

BLS WORKING PAPERS



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

National Chains, Online Shopping, and Trends in Retail Productivity Dispersion

Dominic Smith
U.S. Bureau of Labor Statistics

G. Jacob Blackwood
Amherst College

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

Cheryl Grim
U.S. Census Bureau

Jay Stewart
W.E. Upjohn Institute for Employment Research

Zoltan Wolf
BLN24

Working Paper 594
October 2026

National Chains, Online Shopping, and Trends in Retail Productivity Dispersion*

Dominic Smith, U.S. Bureau of Labor Statistics

G. Jacob Blackwood, Amherst College

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

Cheryl Grim, U.S. Census Bureau

Jay Stewart, W.E. Upjohn Institute for Employment Research

Zoltan Wolf, BLN24

October 20, 2026

Abstract

Retail trade was transformed between 1987 and 2017: first, national chains displaced single-store firms; then, activity began to shift to online sellers. We show how this transformation reshaped the labor productivity distribution by developing and analyzing new public-use dispersion statistics from the Census of Retail Trade. Dispersion in the lower tail rose sharply, consistent with a surviving fringe of small, less productive retailers. Dispersion in the middle declined once stores were weighted by labor inputs—consistent with chains expanding by opening large stores that resemble one another. Moreover, industries with larger increases in the share of large firms experienced a bigger decrease in productivity differences between the 90th and 50th percentiles, but not in the tails of the productivity distribution. Our results imply that business models account for a substantial share of productivity differences. Evaluating these patterns in a single-input Hsieh–Klenow framework, we find labor allocative efficiency rises and then falls below its initial level—but the decline is entirely attributable to e-commerce, an industry whose NAICS definition pools disparate production technologies.

Key words: retail, reallocation, business cycles, productivity dispersion, misallocation

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: CBDRB-FY23-CED006-0016, CBDRB-FY25-0421, CBDRB-FY24-0303, CBDRB-FY26-0226). At the time of the writing of this paper, Blackwood was also a part-time employee of the U.S. Census Bureau and Giandrea was an employee of the U.S. Bureau of Labor Statistics. We thank Emek Basker, Cindy Cunningham, Matthew Dey, Lucia Foster, John Haltiwanger, Rachel Nesbit, Sabrina Pabilonia, Kirk White, and audience members at the 2023 FSRDC Conference and U.K. Competition and Markets Authority for their helpful comments. Previous versions of this paper circulated as CES-WP-25-64, CES-WP-23-60 and BLS Working Paper 570.

1 Introduction

Since the 1980s, the U.S. retail sector has transformed from a sector with many small single-store firms to one dominated by large firms with stores in many markets (Foster et al., 2016), with e-commerce contributing to this trend in recent years (Chava et al., 2024). This transformation has resulted in increased concentration, at both the national and local levels, and in consumers from different markets increasingly shopping at the same firms (Smith and Ocampo, 2025). At the same time, the retail sector saw steady and substantial labor productivity growth (US Bureau of Labor Statistics, 1987-2017). The growth of national chains led to the replacement of less-productive single-store firms with more-productive national chain stores (Foster et al., 2006). Looking at the variation in productivity across stores within retail industries can further shed light on this transformation and its implications for allocative efficiency (Hsieh and Klenow, 2009). It is not clear *a priori* whether this transformation would increase or decrease productivity differences across stores.

To investigate this, we use confidential micro data from the Census of Retail Trade to develop new statistics that summarize the firm- and establishment-based labor productivity differences across stores in the 27 retail industries. These dispersion statistics show the productivity differences in stores at different percentiles in the productivity distribution and allow us to get a more complete picture of the retail transformation.¹ We find substantial changes in productivity dispersion in the tails *and* middle of the productivity distribution. Dispersion in the tails of the productivity distribution, both the difference between the 10th and 1st percentiles (the 10–1 range) and the 99th and 90th percentiles (the 99–90 range), increased substantially, with most of the increase occurring between 1987 and 1992. This increase in dispersion, particularly in the upper tail, is consistent with other work that has focused on the importance of “superstar” firms (Autor et al., 2020). However, we find no

¹The dispersion statistics are available as a joint BLS-Census data product and are available at <https://www.bls.gov/productivity/articles-and-research/dispersion-statistics-on-productivity/retail-trade-dispersion.htm> and <https://www.census.gov/programs-surveys/ces/news-and-updates/updates/09302025.html>.

evidence of a relationship between the growth of large firms and changes in the productivity distribution in the upper tail. Where we do find evidence is in the middle. Specifically, industries where the share of large firms increased saw decreases in dispersion in the upper middle (the 75–50 range) of the productivity distribution. Additionally, changes in the share of small firms are positively correlated with dispersion in the same range. Our results suggest that the success of national retailers is due to their ability to open and operate many stores that have above-average productivity but are not necessarily the most productive.

Productivity dispersion patterns are closely related to standard measures of misallocation—less dispersion usually implies a more efficient allocation of inputs (Hsieh and Klenow, 2009; Bils et al., 2021; Blackwood et al., 2021). A natural question then arises from our findings: if industry consolidation leads to less dispersion, is the sector becoming more efficient? The answer is complicated by the fact that, in the middle of the distribution, dispersion is *rising* on an unweighted basis (as in the manufacturing sector), but *falling* when we weight establishments by hours (unlike in manufacturing). Previous literature has emphasized that unweighted statistics are sufficient for measuring misallocation under certain conditions (Hsieh and Klenow, 2009). We explore whether this holds in U.S. retail. We incorporate these data into a model of misallocation that quantifies the impact of these changes on how efficiently labor is allocated across retail establishments. The model features stores using multiple inputs to produce their output with a Cobb-Douglas production function along the lines of Hsieh and Klenow (2009), although we focus on misallocation of labor at the store level. We find that the resulting model-implied allocative efficiency measure rose initially and then fell.

One industry that requires special consideration is NAICS 4541 (Electronic Shopping and Mail-Order Houses) because of variation in the products sold by different retailers. Stores are classified into this industry based on how they sell (no physical store), instead of being grouped based on the products they sell. As a result, these stores have different dispersion patterns than brick-and-mortar retailers. By the end of our sample period, when the industry is large and dominated by e-commerce, this industry has the largest dispersion for most of

the percentile ranges we consider. In fact, the entire decrease in allocative efficiency is caused by the increasing dispersion and weight associated with e-commerce. When we remove e-commerce, the trend in allocative efficiency looks like the inverse of our hours-weighted dispersion measures and rises throughout the period of our study.

We build on Hsieh and Rossi-Hansberg (2023), who show that top service-sector firms grow by replicating their operations across many local markets, by characterizing the productivity of the stores through which that replication happens. We observe the full within-industry productivity distribution for U.S. retail, which lets us ask where in that distribution the growth of national chains appears. We find that growth of national chains primarily impacts the middle of the productivity distribution, but (perhaps surprisingly) not the upper tail. The 75–50 range generally tightened in industries where large firms gained employment share, but dispersion in the upper tail was unrelated to large-firm growth. This differs from the mechanism in Autor et al. (2020) and Gabaix (2011), which emphasized the role of the tails of the firm or productivity distribution, by showing that in retail the firms that grew did not do so by operating the most productive stores. The share of productivity variation occurring within firms rose only from 11.9 to 14.5 percent, while national chains’ share of retail hours rose from 32.0 to 55.8 percent—the pattern to be expected if chains replicate a standardized store, rather than acquiring existing stores or exploiting favorable locations. Rising concentration in retail therefore reflects the diffusion of a replicable operating model rather than the expansion of exceptionally productive establishments, and the two carry different implications for whether further reallocation of activity toward large firms would raise output.

This paper extends the literature asking what dispersion in revenue productivity actually measures (Haltiwanger et al., 2018; Bils et al., 2021; Blackwood et al., 2021) in two directions. First, we show that the mapping from dispersion to allocative efficiency depends on how establishments are weighted: unweighted dispersion in the middle of the retail distribution rises slightly over our sample, while hours-weighted dispersion falls substan-

tially, and the weighted measures track our model-based allocative efficiency more closely. Because published dispersion statistics for manufacturing move similarly whether they are unweighted or weighted (Cunningham et al., 2023), this distinction has not mattered in the past. Second, we show that the industry classification system is itself a source of measured misallocation. Establishments in NAICS 4541 are classified as such based on *how* they sell rather than *what* they sell, and this single industry accounts for the entire decline in measured allocative efficiency in retail between 1987 and 2017; excluding it, allocative efficiency rises throughout. Since allocative efficiency is estimated within industries, any industry that includes establishments with different products and technologies will be estimated to have distortions and low allocative efficiency—a prediction that the 2022 NAICS revision, which reassigns e-commerce establishments to product-defined industries, will allow future work to test directly.

The remainder of the paper describes how we measure productivity and then documents trends in productivity levels and dispersion. Section 2 describes our data and compares our micro-aggregated statistics to published BLS statistics. Section 3 explores different ways of measuring productivity dispersion across stores and firms in the retail sector. Section 4 introduces a model of misallocation, and Section 5 presents the empirical implementation of the model. Section 6 concludes.

2 Measuring Productivity

Our analysis is restricted to labor productivity (LP), because the Census of Retail Trade (US Census Bureau, 1987-2017c) lacks data on capital and other inputs that are needed to calculate TFP. To calculate labor productivity, we need four things: 1) sales, 2) employment, 3) hours per employee, and 4) industry deflators. In brief, we take sales and employment from the census microdata, and we use the same values of hours per employee and industry deflators as the BLS. In the next subsection, we provide more details on our method of

calculating productivity concepts at the micro level, providing points of comparison to official BLS measures (US Bureau of Labor Statistics, 1987-2017).² Then, we take a closer look at comparisons of aggregate measures of productivity in our data to official measures.

2.1 Establishment-Level Productivity Using Census Data

To measure establishment-level LP, we combine restricted-use establishment-level data from the Census Bureau’s Census of Retail Trade (US Census Bureau, 1987-2017c)—henceforth CRT—with public-use industry-level data on prices and hours worked from the BLS (US Bureau of Labor Statistics, 1987-2017). The CRT is conducted every five years (those ending in “2” and “7”) as part of the Economic Census. It includes all retail establishments of multi-store firms and a sample of single-store firms, and it collects data on sales and employment, along with other relevant information. To improve the coverage and representativeness of single-store firms, the Census Bureau uses information from administrative records to impute sales data. We use the Longitudinal Business Database (US Census Bureau, 1987-2017e) to add longitudinal links and time-consistent industry codes at the establishment level (Chow et al., 2021). Our sample consists of all establishments in the Economic Census—CRT and Census of Wholesale Trade (US Census Bureau, 1987-2017d) data files—that are assigned a NAICS code in the retail sector. In the end we have between 950,000 and slightly over one million establishments in our sample in all years, while the number of firms falls from almost 700,000 to 570,000 over the period.

Our goal is to construct an output measure from the microdata that matches the official BLS measure as closely as possible. The underlying data source for our output measure, sales in the Census of Retail Trade, is the same as the one used by the BLS. When industry codes change, the BLS re-estimates sales by industry for earlier years to create a consistent series over time. In contrast, our data have NAICS codes for each establishment in every

²We refer the reader to official BLS documentation for more information on how the official measures are created. See <https://www.bls.gov/opub/hom/opt/home.htm> for a full description of how the BLS calculates industry statistics.

year (Chow et al., 2021).³ These codes are based on more information than is available in the aggregate data. These differences lead our estimate of sales to differ more before the SIC-to-NAICS transition in 2002, and for industries where NAICS codes had changed between 2002 and 2017. We measure labor input as total annual hours worked, which we calculate by multiplying total employment by the average annual hours worked per employee in the most detailed NAICS code available from the Division of Industry Productivity at the BLS. Thus, almost all of the between-establishment variation in hours worked in a four-digit industry is due to variation in employment across establishments in the industry.⁴ For establishment e in industry j , log labor productivity is defined as (time index is dropped for simplicity):

$$LP_{ej} = \ln \left(\frac{\text{real sales}_e}{\text{hours}_j \times \text{employment}_e} \right)$$

One challenge with measuring employment inputs for retail establishments is that employment can be seasonal, while CRT employment is measured for the pay period that includes March 12. In addition, part-time work is common in retail, but the CRT does not differentiate between full-time and part-time employees. Because the number of employees in the CRT in March may not reflect each establishment’s use of labor inputs throughout the year, we also present estimates using real annual payroll (instead of hours) in the denominator in Appendix B.1.⁵ For our weighted statistics, we weight stores or firms by their share of industry hours and weight industries by their share of total retail hours. Store and firm weights are equivalent to employment shares within an industry, because hours are imputed as store or firm employment times an industry-level constant. We collectively refer to store, firm, and industry weights throughout as “hours weights.”

³When we present firm-level statistics we treat each firm-NAICS as a unique entity. This rarely matters because almost all retail firms operate in only one NAICS code.

⁴A very small amount of labor input variation can come from variation in hours worked across six-digit industries within a four-digit industry.

⁵Payroll-based productivity is potentially a more comprehensive measure of labor for establishments with both full- and part-time workers. However, it is important to keep in mind that observed differences in payroll may reflect unobserved differences in occupational mix, overtime, worker quality, and worker-firm effects.

2.2 Retail Trends

Our first step is to document trends and compare our data to the official BLS statistics. Figure 1 documents trends in the retail sector. Panels (a) and (b) compare our micro-aggregated sales, employment, and labor-productivity growth data to the official BLS data. Panels (c) and (d) show trends in employment share by firm size and the nonstore retail share.

Panel (a) compares indexes of sales and employment. Both data sources indicate the retail sector has grown over our sample period. Sales growth closely tracks the BLS aggregate, which is to be expected because both derive from the Economic Census and use the same deflators. Aggregate employment in the two series also tracks closely, even though the two series use different source data. It is worth noting that our employment estimates are 14 percent lower in 1987—possibly because the BLS also accounts for labor at nonemployer establishments—but our series grows more quickly, and the gap falls to less than 5 percent by 2017. Panel (b) shows the implied growth in labor productivity given these inputs. Retail labor productivity grew at an average annual rate of 2.0 percent over our sample period. Our series qualitatively tracks the BLS series and captures the same variation over time, with an increase in the 1990s, a decline in the late 2000s, and a rebound in 2017. Our series exhibits slower productivity growth than the BLS series, reflecting the slower growth in sales and faster growth of employment seen in panel (a). We conclude that our micro-aggregated data are broadly consistent with the official BLS data.

Panel (c) shows the share of employment hours by firms of different sizes. Over time, national firms, those with retail establishments in more than five states, are the only category of firms that increased its share. Single-unit firms, those with only one retail store, have fallen the most—from 44 percent in 1987 to 28 percent in 2017 (at the same time, the share of sales fell from 45 percent to 31 percent). The employment shares of local and regional chains have also shrunk, by about 6 percentage points and 3 percentage points, respectively.

Panel (d) shows the share of retail sales and hours accounted for by nonstore retailers.

By both measures, nonstore retailers were a small part of retail in 1987. However, they experienced rapid growth and by 2017 accounted for almost 10 percent of retail sales. Their employment share remained small, only 4 percent, implying significant productivity growth.⁶

Together these results illustrate the tremendous growth and changing nature of the retail sector between 1987 and 2017. The expansion of employment and output in the sector coincides with the transition to larger chain retailers. At the same time, nonstore retail has taken off, fundamentally changing the retail landscape. As we present our results in the following sections, we discuss what information measures of dispersion and misallocation can provide about this transformation.

