

# The Industrial Revolution in Services

---

Chang-Tai Hsieh

*University of Chicago*

Esteban Rossi-Hansberg

*University of Chicago*

US firms in service industries increasingly operate in more local markets. Employment, sales, and spending on fixed costs have increased rapidly in these industries. These changes have favored top firms, leading to increasing national concentration. Top firms in service industries have grown by expanding into new local markets, predominantly small and mid-sized US cities. Market concentration at the local level has decreased in all US cities, particularly in cities that were initially small. These facts are consistent with the availability of new fixed-cost-intensive technologies that yield lower marginal costs in service sectors. The entry of top service firms into new local markets has led to substantial unmeasured productivity growth, particularly in small markets.

## I. Introduction

Modern production relies on scale: the ability to use a technology to produce the same product or service innumerable times. In manufacturing industries, inventions such as the steam engine, electricity, and Ford's assembly line allowed firms to scale up production in a single large plant.

We thank Adarsh Kumar, Feng Lin, Harry Li, and Jihoon Sung for extraordinary research assistance. We also thank Rodrigo Adão, Dan Adelman, Audre Bagnall, Jill Golder, Bob Hall, Pete Klenow, Hugo Hopenhayn, Danial Lashkari, Raghuram Rajan, Richard Rogerson, and Chad Syverson for helpful discussions. Any views expressed are those of the authors and not

Electronically published March 30, 2023

*Journal of Political Economy Macroeconomics*, volume 1, number 1, March 2023.

© 2023 The University of Chicago. All rights reserved. Published by The University of Chicago Press.

<https://doi.org/10.1086/723009>

For many goods, the cost advantages of a larger scale overwhelmed the cost of transporting the goods to final consumers, leading to great reductions in total average costs. This ability to scale up production in a single plant was, however, of little use outside of manufacturing. Producing many cups of coffee, retail services, or health services in the same location is of no value, since it is impractical to bring them to final consumers. Modern large-scale production in these industries had to wait for a different technology, one that allowed firms to replicate cheaply the same production process in multiple locations close to consumers.

We argue that new ICT (information and communication technology)-based technologies, together with the adoption of new management practices, have finally made it possible for firms outside of manufacturing to scale up production over a large number of locations. This expansion in the number of markets per firm has been particularly pronounced for the top firms in nontradable industries and has led to an increase in their national market share, a central fact about the US economy in the past 3 or 4 decades documented by Autor et al. (2017). This evolution is the result of a new industrial revolution, one that has taken place in many nontraded-service sectors.

Consider Gawande's (2012, 53) account of how the Cheesecake Factory brought "chain production to complicated sit-down meals." The Cheesecake Factory has invested in technologies that determine optimal staffing and food purchases for each restaurant and each day. The company also has a well-oiled process via which they introduce new items on their menu. This process starts in a centralized "kitchen" in Calabasas, California—their R&D facility, so to speak—where Cheesecake's top cooks cull ideas for new dishes and "figure out how to make each recipe reproducible, appealing, and affordable" (60). The cooks in the R&D facility then teach the new recipes to the kitchen managers of each restaurant at a biannual meeting in California. The kitchen managers then follow a finely honed procedure to teach the new recipes to the cooks in each restaurant. The rollout time, from the time the kitchen managers arrive at Cheesecake's central kitchen in California to when the new dishes are put on the menu in each restaurant, is 7 weeks.

The standardization of production over a large number of establishments in sit-down restaurant meals due to companies such as the Cheesecake Factory has taken place in many nontraded sectors. Take hospitals as another example. Four decades ago, about 85% of hospitals were single-establishment nonprofits. Today, more than 60% of hospitals are owned by

---

those of the US Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2441. This paper was edited by Ariel Burstein.

for-profit chains or are part of a large network of hospitals owned by an academic institution (such as the University of Chicago Hospitals).<sup>1</sup> As an example of the former, consider the Steward Health Care Group. This company was created by the Cerberus private equity fund in 2010 when it purchased six Catholic hospitals in Boston. In Gawande's (2012) account, Cerberus's goal was to create the "Southwest Airlines of healthcare" by figuring out and codifying best practices and implementing these practices over a large scale. Gawande (2012, 61) describes the scene in Steward's remote intensive-care unit (ICU) in a Boston suburb that monitors the ICUs in all of Steward's hospitals:

Banks of computer screens carried a live feed of cardiac-monitor readings, radiology-imaging scans, and laboratory results from I.C.U. patients throughout Steward's hospitals. Software monitored the stream and produced yellow and red alerts when it detected patterns that raised concerns. Doctors and nurses manned consoles where they could toggle on high-definition video cameras that allowed them to zoom into any I.C.U. room and talk directly to the staff on the scene or to the patients themselves.

