# Information Technology, Firm Size, and Industrial Concentration

Erik Brynjolfsson (Stanford University)
Wang Jin (Stanford University)
Georgios Petropoulos (University of Southern California)
Xiupeng Wang (Northeastern University)

CRETE
Tinos, July 8, 2025

# FEDERAL STATISTICAL RESEARCH DATA CENTERS DISCLAIMER

Any views expressed are those of the authors and not those of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

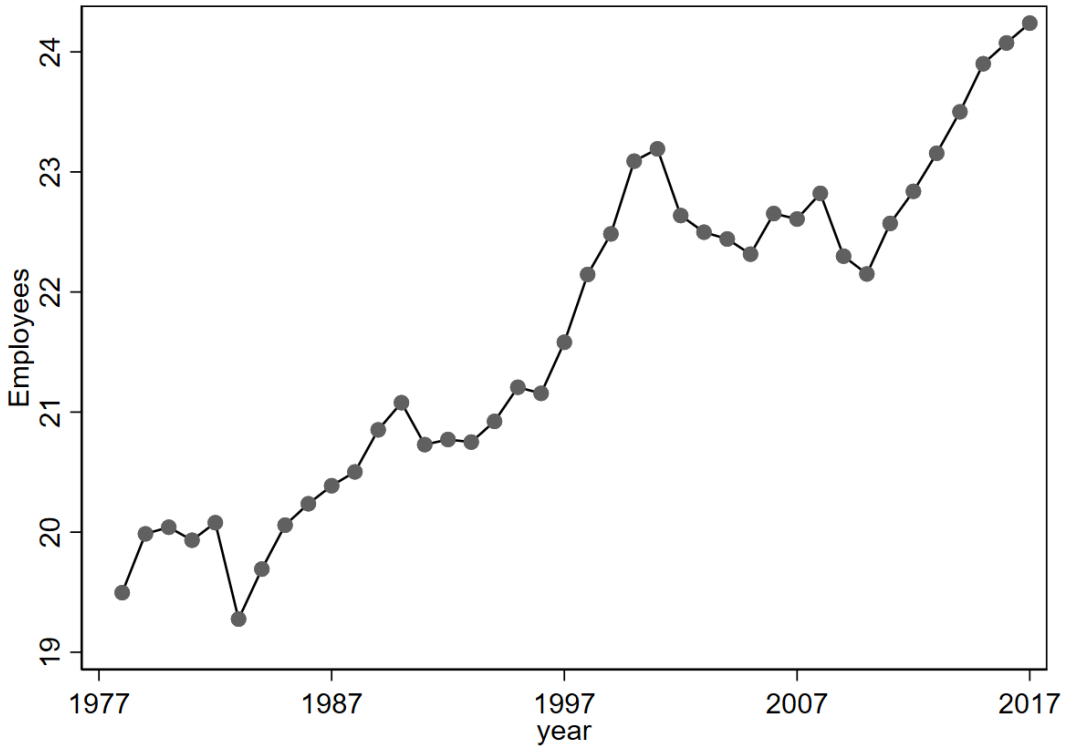
# RESEARCH QUESTIONS

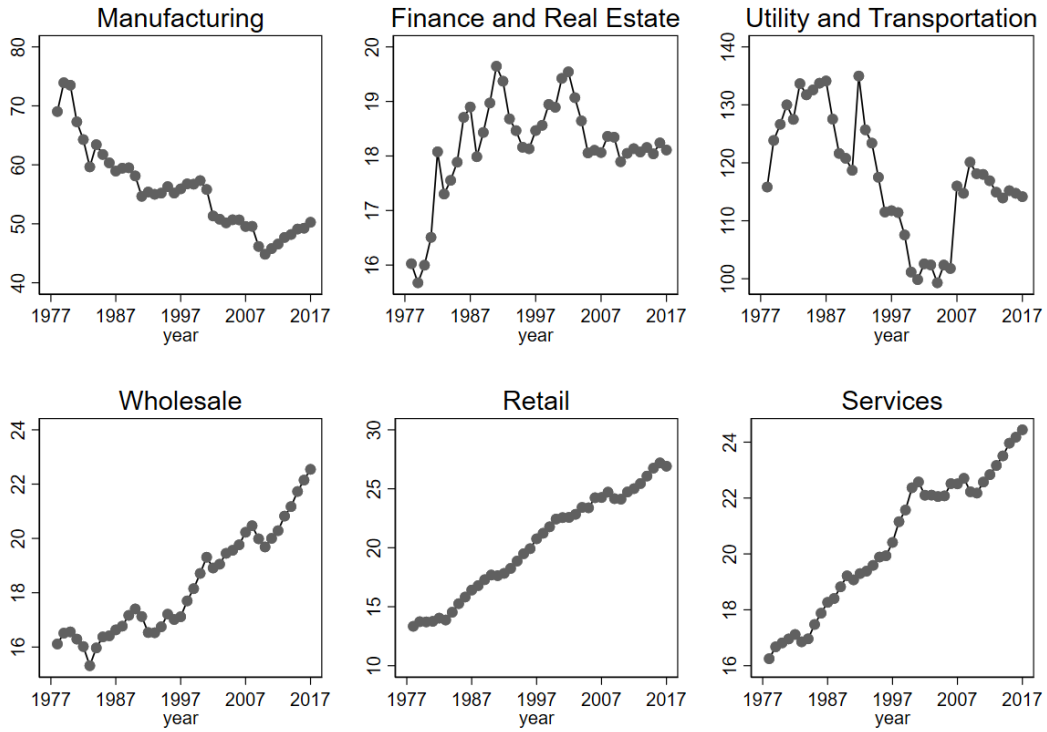
- A theoretical model that points towards a scaling without mass story

