BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Prices and Living Conditions

Productivity Dispersion and Structural Change in Retail Trade

Dominic Smith,
U.S. Bureau of Labor Statistics
G. Jacob Blackwood,
Amherst College
Michael D. Giandrea,
U.S. Bureau of Labor Statistics
Cheryl Grim,
U.S. Census Bureau
Jay Stewart,
U.S. Bureau of Labor Statistics
Zoltan Wolf,
New Light Technologies

Working Paper 570
November 20, 2023

Productivity Dispersion and Structural Change in Retail Trade

By

Dominic Smith
U.S. Bureau of Labor Statistics

G. Jacob Blackwood
Amherst College

Michael D. Giandrea
U.S. Bureau of Labor Statistics

Cheryl Grim
U.S. Census Bureau

Jay Stewart
U.S. Bureau of Labor Statistics

Zoltan Wolf
New Light Technologies

November 20, 2023

Abstract

Official Bureau of Labor Statistics (BLS) estimates of productivity growth in the retail trade sector indicate that productivity has grown at a moderate rate of 2.8 percent per year between 1987 and 2017, and that there is considerable variation in growth rates across 4-digit industries. But the official data, which can be thought of as weighted averages of establishment-level productivity, tell us nothing about what goes on within industries. Given the transformation of retail trade over the past three decades, this information could provide more insight. In this paper, we present productivity dispersion statistics for industries in the retail trade sector. These statistics are similar to the BLS-Census Bureau Dispersion Statistics on Productivity (DiSP) for manufacturing industries and complement the official BLS industry-level productivity statistics. We find that from 1987 through 2017, productivity dispersion increased slightly on average. Surprisingly, the tails of the retail productivity distribution have similar dispersion as we find in the middle. Firm dispersion has increased more than establishment dispersion.

Keyword: retail, reallocation, business cycles, productivity dispersion

JEL Classification: D24, E24, L81

---

* Corresponding Author: Dominic Smith smith.dominic@bls.gov. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau or the Bureau of Labor Statistics. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7526913: CBDRP-FY23-CED006-0016). Blackwood was also a part-time employee of the U.S. Census Bureau at the time of the writing of this paper. We thank Emek Basker, Cindy Cunningham, Matthew Dey, Lucia Foster, John Haltiwanger, Rachel Nesbit, Kirk White, and audience members at the 2023 FSRDC Conference for their helpful comments.

1. Introduction

The U.S. retail trade sector has changed significantly since the 1980s, when single-unit firms accounted for most retail sales. The share of single-unit firms decreased as national chains expanded (Foster, Haltiwanger, Klimek, Krizan, Ohlmacher, 2016). These national chains operate similar stores in many markets, allowing them to reach many consumers and reduce costs through streamlined purchasing and distribution. More recently, e-commerce has upended retail trade (hereafter "retail") again as retail firms use the internet to reach consumers in markets where they have no physical presence.

Official data from the Bureau of Labor Statistics show that this transformation has been accompanied by moderate labor productivity growth. Between 1987 and 2017, retail labor productivity grew by 2.8 percent per year. However, productivity growth was uneven across industries as certain industry groups such as Non-Store Retailers (NAICS 4541)—which includes e-commerce establishments—experienced productivity growth of nearly 10 percent per year, while more traditional industry groups such as Grocery Stores (NAICS 4451) experienced almost no growth despite adoption of new technologies such as self-checkout machines.

In this paper, we examine the patterns of within-industry productivity dispersion. It is not clear, *a priori*, whether the retail transformation would result in decreased or increased dispersion across establishments. On one hand, stores within an industry may have become more similar over time as a few national chains have come to dominate many retail industries. If stores within a chain have similar productivity, we would expect within-industry productivity dispersion to decrease. On the other hand, the entering chains may be much more productive than incumbent stores, which may have survived because of their location or because they operate in niche markets. In this case, we would expect productivity dispersion to increase. Which of these forces dominate is important for understanding allocation of inputs and the distribution of productivity increases.

To answer these questions, we use the Census Bureau's Census of Retail Trade (CRT) and the Longitudinal Business Database (LBD) for 1987 through 2017 to show that, on average, labor productivity dispersion has increased slightly when measured by the interquartile range (about 10 log points) or standard deviation. This suggests that entering establishments may have different levels of productivity than incumbent establishments, which is consistent with prior work (Foster, Haltiwanger, and Krizan, 2006). However, further examination reveals a more nuanced story.

Dispersion in the tails of the productivity distribution has increased substantially over this period. The difference between the productivity of an establishment at the 99th percentile of the distribution and that of an establishment at the 90th percentile increased by about 20 log points on average. At the other end of the distribution, the difference between establishments at the 10th and 1st percentiles increased by a similar amount. Firm productivity dispersion has also increased substantially—and by more than establishment-level dispersion.

The remainder of the paper describes how we measure productivity and then documents trends in productivity levels and dispersion.

2. Measuring Productivity

Our primary goal is to create statistics that provide insights about productivity that complement the official BLS industry-level productivity measures. There are many reasons why our micro-aggregated estimates might not match the official estimates. So, our first step is to construct estimates using our microdata using a similar methodology to the official estimates. We first describe how BLS constructs industry-level estimates of productivity growth and then use the Census microdata to construct measures of output, hours worked, and labor productivity that match the BLS concepts as closely as possible.

2.1 BLS Industry-level Productivity

BLS publishes annual measures of real sectoral output, employment levels, hours worked, and labor productivity growth for two 2-digit, 12 3-digit, and 27 4-digit North

American Industry Classification System (NAICS) retail industries.[1] BLS also makes available a dataset that includes values for selected 5-digit and 6-digit NAICS Retail industries for which the underlying data are of high quality. For each industry, labor productivity (LP) growth is measured as the change in the ratio of indexes of sectoral output and labor. These LP growth rates are chained to construct productivity indexes.

BLS estimates real sectoral *output* using several data sources. Three surveys from the Census Bureau provide nominal output estimates: the Economic Census (conducted every 5 years), the Annual Retail Trade Survey (ARTS), and the Monthly Retail Trade Survey (MRTS). The ARTS collects data on total annual sales, e-commerce sales, sales taxes, end of year inventories, purchases, total operating expenses, and gross margins. The ARTS includes employer businesses classified in the retail sector. Firms without paid employees are included in the BLS estimates through imputation or administrative data provided by other federal agencies. Annual values from the ARTS are adjusted based on the Merchandise Line Sales from the Economic Census (or Product Line Sales in later years). BLS uses the most detailed ARTS sales data available for deflating, and then aggregates to the 4-digit NAICS industry level. BLS does not adjust for resales or changes in inventories.[2]

The industry-specific implicit deflators for output (sales) are constructed by dividing the index of current-dollar sales for all establishments in the industry, which are available annually, by the corresponding Tornqvist index of annual constant-dollar sales constructed using product deflators. The Tornqvist index is constructed by combining product-line deflators (available for each industry annually) with product-line-by-industry sales (available for each industry every five years and interpolated for the intervening years).[3] The first step is to match sales for each product line to the appropriate Consumer Price Index

[1] Real sectoral output is output that is sold to entities outside of the industry. For retail trade industries, real sectoral output is virtually the same as real gross output (gross sales).
[2] Resales in retail trade are negligible.
[3] Data on industry-by-product-line sales are published by the Census only every five years, and BLS interpolates values for the intervening years.

(CPI) or Producer Price Index (PPI),[4] and then deflate the individual product-line sales.

BLS measures *labor input* as the total annual hours worked by all persons in an industry. This measure is constructed by combining data from three surveys: the Current Employment Statistics (CES) survey, the Current Population Survey (CPS), and the National Compensation Survey (NCS). The CES provides detailed information on the employment and average weekly hours *paid* for production and nonsupervisory employees (henceforth referred to as nonsupervisory workers), which is used to calculate total hours paid for nonsupervisory employees. The NCS data, which include information on paid leave and other types of compensation, are used to estimate an hours-worked-to-hours-paid ratio that adjusts total nonsupervisory hours from an hours-paid to an hours-worked basis by removing paid vacation accrued and sick leave taken. To estimate supervisory employee average weekly hours, BLS uses data from the CPS to calculate a ratio of supervisory to nonsupervisory employee average weekly hours worked, which is then multiplied by the adjusted CES nonsupervisory employee hours (worked). Total hours worked by all employees is the sum of nonsupervisory worker hours worked and supervisory worker hours worked.[5] Self-employed and unpaid-family-worker (SE) hours are then added to this total. SE employment and average weekly hours are obtained from the CPS and multiple data sources (including nonemployer statistics, the Internal Revenue Service, ARTS, among others) are used to allocate SE hours to industries. The level of aggregation for the inclusion of SE workers is between the 6-digit and 4-digit NAICS industry level.

[4] BLS uses the CPI Research Series because the series is more consistent over time. The official CPI is never revised, whereas the CPS Research Series is revised to incorporate the current methodology into the historical data. In cases where there are multiple deflators for a product-line definition, BLS creates a deflator for that product line using relative importance values as weights and then uses that to deflate the single product line. This situation occurs when the deflators are defined at a more detailed level than the product lines.

[5] BLS recently changed its method for estimating employee hours. The new method uses the CES all-employee hours series as the main data source. It adjusts these data from an hours-paid concept to an hours worked concept using data from the NCS (to account for paid time off) and the CPS (to account for off-the-clock work). In the next draft of this paper, we will use these estimates for our comparison.

2.2 Establishment-level Productivity using Census Data

To measure establishment-level labor productivity, we combine restricted-use establishment-level microdata files from the Census Bureau with public-use industry-level data on prices and hours worked from BLS.

Our establishment-level microdata come from the Census of Retail (CRT) and the Longitudinal Business Database (LBD). The CRT is collected every 5 years in years ending in "2" and "7" as part of the Economic Census. The frame includes all retail establishments of multi-unit firms and a sample of single-unit retail firms, and it collects data about sales, employment, and other relevant information.[6] To increase the coverage of single-unit firms, the Census Bureau uses information from administrative records to impute sales data. The LBD is a longitudinally linked version of the Census Bureau's Business Register that covers the non- agricultural employer universe of business establishments (see Jarmin and Miranda, 2002 and Chow et al., 2021). The LBD provides us with employment and high-quality longitudinal links.

Ideally, we want to construct an *output* measure using the microdata that matches the BLS measure. The underlying data source for our output measure, the Census of Retail Trade, is the same as the one used by the BLS. However, the BLS uses each year's published aggregates, which are based on industry codes available in that year. When industry codes change, BLS re-estimates sales by industry to create a consistent series over time. In contrast, our data contain NAICS codes for each establishment in every year (Fort and Klimek, 2018). These codes are based on more detail than is available in the aggregate data. These differences lead our estimate of sales to differ more before the SIC-to-NAICS transition and for industries where NAICS codes have changed over time.