Technologies such as the remote ICU have enabled Steward to provide consistent care in all the ICUs in its hospitals. Steward also adopted a common medical data platform in all its hospitals and outpatient clinics.<sup>2</sup> By 2019, Steward had expanded from its six original hospitals in Boston to 38 hospitals and 271 outpatient clinics located in 11 states and Malta.<sup>3</sup>

The rise in industry concentration is due to companies similar to the Cheesecake Factory and Steward Healthcare that have adopted technologies that enable them to standardize and scale up the delivery of nontraded services. In this sense, what has happened in nontraded services is akin to the industrial revolution unleashed by Henry Ford more than a hundred years ago, when Ford introduced mass production to a car industry dominated by independent artisans.

We use microdata from the Longitudinal Business Database from 1977 to 2013, supplemented with sales data at the establishment level from the microdata of the economic censuses from 1977 to 2012, to document six main facts. First, we show that growth in the number of markets per firm has been large and heterogeneous across industries. We measure a market

<sup>1</sup> The employment-weighted share of multiestablishment hospitals in the Longitudinal Business Database increased from 15% in 1977 to 62% in 2013.

<sup>2</sup> Steward uses software by Meditech. The dominant medical software company is EPIC in Madison, Wisconsin.

<sup>3</sup> Steward's hospitals and outpatient clinics are in Massachusetts, New York, Ohio, Florida, Arkansas, Louisiana, Texas, Arizona, Pennsylvania, New Hampshire, Utah, and Malta.

as an establishment, county, zip code, or metropolitan statistical area (MSA). The growth in the number of local markets served by a typical firm has been much more pronounced outside the broad construction and manufacturing sectors, but broad sectoral classifications are imperfect. Nontraded-service industries that exhibit large expansions in markets per firm can be found in all sectors of the economy, including sectors that are classified as “manufacturing.”

Second, service industries where the number of markets per firm has increased have grown faster than other industries in the US economy. The larger growth is evident for all our definitions of a market and when we use either employment or sales. This evidence is consistent with our view that the rise of markets per firm is driven by forces such as the adoption of new technologies or management practices that ultimately raise aggregate industry total factor productivity (TFP).

Third, industries in which the number of markets per firm has increased also experienced large increases in observable fixed-cost expenditures, such as total employment in R&D and headquarters establishments. The measured elasticity of these fixed costs to establishments per firm across industries is as large as 1.5, and it is even larger with respect to MSAs per firm.

Fourth, the number of markets per firm is driven by the top firms in the industry. For example, in the industries that experienced the fastest growth in markets per firm, the top 1% of firms in an industry expanded the number of markets per firms more than twice as fast as the average firm.

Fifth, the increase in national industry concentration, documented by Autor et al. (2017) and others, is driven by the expansion in markets per firm by top firms. National employment and sales concentration, measured by the share of the top 1% or top 10% of firms or by the Herfindahl-Hirschman index (HHI), has risen much more significantly in sectors with higher establishments per firm or MSAs per firm. In fact, more than 100% (155%) of the employment expansion of the top 10% of firms in an industry is driven by an increase in the number of establishments, since the average establishment has shrunk over time. When we define a market as an MSA, this finding is less pronounced but still large: 94% of the expansion of top-10% firms is across, rather than within, MSAs.<sup>4</sup>

Sixth, the new local markets where top firms enter tend to be smaller. The share of top firms in local employment has grown significantly in small and mid-sized US cities. In contrast, in the very largest US cities, there is no change in the employment share of top firms. The increasing

<sup>4</sup> Similar results hold when we measure expansion on the basis of sales rather than employment.

presence of top firms has decreased local concentration, as the new establishments of top firms gain market share from local incumbents. The share of the top firms in the local market and the HHI have declined throughout the city distribution, but the decline has been much more pronounced for smaller cities.<sup>5</sup>

We use a simple theory of firm size and local market entry to show that a key ingredient of the industrial revolution in services, documented by our six main facts, is the availability of new fixed-cost-intensive technologies that lower the marginal cost of production in all markets served by the firm.<sup>6</sup> The adoption decision of firms involves a trade-off between a proportional reduction in all establishments' variable costs and an increase in the firm's fixed costs. Firms that adopt the new fixed-cost-intensive technology in an industry expand by serving new markets that are now viable because of their lower marginal cost. Top firms, which are more productive, find the trade-off between fixed and variable costs more beneficial, and so they adopt the new technology more intensively, which leads to a rise in industry concentration. It also leads to industry expansion relative to industries where these new technologies are less useful or more costly. For example, we show that in industries where goods are easily tradeable and so geographic replication is unnecessary (as in many manufacturing industries), firms adopt these new fixed-cost-intensive technologies less.

