

# Online Appendix for

## Labor Market Concentration, Earnings, and Inequality:

Kevin Rinz

### Appendix A Additional Data Details

Employers use Form W-2 to report their employees' earnings to the IRS. The form includes identifying information for both the employer and the employee, the amount of taxable wages paid to the employee, the amount of tax withheld, and some information about certain non-taxable compensation. The extract available through the Center for Economic Studies (CES) at the U.S. Census Bureau contains the Employer Identification Number (EIN, sometimes also called the Tax Identification number, or TIN), the (uncapped) amount of wages paid, and the amount of deferred compensation paid from each W-2 filed from 2005 through 2015.<sup>1</sup> The personally identifiable information (PII) contained on each form is used to assign a unique person identifier called a Protected Identification Key (PIK) through Census Bureau's Person Identification Validation System (PVS) and is then removed from the files.<sup>2</sup>

---

<sup>1</sup>The form reports the amount of wages that are subject to the Social Security and Medicare payroll taxes. The Social Security payroll tax is capped, for example it was levied on only the first \$117,000 of wage income in 2014. The Medicare payroll tax is uncapped, i.e. it is levied on all wage income. I use the uncapped measure of wages subject to the Medicare payroll tax in this analysis. Note that the extract does not include all information available on Form W-2; for example, information about employer-sponsored health insurance is not available.

<sup>2</sup>In general, PVS assigns PIKs based on PII like social security numbers, date of birth, place of birth, name, and address. Not all records can be assigned a PIK if the available PII is of low quality, contains contradictory information, or is missing important elements, but when social security numbers are available,

Below, I analyze the response of the earnings distribution among people employed in various geography by industry labor markets to changes in industrial concentration. In order to use W-2s for this purpose, I need to assign each form to a person, a place, and an industry. I aggregate earnings to the person level by summing wage and salary earnings and deferred compensation across W-2s within PIKs. The W-2 data I have access to contains the employer’s EIN, but no other information about the employer, so industrial classification is not readily available. For individuals who receive multiple W-2s, I retain the EIN associated with their highest-income W-2. I use the EIN to assign an industrial classification obtained from the LBD and data described in the next section. The W-2 data also do not contain any information about the geographic location of the recipients. I obtain person-level address information from other tax data described in the next subsection. Both industry and geographic information are assigned to W-2s through a process described in Appendix B.

As mentioned above, an important limitation of the W-2 data is that they do not contain any information on the geographic location of the forms’ recipients. They do contain the same individual identifier available on other tax forms that include geographic information. Specifically, I have access to extracts from Form 1040 and a collection of Form 1099 information returns. The 1040 data are available annually beginning in 1998 and contain the address from which they were filed. The 1099 data are available annually beginning in 2003 and contain the address to which they were sent. For my purposes, I am interested in each W-2 recipient’s county of residence (from which the commuting zone of residence is determined). I obtain this information from this tax forms using a prioritization scheme described in Appendix B.

Similarly, the W-2 data do not contain the industry of the employer. They do contain employer EINs, which could be used to link them to other sources of business data. The LBD,

---

as they are on Form W-2, PIKs can be assigned to virtually all records. On other forms, where address information is available, the process also assigns a location identifier called a Master Address File Identifier (MAFID). See Wagner and Layne (2014) for a more detailed description of the PVS process.

which contains a relatively limited set of consistently available variables, does not include business's EINs. During the period relevant to this analysis, however, EINs are available from the BR, another source of administrative data on businesses that is linkable to the LBD. With EINs obtained from the BR added to the LBD, the same industrial classifications available in the LBD are assignable to W-2s using a process described in Appendix B. I can therefore use the Fort-Klimek industrial classification system to consistently construct both measures of industrial concentration within the LBD and statistics summarizing local industry earnings distributions by linking to W-2s.

Finally, in order to conduct an analysis of earnings outcomes for various demographic groups, I obtain data on date of birth and gender from the 2016 Census Numident file, which is generated from the Social Security Administration's Numident file and contains one record for every person issued a Social Security number. I place people into three age categories: under 25, 25-54, and 55 and older. I also obtain data on race and Hispanic origin from the 2000 and 2010 Decennial census and from the ACS from 2005 through 2015. For the sake of ensuring sample sizes are large, I use the race and Hispanic origin variables to create three mutually exclusive categories: non-Hispanic White, non-Hispanic Black, and Hispanic.<sup>3</sup> I exclude other, much smaller race and ethnicity groups from my analysis. Finally, I obtain information on educational attainment from the ACS. I use education information only for individuals who are at least 25 years of age when they appear in the ACS data. Because education information is not collected on the Decennial short form and only about 15 percent of population is covered by the ACS over the available period, education is much more sparsely available. As a result, I use only two education categories: high school or less (low education) and some college or more (high education).

---

<sup>3</sup>I use the most recently reported race and Hispanic origin values for individuals who appear in multiple surveys. For example, for an individual who responded to the 2010 Decennial short form and the 2013 ACS, I use the values reported on the 2013 ACS. Individuals who report being of Hispanic origin are assigned to the Hispanic category regardless of race. Non-Hispanic individuals who report multiple races are categorized according to the first reported race.

## Appendix B W-2 Geography and Industry Assignment

I begin with the universe of IRS Form W-2 information returns for each year from 2005 through 2015. The W-2 data available at CES do not include geographic information, so I obtain address data from IRS Form 1040 and other information returns and merge it onto the W-2 using PIKs. The vast majority of these forms can be matched to a unique address on the Census Bureau's Master Address File (MAF) and assigned a MAFID on that basis. Virtually all forms include the ZIP code of the address from which they were filed/to which they were sent. The particular geography I am interested in is county of residence. I use the available address information to assign county of residence according to the following prioritization scheme:

1. Single or consensus address from Form 1040
2. Modal address from Form 1040
3. Randomly selected address from Form 1040
4. Single or consensus address from information returns
5. Modal address from information returns
6. Randomly selected address from information returns
7. Single or consensus ZIP code from Form 1040
8. Modal ZIP code from Form 1040
9. Randomly selected ZIP code from Form 1040
10. Single or consensus ZIP code from information returns
11. Modal ZIP code from information returns

## 12. Randomly selected ZIP code from information returns

I exclude W-2s that I cannot successfully match to a county, or that belong to individuals residing in outlying U.S. territories.

Individuals who hold multiple jobs in a year commonly receive multiple W-2s. However, the raw data also contain instances of individuals receiving multiple W-2s from the same employer. As workers may have multiple employment spells with a single employer or work at more than one establishment in a given firm in a single year, and employer tax filing practices surely vary, it is not obvious that each person-employer pair should have exactly one W-2. On the other hand, if firms correct initially misfiled W-2s or inadvertently file identical forms multiple times, duplicates should be excluded.

I take several steps to exclude duplicate or erroneously filed records while retaining potentially legitimate observations of multiple W-2s within person-employer pairs. First, in sets of observations that are identical in all variables, I delete all but one. I also drop all but one record from sets of duplicates that are identical on all variables except the date on which they were processed. Second, I drop all W-2s that report zero compensation paid. Third, for each person-employer pair, I retain only W-2s filed on the most recent date on which any W-2 was processed. Finally, I exclude all W-2s from person-employer pairs that have more than five records remaining after the initial restrictions have been imposed.

I then assign a six-digit NAICS code to each W-2 by linking them to records from the Longitudinal Business Database (LBD). The LBD is an establishment level panel that begins in 1976. Industry is assigned at the establishment level. Industry coding schemes have changed several times over the years covered by the LBD, but work previously undertaken at the Census Bureau has led to the creation of crosswalks that assign consistent industry codes to establishments across all years. I assign a 2012 NAICS code to each establishment, using the industrial classification from the most recent observation of each establishment in

all years.<sup>4</sup>

Employers are identified on W-2s by their EIN. Since a single firm may operate multiple establishments under a single EIN, and those establishments may operate in different industries (e.g. a firm could produce its goods at one establishment in a manufacturing industry and sell them at another in a retail industry), assigning industry codes to W-2s is not as simple as matching EINs across datasets.<sup>5</sup>

I assign industry codes to W-2s in four stages. The key merge variables are EIN and county. I use W-2 and LBD data that correspond to the same calendar year. First, I identify EIN-county pairs in which all establishments are in the same industry (I will refer to these as non-conflicted EIN-county pairs) and assign those industries to all W-2s belonging to employees of those firms who live in those counties. Next, I merge remaining unmatched W-2s with non-conflicted EIN-county pairs using EIN only, and retain the match from the county that is closest to the county of residence of each employee, assigning the industry of the establishments in that county to the matched W-2.

Third, I merge the remaining unmatched W-2s with all establishments from industry conflicted EINs located in the employee’s county of residence. I then randomly assign each matched W-2 to an establishment within its EIN (and by extension to an industry), using establishment-level employment to determine the probability of being assigned to each establishment.

Finally, I link the remaining unmatched W-2s with all establishments from industry conflicted EINs located outside the employee’s county of residence, retaining all matches from the county that is closest to the employee’s county of residence. As above, I again randomly assign each matched W-2 to an establishment within its EIN, with the probability

---

<sup>4</sup>Using consistent industry codes assigned contemporaneously with each year of data still produces mechanical changes in industrial classification within EIN in years in which new NAICS coding schemes are introduced. Using the most recently assigned industrial classification eliminates this issue.

<sup>5</sup>The LBD does not itself contain EINs. I obtain EINs from the Business Register and match them to the LBD.

of being assigned to a given establishment being equal to its share of EIN-county employment.

After capturing matches from these four stages using contemporaneous W-2 and LBD data, I then repeat each stage of the matching procedure using LBD data from the calendar year prior to the year the W-2 data refer to, and then again using LBD data from the calendar year after the W-2 year. I do this in case the construction of the LBD, which includes only one EIN per establishment per year, omits some EINs belonging to, for example, establishments that opened or closed in the year covered by the W-2s in question.

