Reallocation, Startups, Innovation and Productivity Growth in the U.S. Economy: Facts, Models and Open Questions

John Haltiwanger

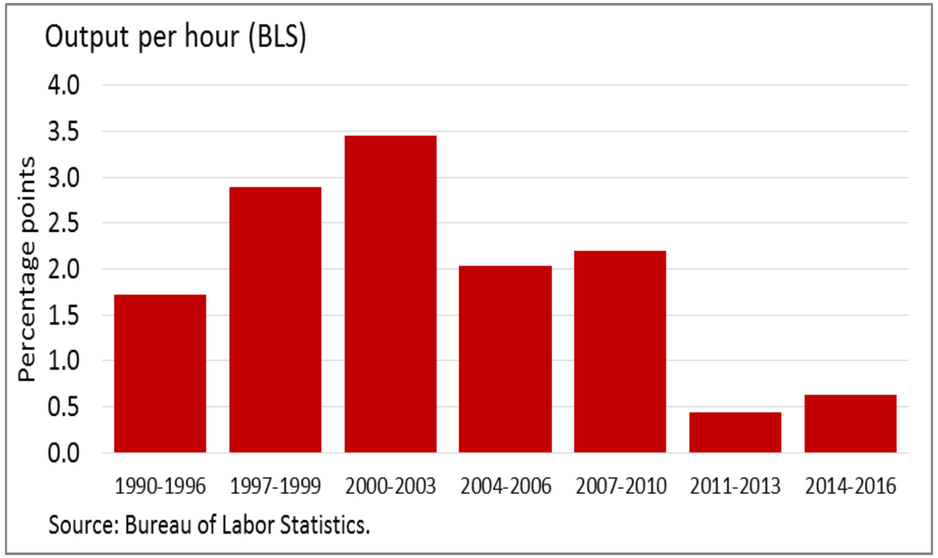
University of Maryland

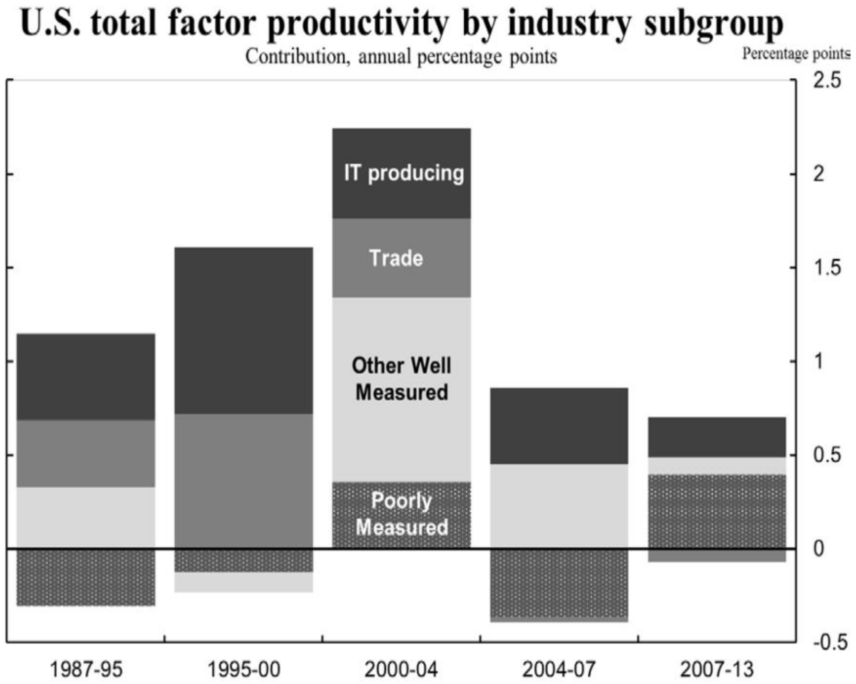
*This talk draws heavily on collaborative work with Steven Davis, Ryan Decker, Lucia Foster, Cheryl Grim, Ron Jarmin, Javier Miranda and Zoltan Wolf

Overview

- Reallocation and Productivity
  - Frictions and Shocks
    - Large dispersion in idiosyncratic shocks (productivity and demand).
    - Frictions for adjusting factors of production (focus today on labor).
    - Rise in frictions (distortions?) will lower productivity (levels vs. growth rates?)
  - Creative Destruction
    - Startups/young firms play outsized role.
    - Major innovations made by young firms.
    - Mature firms due more defensive innovation.
    - Decline in startups and increase in share of activity by more mature firms (increased concentration) can contribute to lower productivity growth.
- Application: U.S. has experienced lower productivity growth in post 2000 period accompanied by decline in pace of reallocation and startups.

Aggregate productivity has slowed since early 2000s

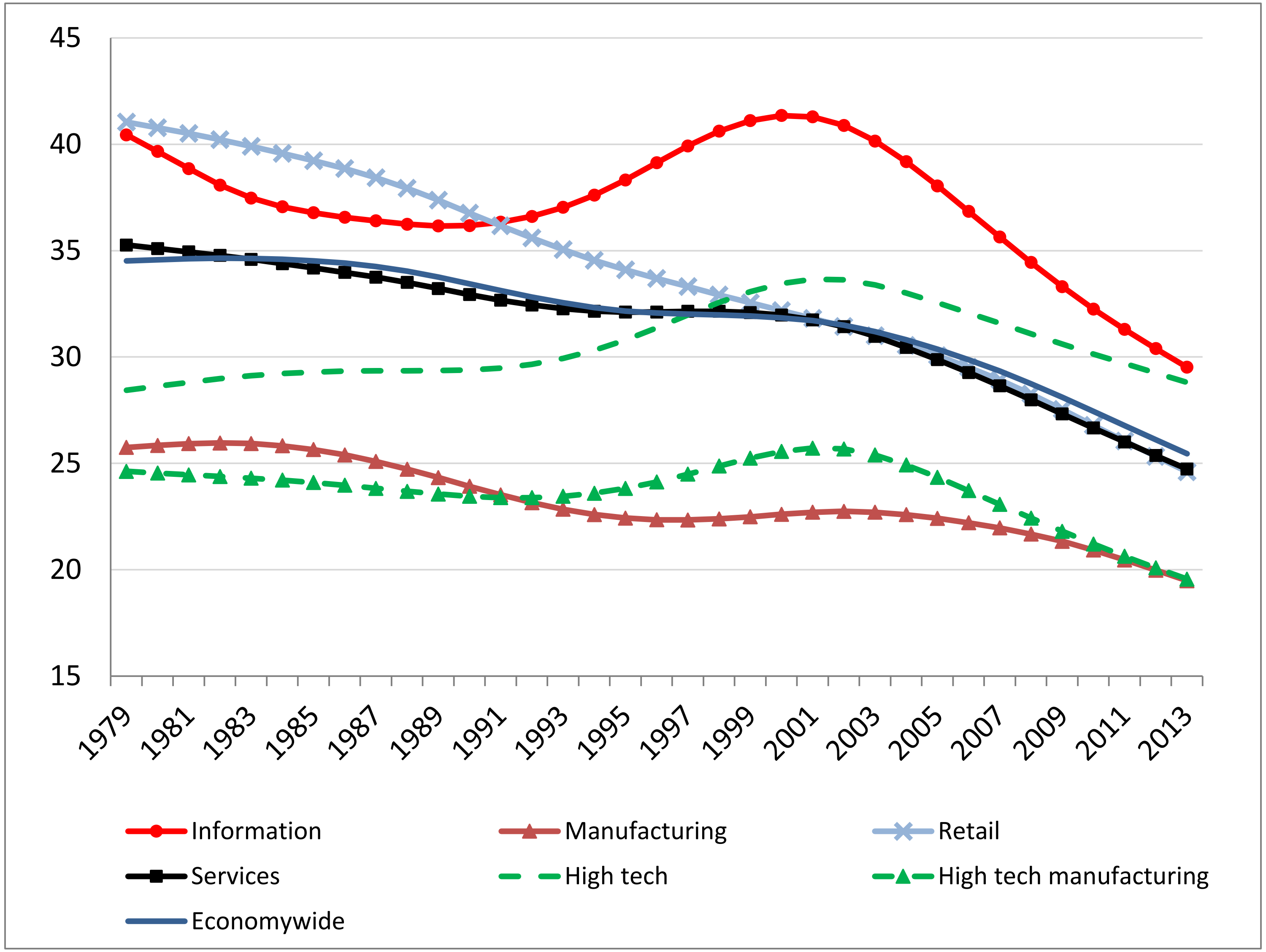




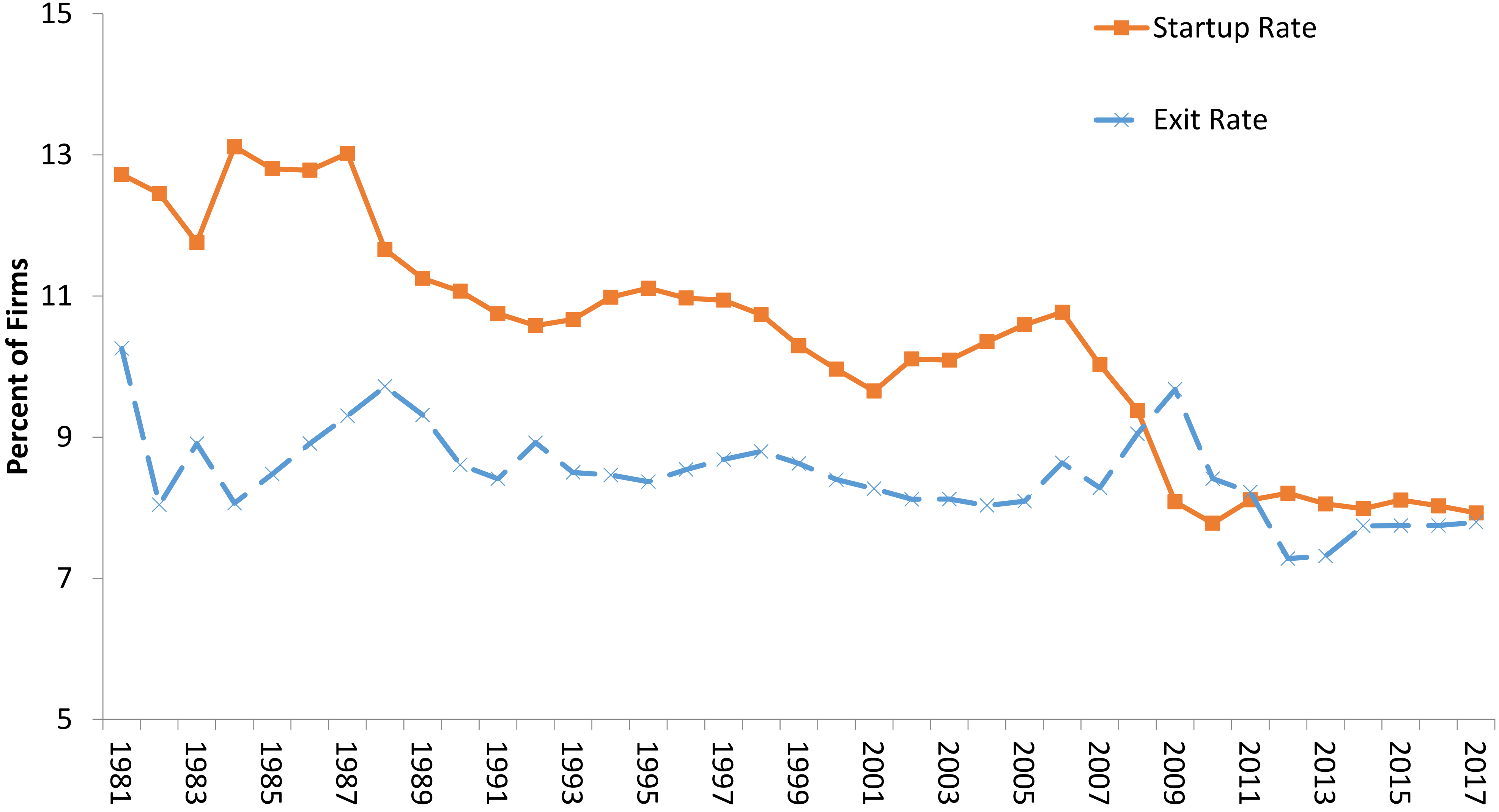
Source: Byrne et al. (2016)