The industrial revolution in services has aggregate and local implications that we also corroborate in the data. Since top firms expand by entering new markets and these markets tend to be smaller, we see the share of top firms grow particularly in small markets. The increasing presence of top firms has decreased local concentration in local markets as the new establishments of top firms gain market share from local incumbents. We see the share of the largest firm in a location and the local HHI decline everywhere, but the decline is much more pronounced in small cities. Contrary to popular narratives, the entry of these top firms has been accompanied by significantly faster employment growth in small cities. As a result, we see that job destruction due to exit or incumbents' employment decline does not vary much by city size. The larger increase in the share of top firms in most cities, but most markedly in small ones, implies that consumers opted to buy from them and so probably gained from their presence. The gain from entry by top national firms into local markets is not measured in official price statistics, because current statistical procedures measure only prices from incumbent establishments. Following the methodology in Aghion et al. (2019a), we calculate "missing growth"

<sup>5</sup> Using a different data set (the National Establishment Time Series), Rossi-Hansberg, Sarte, and Trachter (2020) also find that local concentration has fallen significantly.

<sup>6</sup> Our theory is reminiscent of Gaubert (2018), but it allows firms to serve multiple local markets, as Ramondo (2014) does in an international context. See also Cao et al. (2019) and Oberfield et al. (2020).

to be 1.2% per year in the smallest cities, as low as 0.2% in the largest ones, and 0.5% in the aggregate.

Previous work has identified elements of the technological changes we underscore here. Sutton (1991) argues for the presence of new sunk-cost technologies and describes their effect on market concentration, although he does not emphasize the increasing geographic scope of firms or their resulting specialization. Hortaçsu and Syverson (2015) provide a description of the evolution of concentration and scale in the retail industry consistent with the geographic expansion we emphasize. Holmes (2011) focuses on a single firm (Walmart) and studies its geographic expansion to form a distribution network and inventory system. Similarly, Ganapati (2018) studies the wholesale industry and the expansion of the warehouse and international input use of the top firms. We view these industry studies as examples of the general evolution we document.

It is perhaps hard to set apart a number of concurrent technological changes, all of which are naturally intertwined. ICT started in the 1960s with the systematic use of corporate databases, then continued with the invention and rapid adoption of personal computers, electronic communication technologies, and the internet and the invention and subsequent explosion in the use of smartphones.<sup>7</sup> There is a vast literature on the effect of these changes on the organization of production.<sup>8</sup> The form of technological change we emphasize here was certainly enabled by ICT, at least partly, which explains its timing. The examples of fixed-cost-based technologies described above all have a component that was facilitated either by better data collection and analysis or by better communication and diffusion of information. It is undoubtedly the case that new business processes that reduce the cost of managing many different establishments require easy communication, as well as cheap data gathering and processing. Managing many hospitals and exploiting the synergies between them would be impractical without the heavy use of ICT-based systems. Thus, ICT is an essential part of the industrialization of services. It is the general-purpose technology, as defined by Rosenberg and Trajtenberg (2004), that has enabled the geographic expansion of firms (particularly in retail, services, and wholesale) by allowing them to replicate and control establishments dispersed across space. Perhaps this is where the gains from ICT have been hiding.<sup>9</sup>

<sup>7</sup> See Hobijn and Jovanovic (2001) for the diffusion of ICT-based technologies.

<sup>8</sup> A number of papers have studied the way in which ICT has changed the organization of production (Caroli and Van Reenen 2001), the decentralization of decision-making (Bresnahan, Brynjolfsson, and Hitt 2002), the span of control of managers (Garicano and Rossi-Hansberg 2006; Rajan and Wulf 2006), and the distribution of firm sizes (Garicano and Rossi-Hansberg 2004). More recently, Aghion et al. (2019b) study the growth implications of the ability of firms to manage more establishment due to improvements in ICT.

<sup>9</sup> Syverson (2017) argues that if we were mismeasuring the gains from ICT, the high-tech sector would have to be much larger than it is. If ICT is used for fixed-cost investments, as we argue, this is not necessarily the case.

