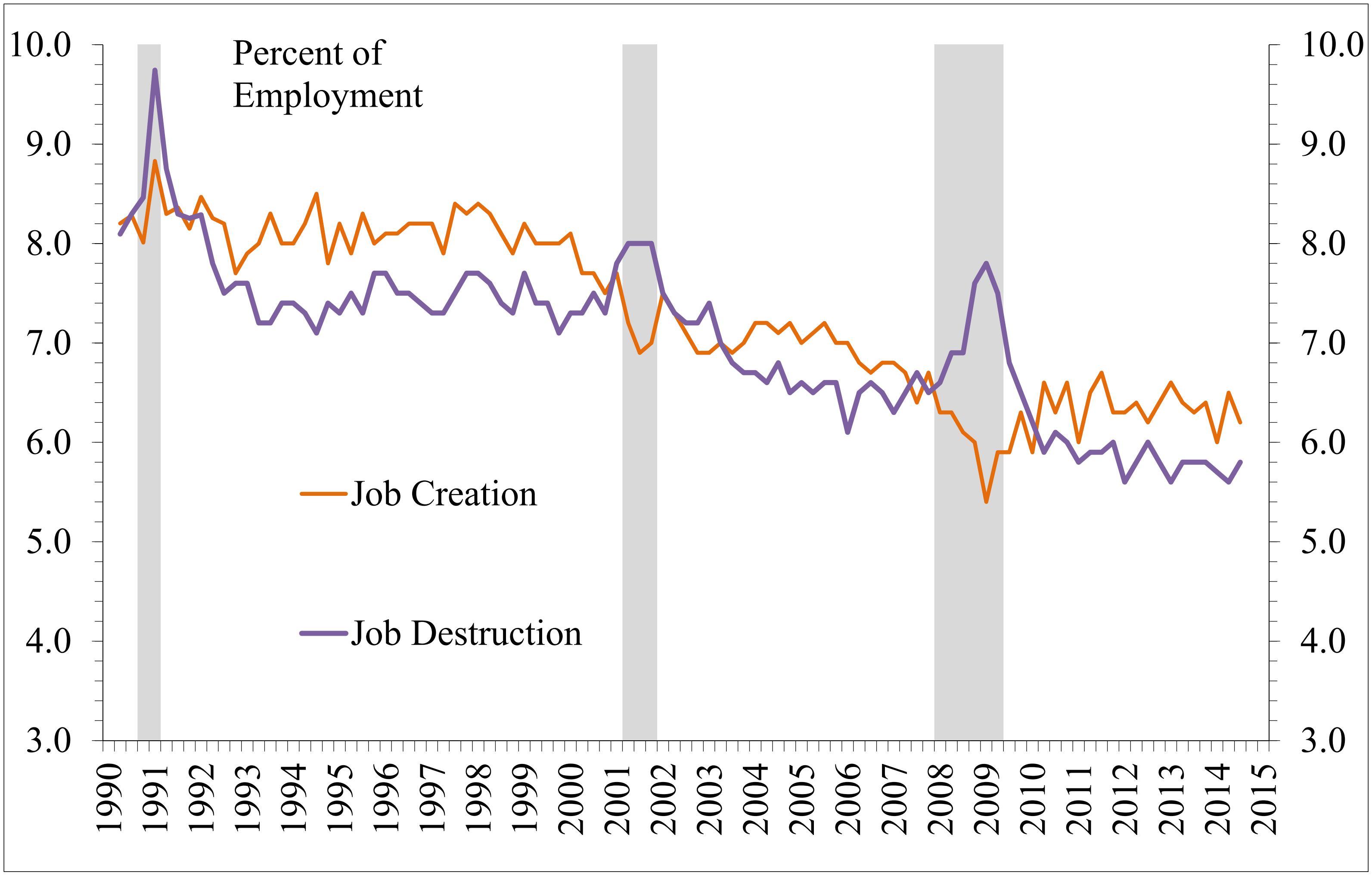

Change in Covariance (within industries)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
United States
United Kingdom
Germany
France
Netherlands
Hungary
Romania
Slovenia


Components of Total Factor Productivity Growth over Five-Year Horizons, 1977-1997, Selected Manufacturing Industries
Total
Within
Reallocation Among Existing Establishments
Net Entry
5.13
3.44
0.35
1.35
Total
Within
Reallocation Among Existing Establishments
Net Entry


Productivity of Young Businesses Relative to Mature Surviving Incumbents, U.S. Retail Trade
2.80%
3%
5%
1.20%
Young Exits
Mature Exits
Young Survivors
Young Survivors Five Years Later
-26.20%
-27%
-31.60%
-32%
Single Unit Establishment Firms
All establishments

Some Disturbing Trends?

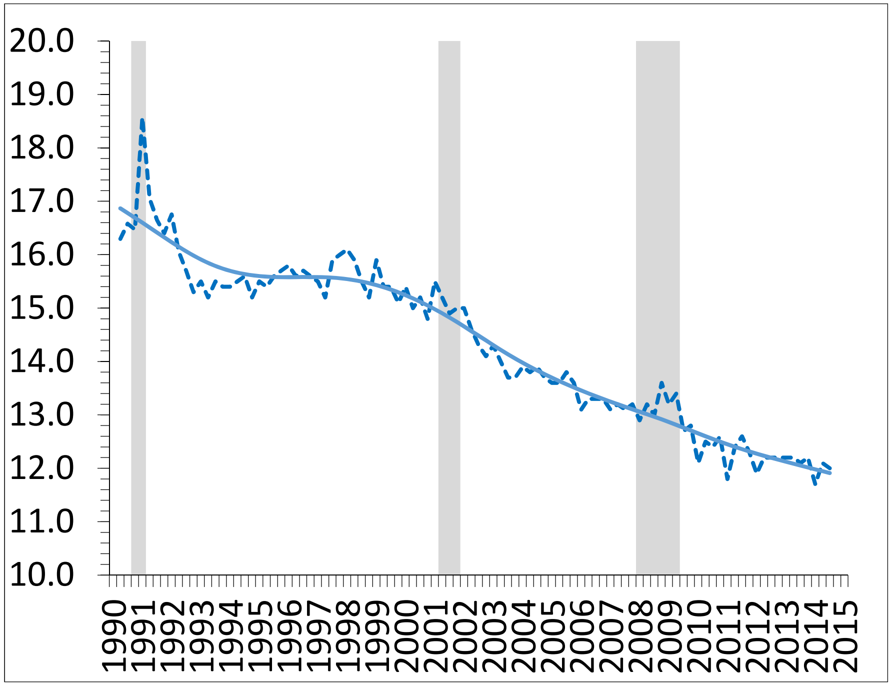
Declining Pace of Creation and Destruction in BED



Source: BLS BED DATA

Declining Business Dynamism is Evident from Multiple Data Sources

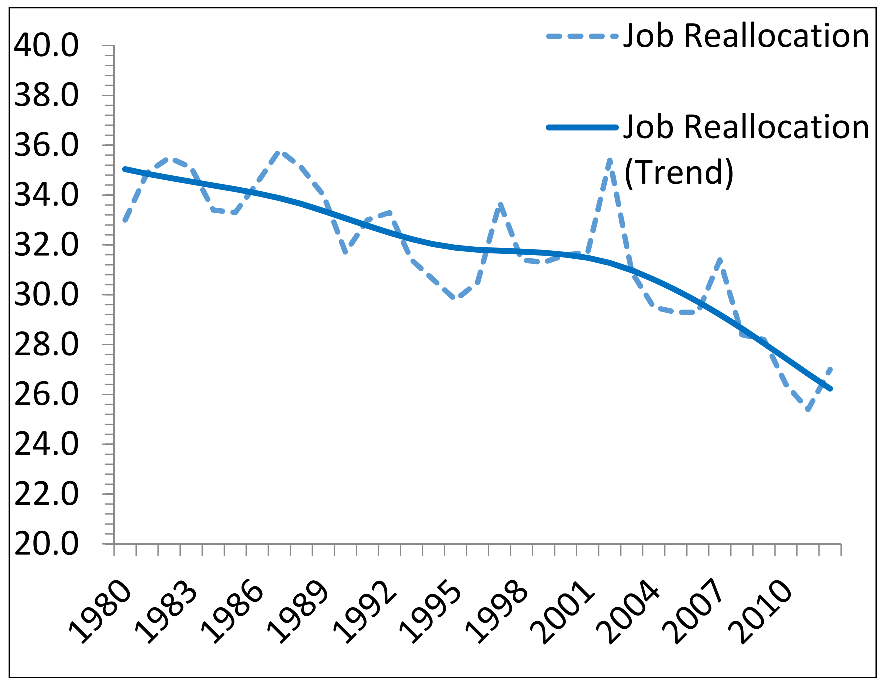
Job Reallocation Rate, U.S. Private Non-Farm (Quarterly)



Source: BED

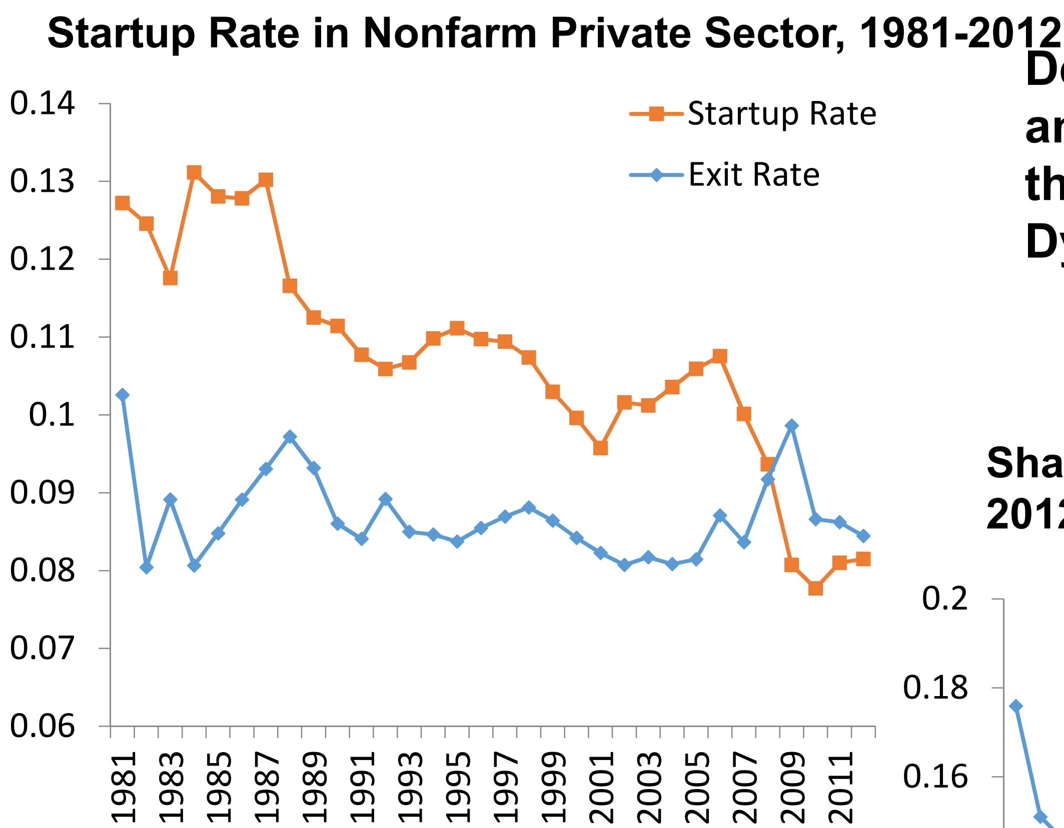
Declining Trend in Job Reallocation Accelerated in Post-2000 Period. Trend decline continues in post Great Recession period.

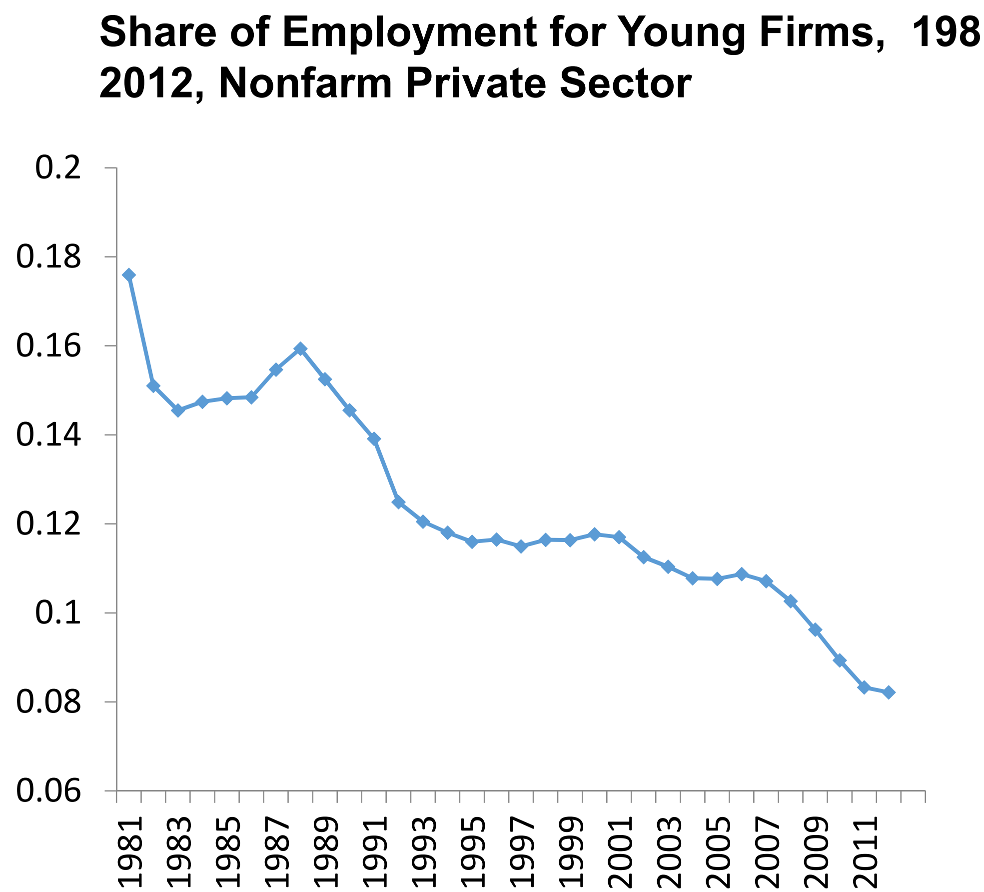
Job Reallocation Rate, U.S. Private Non-Farm (Annual
Source: BDS



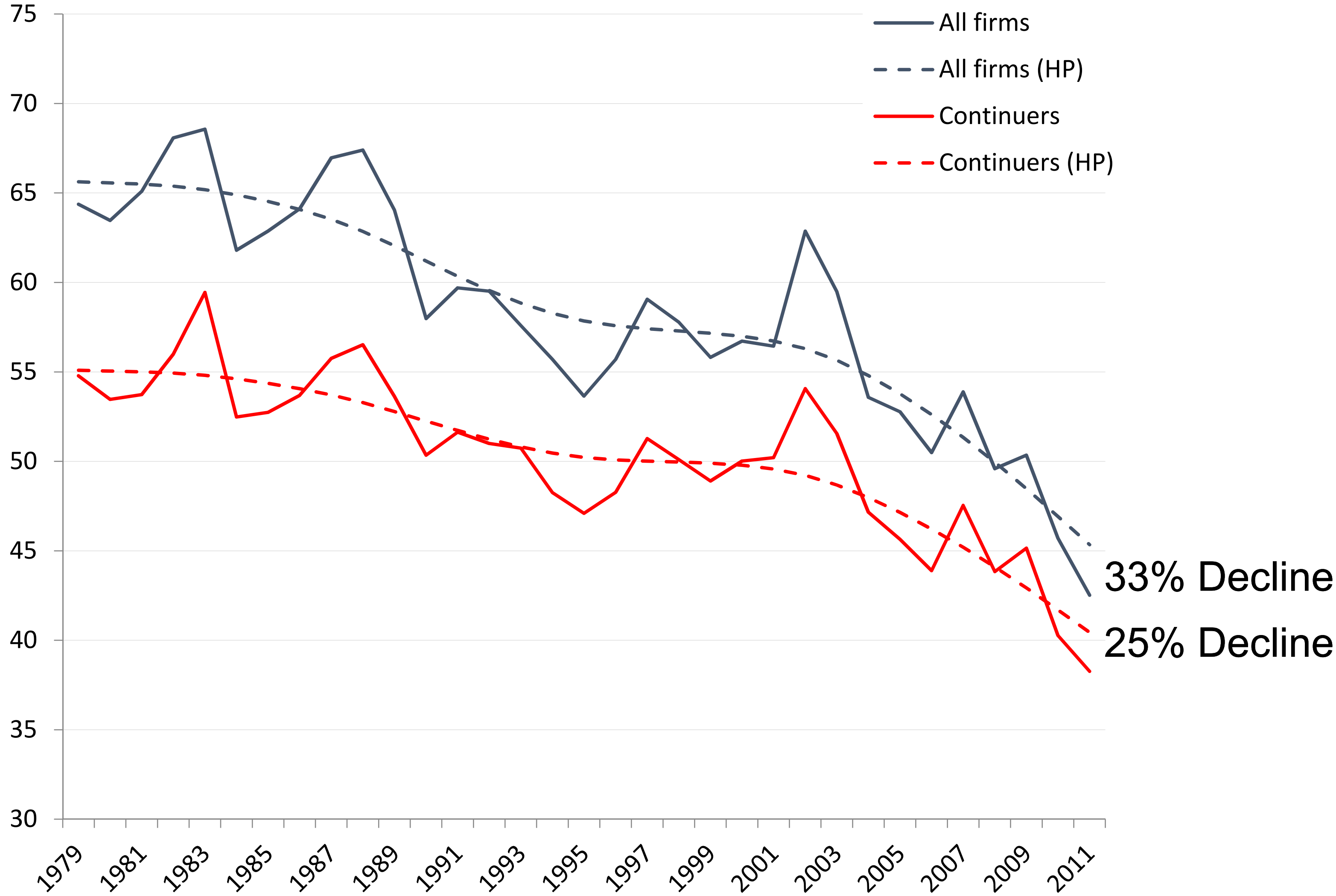
Solid lines are Hodrick-Prescott Trends

Declining Entrepreneurship Accompa and "Accounts" for Substantial Fracti the Decline in Measures of Business Dynamism





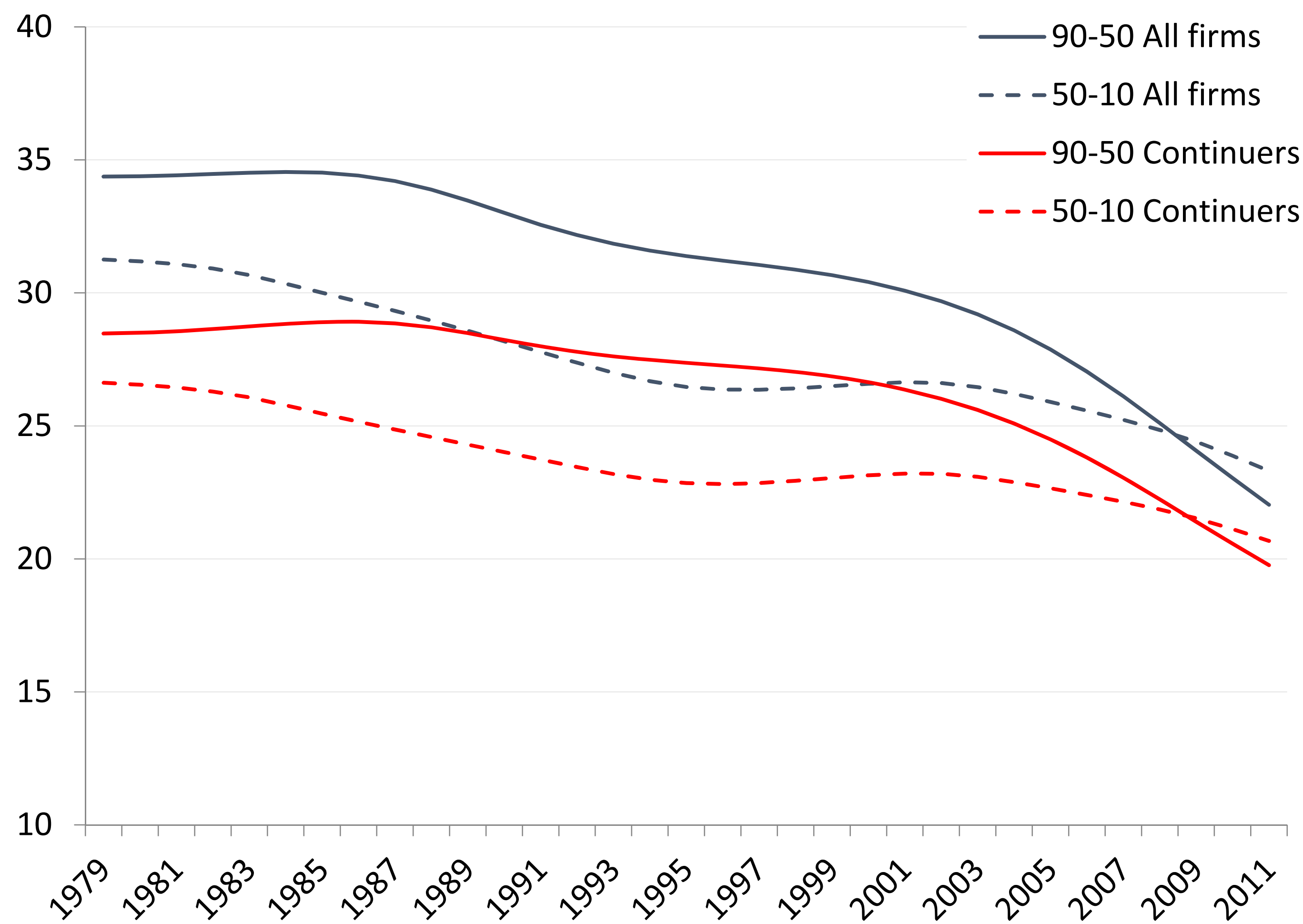
Declining Dispersion (90-10 differential)