- Take the advantage of micro-level census data to study

1. Does IT increase firm size in general?
2. Does the effect particularly favor larger firms?
3. Does it contribute to the increased market concentration?
4. How does IT affect labor share?
5. Do firms with higher IT investment enter in more local markets?

# STYLIZED FACTS: AVERAGE SIZE

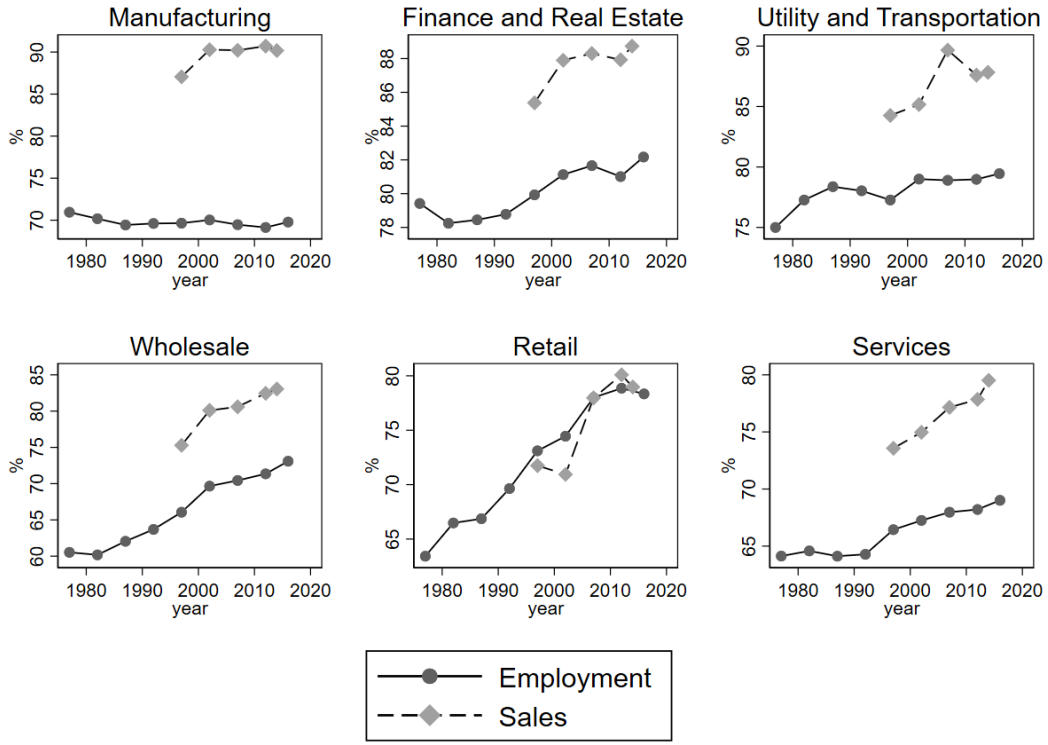



1. On average, the US economy has shown a increasing average employment per firm.
2. Manufacturing sector has shown an overall declining firm size.
3. Wholesale, Retail, and Services show a significant increase of average firm size.