3 Labor Productivity Dispersion in Retail

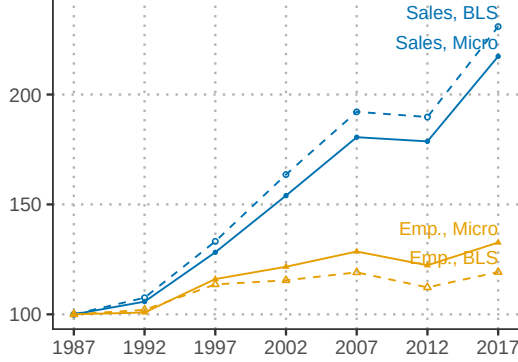
In this section, we explore several aspects of productivity dispersion in retail trade with the goal of shedding light on the transformation that has occurred since 1987. Because this transformation involved a reallocation of economic activity towards large firms and towards industries that sell multiple products, it has created a more homogeneous in-store shopping experience for consumers in different retail markets (Smith and Ocampo, 2025). But the rise of online retail has also fundamentally changed the overall shopping experience, especially in the later part of our sample. In what follows, we trace out how these trends coincide with broader patterns of productivity dispersion in Section 3.1.

Given the increasing importance of large firms, we study what has happened to the tails of the store-level retail productivity distribution in Section 3.2. Previous work has emphasized the roles of the tails of the productivity distribution in understanding aggregate productivity

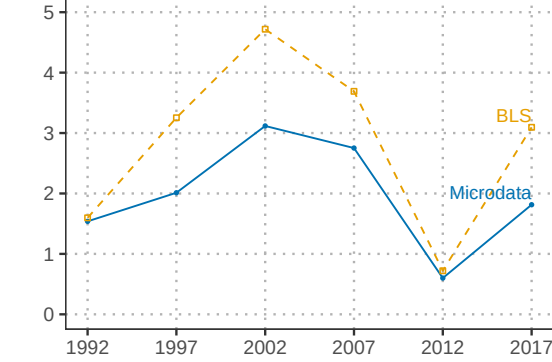
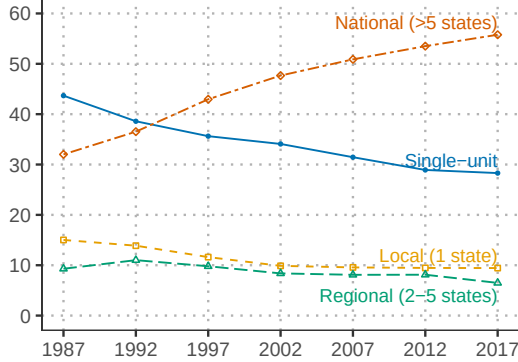
⁶Part of this gap may also be from a measurement artifact rather than pure productivity growth. Warehousing, fulfillment, and delivery labor supporting e-commerce sales is often classified outside NAICS 44–45 (e.g., in Warehousing and Storage or Couriers and Messengers), so our CRT-based hours measure for NAICS 4541 may not capture all of the labor involved in fulfilling online orders. National Academies of Sciences, Engineering, and Medicine (2021) recommend that BLS, in collaboration with BEA and Census, develop a satellite account for an expanded retail sector that would incorporate these supporting activities; see <https://www.bls.gov/productivity/articles-and-research/retail-transformation-measures/retail-trade-supporting-industries/home.htm> for BLS’s ongoing work in this area.

Figure 1: Retail Trends, 1987–2017

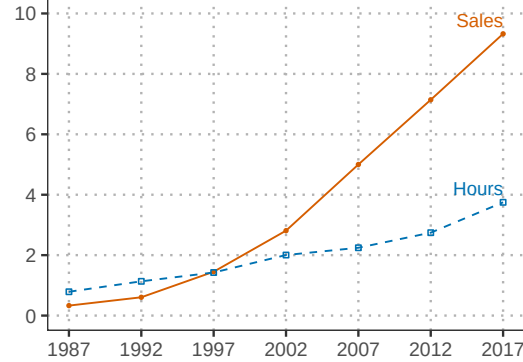
(a) Sales and Employment (Index, 1987 = 100) (b) Labor Productivity Growth (Annualized %)



(c) Employment Share by Firm Size (%)



(d) Non-Store Retail Share (NAICS 4541, %)



Notes: Sample series combine the CRT, Census of Wholesale Trade, and LBD. Panel (a) indexes aggregate real sales and employment to 100 in 1987 for our microdata and for the published BLS series. In 1987, microdata sales are 2.4 percent above BLS and employment is 14 percent below BLS. Panel (b) plots five-year labor productivity growth, computed for each establishment and then averaged, excluding the top and bottom 1 percent of each industry-year productivity distribution; the official BLS measure is the percentage change in published index levels. Panel (c) gives each firm type’s share of retail employment; firm types are single-store firms and multi-store firms operating within one state (Local), two-to-five states (Regional), or more than five states (National). Panel (d) gives NAICS 4541 (Electronic Shopping and Mail-Order Houses) as a share of retail sales and of retail hours. Authors’ calculations from the Census of Retail Trade and BLS industry data.

dynamics (Cunningham et al., 2023). More broadly, research has emphasized the granular nature of the firms in the U.S. economy, with a small number of firms being important for aggregate dynamics (Gabaix, 2011; Autor et al., 2020; Cox et al., 2024; Ma et al., 2025). We explore whether this holds in the retail sector.

After discussing how the movement toward large firms interacts with store productivity dispersion, we investigate the impact on firm-level measures in Section 3.3. We then exploit

time and industry variation in our data to further illuminate the relationship between productivity dispersion and industry consolidation in Section 3.4 with panel regression analysis. Finally, we examine how the e-commerce sector differs from retail in general (3.5).⁷

3.1 Store-Level Dispersion Patterns

Our standard measure of productivity dispersion is the interquartile range (IQR).⁸ Figure 2 shows the IQR, expressed as a multiple, using different weights. Starting with the top line, the unweighted IQR indicates that, on average, an establishment at the 75th percentile is about 150 percent more productive than one at the 25th percentile. On an unweighted basis, the middle of the productivity distribution has remained essentially unchanged over time. The upward trend from 1987 to 2007 was reversed in the 10 years that followed, leaving dispersion in 2017 only slightly above its 1987 level.

However, these unweighted measures miss an important aspect of the retail sector. There is substantial variation in industry size and in the size of stores within an industry. In particular, stores of national chains are on average almost three times larger in terms of employment than single-store firms. Giving those stores equal weight in productivity-dispersion measures risks missing important aspects of reallocation.

Figure 2 also shows the three combinations of store and industry weighting. Weighting by industry share of employment gives substantially more weight to Grocery Stores and General Merchandise Stores, the two largest four-digit industries for the majority of the time we study. This shift lowers the IQR multiple by about 20 percentage points on average and generates a more noticeable upward trend.

The more interesting story arises when stores (establishments) are weighted by hours—the IQR decreases by 10 to 20 percentage points over our sample period.⁹ This result

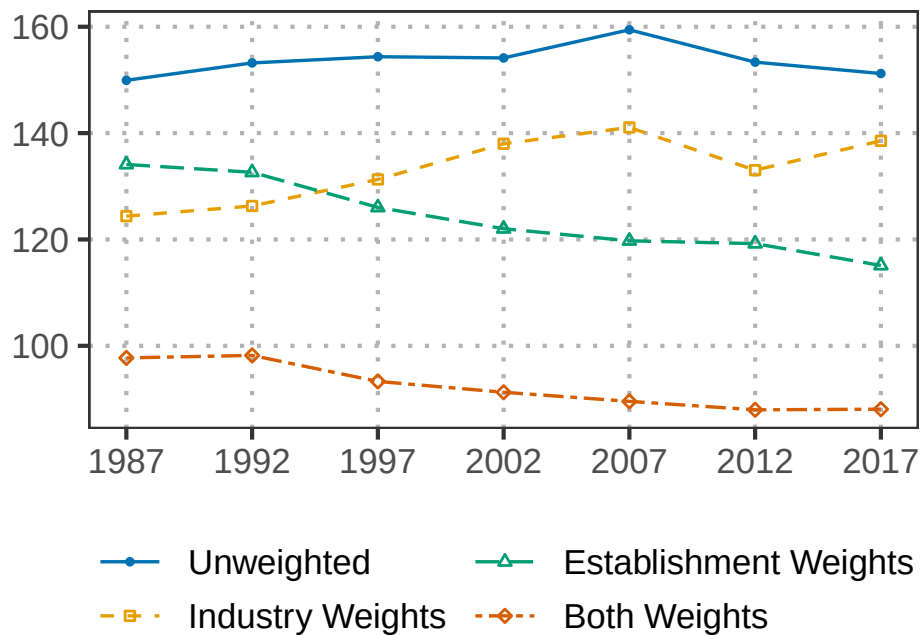
⁷We compare to manufacturing in Appendix A.3.

⁸The interquartile range (IQR) is the difference in log labor productivity between establishments at the 75th and 25th percentiles of the distribution, calculated for each industry-year pair. Other percentile ranges are calculated similarly. All percentiles are calculated using an interval method for disclosure avoidance reasons.

⁹Another way to think about weighted distribution statistics is that they (roughly) measure the dispersion

suggests that the reallocation of economic activity toward large stores, perhaps to those of national chains, has decreased the differences between establishments. However, there is still substantial variation in productivity across stores. Even a 90 percent difference in productivity between the 75th and 25th percentile stores is substantial for stores that are presumably selling similar products at a similarly constructed store. These productivity differences can reflect variation in how busy stores are, the size of the stores, or investments in technologies such as self-checkout machines. They could also reflect variation in prices across geographic areas, as our deflators are national.

Figure 2: Average Store-Level IQR (75–25 Range, % Difference) of Labor Productivity



Notes: Graph lines represent the average ranges (IQR) of the store-level labor productivity distribution across industries when using different weighting schemes. “Unweighted” is a simple arithmetic average across industries, with establishments weighted equally within each industry. “Industry Weights” weight each industry by its share of total retail hours. “Establishment Weights” weight each establishment by its share of hours within its industry. “Both Weights” applies both, weighting establishments by their share of industry hours and industries by their share of total retail hours. Figures are based on authors’ calculations using CRT and BLS data.

Thus, we have two headline results for overall dispersion: 1) it is flat or mildly increasing of worker productivity, in which each worker in a store is assumed to have the same productivity (that of the store).

when stores are weighted equally, and 2) it is notably *decreasing* when stores are weighted by hours. As noted in Hsieh and Klenow (2009), the first measure should be sufficient to characterize misallocation, but only under specific assumptions. If those assumptions are not met in the retail sector, it may be the case that the weighted statistics are more informative. We explore this issue further in Sections 4 and 5.

3.2 Tail Dispersion Measures

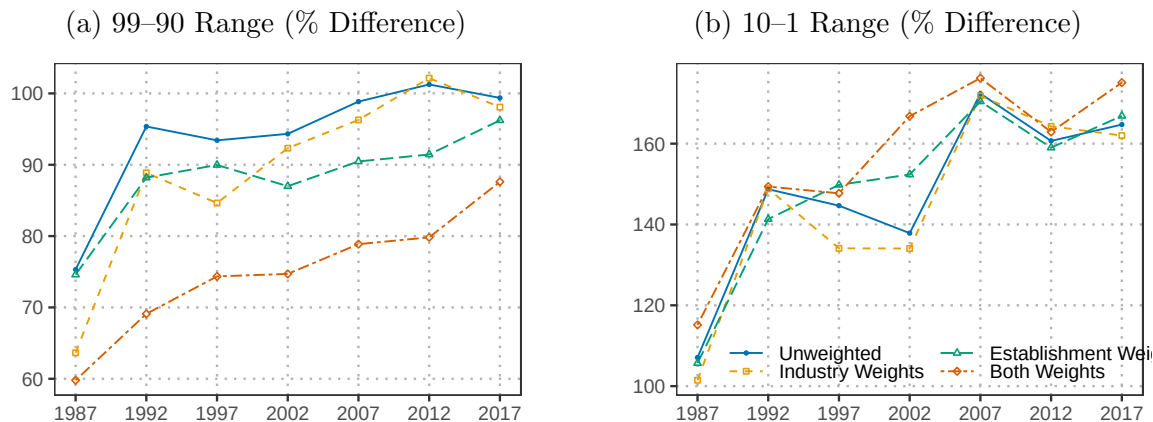
Recent literature has emphasized the importance of influential (“superstar”) firms and establishments on questions related to economic growth and consumer welfare. In the context of misallocation, influential observations can affect measures of allocative efficiency, and so it is worth characterizing the productivity distribution beyond the middle half of the productivity distribution, especially the top. However, it is not clear that large firms in retail are necessarily productivity outliers pushing the frontier of the productivity distribution, as in other industries. First, high productivity in retail does not imply large scale, and many highly productive stores are not particularly replicable due to market size, brand value of exclusivity, or the high-end nature of particular products. Second, retail differs from other sectors in that the largest firms operate hundreds of stores, and it may not be numerically possible for all of these stores to be in the top 10 percent of stores.¹⁰ Even if such firms are outliers in terms of productivity, it is not clear that upper tail dispersion *at the store level* would increase due to the rise of “superstar” retailers. With these considerations in mind, we further investigate the retail-productivity distribution by first looking at the 99–90 (right tail) dispersion measure in the aggregate. We then explore the relationship between tail dispersion and large-firm growth more explicitly in Section 3.4.

Additionally, we also describe trends in the 10–1 (left tail) dispersion measure. Smith et al. (2023) show that single-unit retail firms are disproportionately represented in the

¹⁰We also note that labor inputs are a relatively small proportion of retail costs—we estimate 11 percent—so there is substantial scope for stores with average labor productivity to be highly profitable. The major cost for retailers is cost of goods sold.

lowest productivity quintile, which implies that 10–1 dispersion should be influenced by small firms. To shed light on this segment of the sector, we begin by exploring the 10—1 (left tail) dispersion measure in the aggregate here, turning to its relationship to small-firm activity in more detail in Section 3.4.

Figure 3: Tail Dispersion of Store-Level Labor Productivity



Notes: Graph lines represent the average ranges (99–90 and 10–1) across industries when using different weighting schemes. “Unweighted” is a simple arithmetic average across industries, with establishments weighted equally within each industry. “Industry Weights” weight each industry by its share of total retail hours. “Establishment Weights” weight each establishment by its share of hours within its industry. “Both Weights” applies both, weighting establishments by their share of industry hours, and industries by their share of total retail hours. Figures are based on authors’ calculations using CRT and BLS data.

The right tail (Figure 3a) shows an increase overall between 1987 and 2017, with the timing and pace of this increase depending on weighting choice. For example, the unweighted series rose quickly from 1987 to 1992 before flattening out to 100 for the final 25 years of the study, indicating that, in the average industry, the 99th percentile store is twice as productive as the one at the 90th percentile. When weighting by industry and establishment, the increase was much more smooth but still exhibited a similar upward trend. Weighting by one or the other yields a more uneven series, with a similar end result. We further explore the relationship of this change in dispersion with the shift toward large firms in Section 3.4.