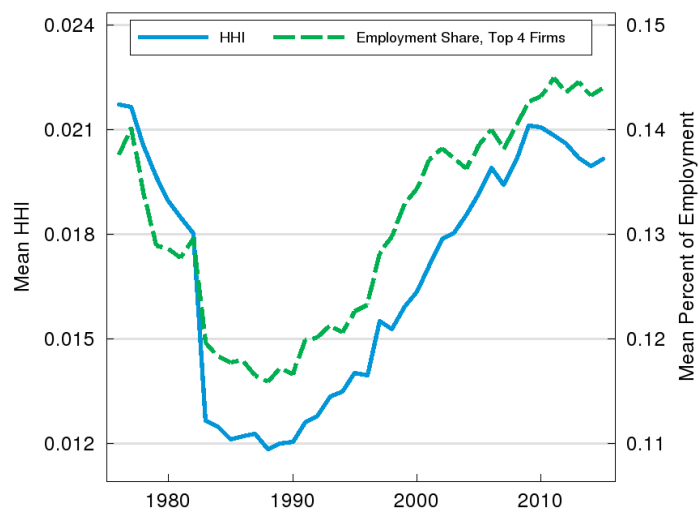




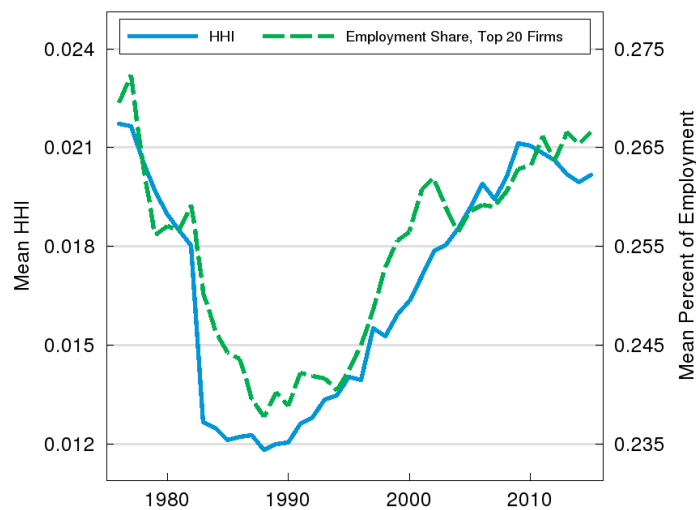
## Appendix C Additional Figures

Figure C1: Trends in National Industrial Concentration, Concentration Ratios

(a) Top Four Firms



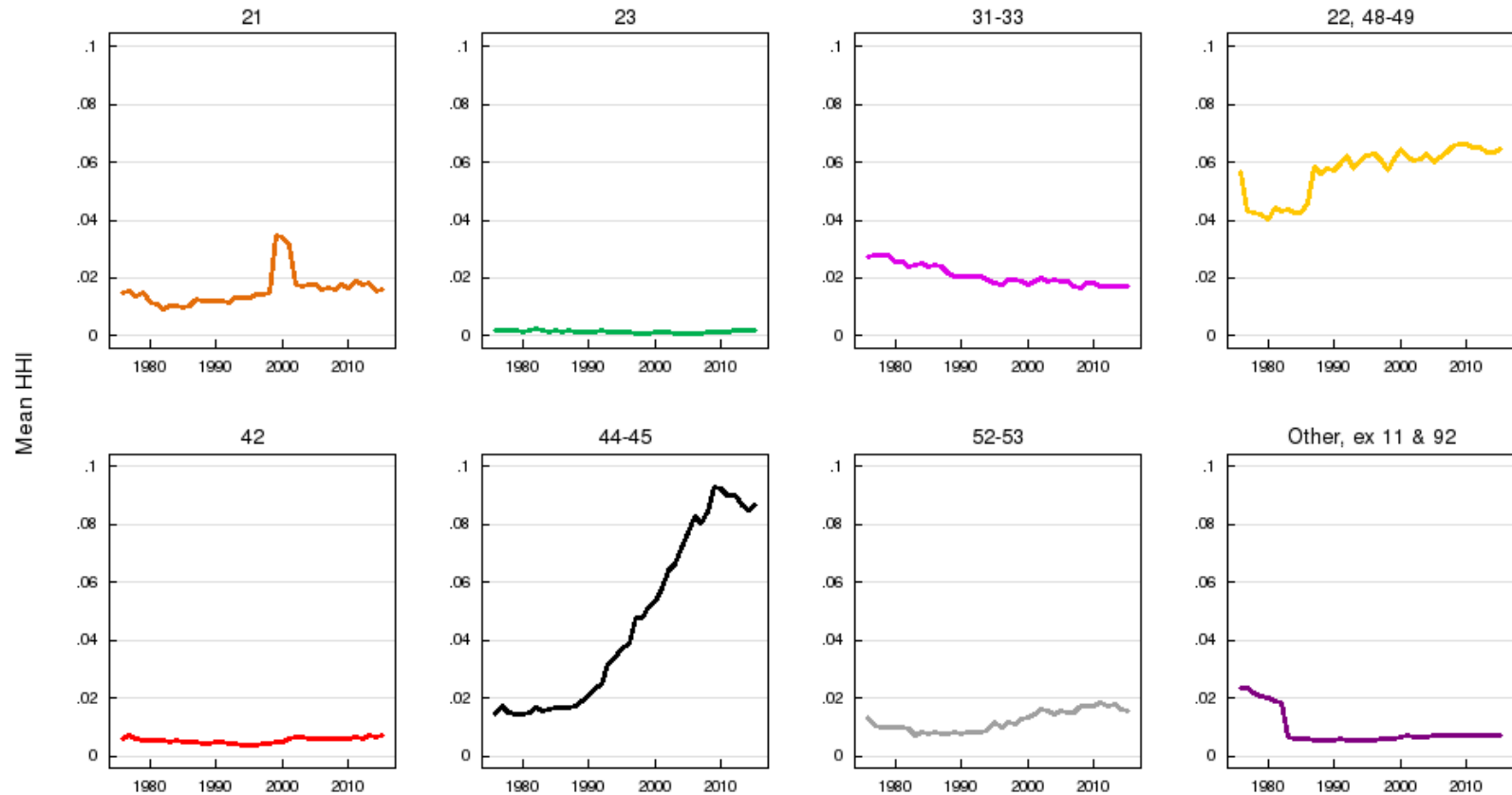
(b) Top 20 Firms



Source: Longitudinal Business Database, 1976–2015

Note: Figure plots the mean Herfindahl-Hirschman Index (left axis) alongside the concentration ratios based on the (a) top four firms and (b) top 20 firms (right axis) across national four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 to 2015.

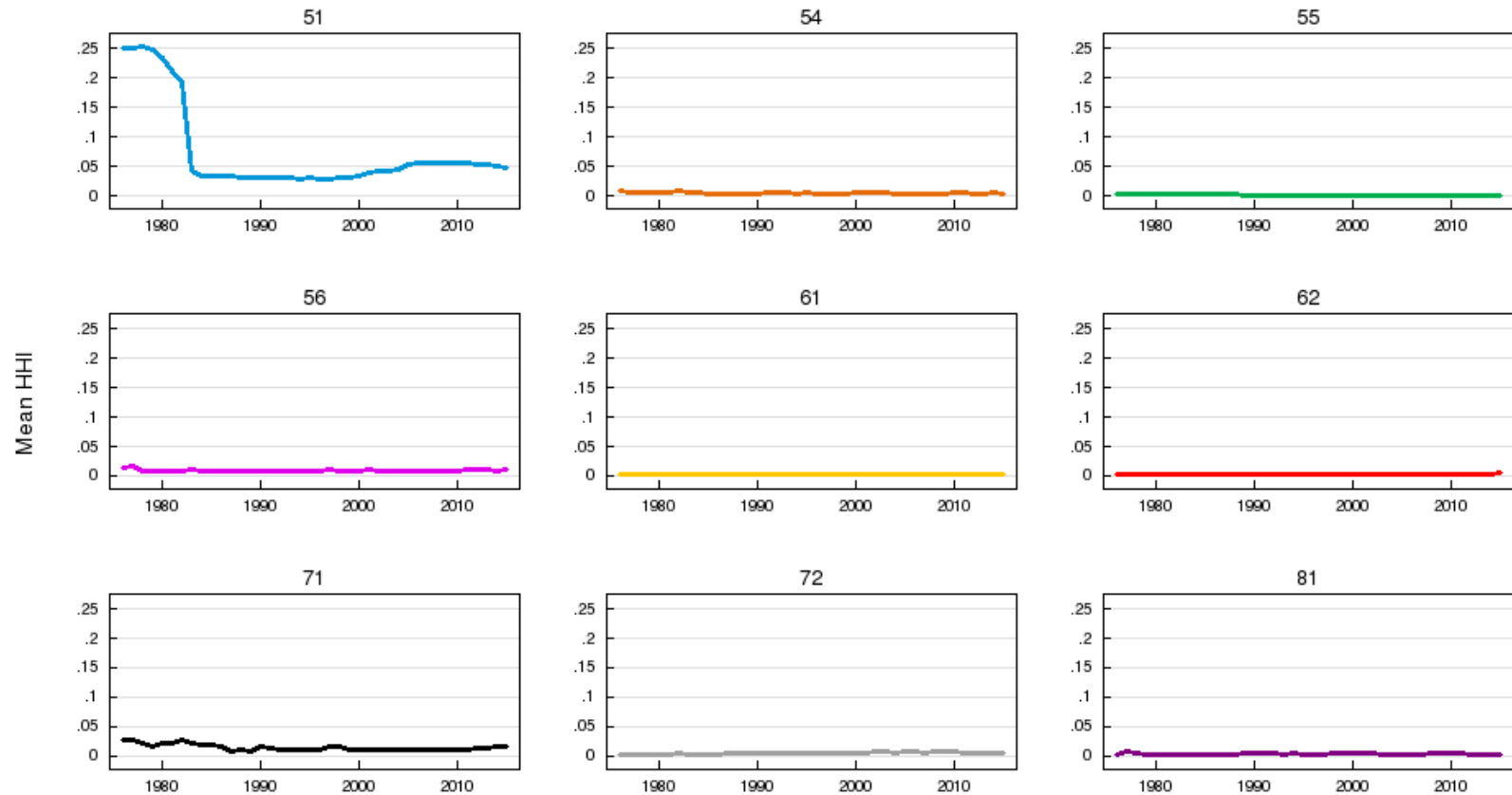
Figure C2: National Industrial Concentration Trends by Major Industry



Source: Longitudinal Business Database, 1976–2015

Note: Figure plots the mean Herfindahl-Hirschman Index across national four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015, by major industry, defined by collections of two-digit NAICS codes. Panels are labeled using the two-digits NAICS codes of the industries presented. Means are calculated using total industry employment as weights.

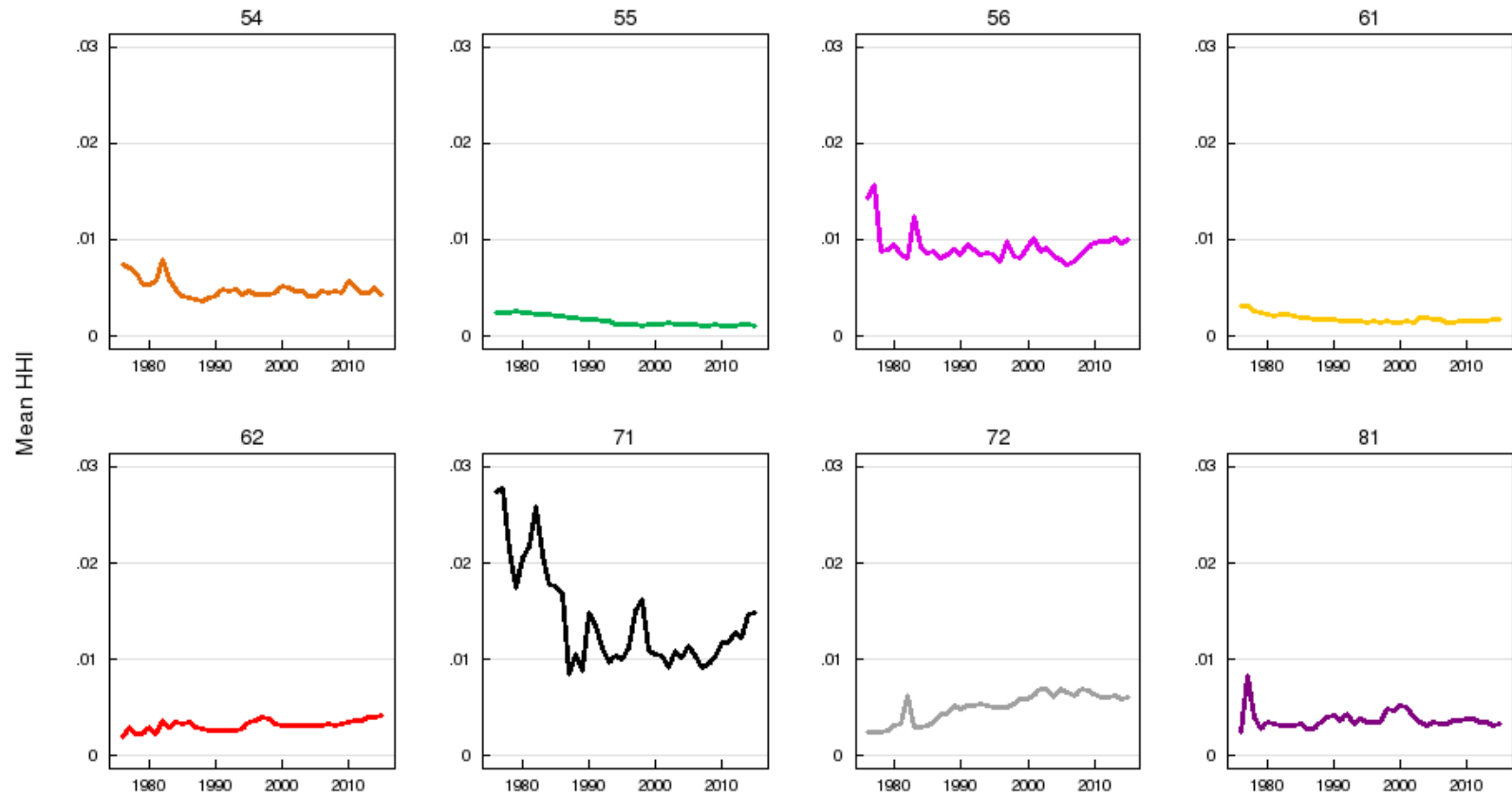
Figure C3: National Industrial Concentration Trends by Two-Digit NAICS Industry, Services



Source: Longitudinal Business Database, 1976–2015

Note: Figure plots the mean Herfindahl-Hirschman Index across national four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015, by major two-digit NAICS industry. Panels are labeled using the two-digits NAICS codes of the industries presented. Means are calculated using total industry employment as weights.