We measure *labor input* as total hours worked. For each establishment, the LBD

[6] A significant caveat is that the CRT does not collect information about capital stocks or investment.

collects the total number of employees on the payroll for the March 12 pay period. We calculate total annual hours worked by multiplying total employment by the average annual hours worked per employee in the most detailed NAICS code available in CES data. Thus, all of the between-establishment variation in hours worked in a 4-digit industry is due to variation in employment across establishments in the industry and variation in hours across 6-digit industries within the 4-digit industry. We calculate establishment-level (denoted by subscript e) log labor productivity as:

$$productivity_e = \ln\left(\frac{real\ sales_e}{hours_j \times employment_e}\right)$$

where j is the most detailed industry code for which data are available.

Two concerns with our retail employment measures are: (1) employment can be seasonal and (2) many employees may be part-time. Because the number of full-time employees in the LBD in March may not reflect each establishment's use of labor inputs, we also measure establishment-level log wage productivity as: real sales divided by total payroll for the year deflated by the CPI. To illustrate, if one establishment uses one full-time worker and another uses two part-time employees, labor input will be more-accurately reflected in the wage bill than in the employment count.[7]

3. Comparing Micro-Aggregated Data to Published Industry Data

In this section, we compare our micro-aggregated estimates to the official data published by BLS, covering the 1987–2017 period. Based on earlier work comparing similar business data across the two government agencies, we expect that there will be some systematic differences between these measures (Elvery et al., 2006). Even though differences in the levels of the micro-aggregated and published first moments do not directly affect our conclusions about dispersion (because our measures are mean invariant), it is useful to

[7] The wage bill is not a perfect measure as it is contaminated by differences in the occupational mix, overtime, and other factors that affect wages.

determine how far apart the two sets of estimates are. If the first moments are close, then it is more reasonable to think of the micro-based second moments as measuring variation around the published first moments. We start by comparing employment and nominal sales levels and then compare productivity levels. We use the BLS average weekly hours series and deflators directly from BLS industry productivity data, which implies that all level differences come from differences in sales, employment, or sample coverage.

3.1 Input and Output Measures

We compare total nominal retail sales growth across the BLS and Census microdata series (Figure 1). We do not deflate these series since the same deflator would be used for both. The two series follow each other very closely in both changes and levels. Both series start at about 2.3 trillion in 1987. Sales growth was modest between 1987 and 1992, accelerated between 1992 and 2007, fell between 2007 and 2012, resuming the accelerated growth between 2012 and 2017. The series diverged starting in 1997, with the BLS series growing at a slightly faster rate.

Next, Figure 2 shows the total number of employees in the retail sector from each series. BLS Retail industry employment in 1987 was just below 14 million workers. Employment grew slowly between 1987 and 1992 and then accelerated between 1992 and 1997, when it grew by over 10 percent. Employment growth slowed after 1997, with a dip in employment between 2007 and 2012. The aggregated Census microdata follow this general trend, but at somewhat lower levels than those reported in BLS data. The vast majority of the difference between the series comes from BLS including self-employed and unpaid family workers in the employment totals according to an analysis using unreleased BLS estimates.

3.2 Productivity Growth

To compare growth rates, we start by calculating 5-year growth rates from the BLS indexes of labor productivity described above. From the sample microdata, we calculate aggregate labor productivity as average labor productivity across all establishments in the sample, and then calculate 5-year growth rates. Figure 3 compares the growth rates of these two series. The series exhibit nearly identical growth rates in three of the six years. In the 1997-2002 and 2002-2007 periods, the microdata show five-year growth rates around 10 percent while the BLS estimates are almost 20 percent and over 25 percent.[8] Between 2002 and 2007 the microdata show productivity declines while the BLS data show slight increases. On average, over the entire period the BLS estimates larger productivity growth than in the microdata.

These comparisons suggest that the micro-aggregated productivity estimates are broadly consistent with published BLS data.

4. Productivity Dispersion

Our analysis of productivity dispersion focuses on log levels rather than growth rates. Because we are interested in comparing within-industry productivity dispersion across industries and over time, we normalize establishment-level productivity values by removing industry and year effects.

Our primary measure of productivity dispersion is the interquartile range (IQR) of establishment-level log-productivity calculated for each 4-digit NAICS industry in each year. The IQR shows the difference between the 75th and 25th percentiles of the log-labor-productivity distribution in an industry-year cell. In addition, we calculate the difference between the 99-90, and 10-1 percentiles, which together provide additional detail about the productivity dispersion distribution. The IQR provides information about the center of the support of the distribution, while the 10-1 and 99-90 differentials inform about dispersion in the left and right tails.

We calculate the productivity dispersion of establishments in each of the 27 4-digit

[8] The conversion from SIC to NAICS codes seems like a logical place to look for the reasons behind these differences. Also, the BLS and Census business registers do not always assign the same industry codes to the same establishments. In the official BLS productivity statistics, the employment data are based on BLS industry coding while output data are based on Census industry coding. Both employment and output are based on Census industry coding in our micro-aggregated estimates.

industries in retail for each year of our sample. We also take a closer look at several of these industries, in part to illustrate the variation across selected industries, but also because of their relative importance and because of their growth over the last 30 years. The two largest 4-digit industries in retail are Grocery Stores (4451) and General Merchandise Stores, including Warehouse Clubs and Supercenters (4523). These two combined account for approximately one-fifth of retail sales in 2017.

Figure 4a shows the IQR, the 99-90 range and 10-1 range of labor productivity of establishments for the Grocery Stores (NAICS 4451). The IQR was essentially flat over the entire sample period, increasing slightly from 0.73 in 1987 to 0.79 in 2017. That is, the grocery store at the 75th percentile was 2.1 ($e^{0.73}$) times as productive as a store in the 25th percentile in 1987 and 2.2 ($e^{0.79}$) times as productive in 2017. Dispersion in the tails followed a different pattern. Dispersion in both tails grew more significantly between 1987 and 2012. In the upper tail of the distribution, productivity differences among the most productive stores grew about 25 log points such that, in 2017, the store at the 99th percentile generated about twice as much revenue as the store at the 90th percentile, with most of the increase occurring between 1987 and 1992. Productivity dispersion is greater among the least productive stores: except for 1987 and 1997, the 10-1 difference was about 100 log points, which implies that establishments at the tenth percentile are 2.7 times as productive as those at the 1st percentile. Dispersion in the upper and lower tenth of the distribution is about the same as for the middle half of the distribution.

Figure 4b presents analogous results for the other large retail industry, General Merchandise Stores (NAICS 4523). The time trends of the three lines are similar to those in Figure 4a, although there are more fluctuations in the IQR and 99-90 range in general merchandise stores, sharper declines in the tails in 2017, and there is a slight decline in the IQR over the sample period. And as with grocery stores, the dispersion in the top and bottom tenth of the distribution is at least as great as dispersion in the middle half.

Another retail industry of particular interest is Electronic Shopping and Mail-Order Houses (NAICS 4541) which has exhibited substantial growth since 1987. Real sales in this industry have grown from $10.5 billion in 1987 to $532.9 billion in 2017 (2012 constant dollars). Along with the dramatic increase in sales, there has been an increase in all three dispersion measures (Figure 4c). The IQR increased from about 1.1 in 1987 to almost 1.5 in 2017. In other words, in 1987, the 75th percentile establishment was 3 times as productive as the establishment at the 25th percentile. This multiple grew over thirty years to 4.5 in 2017. With the dramatic growth in industry sales combined with the transformation of this industry from catalogs to online, some establishments may have been able become more efficient over time. However, at least some of the dispersion in this industry is due to differences in which product categories are sold by each retailer and retailers in this industry may have substantial variation in their tendency to outsource sales related activities such as delivery logistics.

The final industry we consider is Clothing Stores (NAICS 4481). Figure 4d shows that the interquartile range for establishment productivity was relatively constant over the sample period, ranging between 0.76 and 0.83 (multiples of 2.1 and 2.3), while both the 99-90 and 10-1 ranges increased substantially. Compared with the other three industries, the dispersion ranges for clothing stores exhibited less variation over 5-year intervals.

These figures highlight some of the similarities and differences across industries. The variation among industry productivity distributions can be substantial and can also exhibit different trends over time. We saw that grocery stores and general merchandise stores were fairly similar. It is also worth reiterating that in some industries there is as much or more dispersion in the upper and lower tenth of the distribution as there is in the middle half.

4.1. Dispersion in Dispersion

We now describe the evolution of the cross-industry distribution of dispersion measures, or dispersion in dispersion. Because activity is heavily concentrated in a few industries, we weight industries by their share of total annual hours in the retail sector when calculating these cross-industry statistics, but we do not weight establishments when constructing industry IQRs.[9] Figure 5a shows the industry-weighted mean, median, and 25th and 75th percentiles of the IQRs for the 27 4-digit retail industries. The median IQR is 0.80 in 1987, which means that the industry with an IQR that is at the median across all retail industries, the establishment in the 75th percentile is approximately 2.2 times as productive as the one at the 25th percentile. The median IQR then increases over time before peaking in 2007 at 0.98 (multiple of 2.7). The median IQR decreases over the last ten years of our sample to below 0.90 in 2017. The mean IQR shows similar patterns but exhibits less variability. Comparing the 25th and 75th percentiles of the IQR distribution, we can see that the IQR varies considerably across industries. In particular, the difference between high-dispersion and low-dispersion industries is about 15 log points in 1987 (the difference between the long-dashed line and the short-dashed lines is 0.88 − 0.73 = 0.151, which translates to a multiple of 1.2). This value increases to almost 30 log points in 2012 and then decreases to 19 log points in 2017. These dynamics indicate that productivity differences across establishments in already high-dispersion industries grew more until 2012, while productivity differences in low-dispersion industries exhibited a smaller increase.

Figures 5b and 5c show the same statistics for the right and left tails of the distribution of IQRs (the 99-90 and 10-1 ranges). On average, compared to Figure 5a, dispersion in dispersion is lower in the right tail of the IQR distribution (Figure 5b) and slightly greater in the left tail of the distribution (Figure 5c). But they still indicate substantial dispersion when one considers that the IQR is 5 times as large as either of the tails. In both Figures 5b and 5c, the lines for the 25th and 75th percentiles are closer to each other than is the case in Figure 5a. Thus, in both the right and left tails of the IQR distribution, at a given point in time, there is

[9] As previously noted, the two largest industry groups account for about one-third of retail activity in 2002 (Smith and Ocampo, 2023).

less cross-industry dispersion in dispersion than is the case for the middle of the distribution. Comparing the same lines at different points in time, we see that dispersion in dispersion has increased in both tails over our sample period, although the lines for both tails of the IQR distribution turn down toward the end of our sample period. For example, the median for establishments in the right tail of the productivity distribution increased from 0.54 in 1987 to 0.70 in 2017, which implies that the difference between establishments at the 75th percentile and the 25th percentile (of the right tail) increased from a multiple of 1.7 to a multiple of 2.0. And the median IQR in the 10-1 range increased from 0.72 in 1987 to 0.98 in 2017, which implies that the median IQR multiple increased from 2.1 to 2.7.