The lower tail (Figure 3b) tells a roughly similar story that is consistent across weighting schemes, with dispersion rising from 100 to over 160 between 1987 and 2017. The timing of the increase is less front-loaded than the 99–90 increase (depending on weights), but the

changes from 1987 to 1992 and 2002 to 2007 account for much of the overall increase across weighting schemes. Overall, the increase in the difference between the 10th and 1st percentile stores was substantial, making it notably larger than the 99–90 difference. By the end of our sample period, unweighted dispersion in the lower 10 percent of the productivity distribution is about the same as in the middle half, likely reflecting significant variation among a large segment of smaller retailers, possibly due to the heterogeneous impacts of the entry of large chains (Jarmin et al., 2009). More specifically, reasons for this change could include increased diversification across small stores that serve specific niches, persistent distortions that keep unproductive stores afloat, and mismeasurement associated with family members who work without pay.

3.3 Firm-Level Dispersion

To better understand the role of ownership consolidation in retail, we shift our focus from store-level to firm-level statistics. Between 1987 and 2017, retail trade evolved from a sector of mostly single-store firms to one dominated by national chains. So far, we have seen evidence of this evolution when we compare store-weighted and unweighted dispersion, because stores operated by national chains tend to be almost three times larger than stores run by single-store firms. The retail-trade sector differs from manufacturing in that firms tend to expand by opening new stores to reach new consumers (Foster et al., 2006). While it may be the case that local market conditions play a large role in store performance, variation in firm practices also matter. Much new technology—such as self-checkout, inventory control systems, etc.—can be relatively easily incorporated into existing stores, and so are likely to be easily replicated across stores within a firm. If these shared technologies and practices are the main drivers of labor productivity, we would expect much of the variation in productivity across establishments to occur as a result of between-firm influences rather than within-firm. In this case, we might expect even larger differences between weighted and unweighted firm-level statistics. Large firms with many similarly productive establish-

ments create concentration points in the weighted firm-productivity distribution, while large firms have almost no influence on the unweighted distribution, because there are more than 500,000 single-store firms and about 1,000 national firms.

Figure 4 compares store-level dispersion (from Figure 2) and firm-level dispersion, both hours-weighted and unweighted. Focusing on Panel (a), both firm-level and store-level IQRs are higher than their respective weighted IQRs. Unweighted firm-level IQRs are higher than unweighted store-level IQRs, as is consistent with overall dispersion being driven primarily by differences across firms. This should not be too surprising, because the existence of large chains that have many similar stores has qualitatively the same effect on store-level dispersion as weighting. The relationship between store- and firm-level IQRs reverses when we weight by hours, with firm-level IQRs being lower. As with the store-level dispersion statistics, the downward slopes of the weighted IQRs reflect the shift from single stores to national chains.

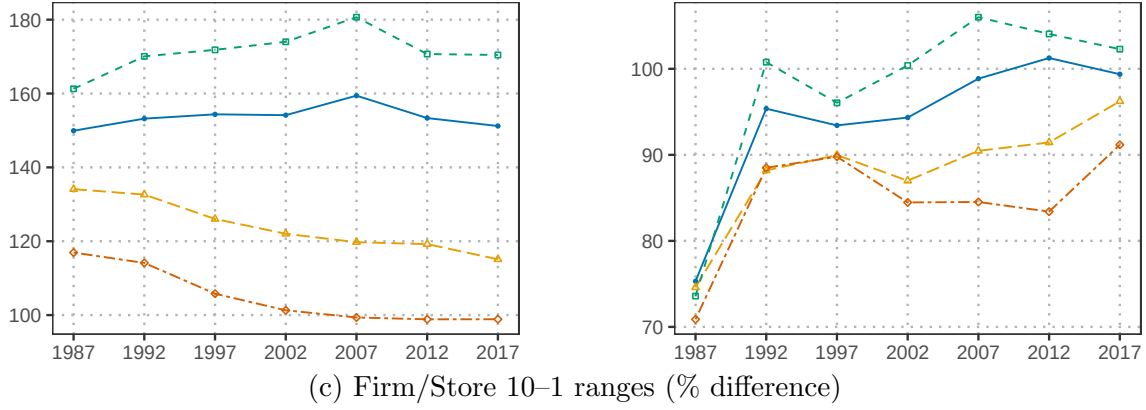
The four 99–90 ranges in Figure 4b are fairly similar to each other. They have the same ordering as the IQRs in Panel (a) and are closer to each other, indicating that there is either less variation in size in the upper tail or that highly productive large and small firms have similar productivities. As before, we see significantly more dispersion in the lower tail than in the upper tail in Figure 4c. Store- and firm-level 10–1 ranges, both weighted and unweighted, are quite similar to each other, suggesting that much of the dispersion is due to productivity differences across single-store firms.

3.3.1 Within-Between Decomposition

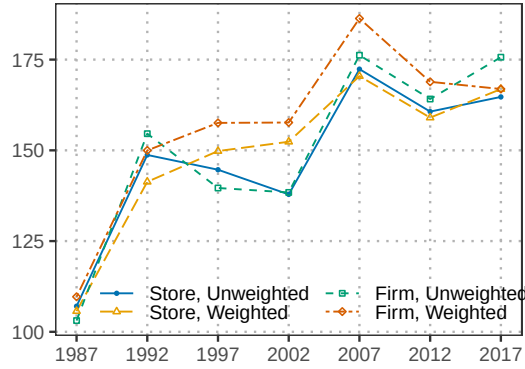
To further explore the role of firms in driving dispersion patterns, we perform a variance decomposition to determine the relative contributions of within-firm and between-firm variance to total variance. The impact of firm consolidation on this decomposition is unclear, *a priori*. For example, if chains were simply acquiring stores and operating them without modification, or if productivity differences were primarily due to geographic location and local

Figure 4: Labor Productivity, Firms vs. Stores

(a) Firm/Store IQRs (75–25 range, % difference) (b) Firm/Store 99–90 Ranges (% difference)



(c) Firm/Store 10–1 ranges (% difference)



Notes: Graph lines represent the average ranges of labor productivity across industries for stores and firms when using different weighting schemes. Unweighted ranges are unweighted arithmetic averages. Establishment- and firm-weighted ranges use each establishment’s or firm’s share of hours within that industry, with industries entering the average equally. Figures are based on authors’ calculations using the Retail DiSP data.

market conditions, we would expect the within-firm share of dispersion to increase. However, if chains replicate practices in new and acquired stores, we would expect the within-firm share of total variance to decrease as weight shifts to large, homogeneous firms with lower variance.

In our data, just 14.5 percent of the variation in productivity was within-firm in 2017—only 2.6 percentage points higher than its 1987 level. This is despite the fact that the share of employment attributed to multi-store firms rose from 56.3 percent to 71.7 percent, and the share of national firms (those with establishments in more than five states) rose from

32.0 percent in 1987 to 55.8 percent in 2017, with most of this share coming from single-store firms.

Overall, despite the shift of retail employment toward large and increasingly geographically dispersed firms, between-firm differences remain the primary driver of variance. While geography and market-specific factors may play a role in explaining heterogeneity among smaller firms, and perhaps between small firms and large chains, this result suggests that large retailers bring their practices—and productivity—with them when they enter a market. We view the relatively mild increase in within-firm productivity dispersion alongside the large shift to chains as corroborating evidence of the narrative that industry consolidation tended to lower dispersion.

3.4 Industry-Level Relationship between Large Firms and Dispersion

In this subsection, we relate changes in dispersion at the industry level to changes in the share of firms of different sizes. We combine the published DiSP industry dispersion measures with public data from the Business Dynamics Statistics (US Census Bureau, 1987-2017a) on employment by firm size. Figure 5 shows the long-run changes (1987–2017) in employment shares against changes in dispersion measures for the 27 four-digit retail industries.

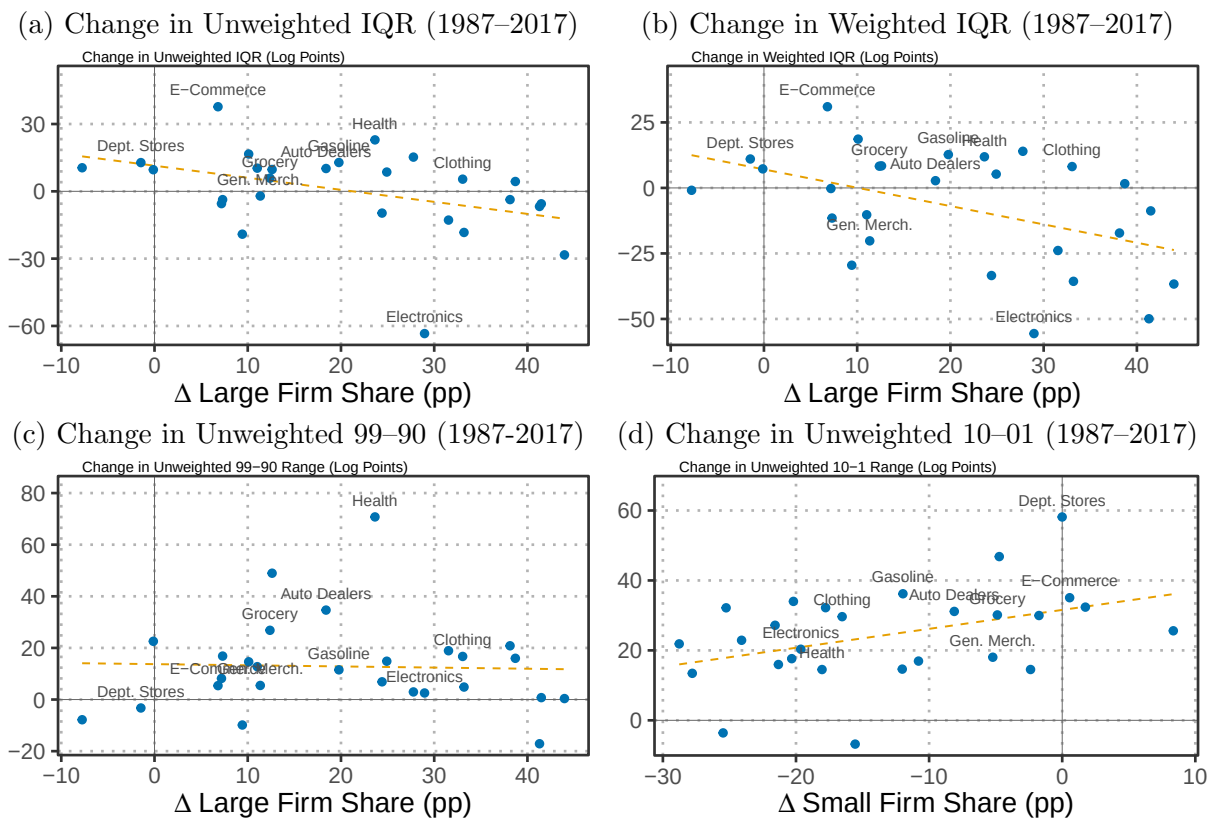
Figures 5a and 5b show the relationship between the percentage-point change in the large-firm (500+ employees) shares of employment and the change in weighted and unweighted IQRs.¹¹ Industries that experienced a large increase in the share of employment in large firms also experienced a large decrease in dispersion in the middle of the productivity distribution, and the relationship is stronger when we weight by employment.

Figure 5c shows the analogous relationship between large firms and dispersion in the

¹¹The BDS defines large firms as those with 500 or more employees, a threshold that captures all national chains plus some local and regional firms. We confirm that the BDS large-firm employment share follows the same time trend as national chain hours in our microdata, running about 10 percentage points higher throughout the period.

upper tail of the productivity distribution. In contrast to the middle of the distribution, industry consolidation appears to have little relationship with right-tail dispersion. And in Figure 5d, we see that industries where the employment share of small firms (fewer than 20 employees) increased saw increased productivity dispersion in the lower tail. This pattern is consistent with the fringe of very small stores driving the left tail: as the small-store sector diversifies into specialized niches, the gap between the tenth and first percentile widens even as chain-driven consolidation compresses the middle of the distribution.

Figure 5: Industry-Level Dispersion and Firm-Size Shares



Notes: Each point represents one four-digit NAICS retail industry ($N = 27$). Panel (a) plots the 1987–2017 change in unweighted IQR against the change in the share of employment in firms with 500 or more employees. Panel (b) plots the 1987–2017 change in hours-weighted IQR against the change in the share of employment in firms with 500 or more employees. Panel (c) plots the change in the upper-tail (99–90) against the change in the share of employment in firms with 500 or more employees. Panel (d) plots the change in lower-tail (10–01) dispersion against the change in the share of employment in firms with fewer than 20 employees. Large- and small-firm employment shares are from the Census Bureau’s Business Dynamics Statistics (public use). Dispersion measures are from the public DiSP data product. Fitted lines are OLS.

The results above are consistent with our hypothesis that the increased presence of large firms is related to shrinking dispersion in the middle of the distribution, and that small firms are responsible for dispersion at the bottom of the productivity distribution. We can now establish that connection more conclusively by studying the relationship between the share of large and small firms and dispersion in the six percentile ranges available in the DiSP statistics. Specifically, we run separate panel regressions for each of the six consecutive DiSP sub-ranges spanning the distribution—99–90, 90–75, 75–50, 50–25, 25–10, and 10–1—on the large- and small-firm employment shares, with industry and year fixed effects.

Table 1 further supports our hypothesis. Panel A shows establishment-level dispersion. For large-firm employment share, the negative and significant coefficients are concentrated in the 90–75 and 75–50 sub-ranges, while the 99–90 upper tail and the 25–10 range show no significant relationship. For small-firm share, in addition to the 90–75 and 75–50 weighted measures being significant, the 10–1 lower tail is significant for unweighted dispersion. This lower tail finding is consistent with small-store fragmentation widening both the center and the very bottom of the distribution. The 99–90 and 25–10 sub-ranges are null throughout.

Panel B repeats this exercise using firm-level dispersion. The pattern finding that large firms impact the upper-middle of the productivity distribution is maintained in the weighted statistics, but unweighted firm-level dispersion shows no significant relationship with either firm-size share. This is consistent with our firm-level results in Section 3.3, because weighting collapses each firm’s stores to a single observation, the firm-level weighted measure treats large firms as occupying substantial mass at the average productivity level of their stores, which is typically the upper-middle of the productivity distribution.