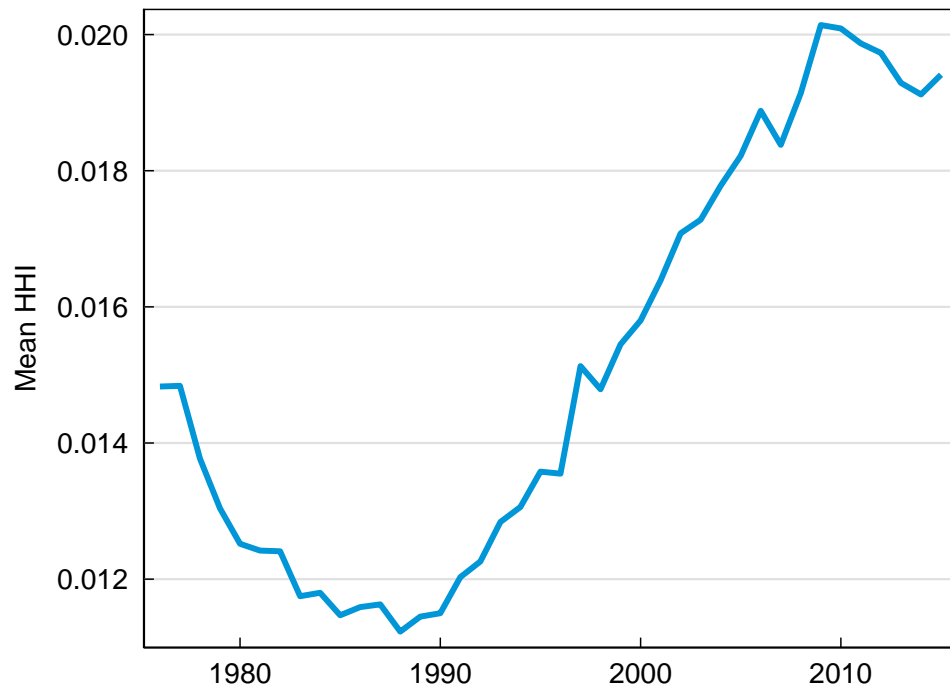
Figure C4: National Industrial Concentration Trends by Two-Digit NAICS Industry, Services, Excluding NAICS 51



Source: Longitudinal Business Database, 1976–2015

Note: Figure plots the mean Herfindahl-Hirschman Index across national four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015, by major two-digit NAICS industry. Panels are labeled using the two-digits NAICS codes of the industries presented. Means are calculated using total industry employment as weights.

Figure C5: Trend in National Industrial Concentration, All Industries Except NAICS 51

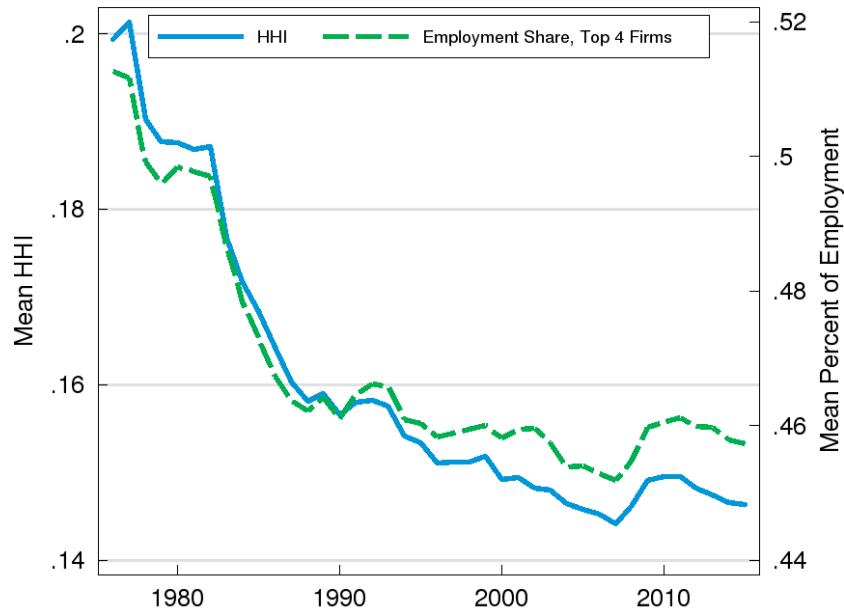


Source: Longitudinal Business Database, 1976–2015

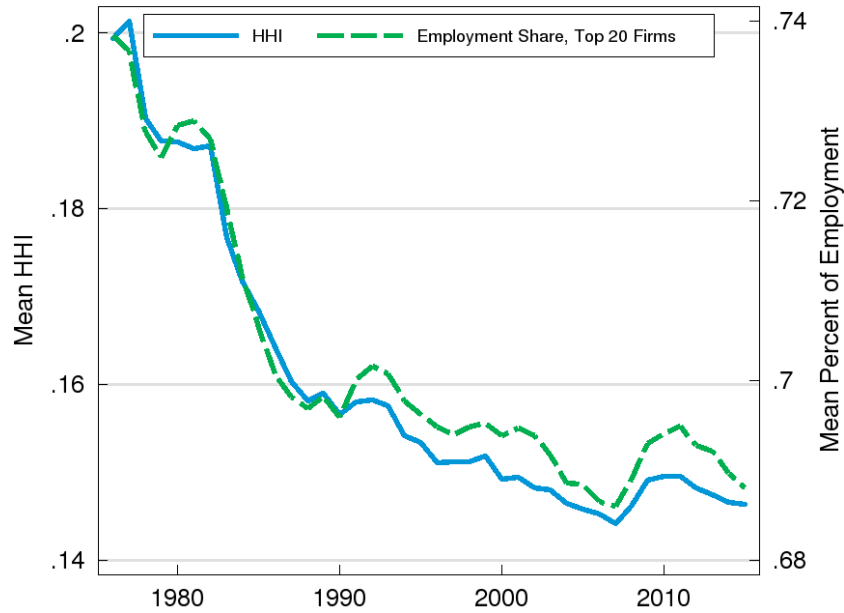
Note: Figure plots the mean Herfindahl-Hirschman Index across all national four-digit NAICS industries, standardized according to Fort and Klimek (2018), except those within the “information” sector (NAICS 51), for each year from 1976 through 2015. Means are calculated using total industry employment as weights.

Figure C6: Trends in Local Industrial Concentration, Concentration Ratios

(a) Top Four Firms



(b) Top 20 Firms

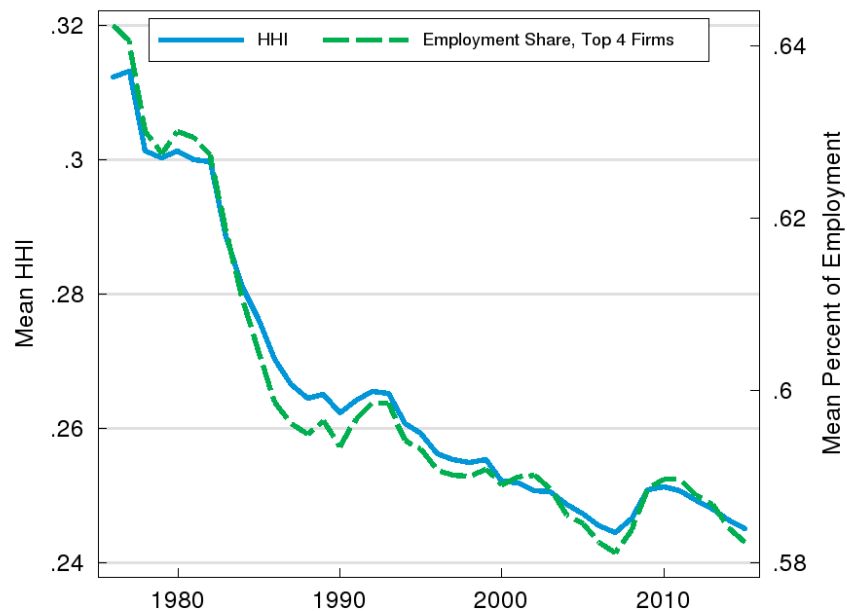


Source: Longitudinal Business Database, 1976–2015

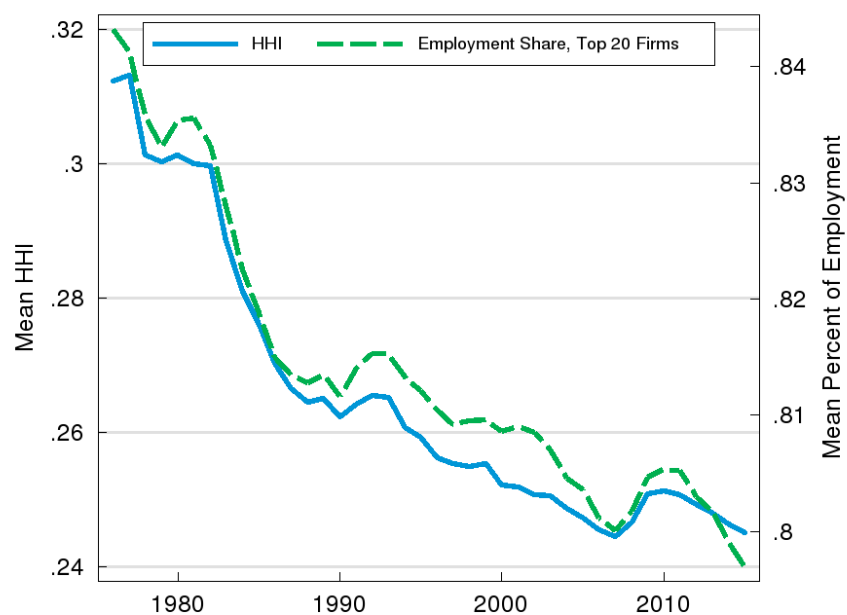
Note: Figure plots the mean Herfindahl-Hirschman Index (left axis) alongside the concentration ratios based on the (a) top four firms and (b) top 20 firms (right axis) across commuting zone-level four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 to 2015.

Figure C7: Trends in Local Industrial Concentration, County Definition, Concentration Ratios

(a) Top 4 Firms



(b) Top 20 Firms

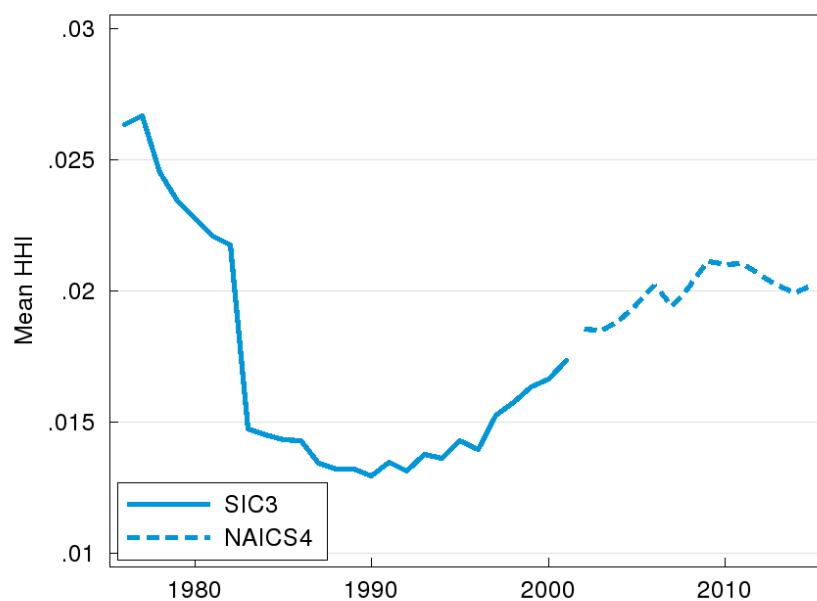


Source: Longitudinal Business Database, 1976–2015

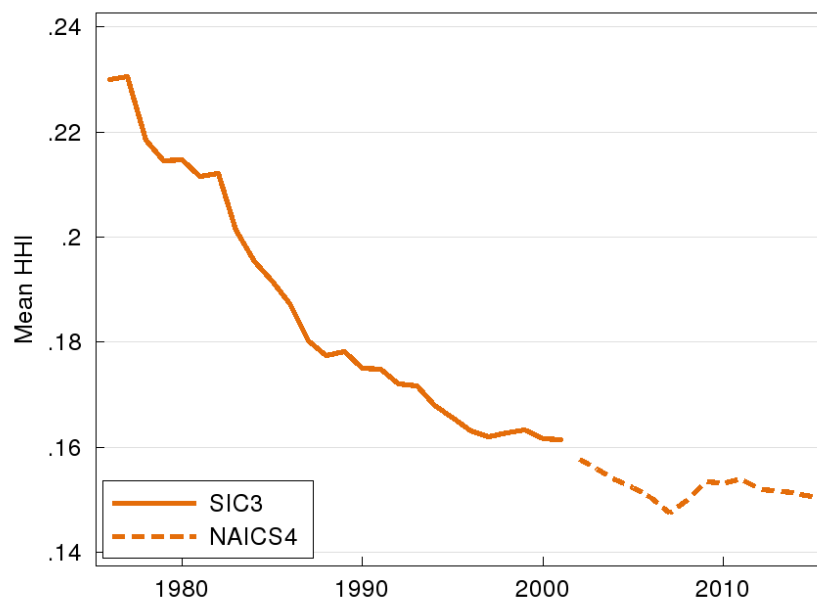
Note: Figure plots the mean Herfindahl-Hirschman Index (left axis) alongside the concentration ratios based on the (a) top four firms and (b) top 20 firms (right axis) across county-level four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 to 2015.