Another phenomenon closely related to the new industrial revolution in services is the rise in intangible capital. As Haskel and Westlake (2017) and Crouzet and Eberly (2018) document, intangible investments became increasingly important during the period of our analysis. Intangible investments in marketing, technology, information, or training all facilitate scale and replication and as such amount to the use of new technologies with higher fixed (or sunk) costs. Hence, the rapid expansion of intangibles is a consequence of the type of technological change we suggest has occurred.

Finally, there is a large recent literature that has interpreted the increase in industry concentration as an indication of the augmented market power of top firms, perhaps facilitated by entry barriers or regulatory capture. This view has been supported by evidence that points to increasing profits and markups (Gutiérrez and Philippon 2017; De Loecker, Eeckhout, and Unger 2018) and a decrease in market dynamism (Decker et al. 2017). Together with a number of other papers in the literature (Autor et al. 2017; Hopenhayn, Neira, and Singhania 2018; Syverson 2019; Edmond, Midrigan, and Xu 2023), we argue that the industrialization of services that we document is technological, not institutional. Nevertheless, although we chose to model this process in a world with CES (constant elasticity of substitution) preferences and, therefore, fixed markups, in a model with variable markups these same technological changes could generate increases in markups. We do not focus on this dimension of the industrial revolution of services, partly because we do not have the data to estimate markups and partly because we find that the geographic expansion of top firms leads to declines in local concentration, as in Rossi-Hansberg, Sarte, and Trachter (2020).<sup>10</sup>

The rest of the paper is organized as follows. Section II describes the data sets we use and their construction. Section III presents our empirical findings, organized in six facts. Section IV presents the theory and derives the implications of the availability of a menu of new technologies offering combinations of fixed and variable costs. Section V discusses the implications of the industrial revolution in services for local outcomes and presents computations of its contribution to aggregate and local TFP growth. Section VI concludes. The appendix includes more details on our data and a number of additional empirical exercises that establish the robustness of our results, as well as the proofs of propositions in section IV.

## II. Data

Our main data set consists of the microdata from the US Census Longitudinal Business Database (LBD). The LBD is based on administrative

<sup>10</sup> The magnitude of the trend in markups is still controversial. See the discussions in Karabarbounis and Neiman (2018) and Traina (2018).

employment records of every nonfarm private establishment in the US economy. The establishment-level variables we use are employment, geographic location (county and zip code), industry (4-digit standard industrial classification codes from 1977 to 2000, 6-digit NAICS [North American Industry Classification System] codes from 2001 to 2013, and 6-digit 2002 NAICS codes provided by Fort and Klimek [2018] from 1977 to 2013), the establishment's ID, and the ID of the firm that owns the establishment. We restrict the sample to observations from 1977 to 2013 and drop establishments in the public, educational, agricultural, and mining sectors.<sup>11</sup>

We aggregate the 2002 NAICS industry classifications provided by Fort and Klimek (2018) into 445 consistently defined industries from 1977 to 2013. Hereafter, when we refer to an industry we mean one of these 445 industries. Appendix G provides additional details on Fort and Klimek's (2018) industry classifications and how we aggregate them into our industry codes.

We group counties into metropolitan areas (MSAs) defined on the basis of the 1980 Population Census.<sup>12</sup> For the counties that were not part of an MSA in 1980, we group them into "pseudo-MSAs" corresponding to their respective states. We end up with a total of 329 MSAs. We therefore have four measures of local markets: establishments, zip codes, counties, and MSAs (based on the 1980 census).

We supplement the LBD with sales data at the establishment level from the microdata of the economic censuses taken every 5 years from 1977 to 2012. Specifically, we use the microdata from the Censuses of Auxiliary Establishments, Construction Industries, Manufactures, Retail Trade, Services, and Wholesale Trade for every 5 years from 1977 to 2012. We also use the microdata of Census of Finance, Insurance, and Real Estate every 5 years from 1992 to 2012 and the Census of Transportation, Communications, and Utilities for every 5 years from 1987 to 2012. We use the establishment ID to match the establishments in the economic censuses to the establishments in the LBD. Our final sample from the economic censuses consists of establishments with sales data that are matched to the LBD. Appendix A shows the summary statistics of our samples from the LBD and the economic censuses.