Source: Census LBD Da

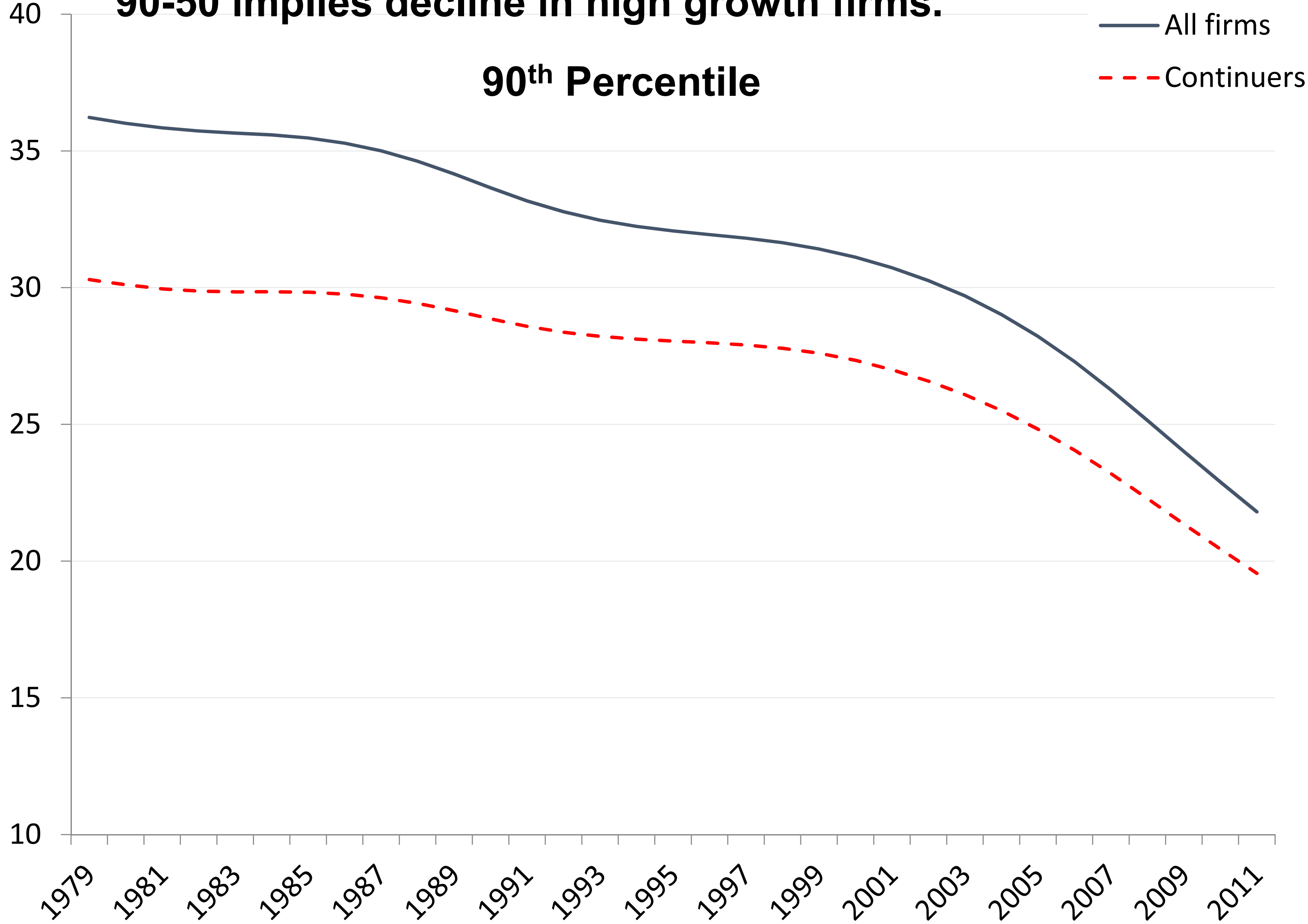
Employment-weighted Distribution of Firm Net Employment Growth

Pre-2000, decline in 90-10 due to decline in 90-50 and 50-10.
Post 2000, sharp decline in 90-50 relative to 50-10.

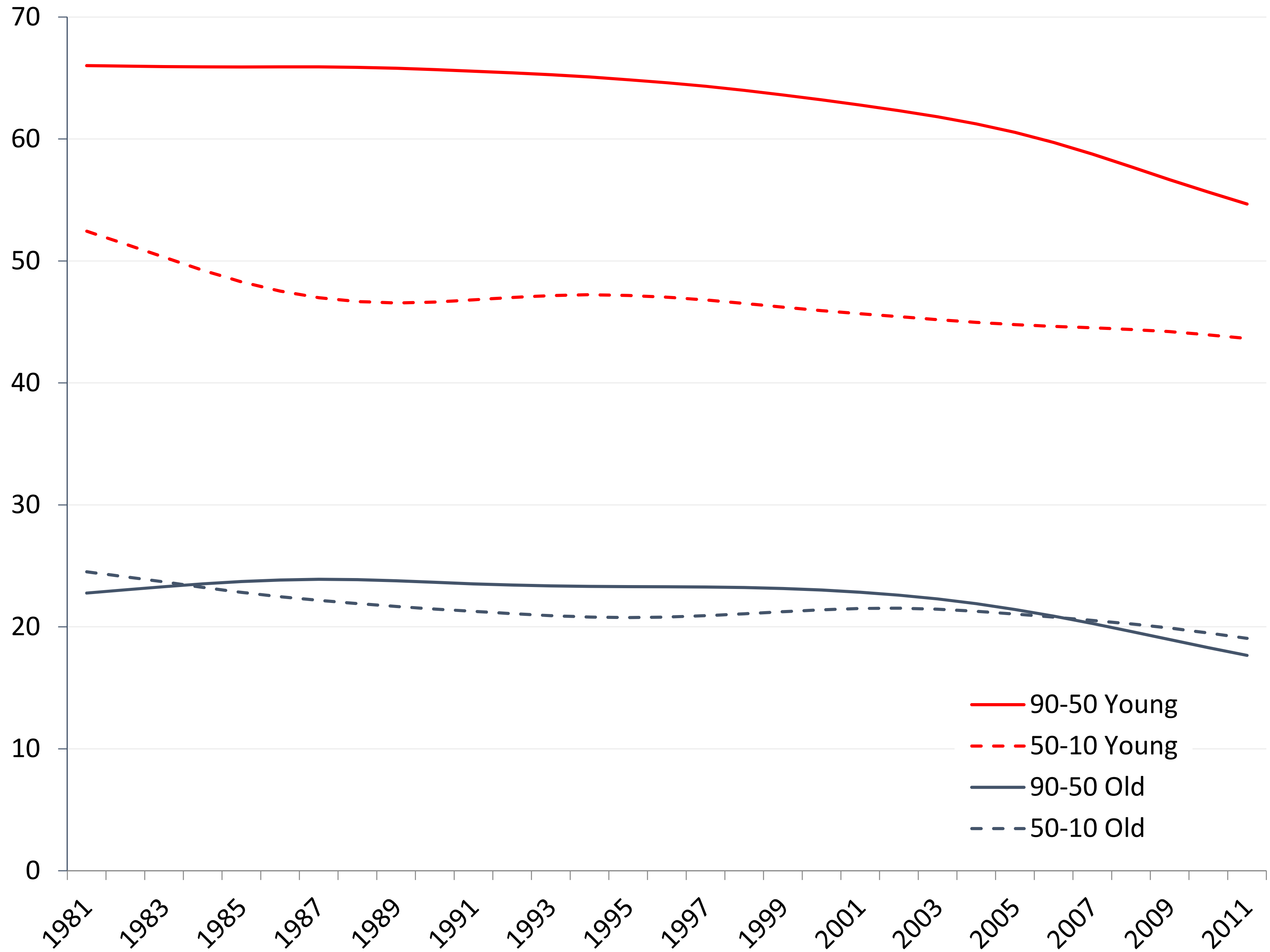


Showing HP Trends

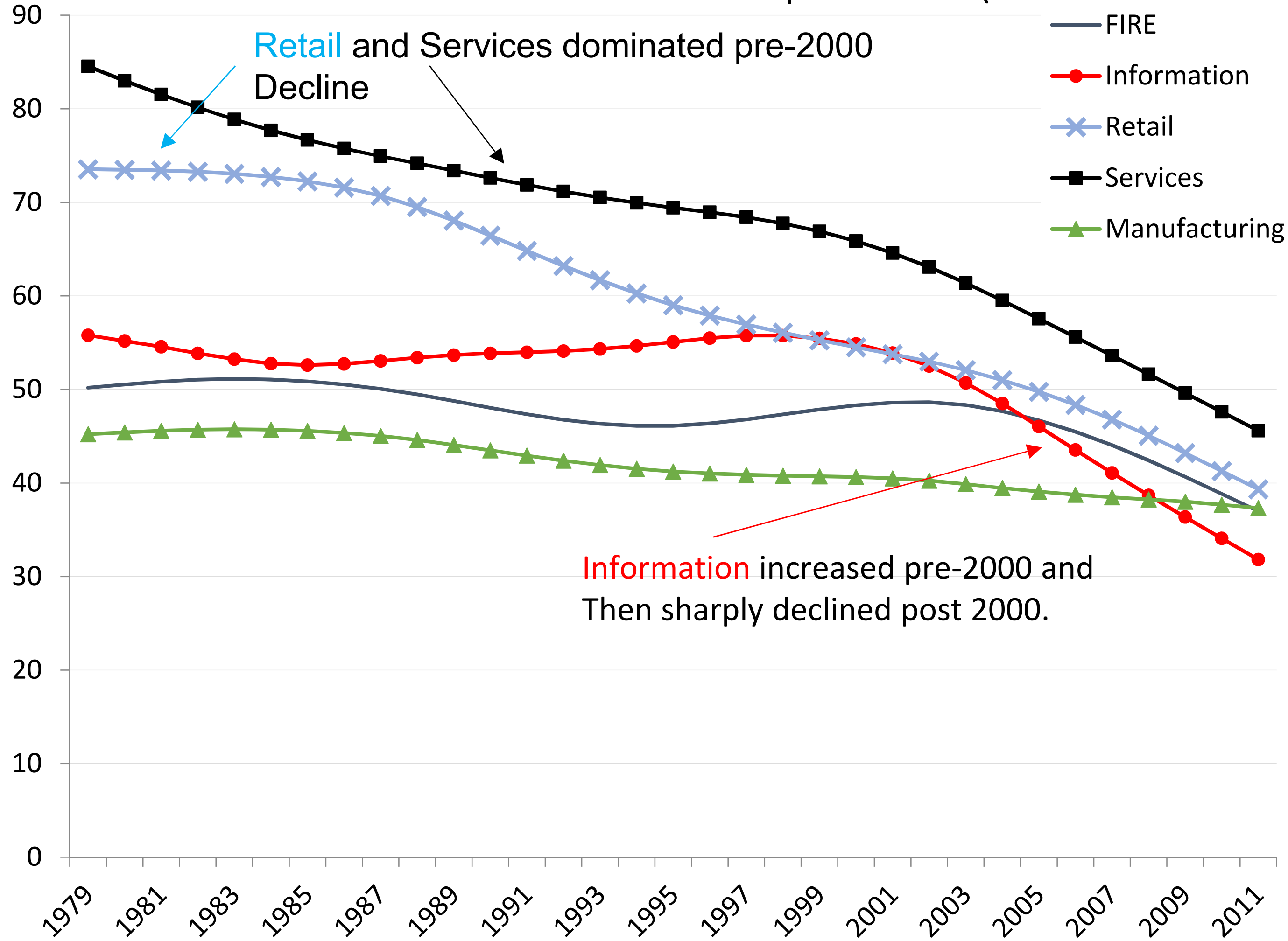
Decline in 90-50 is mimicked by decline in 90th percentile given 50th percentile is approximately zero – in other words, decline in 90-50 implies decline in high growth firms.



Decline in Overall Skewness is Both Composition Effect and Decline in Skewnes Among Age Groups (Young < 5 years old).

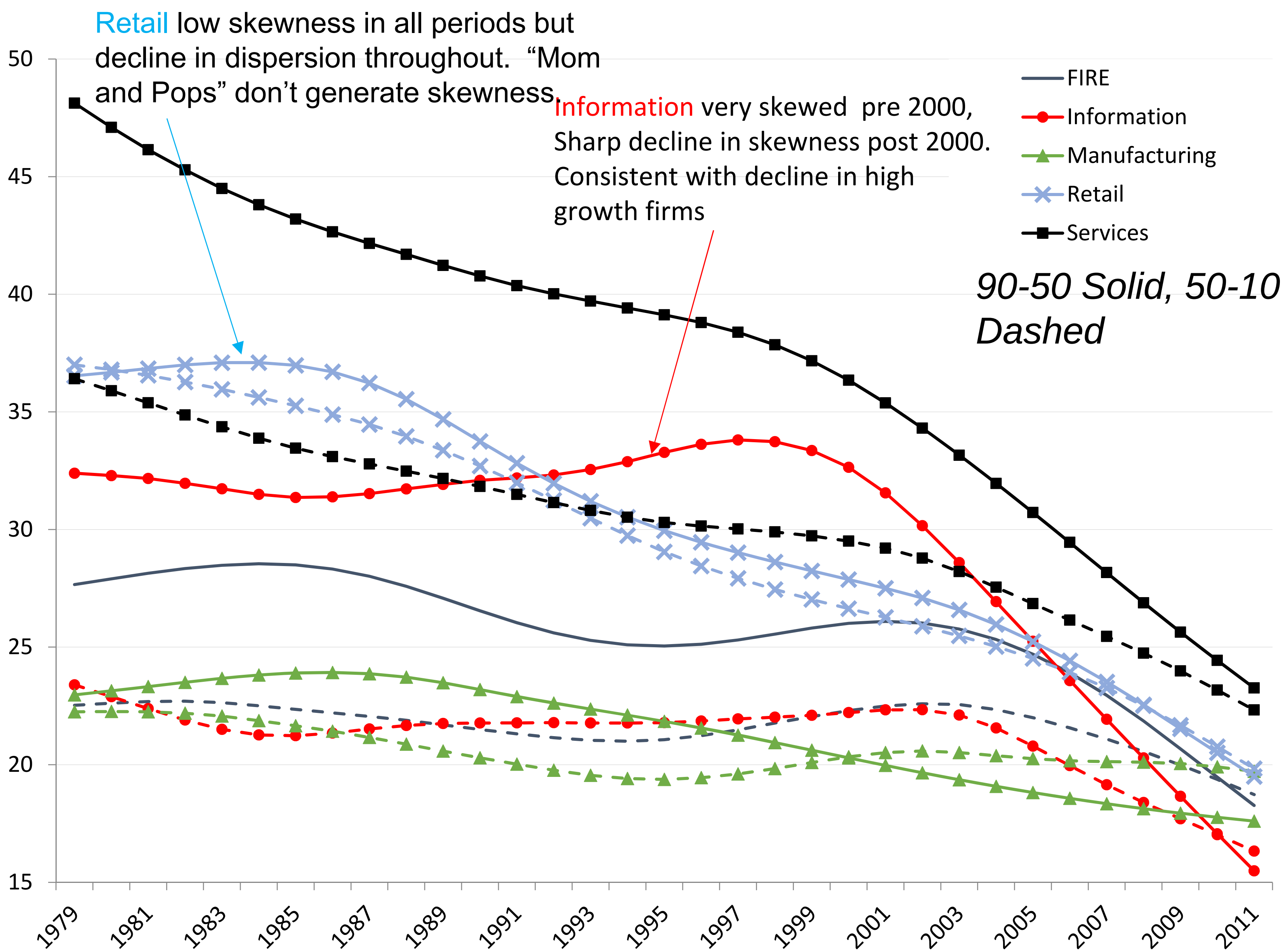


Sectoral Differences in Decline in Dispersion (90-10)



Showing HP Trends

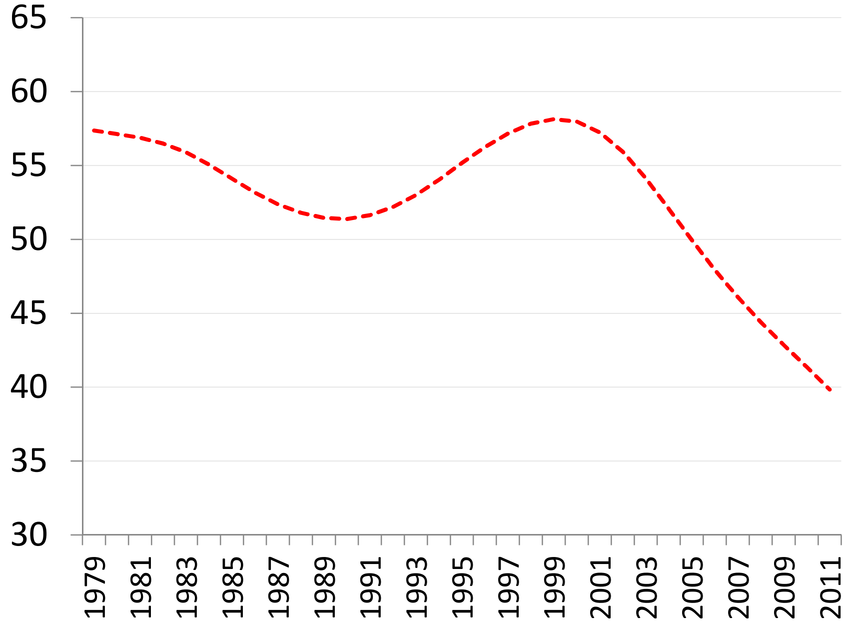
Large Sectoral Differences in Skewness in the Cross Section and Over Time



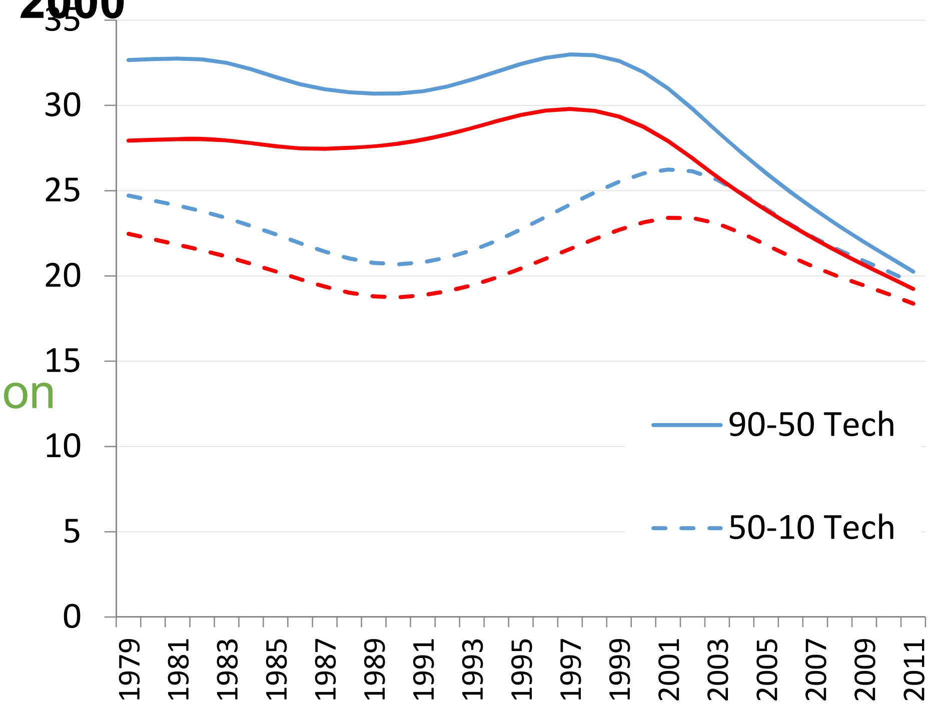
Differences for Information Sector Striking. But High Tech is Spread Across Numerous Broad Sectors including Information, Services and Manuf

NAICS Code	Industry
Information and Communications Technology (ICT) High-Tech	
3341	Computer and peripheral equipment manufacturing
3342	Communications equipment manufacturing
3344	Semiconductor and other electronic component manufacturing
3345	Navigational, measuring, electromedical, and control instruments manufacturing
5112	Software publishers
5161	Internet publishing and broadcasting
5179	Other telecommunications
5181	Internet service providers and Web search portals
5182	Data processing, hosting, and related services
5415	Computer systems design and related services
Miscellaneous High-Tech	
3254	Pharmaceutical and medicine manufacturing
3364	Aerospace product and parts manufacturing
5413	Architectural, engineering, and related services
5417	Scientific research-and-development services

90-10 Gap for High Tech shows sharp decline post 2000



Sharp Decline in 90-50 vs. 50-10 post 2000

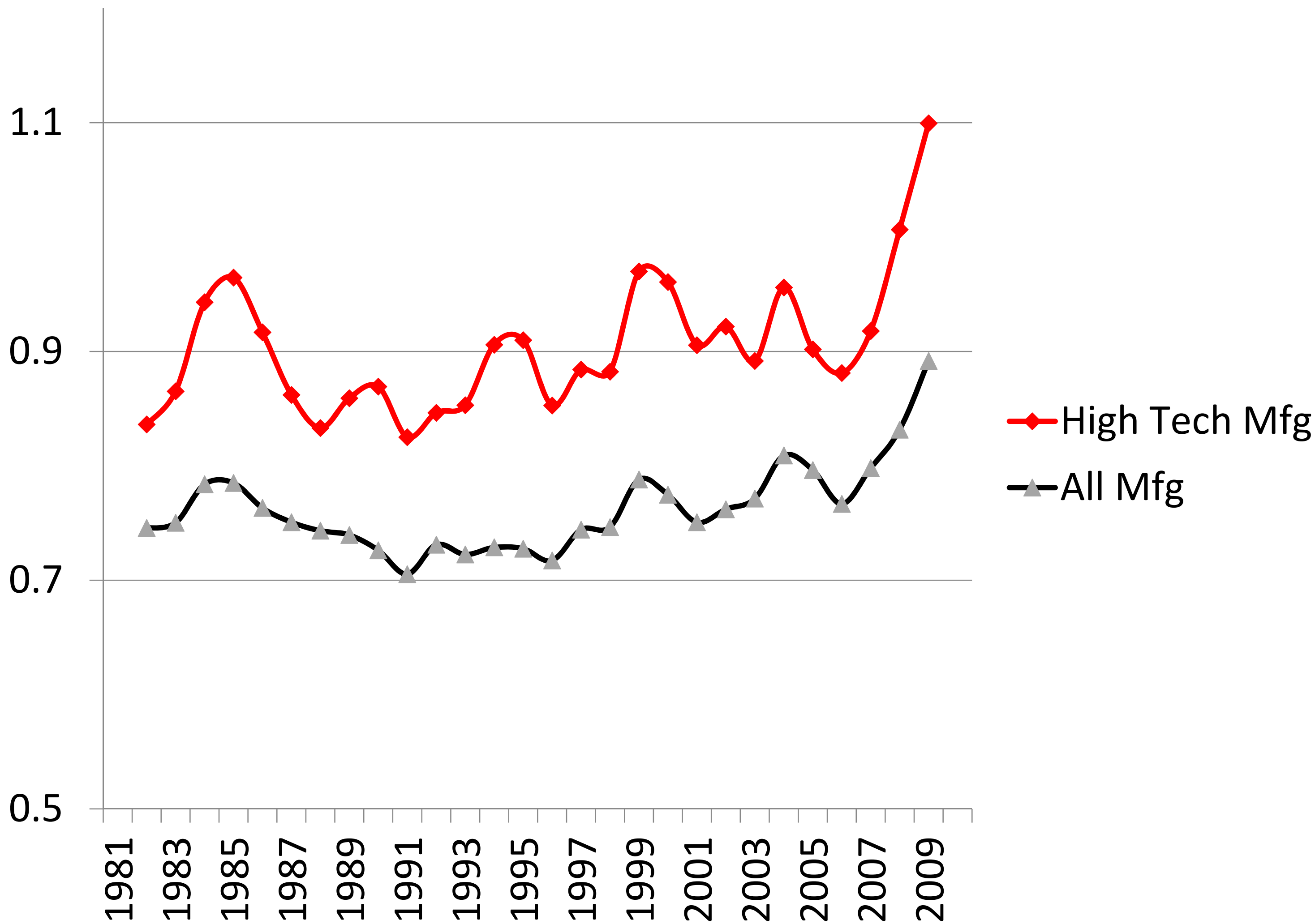


High Tech Sector Exhibits Rising Dispersion and Skewness in 1990s and then Sharp Declines Post 2000.