4.2 Weighted vs. Unweighted Statistics

Given that grocery stores and general merchandise stores employ a large share of workers in the retail trade sector, they have an outsized impact on our dispersion statistics. To examine this impact, we recalculated the establishment-level dispersion statistics on an unweighted basis. That is, we did not weight industries by their share of total hours worked. Results are shown in Figures 6a through 6c. Comparing Figure 6a to Figure 5a shows that every data point in Figure 6a is above the corresponding data point in Figure 5a. For example, in 2017 in the median industry, the 75th percentile establishment was approximately 2.6 times as productive as the 25th percentile establishment compared to the industry-weighted factor of 2.4 in Figure 5a. The fact that measured dispersion is greater when industry weights are not used implies that small industries exhibit greater dispersion in the middle half of the productivity distribution relative to large industries. An interesting difference between Figures 5a and 6a is that the graphs for the 75th percentile of IQRs start at different levels and initially move in opposite directions. The lower initial dispersion in the weighted graph implies that, among industries that have the greatest dispersion in their IQRs, large industries exhibit less dispersion. The increase in the 75th percentile of weighted IQRs combined with the decline in the 75th percentile of unweighted IQRs suggests that dispersion increased in large industries and decreased in small industries.

In contrast, Figures 6b (right-tail dispersion) and Figure 6c (left-tail dispersion) are more similar to Figures 5b and 5c, which implies that dispersion in the tails of the productivity distribution is more similar across small and large industries. The main difference between Figures 5b and 6b is that there is more time-series variation in the weighted statistics compared to the unweighted statistics. The dispersion statistics for the left tail of the IQR distribution (Figures 5c and 6c) look remarkably like each other. This suggests that establishments in the lower tail of the productivity distribution do not exhibit a systematic pattern according to industry size.

There is significant variation within industries in establishment size and in the size of retail industries. Thus, dispersion measures that weight establishments within industries or that give industries may exhibit different patterns. Figure 7 shows the mean industry-weighted IQR, mean unweighted IQR, and the mean IQR where with both establishment and industry weights are used. The top two lines show that industry-weighting yields lower dispersion, which suggests larger industries typically exhibit lower dispersion than smaller industries. In addition, the cross-industry mean of activity weighted dispersion is lower and seems to exhibit a negative trend, which means larger establishments tend to have lower and decreasing dispersion relative to smaller establishments.

The difference between the industry-weighted and the industry-establishment-weighted means can be thought of as difference between establishments and workers. That is, establishment-weighted distribution statistics can be thought of (roughly) as dispersion of worker productivity, where each worker is assumed to have the productivity of the establishment. Thus, we would expect establishment-weighted IQRs to exhibit less dispersion than IQRs that are not establishment weighted, which we see is the case in Figure 7

.

4.3 Establishments vs. Firms

The rich dynamics described above could be due to several factors. One is the appearance and eventual dominance of national chains. The entry of a large firm, with potentially hundreds or even thousands of similar establishments, could create a mass point in the establishment productivity distribution if most productivity variation is across firms. This would reduce dispersion across establishments. However, these establishments may also have very different productivity than existing establishments which would raise dispersion. For some questions, it is more useful to examine the dispersion of productivity across *firms*, rather than establishments. The firm then becomes a single observation within an industry rather than hundreds of establishment observations and is treated the same as a firm that operates as a single establishment. Further, by comparing establishment-based and firm-based dispersion it is possible to make inferences about the nature of the retail trade transformation. To shed some light on the importance of such effects, we calculate dispersion among firms in 4-digit industries.

Figure 8a shows descriptive statistics for the cross-industry distribution of firm-level dispersion. The median IQR (0.83 in 1987) rises to 1.0 in 2017. This means that in the median industry, the firm at the 75th percentile was approximately 2.3 times as productive as the firm at the 25th percentile in 1987 and that the multiple increased to 2.7 in 2017. Comparing these dispersion statistics to the establishment-level statistics in Figure 5a, we see that firm-level dispersion is greater at each percentile. A possible explanation for the greater productivity dispersion is that large, multi-establishment retail firms operate establishments that have similar productivity levels. Thus, the mass of the establishment-level productivity dispersion is grouped together into a single observation, which means that there are fewer observations that are like each other resulting in larger productivity differences. Put another way, large national firms show up as an interval in the establishment distribution but as a single point in the firm distribution.

Figures 8b and 8c present the dispersion-in-dispersion data for the right and left tails, respectively. Right-tail dispersion is generally smaller than the IQR: after rising steeply from 1987 to 1992, median right-tail dispersion averages about 0.7 in 2017, with generally small cross-industry differences. In contrast, left-tail dispersion (Figure 8c) is more similar to IQR dispersion, both in terms of level of dispersion and also cross-industry differences. For example, median dispersion increased from 0.72 in 1987 to 1.02 in 2017. The large dispersion in the left-tail of the firm distribution may indicate that there may be a significant number of low-productivity firms that remain in business over time.

4.4 Robustness

We use industry averages for hours per employee, which means that the variation in total hours across establishments is driven entirely by variation in employment. Thus, we may overestimate dispersion if there is significant variation in the use of part-time vs. full-time employees because we assume that hours per employee are the same across establishments in an industry. Given the difficulty of measuring hours in the CRT, we replicated Figures 5a,b,c using total payroll in the denominator rather than total hours (Figures 9a,b,c). Payroll has the advantage of accounting for differences in part-time vs. full-time work. But payroll also reflects other things besides hours worked such as the mix of occupations (which could also reflect the substitution of other factors for labor), leave policies, and any compensating wage differentials.

There are some fairly significant differences between Figures 5 and 9. Starting with the “a” panels, there is considerably less dispersion when using payroll. All three lines in Figure 9a are well below their counterparts in Figure 5a. The second notable difference is that, starting in 1992, payroll-based productivity dispersion decreased (except for the median IQR) whereas hours-based productivity dispersion increased. Turning to the 99-90 ranges in Figures 5b and 9b, we see similar differences in time trends but relatively little differences in levels. Finally, looking at the lower tails, the lines in both Figures 5c and 9c trend up. The payroll-based productivity dispersion statistics are quite a bit lower than the hours-based measures and they exhibit significantly less variation.

Finally, we calculate the standard deviation of labor productivity for each industry and plot the distribution of this statistic across industries in Figure 10. This measure may be sensitive to outliers which is why it is not our preferred statistic. Regardless, we find that dispersion with standard deviations exhibits the same patterns that we documented above.

4.5 Comparison to Manufacturing Dispersion

This paper marks the second description of labor productivity dispersion data. The first release focused on the manufacturing sector, summarized in Cunningham et al. (2022). We replicate (and extend) their analysis to compare and contrast with our results in retail, and find they are similar to and different from the manufacturing results in interesting ways.[10] Figure 11 displays means and medians of IQRs for both sectors. On average, dispersion is similar in both sectors, with somewhat more dispersion on average in retail. In the average year over the sample, the 75th percentile plant exhibits approximately 2.5 times as much output per hour as the 25th percentile plant in the typical industry in the manufacturing sector, while the 75th plant is roughly 2.55 times as productive as the 25th percentile plant in the typical retail industry. Average dispersion has risen modestly in both sectors when plants and sectors are weighted equally, but the increase is more substantial in manufacturing, and the trend has reversed somewhat in the past decade for retail.

Figures 12a and 12b present a tale of two tails in retail and manufacturing. Right/upper tail dispersion, as captured by the difference between the 99th percentile and 90th percentile plant (Figure 12a), are remarkably similar across the retail and manufacturing sectors when

[10] We compare unweighted statistics, both across and within industries, because the Dispersion in Productivity Statistics (DiSP) data do not include industry weights.

industries are weighted equally, both in levels and time series patterns. Broadly speaking, dispersion in the right tail is substantial in both sectors, with the 99th percentile plant exhibiting about twice as much productivity as the 90th percentile plant in the typical industry. This dispersion has risen notably in both broad sectors, from a multiple of 1.8 between the 99th and 90th percentile firms in the typical retail industry and 1.9 in manufacturing in 1987, to roughly 2.1 in both by 2017. For the lower/left tail (Figure 12b), however, dispersion in substantially lower in the manufacturing sector, where the typical industry has a 10th percentile plant that is 1.7 times as productive as the 1st percentile plant, than in the retail sector, where the 10th percentile plant is more than 2.5 times as productive as the 1st percentile plant. In fact, the lower tail is more dispersed than the upper tail in the typical retail industry, while the opposite is true in manufacturing. The average industry in both sectors exhibits rising dispersion in both tails, with a more pronounced increase in the typical retail industry.

Figure 13 illustrates how dispersion across industries in measures of within industry dispersion have changed over time. For example, the black lines demonstrate the difference between the 75th and 25th percentile industries in terms of IQR within each sector (the IQR of the IQR). Dispersion (as measured by the IQR) in the 75th percentile industry was roughly 1.32 times that of the 25th percentile (unweighted) industry over the sample period in retail, similar to the multiple of 1.4 in manufacturing. However, cross-industry dispersion has fallen notably in retail, with the ratio of the interquartile endpoints falling to 1.2 by 2017 as the ratio rose to 1.5 in manufacturing. Generally speaking, cross-industry dispersion in tail dispersion is higher in the manufacturing sector, suggesting more heterogeneity across industries within the manufacturing sector relative to retail both in the center of the establishment labor productivity distribution and in the tails.

5. Concluding Remarks

In this paper we extend the description of dispersion statistics on productivity from the

Manufacturing sector to the Retail Sales sector. We discussed the measurement issues relating to retail industries and detailed the methods by which labor productivity is measured at the retail industry group level. We found that there is significant within-industry variation in productivity. Dispersion has increased since 1987, particularly in the earlier years of our sample. Dispersion measured at the firm level is larger than establishment dispersion in both levels and changes.

Our results suggest that the structural transformation of the retail sector has resulted in increased dispersion, particularly in the tails of the productivity distribution. In most industries the largest increases were in the 1990s and 2000s when large national chains were expanding and consolidating in local markets. The increases have been smaller more recently when e-commerce has become an important trend in the retail sector.

Future work will focus on exploring the spatial variation in productivity dispersion and decomposing productivity growth into contributions by entering, continuing, and exiting establishments and firms.