We use the establishment's firm ID to do two things. First, we use the firm ID to aggregate employment and sales of establishments to a firm in an industry.<sup>13</sup> Second, we use the firm ID to measure employment in

<sup>11</sup> We also drop commercial banking (2002 NAICS code 522110) because the 1994 Riegle-Neal interstate banking law removed restrictions on interstate banking.

<sup>12</sup> <https://www2.census.gov/programs-surveys/metro-micro/geographies/reference-files/1983/historical-delineation-files/83mfips.txt>.

<sup>13</sup> For establishments that are franchises, the firm ID in the LBD refers to the owner and not the franchisee. For such establishments, using the firm ID will understate the share of firms for which franchising is an important margin.

establishments that provide R&D and headquarters services for the firm's establishments in a given industry. Specifically, we identify a firm's R&D centers and "headquarters" (HQ) as establishments with Fort and Klimmek's (2018) NAICS codes beginning with 54 (R&D) and 55 (HQ) that have the same firm ID. For firms with establishments in multiple industries (outside of R&D and HQ), we split employment in R&D and HQ into the industries served by the firm, using the firm's employment share in each industry (omitting employment in R&D and HQ).

From these data, we calculate the change from 1977 to 2013 in three variables at the industry level:

1. number of markets (establishment, county, zip code, and MSA) per firm in an industry; we focus on the number of markets served by an average firm in an industry and by the top 1% and top 10% of firms in the industry, where top firms are defined by the number of markets they serve;
2. total sales and employment of all firms in an industry and total employment in R&D centers and HQs of these firms; and
3. economic concentration in an industry, measured as the sales and employment share of the top 1% and top 10% of firms in an industry, where the top firms are defined by sales and employment, and the HHI, measured by sales and employment.

We weight the variables at the industry level by the Sato-Vartia weights of each industry between 1977 and 2013.<sup>14</sup>

We also calculate changes from 1977 to 2013 in two variables at the MSA-industry level:

1. employment share in an MSA-industry of top firms, where "top" is measured by the number of establishments or employment in the industry in all MSAs; and
2. employment and sales concentration in an MSA-industry, measured as the employment and sales share of the top firm in the MSA-industry and the employment and sales HHI in an MSA-industry.

We aggregate the variables at the MSA-industry level to the MSA level, using the Sato-Vartia weights of the MSA-industry in 1977 and 2013. Appendix A shows the summary statistics of these industry- and MSA-level statistics.

<sup>14</sup> The Sato-Vartia weight of industry  $j$  is defined as  $(\Delta L_j / \Delta \log L_j) / \sum_{k=1}^J (\Delta L_k / \Delta \log L_k)$ , where  $J$  denotes the set of industries,  $L_j$  denotes employment in industry  $j$ , and  $\Delta$  is the change between 1977 and 2013.

### III. Facts

We highlight six facts from the LBD and economic census data.

#### A. *Fact 1: Growth in Markets per Firm Has Been Large and Heterogeneous across Industries*

Our first fact is the increase in the number of markets per firm. The left-hand panel of figure 1 shows that the number of establishments per firm grew by 0.093 log points from 1977 to 2013 in the median 4-digit industry. This increase in establishments per firm was not uniform across industries. The top quintile of industries, with the fastest increase in establishments per firm, saw an increase of more than 0.4 log points in the same period, while the bottom quintile saw a reduction of about 0.12 log points. The same pattern is evident in the right-hand panel of figure 1, where we plot the change in the number of MSAs in which a firm is present. Again, we see an increase in the number of MSAs per firm in the median industry between 1977 and 2013, albeit smaller than that for establishments, and a large increase of more than 0.15 log points for the top quintile.<sup>15</sup>

Table 1 presents the change in log average markets per firm in each 1-digit sector. The expansion in the number of markets per firm was fastest in finance, retail, and “other” services (which includes industries such as business services, restaurants, gyms, and health care) and slowest, on average, in construction and manufacturing. The table presents four different geographic units, going from individual establishments to zip codes, counties, and MSAs. All of these measures show similar patterns. Thus, in what follows, we present results for establishments and MSAs in the main text and relegate results for zip codes and counties to appendix B.

The expansion in the number of markets per firm varies tremendously across broad sectors but also across industries within them. This can be seen in the large standard deviation in the change in markets per firm within 1-digit sectors in table 2.<sup>16</sup> Figure 2 adds to this evidence by showing the cumulative distribution function (CDF) of log changes in markets per firm. It shows that all sectors include some “service” industries where the expansion in the number of markets per firm has been substantial.