Decline in shocks or decline in responsiveness to shocks?

- Canonical firm dynamic models (e.g., Hopenhayn (1992), Hopenhayn and Rogerson (1993), Ericson and Pakes (1995)) imply decline should be from either:
  - A decline in the volatility of idiosyncratic shocks.
  - A decline in the response to such shocks.
- For high tech (manufacturing sector):
  - We find no evidence of a decline in the volatility of idiosyncratic shocks but a notable decline in the response to such shocks in the post 2000
  - This implies declining contribution of reallocation to productivity growth post 2000.

Within Industry Dispersion in TFP over time in High Tech Mfg vs. All Mfg (3-year MA, 90-10)

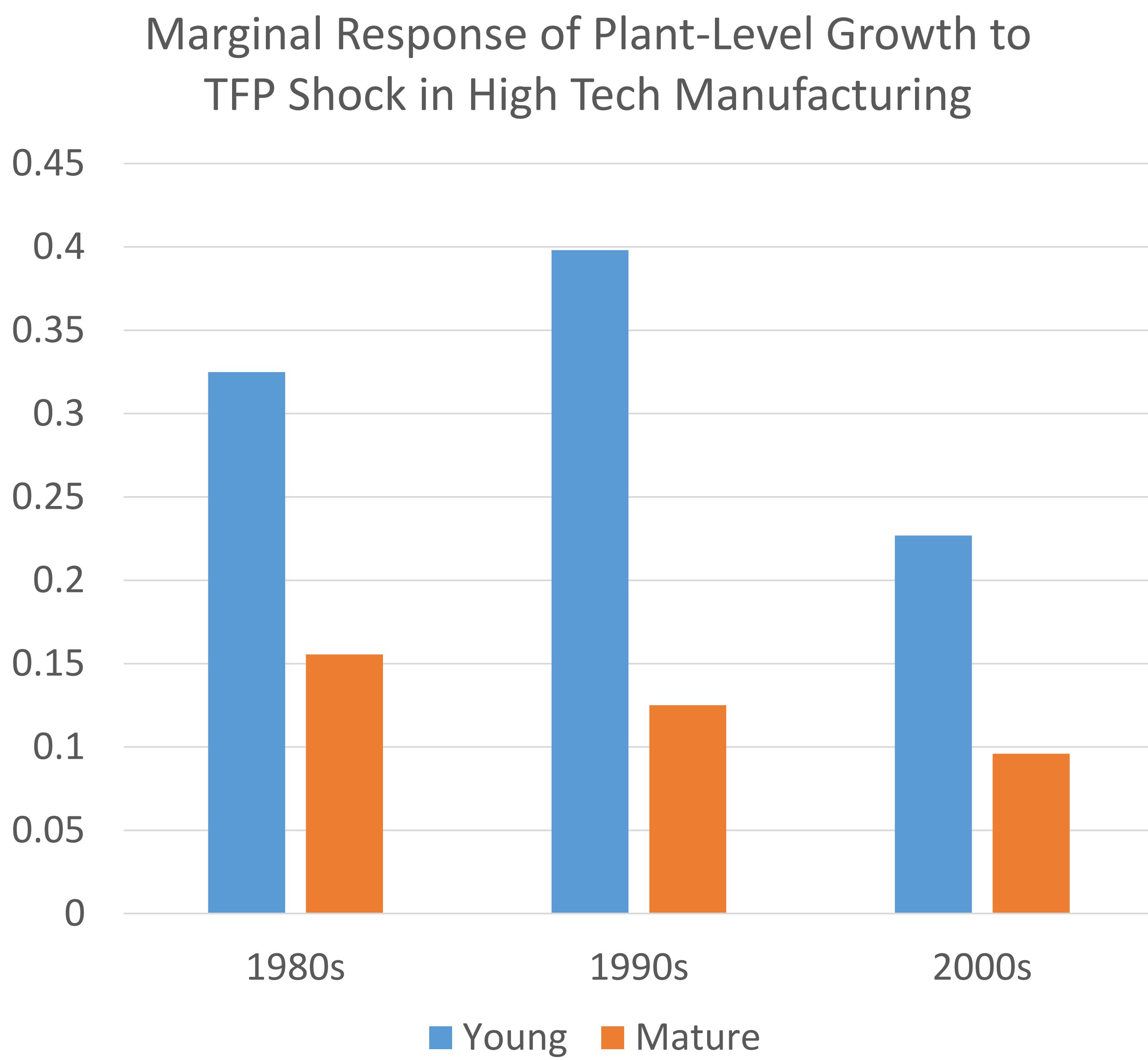


Analysis of Changing Response to Shocks

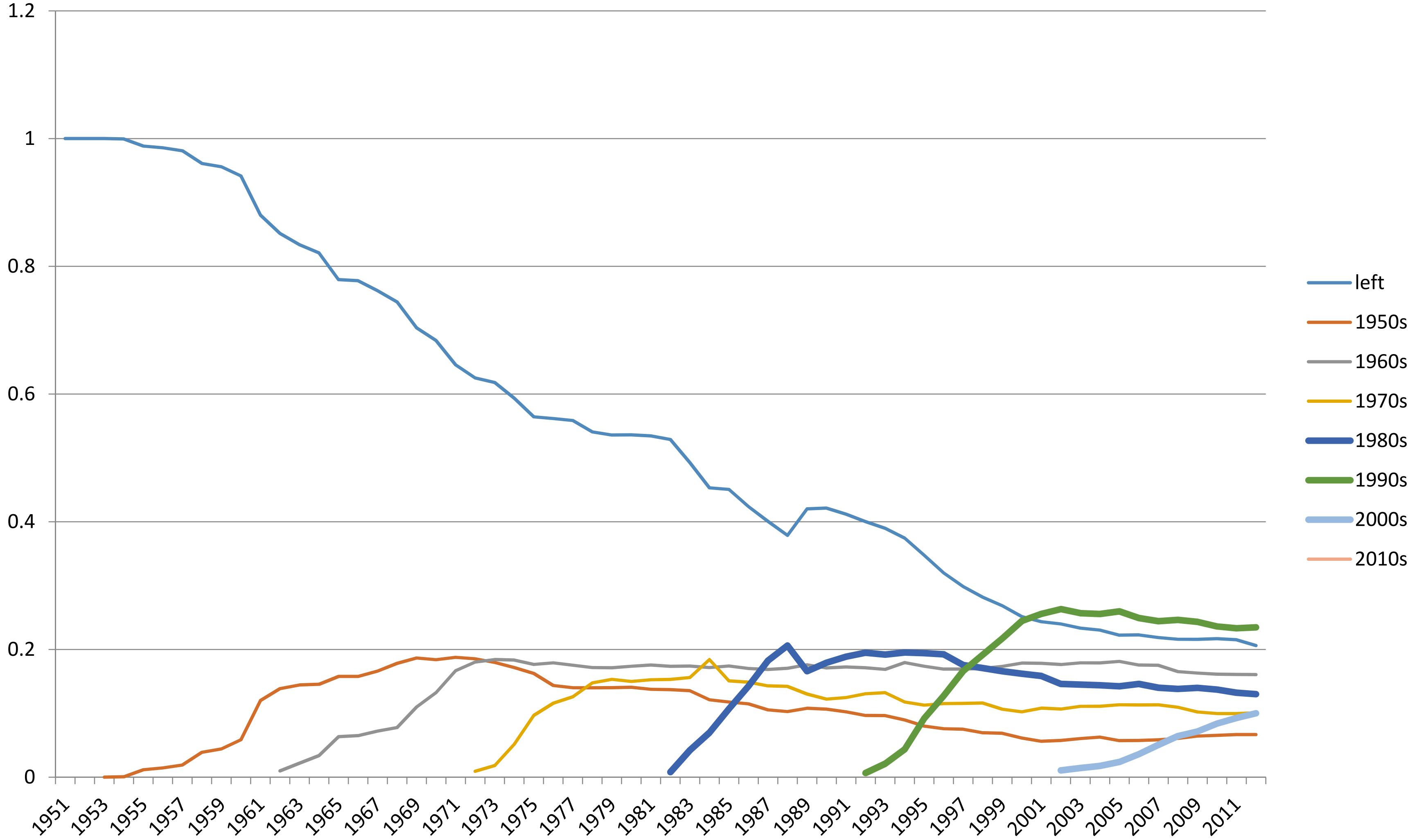
Estimating simple specifications such as:

$$Y_{e,t+1} = \lambda_{t+1} + \beta * TFP_{et} + \delta * TFP_{et} * Trend_{t+1} + X'_{et}\Theta + \varepsilon_{e,t+1}$$

e = establishment,
Y (outcome) = overall growth (or components) from *t* to *t*+1 from *t* to *t*+1
Trend = simple time trend with breaks allowed by decade (arbitrary).
TFP = log TFPR at the establishment level (deviated from industry*year mean)
X includes establishment and firm level controls, cyclical controls including interactions with TFP (FGH (2013))

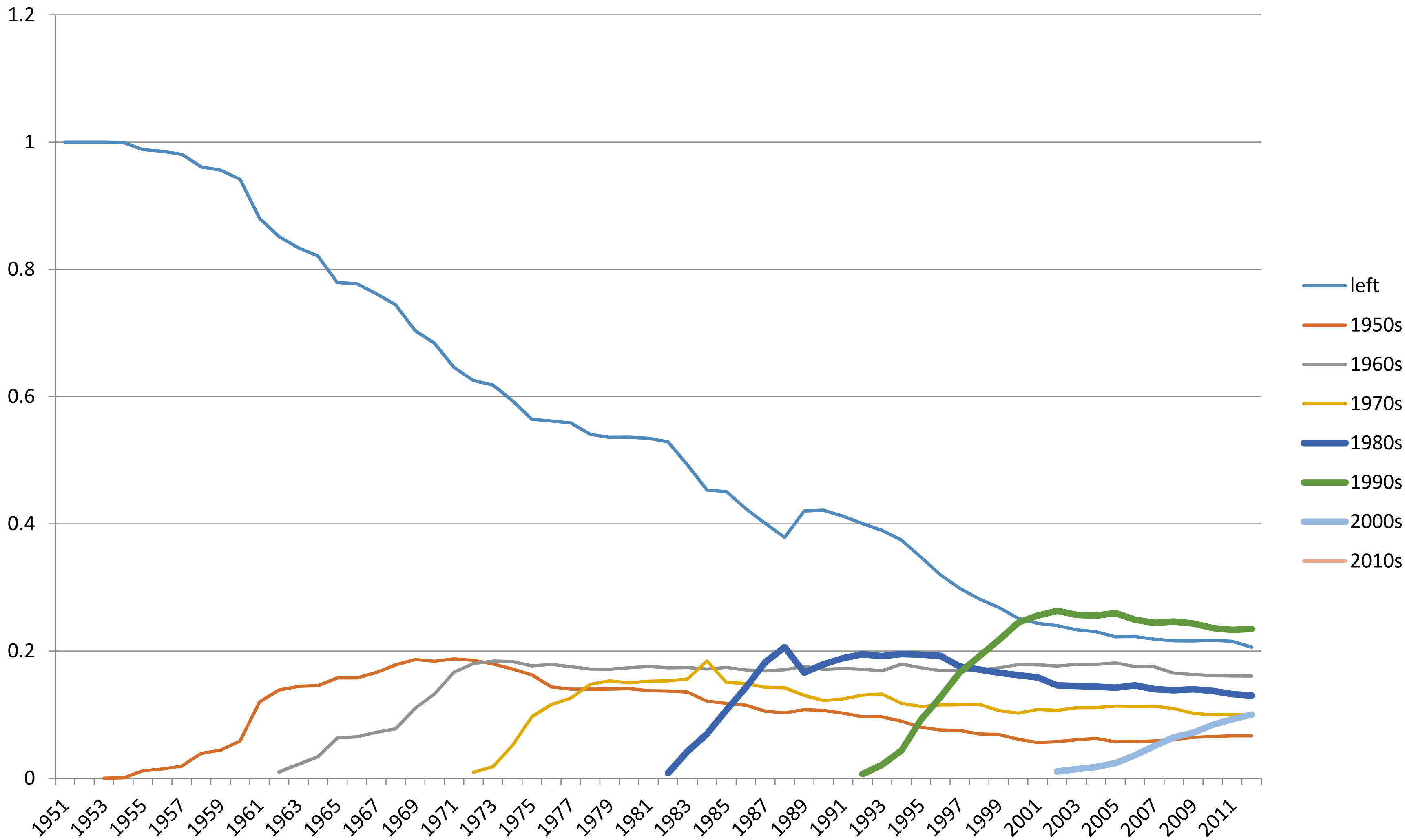

Marginal Response of Plant-Level Growth to
TFP Shock in High Tech Manufacturing
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1980s
1990s
2000s
Young
Mature

Employment shares by Cohorts of Publicly Traded Firms

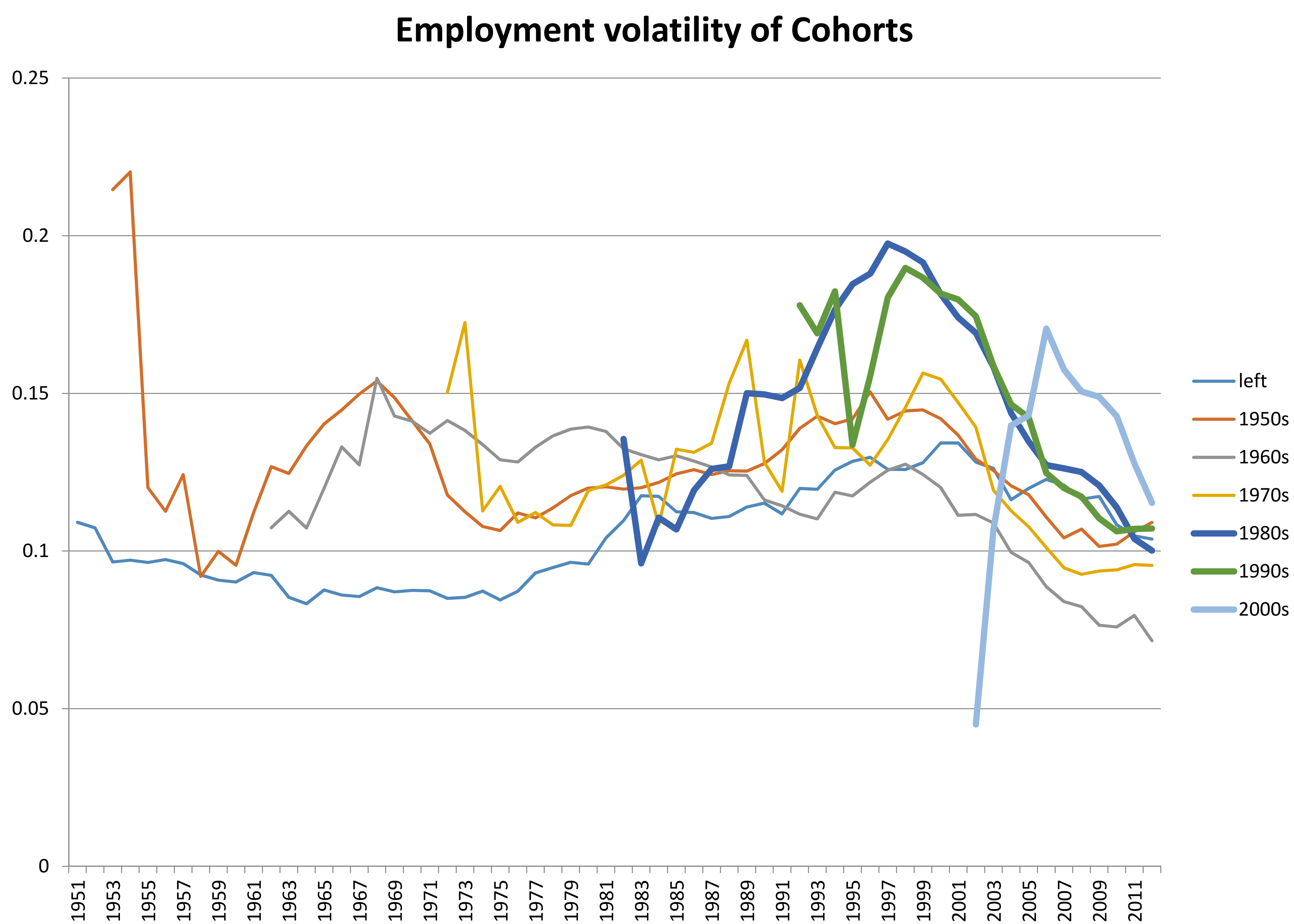


Source: COMPUSTAT – LBD shows similar patterns post 1980

Employment shares by Cohorts of Publicly Traded Firms



Source: COMPUSTAT – LBD shows similar patterns post 1980



Employment volatility of Cohorts
0.25
0.2
0.15
0.1
0.05
0
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
left
1950s
1960s
1970s
1980s
1990s
2000s

NUTS AND BOLTS of PRODUCTIVITY MEASUREMENT

Measurement of Plant-level Productivity

$$tfp_i = y_i - \alpha_l l_i - \alpha_k k_i - \alpha_m m_i - \alpha_e e_t$$

All variables in logs, difficult measurement Issues on outputs and inputs and factor elasticities

Typical to assume Cobb-Douglass or to have Divisia index approach approximation

Measurement issues

- Factor inputs:
  - Labor quality
  - Capital stock (book value vs. perpetual inventory). Typically do not have by asset class. Very different from aggregate measurement. Open question as to how problematic.
- Factor elasticities:
  - Cost shares, estimated elasticities using OLS, IV, proxy methods
  - All typically estimate factor elasticities at the industry level
    - Time invariant with estimated approach typically given Cobb-Douglass assumptions
  - Estimates vary in literature but measures of TFP highly correlated across these methods. Other issues (below) appear to matter more.
- Plant-level heterogeneity in output and input prices
- Plant-level heterogeneity in factor elasticities

Details matter

- For cost share approach, two different methods:
  - Shares of revenue vs. shares of total costs.
  - Shares of revenue requires profit maximization, first order conditions hold, perfect competition, CRS
  - Shares of total costs requires cost minimization, first order conditions hold, CRS.
    - Advantage of this method is that even with imperfect competition this method yields production elasticities. Can also be used without CRS but need to pin down RTS with some other method.
  - Often take averages across plants over time or within industry, so first order conditions hold on average.
- Estimation methods don't require so many assumptions but are limited by two key issues:
  - Typically do not observe y but rather p*y (revenue). Deflate by industry deflator. So estimating revenue function. NOT factor elasticities if prices are endogenous with firms facing downward sloping demand curve. Difficult to recover production elasticities without much more structure.
  - Proxy methods use high order polynomial approximations. These are sensitive to measurement error.