References

Bureau of Labor Statistics, U.S. Department of Labor, and the Census Bureau, U.S. Department of Commerce, "Dispersion Statistics on Productivity," 1997–2016. https://www.bls.gov/productivity/articles-and-research/dispersion-statistics-on-productivity/ and https://www.census.gov/disp.

Chow, Melissa C., Teresa C. Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White, 2021). "Redesigning the Longitudinal Business Database." NBER Working Paper 28839. https://doi.org/10.3386/w28839.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf. 2021a. "Dispersion in Dispersion: Measuring Establishment Level Differences in Productivity." IZA Discussion Paper No. 14459. http://ftp.iza.org/dp14459.pdf.

Elvery, Joel, Lucia Foster, C.J. Krizan, and David Talan. 2006. "Preliminary Micro Data Results from the Business List Comparison Project." In 2006 Proceedings of the American Statistical Associations, Business and Economics Statistics Section. Alexandria, VA: American Statistical Association.

Fort, Teresa and Klimek, Shawn, 2018) The Effects of Industry Classification Changes on US Employment Composition, Working Papers, U.S. Census Bureau, Center for Economic Studies, https://EconPapers.repec.org/RePEc:cen:wpaper:18-28.

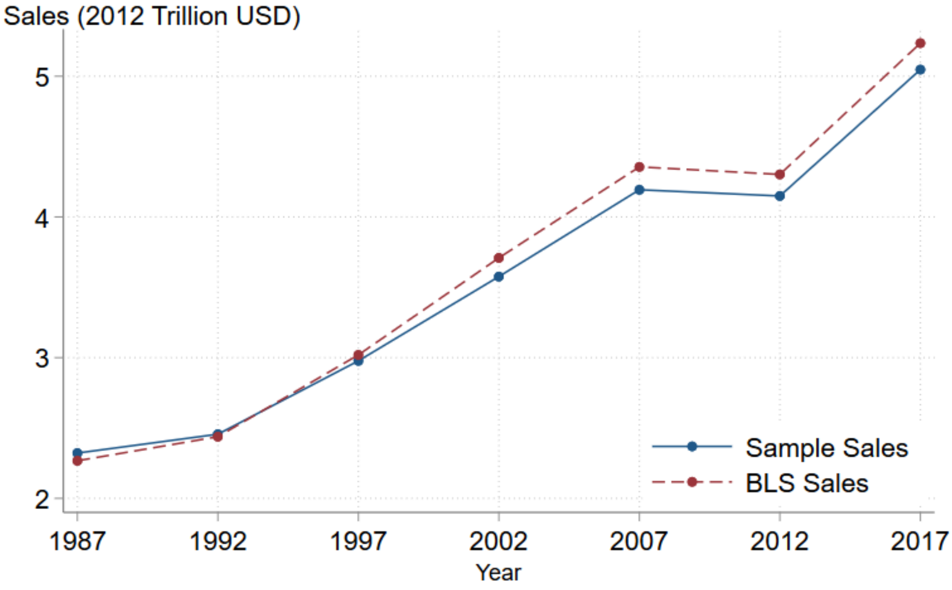
Foster, Lucia, John Haltiwanger, Shawn Klimek, Cornell J. Krizan, and Scott Ohlmacher. 2016. "The Evolution of National Retail Chains: How We Got Here." *Handbook on the Economics of Retailing and Distribution*, edited by Emek Basker pp. 7-37, ch 1, https://dx.doi.org/10.4337/9781783477388.

Foster, Lucia, John Haltiwanger, and Cornell J. Krizan. 2006. "Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s." *The Review of Economics and Statistics*, 88(4): 748-58. https://doi.org/10.1162/rest.88.4.748

Jarmin , Ron S and Javier Miranda, 2002. "The Longitudinal Business Database," Working Papers 02-17, Center for Economic Studies, U.S. Census Bureau. https://ideas.repec.org/p/cen/wpaper/02-17.html

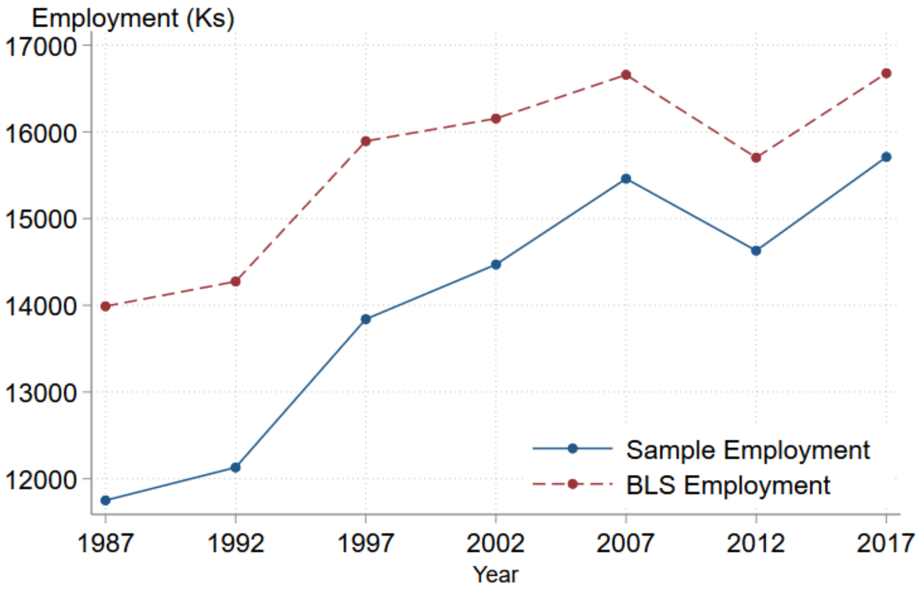
Smith, Dominic A. and Sergio Ocampo. 2023. "The Evolution of US Retail Concentration." https://doi.org/10.48550/arXiv.2202.07609

Figure 1: Retail Sales



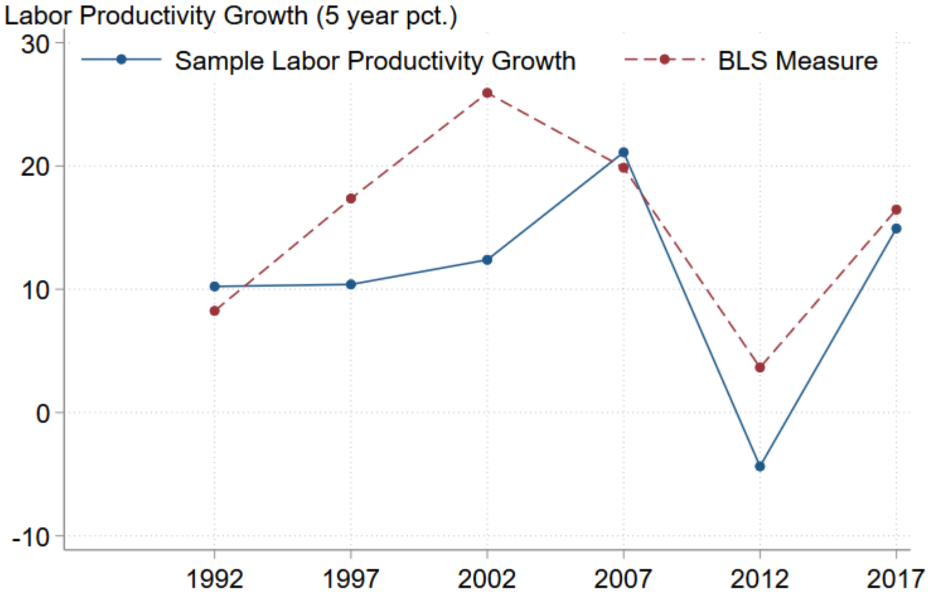
Notes: BLS sales is the sales according to the BLS industry statistics. Sample sales is the sales of establishments in the sample created by combining the CRT and LBD. Sample sales come from the CRT.

Figure 2: Retail Employment Levels



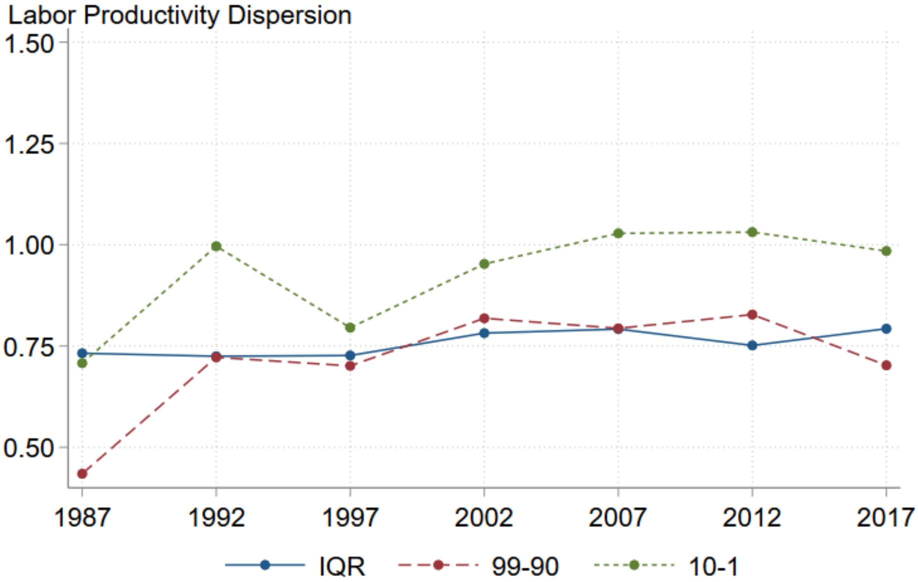
Notes: BLS employment is the employment according to the BLS industry statistics. Sample employment is the employment of establishments in the sample created by combining the CRT and LBD. Sample employment comes from the LBD. The two employment measures differ, in part, due to their treatment of nonemployer establishments.