3.5 Dispersion in E-Commerce

The Electronic Shopping and Mail-Order Houses industry (Figure 6) is fundamentally different from other retailers—they have no physical store, and they serve customers over a wider geographic area. Thus, online retail represents a different production technology from

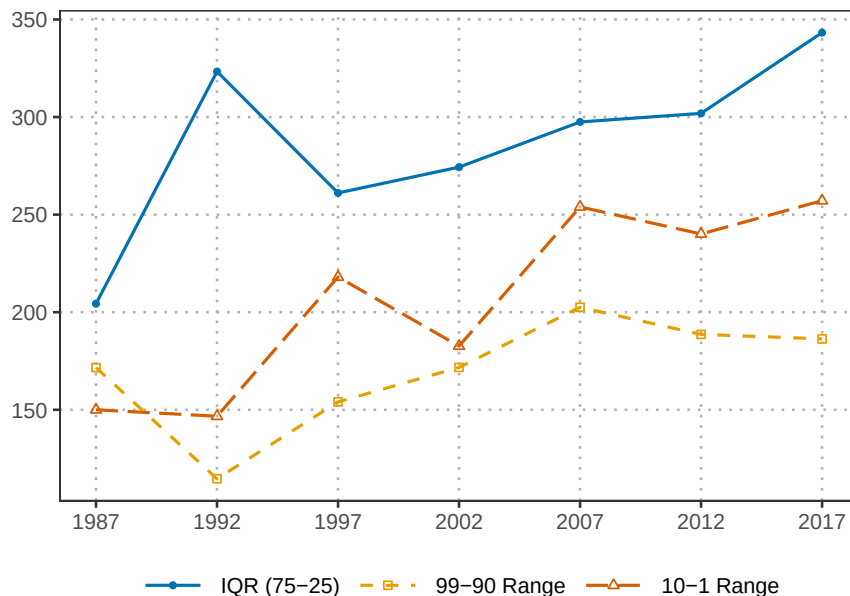
Table 1: Firm-Size Employment Shares and the Productivity Distribution

	DiSP Sub-Range					
	99–90	90–75	75–50	50–25	25–10	10–1
<i>Panel A: Establishment-Level</i>						
Large-firm share (≥ 500 emp)						
Unweighted	−0.180 (0.164)	−0.105 (0.114)	−0.279* (0.146)	−0.125 (0.077)	−0.026 (0.095)	−0.237** (0.104)
Weighted	−0.183 (0.178)	−0.416*** (0.139)	−0.438*** (0.162)	−0.154 (0.137)	0.095 (0.208)	−0.304* (0.172)
Small-firm share (< 20 emp)						
Unweighted	0.255 (0.246)	0.213 (0.158)	0.381* (0.197)	0.095 (0.128)	0.004 (0.150)	0.341** (0.148)
Weighted	0.337 (0.273)	0.615*** (0.206)	0.625*** (0.216)	0.207 (0.171)	−0.208 (0.326)	0.348 (0.286)
<i>Panel B: Firm-Level</i>						
Large-firm share (≥ 500 emp)						
Unweighted	−0.132 (0.169)	−0.081 (0.087)	−0.163 (0.135)	−0.078 (0.114)	−0.027 (0.074)	−0.092 (0.092)
Weighted	−0.232 (0.169)	−0.280 (0.221)	−0.617** (0.274)	−0.299 (0.238)	0.015 (0.262)	−0.232 (0.262)
Small-firm share (< 20 emp)						
Unweighted	0.214 (0.218)	0.182 (0.131)	0.244 (0.178)	0.051 (0.164)	0.026 (0.111)	0.172 (0.143)
Weighted	0.398* (0.232)	0.475 (0.326)	0.933*** (0.347)	0.309 (0.328)	0.014 (0.405)	0.250 (0.405)
Observations	189 (Panel A) / 175–186 (Panel B); varies by disclosure suppression					
Industries (G)	27 (Panel A) / 26–27 (Panel B)					
Fixed effects	Industry + Year					

Notes: Each cell reports the coefficient from a separate two-way fixed-effects panel regression of the DiSP subrange dispersion measure (in log points) on the indicated firm-size employment share, using up to 27 four-digit NAICS retail industries over the seven census years 1987–2017. Panel A uses establishment-level dispersion; Panel B uses firm-level dispersion. “Unweighted” and “Weighted” refer to whether establishments (Panel A) or firms (Panel B) are weighted by their share of industry hours (activity-weighted in DiSP notation). Firm-size employment shares are from Census BDS public-use data. Cluster-robust standard errors (clustered by NAICS four-digit industry) in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

brick-and-mortar retail, and since NAICS categories were primarily meant to capture differences in production rather than product, online retailers have been categorized together (in NAICS 4541) regardless of the products they sell. Despite the fact that nonstore retailers generally take orders and fulfill them by shipping products in boxes, there may be significant productivity variation across these retailers, depending on what types of products they sell.¹² And given the dramatic growth in industry sales combined with the transformation of this industry from catalog to online sales, it seems likely that some of the increase in dispersion is due to differences in productivity between entering and incumbent establishments. There may also be variation in the tendency of these retailers to outsource sales-related activities such as delivery.

Figure 6: Dispersion in Electronic Shopping and Mail-Order Houses (%)



Notes: Lines represent the unweighted difference between establishments at different percentiles of the within-industry labor productivity distribution in each year. The y-axis is expressed in percentage differences. Based on authors' calculations using CRT and BLS data.

Perhaps unsurprisingly, then, there was an increase in all three dispersion measures. The

¹²For this reason, the NAICS has been updated in 2022 to classify online retailers with producers of similar products. This rectifies the heterogeneity in product markets but opens the door for within-industry heterogeneity in production technologies—another source of productivity dispersion.

IQR multiple increased from about 200 percent in 1987 to almost 350 percent in 2017. In addition to the substantial change over time, the level of dispersion is generally much higher than in other industries. It seems plausible that much of the dispersion in this industry is due to differences in the product categories that are sold by the different retailers. Our data do not allow us to calculate gross margins (revenue minus cost of goods sold), which would allow us to account for differences in the types of goods sold. Prices and gross margins are likely to be more similar for brick-and-mortar stores in the same four-digit industry. But that is not the case for e-commerce. Given the many potential contributing factors to the substantial dispersion in this industry, we return to its influence on overall measures of allocative efficiency in Section 5.

4 A Model of Labor Misallocation

A common theme from the literature is that productivity dispersion measures are useful—sometimes even sufficient—statistics for characterizing misallocation (Hsieh and Klenow, 2009; Bils et al., 2021; Blackwood et al., 2021). However, what are the implications for misallocation when alternative dispersion statistics display different patterns? Previous literature gives theoretical conditions under which unweighted dispersion is a sufficient statistic, but those conditions are unlikely to hold in this context. We shed light on this issue by specifying a model of allocative efficiency that maps onto our data. The model is similar to the appendix of Hsieh and Klenow (2009), but with a few modifications. We derive a measure of allocative efficiency and then empirically implement it using our data.

Critically, our data do not permit measurement of other inputs. This does not require us to write down a single-input model—we incorporate the relative weight of labor in production through its elasticity—but it does require us to treat other inputs as fixed. So the measure of allocative efficiency is one in which the distribution of other inputs is held constant.¹³

¹³A potential productivity-enhancing change could be to reallocate nonlabor inputs, rather than reallocating labor.

Thus, the focus is on how misallocated labor is *given the distribution of other inputs*.

4.1 Aggregate and Sectoral Production

Following Hsieh and Klenow (2009), we begin with total retail output, which we will refer to as aggregate output, which is a Cobb-Douglas across subsectors s (which we will refer to as sectors):

$$Y = \prod_s Q_s^{\theta_s}. \quad (1)$$

Production is overseen by a single aggregate firm that maximizes profit $\pi = Y - \sum_s P_s Q_s$, where the aggregate price, P , is normalized to 1. The first order conditions imply sectoral revenue is a share, θ_s , of total output: $R_s \equiv P_s Q_s = \theta_s Y$. We then specify that sectoral output is a CES composite of output across stores, indexed by i :

$$Q_s = \left(\sum_i q_{is}^{\rho_s} \right)^{\frac{1}{\rho_s}} \quad (2)$$

Similar to aggregate production, there is a single within-industry aggregator firm that maximizes profit: $\pi_s = P_s Q_s - \sum_i p_{is} q_{is}$. First-order conditions imply that

$$p_{is} = P_s \left(\frac{Q_s}{q_{is}} \right)^{1-\rho_s}. \quad (3)$$

For plant-level production, we would ideally like to specify a general production function that is Cobb-Douglas:

$$q_{is} = Z_{is} \prod_j X_{ijs}^{\alpha_{js}},$$

where $\sum_j \alpha_{js} = 1$. We do not have data on inputs other than labor, so we define the composite of technical efficiency Z_{is} and inputs other than labor X_{ijs} as our effective productivity term $A_{is} = Z_{is} \prod_{j \neq L} X_{ijs}^{\alpha_{js}}$. Our reduced-form production function is then given by

$$q_{is} = A_{is} L_{is}^{\alpha_{L_s}}. \quad (4)$$

Profit maximization (along the labor input margin) is subject to a distortion, τ_{is} , and is given by

$$\pi_{is} = p_{is}q_{is} - (1 + \tau_{is}) w_s L_{is}. \quad (5)$$

The resulting first-order condition implies that

$$\rho_s \alpha_{Ls} p_{is} q_{is} = (1 + \tau_{is}) w_s L_{is}. \quad (6)$$

Denoting revenue as $R_{is} = p_{is}q_{is}$ and re-arranging the FOC, we can show that within-sector variation in labor productivity is due solely to distortions when input prices, demand elasticities, and production technologies are the same across plants.¹⁴

$$LP_{is} = \frac{R_{is}}{L_{is}} = \frac{w_s}{\rho_s \alpha_{Ls}} (1 + \tau_{is}). \quad (7)$$

4.2 Allocative Efficiency

A key object for allocative efficiency in this context is sectoral labor productivity, defined as sectoral output Q_s divided by sectoral labor input L_s :

$$LP_s = \frac{Q_s}{L_s}. \quad (8)$$

We show in appendix C.2 that labor productivity can be expressed as

$$LP_s = \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s \alpha_{Ls}}} \left(\frac{1+\tau_{is}}{1+\tau_s} \right)^{\frac{-\rho_s \alpha_{Ls}}{1-\rho_s \alpha_{Ls}}} \right)^{\frac{1-\rho_s \alpha_{Ls}}{\rho_s}}}{L_s^{1-\alpha_{Ls}}} \quad (9)$$

where $1 + \tau_s \equiv \left(\sum_i \frac{R_{is}}{R_s} \frac{1}{1+\tau_{is}} \right)^{-1}$ is a sectoral distortion, $L_s = \sum_i L_{is}$ and $R_s = \sum_i R_{is}$.

Allocative efficiency is defined as the ratio of labor productivity (with distortions) to

¹⁴See Blackwood et al. (2026) for further discussion about the effect of relaxing these assumptions.

counterfactual productivity (when all distortions are equalized):

$$AE_s = \frac{LP_s}{LP_s^*} = \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s\alpha_{Ls}}} \left(\frac{1+\tau_{is}}{1+\tau_s} \right)^{\frac{-\rho_s\alpha_{Ls}}{1-\rho_s\alpha_{Ls}}} \right)^{\frac{1-\rho_s\alpha_{Ls}}{\rho_s}}}{LP_s^* L_s^{1-\alpha_{Ls}}}, \quad (10)$$

where $LP_s^* \equiv \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s\alpha_{Ls}}} \right)^{\frac{1-\rho_s\alpha_{Ls}}{\rho_s}}}{(L_s^*)^{1-\alpha_{Ls}}}$, and L_s^* is the level of labor in the undistorted counterfactual. We can then define $\tilde{A}_s = \left(\frac{1}{N_s} \sum_i A_{is}^{\frac{\rho_s}{1-\rho_s\alpha_{Ls}}} \right)^{\frac{1-\rho_s\alpha_{Ls}}{\rho_s}}$, where N_s is the number of stores in the sector, to rewrite allocative efficiency in a more familiar form, similar to Blackwood et al. (2021):

$$AE_s = \underbrace{\left(\frac{1}{N_s} \sum_i \left(\frac{A_{is}}{\tilde{A}_s} \right)^{\frac{\rho_s}{1-\rho_s\alpha_{Ls}}} \left(\frac{1+\tau_{is}}{1+\tau_s} \right)^{\frac{-\rho_s\alpha_{Ls}}{1-\rho_s\alpha_{Ls}}} \right)^{\frac{1-\rho_s\alpha_{Ls}}{\rho_s}}}_{AE_s^{cov}} \underbrace{\left(\frac{L_s^*}{L_s} \right)^{1-\alpha_{Ls}}}_{\text{Sector Input Term}} \quad (11)$$

We focus on AE_s^{cov} , which captures how the covariance structure of distortions and “effective productivity” A_{is} affects efficiency. Intuitively, the more highly correlated distortions and A_{is} are, the less efficient the economy will be. Since the sectoral input term is not dependent on within-industry distribution of distortions, but only the sectoral level distortions τ_s , the term cancels. In other words, τ_s captures cross-sector misallocation, while the distribution of τ_{is} captures within-industry misallocation. Furthermore, since our focus is on within-industry, the distribution of labor across sectors is invariant to the within-industry distribution of distortions if aggregate labor is in fixed supply. We show in Appendix C.1 that sectoral input demand

$$L_s = L \frac{\rho_s \alpha_{Ls} \theta_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}}, \quad (12)$$

where $L = \sum_s L_s$, and thus is solely dependent on sectoral distortions.

4.3 Aggregate AE

Define aggregate labor productivity as $LP \equiv \frac{Q}{L}$. Then, using our expression for L from before and the definition of Q , we can write aggregate labor productivity as

$$LP = \left(\prod_s LP_s^{\theta_s} \right) \underbrace{\prod_s \left(\frac{\rho_s \alpha_{Ls} \theta_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}} \right)^{\theta_s}}_{\bar{T}} = \bar{T} \times \prod_s LP_s^{\theta_s}. \quad (13)$$

Allocative efficiency at the aggregate level is then the ratio of this value to the efficient level.

$$AE = \frac{LP}{LP^*} = \frac{\bar{T} \times \prod_s LP_s^{\theta_s}}{\bar{T} \times \prod_s (LP_s^*)^{\theta_s}} = \prod_s \left(\frac{LP_s}{LP_s^*} \right)^{\theta_s} = \prod_s AE_s^{\theta_s}. \quad (14)$$

Note that since labor allocation across sectors is invariant to within-industry redistribution and sectoral distortions are unchanged, the $\bar{T}$ term is also unchanged.

4.4 Empirical Implementation

To calculate AE , we need the components of labor productivity at the store level, which we have described previously, and which chiefly include revenue and labor input data. Additionally, we require parameter values for ρ_s , α_s , and θ_s . We obtain θ_s from revenue shares within the CRT (US Census Bureau, 1987-2017c). For the demand parameter ρ_s and labor elasticity α_s for each subsector s , we consider several alternatives. Our primary approaches involve calibrating α_s from the average cost-share of labor in total costs in subsector s , and ρ_s from markups over costs in the subsector to publicly available Annual Retail Trade Survey (US Census Bureau, 1987-2017b) data. Our baseline is to consider aggregate parameters that are constant over time and across industries based on overall sectoral labor shares and markups. We then allow for time-varying parameters that are constant across industries, and then consider industry-specific parameters that are fixed to 2007 levels. Fully industry-time-varying parameters are not feasible due to data limitations. Finally, we also consider

alternatives to the baseline in which α_s is set to 0.33 for all industries, and one where ρ_s is set to 0.75 for all industries, both of which are common calibration values in the literature.

With these in hand, we are now ready to find our key empirical analogues to the theoretical concepts. Note that in the model:

$$(1 + \tau_{is}) = \frac{\rho_s \alpha_{Ls}}{w_s} \frac{R_{is}}{L_{is}}, \quad (15)$$

and

$$(1 + \tau_s) = \frac{\rho_s \alpha_{Ls}}{w_s} \frac{R_s}{L_s}. \quad (16)$$

So, when the ratio $\frac{1+\tau_{is}}{1+\tau_s}$ is taken, the parameter values and prices cancel out.