- Surge in productivity in 1990s led by ICT-producing and using industries.
- Slowdown from early 2000s especially in ICT industries

Patterns of Decline in Reallocation Differ Substantially Across Sectors

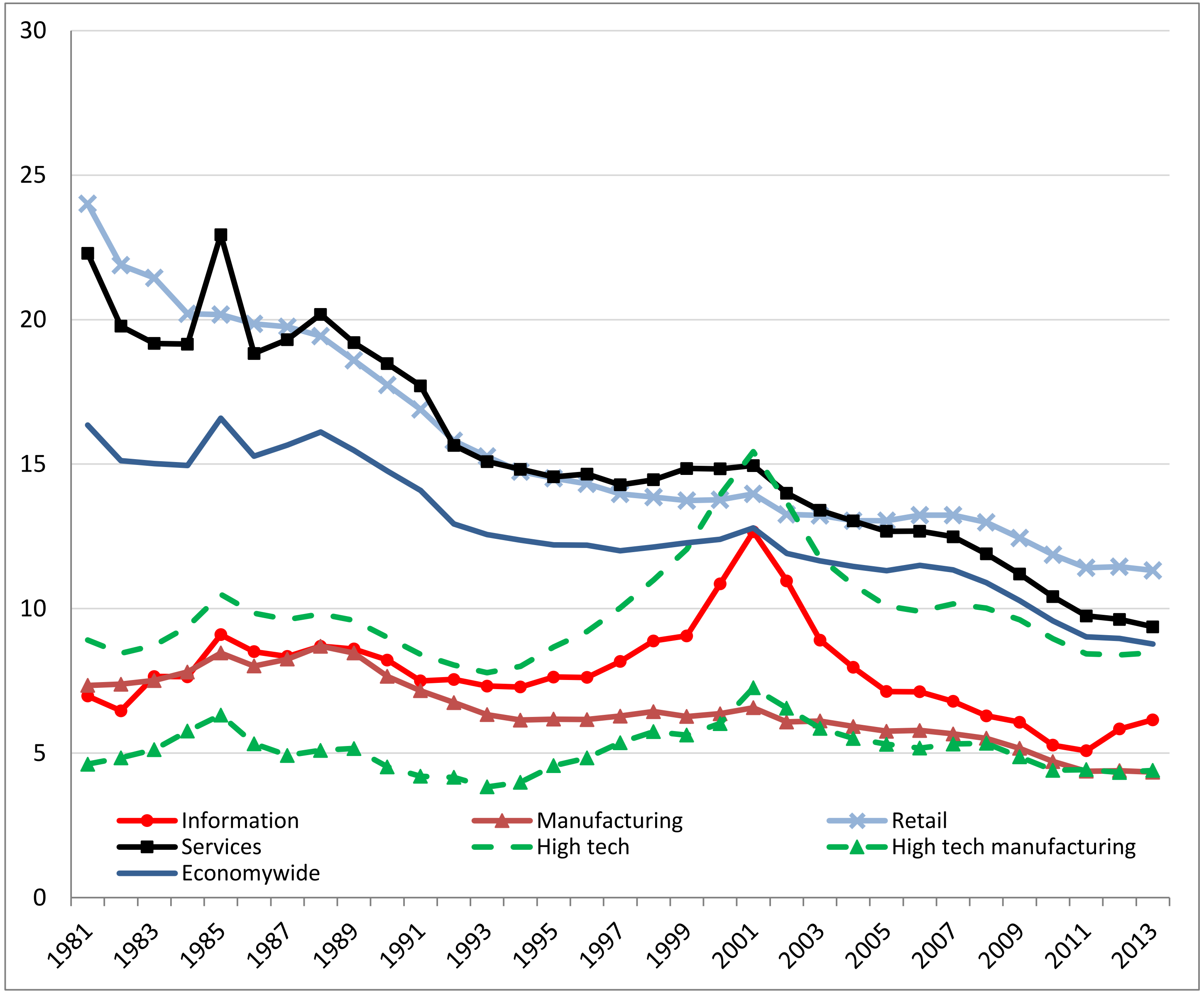


Startup and Exit Rates in U.S. Private Sector, 1981-2017



Source: Business Dynamic Statistics (Census) Spliced with Business Employment Dynamics (BLS)

Changing Share of Employment Accounted for by Young Firms



Retail and Services Dominated Pre 2000. Information Sector increased In 1990s. Broad based decline Post 2000.

Underlying these Patterns are Changing patterns Of Startup Rates.

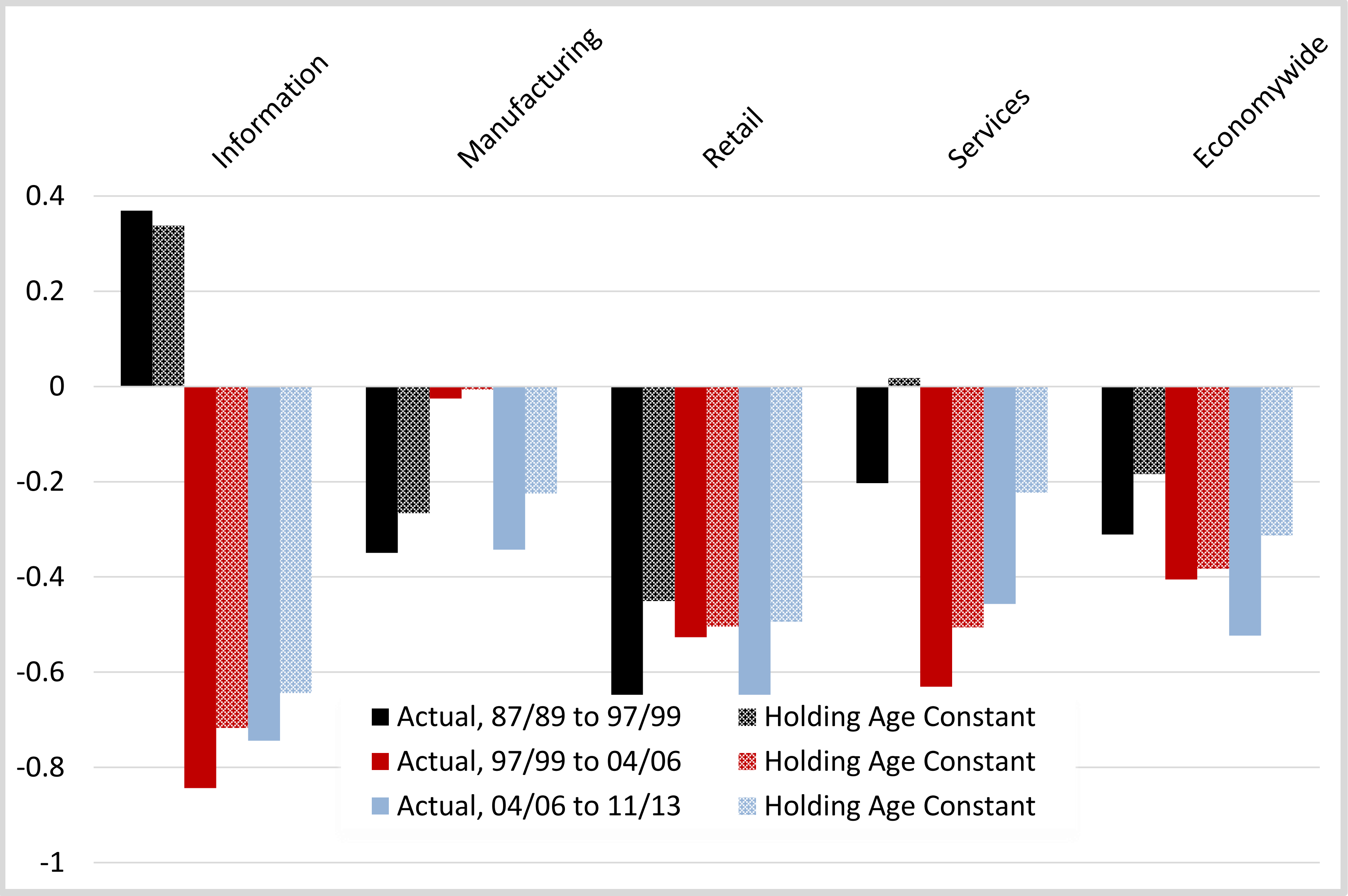
Possible connections between indicators of business dynamism and productivity ?

1. Increase in frictions has reduced pace of dynamism and entrepreneurship.
   - Large, within industry dispersion in productivity. Stochastic with AR1<1.
   - In healthy economy, within industry reallocation moving resources towards firms with more positive realizations of idiosyncratic productivity relative to industry mean.
   - An increase in frictions (e.g., Hopenhayn and Rogerson (1993)) will yield a decline in reallocation and productivity
2. Decline in pace of innovation/technological change (Gordon (2016)) has led to decline in dynamism/entrepreneurship (Gort and Klepper (1982) and Jovanovic (1982))

- Innovation/entry → Experimentation/Dispersion → Reallocation/Productivity Growth
- This is reflected in variation in the startup rate.

3. Startups are critical for major innovations (Acemoglu et. al. (2017)).
4. Structural changes due to demographics, changes in business model.
   - Aging of population yields decline in startup rate.
   - Change in startup rate might be due to change in business model (retail trade).
5. Latter three focus on changing startups and age structure. Important but far from whole story.

Changing Firm Age Composition Accounts for Substantially Less than Half of Decline in Job Reallocation. More in 1990s (Retail and Services). Less in post 2000s.



Useful to both understand Decline in startups but also Decline in Reallocation Rates Within Age Groups

Illustrative model

Firm's problem:

$$V(A_{et}, E_{et-1}) = A_{et}E_{et}^{\varphi} - w_t E_{et} - C(H_{et}) + \beta V(A_{et+1}, E_{et+1})$$

$$C(H_{et}) = \frac{\gamma}{2}\left(\frac{H_{et}}{E_{et-1}}\right)^2 + F_+ \max(H_{et}, 0) + F_- \max(-H_{et}, 0)$$

$$\ln(A_{et}) = \rho A_{et} + \eta_{et}$$
$$H_{et} = E_{et} - E_{et-1}$$

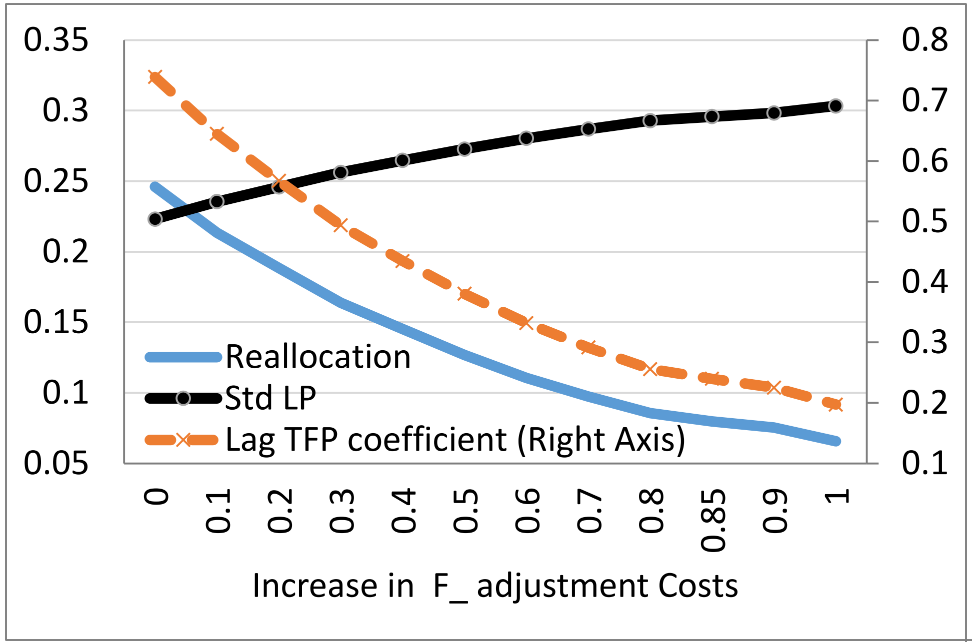
where

- $\varphi \leq 1$ due to product differentiation.
- $A_{et}E_{et}^{\varphi}$ is the revenue function with $P_{et} = Q_{et}^{\varphi-1}$ and $Q_{et} = \bar{A}_{et}E_{et}$
- Study qualitative predictions of changes in $\boldsymbol{\sigma_A}, \boldsymbol{\gamma}, \boldsymbol{F_+}, \boldsymbol{F_-}$
  - For brevity, focus on kinked adjustment costs
  - Note: Common model setup (similar to Cooper & Haltiwanger 2000/2006, Cooper, Haltiwanger, Willis 2007/2014), Elsby & Michaels (2013))