Figure 3: Productivity Growth by Data Source and Measure (in percent)

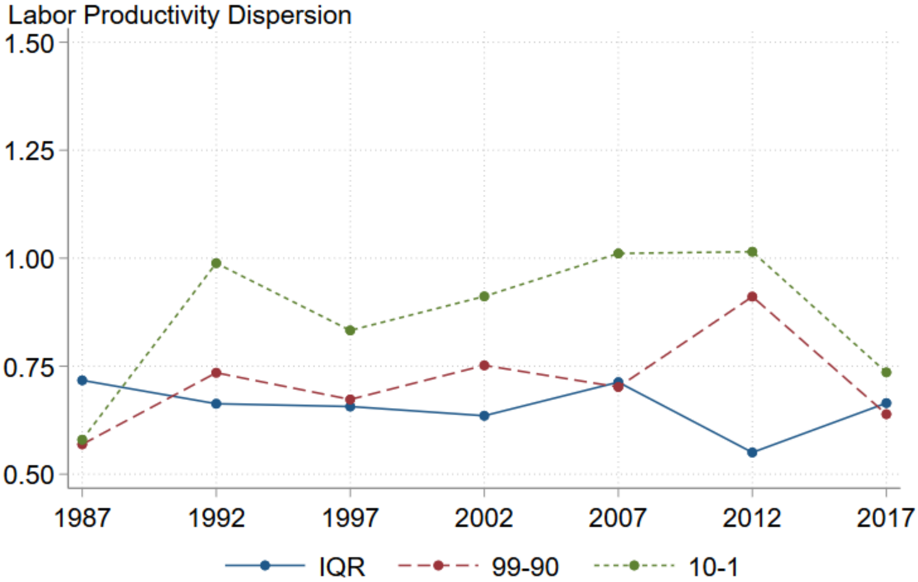


Notes: Microdata labor productivity growth plots five-year growth rates using sales from the CRT, employment from the LBD, and hours from the BLS industry productivity statistics. Productivity is calculated for each establishment and then averaged to calculate aggregate productivity. The establishments in the top and bottom one percent of the productivity distribution for each industry group are excluded from the calculation. Official BLS Measure is the percent change in index levels.

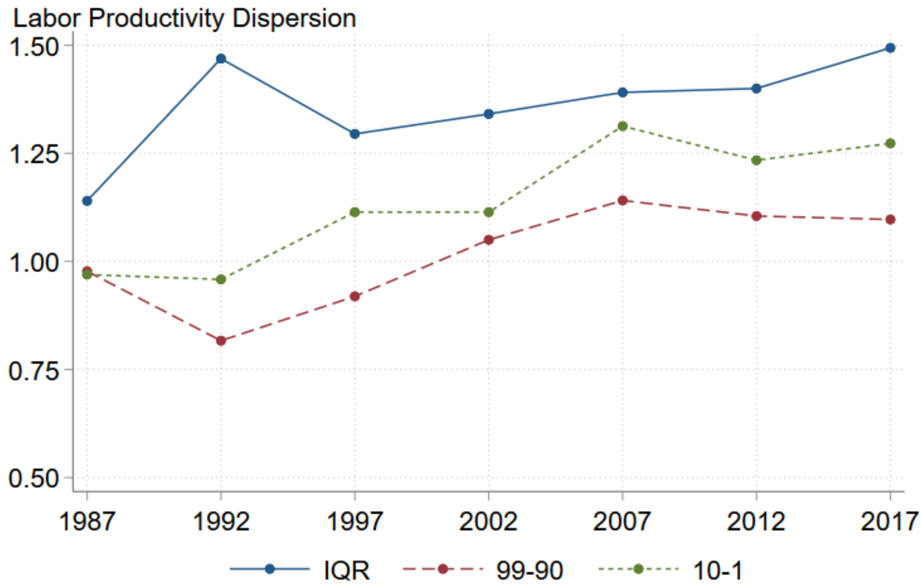
Figure 4: Within Industry Dispersion



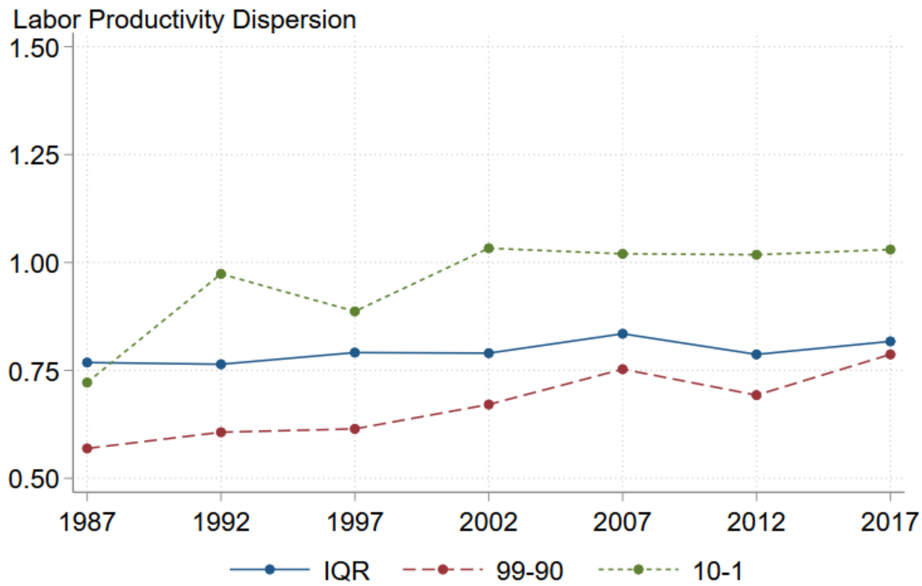
a) 4451 – Grocery Stores



b) 4523 – General Merchandise Stores, including Warehouse Clubs and Supercenters



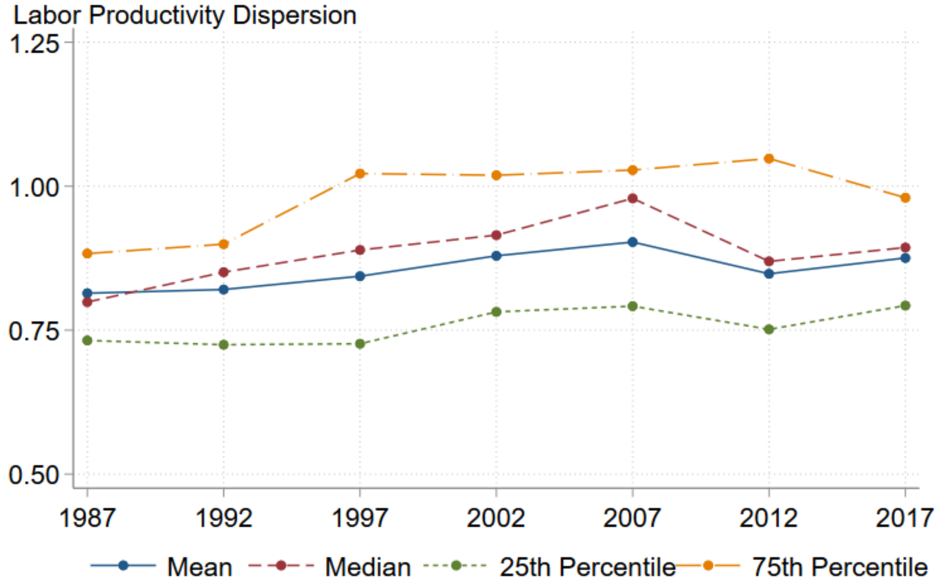
c) 4541- Electronic Shopping and Mail-Order Houses



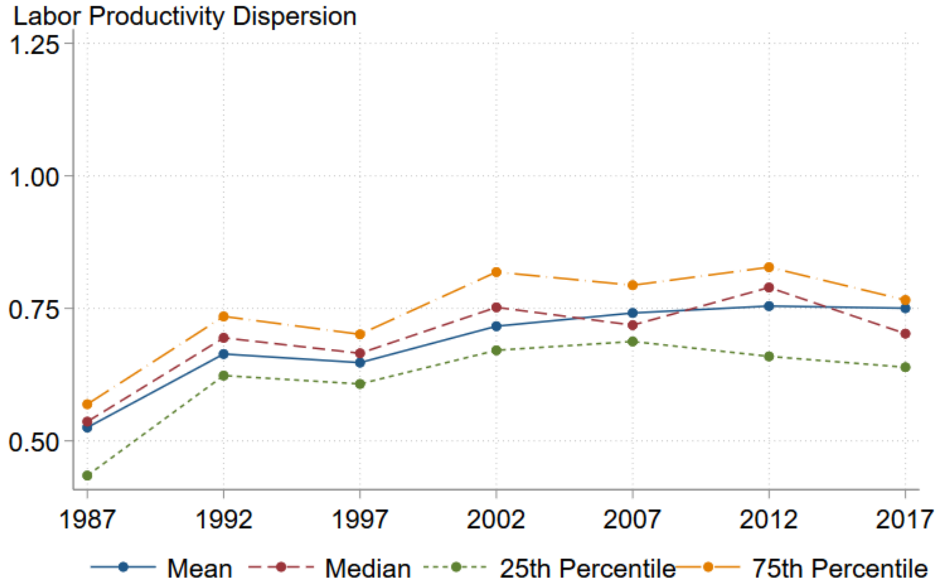
b) 4481 - Clothing Stores

Notes: The figures represent the difference between establishments at different percentiles of the within-industry labor productivity distribution in each year. The IQR is the difference between the 75th and 25th percentile. The y-axis is expressed in log points. Calculations use data from the CRT, LBD, and BLS industry-productivity statistics.

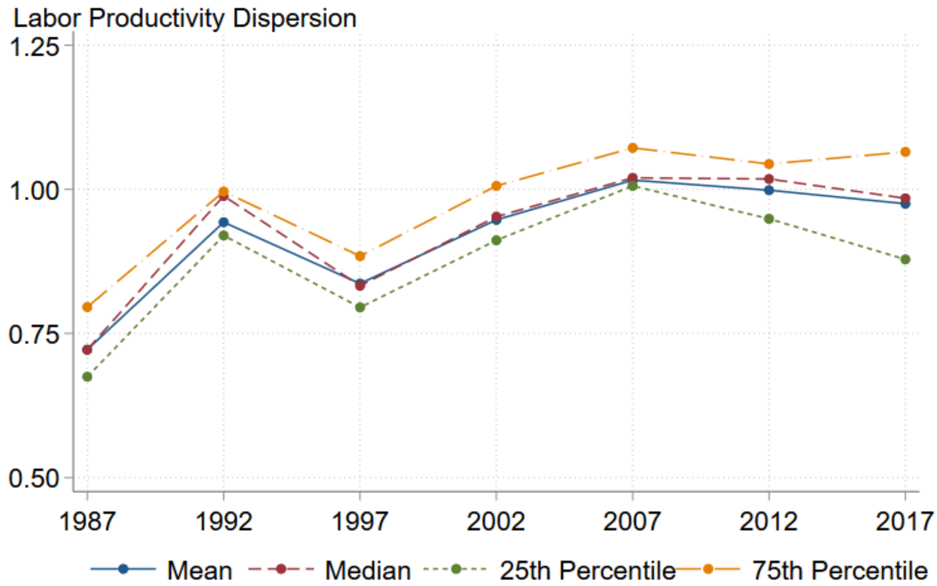
Figure 5: Distribution of Establishment Labor Productivity