(Data source: Business Dynamics Statistics; by authors' calculation)

# STYLIZED FACTS: INDUSTRIAL CONCENTRATION

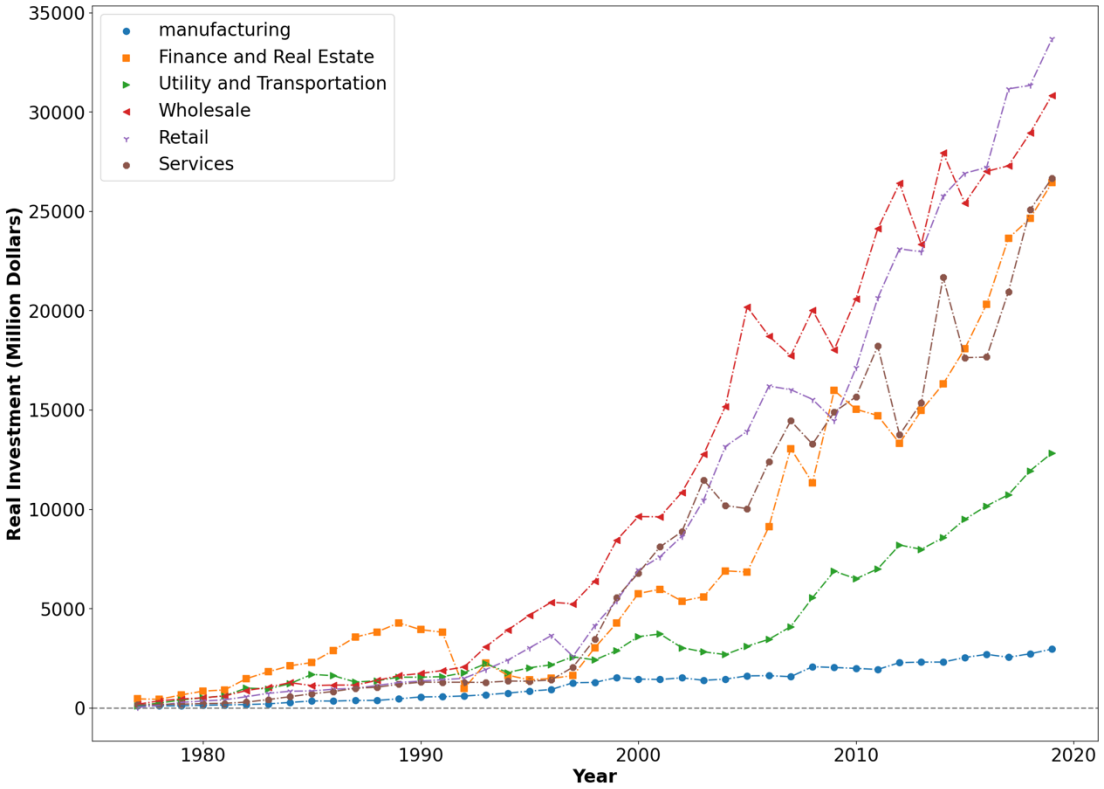



1. Share of top 10% firms

2. Δ log employment share from 1977-2013 of top 10% firms:
   - Wholesale: 0.175
   - Retail: 0.205
   - Service: 0.079

3. Similar to Hsieh and Rossi-Hansberg 2020

# STYLIZED FACTS: IT INVESTMENT




1. Sectors with high growth in industry concentrations are also the ones with the most rapid growth of IT expenditure.
   - Retail
   - Wholesale
   - Services
   - Finance

2. Research at the industry level finds a strong link between IT and rising industry concentration (Bessen, 2020) .

(Data source: Bureau of Economic Analysis; by authors' calculation)

# THEORETICAL MODEL

- Consider an economy with $N$ distinct locations or cities and a population of n firms.
- A total number of $n_{it} \geq 2$ firms out of $n$ is present (with a local branch) in city $i$ and in period $t$, where $i = 1, ..., N$ and $t = 0, 1, ..., \infty$.
- In each city $i$ and period $t$, active firms compete à la Cournot and sell a homogeneous good to consumers.
- Each firm $j$'s local branch in city $i$ and in period $t$ simultaneously chooses its individual output level $q_{ijt}$.