Let $\widehat{LPR}_{is} = \frac{\widehat{R}_{is}}{\widehat{L}_{is}}$, where $\widehat{X}_{is}$ is the empirical equivalent of the plant-level concept X_{is} in the data. Let $\widehat{LPR}_s = \frac{\widehat{R}_s}{\widehat{L}_s}$ denote the same empirical notation for sectoral aggregates ($\widehat{L}_s = \sum_i \widehat{L}_{is}$, e.g.). Then, we can say

$$\frac{1 + \widehat{\tau}_{is}}{1 + \widehat{\tau}_s} = \frac{\widehat{LPR}_{is}}{\widehat{LPR}_s}. \quad (17)$$

Since A_{is} and A_s are proportional to the same industry-level constant, the following equivalencies hold:

$$\widehat{LPQ}_{is} = \frac{\widehat{R}_{is}^{\frac{1}{\rho_s}}}{\widehat{L}_{is}^{\alpha_{Ls}}}, \quad \widehat{LPQ}_s = \left(\sum_i \left(\frac{\widehat{R}_{is}^{\frac{1}{\rho_s}}}{\widehat{L}_{is}^{\alpha_{Ls}}} \right)^{\frac{\rho_s}{1-\rho_s \alpha_{Ls}}} \right)^{\frac{1-\rho_s \alpha_{Ls}}{\rho_s}}. \quad (18)$$

We can then write the empirical estimate of allocative efficiency as follows:

$$\widehat{AE}_s = \left(\sum_i \left(\frac{\widehat{LPQ}_{is}}{\widehat{LPQ}_s} \right)^{\frac{\rho_s}{1-\rho_s \alpha_{Ls}}} \left(\frac{\widehat{LPR}_{is}}{\widehat{LPR}_s} \right)^{\frac{-\rho_s \alpha_{Ls}}{1-\rho_s \alpha_{Ls}}} \right)^{\frac{1-\rho_s \alpha_{Ls}}{\rho_s}}. \quad (19)$$

Finally, we can obtain an aggregate estimate:

$$\widehat{AE} = \prod_s \widehat{AE}_s^{\theta_s}. \quad (20)$$

A key counterfactual we consider is the effect of “removing” a particular industry’s contribution to allocative efficiency. In reality, we do not remove the industry entirely, since it still attracts its share of aggregate labor (which is in part determined by its sectoral distortion) and maintains its revenue share, but we remove plant-level distortions from that industry only. Thus, if industry v is “removed” in this fashion, we have

$$\widehat{AE}_{-v} = \prod_{s \neq v} \widehat{AE}_s^{\theta_s}. \quad (21)$$

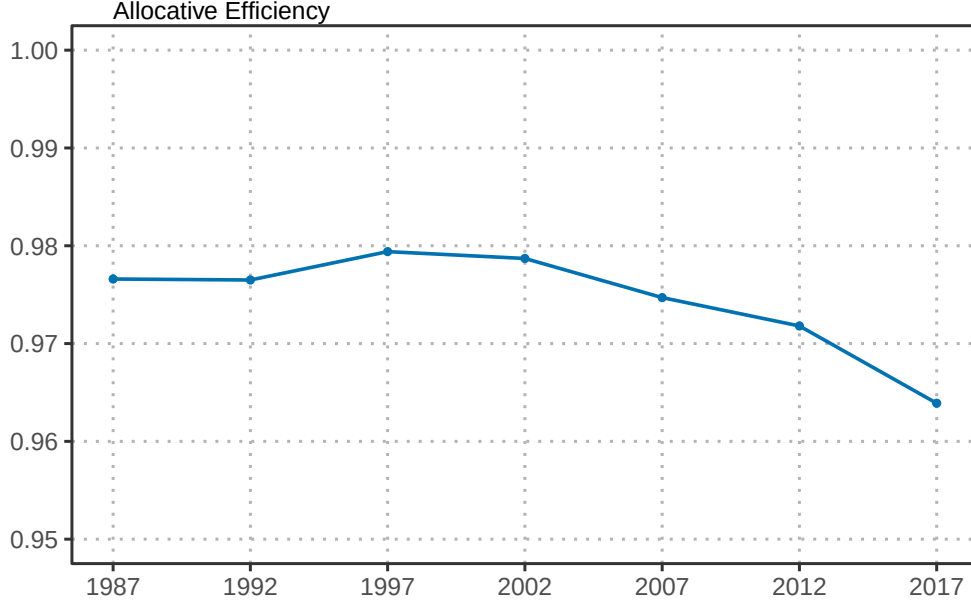
In particular, we consider this measure when the effect of mail-order and e-commerce retailers is eliminated.

5 Allocative Efficiency Results

Figure 7 presents our main results. Labor allocative efficiency is generally above 96 percent. This is high, relative to the allocative efficiency measures based on revenue TFP, due to the fact that our framework takes into account that the influence of labor is relatively small. So, the potential gains from reallocating labor are muted, relative to the gains from reallocating all inputs, which typical AE measures capture.

We also find that allocative efficiency has a hump-shaped pattern, initially rising and then falling over the sample period considered. These results are robust to parameterization choice, as we explore in Appendix B.2. At first glance, then, it is difficult to say which dispersion measure best captures allocative efficiency. To correspond more closely with the early empirical results, we compare to industry-weighted statistics, since our measure of AE also weights by industry. The rise in allocative efficiency early on is consistent with the

Figure 7: Labor Allocative Efficiency in Retail



Notes: Labor allocative efficiency for the retail sector, computed as described in Section 4 under the baseline parameterization (aggregate parameters constant across industries and years). Authors’ calculations from the CRT microdata.

steady decline in weighted statistics, while the sharp decline is more closely aligned with the unweighted statistics (although the time series pattern for unweighted dispersion is more uneven in the last part of the sample).

However, we make two observations here that qualify these results. The first is that allocative efficiency is not solely a function of revenue productivity dispersion, but also of its relationship with “effective productivity” A_{is} . Blackwood et al. (2021) show that the positive correlation between effective productivity and revenue productivity is a substantial contributor to allocative efficiency in the manufacturing context, and that is likely the case in retail as well. Additionally, A_{is} captures variation in other inputs like capital and intermediate inputs. Given the importance of physical space and intermediates for the sector, these likely contribute substantially to variation in our A_{is} measure. Furthermore, previous literature has discussed how mechanisms like differential demand, adjustment costs, and production-function heterogeneity can generate measured differences in A_{is} and τ_{is} (Blackwood et al.,

2026).

Second, time-series patterns in allocative efficiency could be sensitive to structural change—shifts in cross-industry revenue weights—that the dispersion statistics do not pick up. The geometric mean is more sensitive to outliers than arithmetic averages, which can give individual industries more influence. For example, Blackwood et al. (2026) show the influence of three industries in explaining more than half of the decline in measured allocative efficiency in U.S. manufacturing. We focus on e-commerce in this context in the next section.

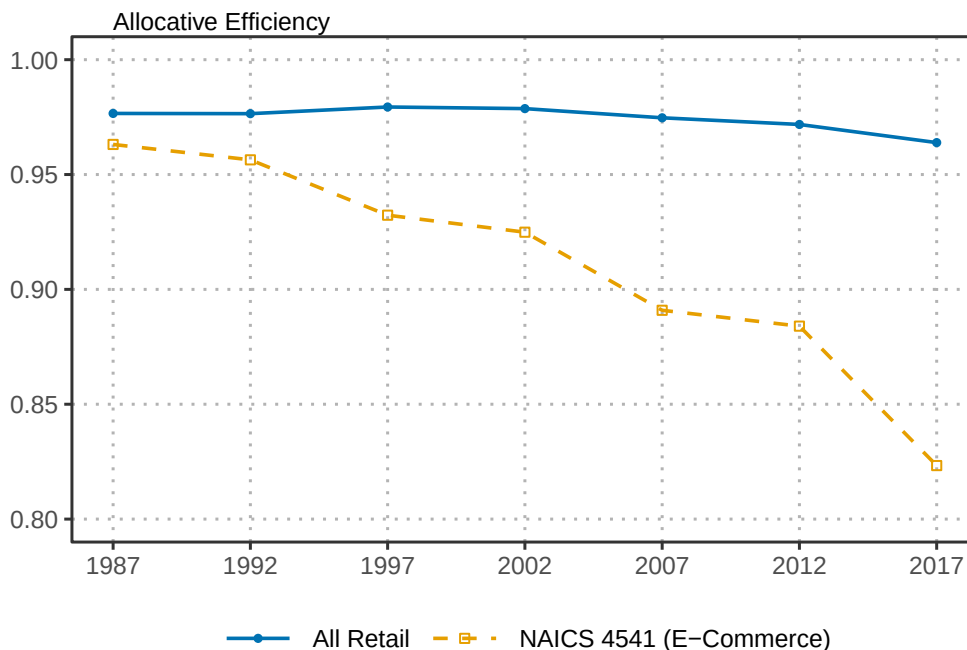
5.1 The Role of E-commerce in Allocative Efficiency

The revenue share for Electronic Shopping and Mail-Order Houses (NAICS 4541) rose from 0.3 percent in 1987 to over 9 percent in 2017 in our data. Furthermore, this industry had notably high dispersion, which increased substantially over the sample period, likely due to the variety of products offered in this category, as discussed in Section 3.5.

Consistent with those findings, we find allocative efficiency to be low in NAICS 4541, even in 1987, prior to the rise of online shopping. Moreover, as the industry became more prominent, it also became less efficient, as seen in Figure 8. Thus, the contribution of this sector to overall allocative efficiency is due to both a “between” effect—it was a low-efficiency industry that gained revenue share—and a “within” effect—its efficiency fell substantially during the period. Furthermore, these were correlated, which is often captured by a “covariance” term in standard decompositions.

Rather than perform a decomposition, as in Foster et al. (2001), we instead perform a similar exercise to Blackwood et al. (2026), where we remove the influence of the industry. The result of this exercise is displayed in Figure 9. We find that the downturn in allocative efficiency can be eliminated by removing e-commerce from consideration, and the resulting trend is consistently positive. This positive trend in allocative efficiency is then consistent with hours-weighted dispersion measures rather than unweighted measures.

Figure 8: Allocative Efficiency in E-commerce



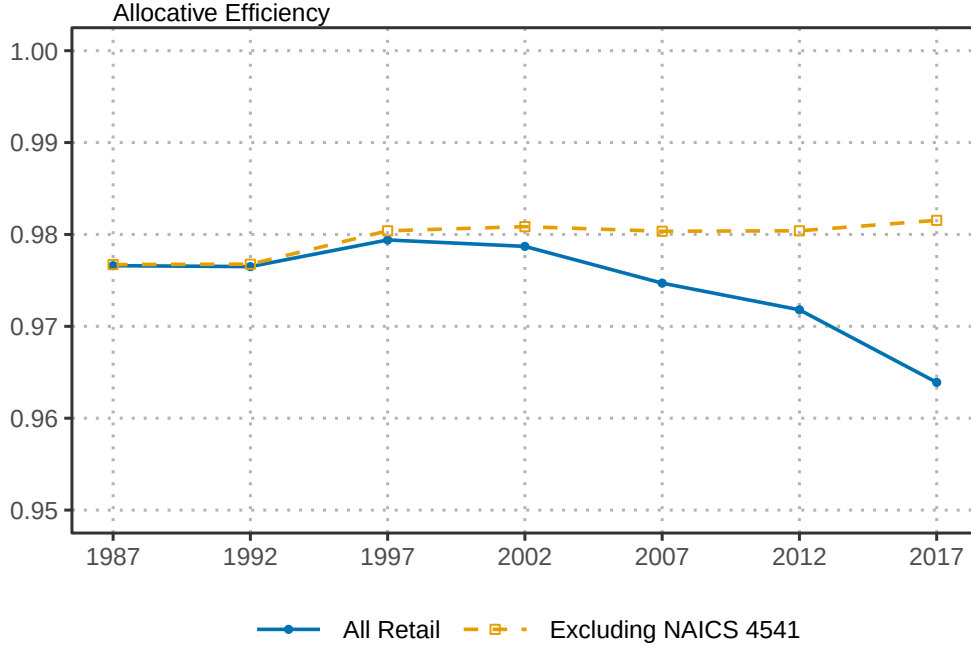
Notes: Labor allocative efficiency for all retail and for Electronic Shopping and Mail-Order Houses (NAICS 4541), baseline parameterization. The revenue share of NAICS 4541 rose from 0.3 percent in 1987 to 9.3 percent in 2017. Authors' calculations from census microdata.

5.2 Discussion

Depending on how establishments are weighted, productivity-dispersion statistics display very different trends in the retail sector (Figure 2). Based on our allocative efficiency results, we conclude that it is generally the case that hours-weighted statistics better capture patterns of allocative efficiency in the retail sector. This stands in contrast to previous focus on unweighted dispersion in the literature. Critically, Hsieh and Klenow (2009) note that LP dispersion is a sufficient statistic under the assumption that A_{is} and LP_{is} are jointly log normally distributed. The consistency of the allocative efficiency results with the weighted dispersion measure suggests that this condition does not hold in retail.

However, the clear preference for weighted dispersion measures remains only as long as stores within industries are suitably homogeneous in their process and product, which is not the case in online retail. This poses a measurement challenge going forward, as online

Figure 9: Allocative Efficiency Excluding E-commerce



Notes: Labor allocative efficiency for the retail sector including and excluding Electronic Shopping and Mail-Order Houses (NAICS 4541), baseline parameterization. Authors' calculations from census microdata.

retailers will be reassigned to industries more closely aligned with their product, but perhaps at the expense of process homogeneity. The effect of this shift is an area for future study as the 2022 CRT data are made available.

We also note that these broad findings are not sensitive to parameter choices, as displayed in Appendix B.2. The most notable change is the drop in level when $\alpha = 0.33$, to demonstrate the effect of a larger influence of the labor margin. This results in much lower levels of AE and more pronounced changes. Alternatively, setting $\rho_s = 0.75$ for all industries does not influence our findings in any substantial way.

6 Concluding Remarks

In this paper we use newly developed productivity-dispersion statistics for retail trade industries to examine how the transformation in the retail sector impacted the productivity

distribution. Our work builds on earlier research covering manufacturing industries. We discuss measurement issues and detail the methods by which labor productivity is measured at the retail industry level. The retail trade sector has seen dramatic changes since the late 1980s as single stores have been replaced by stores in national chains, which are larger and more homogeneous. At the same time, a movement to online shopping has led to a substantial change in shopping experiences. Given the structure of NAICS codes prior to 2022, much of this activity is allocated to a single industry.