Figure C8: Trends in Industrial Concentration, Contemporaneous Industrial Classifications

(a) National



(b) Local

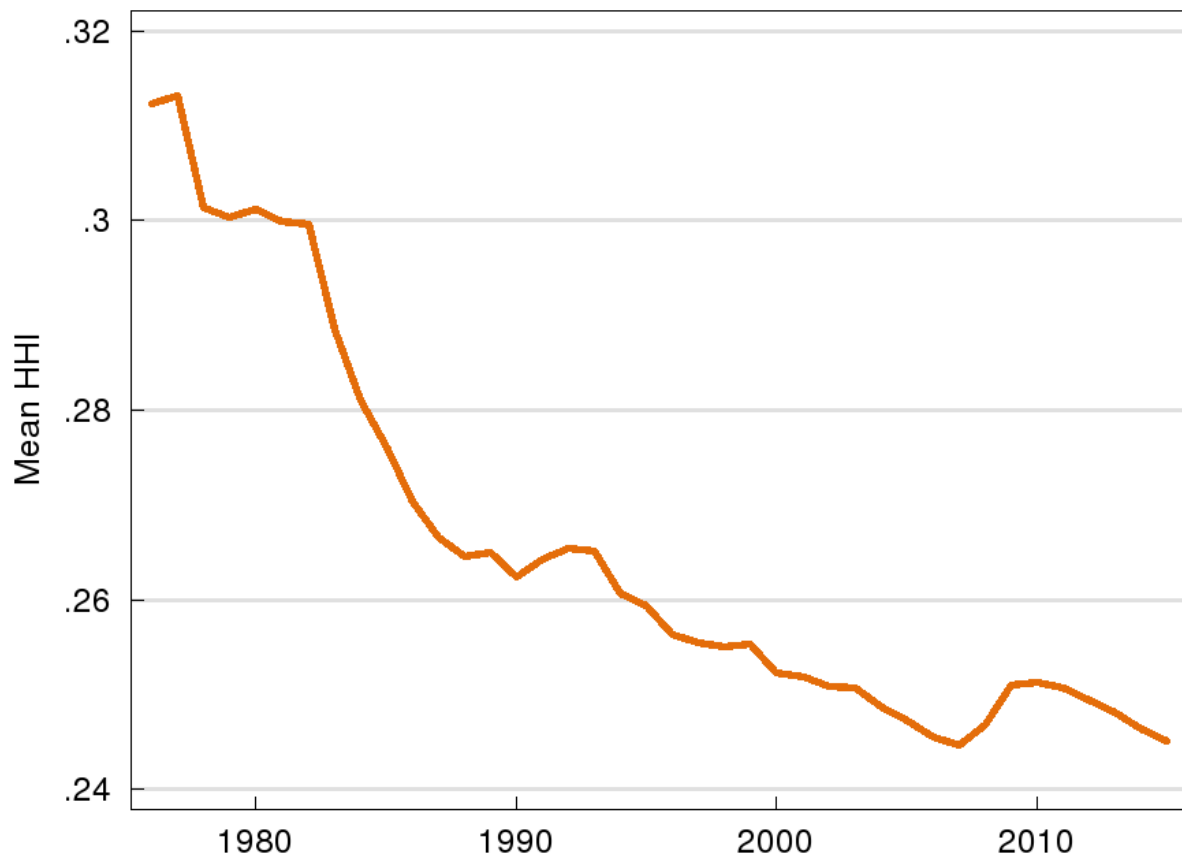


Source: Longitudinal Business Database, 1976–2015

Note: Figure plots the mean Herfindahl-Hirschman Index across (a) national and (b) commuting zone-level four-digit NAICS industries for each year from 1976 through 2015. Means are calculated using total market employment as weights. Firms are classified into industries using contemporary industrial classifications rather than the standardized classifications from Fort and Klimek (2018). From 1976–2001, firms are classified into three-digit SIC industries. From 2002–2015, firms are classified into four-digit NAICS industries.



Figure C9: Local Industrial Concentration Trend, County-based Market Definition

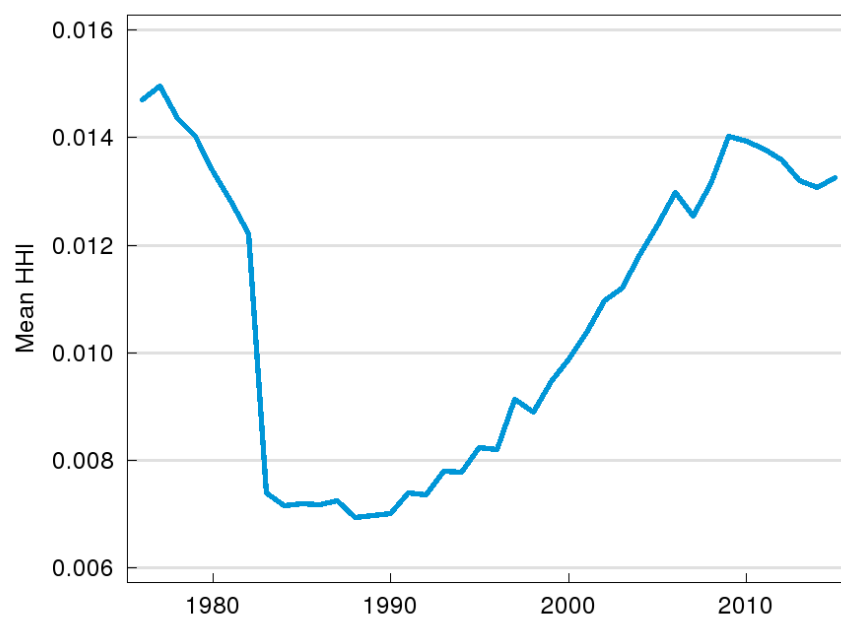


Source: Longitudinal Business Database, 1976–2015

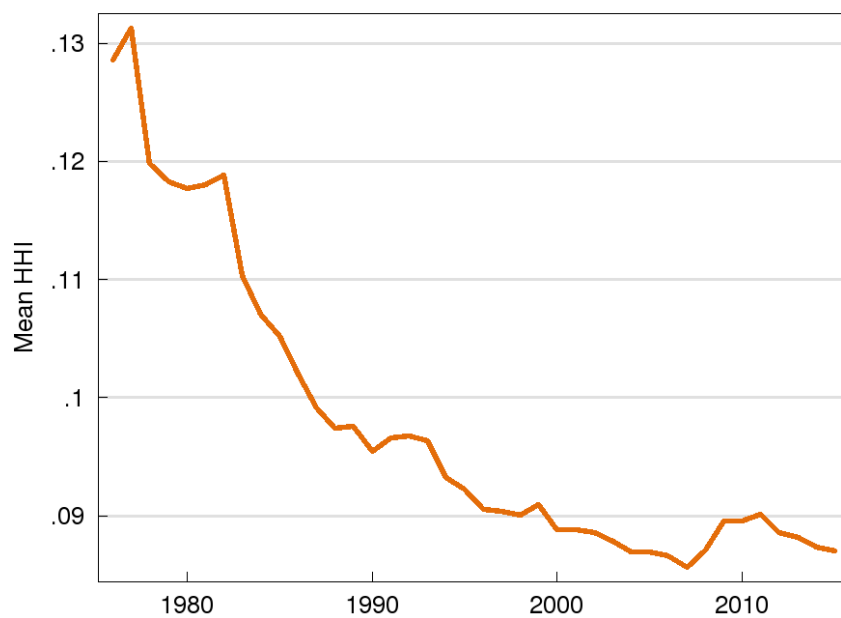
Note: Figure plots the mean Herfindahl-Hirschman Index across county-level four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015. Means are calculated using total market employment as weights.

Figure C10: Trends in Industrial Concentration, Broader Industrial Classification

(a) National



(b) Local

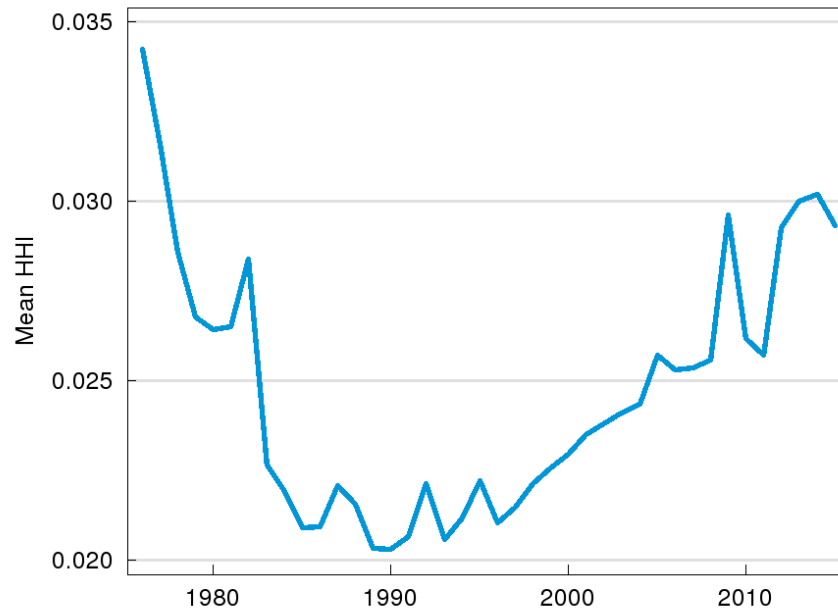


Source: Longitudinal Business Database, 1976–2015

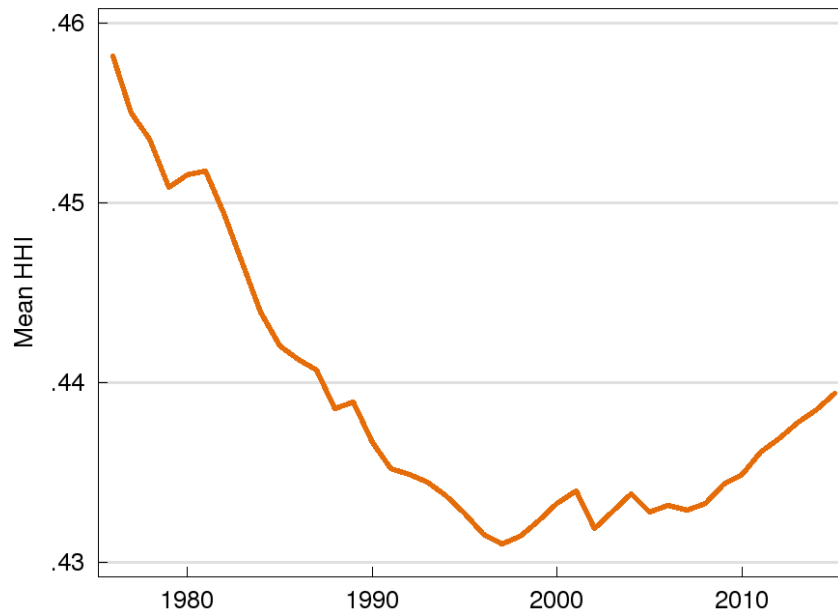
Note: Figure plots the mean Herfindahl-Hirschman Index across (a) national and (b) commuting zone-level three-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015. Means are calculated using total market employment as weights.