# OUTPUT PRODUCTION COST

- Output production requires each firm to incur a production cost that is equal to its marginal cost $c_{ijt}$ multiplied by the output level produced by its local branch.

- This marginal cost of firm $j$ in city $i$ and in each period $t$ is drawn from a known continuous distribution, $F(c_{ijt})$ and a positive probability density function $f(\cdot)$ only within the finite support $c_{ijt} \in [\underline{c}, \bar{c}]$, where $\bar{c} > \underline{c} \geq 0$.

- If firm $j$ has active branches in period $t$ in $N_{jt}$ cities, then, it draws $N_{jt}$ marginal costs, one for each city it is active.

# IT AND KNOWLEDGE TRANSFER

- Let $c_{jt}^{min} = \min\{c_{1jt}, c_{2jt}, ..., c_{N_{jt}jt}\}$.
- Let also $\varphi_{jt} \in [0, 1]$ be the ability of firm $j$ in period $t$ to transfer knowledge from its most efficient branch with $c_{jt}^{min}$, say $k$, to its other branches.
- For each city $i \neq k$, the marginal cost can be reduced from the initially drawn value $c_{ijt}$ to value $\tilde{c}_{ijt}$, such that: $c_{ijt} - \tilde{c}_{ijt} = \varphi_{jt}(c_{ijt} - c_{jt}^{min})$.
- Variable $\varphi_{jt}$ is strictly increasing in IT investment.

# IT INVESTMENT PHASE

- Firms commit to their spending on IT capital before they observe the marginal costs of competitors and before they set quantities.

- IT capital expenditure is considered as a sunk cost when firms make their output decisions.

- Two considerations (inspired by recent literature)
  - Firms invest in IT in each period t. Justification: High depreciation rate and cloud-enabled IT services (De Ritter, AER, 2024).
  - Firms have to pay a per-period overhead cost which is a convex function of the number of cities they span (Aghion et al., RES, 2023)

# TIMING

- In each period t

1. Firms invest in IT

2. After each firm j observes the marginal costs of its competitors and the realization of its own marginal cost in each city, it chooses (simultaneously with its competitors) its output level per city qijt and chooses in how many cities it will be active. It accordingly hires labor.

# PRODUCTION AND DEMAND FUNCTIONS

- The production function of firm $j$ in city $i$ and in period $t$ is:

$$q_{ijt} = \frac{wl_{ijt}}{\tilde{c}_{ijt}}.$$

- The inverse demand function in city $i$ and time period $t$ is $P_{it}(Q_{it}) = A_i - b_i Q_{it}$, where $Q_{it} = \sum_{j=1}^{n_{it}} q_{ijt}$ is the total market output and $A_i, b_i > 0$ are city-specific demand parameters.

- the profit function of firm $j$ from city $i$ in time period $t$ is

$$\pi_{ijt}(q_{ijt}, c_{ijt}; \varphi_{jt}) = (A_i - b_i Q_{it} - \tilde{c}_{ijt}) q_{ijt}.$$

# COURNOT EQUILIBRIUM

**Proposition (Cournot equilibrium output and profit.)**

*The equilibrium output of firm $j$ in city $i$ in time period $t$ is*

$$q^*_{ijt} = \frac{1}{b_i}\left(\frac{A_i + \sum_{y \neq j} \tilde{c}_{iyt} - n_{it}\tilde{c}_{ijt}}{n_{it}+1}\right). \quad (1)$$

*The corresponding profit for firm $j$ in city $i$ in time period $t$ is*

$$\pi^*_{ijt} = (A_i - b_i Q^*_{it} - \tilde{c}_{ijt}) q^*_{ijt}, \quad (2)$$

*where,* $Q^*_{it} = \sum_{j=1}^{n_{it}} q^*_{ijt}$.

# EMPIRICAL PREDICTIONS

**Empirical Prediction 1 (Firm sales and employment):**
Investment in IT allows the firm to operate on a greater scale with more sales. The effect of IT on firm employment is ambiguous.