The effects of these shifts can be seen in the dispersion statistics by looking at the retail sector along different dimensions. Most of the impact of national chains on dispersion shows up in the middle of the productivity distribution, which is consistent with other findings that much of the expansion of national chains was in the middle of the productivity distribution (Smith et al., 2023). This is also consistent with the narrative that big-box retailers expand by replicating productive (but not extremely productive) stores. Comparing hours-weighted to unweighted mean IQRs and store-based to firm-based mean IQRs, we see the impact of the expansion of national chains. Store-based dispersion is higher and exhibits a slight upward trend with no weighting. But when stores are weighted by hours, the mean IQR has a downward trend. Comparing firm-based and store-based dispersion statistics, a similar story emerges. We then explore productivity patterns on a more granular level, and highlight that the share of employment in large firms in an industry is negatively correlated with industry dispersion. Likewise, a larger share of small firms is associated with more dispersion. Together, these suggest that the shift to larger chains is associated with less dispersion. Finally, we demonstrate that e-commerce is substantially more dispersed, with a consistent and substantial upward trend.

We then explore the implications for efficiency in a standard model of misallocation similar to Hsieh and Klenow (2009). We find that measured allocative efficiency first rises in the '90s, then falls starting in the 2000s, but all of this decline can be accounted for by the e-commerce sector. Removing the effect of this sector yields a consistent upward trend,

which is consistent with the weighted statistics rather than the unweighted statistics. The influence of this single sector on aggregate trends suggests that classification choices could be critical for inferred allocative efficiency measures.

References

- Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, 2020. “The Fall of the Labor Share and the Rise of Superstar Firms.” *The Quarterly Journal of Economics* 135(2):645–709, <https://dx.doi.org/10.1093/qje/qjaa004>.
- Bils, Mark, Peter J. Klenow, and Cian Ruane, 2021. “Misallocation or Mismeasurement?” *Journal of Monetary Economics* 124:S39–S56.
- Blackwood, G. Jacob, Cindy Cunningham, Matthew Dey, Lucia Foster, Cheryl Grim, John Haltiwanger, Rachel Nesbit, Sabrina Wulff Pabilonia, Jay Stewart, Cody Tuttle, and Zoltan Wolf, 2025. “Job Tasks, Worker Skills, and Productivity.” Tech. Rep. CES Working Paper 25-63 and BLS Working Paper WP-585, CES and BLS. <https://www2.census.gov/library/working-papers/2025/adrm/ces/CES-WP-25-63.pdf>.
- Blackwood, G. Jacob, Lucia S. Foster, Cheryl A. Grim, John Haltiwanger, and Zoltan Wolf, 2021. “Macro and Micro Dynamics of Productivity: From Devilish Details to Insights.” *American Economic Journal: Macroeconomics* 13(3):142–172.
- Blackwood, G. Jacob, John Haltiwanger, and Zoltan Wolf, 2026. “Allocating Misallocation: Decomposing Measures of Aggregate Allocative Efficiency.” Tech. Rep. 26-26, CES. <https://www.census.gov/library/working-papers/2026/adrm/CES-WP-26-26.html>.
- Chava, Sudheer, Alexander Oettl, Manpreet Singh, and Linghang Zeng, 2024. “Creative Destruction? Impact of E-Commerce on the Retail Sector.” *Management Science* 70(4):2168–2187.
- 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.” Tech. Rep. 28839, NBER. <https://doi.org/10.3386/w28839>.

Cox, Lydia, Gernot J. Müller, Ernesto Pastén, Raphael Schoenle, and Michael Weber, 2024. “Big G.” *Journal of Political Economy* 132(10):3260–3297, <https://dx.doi.org/10.1086/730426>.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf, 2023. “Dispersion in Dispersion: Measuring Establishment-Level Differences in Productivity.” *Review of Income and Wealth* 69(4):999–1032.

Foster, Lucia, John Haltiwanger, Shawn Klimek, Cornell J. Krizan, and Scott Ohlmacher, 2016. “The Evolution of National Retail Chains: How We Got Here.” In **Emek Basker**, editor, *Handbook on the Economics of Retailing and Distribution*. Edward Elgar. <https://dx.doi.org/10.4337/9781783477388>.

Foster, Lucia, John Haltiwanger, and C. 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):748–758, <https://dx.doi.org/10.1162/rest.88.4.748>.

Foster, Lucia S., John Haltiwanger, and C.J. Krizan, 2001. “Aggregate Productivity Growth: Lessons from Microeconomic Evidence.” Nber chapters, National Bureau of Economic Research. In *New Developments in Productivity Analysis*.

Gabaix, Xavier, 2011. “The Granular Origins of Aggregate Fluctuations.” *Econometrica* 79(3):733–772, <https://dx.doi.org/https://doi.org/10.3982/ECTA8769>.

Haltiwanger, John, Robert Kulick, and Chad Syverson, 2018. “Misallocation Measures: The Distortion That Ate the Residual.” NBER Working Paper 24199, National Bureau of Economic Research.

- Hsieh, Chang-Tai and Peter J. Klenow**, 2009. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics* 124(4):1403–1448, <https://dx.doi.org/10.1162/qjec.2009.124.4.1403>.
- Hsieh, Chang-Tai and Esteban Rossi-Hansberg**, 2023. “The Industrial Revolution in Services.” *Journal of Political Economy Macroeconomics* 1(1):3–42.
- Jarmin, Ronald S., Shawn D. Klimek, and Javier Miranda**, 2009. “The Role of Retail Chains: National, Regional and Industry Results.” In **Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts**, editors, *Producer Dynamics: New Evidence from Micro Data*, pp. 237–262. University of Chicago Press, Chicago. <http://www.nber.org/chapters/c0487>. National Bureau of Economic Research Conference, April 8–9, 2005.
- Ma, Yueran, Benjamin Pugsley, Haomin Qin, and Kaspar Zimmermann**, 2025. “Superstar Firms through the Generations.” Working Paper 34194, National Bureau of Economic Research. <https://dx.doi.org/10.3386/w34194>.
- National Academies of Sciences, Engineering, and Medicine**, 2021. *A Satellite Account to Measure the Retail Transformation: Organizational, Conceptual, and Data Foundations*. The National Academies Press, Washington, DC. <https://dx.doi.org/10.17226/26101>.
- Smith, Dominic, G. Jacob Blackwood, Michael Giandrea, Cheryl Grim, Jay Stewart, and Zoltan Wolf**, 2023. “Productivity Dispersion and Structural Change in Retail Trade.” CES Working Paper 23-60R, U.S. Census Bureau. <https://www2.census.gov/library/working-papers/2023/adrm/ces/CES-WP-23-60R.pdf>.
- Smith, Dominic A. and Sergio Ocampo**, 2025. “The Evolution of US Retail Concentration.” *American Economic Journal: Macroeconomics* 17(1):71–101, <https://dx.doi.org/10.1257/mac.20220249>.

US Bureau of Labor Statistics, 1987-2017. “Labor Productivity and Cost Measures.” <https://www.bls.gov/productivity/tables/>. Accessed 2025.

US Census Bureau, 1987-2017a. “2022 Business Dynamics Statistics.” <https://www.census.gov/programs-surveys/bds.html>. Accessed 2026.

US Census Bureau, 1987-2017b. “Annual Retail Trade Survey.” <https://www.census.gov/programs-surveys/arts/data/tables.html>. Accessed 2025.

US Census Bureau, 1987-2017c. “Census of Retail Trade.” <https://www.census.gov/topics/research/guidance/restricted-use-microdata/economic-data.html>. Accessed 2026.

US Census Bureau, 1987-2017d. “Census of Wholesale Trade.” <https://www.census.gov/topics/research/guidance/restricted-use-microdata/economic-data.html>. Accessed 2026.

US Census Bureau, 1987-2017e. “Longitudinal Business Database.” <https://www.census.gov/programs-surveys/ces/data/restricted-use-data/longitudinal-business-database.html>. Accessed 2026.

Appendix

A Additional Dispersion Results

A.1 Dispersion across Industries

The dispersion results in Section 3 tell us about average dispersion across industries but say little about how dispersion varies across industries. If a small number of industries have high dispersion while others have minimal dispersion, that could indicate that our results are driven by industry classification; some industry codes are more broadly defined than others. We find that this is not the case. The level of dispersion for even the industry at the 25th percentile of the distribution tends to be quite high, and the time trends of industries at the 25th and 75th percentiles of the distribution qualitatively follow the trends of the mean and median industries. Figure 10 shows percentiles of the across-industry distribution of IQRs, 99–90 ranges, and 10–1 ranges. We plot the 25th, 75th, and 50th percentiles of the cross-industry distribution in addition to the means that were plotted in Figure 2. We present results with industry and store-hour weights and without any weighting in Figure 10.¹⁵ With 27 industries, the 75th percentile corresponds to the industry with the seventh-highest dispersion in a given year, the 50th percentile corresponds to the fourteenth, and the 25th corresponds to the twentieth. In Panel (a) we see that a wide range of industries exhibit IQRs over 100. Comparing the two panels, we see that the range between the 75th and 25th percentiles is similar for weighted and unweighted dispersion statistics, but the mean IQR is lower when weights are used and their trends differ. The 75th and 25th percentile lines converge when no weights are used (Figure 10a), indicating that industries are looking more similar in terms of dispersion. The convergence is due to both a decrease in the 75th percentile IQR and an increase in the 25th percentile IQR. But 75th and 25th percentile trends are roughly parallel in the industry-store weighted graph (Figure 10b), even though

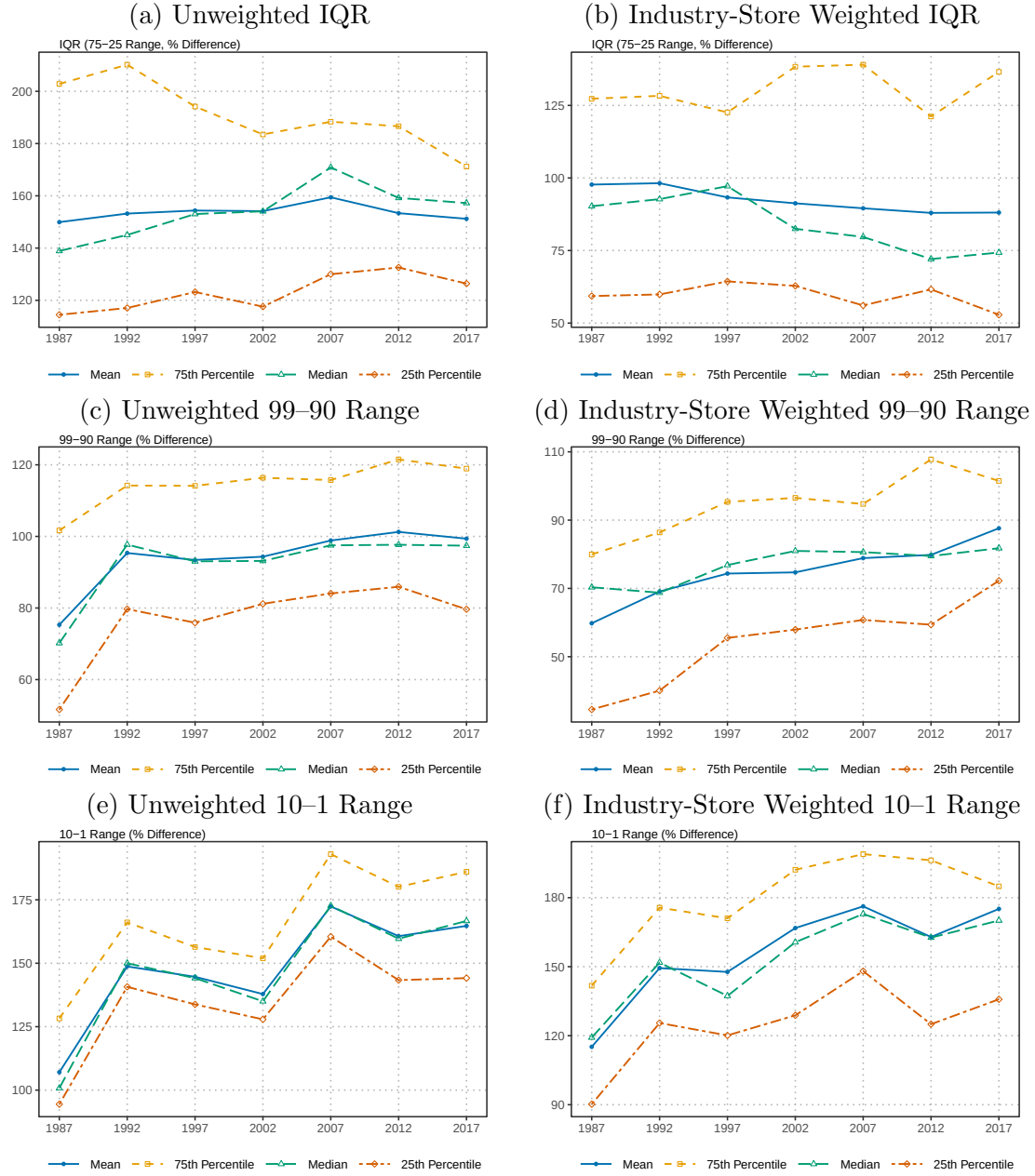
¹⁵Results for the “Industry Weighted” and “Store Weighted” methods fall between these two extremes.

the mean IQR is decreasing.¹⁶

Turning to the tails of the productivity distribution, weighting continues to have a minimal impact in the 99–90 dispersion statistics. In all cases, there is a slight upward time trend, with the weighted graph having slightly lower levels. These findings are consistent with the reallocation of economic activity toward large firms, mostly impacting the middle of the productivity distribution. At the other end of the productivity distribution (Panels e and f), all percentiles and weighting methods tell the same story of increasing dispersion. On a weighted basis, there is more variation across industries, and the 75th percentile is flat or falling for the final 10 years of the sample.