Experiments

- Basic calibration:
  - Standard parameters
  - Choose σ_A to match empirical TFP dispersion
  - Choose F_+(hiring cost) to match 1980s job reallocation rate
    - Qualitative results not sensitive to fixing F_- or γ
- Experiments:
  1. Vary F_- (kinked adjustment cost)
  2. Vary σ_A (TFP dispersion)
- Observe effects on
  - Job reallocation
  - Regression of employment growth on productivity ($g_{it+1} = \alpha + \beta A_{it} + \gamma X_{it} + \varepsilon_{it+1}$)
    - use this moment for data/econometric reasons
  - Dispersion of labor productivity
  - Olley-Pakes covariance (between *size* and *productivity*)
- As we will see, key is focusing on *multiple* moments

Changes in Adjustment frictions



Holding TFP dispersion σ_A and persistence ρ constant, increasing adjustment cost $F\_$ ➔

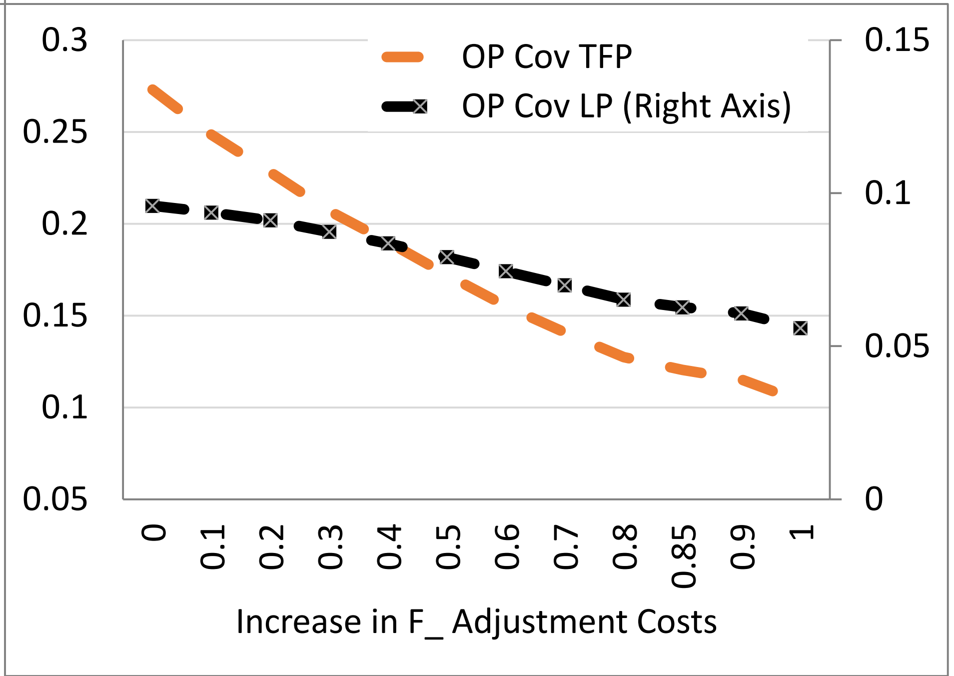
1. **Lower** job reallocation
2. **Lower** *responsiveness* of emp growth to TFP ("Lag TFP coefficient")
   - Note: Same implication for LP
3. **Higher** dispersion of labor productivity

Olley-Pakes productivity decomposition:

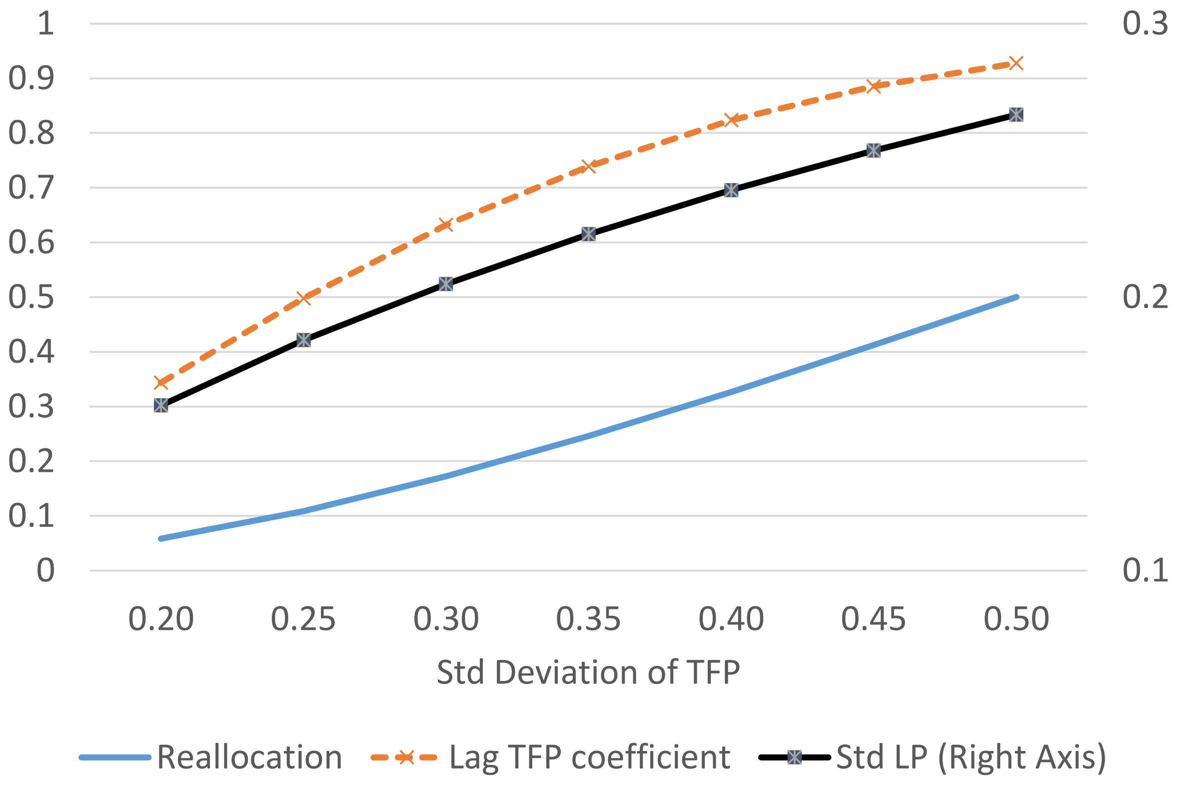
$$A_t = \bar{A}_t + cov(\theta_{it}, A_{it})$$

where θ_{it} is size

➔ Higher adjustment cost reduces aggregate productivity via weaker OP covariance



Changes in Shock Process (e.g., Dispersion and Persistence of Shocks) also can change key moments:



Increase in dispersion of TFP (holding persistence Constant):

1. Increase in job reallocation
2. Increase in dispersion in LP
3. Increase in responsiveness (volatility effect dominates real options effect).

Key difference between frictions vs. shocks is the difference in accounting for a decline in job reallocation is implied changes in labor productivity dispersion. Rise in frictions yields increase in LP dispersion while decline in dispersion of shocks yields decline in LP dispersion.

Toward implications for aggregate (industry-level) productivity

- Olley-Pakes (1996) aggregate (industry) productivity:

$$A_t^{OP} = \sum_i \theta_{it} A_{it} = \bar{A}_{et} + cov(\theta_{et}, A_{et})$$

θ_{it} = input weight
A_{it} = plant TFP

- True aggregate productivity: $A_t = Q_t / E_t$

Q_t = agg output
E_t = agg input

- $A_t = A_t^{OP}$ iff CRS and perfect competition!
  - Maximize A_t^{OP} by moving all resources to most-productive firm
  - Not optimal with revenue curvature because $Q_t = \left(\sum_i Q^{\varphi}\right)^{1/\varphi}$ (focus on *marginal* products)
- Corollary: rise in adjustment frictions (or static distortions) has smaller effect on A_t than A_t^{OP}
- But… A_t and A_t^{OP} highly correlated

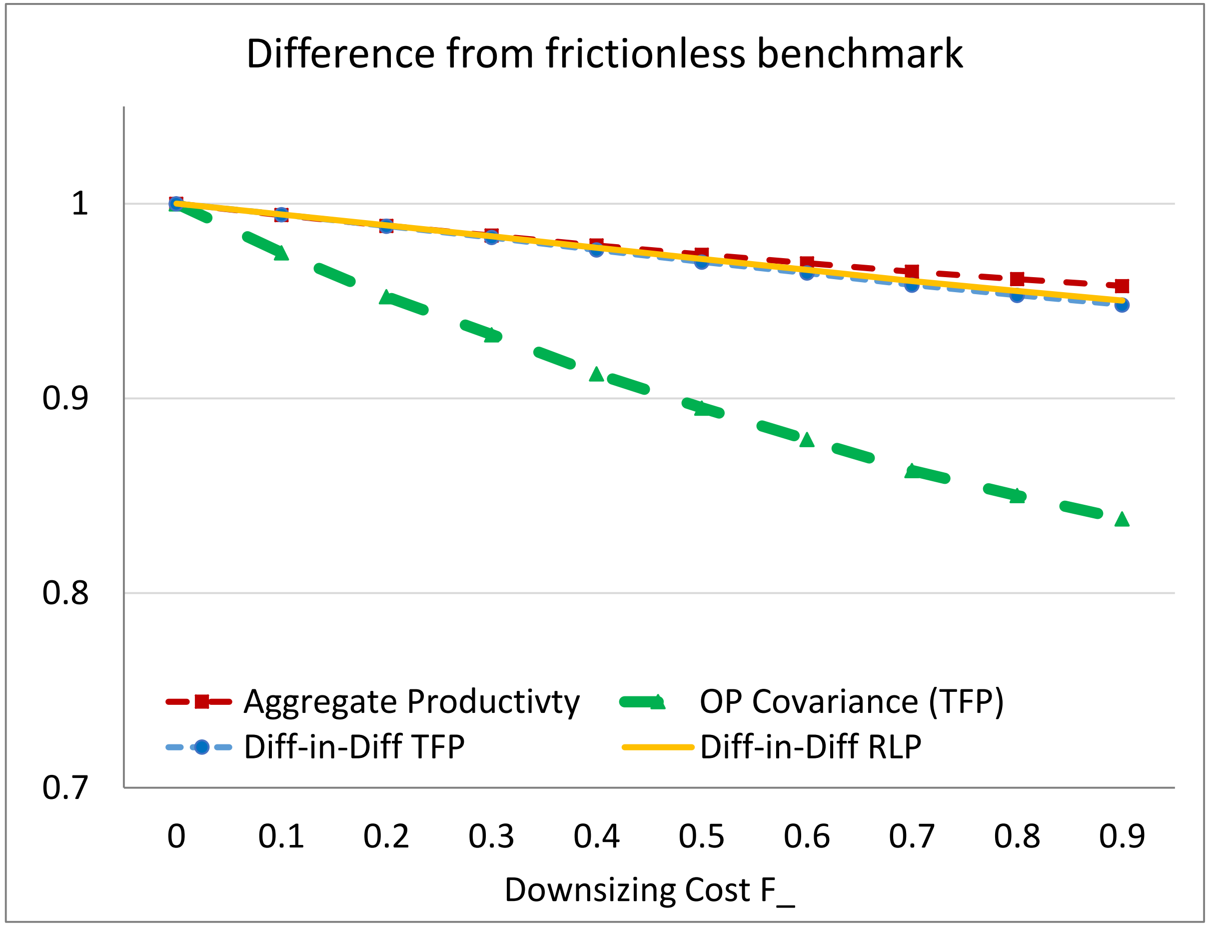
Toward implications for aggregate (industry-level) productivity

- Changing productivity responsiveness implies changes in Olley-Pakes $cov(\theta_{it}, A_{it})$ and therefore A_t^{OP}
- In CRS/perfect competition world, could study misallocation with simple comparisons of $cov(\theta_{it}, A_{it})$ for different responsiveness regimes
- Instead we focus on a 'diff-in-diff':

$$\Delta_t^{t+1} = \sum_i \left(\theta_{i,t+1}^{Trend} - \theta_{i,t+1}^{No\,Trend}\right) A_{it}$$

holding constant everything but responsiveness

Comparing aggregate productivity measures (model)