Figure C11: Trends in Industrial Concentration, Unweighted

(a) National



(b) Local

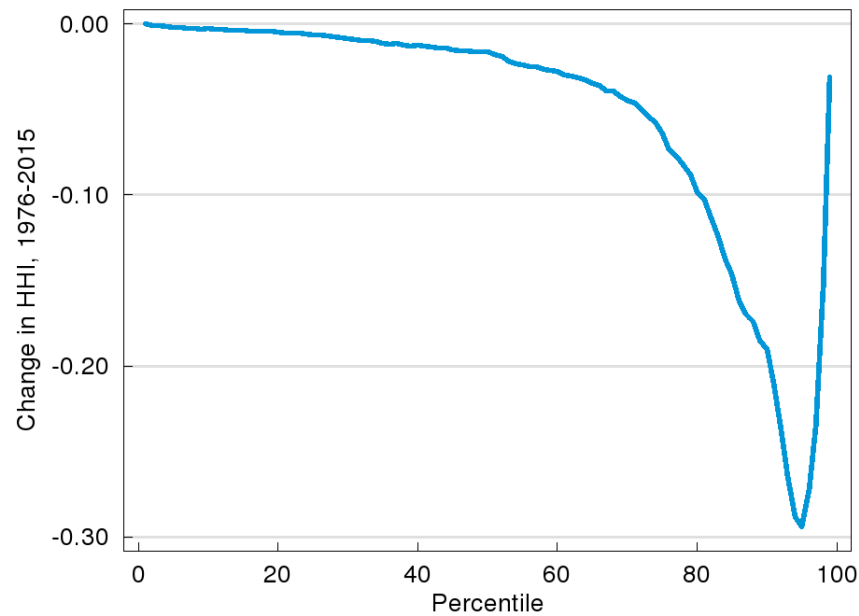


Source: Longitudinal Business Database, 1976–2015

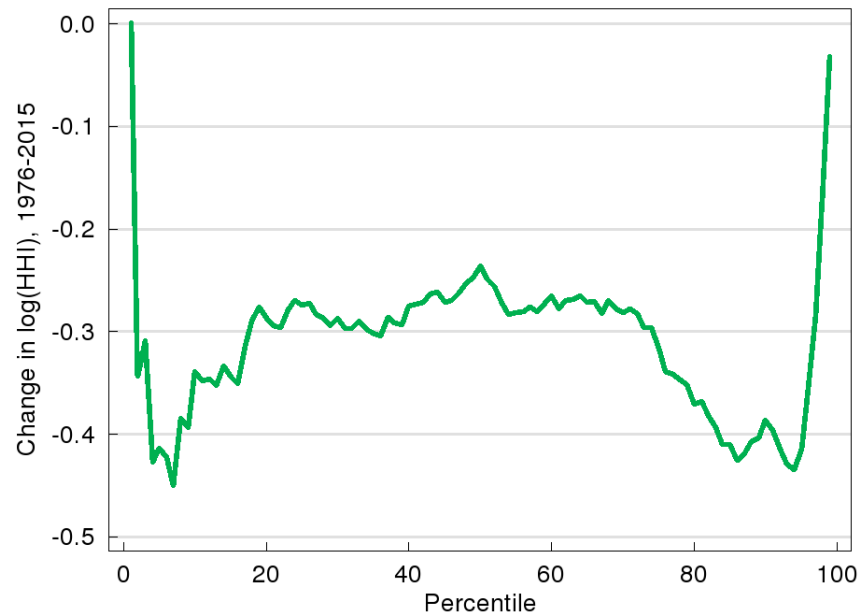
Note: Figure plots the mean Herfindahl-Hirschman Index across (a) national and (b) commuting zone-level four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015. Means are calculated with each market receiving equal weight, regardless of employment.

Figure C12: Change in Local Industrial Concentration by Percentile, 1976–2015

(a) Levels



(b) Logs

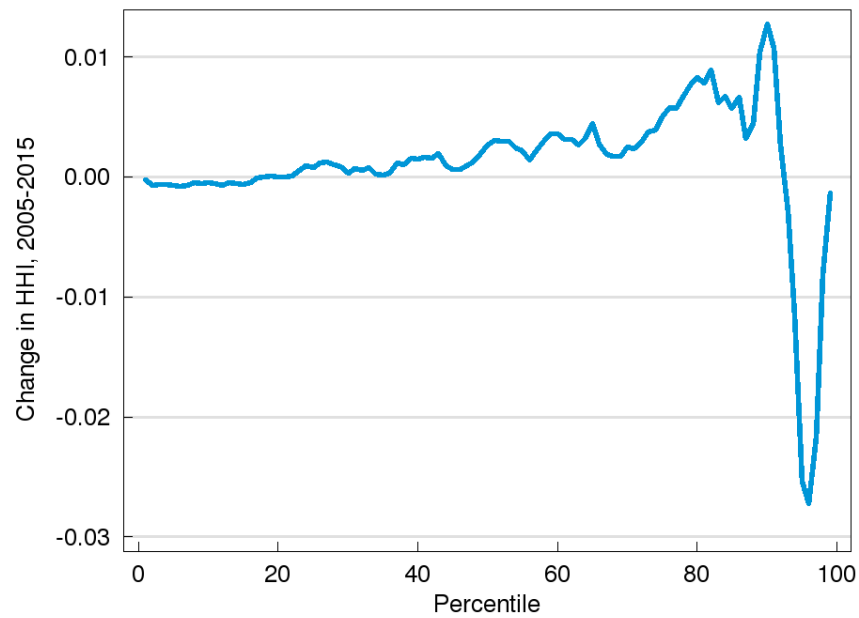


Source: Longitudinal Business Database, 1976 and 2015

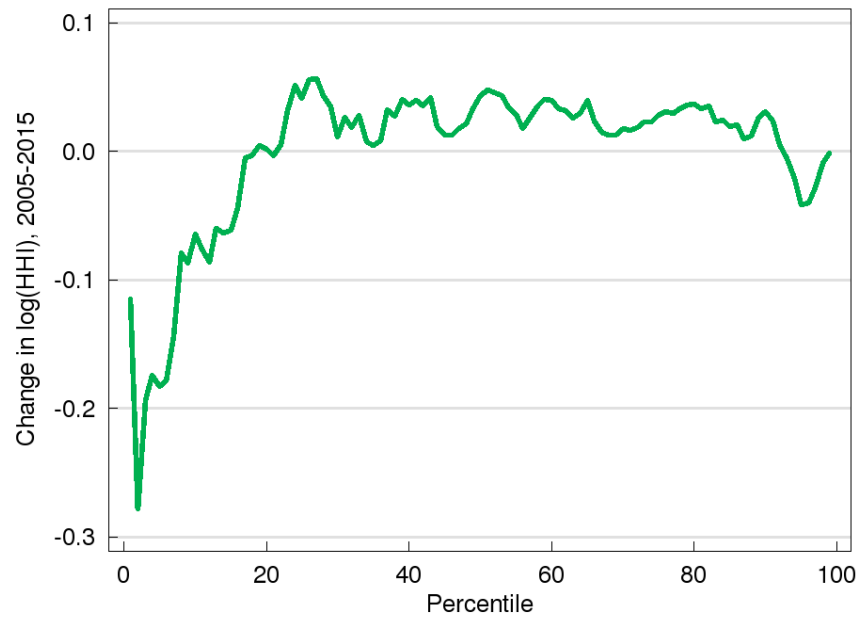
Note: Figures report changes in percentile values of the local industrial concentration distribution, as measured using the Herfindahl-Hirschman Index, between 1976 and 2015 in (a) levels and (b) logs. The unit of analysis is the commuting zone-level four-digit NAICS industry.

Figure C13: Change in Local Industrial Concentration by Percentile, 2005–2015

(a) Levels



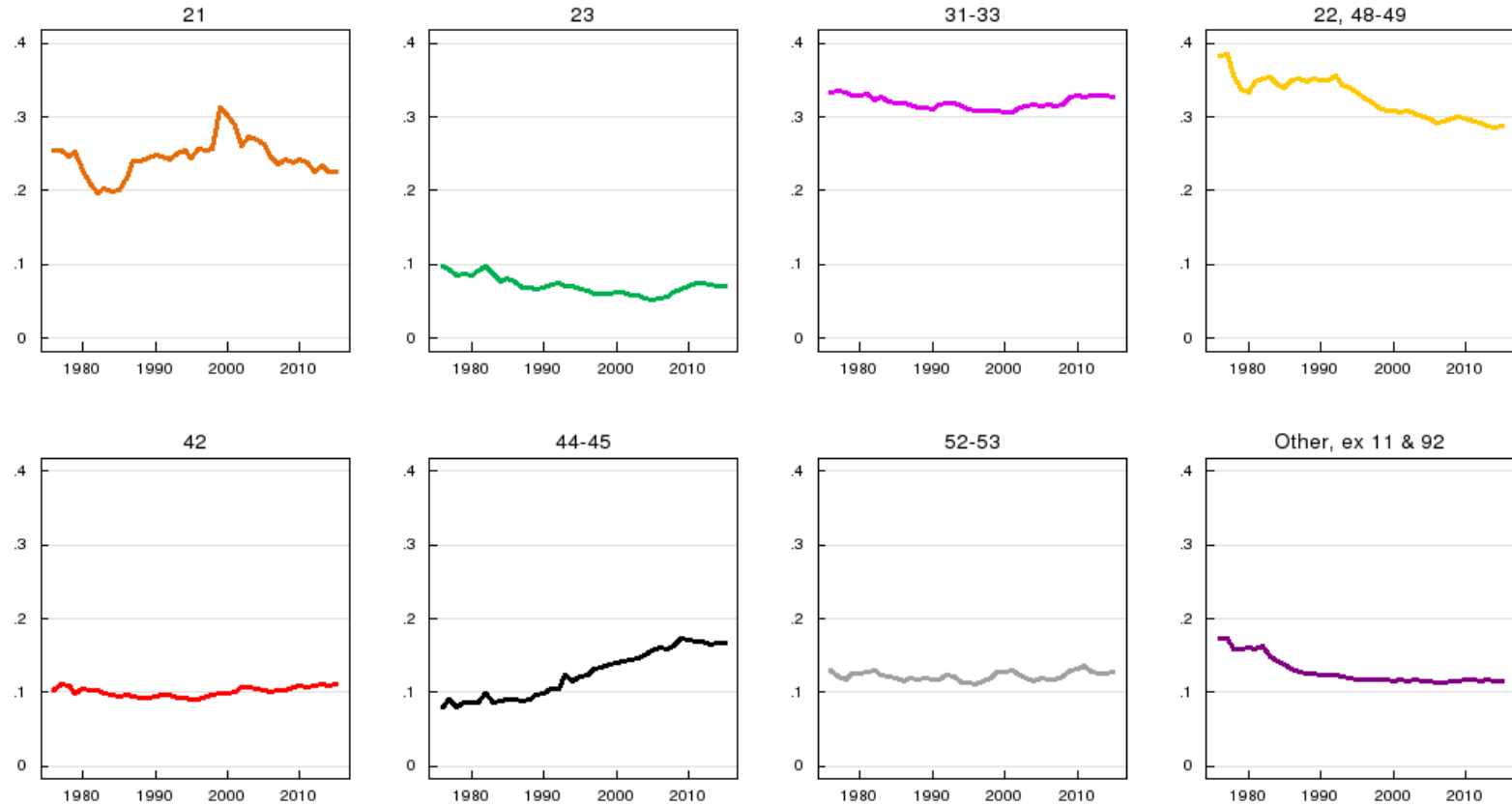
(b) Logs



Source: Longitudinal Business Database, 2005 and 2015

Note: Figures report changes in percentile values of the local industrial concentration distribution, as measured using the Herfindahl-Hirschman Index, between 2005 and 2015 in (a) levels and (b) logs. The unit of analysis is the commuting zone-level four-digit NAICS industry.