<sup>15</sup> Figure B1, in app. B, shows that a similar pattern holds for the number of counties and zip codes per firm.

Consistent with our findings, Cao et al. (2019) use data from the Quarterly Census of Employment and Wages between 1990 and 2015 to document an increase in the average number of establishments per firm. They also show that the increase is more pronounced for larger firms and in the service sector.

<sup>16</sup> A regression of the change in log establishments per firm of the average firm in a 4-digit industry on indicator variables for 1-digit sector has an  $R^2$  of 0.124. A similar regression of the change in log MSAs per firm on indicator variables for 1-digit sector has an  $R^2$  of 0.039.

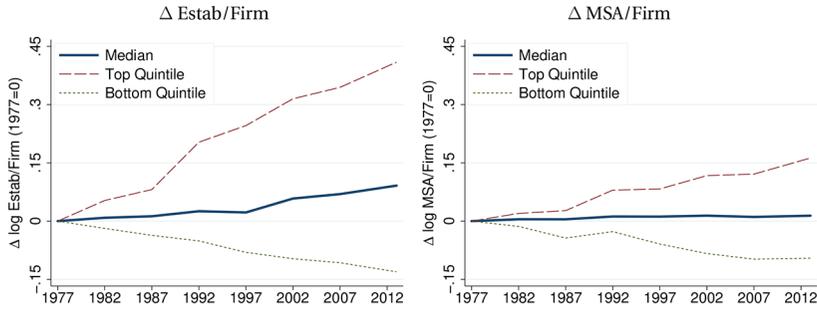


FIG. 1.—Unit of observation is a 4-digit industry ( $N = 445$ ). The figure shows cumulative change from 1977 to 2013 of log establishments/firm and MSAs/firm of the average firm in the median industry, top-quintile industry, and bottom-quintile industry, weighted by Sato-Vartia employment share of each 4-digit industry in 1977 and 2013.

Naturally, sectors such as other services, wholesale, retail, utilities and transportation, and finance include many more of these service industries. For example, in retail, about 43% of industries expanded the number of establishments per firm by more than 0.425 log points between 1977 and 2013. The large heterogeneity within 1-digit sectors in the change in markets per firm indicates that it is inaccurate to simply define an industry as a nontradable service according to the sector to which it belongs. For example, four manufacturing industries are among the top quintile

TABLE 1  
WEIGHTED AVERAGE AND STANDARD DEVIATION OF  $\Delta \text{LOG}$  MARKETS  
PER FIRM BY SECTOR, 1977–2013

	Establishments	Zip Codes	Counties	MSAs
Construction	.016 (.034)	.017 (.031)	.015 (.028)	.012 (.020)
Manufacturing	.019 (.141)	.017 (.132)	.012 (.115)	.006 (.089)
Other	.180 (.229)	.128 (.200)	.081 (.150)	.050 (.094)
Wholesale	.156 (.248)	.139 (.239)	.076 (.156)	.030 (.084)
Retail	.216 (.237)	.186 (.185)	.096 (.126)	.040 (.078)
Utilities and transportation	.172 (.234)	.126 (.202)	.101 (.180)	.070 (.148)
Finance	.299 (.215)	.211 (.170)	.099 (.137)	.044 (.109)

NOTE.—Table shows weighted average and standard deviation of  $\Delta \text{log}$  establishments/firm, zip codes/firm, counties/firm, and MSAs/firm of the average firm in each 4-digit industry within each 1-digit sector, weighted by Sato-Vartia average of the employment share of each 4-digit industry in 1977 and 2013. “Other” includes 4-digit industries not included in the other 1-digit sectors.

TABLE 2  
REGRESSION OF INDUSTRY GROWTH ON  $\Delta$ LOG MARKETS PER FIRM, 1977–2013

	$\Delta$ Log Employment	$\Delta$ Log Sales
$\Delta$ Log establishments/firm	.845 (.169)	1.192 (.206)
$\Delta$ Log MSAs/firm	1.444 (.415)	2.926 (.468)

NOTE.—Unit of observation is a 4-digit industry ( $N = 445$ ). Table shows coefficient estimates and standard errors from weighted regressions of  $\Delta$ log aggregate employment (col. 1) and sales (col. 2) in the industry on  $\Delta$ log establishments/firm (row 1) and MSAs/firm (row 2) of the average firm in the industry. Columns 1 and 2 are from 1977 to 2013. Sales growth is from 1977 to 2012, and growth in the number of markets per firm is from 1977 to 2013, except for utilities and transportation and finance, where sales are from 1987 to 2012 and 1992 to 2012 and change in markets per firm are from 1987 to 2013 and 1992 to 2013, respectively. Weights are Sato-Vartia average of industry employment in 1977 and 2013.

of industries with the largest change in the number of establishments per firm.<sup>17</sup> Hence, our approach is to measure whether firms in an industry provide local services using the observed change in the number of local markets per firm between 1977 and 2013. That is, the change in markets per firm is our metric for the extent to which firms in an industry want to be close to their customers and, therefore, the extent to which they are affected by the industrial revolution in services.