- As expected, OP covariance declines too quickly in presence of curvature
- But diff-in-diff closely tracks true productivity losses
- Diff-in-diff can be done in microdata
  - Observed plant-level TFP
  - Roll forward θ_{it} weights using estimated responsiveness coefficient

Measuring productivity

1. TFPR from cost shares

$$\ln(TFPR_{et}) = \ln Q_{et}^{R} - \alpha_K \ln K_{et} - \alpha_L \ln L_{et} - \alpha_M \ln M_{et} - \alpha_E \ln E_{et}$$

   - factor α_is are cost shares from NBER/CES productivity database at 6-digit NAICS level
   - Q_{et}^{R} is plant revenue deflated by 6-digit industry deflator
   - TFPR=P*TFPQ under CRTS; within-industry relative TFPR=TFPQ under price taking
   - TFPR highly correlated with TFPQ in presence of adjustment costs (0.75 empirically, Foster et al. 2008,2016; Eslava 2013)

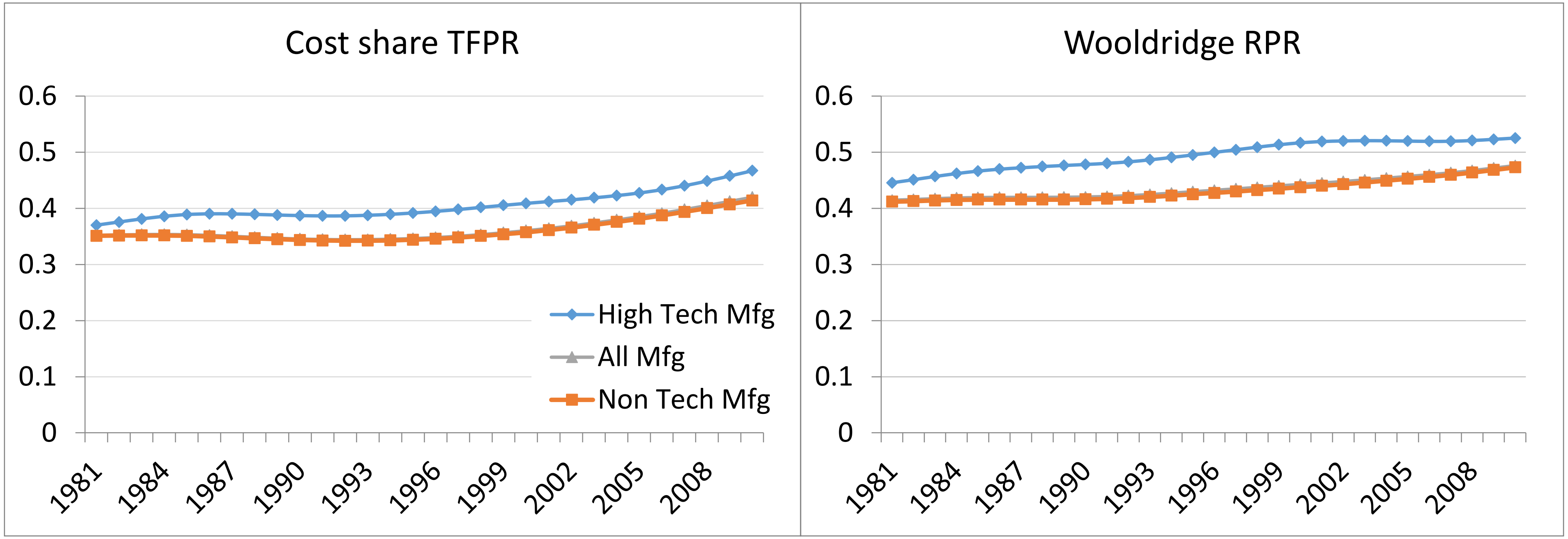
2. Revenue Product Residual (RPR, Wooldridge 2009)
   - Let $P_{et} = D_{et} Q_{et}^{\varphi - 1}$, $Q_{et} = A_{et} \prod_i X_{et}^{\alpha_i}$
   - Then revenue is given by (in logs):

$$p_{et} + q_{et} = \sum_i \beta_i x_{et}^{i} + \varphi a_{et} + d_{et}$$

   - Estimate via proxy methods, parameter estimates are *revenue elasticities* $\beta_i = \varphi \alpha_i$
   - $RPR_{et} = \varphi a_{et} + d_{et}$, only a function of TFPQ and demand shocks!
     - Corr(RPR,TFPR)=0.75

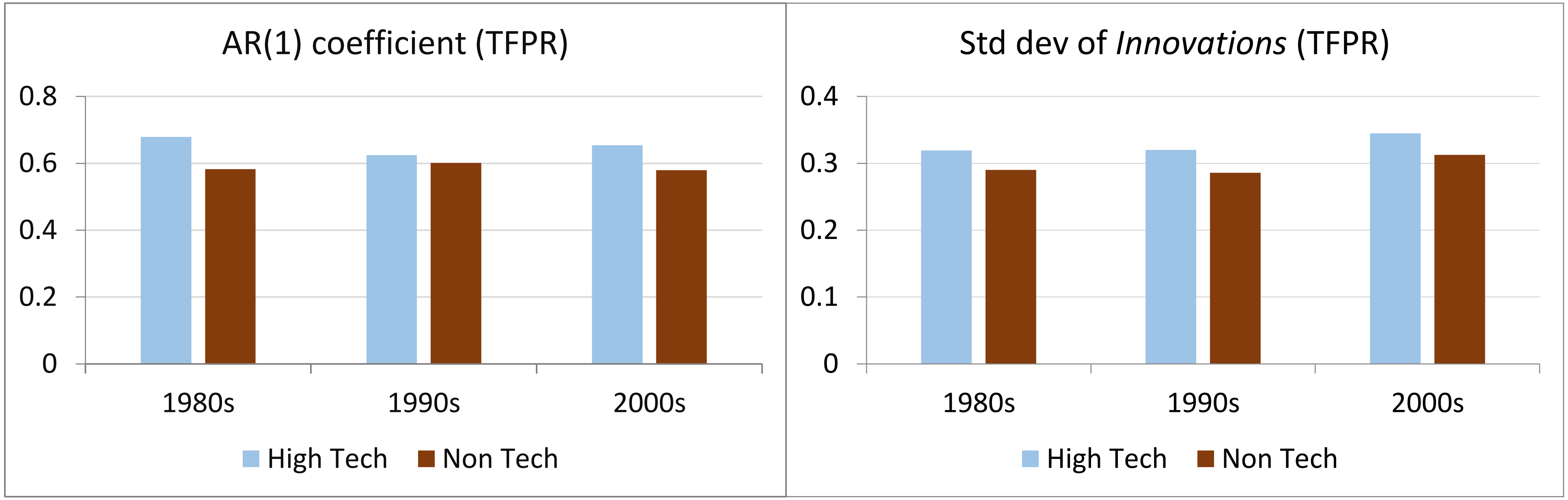
Dispersion of ***shocks***?

Standard deviation of *within-industry* TFP



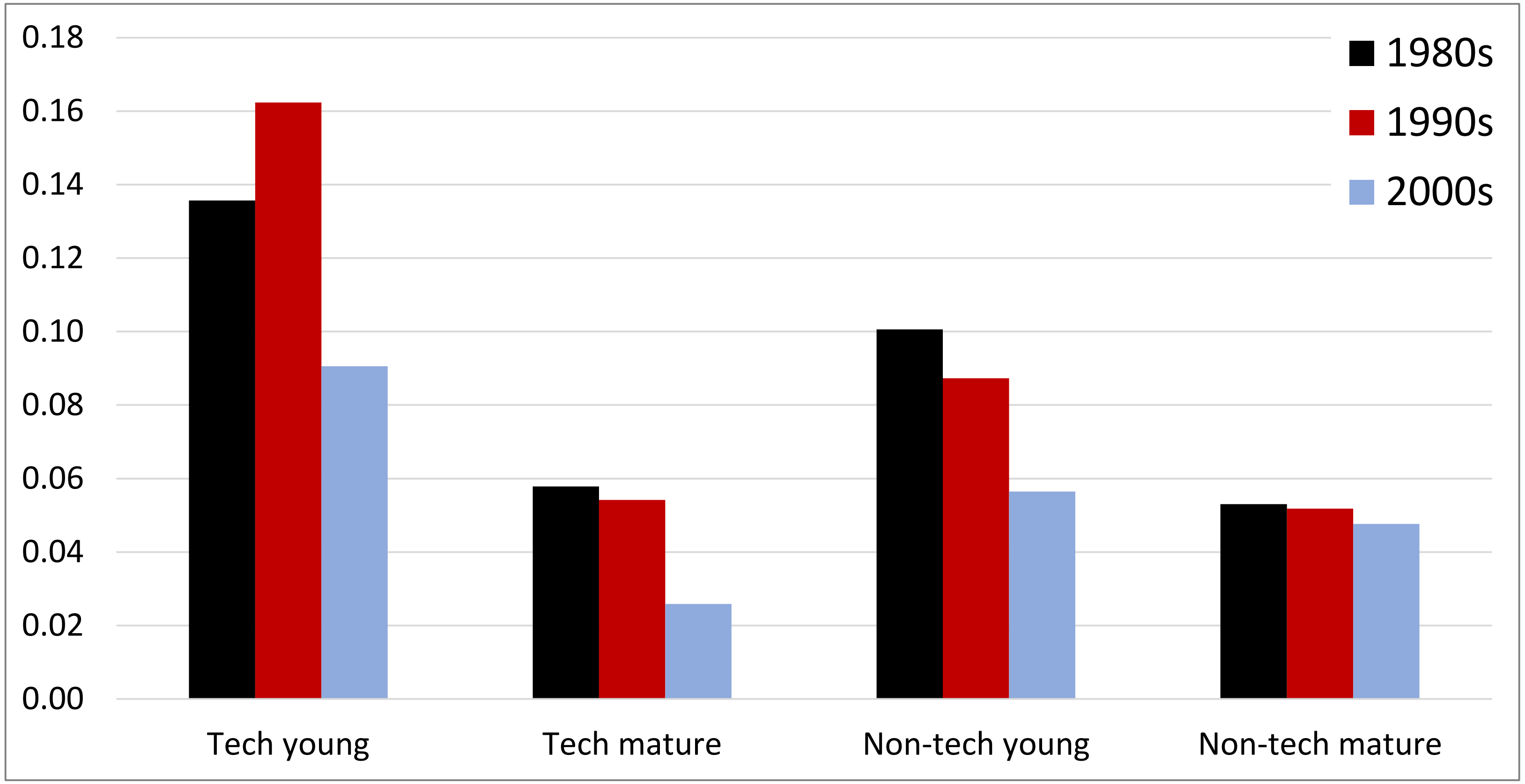
- Similar patterns for TFPR and RPR (corr 0.75)
- Little difference between young and mature firms (not shown)
- **Can "shocks" explain reallocation patterns?**
  - High tech: expect hump-shaped pattern of dispersion
  - Non-tech: expect falling dispersion throughout
- And: for constant adjustment costs, rising TFP dispersion predicts rising 'responsiveness' (β), LP dispersion (is rising, not shown), and OP covariance (later)