**Empirical Prediction 2 (Labor share):**
IT leads to a decline in labor share.

**Empirical Prediction 3 (New markets):**
Firms with higher investment in IT set up more establishments and enter more local markets.

**Empirical Prediction 4 (Increased concentration):**
The effect of IT investment is more pronounced among firms with greater sales increasing the industrial concentration rates.

# DATA SOURCES

## Revenue-enhanced Longitudinal Business Database (LBDRev) (1976-2018)

- Build on Business Register data (often used alongside the BR)
- Over 8.5 million total business establishments in 2013
- Provide key measures on employment, sales, payroll

## Annual Capital Expenditures Survey (ACES) (1996-2019)

- All domestic, private, non-farm businesses
- Annual data on expenditures on new and used structures and equipment
- Source for the BEA industry-level data of private nonresidential fixed assets

## Information and Communication Technology Survey (ICTS) (2003-2013, discontinued)

- Supplemental to the Annual Capital Expenditure Survey (same coverage)
- Collects data on non-capitalized and capitalized business spending for Information and Communication Technology (ICT) equipment and computer software

# EMPIRICAL METHODOLOGY

$$Size_{i,j,t(t+1)} = \alpha + \beta_1 IT_{i,j,t} + \beta_2 nonIT_{i,j,t} + \beta_3 IT_{i,j,t-1} + \beta_4 nonIT_{i,j,t-1} + \beta_x X_{i,j,t} + \varepsilon_{i,j,t}$$

$Size_{i,j,t}$ is the natural logarithm of the size variables for firm $i$ in industry $j$ in year $t$.

Here we use both firm employment and firm sales.

$IT_{i,j,t}$ indicates the measure on firm's IT expenditure.

IT intensity defined as IT expenditure per worker (Tambe et al., 2020)

$nonIT_{i,j,t}$ is the firm's non-IT expenditure.

Non-IT expenditure per worker

$X_{i,j,t}$ includes year, industry, and firm fixed effects.

$\varepsilon_{i,j,t}$ is the error term.

## ❑ Identification strategy

- We follow Bessen (2020) to adopt the share of sedentary employees at the industry level as the instrumental variable.
  - Sedentary workers are more likely to work with computers and other IT technologies.
  - Firms in industries with higher proportion of sedentary workers would therefore tend to adopt new IT technology earlier and more intensively.
  - Building upon DOT, England and Kilbourne (2013) map the sedentariness of nearly 14,000 jobs to the census occupation codes.
  - Bessen (2020) calculates the distribution of sedentary occupations for industries using the 2000 census Public-Use Microdata
  - 264 industries that are mostly at three and four digits of NAICS codes.
  - First stage shows a strong correlation between IT intensity and sedentariness.

# EMPIRICAL RESULTS – BASELINE

| Dep.Var. | Employment (t+1) | | Sales | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| log(IT Intensity) | 0.056*** | 0.050*** | 0.101*** | 0.098*** |
| | (0.007) | (0.008) | (0.008) | (0.008) |
| log(Non-IT Intensity) | 0.023*** | 0.023*** | 0.043*** | 0.042*** |
| | (0.003) | (0.003) | (0.004) | (0.003) |
| log(IT Intensity)(t-1) | | 0.025*** | | 0.010 |
| | | (0.008) | | (0.007) |
| log(Non-IT Intensity)(t-1) | | 0.011*** | | 0.010*** |
| | | (0.003) | | (0.003) |
| Log(Emp) | 0.423*** | 0.368*** | 0.382*** | 0.381*** |
| | (0.018) | (0.023) | (0.020) | (0.020) |
| Firm FE | Y | Y | Y | Y |
| Year FE | Y | Y | Y | Y |
| Industry FE | Y | Y | Y | Y |
| R-squared | 0.966 | 0.969 | 0.965 | 0.965 |
| Observations | ~50,000 | ~35,000 | ~50,000 | ~50,000 |