Example of proxy method

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + \omega_{jt} + \eta_{jt} \tag{24}$$

$$i_{jt} = i(k_{jt}, a_{jt}, \omega_{jt}, \Delta_t) = i_t(k_{jt}, a_{jt}, \omega_{jt}). \tag{27}$$

$$\omega_{jt} = h_t(k_{jt}, a_{jt}, i_{jt}). \tag{28}$$

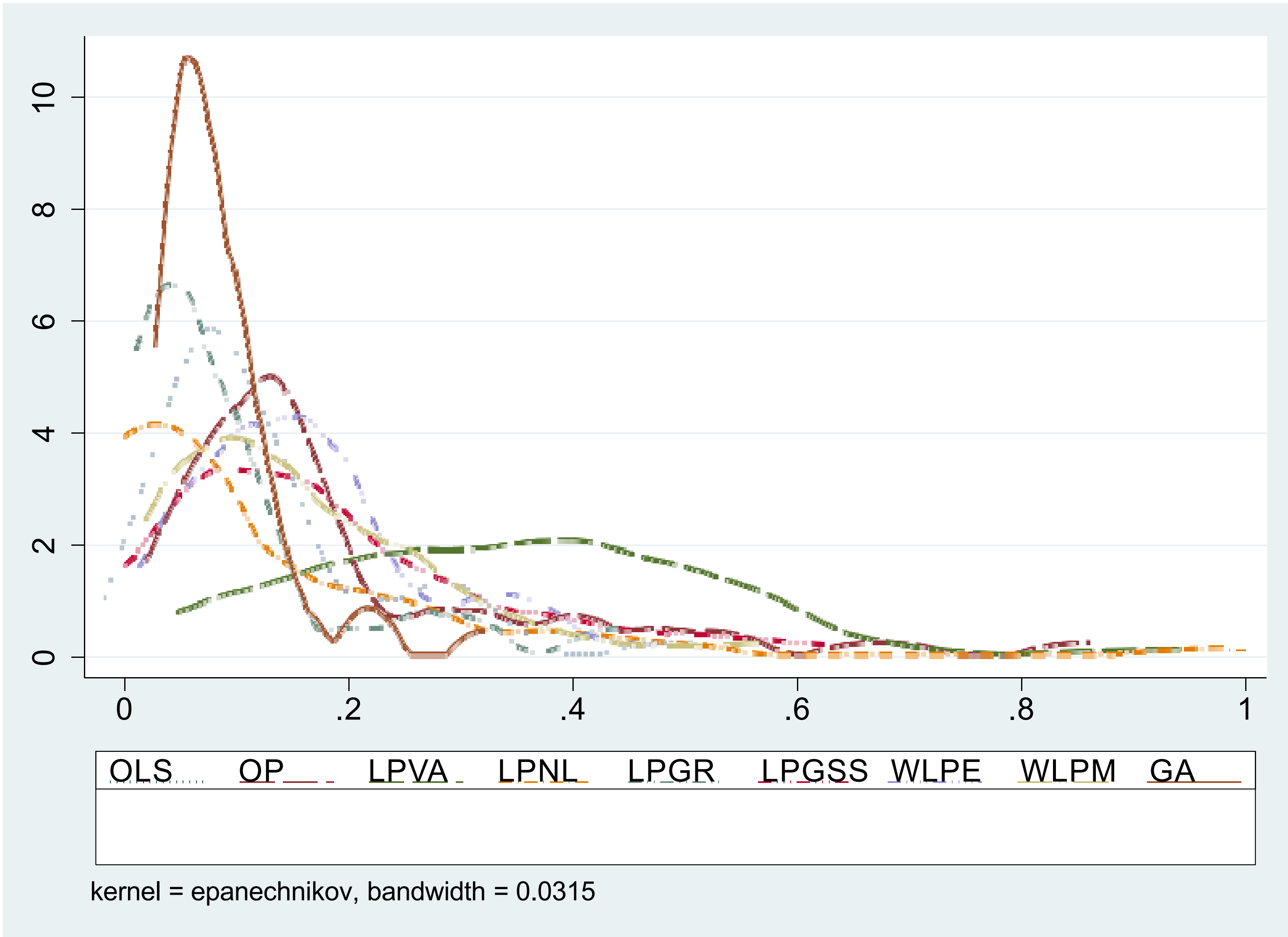
$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + h_t(k_{jt}, a_{jt}, i_{jt}) + \eta_{jt}. \tag{29}$$

$$y_{jt} - \beta_l l_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \omega_{jt} + \eta_{jt}. \tag{33}$$

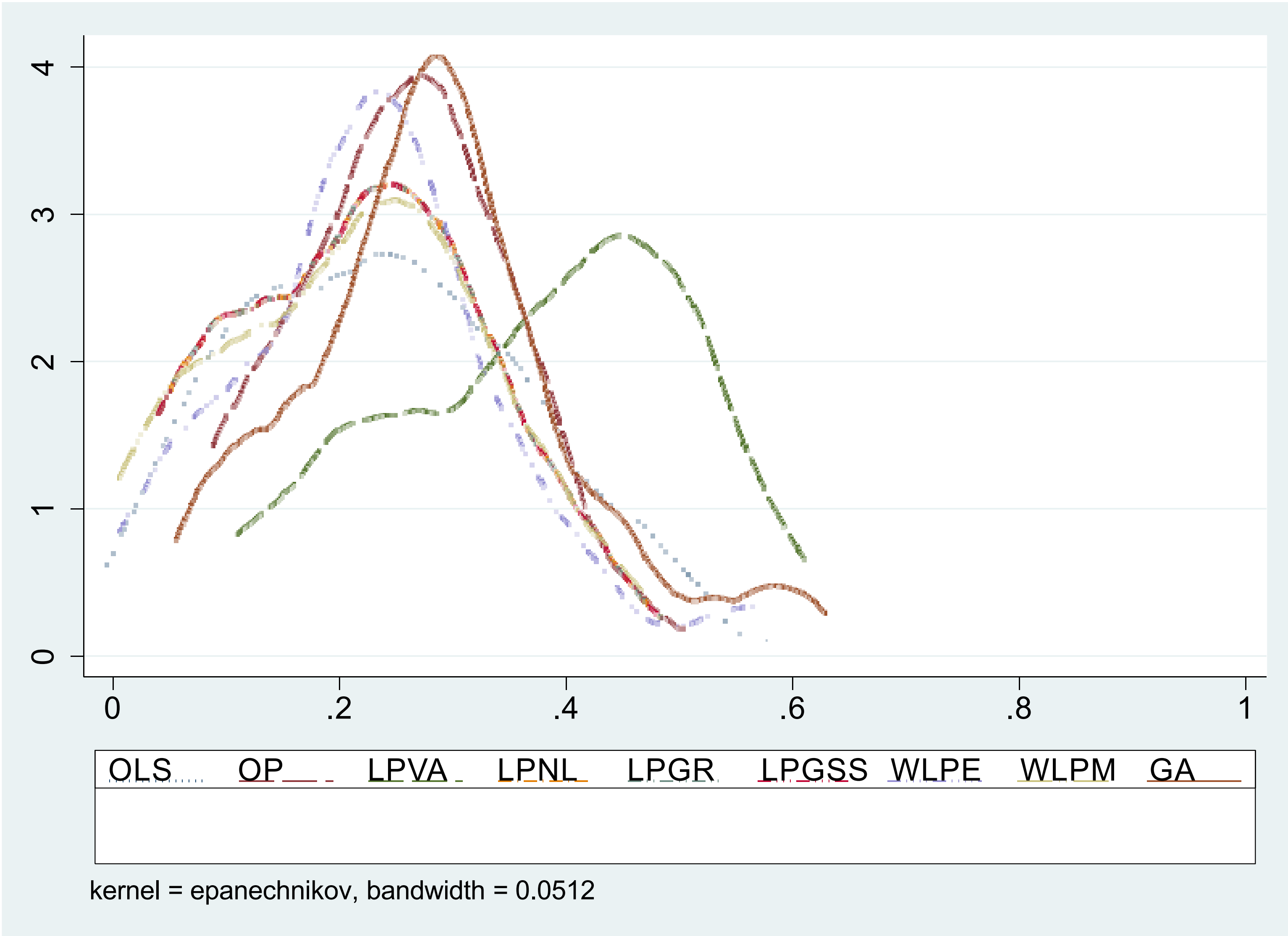
$$\begin{aligned} y_{jt} &- \beta_l l_{jt} \\ &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + g(\omega_{jt-1}) + \xi_{jt} + \eta_{jt} && \text{(34a)} \\ &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + g(\phi_{jt-1} - \beta_0 - \beta_k k_{jt-1} - \beta_a a_{jt-1}) + \xi_{jt} + \eta_{jt} \\ &= \beta_k k_{jt} + \beta_a a_{jt} + \tilde{g}(\phi_{jt-1} - \beta_k k_{jt-1} - \beta_a a_{jt-1}) + \xi_{jt} + \eta_{jt}, && \text{(34b)} \end{aligned}$$

Depends critically on the invertibility amongst other assumptions

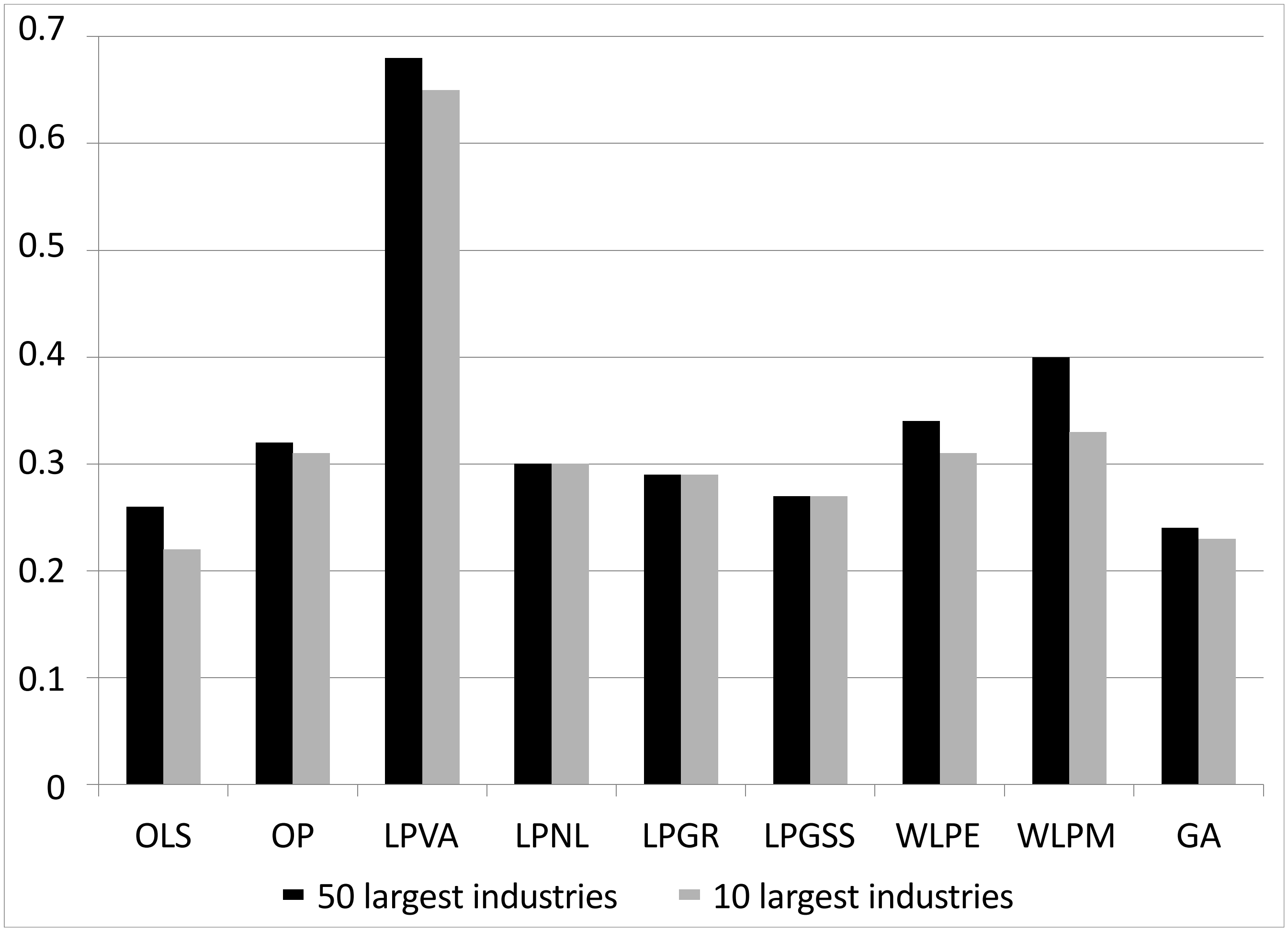
Factor Elasticity of Capital



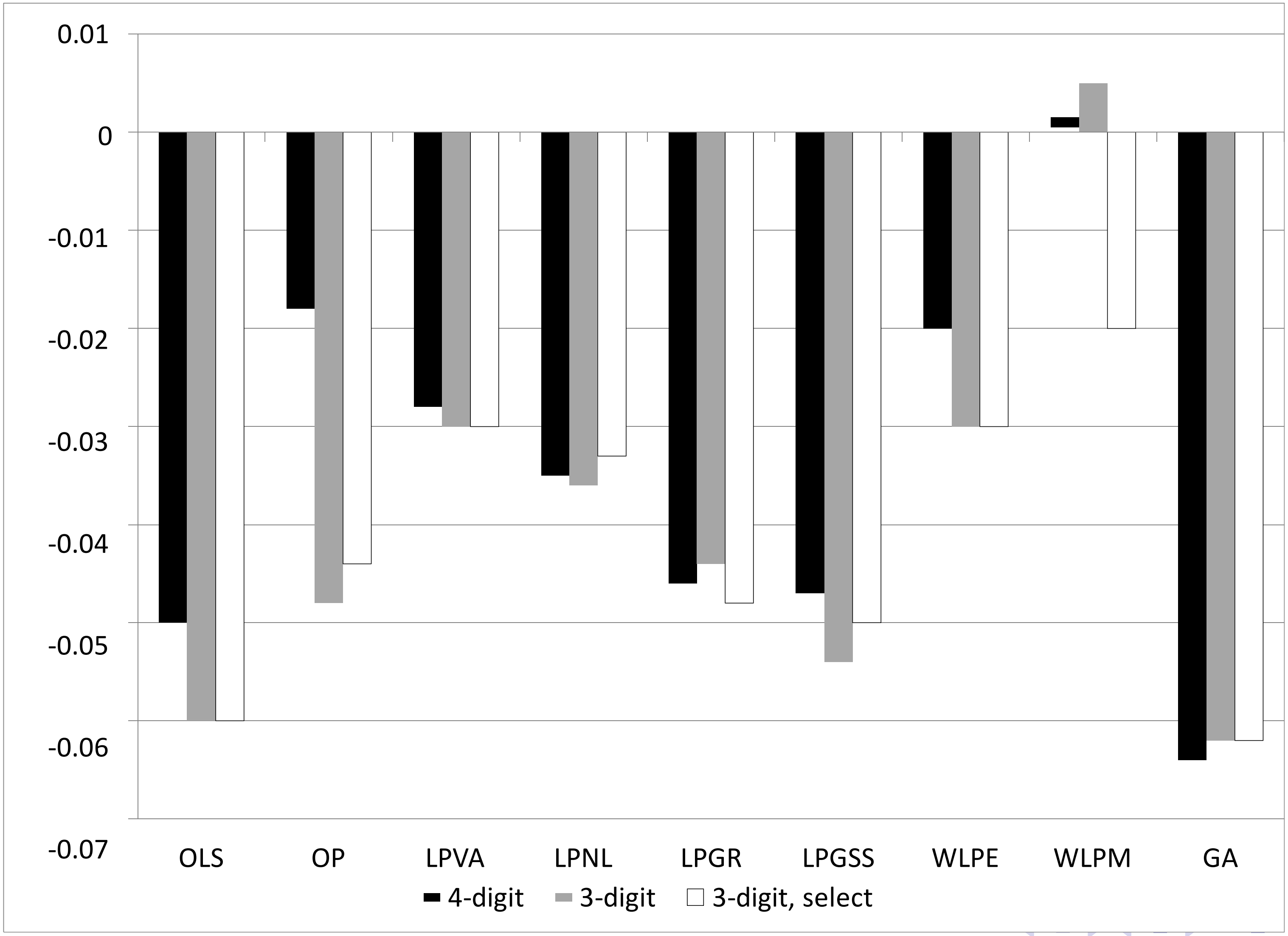
Factor Elasticity of Labor



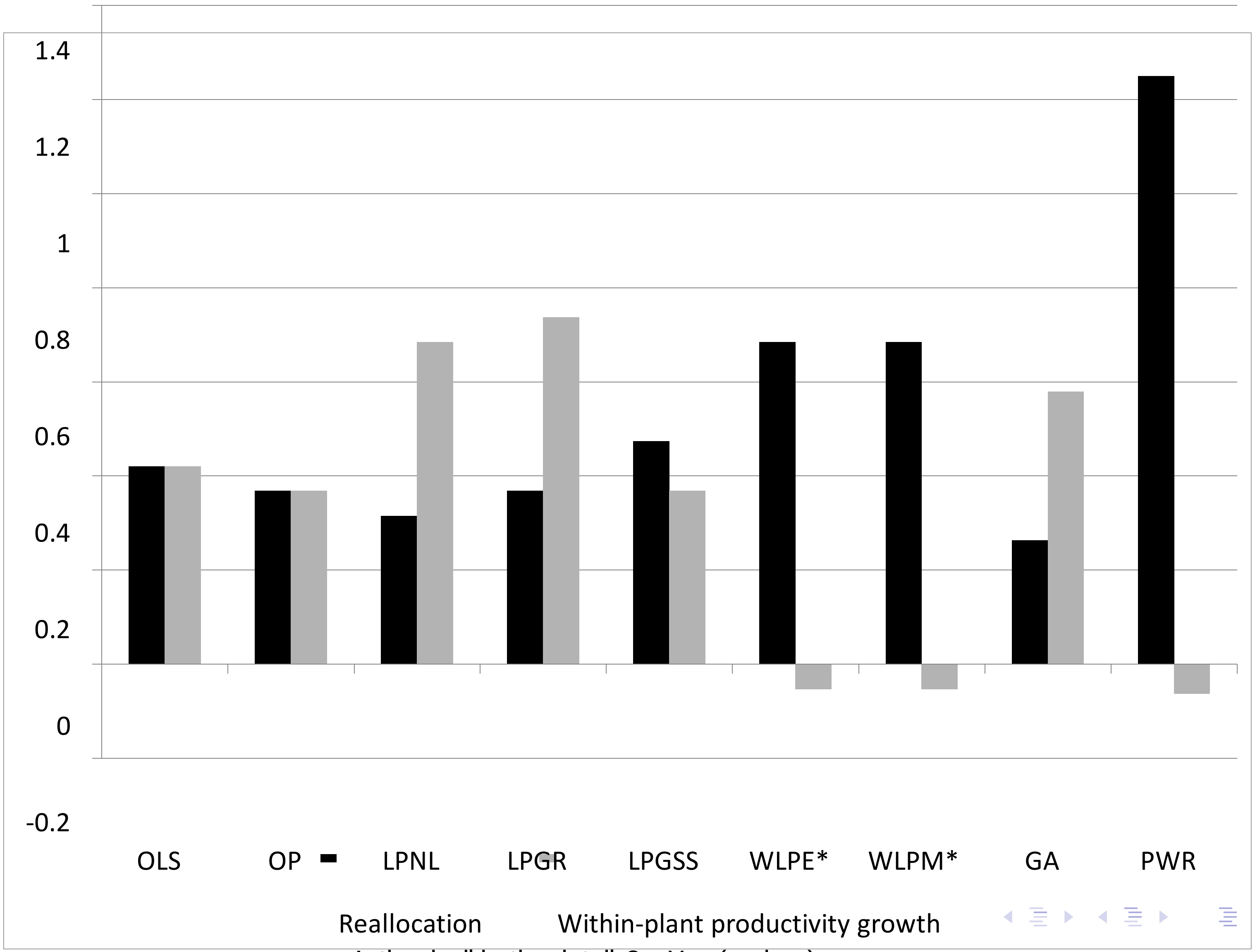
TFPR Dispersion (IQR)



Marginal effect of Productivity on Exit



Contribution of Reallocation to Productivity Growth – Structural Decomposition



More Basic Measures of Productivity Are Often Used

- Labor productivity Measures at the Establishment (or Firm level)
  - Real Value Added Per Worker

$$RLP_{et} = (VA_{et} / TE_{et}) = (Y_{et} - M_{et}) / TE_{et}$$

Where Y_{et} = Real Gross Output
M_{et}= Real Materials (including energy)
Te_{et}= Total Employment
Use detailed industry output and material price deflators

Often best available measure is real gross output per worker – comparable within industries

Dispersion in Productivity – How should we think about this?

$$Y = AK^{\alpha}L^{1-\alpha}$$

$$Y = \frac{D^{\varepsilon-1}}{P^{\varepsilon}}$$

$$\frac{rK}{wL} = \frac{\alpha}{1-\alpha}$$

$$APK = \frac{PY}{K} = \frac{\varepsilon}{\varepsilon-1}\frac{r}{\alpha}$$

$$APL = \frac{PY}{L} = \frac{\varepsilon}{\varepsilon-1}\frac{w}{1-\alpha}$$

$$TFPR = PA = \frac{\varepsilon}{\varepsilon-1}\left(\frac{w}{1-\alpha}\right)^{1-\alpha}\left(\frac{r}{\alpha}\right)^{\alpha}$$

$$TFPQ = A$$

Cobb-Douglas Technology, CRS
Isoelastic Demand, No Frictions,
Price takers in factor markets

No dispersion in factor cost share ratio, Revenue average product of capital, revenue average product of labor, TFPR

Even though there is dispersion In TFPQ

Why is there so much dispersion in productivity across businesses in narrowly defined sectors?