a) IQR



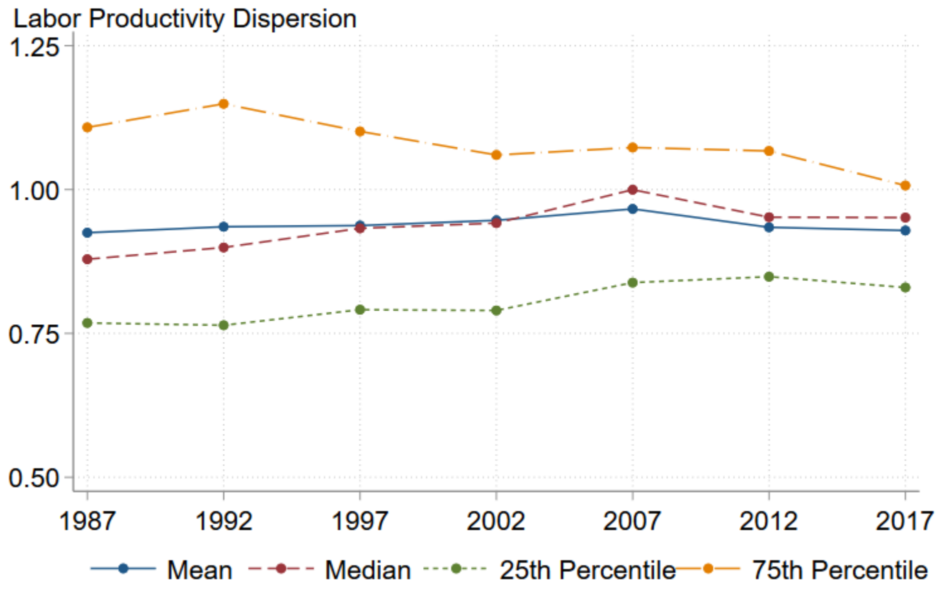
b) 99-90



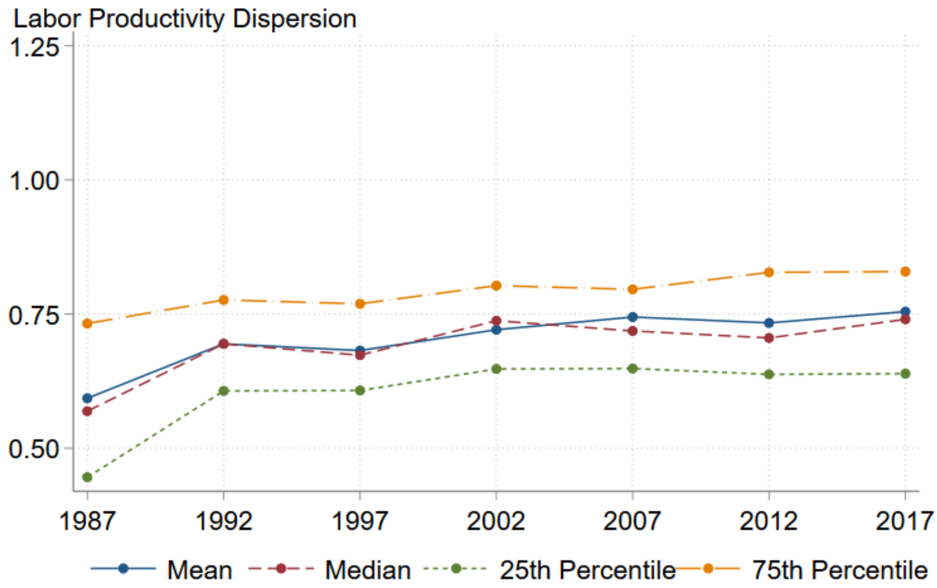
c) 10-1

Notes: Lines represent the variation across industries in the within-industry statistics as indicated by the subfigure caption. The lines indicate the 25th percentile, median, 75th percentile, and mean within-industry IQR across industries. NAICS-4 industries are weighted according to their share of total hours within a year. Establishments within each industry are given equal weight. Calculations use data from the CRT, LBD, and BLS industry-productivity statistics.

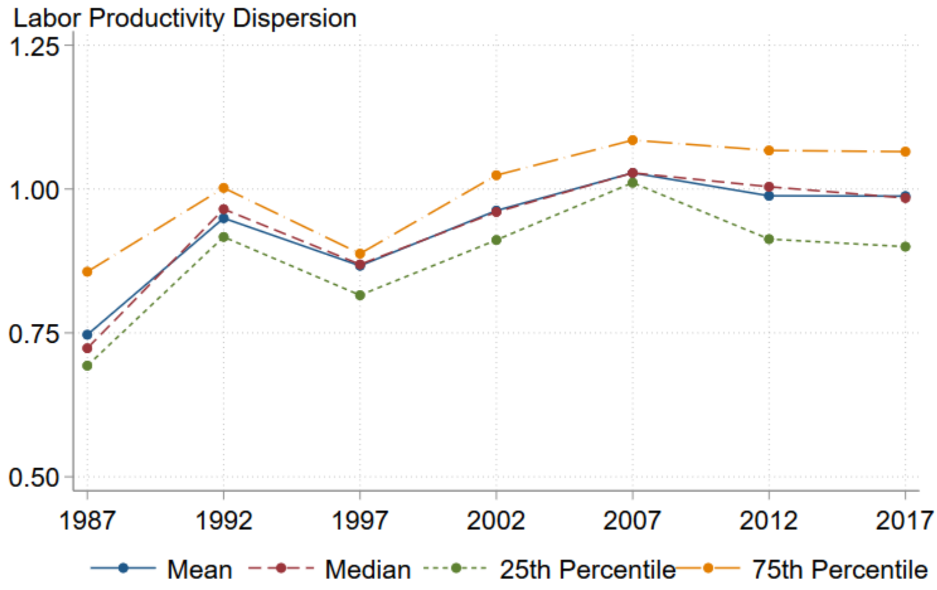
Figure 6: Distribution of Establishment Labor Productivity, Unweighted



a) IQR



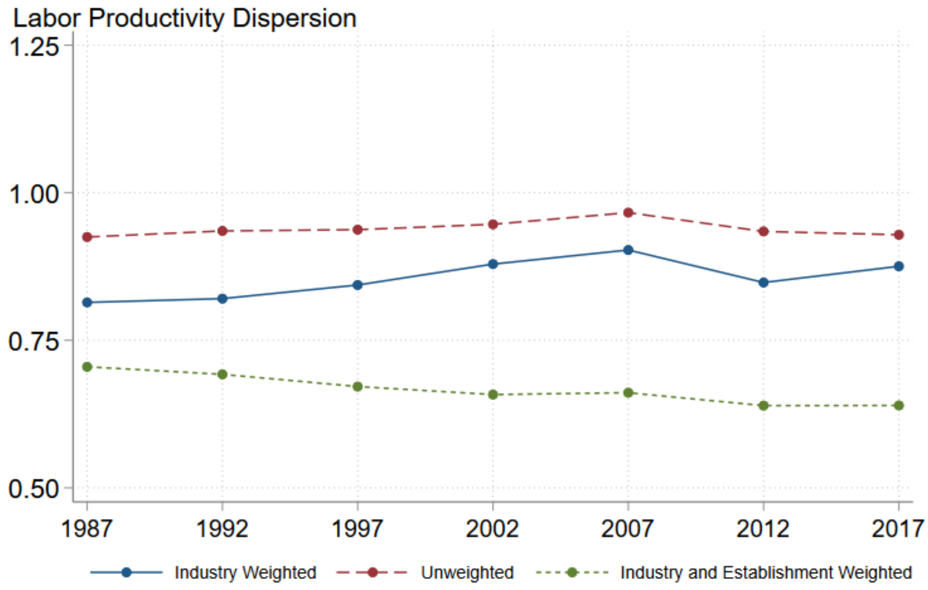
b) 99-90



c) 10-1

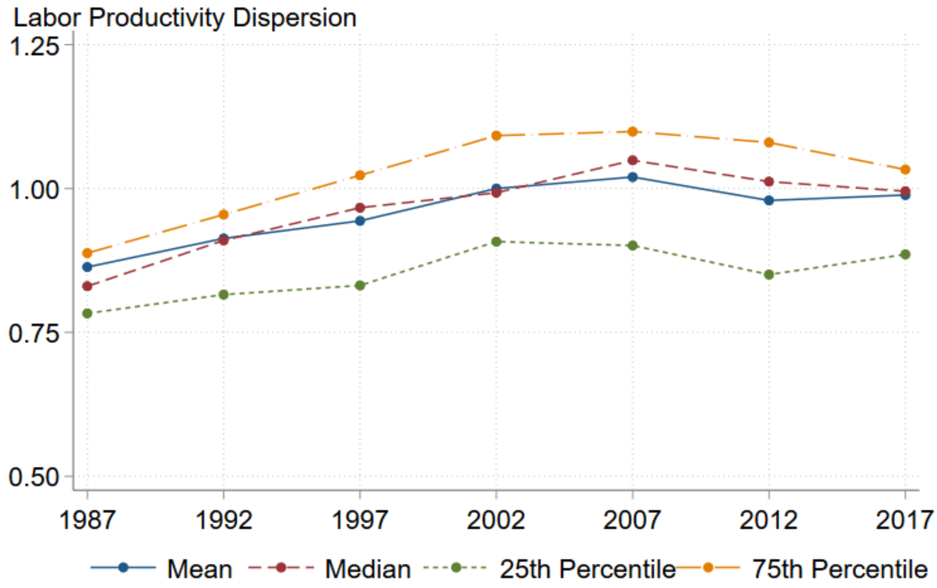
Notes: Lines represent the variation across industries in labor productivity. The lines indicate the 25th percentile, median, 75th percentile, and mean for each statistic. NAICS-4 industries are given equal weight. Establishments within each industry are given equal weight. Calculations use data from the CRT, LBD, and industry-productivity statistics.