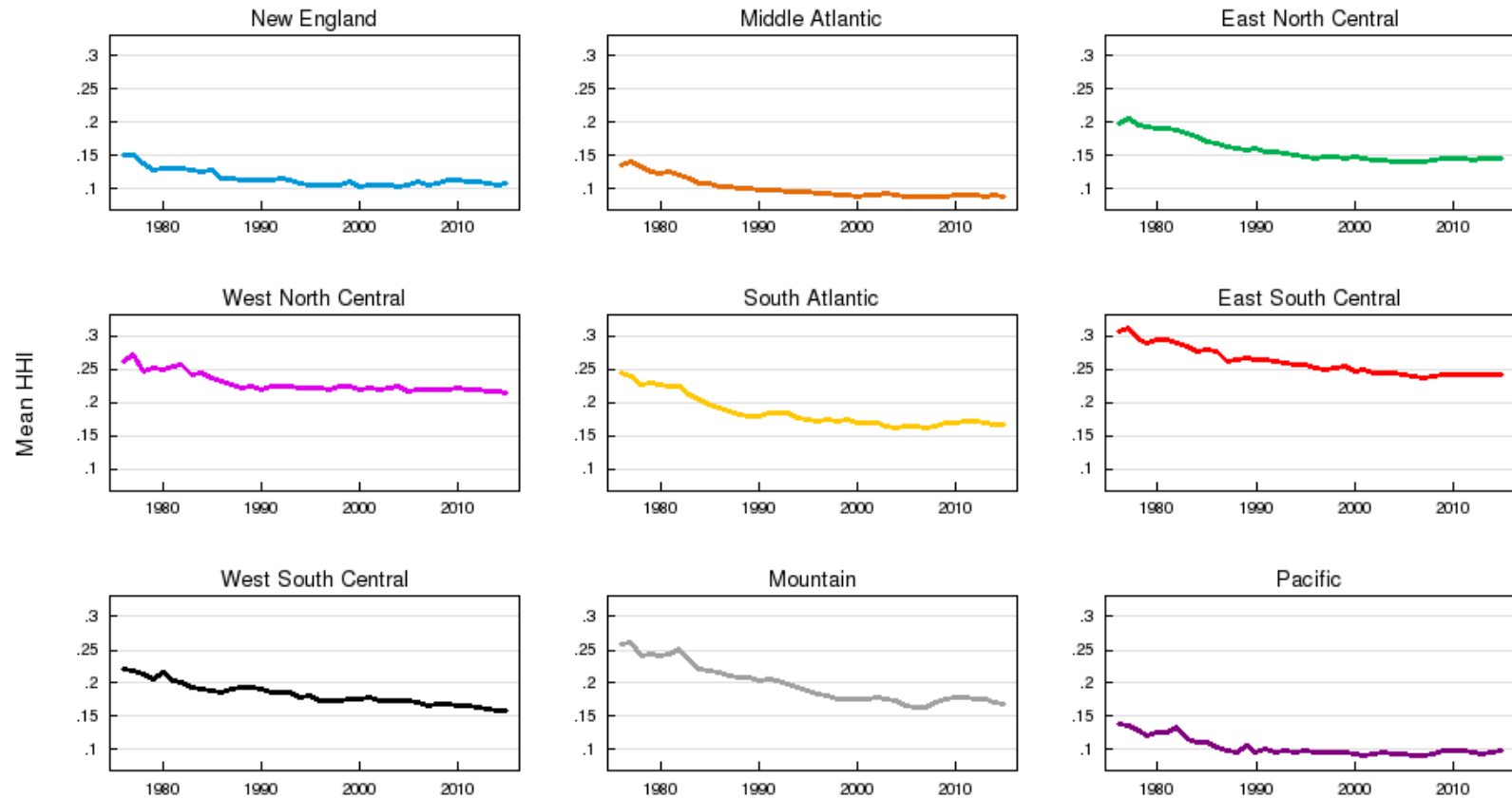
Figure C14: Local Industrial Concentration Trends by Major Industry



Source: Longitudinal Business Database, 1976–2015

Note: Figure plots the mean Herfindahl-Hirschman Index across local four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015, by major industry, defined by collections of two-digit NAICS codes. Panels are labeled using the two-digits NAICS codes of the industries presented. Means are calculated using total market employment as weights.

Figure C15: Local Industrial Concentration Trends by Census Division

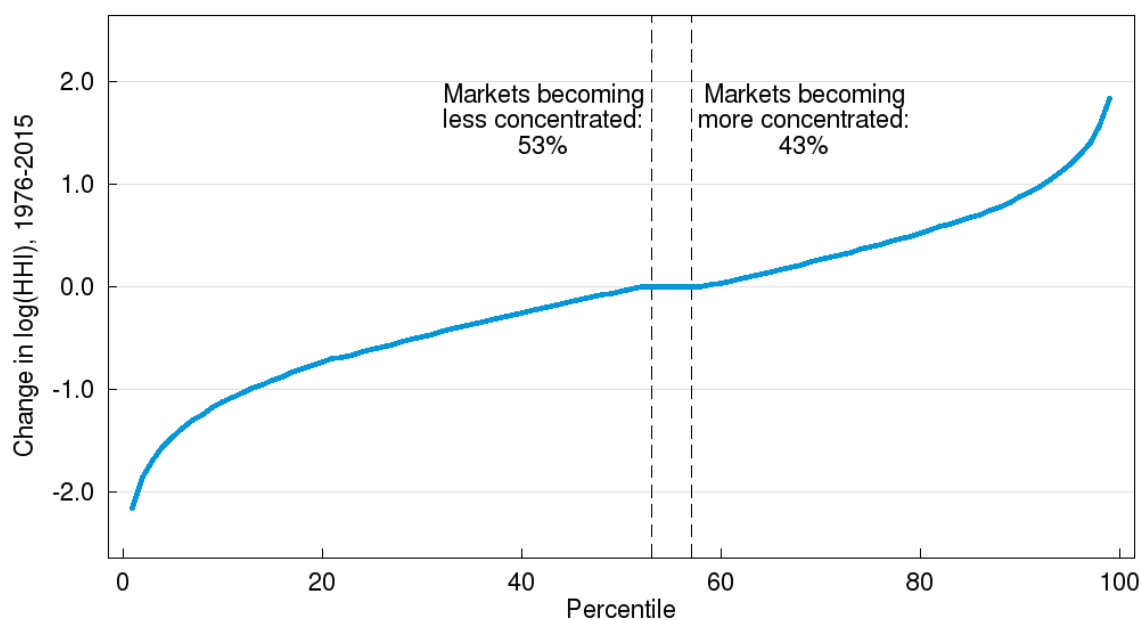


Source: Longitudinal Business Database, 1976–2015

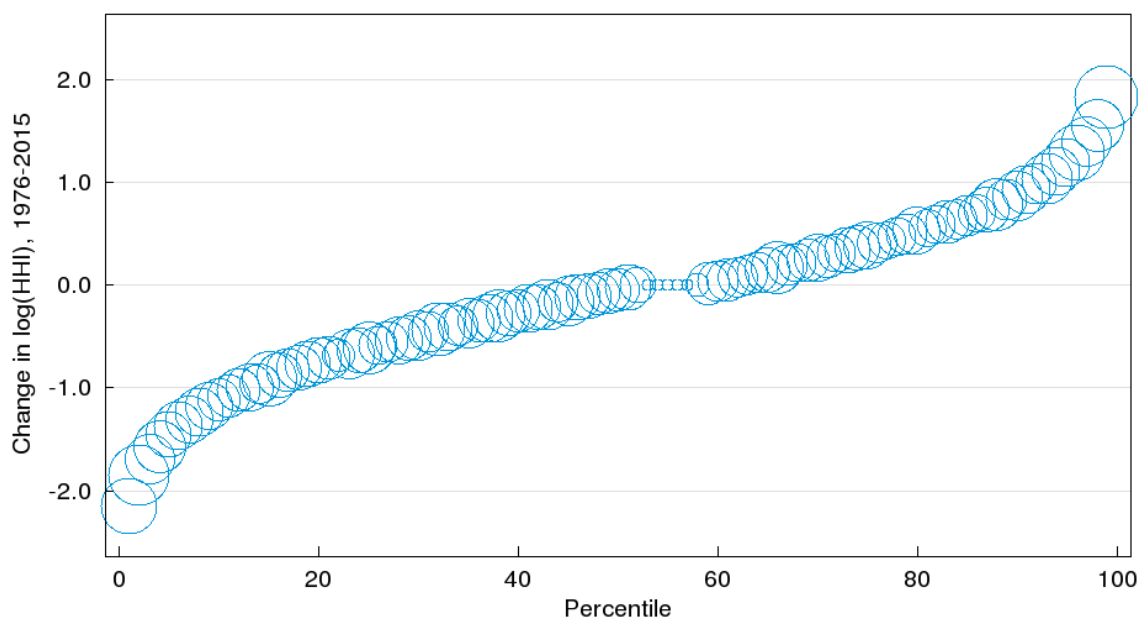
Note: Figure plots the mean Herfindahl-Hirschman Index across national four-digit NAICS industries, standardized according to Fort and Klimek (2018), for each year from 1976 through 2015, by Census division. Means are calculated using total market employment as weights.

Figure C16: Distribution of Changes in Log Local Industrial Concentration, 1976–2015

(a) Mean Changes within Percentile



(b) Markers Scaled by Employment



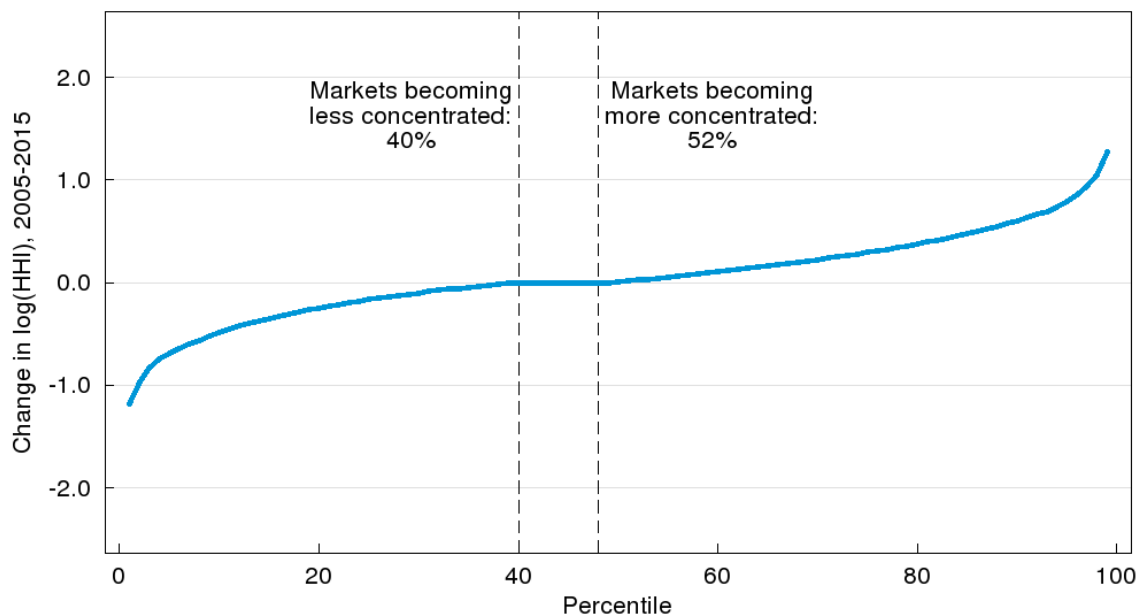
Source: Longitudinal Business Database, 1976 and 2015

Note: Figure plots the mean changes in log local industrial concentration between 1976 and 2015 within percentile bins of the log local industrial concentration distribution. The unit of analysis is the commuting zone-level four-digit NAICS industry. In panel (b), markers are proportional to total employment in markets within each percentile.

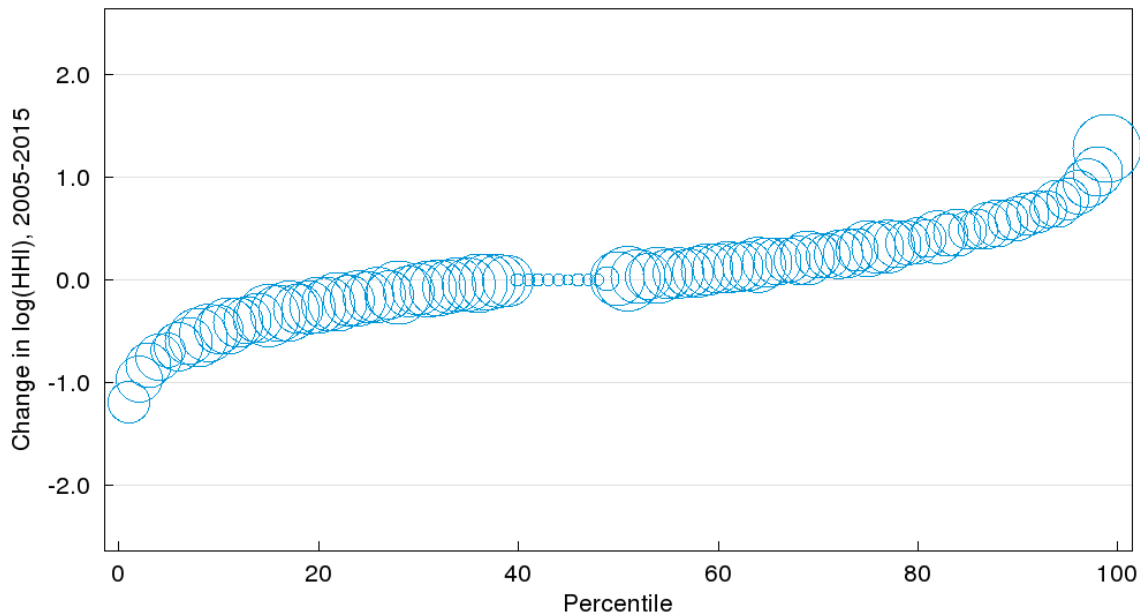


Figure C17: Distribution of Changes in Log Local Industrial Concentration, 2005–2015

(a) Mean Changes within Percentile



(b) Markers Scaled by Employment



Source: Longitudinal Business Database, 2005 and 2015

Note: Figure plots the mean changes in log local industrial concentration between 2005 and 2015 within percentile bins of the log local industrial concentration distribution. The unit of analysis is the commuting zone-level four-digit NAICS industry. In panel (b), markers are proportional to total employment in markets within each percentile.