- Background facts:
  - Interquartile range of log of Revenue TFP (TFPR) is 0.29
  - Interquartile range of log of Revenue Labor Productivity (RLP) is 0.65
  - Dispersion in TFPQ, TFPR, and output price within narrow product classes (7-digit) in U.S. (Source: FHS (2008)):
    - Std. Dev of log(TFPQ) is: 0.26
    - Std. Dev of log(TFPR) is: 0.22
    - Std. Dev of log(RLP) is: 0.65
    - Std. Dev of log(P) is: 0.18
    - Std. Dev of log(Q) is: 1.05
    - Corr(log(TFPQ),log(P)) is: -0.54
    - Corr(log(TFPQ),log(Q)) is: 0.28
    - Corr(log(TFPQ),log(TFPR)) is: 0.75
    - Corr(log(TFPQ),log(RLP)) is: 0.56

Frictions + Distortions

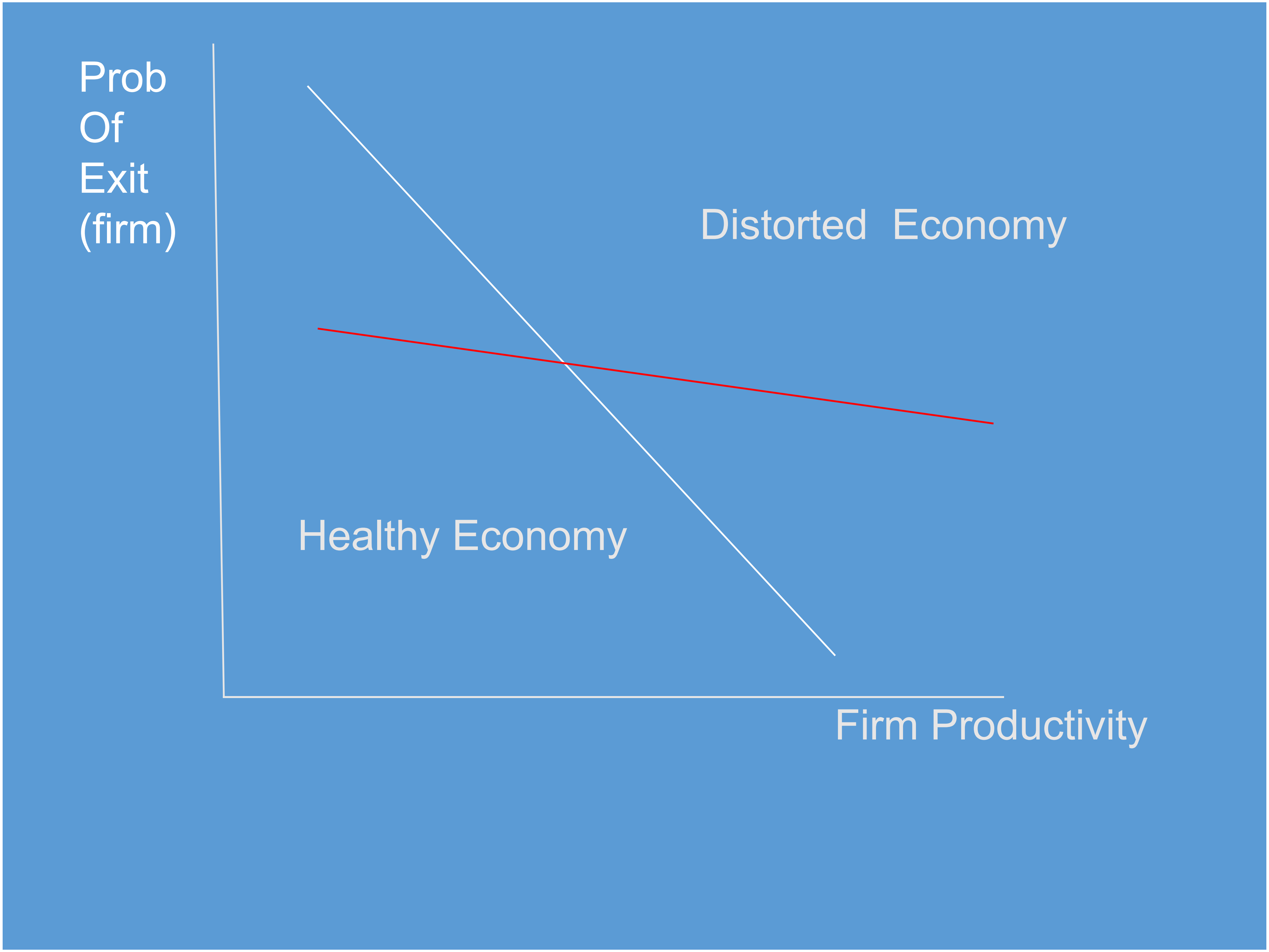
- Costs of Entry (and exit)
  - Including costs of entering new markets
  - Hopenhayn (1992), Melitz (2003), Melitz and Ottaviano (2005)
- Learning (initial conditions and after changing products/processes)
  - Jovanovic (1982) and Ericson and Pakes (1998)
  - Experimentation
- Adjustment costs for factors of production (capital, labor, intangible capital)
  - Convex vs. Nonconvex
- Economies of scope and control
- Product Differentiation:
  - Horizontal (e.g., spatial) vs. Vertical
- Output and input price dispersion and determination
- Imperfections in product, labor, capital, credit markets
- Distortions to all of the above + market institutions
  - Idiosyncratic distortions as in Banerjee and Duflo (2003), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Bartelsman, Haltiwanger and Scarpetta (2013)

What frictions matter the most?

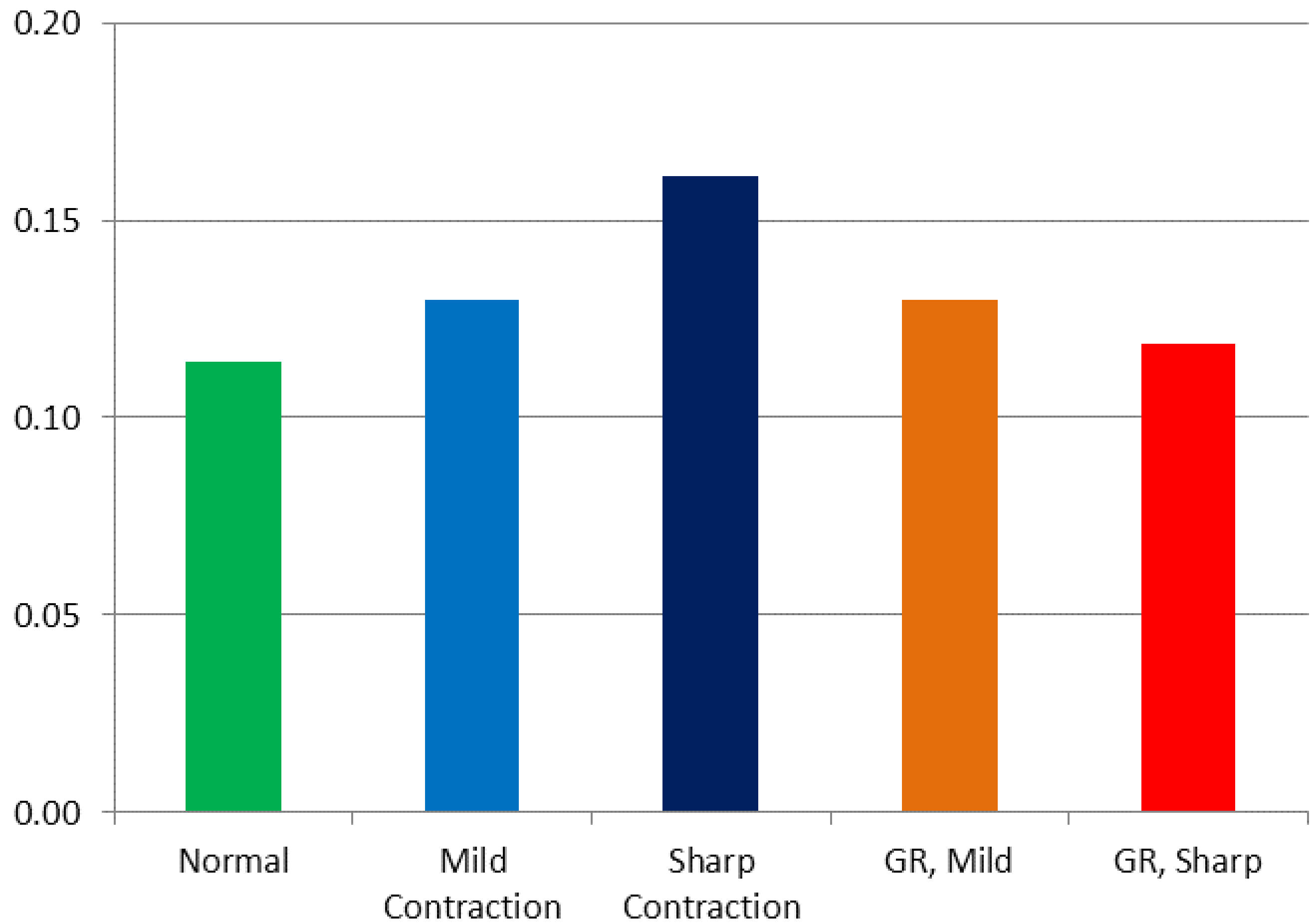
- Many studies showing evidence of entry costs, labor adjustment costs, capital adjustment costs, trade costs, product differentiation, and so on.
- Many open questions and issues:
  - Not practical to include all frictions in all models – but caution about identification since we are all using same data
  - How do frictions vary across advanced vs. emerging vs. transition?
- Important to distinguish between those frictions that yield some plants persistently higher productivity than others as opposed to adjustment dynamics

Lots of margins for distortions…

- Cross sectional misallocation
- Dynamic distortions:
  - Startups
  - Post-entry up or out dynamic
  - Creative Destruction
- Secular vs. Cyclical Distortions


Prob
Of
Exit
(firm)
Distorted Economy
Healthy Economy
Firm Productivity

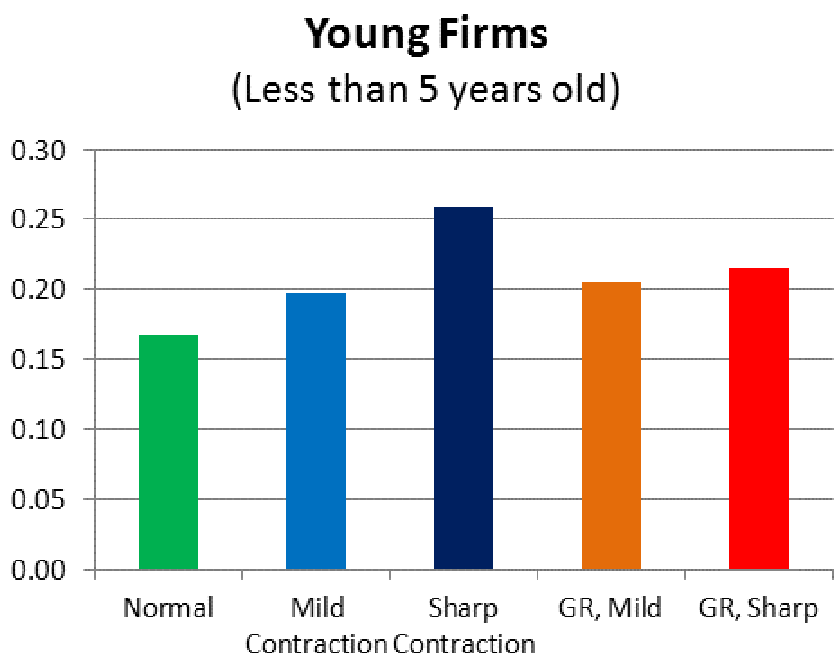
Differences in Overall Growth Rates Over the Business Cycle:
High and Low Productivity Establishments

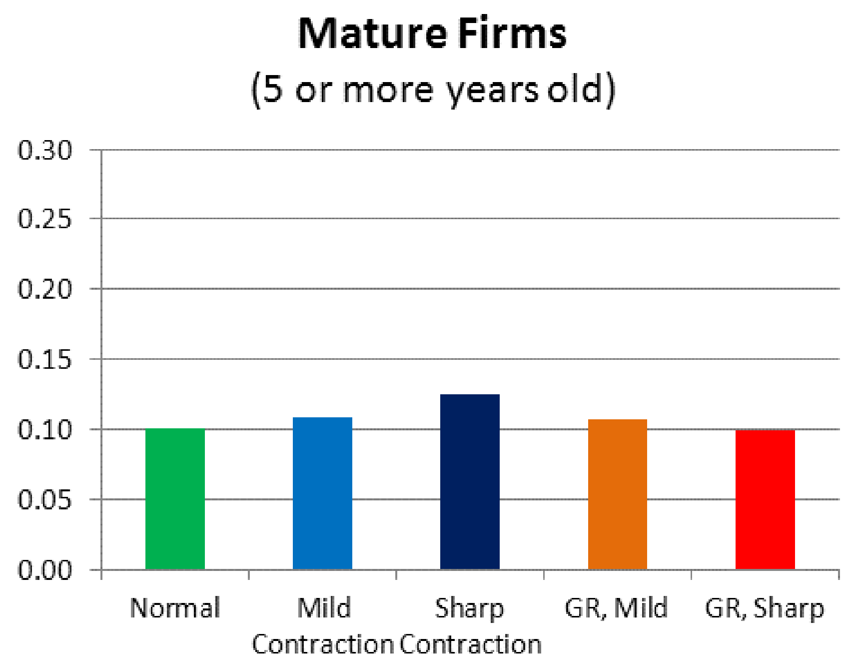


Normal is Zero Change in Unemployment, Mild is 0.01 Change, Sharp is 0.03 Change.
High Productivity is 1 std dev above mean, Low Productivity is 1 std dev below mean.

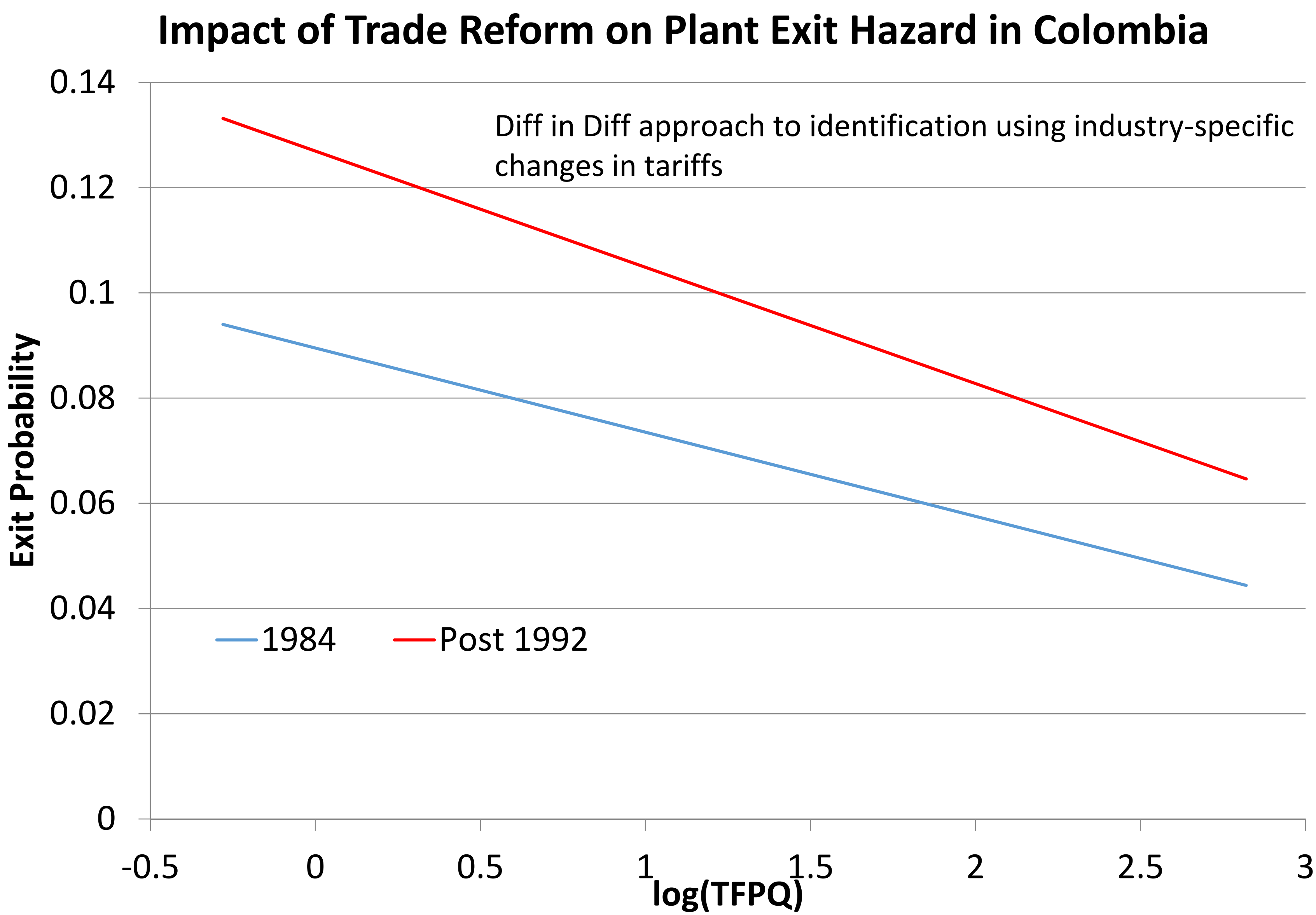
Foster, Grim and Haltiwanger (2013)

Differences in Overall Growth Rates Over the Business Cycle: High and Low Productivity Establishments