Figure 7: Mean IQR by Weighting Methods

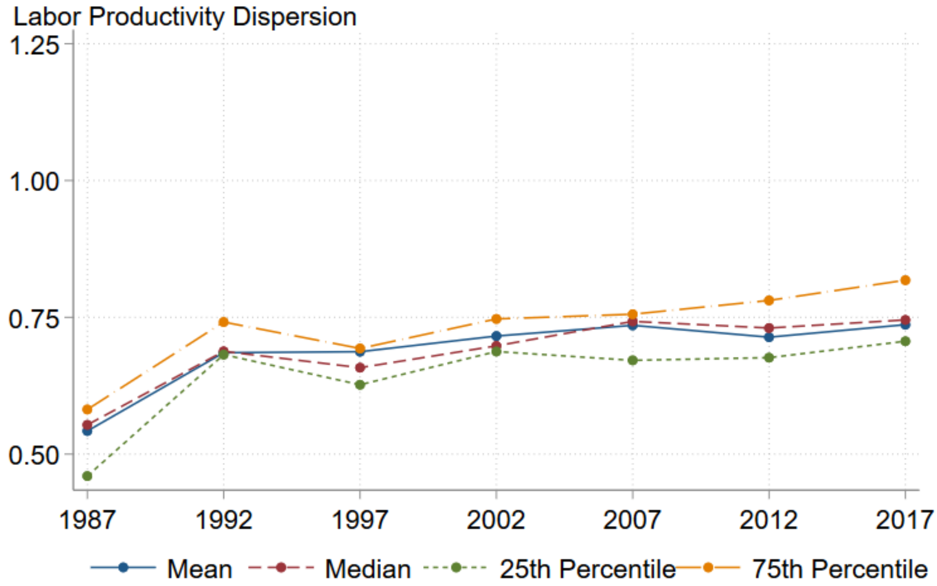


Notes: Lines represent the average IQR across industries with different weighting schemes. Industry weighted combines industries using their share of hours. Unweighted is an unweighted arithmetic average. Industry and establishment weighted calculates the within-industry IQR using each establishment's share of hours within that industry and then aggregates across industries using each industry's share of total hours. *Calculations use data from the CRT, LBD, and industry-productivity statistics.*

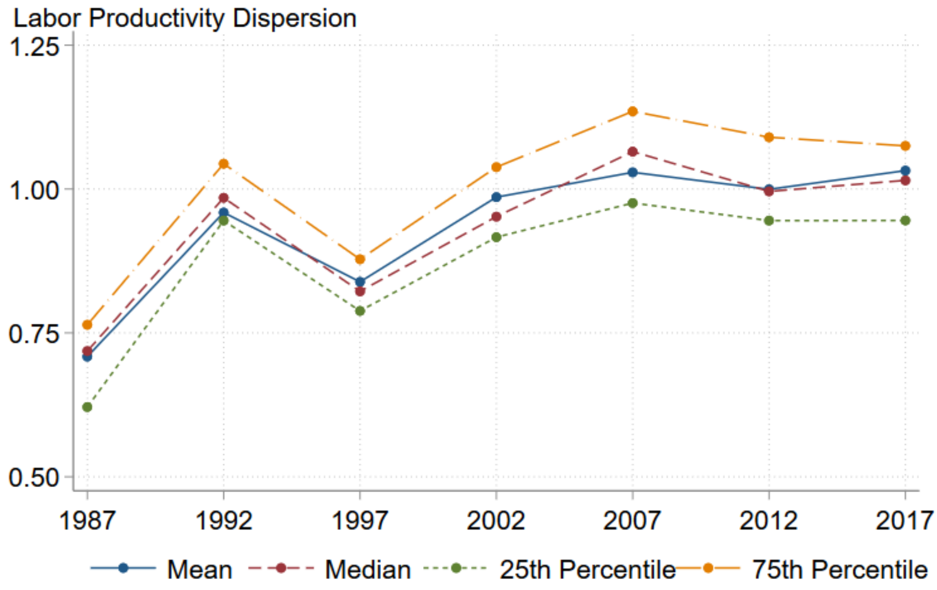
Figure 8: Distribution of Firm Labor Productivity



a) IQR



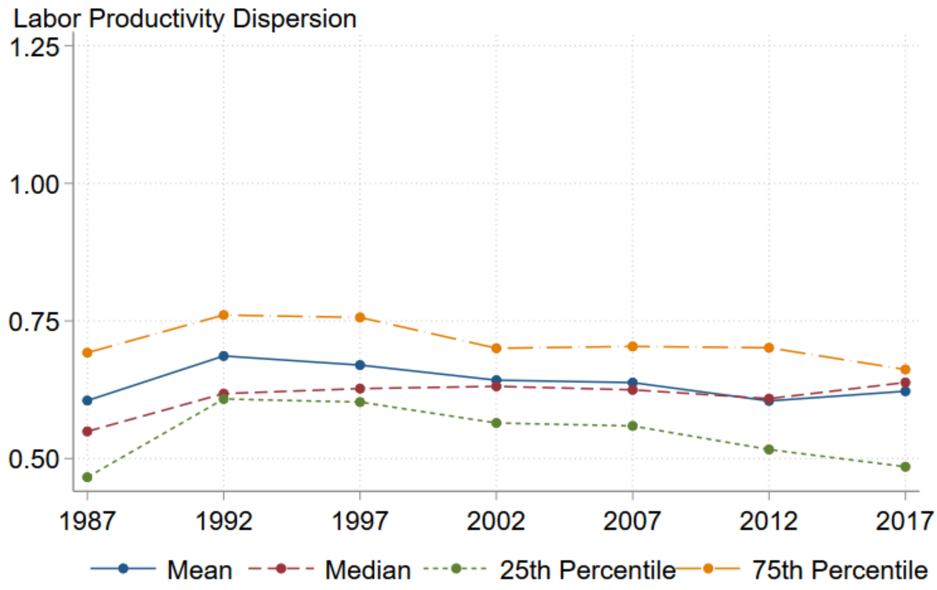
b) 99-90



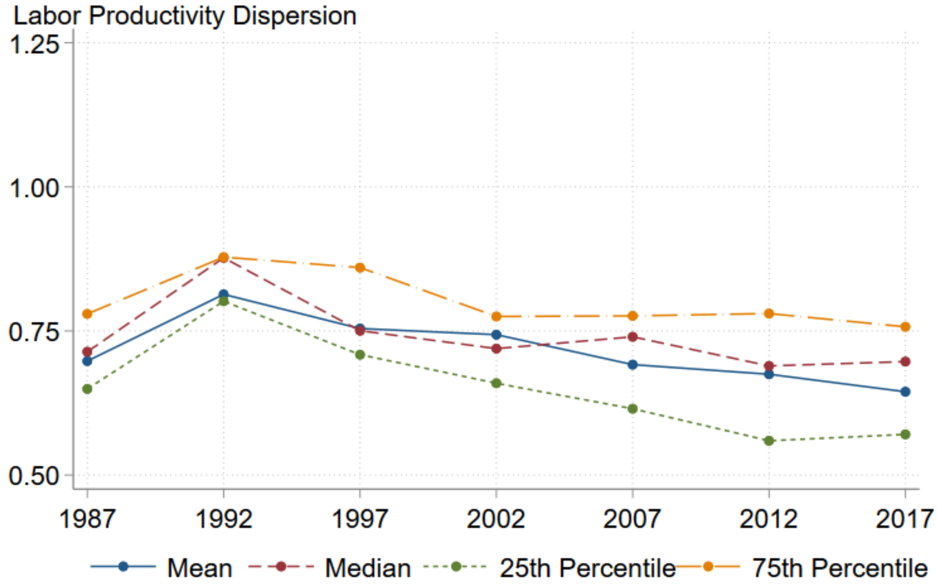
c) 10-1

Notes: Lines represent the variation across industries in the within-industry statistic as noted in the subfigure. The lines indicate the 25th percentile, median, 75th percentile, and mean of the noted statistic. NAICS-4 industries are weighted according to their share of total hours within a year. Firms within each industry are given equal weight. Calculations use data from the CRT, LBD, and industry-productivity statistics.

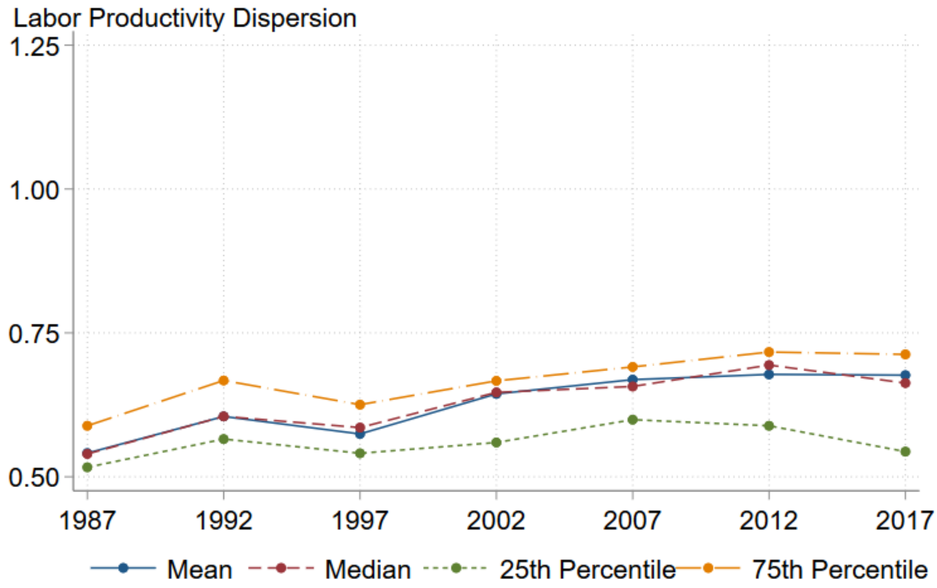
Figure 9: Distribution of Establishment Labor Productivity using Payroll