## Appendix D Additional Tables

Table D1: HHI and Components of Average Earnings

VARIABLES	(1) $\Delta$ Total Earnings LBD	(2) $\Delta$ Total Earnings W-2	(3) $\Delta$ Total Employment LBD	(4) $\Delta$ Total Employment W-2	(5) $\Delta$ Mean Earnings LBD	(6) $\Delta$ Mean Earnings W-2
$\Delta \log(\text{HHI})$	-0.174*** (0.00707)	-0.143*** (0.00483)	-0.161*** (0.00656)	-0.147*** (0.00434)	-0.0126*** (0.00312)	0.00436** (0.00213)
Observations	131,000	130,000	131,000	130,000	131,000	130,000
R-squared	0.008	0.008	0.009	0.011	0.000	0.000

Source: Longitudinal Business Database, 1976–2015

Note: Table reports OLS estimates of the relationship between the change in local industrial concentration, as measured by the HHI, and the change in total earnings, total employment, and average earnings, as measured using the LBD and Form W-2, from 2005 to 2015. Columns represent separate univariate regressions of the indicated outcome on  $\Delta \log(\text{HHI})$ . Regressions are employment-weighted. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D2: Effects of Industrial Concentration on Earnings Outcomes, Industry-Clustered Standard Errors

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HHI	Mean Earnings	90/10	90/50	50/10	Gini
$\log(HHI^{-m})$	0.505*** (0.109)					
$\log(HHI)$		-0.0324 (0.0499)	0.173 (0.113)	0.0659 (0.0780)	0.107 (0.0949)	0.0124 (0.0171)
Observations	1,519,000	1,519,000	1,519,000	1,519,000	1,519,000	1,519,000
R-squared	0.975	-0.014	-0.026	-0.017	-0.011	-0.019
Year FEs	No	No	No	No	No	No
CZ FEs	No	No	No	No	No	No
Industry FEs	No	No	No	No	No	No
Market FEs	Yes	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Market Trends	No	No	No	No	No	No
F-stat	21.42					

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on measures of earnings and inequality, constructed using earnings data from Form W-2. The first column reports the first-stage regression. In the subsequent columns, the dependent variables are the log of mean earnings (Column 2), the logs of the ratios of the 90th and 10th (Column 3), 90th and 50th (Column 4), or 50th and 10th (Column 5) percentiles of the earnings distribution, and the Gini coefficient (Column 6). Columns represent separate regressions, which include the indicated years of data and fixed effects. Regressions are employment-weighted. Coefficients in columns 2-5 represent elasticities, while the coefficient in column 6 is a semi-elasticity. Sample sizes and statistic values have been rounded for disclosure avoidance. Standard errors are clustered at the industry level, rather than at the industry by commuting zone level, as in the main estimates. Coefficients are identical to those in the main estimates.

Table D3: Effects of Local Industrial Concentration on Earnings Outcomes, by Market Size

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HHI	Mean Earnings	90/10	90/50	50/10	Gini
Large Markets						
$\log(HHI^{-m})$	0.339*** (0.00590)					
$\log(HHI)$		-0.0987*** (0.00982)	0.265*** (0.0281)	0.128*** (0.0133)	0.137*** (0.0267)	0.0109*** (0.00256)
Observations	952,000	952,000	952,000	952,000	952,000	952,000
R-squared	0.962	-0.020	-0.015	-0.018	-0.005	-0.006
Small Markets						
$\log(HHI^{-m})$	0.151*** (0.00600)					
$\log(HHI)$		-0.141*** (0.0405)	0.320*** (0.115)	0.189*** (0.0524)	0.131 (0.112)	0.0112 (0.0104)
Observations	567,000	567,000	567,000	567,000	567,000	567,000
R-squared	0.939	-0.012	-0.008	-0.012	-0.002	-0.003
Market FEs	Yes	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports first stage and instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on various earnings outcomes, as measured by Form W-2, from 2005 to 2015. The first column reports the first-stage regression. In the subsequent columns, the dependent variables are the log of mean earnings (Column 2), the logs of the ratios of the 90th and 10th (Column 3), 90th and 50th (Column 4), or 50th and 10th (Column 5) percentiles of the earnings distribution, and the Gini coefficient (Column 6). Columns represent separate regressions, which include the indicated fixed effects. Large markets are those located in commuting zones with above-median populations, and small markets are those located in commuting zones with below-median populations. Regressions are not employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The first-stage F-statistics are 3299 in large markets and 633.6 in small markets.

Table D4: Effects of Local Industrial Concentration on Earnings Percentiles, Recentered Influence Function Regressions

VARIABLES	(1) 10th percentile	(2) 25th percentile	(3) 50th 50th percentile	(4) 75th 75th percentile	(5) 90th 90th percentile	(6) Mean
$\log(HHI)$	-0.4122*** (0.08812)	-0.1628*** (0.04639)	0.01380 (0.01675)	-0.01359 (0.01044)	-0.03255*** (0.01064)	-0.07816* (0.02499)
$\log(HHI)^2$	-0.04228*** (0.01283)	-0.01992*** (0.006691)	-0.001270 (0.002477)	-0.005253*** (0.001587)	-0.006559*** (0.001619)	-0.009468 (0.00372)
Observations	23,360,000	23,360,000	23,360,000	23,360,000	23,360,000	23,360,000
R-squared	-0.0003	-0.0002	0.0000	-0.0000	-0.0001	-0.0003

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on key percentiles of the local earnings distribution, as measured by Form W-2, from 2005 to 2015. Estimates are produced using recentered influence function regressions and a two percent sample of individuals from the data underlying the market-level analyses presented elsewhere in this paper. Influence functions are estimated within markets. Markets with less than 1,000 person-year observations in the sample are excluded from this analysis. Columns represent separate regressions, which include the indicated fixed effects. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D5: Effects of Industrial Concentration on Earnings, 1976–2015, LBD Earnings Measure

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	0.105*** (0.00764)	-0.0423* (0.0235)	-0.0411* (0.0211)	-0.0512** (0.0200)	-0.282*** (0.0282)
Observations	5,450,000	5,450,000	5,446,000	5,446,000	5,446,000
R-squared		0.478	0.524	0.657	0.721
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database, 1976–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on log mean earnings, as measured by payroll divided by employment in the LBD, from 1976 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D6: Effects of Industrial Concentration on Earnings, 2005–2015, LBD Earnings Measure

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	0.184*** (0.00805)	-0.0981 (0.0739)	-0.0120 (0.0127)	-0.00857 (0.0122)	-0.161*** (0.0351)
Observations	1,531,000	1,531,000	1,527,000	1,527,000	1,527,000
R-squared		0.858	0.971	0.972	0.980
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on log mean earnings, as measured by payroll divided by employment in the LBD, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D7: Effects of Industrial Concentration on Earnings, 2005–2015, W-2 Earnings Measure

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	0.194*** (0.00787)	-0.0242 (0.105)	-0.0372*** (0.0122)	-0.0324*** (0.0117)	-0.134*** (0.0282)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared		0.886	0.982	0.983	0.988
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on log mean earnings, as measured by Form W-2, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.



Table D8: Effects of Industrial Concentration on Earnings, 2005–2015, W-2 Earnings Measure, Unweighted

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	0.204*** (0.00255)	-0.256*** (0.0321)	-0.0985*** (0.0122)	-0.109*** (0.0121)	-0.199*** (0.0449)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared		0.585	0.869	0.871	0.911
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on log mean earnings, as measured by Form W-2, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are not employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D9: Effects of Industrial Concentration on the 90/10 Earnings Ratio

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	-0.136*** (0.00787)	-0.676* (0.373)	0.172*** (0.0270)	0.173*** (0.0265)	1.018*** (0.156)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.028	0.420	0.893	0.895	0.890
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 10th percentile of the earnings distribution, as measured by Form W-2, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D10: Effects of Industrial Concentration on the 50/10 Earnings Ratio

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	-0.0720*** (0.00593)	-0.408 (0.283)	0.107*** (0.0207)	0.107*** (0.0210)	0.784*** (0.124)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.012	0.417	0.839	0.841	0.852
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on the log of the ratio of the 50th percentile of the earnings distribution to the 10th percentile of the earnings distribution, as measured by Form W-2, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D11: Effects of Industrial Concentration on the 90/50 Earnings Ratio

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	-0.0641*** (0.00319)	-0.268** (0.129)	0.0655*** (0.0125)	0.0659*** (0.0123)	0.234*** (0.0410)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.038	0.383	0.877	0.880	0.900
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 50th percentile of the earnings distribution, as measured by Form W-2, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D12: Effects of Industrial Concentration on the Gini Coefficient

VARIABLES	(1)	(2)	(3)	(4)	(5)
log(HHI)	-0.0241*** (0.000869)	-0.0822** (0.0406)	0.0123*** (0.00275)	0.0124*** (0.00273)	0.0689*** (0.0105)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.092	0.465	0.938	0.940	0.944
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables regression estimates of the effect of local industrial concentration, as measured by the HHI, on the log of the Gini coefficient, as measured by Form W-2, from 2005 to 2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent semi-elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D13: Effects of Industrial Concentration on Earnings, 1976–2015, LBD Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	0.112*** (0.00785)	-0.0317* (0.0177)	-0.0341* (0.0176)	-0.0424** (0.0166)	-0.131*** (0.0122)
Observations	5,450,000	5,450,000	5,446,000	5,446,000	5,446,000
R-squared	0.015	0.476	0.522	0.655	0.724
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database, 1976–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of earnings within markets, as measured by payroll over employment in the LBD, from 1976-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D14: Effects of Industrial Concentration on Earnings, 2005–2015, LBD Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	0.195*** (0.00806)	0.0322 (0.0208)	-0.00603 (0.00636)	-0.00432 (0.00614)	-0.0310*** (0.00693)
Observations	1,531,000	1,531,000	1,527,000	1,527,000	1,527,000
R-squared	0.116	0.872	0.971	0.971	0.980
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of earnings within markets, as measured by payroll over employment in the LBD, from 2005–2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D15: Effects of Industrial Concentration on Earnings, 2005–2015, W-2 Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	0.204*** (0.00764)	0.00317 (0.0133)	-0.0188*** (0.00583)	-0.0163*** (0.00558)	-0.0251*** (0.00449)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.124	0.888	0.982	0.983	0.990
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of earnings within markets, as measured by Form W-2, from 2005-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.