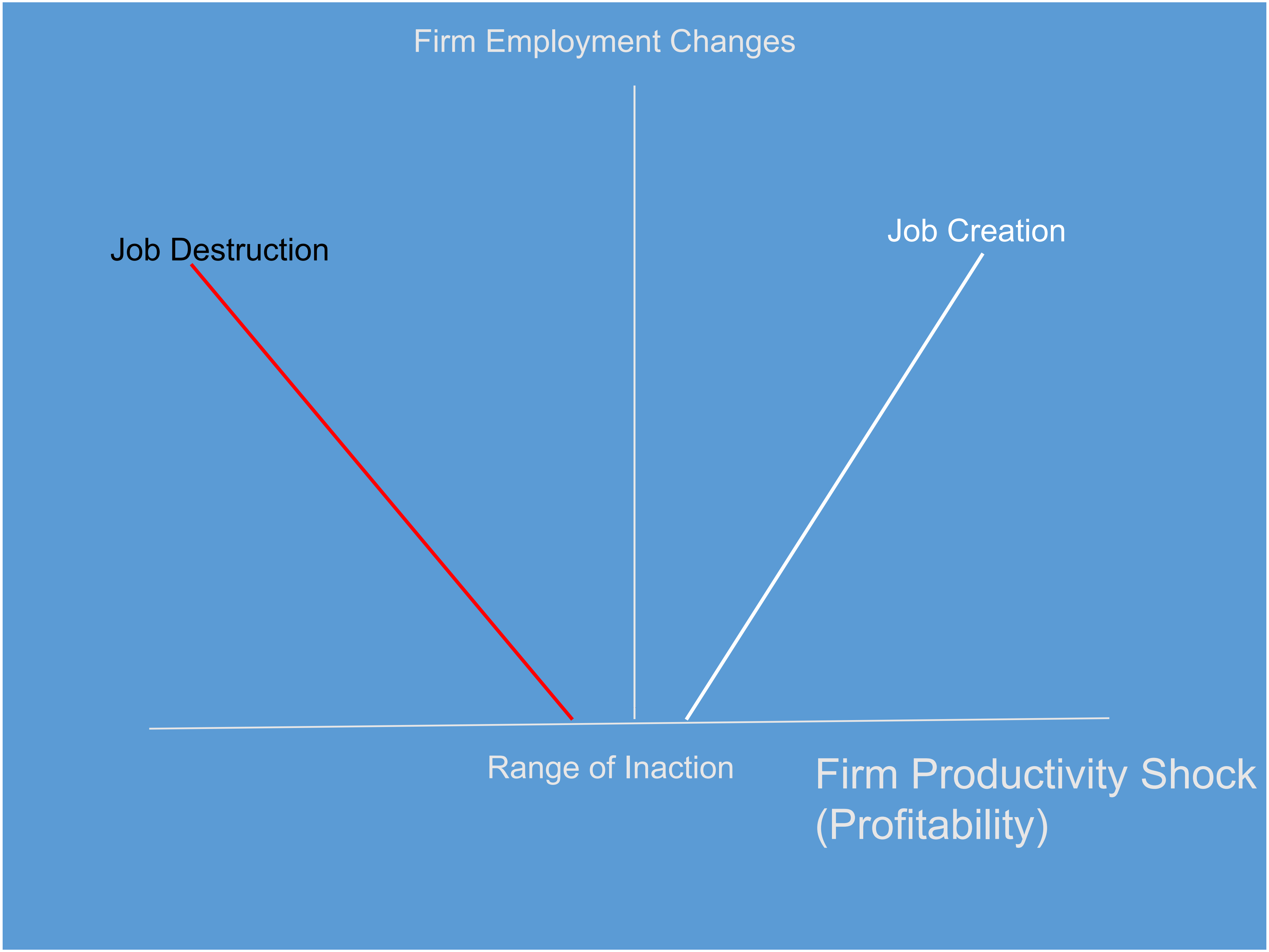


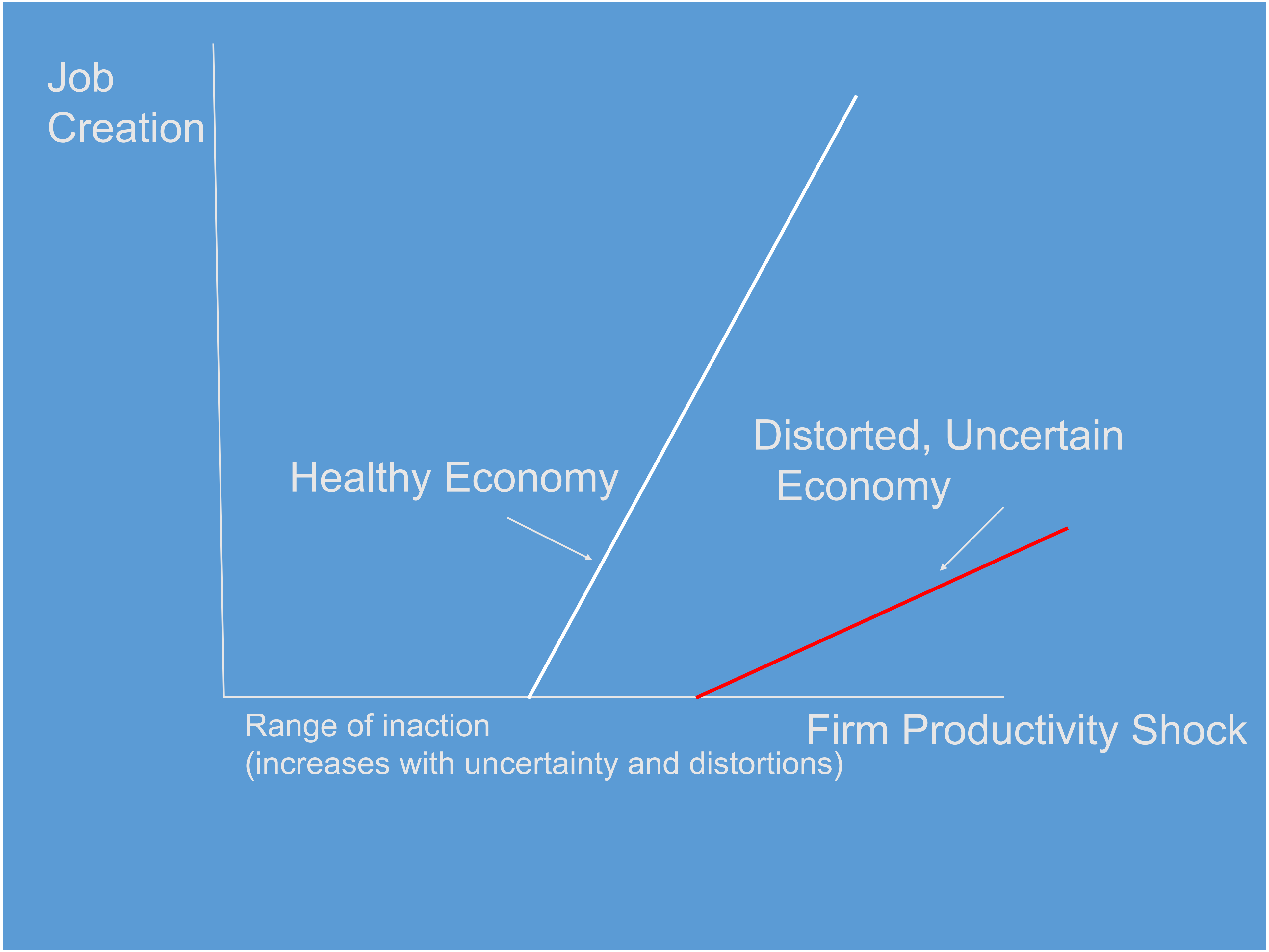


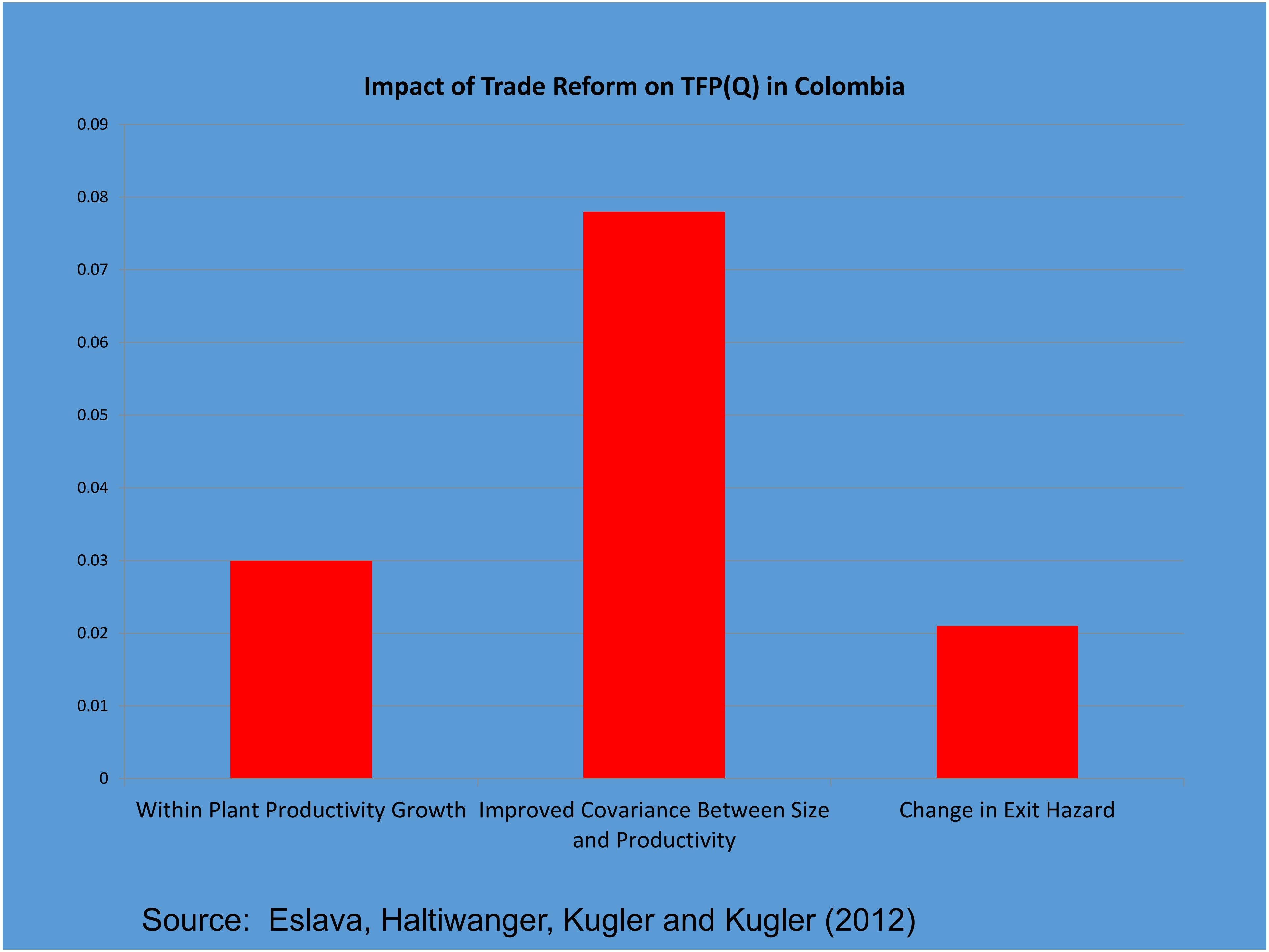
Normal is Zero Change in Unemployment, Mild is 0.01 Change, Sharp is 0.03 Change.
High Productivity is 1 std dev above mean, Low Productivity is 1 std dev below mean.



Source: Eslava, Haltiwanger, Kugler and Kugler (2012)


Firm Employment Changes
Job Destruction
Job Creation
Range of Inaction
Firm Productivity Shock
(Profitability)


Job
Creation
Healthy Economy
Distorted, Uncertain
Economy
Range of inaction
(increases with uncertainty and distortions)
Firm Productivity Shock


Impact of Trade Reform on TFP(Q) in Colombia
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Within Plant Productivity Growth
Improved Covariance Between Size and Productivity
Change in Exit Hazard
Source: Eslava, Haltiwanger, Kugler and Kugler (2012)

Taking Stock

- High pace of churning of businesses within narrowly defined industries
- Startups and young businesses play an important role in these dynamics
- Up or out dynamics
- These dynamics connected to productivity (and demand) dynamics at the micro level
- Identifying the frictions and how they vary across industry, time, and country ongoing activity
- But what about before entry?

How Do We measure the CONTRIBUTION OF REALLOCATION?

Size/productivity relationship within industries

$$\Omega_t = \sum_i s_{it}\omega_{it}$$
$$= (1/N_t)\sum_i \omega_{it} + \sum_i (s_{it} - (1/N_t)\sum_i s_{it})(\omega_{it} - (1/N_t)\sum_i \omega_{it}))$$

Olley and Pakes (1996) decomposition

$$\Delta\Omega_t = \sum_i s_{it}\omega_{it} - \sum_i s_{it-1}\omega_{it-1}$$
$$= \sum_{i\in C} \overline{s_{it}}\Delta\omega_{it} + \sum_{i\in C} \Delta s_{it}\overline{(\omega_{it}} - \overline{\Omega_t}) + \sum_{i\in N} s_{it}\left(\omega_{it} - \overline{\Omega_t}\right) - \sum_{i\in X} s_{it-1}\left(\omega_{it-1} - \overline{\Omega_t}\right)$$
$$= within \quad + \quad reallocati\,on \quad + \quad entry \quad - \quad exit$$

Modified Baily, Hulten and Campbell (1992) and Griliches and Regev (1995) decomposition

Comments on Decomposition in Literature

- Some questions about how to interpret industry-level index defined in this manner
  - Typical check (e.g., BHC and FHK) to see how this index performs relative to standard aggregate ***industry*** measures
    - Common result – magnitudes very similar and correlations high in most studies
    - Cautions:
      - These measures very sensitive to measurement error since depend on measuring within industry productivity (log) level dispersion accurately
      - Not appropriate for decompositions that exploit between industry variation (measurement and index problems)
  - Standard decomposition summarizes changes in activity weighted micro distribution
  - These decompositions can be used as moments to match in a calibration or indirect inference approach (see, e.g., Bartelsman, Haltiwanger and Scarpetta (2009)
  - These decompositions even with crude measures of productivity (like labor or capital productivity) may be more robustly measured than more structural decompositons.

Comments on Decomposition in Literature

- Decompositions more closely tied to aggregate welfare and productivity have been developed (Petrin and Levinsohn (2008), Basu and Fernald (2002)
  - These decompositions highlight that at any instant of time marginal revenue products have not been equalized to factor prices because of adjustment frictions.
  - Reallocation is constantly moving resources that push towards such equalization. Shocks each period continually yield gaps/wedged in marginal revenue products.
  - This approach has considerable potential but implementation is complicated by measurement. Requires measuring marginal revenue products. Accurate estimation of factor elasticities and relevant factor prices critical.
  - For example, variation across plants in factor prices may reflect quality differences in factors.

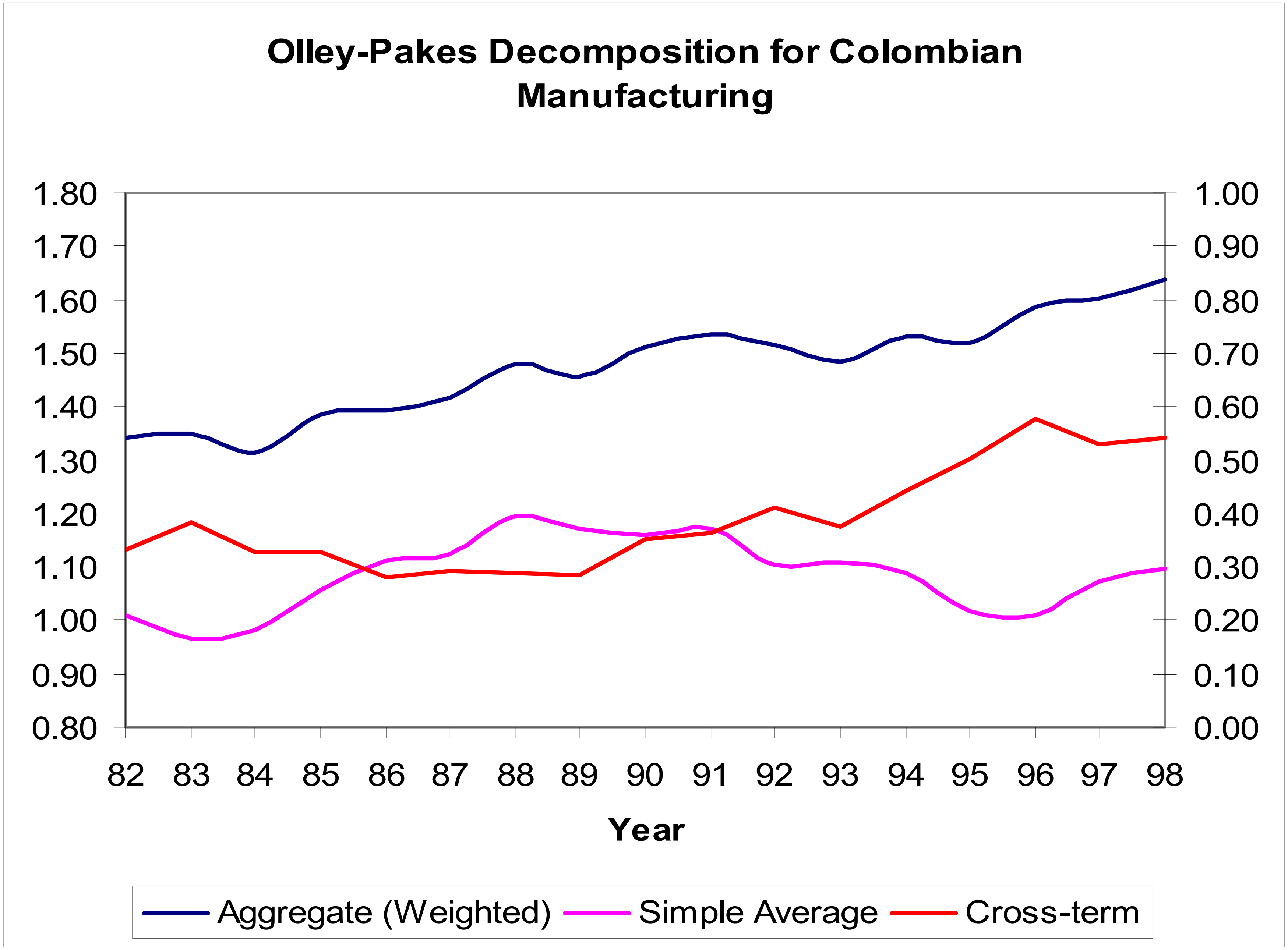
Olley and Pakes (1996) results for Telecommunications equipment

TABLE XI

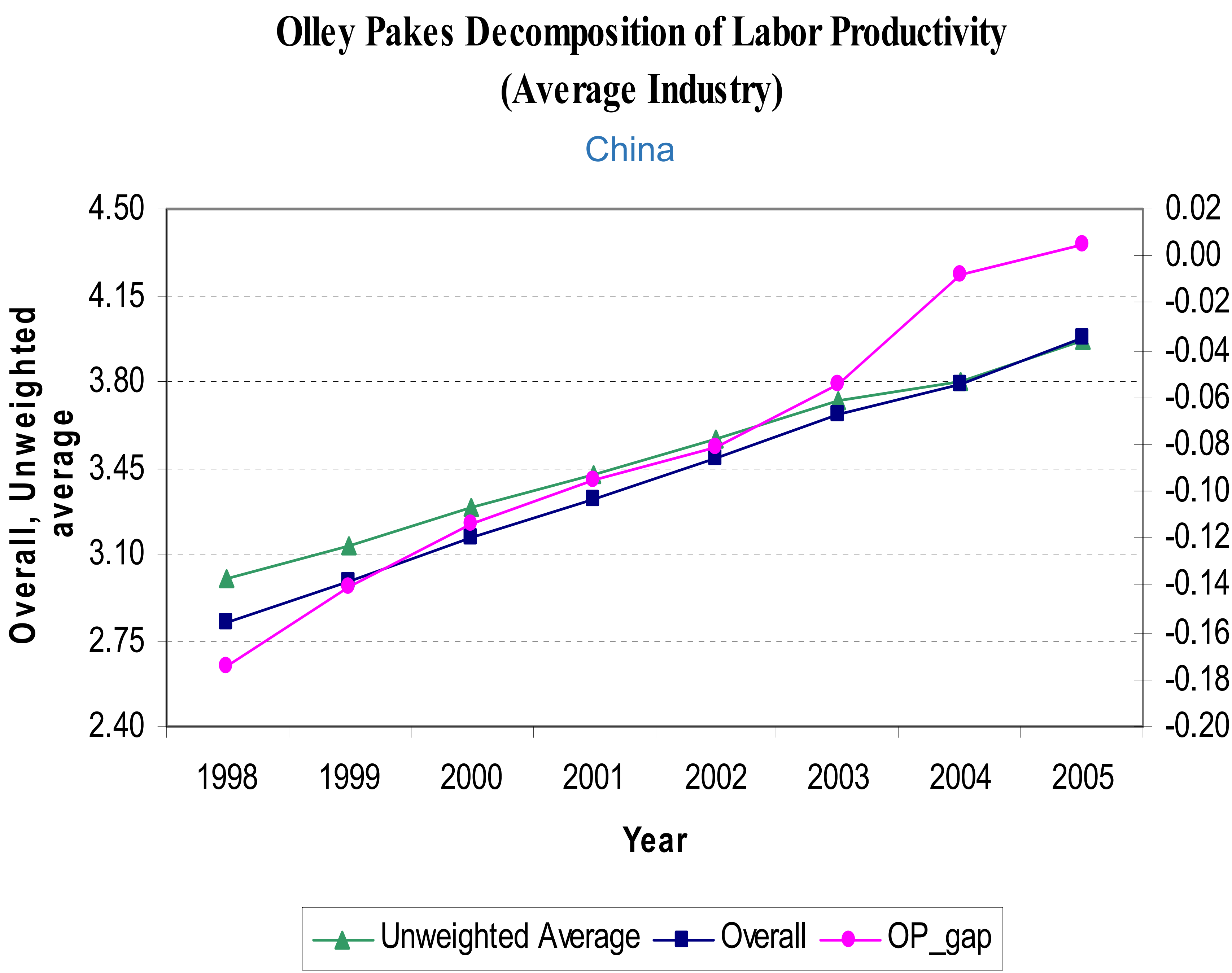
DECOMPOSITION OF PRODUCTIVITY[a]
(EQUATION (16))

Year	p_t	$\bar{p}_t$	$\Sigma_i \Delta s_{it} \Delta p_{it}$	$\rho(p_t, k_t)$
1974	1.00	0.90	0.01	−0.07
1975	0.72	0.66	0.06	−0.11
1976	0.77	0.69	0.07	−0.12
1977	0.75	0.72	0.03	−0.09
1978	0.92	0.80	0.12	−0.05
1979	0.95	0.84	0.12	−0.05
1980	1.12	0.84	0.28	−0.02
1981	1.11	0.76	0.35	0.02
1982	1.08	0.77	0.31	−0.01
1983	0.84	0.76	0.08	−0.07
1984	0.90	0.83	0.07	−0.09
1985	0.99	0.72	0.26	0.02
1986	0.92	0.72	0.20	0.03
1987	0.97	0.66	0.32	0.10

[a] See text for details.



Source: Eslava et al. (2005)



Olley Pakes Decomposition of Labor Productivity
(Average Industry)
China
4.50
4.15
3.80
3.45
3.10
2.75
2.40
0.02
0.00
-0.02
-0.04
-0.06
-0.08
-0.10
-0.12
-0.14
-0.16
-0.18
-0.20
Overall, Unweighted average
OP_gap
1998
1999
2000
2001
2002
2003
2004
2005
Year
Unweighted Average
Overall
OP_gap

Key equations

$$\sum_i \sum_k \left(P_i \frac{\partial Q_i}{\partial X_{ik}} - W_{ik}\right) dX_{ik} + \sum_i \sum_j \left(P_i \frac{\partial Q_i}{\partial M_{ij}} - P_j\right) dM_{ij} - \sum_i P_i dF_i + \sum_i P_i d\omega_i, \quad (9)$$

$$\sum_i D_i \sum_k (\varepsilon_{ik} - c_{ik}) dlnX_{ik} + \sum_i D_i \sum_j (\varepsilon_{ij} - c_{ij}) dlnM_{ij} - \sum_i D_i dlnF_i + \sum_i D_i dln\omega_i, \quad (11)$$

$$c_{ik} = \frac{W_{ik} X_{ik}}{P_i Q_i}$$

ε_{ik} is estimated at the 4-digit level from production function

- Key point: In practice gaps are based on difference between industry-specific, time invariant estimated factor elasticities and plant-specific, time varying cost shares
  - As a rough approximation they are exploiting within industry dispersion in cost shares

What do such gaps capture?

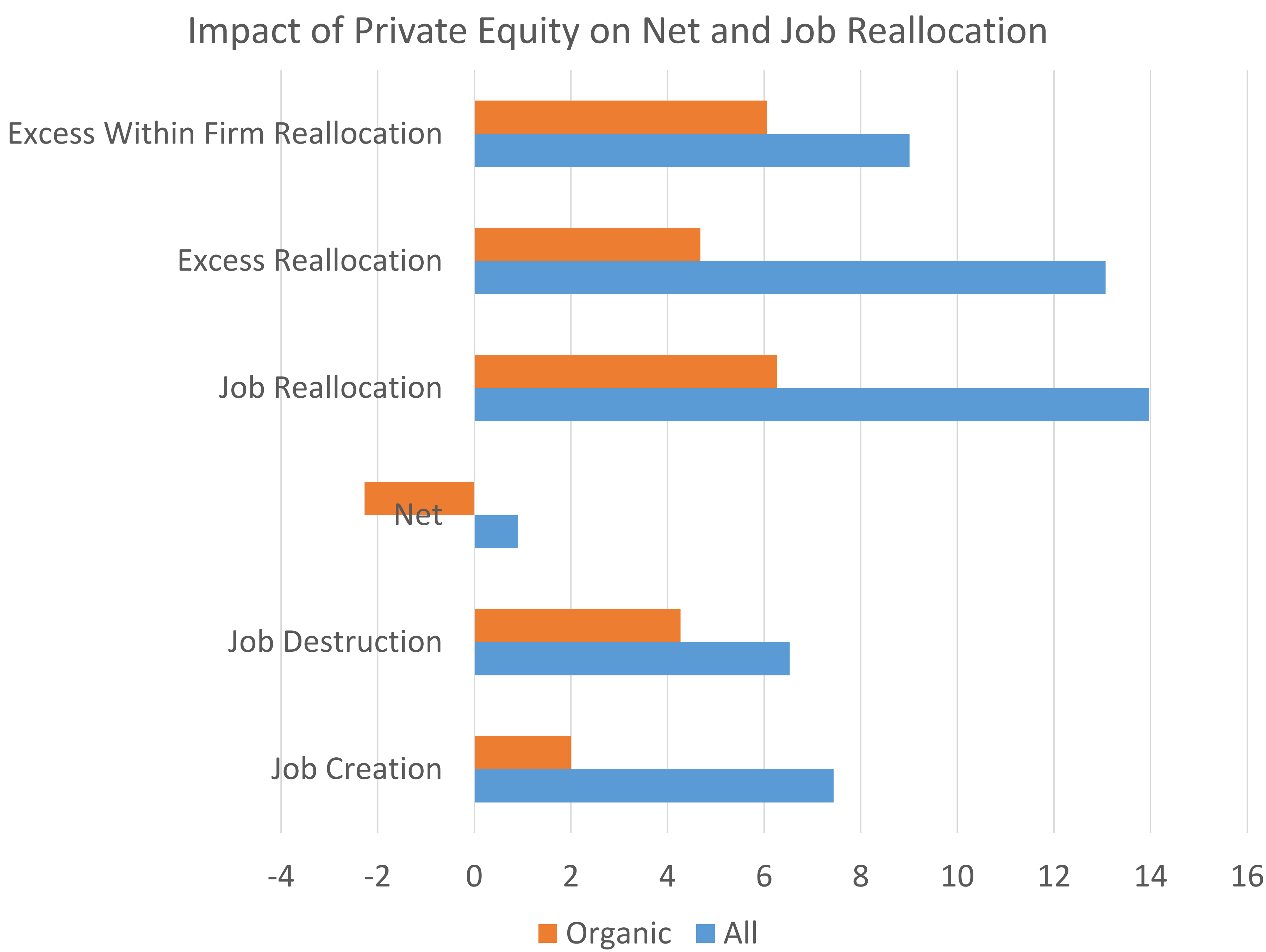
- Frictions and distortions?
  - Gaps are similar in spirit to measures from factor adjustment literature
    - But even here difficulties:
      - Best fitting structural models to plant-level data have proportional nonconvex costs of adjustment (disruption effects) so not additively separable
      - Can't separately estimate production/profit function from adjustment costs
- Unobserved heterogeneity in factor elasticities?
  - Much evidence that K/L ratios, skill mixes exhibit persistent differences across plants in the same industry
    - Matched employer-employee data show persistence of 0.92 of idiosyncratic (within 4-digit industry) in fraction of highly educated workers at establishments
    - At the core of the labor literature on skill biased technical change using micro data

What do such gaps capture?