# EMPIRICAL RESULTS – BASELINE

| Dep.Var. | Employment Share (t+1) | | Sales Share | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| log(IT Intensity) | 0.015*** | 0.011* | 0.023*** | 0.022*** |
| | (0.005) | (0.006) | (0.004) | (0.004) |
| log(Non-IT Intensity) | 0.006** | 0.002 | 0.010*** | 0.009*** |
| | (0.002) | (0.003) | (0.002) | (0.002) |
| log(IT Intensity)(t-1) | | 0.012* | | -0.000 |
| | | (0.006) | | (0.004) |
| log(Non-IT Intensity)(t-1) | | 0.003 | | 0.004*** |
| | | (0.002) | | (0.002) |
| Log(Emp) | 0.118*** | 0.097*** | 0.091*** | 0.090*** |
| | (0.011) | (0.014) | (0.006) | (0.006) |
| Firm FE | Y | Y | Y | Y |
| Year FE | Y | Y | Y | Y |
| Industry FE | Y | Y | Y | Y |
| R-squared | 0.940 | 0.945 | 0.972 | 0.972 |
| Observations | ~50,000 | ~35,000 | ~50,000 | ~50,000 |

# EMPIRICAL RESULTS – INSTRUMENTAL VARIABLE

| Dep.Var. | Employment (t+1) | | Sales | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| log(IT Intensity) | 0.002 | 0.007 | 0.645*** | 0.660*** |
| | (0.006) | (0.027) | (0.043) | (0.135) |
| log(Non-IT Intensity) | 0.017*** | 0.028*** | 0.325*** | 0.168*** |
| | (0.002) | (0.007) | (0.016) | (0.037) |
| Log(Emp) | 0.984*** | 0.978*** | 0.938*** | 0.927*** |
| | (0.001) | (0.002) | (0.008) | (0.009) |
| 1st Stage coefficient | 0.928*** | 0.606*** | 0.935*** | 0.593*** |
| | (0.025) | (0.051) | (0.025 ) | (0.051) |
| KP Chi-sq | 968.1 | 140.9 | 966.6 | 132 |
| CD Wald F | 5800 | 384.3 | 5690 | 355.7 |
| KP Wald F | 1409 | 142.6 | 1408 | 133.2 |
| Year FE | Y | Y | Y | Y |
| Industry FE | | Y | | Y |
| R-squared | 0.954 | 0.945 | 0.674 | 0.692 |
| Observations | ~50,000 | ~50,000 | ~50,000 | ~50,000 |

# EMPIRICAL RESULTS – LABOR SHARE

| Dep.Var. | Labor Share | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Model | Baseline | w/ Lag | Robust | Weighted | IV (Sedentariness) | |
| Log(IT Intensity) | -0.095*** | -0.094*** | -0.094*** | -0.091*** | -0.100** | 0.061 |
| | (0.009) | (0.009) | (0.009) | (0.012) | (0.039) | (0.121) |
| Log(Non-IT Int.) | -0.031*** | -0.030*** | -0.031*** | -0.032*** | -0.282*** | -0.240*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.015) | (0.034) |
| Log(Emp) | 0.418*** | 0.419*** | 0.427*** | 0.420*** | 0.023*** | 0.038*** |
| | (0.020) | (0.020) | (0.022) | (0.023) | (0.008) | (0.009) |
| Firm FE | Y | Y | Y | Y | | |
| Year FE | Y | Y | Y | Y | Y | Y |
| Industry FE | Y | Y | Y | Y | | Y |
| R-squared | 0.882 | 0.882 | 0.882 | 0.883 | 0.178 | 0.063 |
| Observations | ~50,000 | ~50,000 | ~50,000 | ~50,000 | ~50,000 | ~50,000 |