Other shock moments



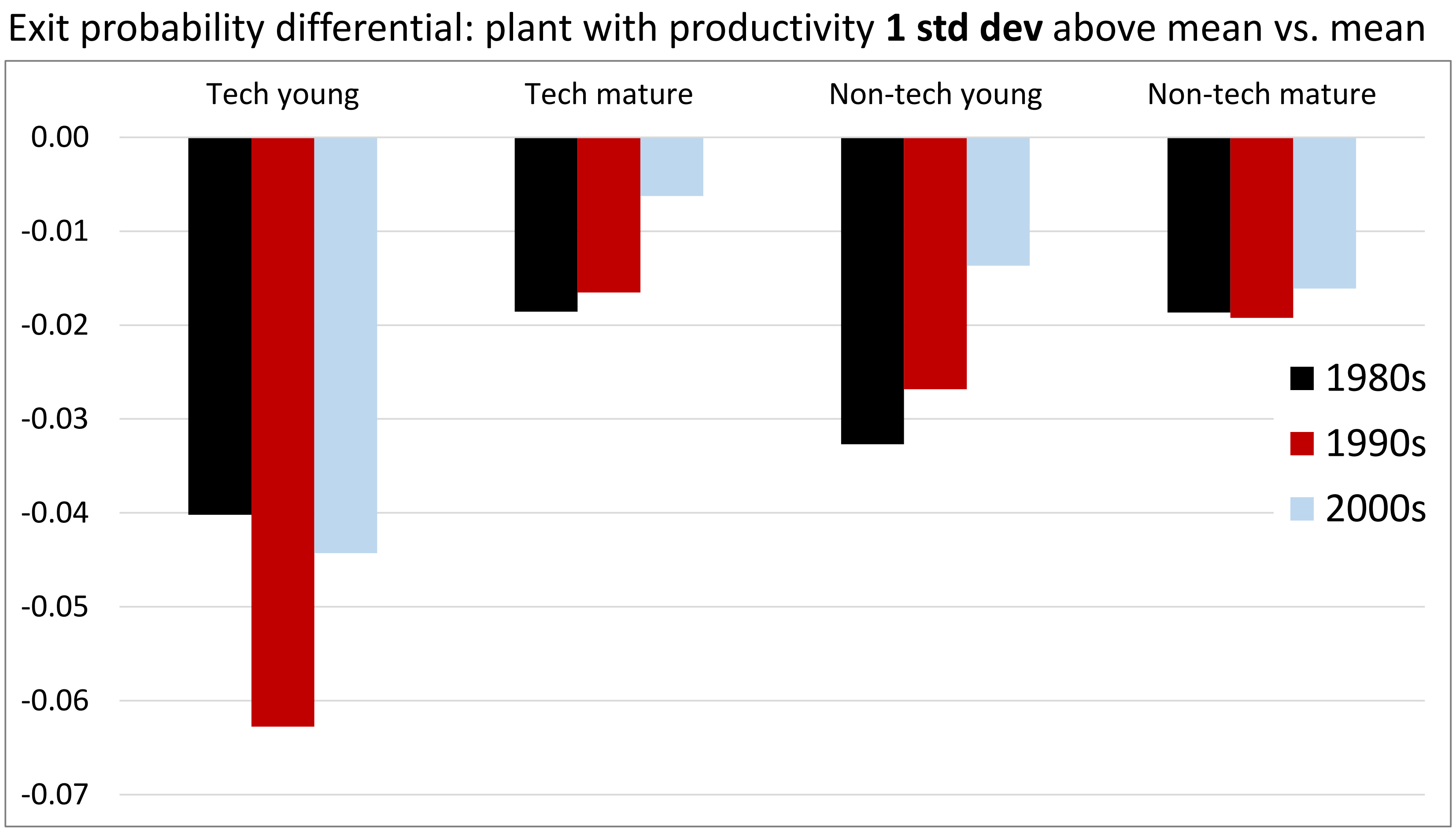
- Little evidence that changes in persistence or other shock moments drive reallocation patterns
- Similar results for RPR

Growth differential: plant with productivity **1 std dev** above mean vs. mean



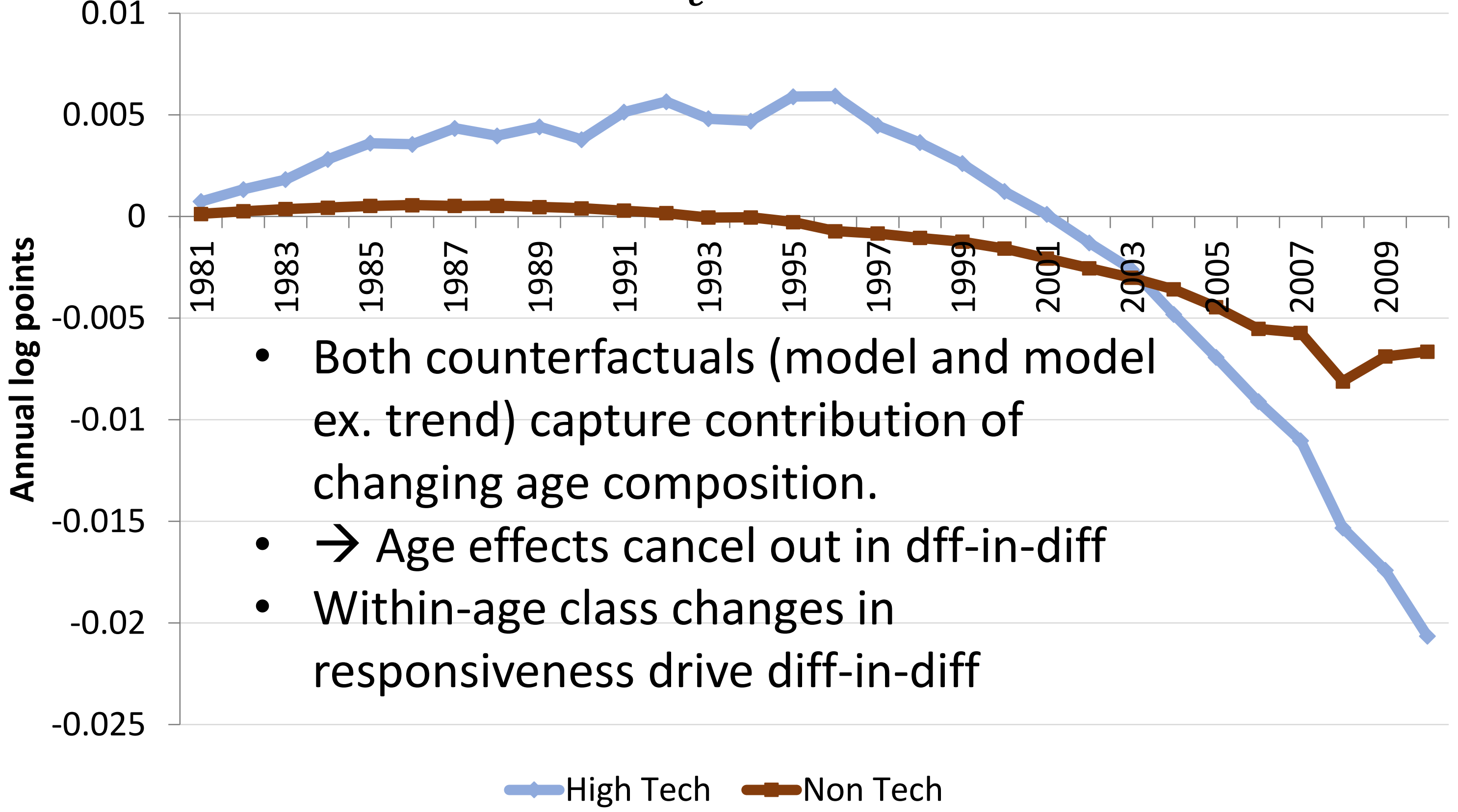
- Non-tech young:
  - **1990s**: TFP 1 std dev above industry mean ➔ **9 ppts** growth differential
  - **2000s**: TFP 1 std dev above industry mean ➔ **6 ppts** growth differential
- Non-tech mature:
  - **1990s**: TFP 1 std dev above industry mean ➔ **5 ppts** growth differential
  - **2000s**: TFP 1 std dev above industry mean ➔ **4.5 ppts** growth differential

Exit selection weaker too



Diff-in-diff accounting counterfactual for changing responsiveness

$$\Delta_t^{t+1} = \sum_e (\theta_{i,t+1}^{T} - \theta_{i,t+1}^{NT}) A_{it}$$



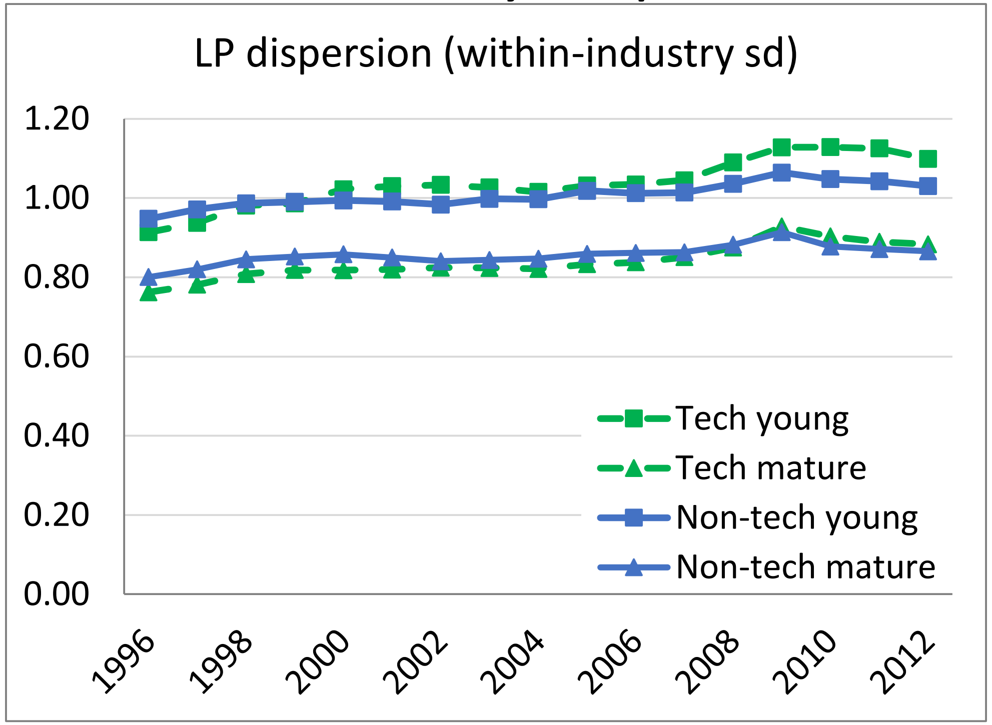
- Both counterfactuals (model and model ex. trend) capture contribution of changing age composition.
- → Age effects cancel out in dff-in-diff
- Within-age class changes in responsiveness drive diff-in-diff

Taking stock

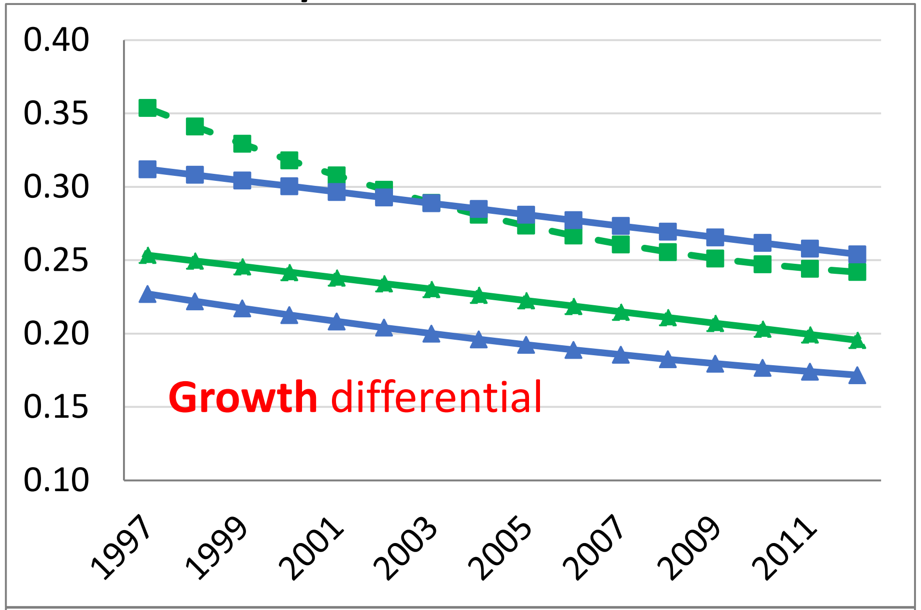
- Reallocation decline not driven by less “shocks” dispersion
  - TFP dispersion rising gradually
- Productivity responsiveness has weakened
  - Consistent with rising adjustment costs
  - Adjustment cost effect dominates effect of rising TFP dispersion
- (Not shown) Labor productivity dispersion also rising
  - Consistent with rising TFP dispersion and/or rising frictions
- Slowing responsiveness means lower activity weights for higher-productivity businesses
  - Next… study implications for aggregate productivity via OP covariance