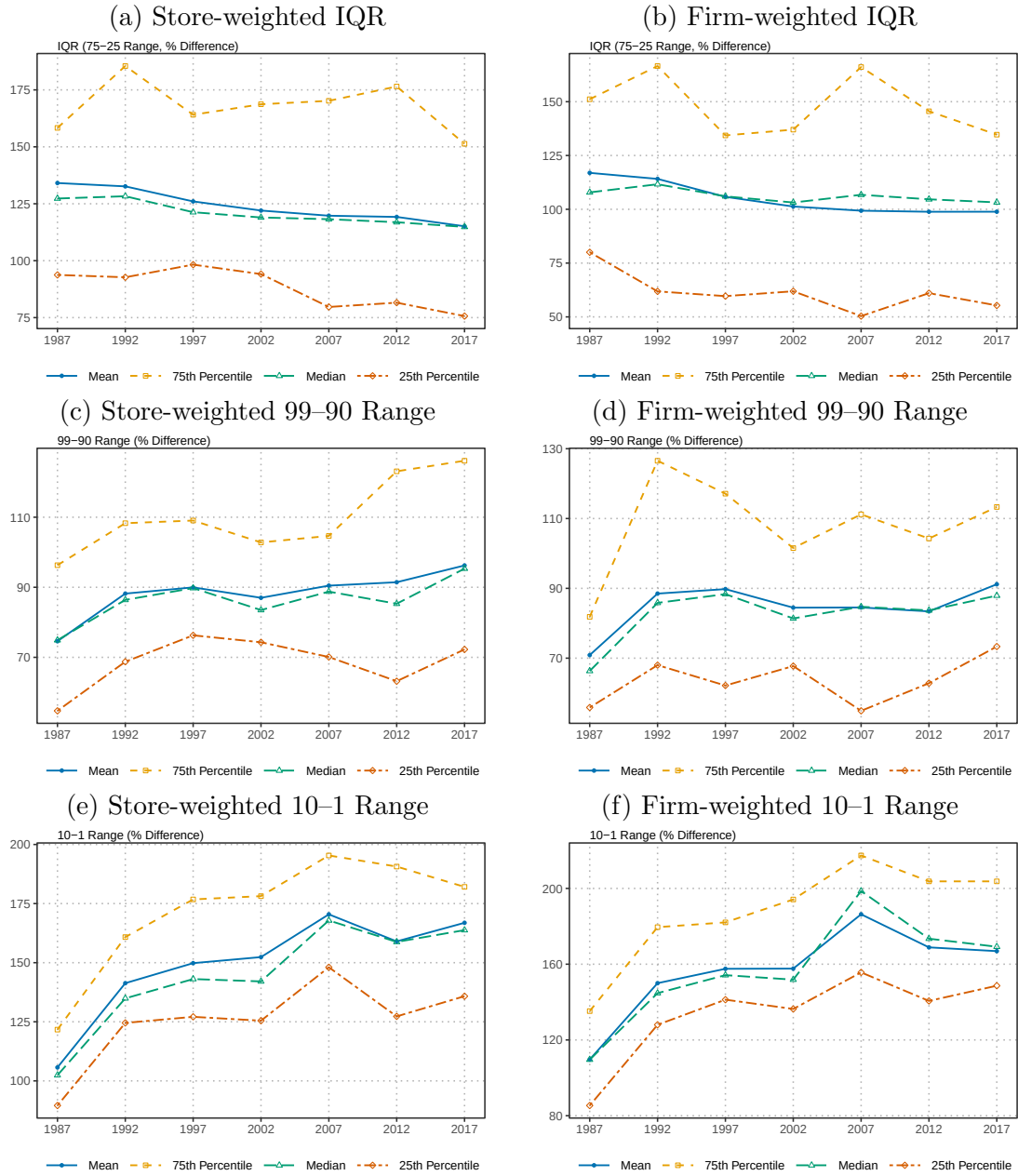
¹⁶It is not necessarily the same industries that occupy these percentiles across time and weighting schemes.

Figure 10: Dispersion across Industries: Unweighted vs. Weighted



Notes: Graph lines represent the variation across industries in the within-industry statistics. The lines indicate the 25th percentile, median, 75th percentile, and mean within-industry range across industries. The y-axis is expressed in percentage differences indicating how much more productive establishments at the top of the range are relative to those at the bottom. A value of 100 indicates stores at the top are twice as productive as stores at the bottom. “Industry-store weighted” is referred to as “both weighted” in the main text. Figures are based on authors’ calculations using CRT and BLS data.

Figure 11: Dispersion across Industries: Store vs. Firm



Notes: Graph lines represent the variation across industries in the within-industry statistics. The lines indicate the 25th percentile, median, 75th percentile, and mean of the distributions of within-industry ranges across industries. The y-axis is expressed in percentage differences. Stores and firms are weighted by share of total hours within each industry. Figures are based on authors' calculations using CRT and BLS data.

A.2 Industry-Specific Dispersion Patterns

Of the 27 four-digit retail industries for each year, we focus on five industries that highlight differences across industries that cannot be seen in the aggregate statistics. Four of these industries were selected because of their relative importance, or growth in importance, over the past 30 years, and because they are representative to varying degrees of the key structural changes discussed above. The two largest, Grocery Stores (4451) and General Merchandise Stores, including Warehouse Clubs and Supercenters (4523), accounted for approximately one-third of retail employment hours in 2017 and are industries in which chain stores feature prominently. Electronics and Appliance Stores (4431) is an industry composed of a mix of single-store firms and chains of various sizes, and it experienced substantial sales growth over the sample period. We include Clothing Stores (4481) because that industry has experienced faster productivity growth than most retail industries from 1987 through 2017. Finally, Electronic Shopping and Mail-Order Houses (4541) was chosen due to its rapid rise to prominence over the sample period, representing a new technology and a fundamental shift in how consumers shop as the emphasis shifted away from mail-order to Internet shopping. As noted above, the measurement issues that arise with the new technologies represented in this industry have implications for dispersion and misallocation, which may be important to isolate. Figures 12a–12e show the IQR, the 99–90 range, and the 10–1 range of labor productivity of establishments for these five industries.

In Figure 12a, the IQR for Grocery Stores (NAICS 4451) was essentially flat over the entire sample period, increasing slightly from about 110 percent in 1987 to 120 in 2017. That is, the grocery store at the 75th percentile was 110 percent more productive than a store at the 25th percentile in 1987, and 120 percent more productive in 2017 (in other words, 2.1 times and 2.2 times as productive). Dispersion in the tails followed a different pattern. Between 1987 and 2012, dispersion in both tails grew by more than the IQR before falling between 2012 and 2017. The 99–90 difference increased from 50 percent to around 100 percent, with nearly all of the increase occurring between 1987 and 1992. In 1992 and later

years, dispersion in the upper tenth of the distribution was about the same as for the middle of the distribution. Dispersion is greater and more variable among the least productive stores (except for 1987) as compared to the middle half and upper tail of the productivity distribution.

Among General Merchandise Stores (NAICS 4523) in Figure 12b, dispersion levels are similar to Grocery Stores, with dispersion being the greatest in the lower tail of the productivity distribution and about the same for the upper tail and middle half of the distribution. The IQR declined over the beginning of the sample period, while dispersion in the lower tail increased substantially. Like Grocery Stores, dispersion was considerably more variable in the lower tail.

The trends in productivity dispersion in the Electronics and Appliances (NAICS 4431) industry differ from those of the other retail industries (Figure 12d) in that the IQR declined sharply between 1987 and 2002 and then increased slightly between 2002 and 2017. Dispersion in the tails was lower or the same as for the IQR before 2002. As with most of the other industries, there is less dispersion in the upper tail than in the other parts of the distribution.

Figure 12e shows that the IQR among Clothing Stores (NAICS 4481) was relatively stable over the sample period at about 120 percent, while both the 99–90 and 10–1 percentile ranges increased substantially. Compared with the other four industries, the dispersion ranges for clothing stores exhibited less volatility even though this industry experienced significant productivity growth.¹⁷

A.3 Comparison to Manufacturing

Although the market conditions facing retail stores differ from those faced by manufacturing plants during this period, the manufacturing sector is a useful comparison, because there has been little research on productivity distribution outside of manufacturing. Figures 13a, 13b,

¹⁷In the 2022 NAICS, e-commerce activities are included with the industry associated with the products sold. For example, online clothing sales will be included in Clothing rather than Electronic Shopping.

and 13c show the percentile ranges (both store-weighted and unweighted) for manufacturing and retail trade.

Starting with the IQRs in Figure 13a, we see the familiar pattern that the unweighted IQRs are greater than store-weighted IQRs in both sectors. Looking at the unweighted lines, we see that there is more dispersion in retail trade and that the trends in the two sectors are similar until 2002. After 2002, the manufacturing IQR increases rapidly, while the retail IQR remains relatively flat. Turning to the weighted IQRs, we see more significant differences. The weighted retail IQR starts at a higher level but exhibits a strong downward trend, while the weighted manufacturing IQR trends strongly upward.

Turning to the right tail, the two sectors have different levels of dispersion, but they have similar trends. Right-tail dispersion is greater in manufacturing in most years whether or not we weight by hours. However, the difference is smaller when establishments are weighted. Furthermore, the difference between the weighted and unweighted lines is much larger in manufacturing.

More investigation into these differences is warranted but is beyond the scope of this paper. However, we describe some plausible explanations briefly. We saw above that the shift from single stores to national chains, which open larger and more homogeneous stores, plausibly explains some of the patterns in retail. Some of this could be due to new technologies in retail, such as scanners and self-checkout, that can be retrofitted into existing stores. In contrast, new plants in manufacturing industries are more likely to be more productive than existing plants because they can more easily incorporate the latest technology, regardless of size. Older plants are less likely to be retrofitted because the existing structure may not be able to accommodate the new technology. Thus, as new manufacturing plants are built, we would expect dispersion to increase. Differences in the left tails of the distributions likely have a different explanation. To speculate, one possibility is that stores in the left tail of the retail distribution may be more likely to survive, relative to manufacturing establishments, if they serve geographically isolated areas where there is little competition or if they

cater to niche markets. In manufacturing, new plants are more productive than older and less-productive plants, making those plants less likely to survive.

B Robustness

B.1 Payroll vs. Hours

Our use of industry averages for hours per employee means that the variation in total hours across establishments is determined by variation in total employment and does not account for variation in the use of part-time versus full-time employees. And hours worked during the March reference period may not be representative of the entire year. As a robustness check, we repeat the dispersion statistics shown in Figure 2 using total payroll in the denominator rather than total hours. It should be noted that 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), worker-firm effects, leave policies, and any compensating wage differentials. In this sense this measure can be considered more comprehensive. Figure 14 shows that using payroll implies significant differences. First, using payroll implies considerably less measured dispersion (Figure 14a): all three lines are below their counterparts in Figure 2. This indicates that payroll measures may capture additional variation in effective labor, including differences in hours and differences in worker skills and occupation, as discussed in Blackwood et al. (2025). 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 3a and 14b, we see similar differences in time trends. The dispersion measures for payroll-based productivity start at a higher level, but the two sets of lines converge and remain at approximately the same level for the last 10 years of the sample. Finally, looking at the left tails, the lines in both Figures 3b and 14c trend up. The payroll-based productivity dispersion statistics are lower than the hours-based

ones, and they exhibit significantly less variation.¹⁸

B.2 Robustness of Allocative Efficiency Results

Figure 15 displays labor allocative efficiency under the five parameterizations described in Section 4. Panel (a) covers all retail, and Panel (b) excludes Electronic Shopping and Mail-Order Houses (NAICS 4541). The baseline holds aggregate parameters constant across industries and years. The alternatives allow time-varying aggregate parameters, industry-specific parameters fixed at 2007 levels, a labor share of 0.33 for all industries, and $\rho_s = 0.75$ for all industries. The time series patterns are similar across parameterizations. Setting the labor share to 0.33 lowers the level of allocative efficiency and amplifies its changes, as discussed in Section 5, but does not alter the qualitative conclusions.

C Model Derivations

C.1 Sectoral input demand and aggregate labor productivity

Note that the standard assumption is that aggregate inputs are in fixed supply, so total L is known, and we need to figure out how it is allocated across sectors. Start with store input demand :

$$w_s L_s = \rho_s \alpha_{Ls} R_s (1 + \tau_s)^{-1}.$$

There are two approaches here. The first is to allow w_s to vary. Summing and multiplying by L/L :

$$L \underbrace{\sum_v \frac{L_v}{L} w_v}_w = \sum_v \rho_v \alpha_{Lv} R_v (1 + \tau_v)^{-1}.$$

¹⁸Although the standard deviation is sensitive to outliers, we repeated the analysis using this measure for robustness. The main implication is that dispersion measured by the standard deviations exhibits the same patterns documented above.

Then, dividing the first equation by the second and re-arranging, we have:

$$L_s = L \frac{w}{w_s} \frac{\rho_s \alpha_{Ls} R_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} R_v (1 + \tau_v)^{-1}}.$$

Noting that $R_s = \theta_s Q$, we have:

$$L_s = L \frac{w}{w_s} \frac{\rho_s \alpha_{Ls} \theta_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}}.$$

Note that we can then rearrange:

$$w = \frac{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}}{\sum_s \frac{\rho_s \alpha_{Ls} \theta_s}{w_s (1 + \tau_s)}}$$

You still need sectoral wage data here, and an assumption that labor cannot move freely across sectors (this shows up as a kind of cross-sectoral misallocation).

Alternatively, if we assume labor can move freely across sectors, and $w_s = w$ for all s , then:

$$L_s = L \frac{\rho_s \alpha_{Ls} \theta_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}}.$$

Then we can write aggregate labor productivity as

$$\begin{aligned} LP &= \frac{Q}{L} = \frac{\prod_s Q_s^{\theta_s}}{L} = \frac{\prod_s (LP_s L_s)^{\theta_s}}{L} = \left(\prod_s LP_s^{\theta_s} \right) \left(\frac{L \prod_s \left(\frac{\rho_s \alpha_{Ls} \theta_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}} \right)^{\theta_s}}{L} \right) \\ &= \left(\prod_s LP_s^{\theta_s} \right) \underbrace{\prod_s \left(\frac{\rho_s \alpha_{Ls} \theta_s (1 + \tau_s)^{-1}}{\sum_v \rho_v \alpha_{Lv} \theta_v (1 + \tau_v)^{-1}} \right)^{\theta_s}}_{\bar{T}} \\ &= \bar{T} \times \prod_s LP_s^{\theta_s}. \end{aligned}$$

C.2 Details for sectoral LP

Start with the expression for LP and re-express as revenue productivity:

$$LP_s = \frac{Q_s}{L_s} = \frac{R_s}{L_s} P_s^{-1}.$$

We can now derive a relationship between sectoral labor productivity and the sectoral distortion by summing over a rearranged first-order condition $\rho_s \alpha_{Ls} R_{is} \frac{1}{1+\tau_{is}} = w_s L_{is}$ and multiplying by $\frac{R_s}{L_s}$:

$$\rho_s \alpha_{Ls} R_s \underbrace{\sum_i \frac{R_{is}}{R_s} \frac{1}{1+\tau_{is}}}_{(1+\tau_s)^{-1}} = w_s \underbrace{\sum_i L_{is}}_{L_s}.$$

Rearranging again, we have

$$\frac{R_s}{L_s} = \frac{w_s (1 + \tau_s)}{\rho_s \alpha_{Ls}}.$$

Plugging in the above, we have

$$\frac{Q_s}{L_s} = \frac{w_s}{\rho_s \alpha_{Ls}} (1 + \tau_s) P_s^{-1}.$$

Multiply by $Q_s^{-(1-\rho_s \alpha_{Ls})} Q_s^{1-\rho_s \alpha_{Ls}}$, to arrive at

$$LP_s = \frac{w_s}{\rho_s \alpha_{Ls}} (1 + \tau_s) P_s^{-1} Q_s^{-(1-\rho_s \alpha_{Ls})} \left(\sum_i A_{is}^{\rho_s} L_{is}^{\rho_s \alpha_{Ls}} \right)^{\frac{1-\rho_s \alpha_{Ls}}{\rho_s}}.$$

We need to substitute for labor demand, rearrange the FOC again:

$$\begin{aligned} \rho_s \alpha_{Ls} P_s Q_s^{1-\rho_s} A_{is}^{\rho_s} L_{is}^{\rho_s \alpha_{Ls}} &= (1 + \tau_{is}) w_s L_{is} \\ \implies L_{is} &= \left(\frac{\rho_s \alpha_{Ls} P_s Q_s^{1-\rho_s} A_{is}^{\rho_s}}{w_s (1 + \tau_{is})} \right)^{\frac{1}{1-\rho_s \alpha_{Ls}}}. \end{aligned}$$