a) IQR



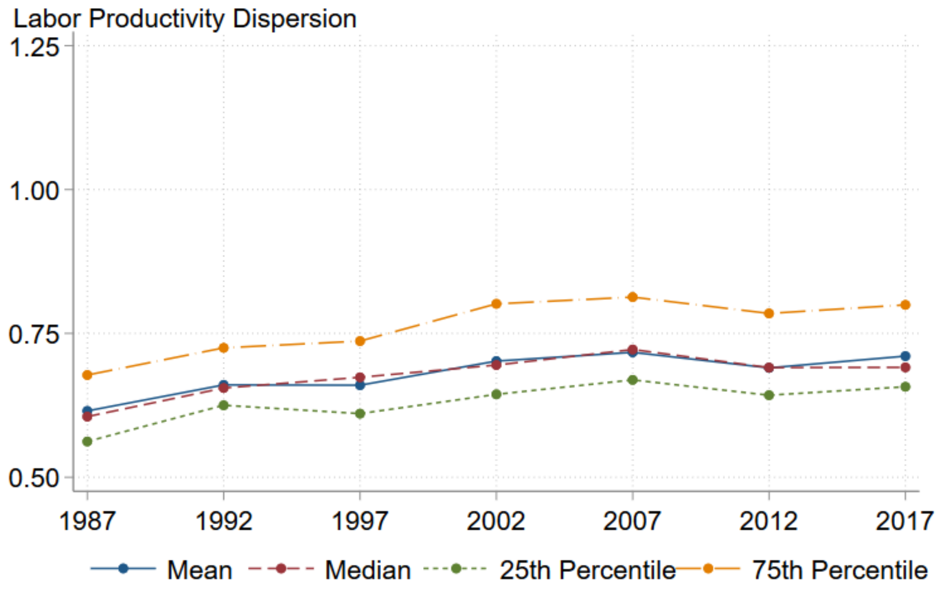
b) 99-90



c) 10-1

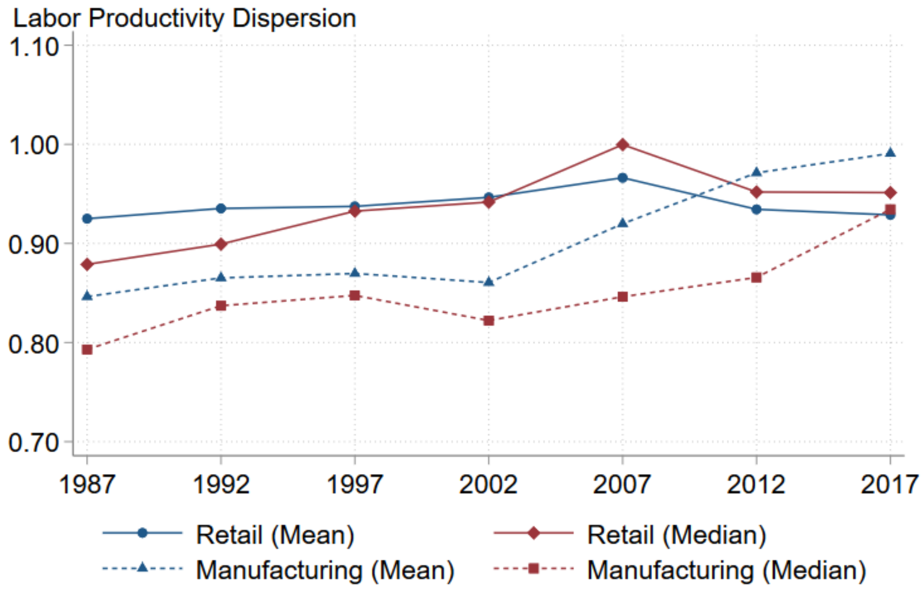
Notes: Lines represent the variation across industries in the statistics as noted for the subfigure. The lines indicate the 25th percentile, median, 75th percentile, and mean within-industry IQR across industries. NAICS-4 industries are weighted according to their share of total hours within a year. Establishments within each industry are given equal weight. Calculations use data from the CRT, LBD, and industry-productivity statistics.

Figure 10: Standard Deviation of Establishment Labor Productivity



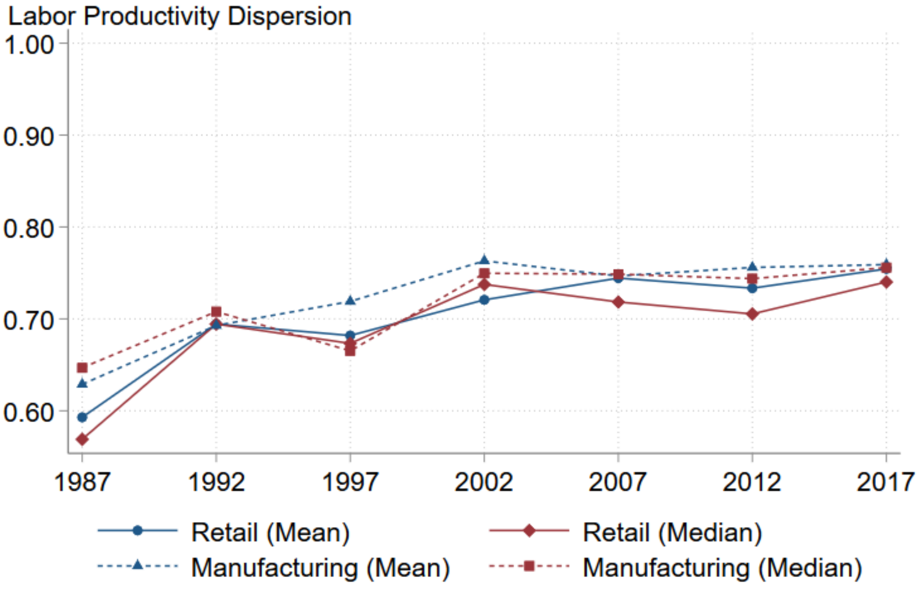
Notes: Lines represent the variation across industries in the within-industry standard deviation. The lines indicate the 25th percentile, median, 75th percentile, and mean within-industry IQR across industries. NAICS-4 industries are weighted according to their share of total hours within a year. Establishments within each industry are given equal weight. Calculations use data from the CRT, LBD, and industry-productivity statistics.

Figure 11 Comparison to Average Manufacturing Dispersion

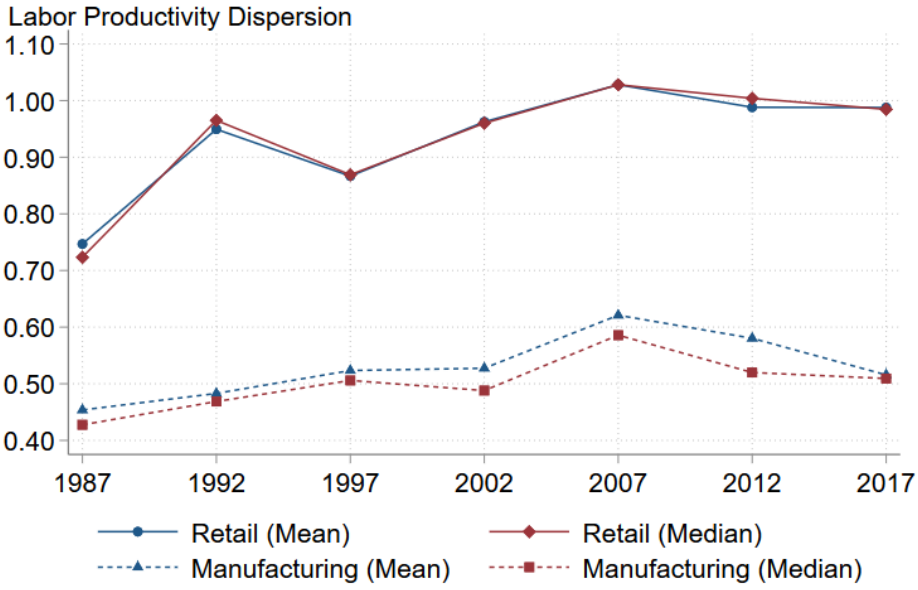


Notes: Lines represent central moments of within-industry IQRs of labor productivity. The lines indicate the median and mean for each statistic and broad sector (manufacturing and retail). NAICS-4 industries are given equal weight within each broad sector. Establishments within each industry are given equal weight. Calculations for retail use data from the CRT, LBD, and industry-productivity statistics. Calculations for manufacturing use the DiSP data product.

Figure 12 Comparison to Tail Dispersion in Manufacturing



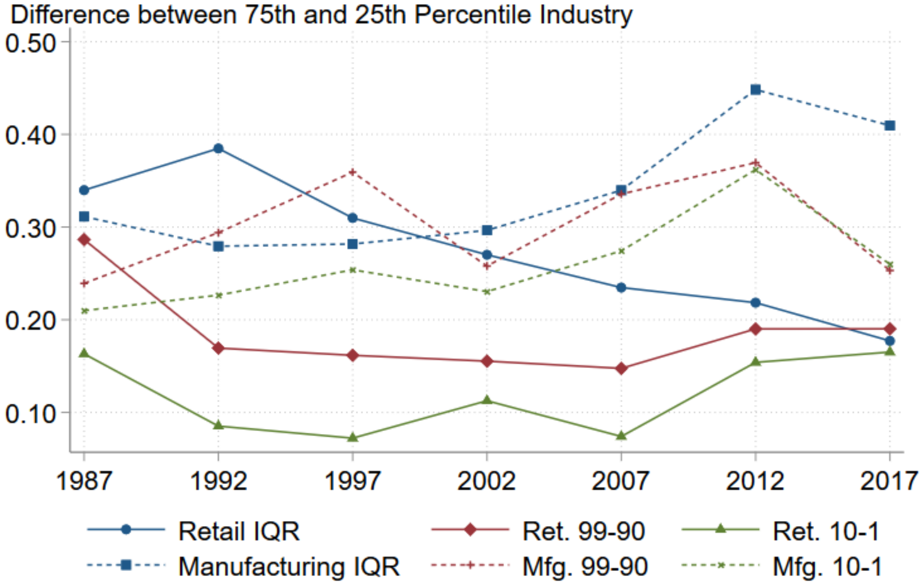
a) 99-90



b) 10-1

Notes: Lines represent the variation across industries in labor productivity. The lines indicate the median and mean for each statistic and broad sector (manufacturing and retail). NAICS-4 industries are given equal weight within each broad sector. Establishments within each industry are given equal weight. Calculations for retail use data from the CRT, LBD, and industry-productivity statistics. Calculations for manufacturing use the DiSP data product.

Figure 13 Comparison of cross-industry dispersion in measures of within industry labor productivity dispersion



Notes: Lines represent the difference between the 75th and 25th percentile industry for each measure within each broad sector (manufacturing and retail). IQR is the interquartile range for within-industry labor productivity. 99-90, and 10-1 are the right and left tail dispersion measures, respectively. NAICS-4 industries are given equal weight within each broad sector. Calculations use data from the CRT, LBD, and industry-productivity statistics. Calculations for manufacturing use the DiSP data product.

Table A.1 – Industry Share of Real Sales (2002)

NAICS	INDUSTRY	2002 SHARE
4411	Automobile dealers	0.200
4471	Gasoline stations	0.156
4451	Grocery stores	0.148
4441	Building material and supplies dealers	0.066
4523	General Merchandise Stores, including Warehouse Clubs and Supercenters	0.066
4461	Health and personal care stores	0.062
4522	Department stores	0.056
4481	Clothing stores	0.033
4541	Electronic shopping and mail-order houses	0.029
4543	Direct selling establishments	0.025
4413	Automotive parts, accessories, and tire stores	0.024
4539	Other miscellaneous store retailers	0.017
4412	Other motor vehicle dealers	0.015
4421	Furniture stores	0.012
4511	Sporting goods, hobby, and musical instrument stores	0.012
4431	Electronics and appliance stores	0.011
4422	Home furnishings stores	0.010
4453	Beer, wine, and liquor stores	0.009
4532	Office supplies, stationery, and gift stores	0.009
4483	Jewelry, luggage, and leather goods stores	0.009
4442	Lawn and garden equipment and supplies stores	0.008
4482	Shoe stores	0.007
4452	Specialty food stores	0.005
4512	Book stores and news dealers	0.005
4542	Vending machine operators	0.003
4533	Used merchandise stores	0.002
4531	Florists	0.002