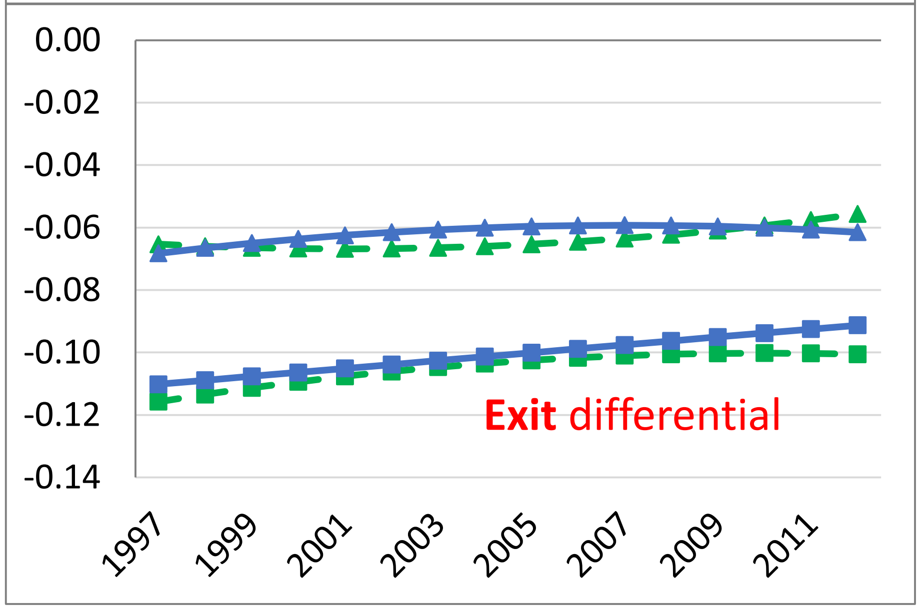
Economywide results

Productivity Dispersion



"Responsiveness"

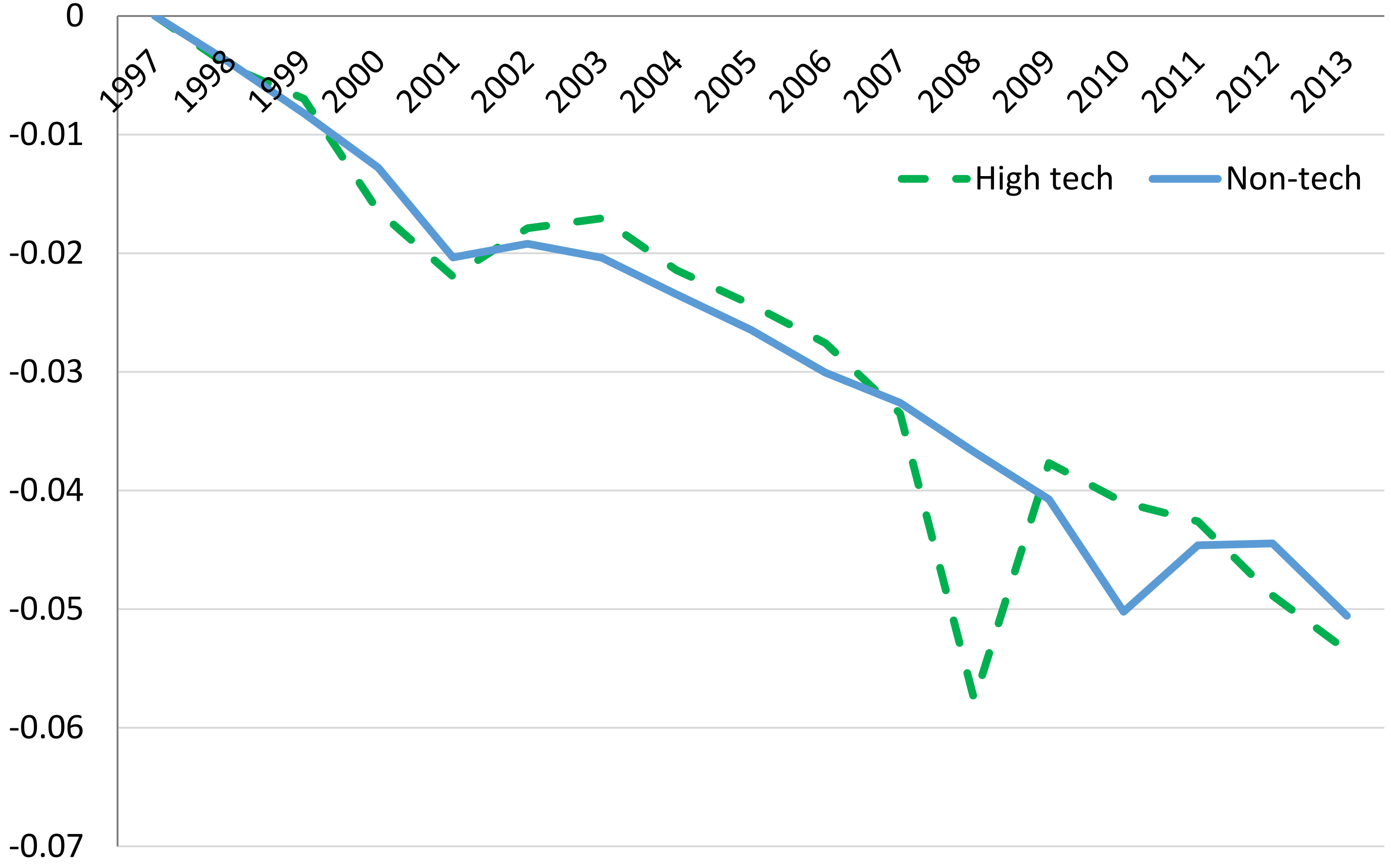




- Rising, not falling, labor productivity dispersion
  - Consistent with rising TFP dispersion
  - *Administrative data*: not a function of sampling/imputation/related error
  - No sign of Gort & Klepper "industry shakeout"
  - Note: rising in Manufacturing too
- Declining *growth* responsiveness
- Weakening *exit* selection

Economywide diff-in-diff counterfactual

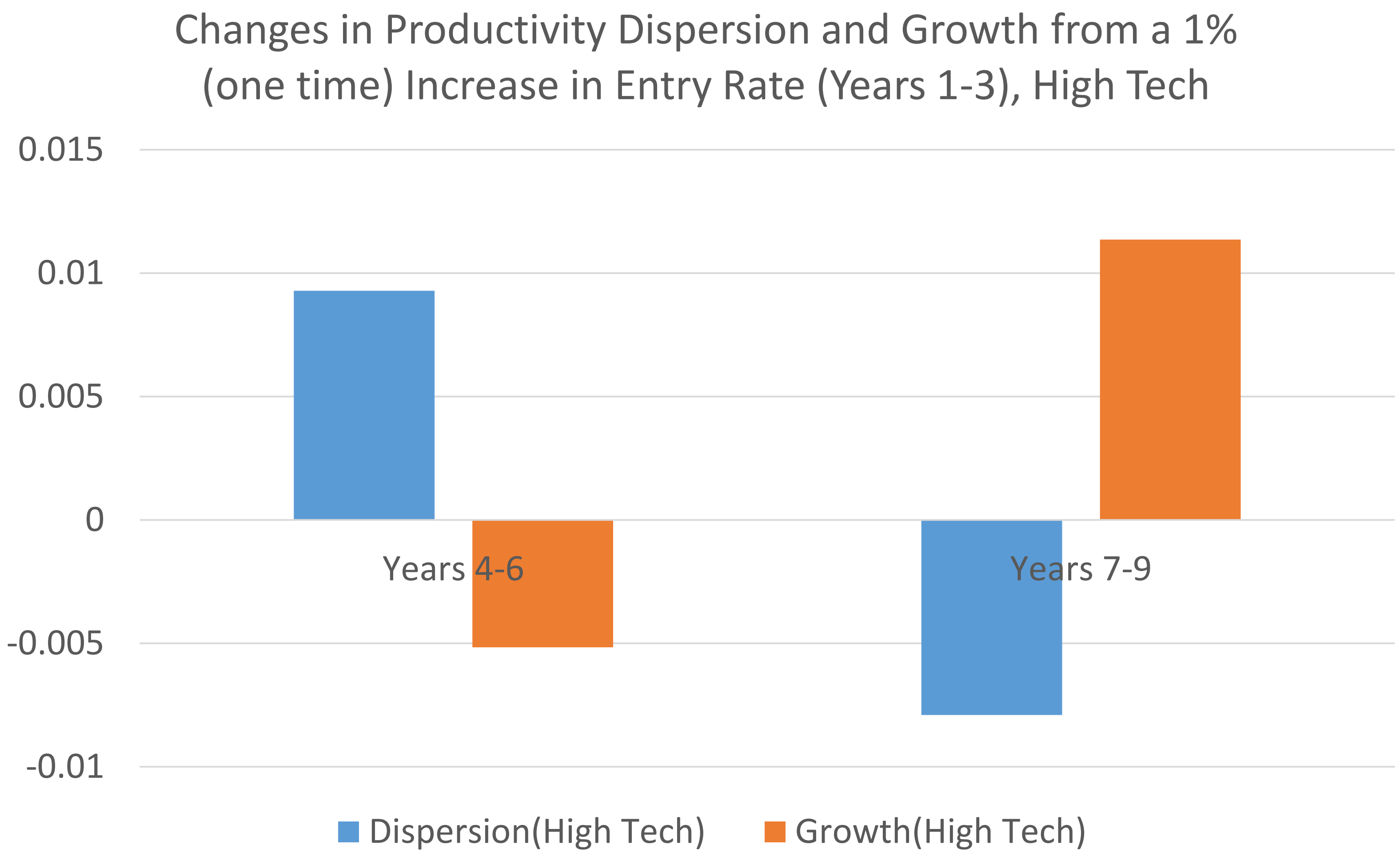
$$\boldsymbol{\Delta_t^{t+1} = \sum_e (\theta^T_{e,t+1} - \theta^{NT}_{e,t+1}) P_{et}}$$



Creative Destruction – Role of Young Firms

- Two related perspectives:
  - Gort and Klepper (1982)
    - Innovation in industry accompanied by surge in entry and experimentation
    - During experimentation phase, high dispersion of productivity and perhaps decline in productivity.
    - Then shakeout/consolidation phase. Productivity growth emerges as successful innovators expand and unsuccessful entrants contract and exit.
    - Their evidence shows business formation and evolution of firm counts with specific innovations (e.g., TVs vs. Tires vs. Lasers)
  - Acemoglu et. al. (2017) and Ackcigit and Kerr (2017):
    - Evidence and model that young firms make major innovations, mature firms minor (defensive) innovations.
- Both perspectives suggest that innovation closely linked to entry/young firm activity.
  - Decline in U.S. Productivity Associated with Decline in Young Firm Activity