- *Labor share is defined as payroll over sales, following Autor et al. 2020.*

# EMPIRICAL RESULTS – LOCAL MARKETS

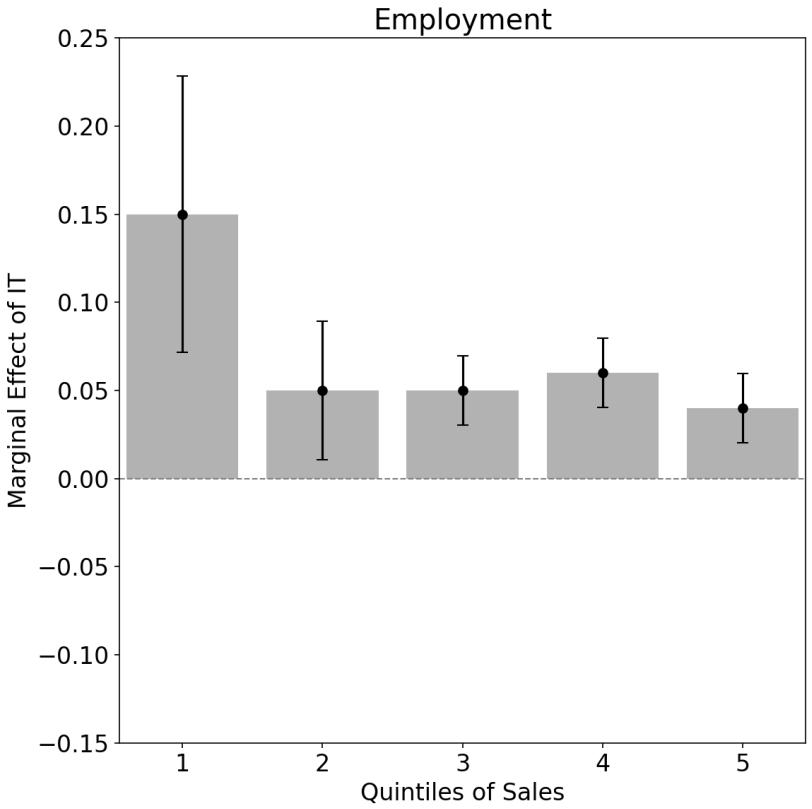
| Dep.Var. | Establishments | | Zip Codes | | Industries | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Model | Baseline | IV | Baseline | IV | Baseline | IV |
| log(IT Intensity) | 0.032*** | 0.964*** | 0.030*** | 0.792*** | 0.011** | 0.426*** |
| | (0.006) | (0.280) | (0.006) | (0.269) | (0.004) | (0.131) |
| Log(Non-IT Int.) | 0.002 | -0.192** | 0.001 | -0.153** | 0.005*** | -0.042 |
| | (0.002) | (0.076) | (0.002) | (0.073) | (0.002) | (0.036) |
| Log(Emp) | 0.352*** | 0.8010*** | 0.332*** | 0.791*** | 0.143*** | 0.351*** |
| | (0.016) | (0.014) | (0.015) | (0.014) | (0.008) | (0.008) |
| Firm FE | Y | | Y | | Y | |
| Year FE | Y | Y | Y | Y | Y | Y |
| Industry FE | Y | Y | Y | Y | Y | Y |
| R-squared | 0.988 | 0.360 | 0.988 | 0.401 | 0.957 | 0.258 |
| Observations | ~50,000 | ~50,000 | ~50,000 | ~50,000 | ~50,000 | ~50,000 |