We also consider the share of counties and MSAs where a given industry has a physical presence in 1977 as an alternative measure of whether firms in an industry provide local services. This measure is not perfect because, as we show below, the technological revolution for replicating nontraded services over a large number of markets also shows up on the extensive margin in the form of the entry of an industry in a locality. Nonetheless, we find that our preferred measure of an industry that provides local services, namely, the change in markets per firm between 1977 and 2013, is highly correlated with the share of counties or MSAs where the industry had a physical presence in 1977. Specifically, a regression of the change in log of average establishments per firm of an industry between 1977 and 2013 on the share of counties where the industry has a physical presence in 1977 yields a coefficient of 0.187 (SE = 0.042). A similar regression on the share of MSAs in 1977 where the industry is present yields a coefficient of 0.193 (SE = 0.036).

Some of the change in the average number of markets per firm, and the dispersion in the change in markets per firm across sectors shown in figures 1 and 2, could be due to the combination of life-cycle growth in

<sup>17</sup> These industries are lime production, asphalt and tar paving materials, concrete manufacturing, and cosmetics and hygiene products.

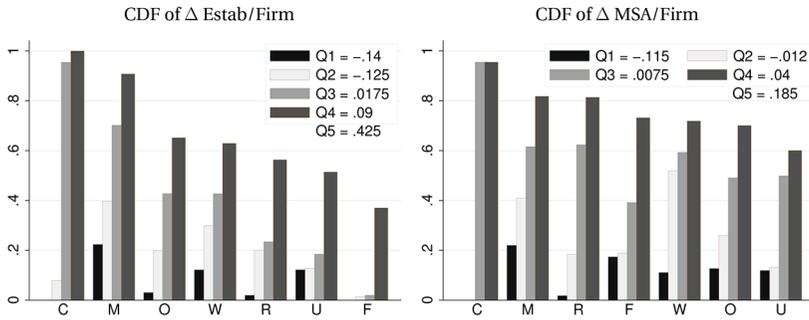


FIG. 2.—Sectors are construction (C), manufacturing (M), other (O), wholesale trade (W), retail trade (R), utilities and transportation (U), and finance, insurance, and real estate (F). “Other” includes 4-digit industries not included in the other 1-digit sectors. Figure shows cumulative distribution of the change in log establishments/firm and MSAs/firm from 1977 to 2013 across 4-digit industries in each broad sector in quintiles of the change of establishments/firm and MSAs/firm (also from 1977 to 2013) across all 445 4-digit industries. Average change in log establishments/firm in each quintile is  $-0.14$ ,  $-0.125$ ,  $0.0175$ ,  $0.09$ , and  $0.425$  in quintiles 1–5, respectively. The corresponding change in log MSAs/firm in each quintile is  $-0.115$ ,  $-0.012$ ,  $0.0075$ ,  $0.04$ , and  $0.185$ . Top quintile is omitted in the figure.

markets per firm with changes in the age distribution between 1977 and 2013. For example, in the sample of incumbent firms between 1977 and 2013 (defined as firms that exist in the same industry in the two years), the average number of establishments per firm increased by 0.027 log points over this period. If all firms in 1977 were entrants, which is obviously not the case, the “aging” effect could account for about 30% of the increase in the number of establishments per firm of the median industry shown in figure 1. Unfortunately, since the LBD does not directly measure age, we do not know whether the average firm is older in 2013, compared to 1977. We can, of course, impute firm age from the year the firm shows up in the data for the first time. Hence, we can impute firm age reasonably well in later years, but not in 1977, the first year of the data.<sup>18</sup> That is, we cannot measure the change in the average age of firms in the sector, but we can impute the average age of firms in a sector in 2013. We can then check whether the dispersion across sectors in the growth in markets per firm between 1977 and 2013 reflects differences in the average age of firms in a sector in 2013. A regression of the change in log establishments per firm between 1977 and 2013 of a 4-digit sector on the average age of firms in the sector in 2013 essentially yields a coefficient of zero (0.001, with a standard error of 0.003). Hence, we conclude that aging does not seem to be driving our first fact.