Dynamics of Entry, Productivity dispersion and Productivity growth



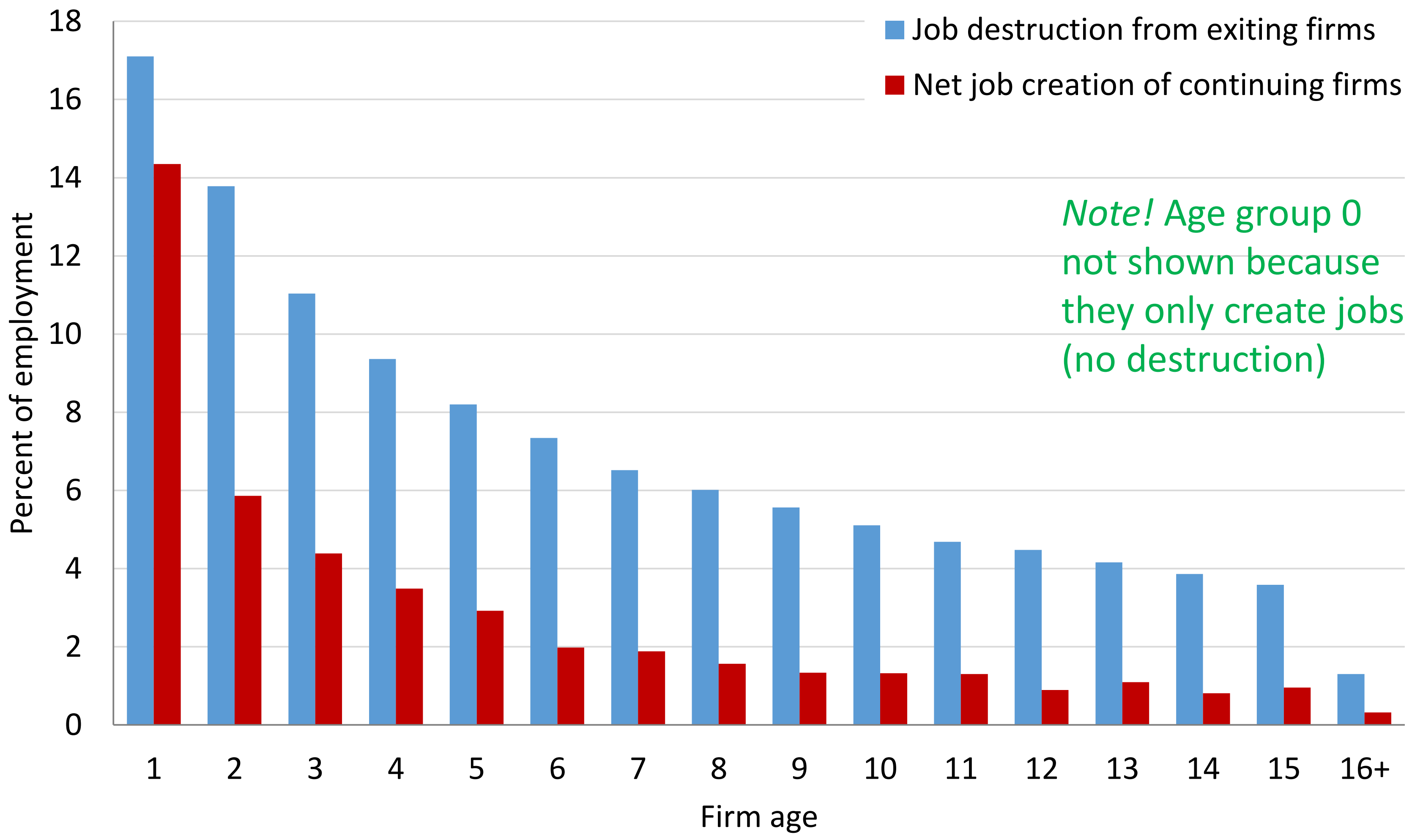
Surge in entry in a given 3-year period leads to:

- Rise in within industry productivity dispersion and decline in industry productivity growth in next 3-year Period
- Decline in within industry productivity dispersion and rise in industry in subsequent 3-year period
- Surge in reallocation following surge in entry as well (not depicted).
- Similar, dampened patterns for Non-Tech

Using 4-digit NAICS data for High Tech sectors (ICT in mfg and non-mfg plus sectors such as Bio Tech)

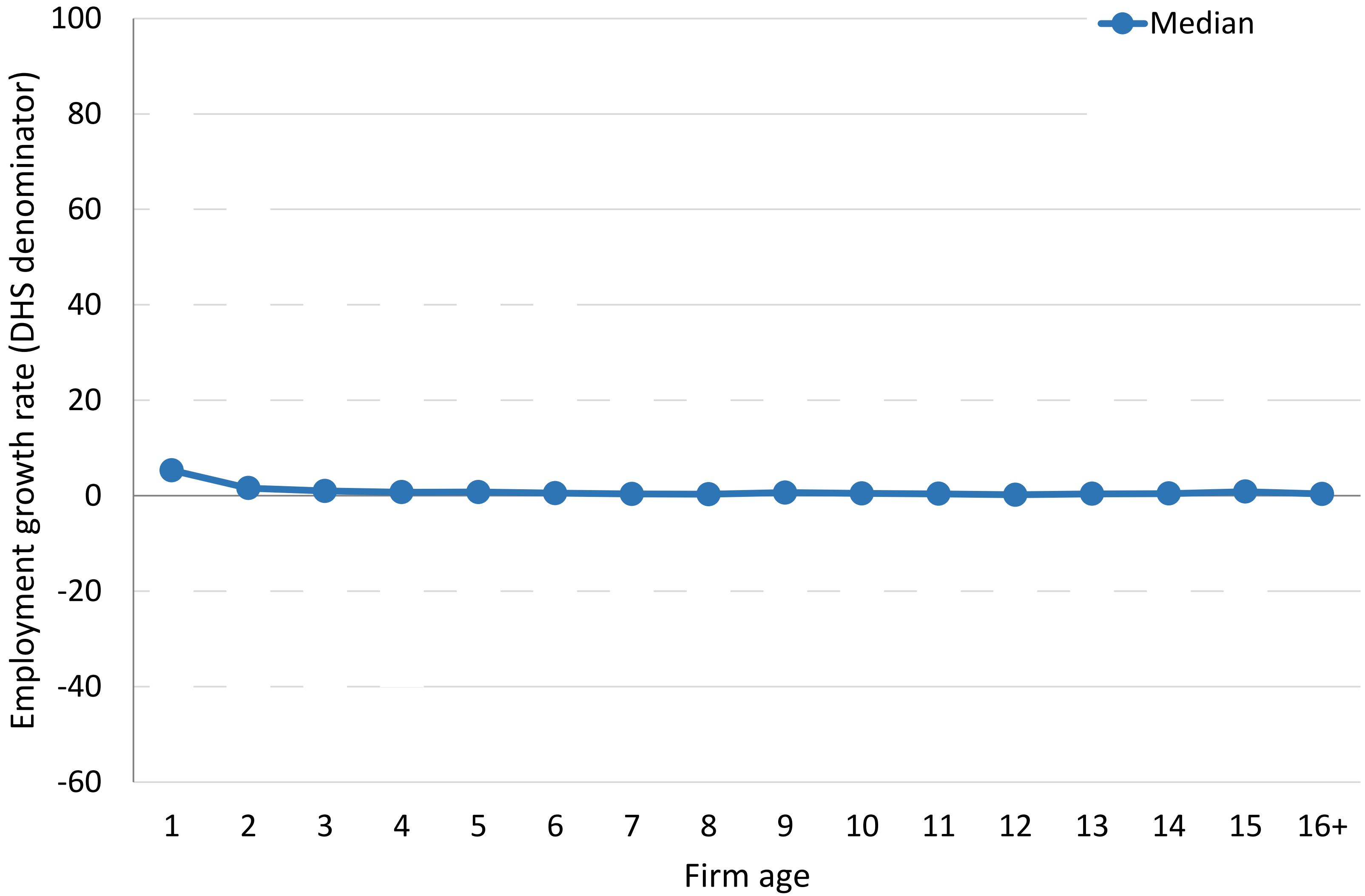
Source: Foster et. al. (2018)

Up or out!



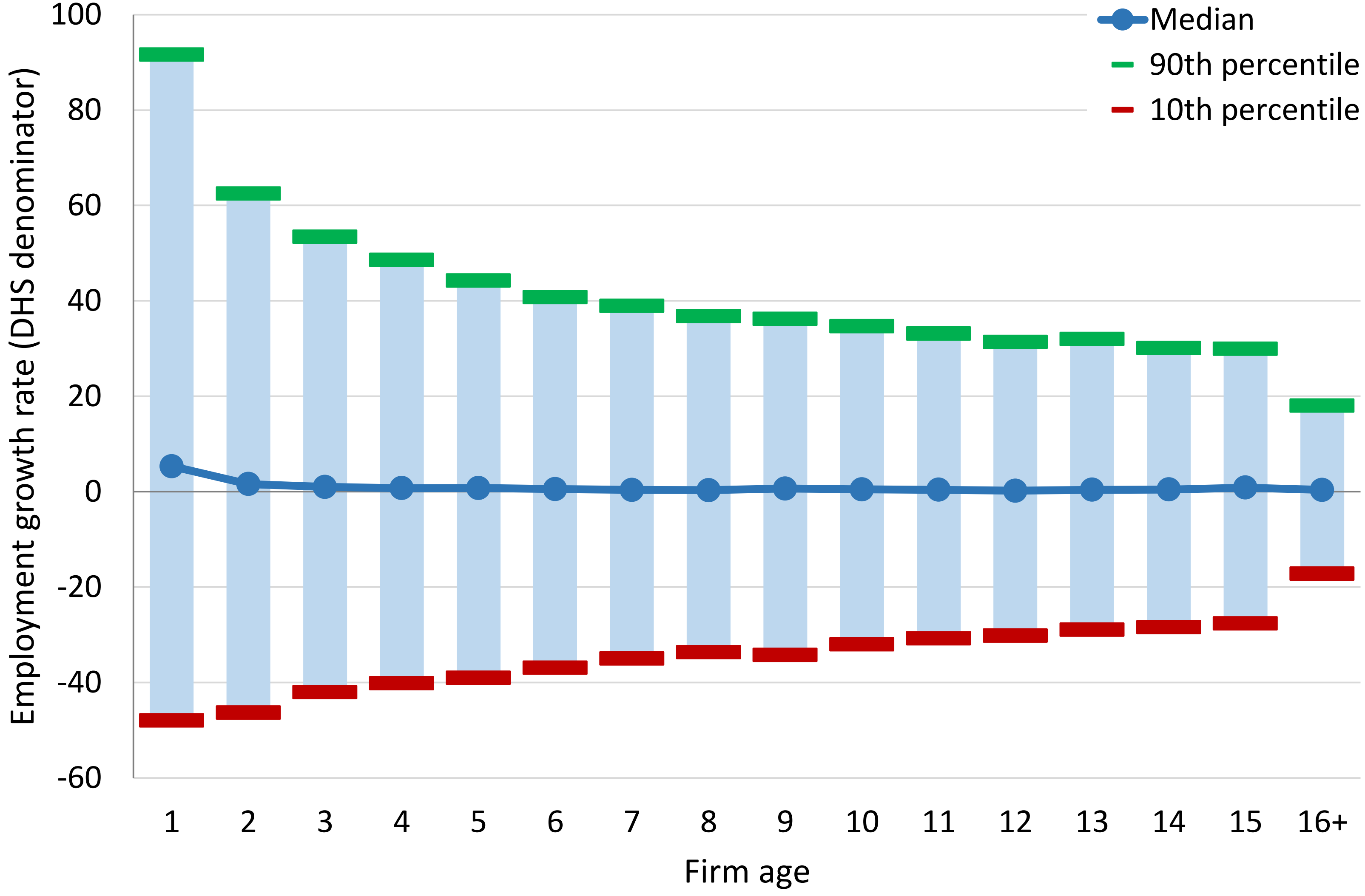
Source: Decker et al. (2014)

A view of the skew



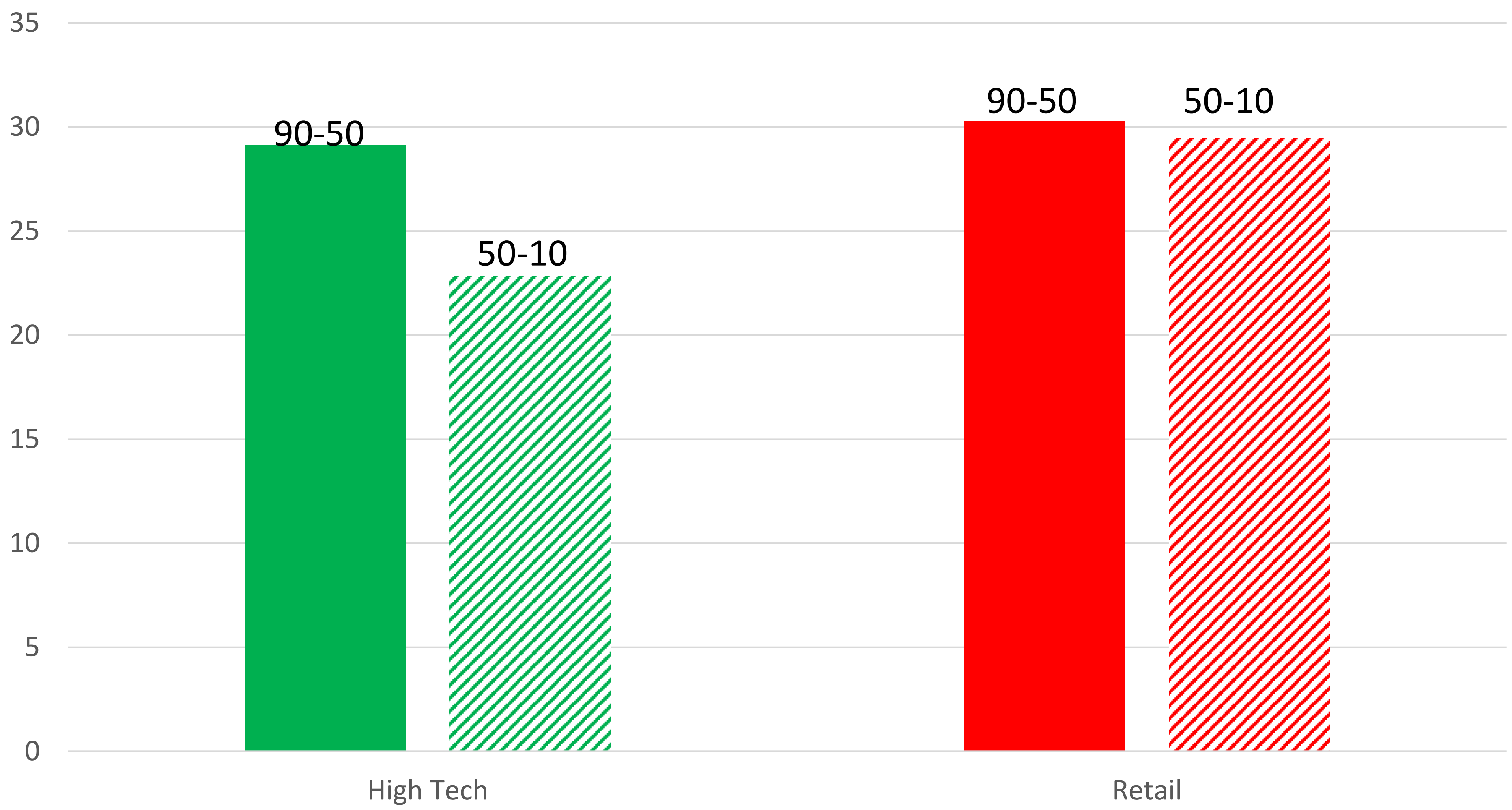
Source: Decker et al. (2014). Employment-weighted distributions.