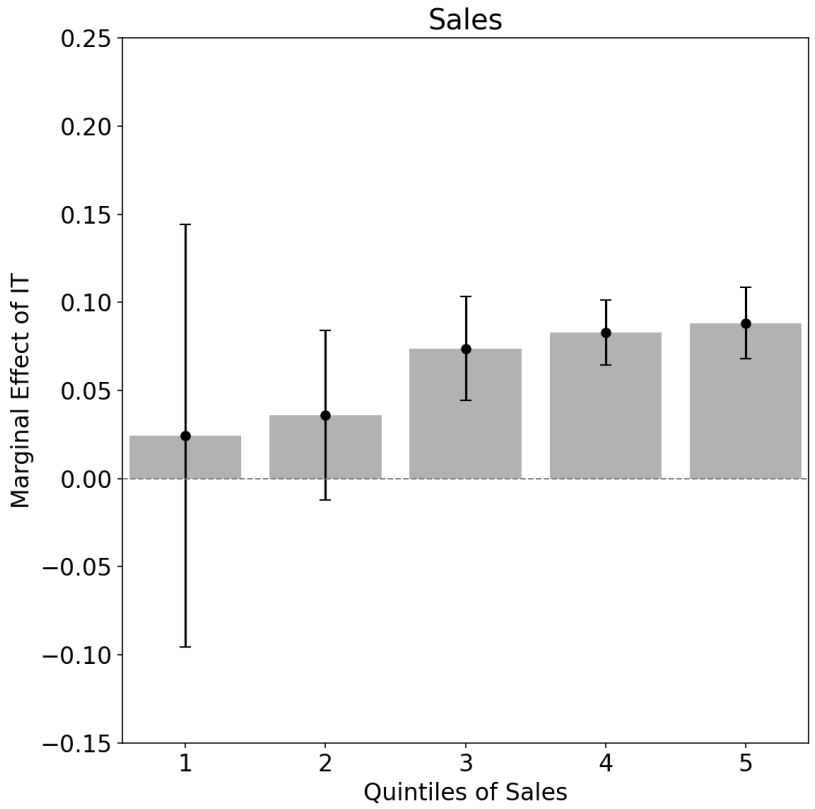
# EMPIRICAL RESULTS – INDUSTRIAL CONCENTRATION

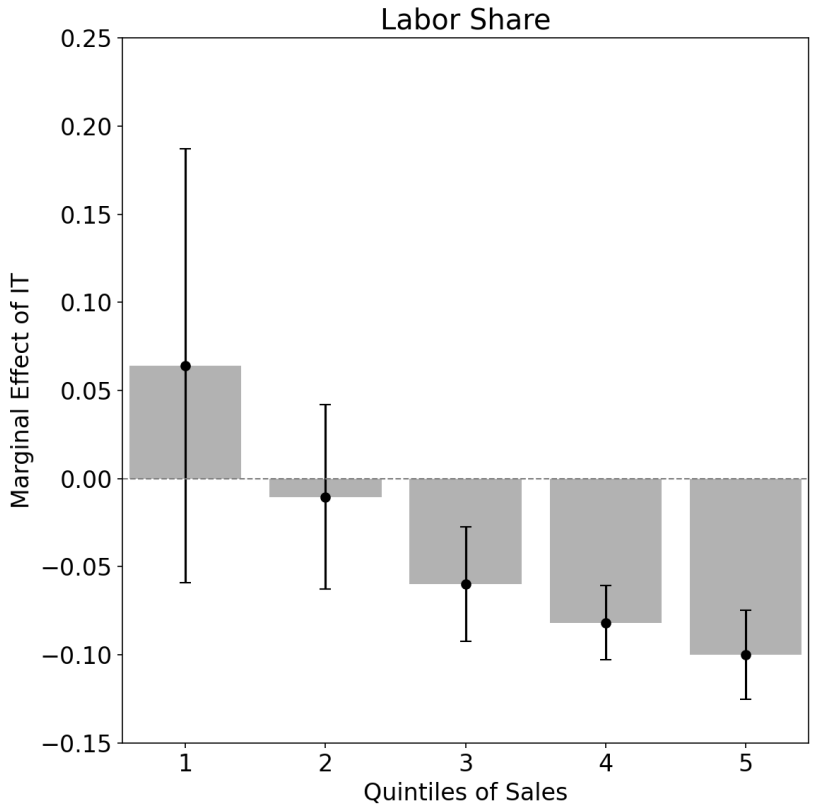
| | 1<br>Δ log(emp HHI) | 2<br>Δ log(emp CR4) | 3<br>Δ log(emp top 10%) | 4<br>Δ log(sale HHI) | 5<br>Δ log(sale CR4) | 6<br>Δ log(sale top 10%) |
|---|---|---|---|---|---|---|
| Δ log(IT) | 0.084*** | 0.012*** | 0.047*** | 0.080*** | 0.010*** | 0.037*** |
| | (0.014) | (0.002) | (0.008) | (0.022) | (0.003) | (0.011) |
| R-squared | 0.042 | 0.042 | 0.043 | 0.015 | 0.015 | 0.014 |
| Observations | 815 | 815 | 815 | 815 | 815 | 815 |

- All Δ are taken the differences between 2004 and 2013, the period of our IT measure covers.
- The results suggest that industries with a faster adoption of IT investment are associated with a faster increase of industrial concentration.
- The results are true for both employment concentration and sales concentration measures.
- The results are robust across all three concentration measures.

# EMPIRICAL RESULTS – HETEROGENEOUS RETURNS BY SIZE CATEGORY




- *The absolute effect on employment is much greater among firms in higher quintiles.*

# SUMMARY AND CONCLUSION

1. Investment in IT is associated with larger firm size and greater sales

2. IV approach and long difference models suggest a causal relation between IT intensity and firm size measures.

3. The effect of IT on sales is significantly larger than its effect on firm employment.
---- Scale without Mass

4. IT intensive firms also present a declining labor share while operate in more local markets.

5. Heterogeneous effects of IT investment towards larger firms may contribute to the increased market concentration in recent years.
---- Winner Take Most