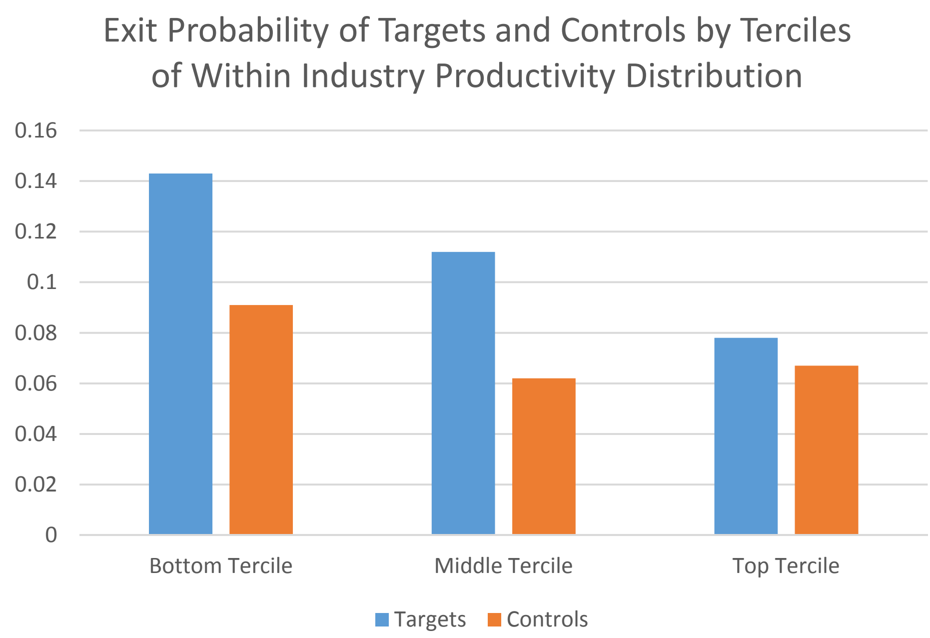
- Unobserved differences in input quality
  - Model assumes factors of type i are homogenous.
  - Large within and between plant wage dispersion
    - In 1992, within plant C.V. for Production Workers 0.21, between plant C.V. for Nonproduction workers 0.47
    - Much due to differences in skill mixes within and across plants
- Differences in factor prices across plants for same quality input
  - Efficiency wages, rent sharing (so not exogenous differences)
    - Abowd et. al. findings suggest about 50 percent of variation is person effects and about 50 percent firm effects

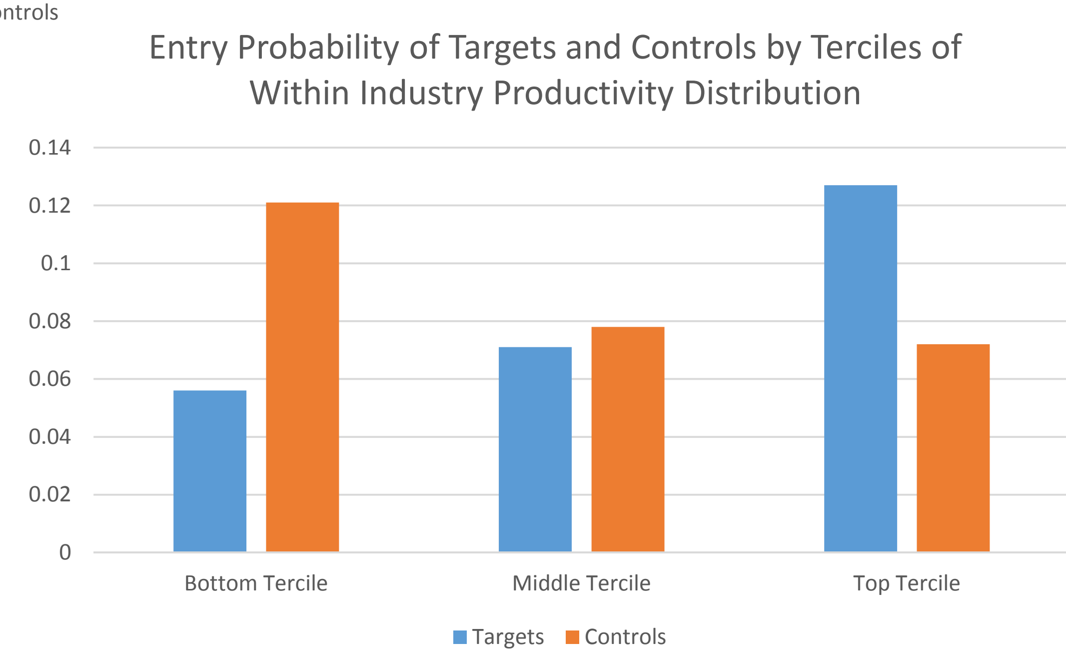
Ownership Change, Management, Financing…

- Many factors underlie the ongoing restructuring and reallocation of businesses
- For allocative efficiency, financial markets need to be facilitating the reallocation of resources to the most productive businesses
- Ownership/management practices and changes are part of these dynamics.
- Example: Private Equity



Impact of Private Equity on Net and Job Reallocation
Excess Within Firm Reallocation
Excess Reallocation
Job Reallocation
Net
Job Destruction
Job Creation
-4
-2
0
2
4
6
8
10
12
14
16
Organic
All


Exit Probability of Targets and Controls by Terciles of Within Industry Productivity Distribution
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Bottom Tercile
Middle Tercile
Top Tercile
Targets
Controls


Entry Probability of Targets and Controls by Terciles of Within Industry Productivity Distribution
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Bottom Tercile
Middle Tercile
Top Tercile
Targets
Controls

Two Year Productivity Growth Impact From Private Equity

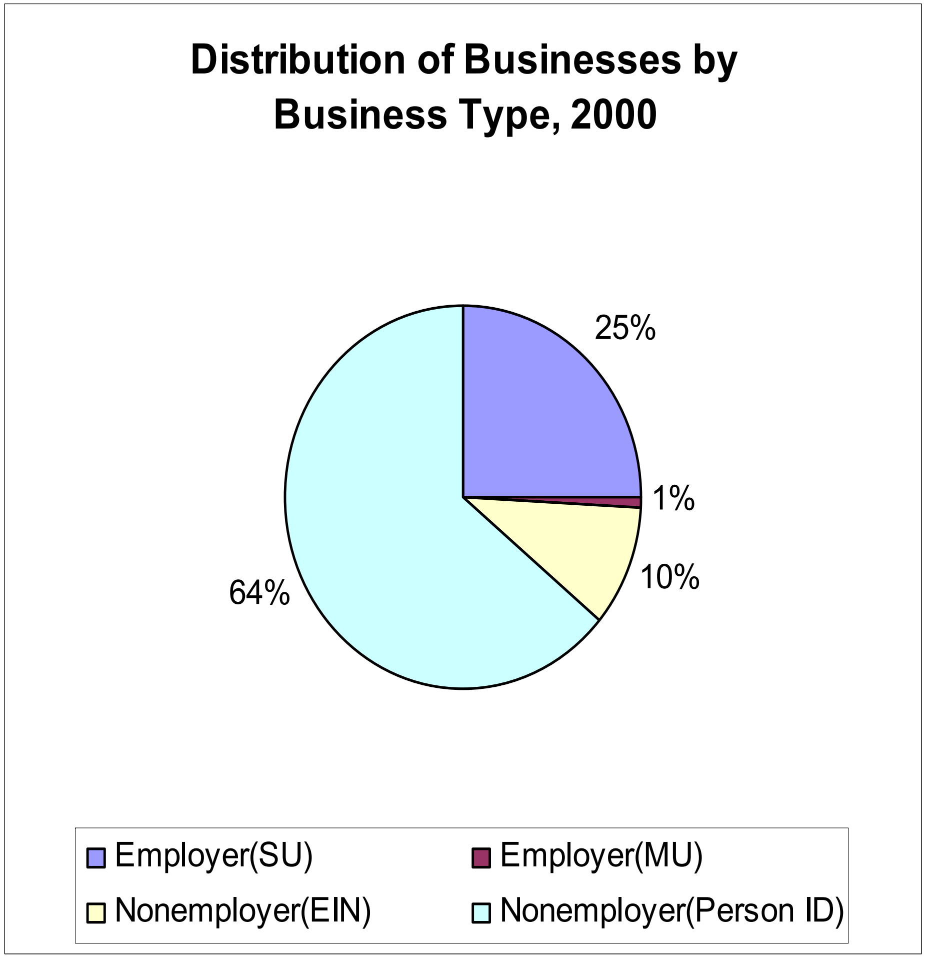
Total Productivity Growth Differential	2.09
Excluding Acquisition/Divestiture	1.96
Share of Total from:	
Continuing Establishments	0.20
Net Entry	0.74
Net Acquisition	0.06

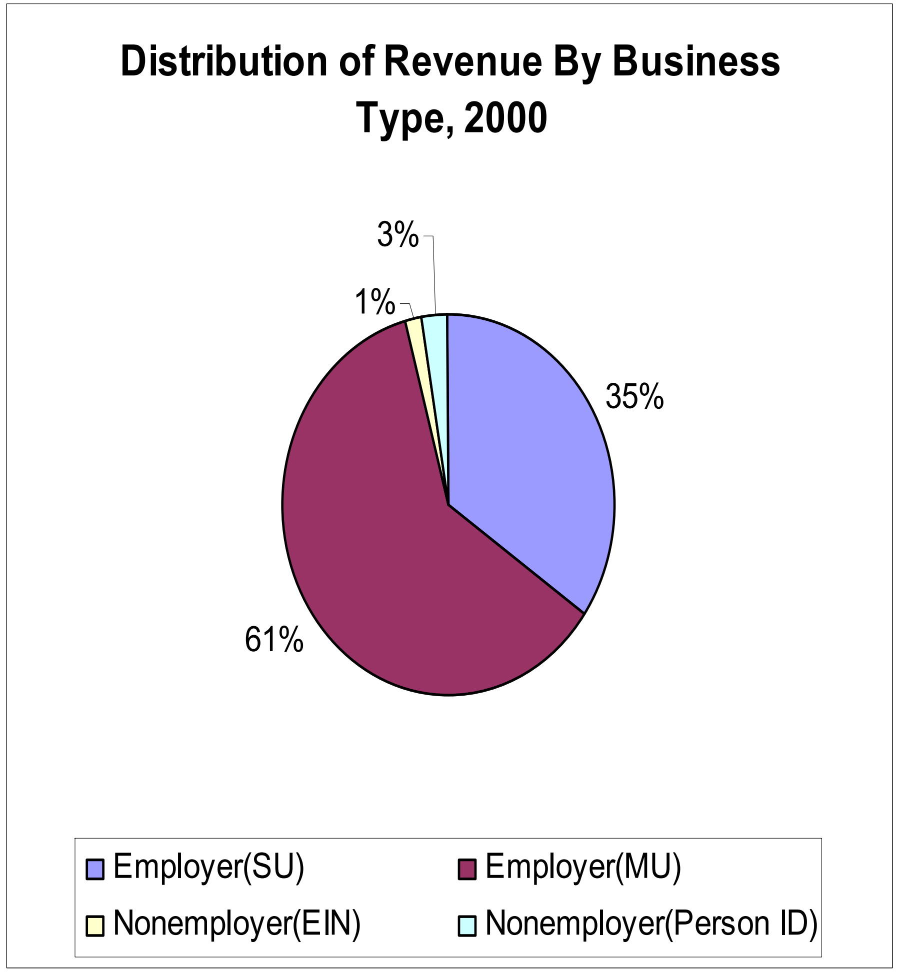
Pre-Entry History of entrepreneurs

“Before” Entry….

- Entrepreneurial dynamics starts at micro business level
- Entrepreneurs start with an idea – often while employed elsewhere
- New longitudinal databases at U.S. Census Bureau tracking this process
  - ILBD: Nonemployers (e.g., sole props without employees) + Employers
  - LEHD/SED: Tracking transitions from W&S jobs to self-employed jobs

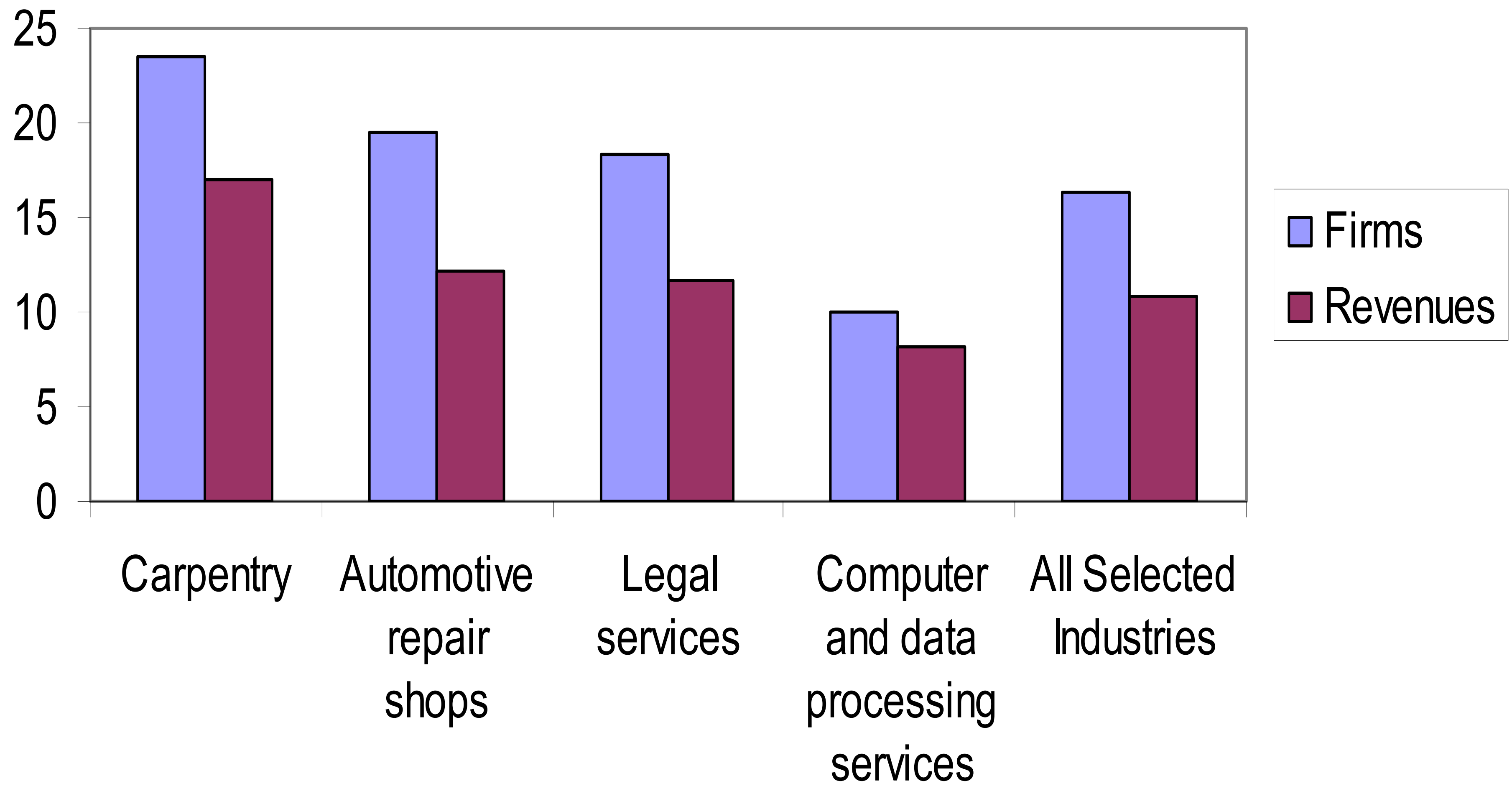
Micro Businesses constitute a large share of businesses and a small share of revenue…





Source: Davis et. al. (2008)

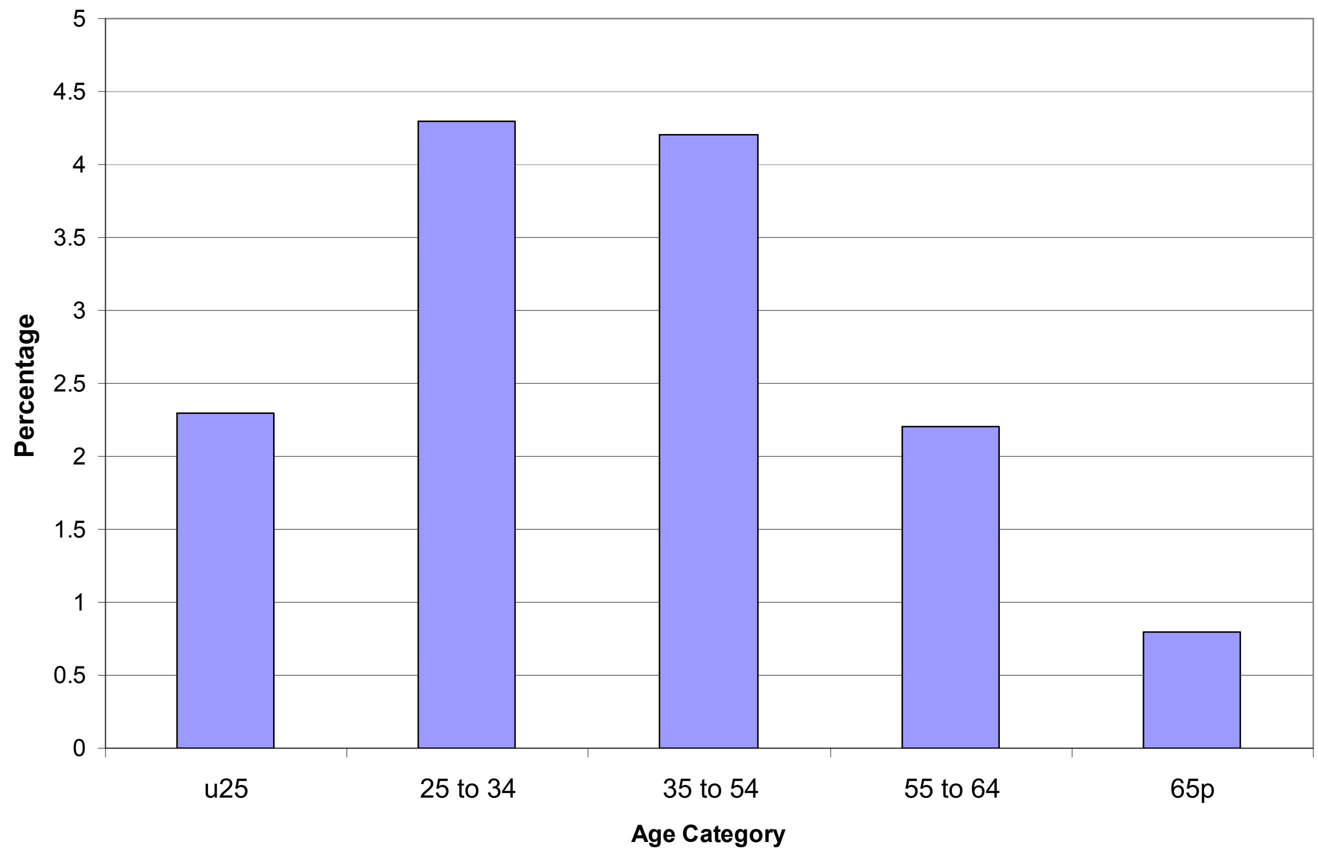
Shares of New Employer Businesses in 1997 with Pre-History as Nonemployer Businesses



Source: Davis et al. (2008)