A view of the skew – High Growth Firms are Disproportionately Young Firms



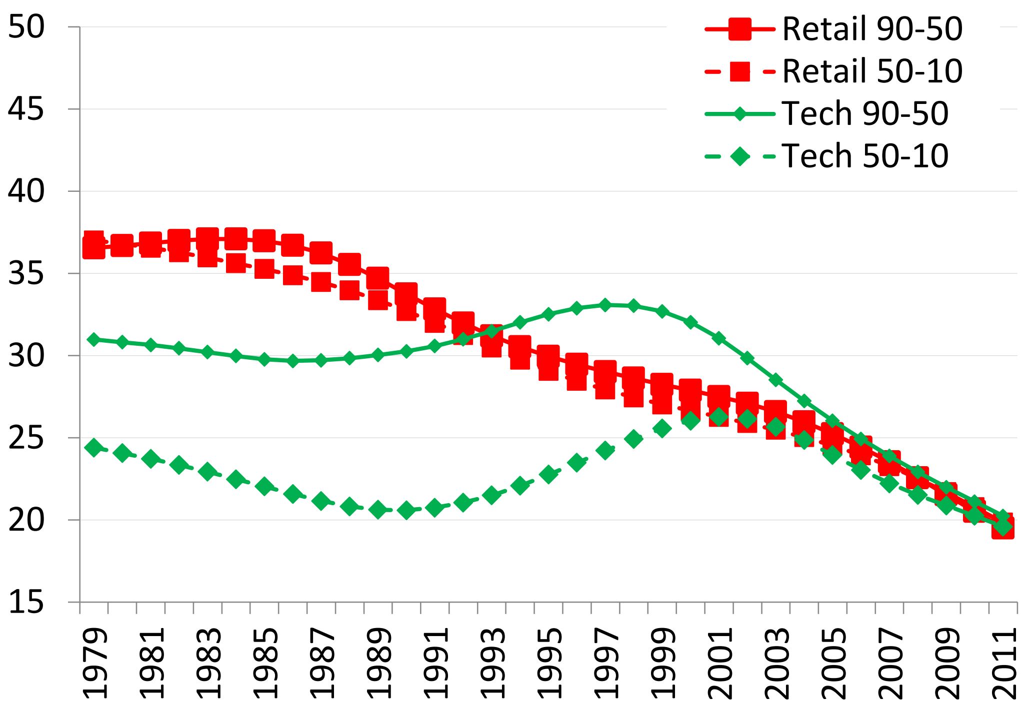
Source: Decker et al. (2014). Employment-weighted distributions.

Large Differences in Skewness Across Sectors – High 90-50 in High Tech Driven by Young Firms

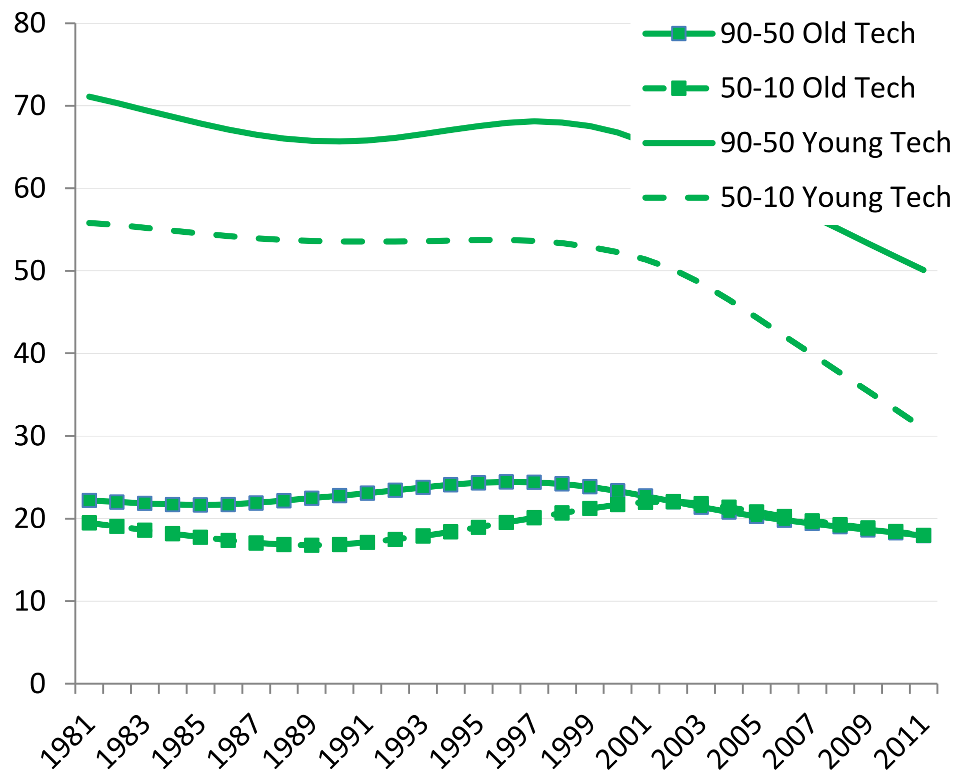


Source: Decker et. al. (2016)

Times series patterns of skewness (high growth) vary dramatically across sectors



Skewness primarily accounted for by Young Firms. In High Tech, Decline in young firms and decline In High Growth Firms in High Tech

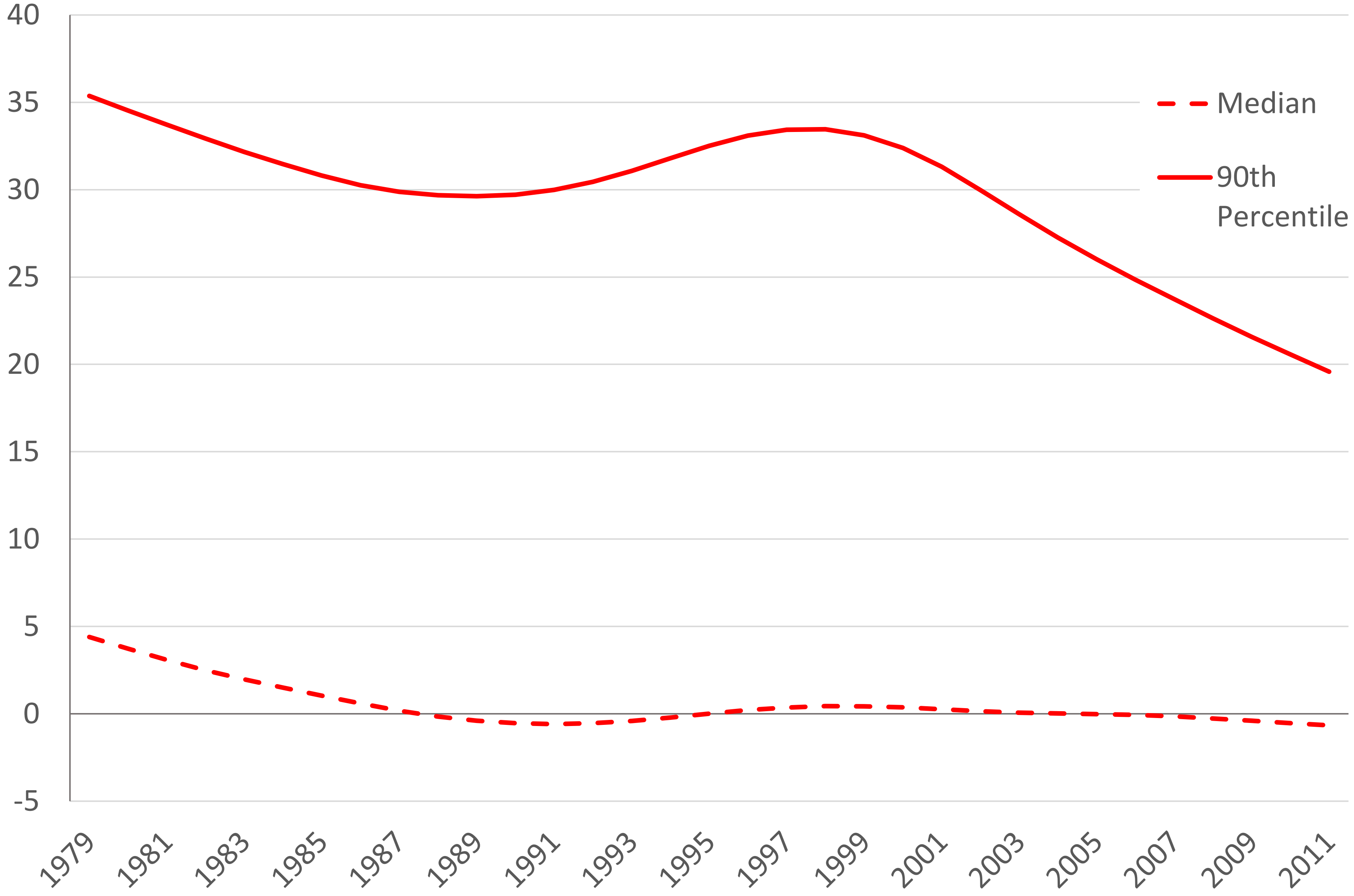


Retail: dispersion decline equal parts 90-50, 50-10

High Tech: Growing Skewness in 1990s, sharp Decline post 2000

High Tech includes (most of) Information but also High Tech Mfg and Services. Source: Decker et. al. (2016)

High Growth vs. Median Growth Firms in High-Tech (Employment-Weighted Distribution)



Source: Tabulations from the Longitudinal Business Database (Census). HP Trends depicted

Facts and Open Questions

- Periods of rapid innovation (especially in innovative intensive industries like High Tech):
  - First surge of entry
  - Then experimentation (dispersion)
  - Then productivity growth
  - Potentially long (and variable) lags
- Both innovative intensive industries (High Tech) and other industries have seen relatively modest entry and productivity growth post 2000.
  - Part of declining entry, dynamism and labor market fluidity post 2000.
- Dispersion in Productivity Growth in High Tech and Non Tech has risen substantially in the post 2000 period
  - Experimentation that has not yet resolved?
  - Diminished Dynamism – Slower diffusion or slower adjustment dynamics?
  - If in experimentation phase why has entry declined? Where is dispersion coming from?

Broader issues

- What is the role of entrepreneurship and reallocation for growth in China?
- One view (Hsieh and Klenow (2009)) is that China has had (still has?) a very large rate of productivity dispersion relative to U.S.
  - This reflects frictions/distortions.
  - But there is a silver lining to this dispersion. As frictions/distortions are reduced, economic and productivity growth can be substantial as allocative efficiency increases.
  - How important has this been empirically?
  - What are the frictions/distortions that have been reduced?
  - Are these reductions in “static” or “dynamic” frictions (or both)?