Plugging in and rearranging, we have:

$$LP_s = \left(\frac{w_s}{\rho_s \alpha_{L_s}} \right)^{1-\alpha_{L_s}} (1 + \tau_s) P_s^{-(1-\alpha_{L_s})} Q_s^{-(1-\alpha_{L_s})} \left(\sum_i A_{is}^{\rho_s} \left(\frac{A_{is}^{\rho_s}}{1 + \tau_{is}} \right)^{\frac{\rho_s \alpha_{L_s}}{1-\rho_s \alpha_{L_s}}} \right)^{\frac{1-\rho_s \alpha_{L_s}}{\rho_s}}.$$

Simplifying, and noting $R_s = P_s Q_s$, we have

$$LP_s = \left(\frac{w_s}{\rho_s \alpha_{L_s}} \right)^{1-\alpha_{L_s}} \frac{(1 + \tau_s)}{R_s^{1-\alpha_{L_s}}} \left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s \alpha_{L_s}}} (1 + \tau_{is})^{\frac{-\rho_s \alpha_{L_s}}{1-\rho_s \alpha_{L_s}}} \right)^{\frac{1-\rho_s \alpha_{L_s}}{\rho_s}}.$$

Multiplying by $\left(\frac{L_s}{L_s} \right)^{1-\alpha_{L_s}}$, we have

$$LP_s = \left(\frac{w_s}{\rho_s \alpha_{L_s}} \right)^{1-\alpha_{L_s}} \left(\frac{R_s}{L_s} \right)^{-(1-\rho_s \alpha_{L_s})} (1 + \tau_s) \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s \alpha_{L_s}}} (1 + \tau_{is})^{\frac{-\rho_s \alpha_{L_s}}{1-\rho_s \alpha_{L_s}}} \right)^{\frac{1-\rho_s \alpha_{L_s}}{\rho_s}}}{L_s^{1-\alpha_{L_s}}}$$

From above, we have that sectoral revenue productivity $\frac{R_s}{L_s} = \frac{w_s(1+\tau_s)}{\rho_s \alpha_{L_s}}$, which we can plug in to obtain

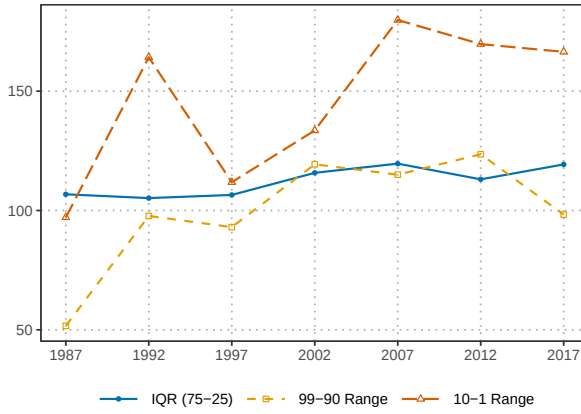
$$LP_s = (1 + \tau_s)^{\alpha_{L_s}} \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s \alpha_{L_s}}} (1 + \tau_{is})^{\frac{-\rho_s \alpha_{L_s}}{1-\rho_s \alpha_{L_s}}} \right)^{\frac{1-\rho_s \alpha_{L_s}}{\rho_s}}}{L_s^{1-\alpha_{L_s}}},$$

and the sectoral distortion can be moved into the brackets

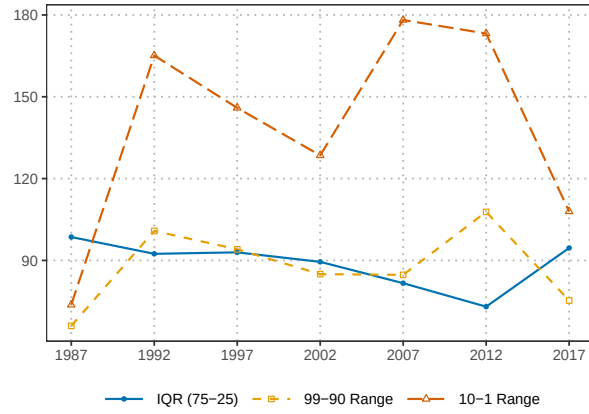
$$LP_s = \frac{\left(\sum_i A_{is}^{\frac{\rho_s}{1-\rho_s \alpha_{L_s}}} \left(\frac{1+\tau_{is}}{1+\tau_s} \right)^{\frac{-\rho_s \alpha_{L_s}}{1-\rho_s \alpha_{L_s}}} \right)^{\frac{1-\rho_s \alpha_{L_s}}{\rho_s}}}{L_s^{1-\alpha_{L_s}}}.$$

Figure 12: Within-Industry Dispersion — Key Industries

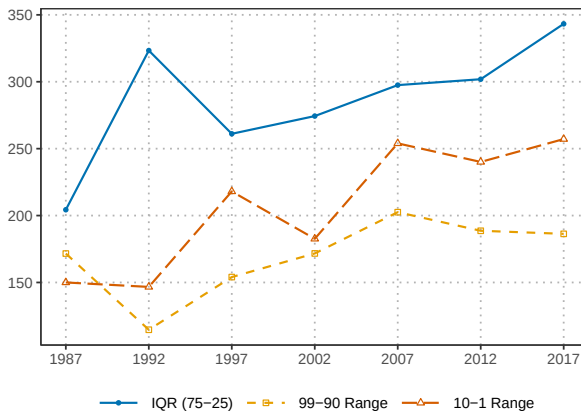
(a) Grocery Stores - 4451



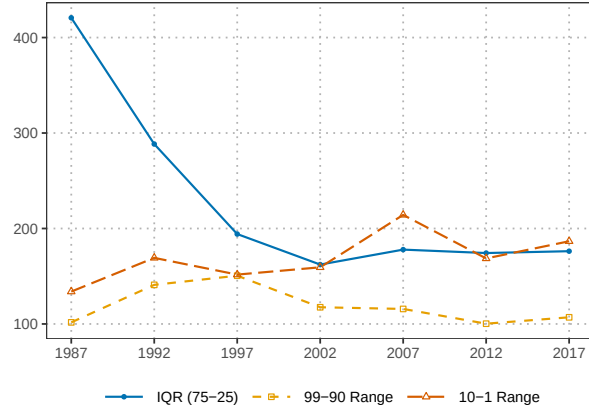
(b) General Merchandise Stores - 4523



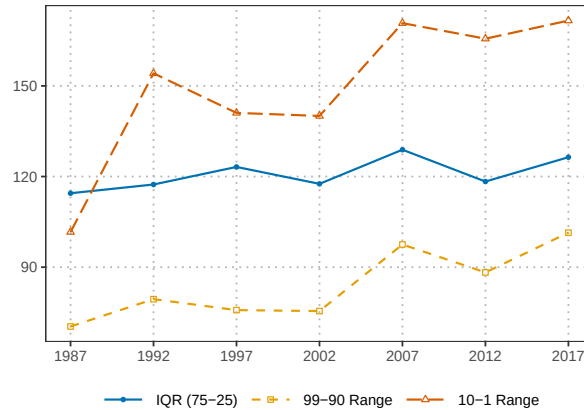
(c) Electronic Shopping and Mail-Order Houses - 4541



(d) Electronics and Appliances - 4431

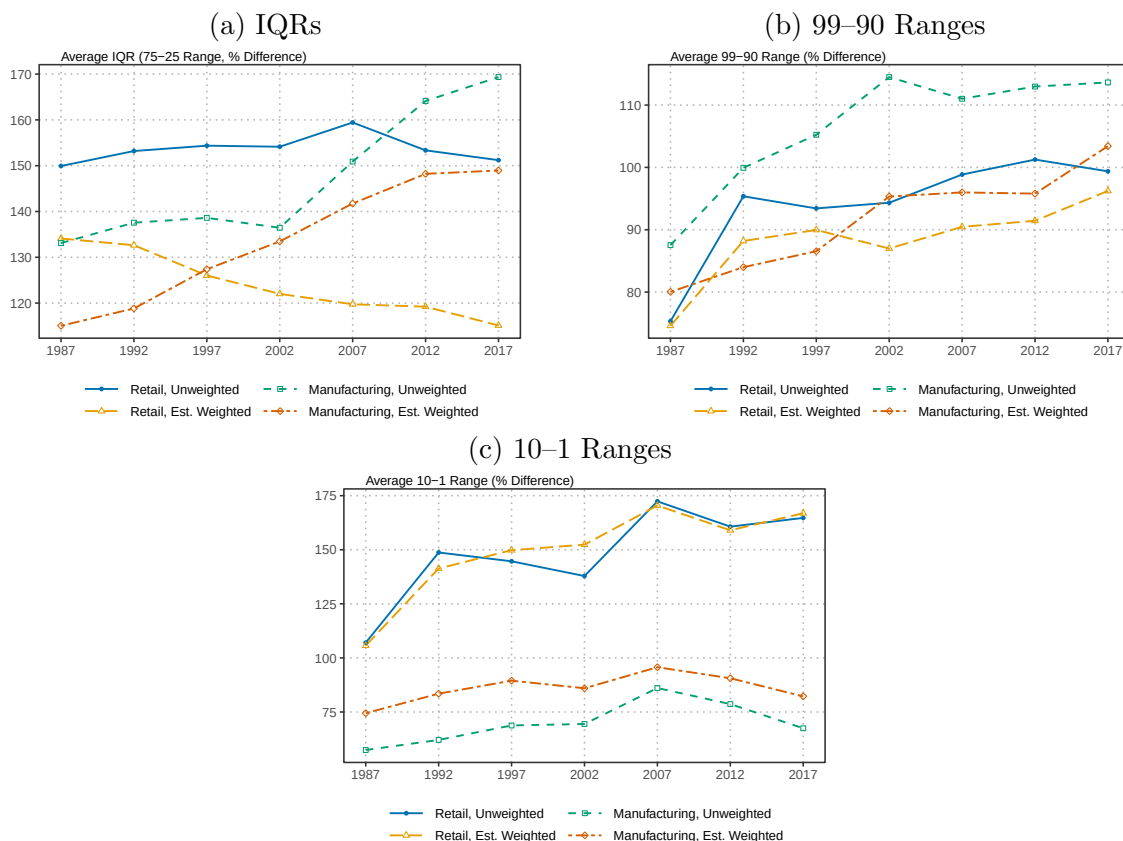


(e) Clothing Stores - 4481



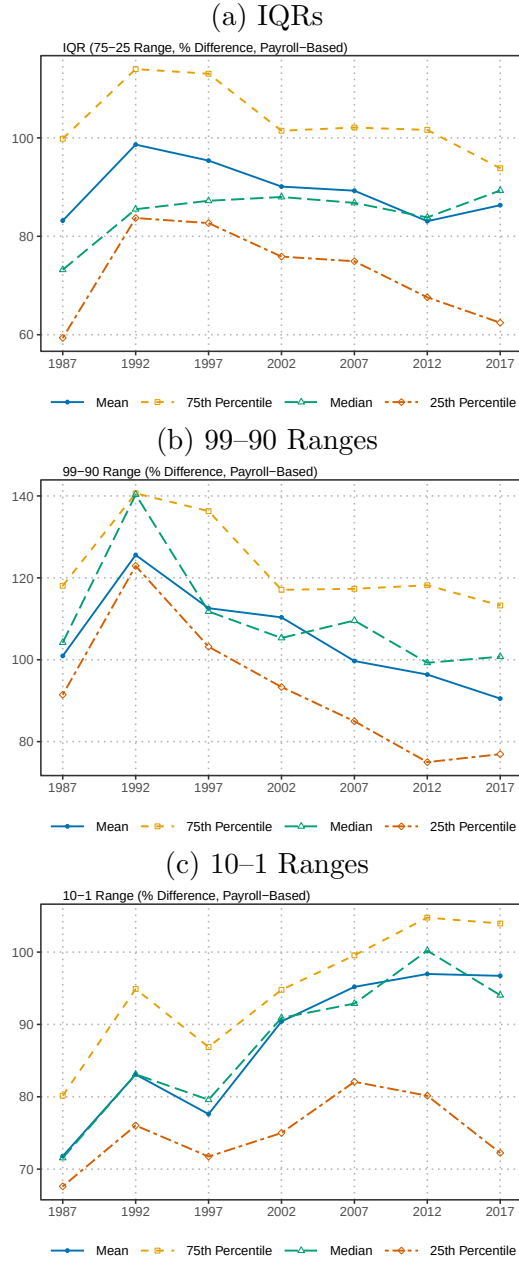
Notes: Lines represent the unweighted difference between establishments at different percentiles of the within-industry labor productivity distribution in each year. The y-axis is expressed in percentage differences. Based on authors' calculations using CRT and BLS data.

Figure 13: Comparison of Retail Trade and Manufacturing



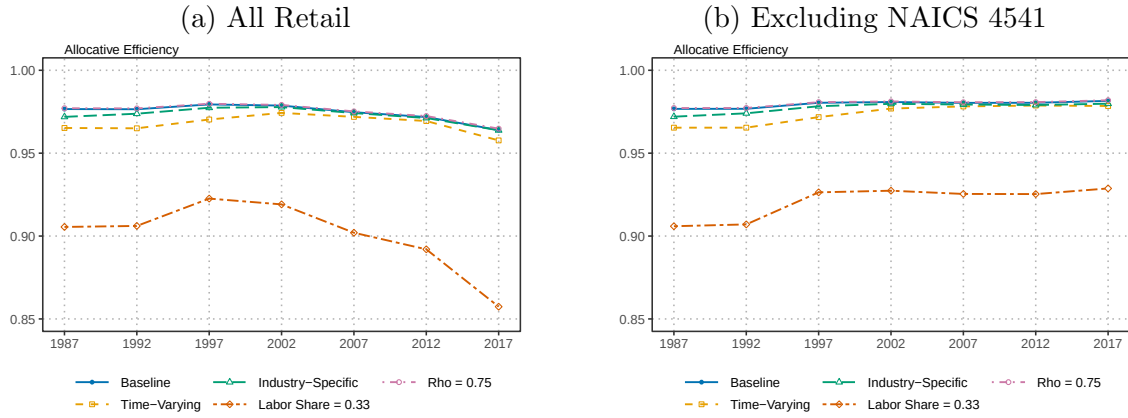
Notes: Graph lines represent the average ranges (IQR, 99–90, 10–1) across industries when using different weighting schemes. “Unweighted” is an unweighted arithmetic average. “Weighted” ranges are calculated using each establishment’s share of hours within that industry. Figures are based on authors’ calculations using CRT and BLS data.

Figure 14: Dispersion across Industries: Establishment-Payroll-Based



Notes: Graph lines represent the variation across industries in the within-industry statistics. The lines indicate the 25th percentile, median, 75th percentile, and mean within-industry range across industries. The y-axis is expressed in percentage differences, consistent with Figure 2. Productivity is calculated as the log ratio of sales to payroll. NAICS-4 industries are weighted according to their share of total hours within a year. Establishments within each industry are given equal weight. Figures are based on authors' calculations using CRT and BLS data.

Figure 15: Allocative Efficiency under Alternative Parameterizations



Notes: Labor allocative efficiency for the retail sector under alternative parameter choices. “Baseline” uses aggregate parameters constant across industries and years; “Time-Varying” allows aggregate parameters to vary by year; “Industry-Specific” uses industry-level parameters fixed at 2007 values; the remaining series set the labor share to 0.33 or ρ_s to 0.75 for all industries. Authors’ calculations from census microdata (CBDRB-FY26-0226).