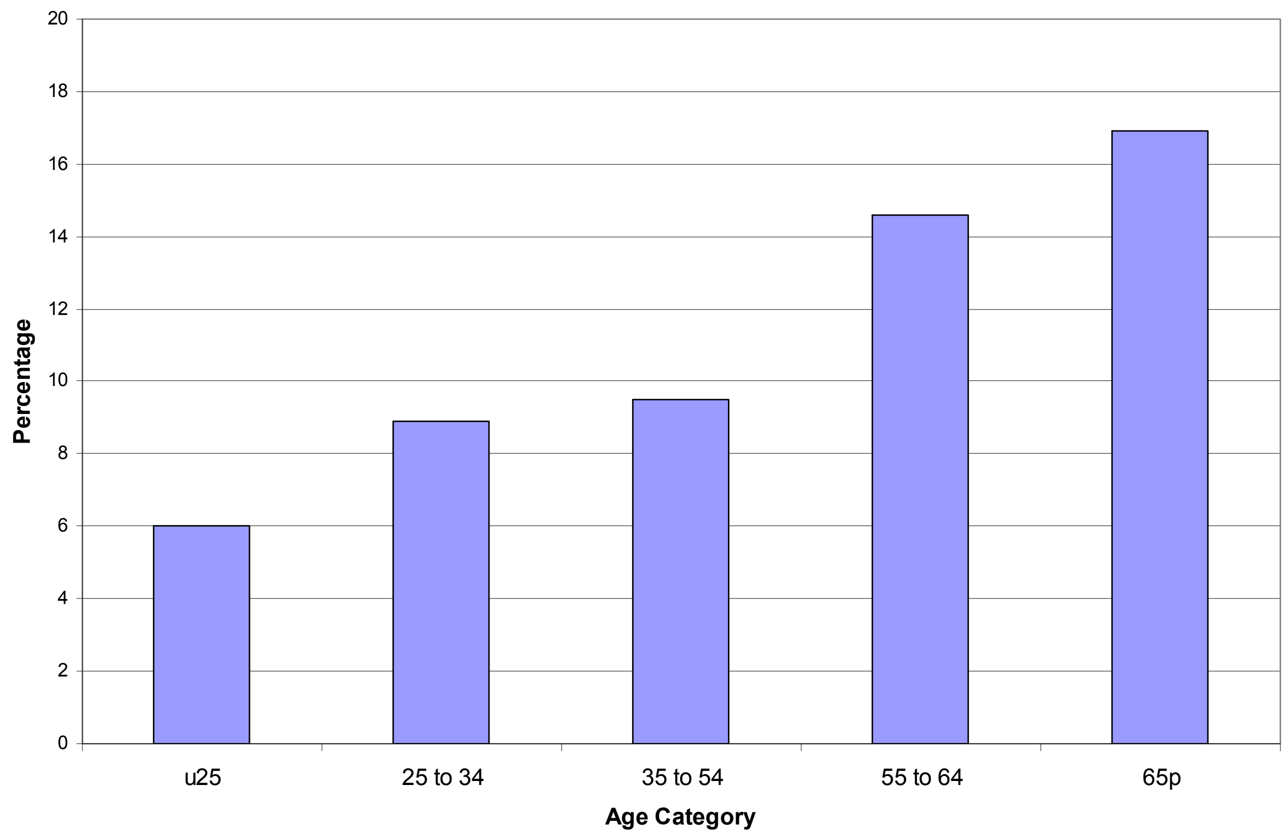
Propensity to Diversify in Labor Market Varies in Important Ways Across Worker Life Cycle

Percent of 1992 Wage and Salary Earners moving to Partial Self-Employment by 1997:
By Age Category

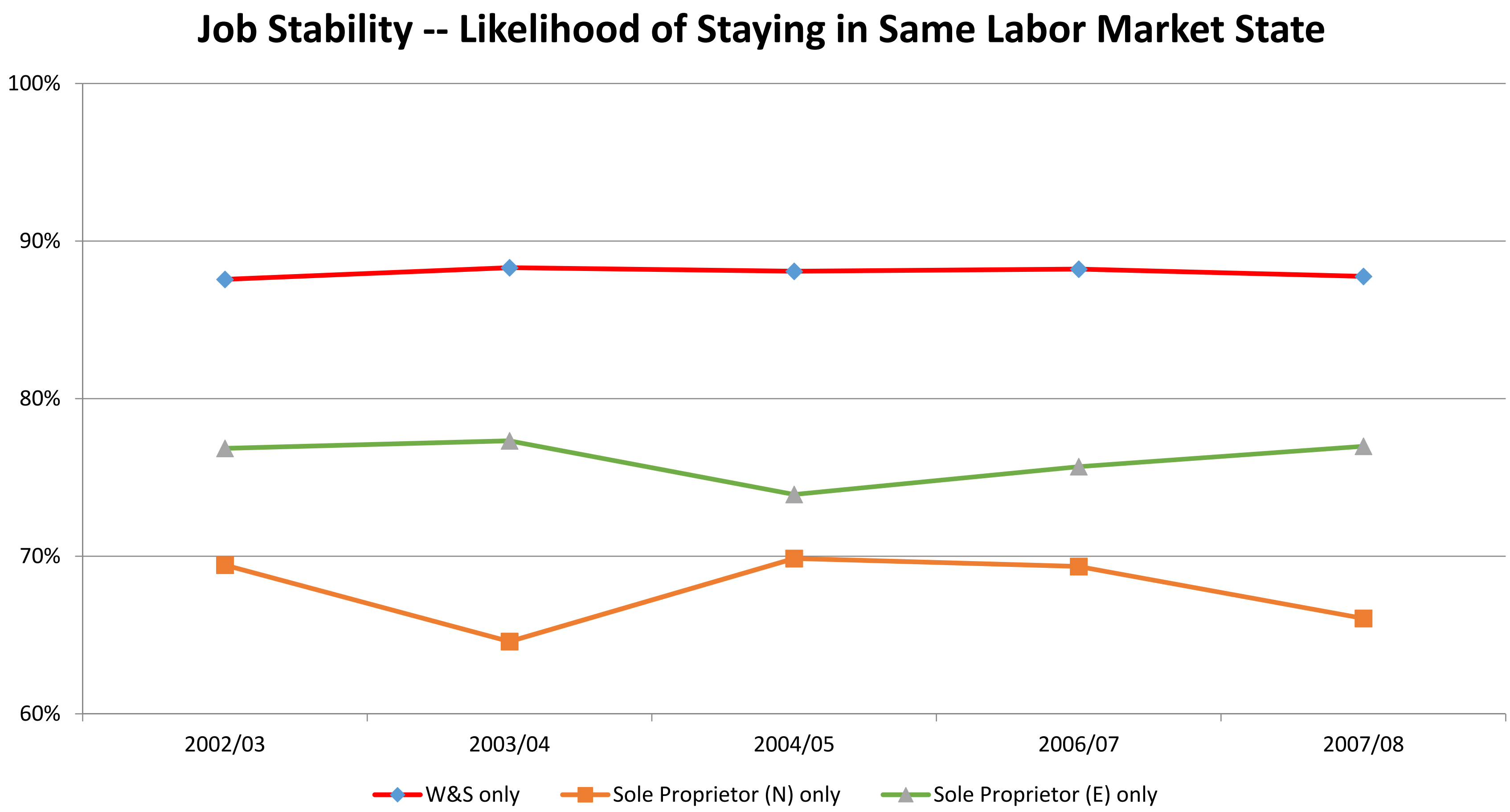


Propensity to Diversify in Labor Market Varies in Important Ways Across Worker Life Cycle

Percent of 1992 Partially Self-Employed moving to Full Self-Employment by 1997: By Age Category



Small Businesses With and Without Paid Employees Differ in Fundamental Ways



Data

- Tracking U.S. Business Dynamics
  - The Longitudinal Business Database
    - 1975-2005 (08) – long time series permits analysis by firm age
    - Private Non Farm Economy
    - Establishment level with Firm identifiers
    - High quality establishment links to identify entry/exit
      - Need both firm and establishment level data to get dynamics right
    - Firm Size: constructed by aggregating employment up to firm
    - Firm Age: constructed from age of oldest establishment at time of firm birth
    - Other: Payroll, Industry, Location (Lat/Lon possible)
    - Can be integrated with data from Economic Censuses and Annual Surveys as well as external data (COMPUSTAT, Venture Capital, Private Equity)

Micro Productivity Data in U.S.

- Manufacturing:
  - Annual Survey of Manufactures and Census of Manufactures
    - Nominal revenue and expenditures
    - Can construct measures of real outputs and inputs
    - Five year panel rotation so longitudinal analysis possible (but requires careful treatment of data)
    - Selected products have physical quantities
- Retail Trade
  - Census of Retail Trade
    - Nominal revenue so a gross output per store measure feasible

New data on micro businesses

- ILBD:
  - Tracks all nonemployer and employer businesses including transitions
- LEHD:
  - Tracks all employer-employee matches in U.S.
  - Can be integrated with ILBD
  - Enables tracking of transitions between W&S, an owner of nonemployer business and owner of employer business

Availability of data

- Public domain tabulations available at:

http://www.ces.census.gov/index.php/bds/bds_home

- Census NSF/RDC access at:

http://www.ces.census.gov/index.php/ces/researchguidelines

- Sensitive data:
  - Must work in enclave (NBER, NYCRDC, Washington, D.C., Chicago Fed, Duke, UCLA, UC-Berkeley, Univ. of Michigan, Cornell, Stanford , Univ. of Minn., Atlanta, …)
  - Predominant purpose must benefit U.S. Census

Extra Slides on Firm Dynamics model

Standard Heterogeneous-Producer Industry Models

The Workhorse:

- Producers i differ in a profitability component ω_i, usually taken to represent costs/productivity

- Profits depend on ω_i and industry state S: $\pi_i = \pi_i(\omega_i, S) \quad \omega_i \sim G(\omega)$

- There is some critical ω^* such that producers with $\omega_i < \omega^*$ have NPVs below outside option and therefore exit the industry

- Industry state S typically depends on endogenously determined distribution of ω_i among producers (add'l free entry assumption)

- Examples: Jovanovic (1982), Hopenhayn (1992), Melitz (2003), Asplund and Nocke (2007)

Closely Related Issue – Size Distribution of Activity

- $\pi_i = \pi_i(\omega_i, S)$ has curvature either from decreasing returns (e.g., Lucas (1978)) or product differentiation (e.g., Melitz (2003))
- Curvature pins down the size distribution of activity and permits studying the evolution of the size distribution of activity
- In healthy market economies, most productive plants are the largest – allocative efficiency
- Active literature attempting to explain cross country differences in productivity (e.g., Hsieh and Klenow (2009)) using distortions on this margin

Model: Melitz/Ottaviano (2005) and FHS (2008)

Industry is comprised of a continuum of producers of measure *N*.
Each produces a single variety (indexed by *i*) of industry product.
Representative consumer's utility function

$$U = y + \int_{i\in I}(\alpha+\delta_i)q_i di - \frac{1}{2}\eta\left(\int_{i\in I} q_i di\right)^2 - \frac{1}{2}\gamma\int_{i\in I} q_i^2 di$$

$$= y + \alpha\int_{i\in I} q_i di - \frac{1}{2}\left(\eta+\frac{\gamma}{N}\right)\left(\int_{i\in I} q_i di\right)^2 + \int_{i\in I}\delta_i q_i di - \frac{1}{2}\gamma\int_{i\in I}(q_i-\bar{q})^2 di$$

$\alpha > 0$, $\eta > 0$, and $\gamma \geq 0$.
y = numeraire good
δ_i = variety-specific, mean-zero taste shifter
q_i = quantity of good *i* consumed

$$\bar{q} = \frac{1}{N}\int_{i\in I} q_i di$$

The implied demand curve:

$$q_i = \frac{1}{\eta N+\gamma}\alpha - \frac{1}{\gamma}\frac{\eta N}{\eta N+\gamma}\bar{\delta} + \frac{1}{\gamma}\frac{\eta N}{\eta N+\gamma}\bar{p} + \frac{1}{\gamma}\delta_i - \frac{1}{\gamma}p_i$$

Model: Supply

Production Function: $q_i = \omega_i x_i$

Producers face (potentially idiosyncratic) factor price w_i

$\Rightarrow$ marginal cost = w_i / ω_i

Profits:

$$\pi_i = \left(\frac{1}{\eta N + \gamma}\alpha - \frac{1}{\gamma}\frac{\eta N}{\eta N + \gamma}\bar{\delta} + \frac{1}{\gamma}\frac{\eta N}{\eta N + \gamma}\bar{p} + \frac{1}{\gamma}\delta_i - \frac{1}{\gamma}p_i \right)\left(p_i - \frac{w_i}{\omega_i} \right)$$

Profit-maximizing price (constant marginal cost c_i):

$$p_i = \frac{1}{2}\frac{\gamma}{\eta N + \gamma}\alpha - \frac{1}{2}\frac{\eta N}{\eta N + \gamma}\bar{\delta} + \frac{1}{2}\frac{\eta N}{\eta N + \gamma}\bar{p} + \frac{1}{2}\delta_i + \frac{1}{2}\frac{w_i}{\omega_i}$$

Deviation from industry-average price:

$$p_i - \bar{p} = \frac{1}{2}\left(\delta_i - \bar{\delta}\right) + \frac{1}{2}\left(\frac{w_i}{\omega_i} - \overline{\left(\frac{w}{\omega}\right)} \right)$$

Maximized profits:

$$\pi_i = \frac{1}{4\gamma}\left(\frac{\gamma}{\eta N + \gamma}\alpha - \frac{\eta N}{\eta N + \gamma}\bar{\delta} + \frac{\eta N}{\eta N + \gamma}\bar{p} + \delta_i - \frac{w_i}{\omega_i} \right)^2$$

Model: Equilibrium

Equilibrium Condition 1: The marginal producer in the industry makes zero profits

Define "profitability index" $\phi_i \equiv \delta_i - \frac{w_i}{\omega_i}$. Then marginal producer has index equal to:

$$\phi^* = -\frac{\gamma}{\eta N + \gamma}\alpha + \frac{\eta N}{\eta N + \gamma}\bar{\delta} - \frac{\eta N}{\eta N + \gamma}\bar{p}$$

Profits can be rewritten in terms of this marginal profitability level

$$\pi_i = \frac{1}{4\gamma}(\phi_i - \phi^*)^2$$

Profits increase in demand (δ_i) and efficiency (ω_i), decrease in factor price (w_i)

Equilibrium Condition 2: Potential entrants decide whether to pay sunk entry cost s to learn δ_i, ω_i, w_i. Expected value of entry is 0.

$$V^e = \int_0^{w_u}\int_{\omega_l}^{\omega_u}\int_{\phi^*+\frac{w}{\omega}}^{\delta_e} \frac{1}{4\gamma}(\phi_i - \phi^*)^2 f(\delta,\omega,w)\,d\delta\, d\omega\, dw - s = 0$$

Selection effect:

- Only high-profitability producers operate in equilibrium
- Low types exit

Sunk costs, market power and dispersion:

- Sunk costs make entry costly
- Curvature yields equilibrium size distribution

Many models of selection also include fixed costs of operating each period

Model: Empirical Implications

Output-based productivity:

$$TFPQ_i = \frac{q_i}{x_i} = \frac{\omega_i x_i}{x_i} = \omega_i$$

Revenue-based productivity (literature standard):

$$TFPR_i = \frac{p_i q_i}{x_i} = p_i \omega_i = \frac{1}{2}\frac{\gamma\alpha}{\eta N + \gamma}\omega_i + \frac{1}{2}\frac{\eta N}{\eta N + \gamma}\left(\overline{p} - \overline{\delta}\right)\omega_i + \frac{1}{2}\delta_i \omega_i + \frac{1}{2} w_i$$

Plant price deviation from industry deflator depends on both demand (enters positively into profits) and costs (enter negatively):

$$p_i - \overline{p} = \frac{1}{2}\left(\delta_i - \overline{\delta}\right) + \frac{1}{2}\left(\frac{w_i}{\omega_i} - \overline{\left(\frac{w}{\omega}\right)}\right)$$

Comparative Statics:

- $\frac{d\phi^*}{d\gamma} < 0$: Lower substitutability (higher γ) lowers ϕ^*
- $\frac{d\phi^*}{ds} < 0$: Higher sunk entry cost lowers ϕ^*

Start with Foster, Haltiwanger and Syverson (2008)

- Source data: Census of Manufactures
  - High quality coverage
  - Limited number of products with physical quantity data

Correlations								
Variables	Trad'l. Output	Revenue Output	Physical Output	Price	Trad'l. TFP	Revenue TFP	Physical TFP	Capital
Traditional Output	1.00							
Revenue Output	0.99	1.00						
Physical Output	0.98	0.99	1.00					
Price	-0.03	-0.03	-0.19	1.00				
Traditional TFP	0.19	0.18	0.15	0.13	1.00			
Revenue TFP	0.17	0.21	0.18	0.16	0.86	1.00		
Physical TFP	0.17	0.20	0.28	-0.54	0.64	0.75	1.00	
Capital	0.86	0.85	0.84	-0.04	0.00	-0.00	0.03	1.00
Standard Deviations								
Standard Deviations	1.03	1.03	1.05	0.18	0.21	0.22	0.26	1.14

Measuring Plant-Level Demand

Estimate product demand curves; plant-specific residual is idio. demand

$$\ln q_{it} = \alpha_o + \alpha_1 \ln p_{it} + \alpha_2 \ln(INCOME_{mt}) + \sum_t \alpha_t YEAR_t + \eta_{it}$$

q_{it}—physical output of plant i in year t
p_{it}—plant unit price
$INCOME_{mt}$—average income in the plant's local market m
$YEAR_t$—year dummy
η_{it}—plant-year disturbance term

Plant demand:

$$\hat{\delta}_{it} = \hat{\eta}_{it} + \hat{\alpha}_2 \ln(INCOME_{mt}) = \ln q_{it} - \hat{\alpha}_o - \hat{\alpha}_1 \ln p_{it} - \sum_t \hat{\alpha}_t YEAR_t$$

I.e., residual is plant quantity sold that can't be accounted for by unit price or local income differences

- Use $TFPQ_{it}$ to instrument for prices (captures production costs)

	IV Estimation			OLS Estimation	
Product	Price Coefficient (α_1)		Income Coefficient (α_2)	Price Coefficient (α_1)	Income Coefficient (α_2)
Boxes	-3.02		-0.03	-2.19	-0.03
	0.17	[0.61]	*0.02*	*0.12*	*0.02*
Bread	-3.09		0.12	-0.89	0.07
	0.42	[0.33]	*0.05*	*0.15*	*0.04*
Carbon Black	-0.52		-0.21	-0.57	-0.21
	0.38	[0.50]	*0.11*	*0.21*	*0.11*
Coffee	-3.63		0.22	-1.03	0.20
	0.98	[0.41]	*0.14*	*0.32*	*0.13*
Concrete	-5.93		0.13	-0.83	0.15
	0.36	[0.10]	*0.01*	*0.09*	*0.01*
Hardwood Flooring	-1.67		-0.20	-0.87	-0.24
	0.48	[0.61]	*0.18*	*0.47*	*0.18*
Gasoline	-1.42		0.23	-0.16	0.23
	2.72	[0.20]	*0.07*	*0.80*	*0.07*
Block Ice	-2.05		0.00	-0.63	0.16
	0.46	[0.32]	*0.11*	*0.20*	*0.07*
Processed Ice	-1.48		0.18	-0.70	0.16
	0.27	[0.37]	*0.03*	*0.13*	*0.03*
Plywood	-1.21		-0.23	-1.19	-0.23
	0.14	[0.89]	*0.10*	*0.13*	*0.10*
Sugar	-2.52		0.76	-1.04	0.72
	1.01	[0.15]	*0.13*	*0.55*	*0.12*

	Five-Year Horizon		Implied One-Year Persistence Rates	
Dependent Variable	Unweighted Regression	Weighted Regression	Unweighted Regression	Weighted Regression
Traditional TFP	0.249 *0.017*	0.316 *0.042*	0.757	0.794
Revenue TFP	0.277 *0.021*	0.316 *0.042*	0.774	0.794
Physical TFP	0.312 *0.019*	0.358 *0.049*	0.792	0.814
Price	0.365 *0.025*	0.384 *0.066*	0.817	0.826
Demand Shock	0.619 *0.013*	0.843 *0.021*	0.909	0.966

Variable	Plant Age Dummies			
	Exit	Entry	Young	Medium
Unweighted Regressions				
Traditional TFP	-0.0211	0.0044	0.0074	0.0061
	0.0042	*0.0044*	*0.0048*	*0.0048*
Revenue TFP	-0.0220	0.0133	0.0075	0.0028
	0.0044	*0.0047*	*0.0051*	*0.0053*
Physical TFP	-0.0186	0.0128	0.0046	-0.0039
	0.0050	*0.0053*	*0.0058*	*0.0062*
Price	-0.0034	0.0005	0.0029	0.0067
	0.0031	*0.0034*	*0.0038*	*0.0042*
Demand Shock	-0.3466	-0.5557	-0.3985	-0.3183
	0.0227	*0.0264*	*0.0263*	*0.0267*

Determinants of Market Selection

Specification:	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Traditional TFP	-0.073						
	0.015						
Revenue TFP		-0.063					
		0.014					
Physical TFP			-0.040			-0.062	-0.034
			0.012			*0.014*	*0.012*
Prices				-0.021		-0.069	
				0.018		*0.021*	
Demand Shock					-0.047		-0.047
					0.003		*0.003*

Note: Much greater dispersion in demand shocks than physical TFP

Establishment-level Productivity Empirical Patterns

- Dispersion (large), persistence (high) evolution (consistent with learning and selection)
- Selection
  - Lower productivity plants exit
  - Other determinants of productivity matter
  - Open questions: Impact of distortions on selection?
    - Models like Melitz (2003) and Restuccia and Rogerson (2007) imply reduced distortions will improve selection
    - Eslava et. al. (2009) find evidence that trade liberalization improves market selection
- These patterns both support basic models and can be used to test and estimate models
- One other approach has to been to explore the covariance between size and productivity within industries.
  - Basic prediction of virtually all models is positive correlation between size and profitability/productivity

Within Industry Dynamic Decomposition Applied to FHS (2008) data

		Components of Decomposition (GR)				
		Within	Between	Entry	Exit	Net Entry
Traditional	2.30	1.40	0.18	0.44	0.27	0.72
Revenue	5.13	4.03	0.16	0.55	0.39	0.94
Physical	5.13	3.82	-0.05	1.04	0.32	1.36

Extra Slides

Growth Identities: Establishment

$$g_{it} = (E_{it} - E_{it-1}) / X_{it}$$

where

$$X_{it} = .5 * (E_{it} + E_{it-1})$$

Then

$$JC_{it} = \max(g_{it}, 0)$$

$$JD_{it} = \max(-g_{it}, 0)$$

From Entry/Exit

$$JC_{it} = \max(g_{it}, 0) * I\{g_{it} = 2\}$$

$$JD_{it} = \max(-g_{it}, 0) * I\{-g_{it} = 2\}$$

Growth Identities: Aggregate Measures (any level)

$$JC_t = \sum_i (X_{it} / X_t) \max\{g_{it}, 0\} \qquad JD_t = \sum_i (X_{it} / X_t) \max\{-g_{it}, 0\}$$

$$JC\_Entry_t = \sum_i (X_{it} / X_t) I\{g_{it} = 2\} \max(g_{it}, 0)$$

$$JD\_Exit_t = \sum_i (X_{it} / X_t) I\{g_{it} = -2\} \max(-g_{it}, 0)$$

$$g_t = JC_t - JD_t$$

$$JC_t = JC\_Cont_t + JC\_Entry_t$$

$$JD_t = JD\_Cont_t + JD\_Exit_t$$