Table D16: Effects of Industrial Concentration on Earnings, 2005–2015, W-2 Earnings Measure, Reduced Form, Unweighted

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	0.151*** (0.00167)	-0.0317*** (0.00367)	-0.0257*** (0.00317)	-0.0285*** (0.00312)	-0.0168*** (0.00371)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.045	0.648	0.870	0.872	0.914
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of earnings within markets, as measured by Form W-2, from 2005-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are not employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D17: Effects of Industrial Concentration on the 90/10 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	-0.143*** (0.00803)	0.0885*** (0.0251)	0.0872*** (0.0129)	0.0872*** (0.0126)	0.191*** (0.0179)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.028	0.652	0.896	0.898	0.932
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the 90th percentile of the earnings distribution to the 10th percentile within markets, as measured by Form W-2, from 2005-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D18: Effects of Industrial Concentration on the 50/10 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	-0.0758*** (0.00609)	0.0534** (0.0240)	0.0541*** (0.00985)	0.0539*** (0.00996)	0.147*** (0.0153)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.013	0.562	0.841	0.843	0.893
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the 50th percentile of the earnings distribution to the 10th percentile within markets, as measured by Form W-2, from 2005-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D19: Effects of Industrial Concentration on the 90/50 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	-0.0675*** (0.00332)	0.0350*** (0.0105)	0.0331*** (0.00630)	0.0333*** (0.00626)	0.0439*** (0.00537)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.036	0.579	0.879	0.882	0.913
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 50th percentile within markets, as measured by Form W-2, from 2005-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D20: Effects of Industrial Concentration on Key Percentiles of the Earnings Distribution, 2005–2015, W-2 Earnings Measure, Instrumental Variables

VARIABLES	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th
log(HHI)	-0.180*** (0.0275)	-0.128*** (0.0220)	-0.0736*** (0.0132)	-0.0171 (0.0111)	-0.00767 (0.0117)
Observations	1,519,000	1,519,000	1,519,000	1,519,000	1,519,000
R-squared	0.936	0.943	0.959	0.975	0.981
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports instrumental variables estimates of the effect of local industrial concentration, as measured by HHI, on the log of key percentiles of the earnings distribution within markets, as measured by Form W-2, from 2005–2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D21: Effects of Industrial Concentration on Key Percentiles of the Earnings Distribution, 2005–2015, W-2 Earnings Measure, Reduced Form

VARIABLES	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th
$\log(HHI^{-m})$	-0.0911*** (0.0125)	-0.0647*** (0.0105)	-0.0372*** (0.00624)	-0.00864 (0.00550)	-0.00388 (0.00584)
Observations	1,519,000	1,519,000	1,519,000	1,519,000	1,519,000
R-squared	0.938	0.944	0.960	0.975	0.981
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of key percentiles of the earnings distribution within markets, as measured by Form W-2, from 2005–2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D22: Effects of Industrial Concentration on the Gini Coefficient, 2005–2015, W-2 Earnings Measure, Reduced Form

VARIABLES	(1)	(2)	(3)	(4)	(5)
$\log(HHI^{-m})$	-0.0254*** (0.000855)	0.0108*** (0.00225)	0.00623*** (0.00141)	0.00627*** (0.00141)	0.0129*** (0.00128)
Observations	1,522,000	1,522,000	1,519,000	1,519,000	1,519,000
R-squared	0.077	0.749	0.940	0.941	0.961
Year FEs	No	Yes	Yes	No	No
CZ FEs	No	Yes	No	No	No
Industry FEs	No	Yes	No	No	No
Market FEs	No	No	Yes	Yes	Yes
CZ by Year FEs	No	No	No	Yes	Yes
Market Trends	No	No	No	No	Yes

Source: Longitudinal Business Database and Form W-2, 2005–2015

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the Gini coefficient within markets, as measured by Form W-2, from 2005-2015. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent semi-elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance.

Table D23: Effects of Industrial Concentration on Earnings, 2005–2015, W-2 Earnings Measure, Instrumental Variables, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
log(HHI)	-0.0366** (0.0162)	0.0347*** (0.00816)	-0.157*** (0.0109)	-0.0476*** (0.0132)	-0.0119 (0.0154)
Observations	1,498,000	1,478,000	1,386,000	1,503,000	1,461,000
R-squared	0.978	0.983	0.950	0.980	0.951
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
log(HHI)	-0.0510*** (0.00909)	0.00227 (0.0128)	-0.0203 (0.0231)	-0.0847*** (0.0136)	-0.0648*** (0.0132)
Observations	1,513,000	972,000	1,135,000	1,373,000	1,417,000
R-squared	0.982	0.966	0.967	0.946	0.961
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports instrumental variables estimates of the effect of local industrial concentration, as measured by HHI, on the log of mean earnings within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.



Table D24: Effects of Industrial Concentration on Earnings, 2005–2015, W-2 Earnings Measure, Reduced Form, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
$\log(HHI^{-m})$	-0.0158** (0.00652)	0.0205*** (0.00493)	-0.0934*** (0.00674)	-0.0229*** (0.00582)	-0.00558 (0.00732)
Observations	1,524,000	1,500,000	1,403,000	1,529,000	1,481,000
R-squared	0.978	0.983	0.955	0.980	0.951
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
$\log(HHI^{-m})$	-0.0249*** (0.00421)	0.00137 (0.00767)	-0.00998 (0.0106)	-0.0365*** (0.00491)	-0.0315*** (0.00606)
Observations	1,541,000	977,000	1,143,000	1,387,000	1,434,000
R-squared	0.983	0.966	0.967	0.947	0.962
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of mean earnings within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D25: Effects of Industrial Concentration on the 90/10 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Instrumental Variables, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
log(HHI)	0.369*** (0.0411)	0.0773*** (0.0203)	0.174*** (0.0208)	0.114*** (0.0236)	0.412*** (0.0640)
Observations	1,498,000	1,478,000	1,386,000	1,503,000	1,461,000
R-squared	0.880	0.891	0.776	0.916	0.813
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
log(HHI)	0.160*** (0.0254)	0.171*** (0.0419)	0.305*** (0.0612)	0.394*** (0.0419)	0.208*** (0.0359)
Observations	1,513,000	972,000	1,135,000	1,373,000	1,417,000
R-squared	0.884	0.861	0.850	0.769	0.801
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports instrumental variables estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 10th percentile of the earnings distribution within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D26: Effects of Industrial Concentration on the 90/10 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Reduced Form, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
$\log(HHI^{-m})$	0.160*** (0.0144)	0.0457*** (0.0120)	0.103*** (0.0125)	0.0546*** (0.0105)	0.198*** (0.0280)
Observations	1,524,000	1,500,000	1,403,000	1,529,000	1,481,000
R-squared	0.888	0.891	0.779	0.918	0.821
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
$\log(HHI^{-m})$	0.0781*** (0.0113)	0.102*** (0.0248)	0.150*** (0.0250)	0.170*** (0.0148)	0.101*** (0.0159)
Observations	1,541,000	977,000	1,143,000	1,387,000	1,434,000
R-squared	0.886	0.862	0.858	0.776	0.805
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 10th percentile of the earnings distribution within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D27: Effects of Industrial Concentration on the 50/10 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Instrumental Variables, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
log(HHI)	0.218*** (0.0337)	0.00351 (0.0163)	0.00353 (0.0162)	0.0988*** (0.0180)	0.375*** (0.0543)
Observations	1,498,000	1,478,000	1,386,000	1,503,000	1,461,000
R-squared	0.814	0.839	0.642	0.894	0.709
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
log(HHI)	0.109*** (0.0212)	0.000628 (0.0340)	0.204*** (0.0451)	0.314*** (0.0354)	0.224*** (0.0315)
Observations	1,513,000	972,000	1,135,000	1,373,000	1,417,000
R-squared	0.823	0.776	0.792	0.662	0.740
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports instrumental variables estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 50th percentile of the earnings distribution to the 10th percentile of the earnings distribution within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D28: Effects of Industrial Concentration on the 50/10 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Reduced Form, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
$\log(HHI^{-m})$	0.0946*** (0.0125)	0.00209 (0.00963)	0.00197 (0.00959)	0.0475*** (0.00788)	0.180*** (0.0237)
Observations	1,524,000	1,500,000	1,403,000	1,529,000	1,481,000
R-squared	0.819	0.839	0.641	0.895	0.719
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
$\log(HHI^{-m})$	0.0535*** (0.00972)	0.000284 (0.0203)	0.100*** (0.0176)	0.136*** (0.0129)	0.109*** (0.0139)
Observations	1,541,000	977,000	1,143,000	1,387,000	1,434,000
R-squared	0.825	0.776	0.797	0.667	0.745
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 50th percentile of the earnings distribution to the 10th percentile of the earnings distribution within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D29: Effects of Industrial Concentration on the 90/50 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Instrumental Variables, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
log(HHI)	0.150*** (0.0154)	0.0738*** (0.00939)	0.170*** (0.0127)	0.0148 (0.0110)	0.0371 (0.0335)
Observations	1,498,000	1,478,000	1,386,000	1,503,000	1,461,000
R-squared	0.863	0.885	0.826	0.925	0.820
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
log(HHI)	0.0502*** (0.0113)	0.170*** (0.0288)	0.102*** (0.0217)	0.0805*** (0.0219)	-0.0160 (0.0120)
Observations	1,513,000	972,000	1,135,000	1,373,000	1,417,000
R-squared	0.882	0.765	0.823	0.781	0.843
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports instrumental variables estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 50th percentile of the earnings distribution within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D30: Effects of Industrial Concentration on the 90/50 Earnings Ratio, 2005–2015, W-2 Earnings Measure, Reduced Form, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
$\log(HHI^{-m})$	0.0650*** (0.00621)	0.0436*** (0.00580)	0.101*** (0.00768)	0.00713 (0.00523)	0.0180 (0.0160)
Observations	1,524,000	1,500,000	1,403,000	1,529,000	1,481,000
R-squared	0.870	0.887	0.837	0.925	0.820
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
$\log(HHI^{-m})$	0.0246*** (0.00535)	0.102*** (0.0168)	0.0500*** (0.0106)	0.0347*** (0.00917)	-0.00770 (0.00590)
Observations	1,541,000	977,000	1,143,000	1,387,000	1,434,000
R-squared	0.882	0.772	0.830	0.782	0.842
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the ratio of the 90th percentile of the earnings distribution to the 50th percentile of the earnings distribution within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

Table D31: Effects of Industrial Concentration on the Gini Coefficient, 2005–2015, W-2 Earnings Measure, Instrumental Variables, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
log(HHI)	0.0291*** (0.00326)	0.0118*** (0.00242)	0.0365*** (0.00281)	0.00477* (0.00261)	-0.00780** (0.00354)
Observations	1,498,000	1,478,000	1,386,000	1,503,000	1,461,000
R-squared	0.930	0.937	0.872	0.937	0.893
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
log(HHI)	0.00758*** (0.00241)	0.0305*** (0.00496)	0.0261*** (0.00576)	0.0269*** (0.00329)	-0.00467 (0.00314)
Observations	1,513,000	972,000	1,135,000	1,373,000	1,417,000
R-squared	0.937	0.909	0.908	0.874	0.897
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports instrumental variables estimates of the effect of local industrial concentration, as measured by HHI, on the log of the Gini coefficient within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent semi-elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.



Table D32: Effects of Industrial Concentration on the Gini Coefficient, 2005–2015, W-2 Earnings Measure, Reduced Form, by Demographic Group

VARIABLES	(1) Men	(2) Women	(3) Age <25	(4) Age 25-54	(5) Age 55+
$\log(HHI^{-m})$	0.0126*** (0.00139)	0.00697*** (0.00146)	0.0216*** (0.00190)	0.00229* (0.00125)	-0.00368** (0.00170)
Observations	1,524,000	1,500,000	1,403,000	1,529,000	1,481,000
R-squared	0.934	0.938	0.882	0.937	0.893
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

VARIABLES	(1) White	(2) Black	(3) Hispanic	(4) LTHS/HS	(5) Some College+
$\log(HHI^{-m})$	0.00372*** (0.00116)	0.0182*** (0.00291)	0.0129*** (0.00281)	0.0116*** (0.00126)	-0.00225 (0.00154)
Observations	1,541,000	977,000	1,143,000	1,387,000	1,434,000
R-squared	0.937	0.913	0.913	0.878	0.897
Market FEs	Yes	Yes	Yes	Yes	Yes
CZ by Year FEs	Yes	Yes	Yes	Yes	Yes

Source: Longitudinal Business Database, Form W-2, and American Community Survey, 2005 through 2015; Decennial Census, 2000 and 2010; Census Numident.

Note: Table reports reduced form estimates of the effect of local industrial concentration, as measured by HHI, on the log of the Gini coefficient within markets, as measured by Form W-2, from 2005-2015, by demographic group. Columns represent separate regressions, which include the indicated fixed effects. Regressions are employment-weighted. Coefficients represent semi-elasticities. Sample sizes and statistic values have been rounded for disclosure avoidance. The White and Black categories refer to non-Hispanic White and non-Hispanic Black. The Hispanic category includes Hispanics of any race. LTHS refers to individuals with less than a high school diploma, HS refers to those with exactly a high school diploma, and “Some College+” refers to those who have at least attended some college.

## References

- Fort, T. C. and Klimek, S. D. (2018). The effects of industry classification changes on us employment composition. CES Working Paper 18-28.
- Wagner, D. and Layne, M. (2014). The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software. CARRA Working Paper No. 2014-01.