<sup>18</sup> It is clearly the case that entry rates have fallen since the late 1970s, but the change in the age distribution depends on the trend in the entry rate before the late 1970s, as well as on changes in the exit rate of incumbent firms.

*B. Fact 2: Industries Where Markets per Firm Increased Grew Faster*

Fact 1 showed that the average firm in service industries has increased significantly the number of markets they serve. Our next fact shows that these industries have also expanded in terms of total employment and sales. Figure 3 presents a kernel regression of the relationship between the change in the log of markets per firm and the change in log employment and sales between 1977 and 2013.<sup>19</sup> The left-hand panel uses establishments as the definition of a market, while the right-hand one uses MSAs. The figure shows clearly that these elasticities are significantly positive and roughly similar throughout the range of industries. Furthermore, the results are almost identical for employment and sales. We do note that the number of industries declines substantially, and therefore the standard errors grow, when we look at industries with expansions in the number of establishments per firm larger than 0.4 log points (0.2 for MSAs per firm). Fact 2 implies that the changes experienced by these industries, which motivated firms to expand their number of markets, have also made the industries larger. This expansion is consistent with positive technological innovations in service industries.<sup>20</sup>

Table 4 summarizes the results in figure 3 when we impose a constant elasticity between the change in log industry employment or sales and the change in log markets per firm. As we noted above, all these elasticities are positive and significant. Both for employment and sales, the elasticity is larger when we define the number of markets using MSAs instead of establishments. This is natural, if the underlying industry change is a technological innovation; expanding across MSAs requires an innovation larger than the one needed to add another local store, and so the innovation also implies a larger expansion of the industry.

*C. Fact 3: Total Employment in R&D and HQ Establishments Grew in Industries Where Markets per Firm Increased*

Incorporating the technological and management innovations that allow firms to expand the number of markets comes at the cost of increasing fixed production costs. Measuring these costs precisely is hard, since the distinction between fixed and marginal costs is conceptual and not directly observable. Firm-level fixed costs should include firm expenditures that benefit all establishments and so are not rival. Two natural examples

<sup>19</sup> The change in log sales is from 1977 to 2012, except for finance (1992–2012) and utilities and transportation (1987–2012).

<sup>20</sup> Of course, service industries might have expanded also because of other concurrent changes. For example, non-homotheticities in preferences for services would result in larger services expenditure shares as incomes grow.

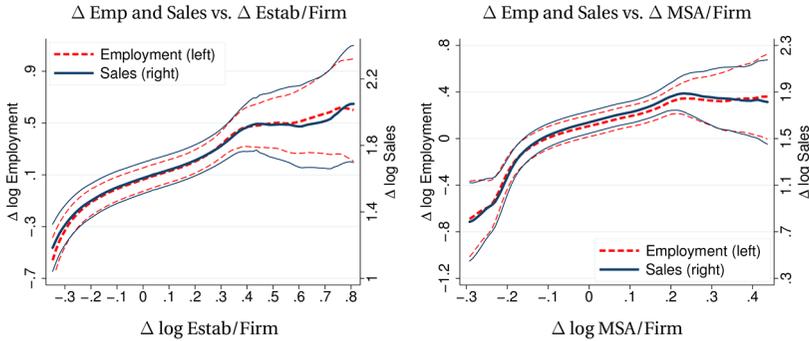


FIG. 3.—Unit of observation is a 4-digit industry ( $N = 445$ ). Figure shows point estimates and 95% confidence intervals of a nonparametric regression of  $\Delta \log$  total employment or total sales of the industry on  $\Delta \log$  establishments/firm (left) and  $\Delta \log$  MSAs/firm (right) of average firm in the industry. Regressions with employment growth use change from 1977 to 2013 for all variables. Regressions with sales growth are from 1977 to 2012 for sales growth and from 1977 to 2013 for the change in markets/firm, except for utilities and transportation and finance, where sales are from 1987 to 2012 and from 1992 to 2012 and change in markets per firm is from 1987 to 2013 and from 1992 to 2013, respectively.

are a firm's management and its R&D expenditures. Management decisions, as well as new designs or product innovations, can be used repeatedly across establishments without depleting them. Of course, the more useful these fixed-cost-intensive technologies are in lowering the cost of operating local establishments, the more firms will choose to adopt them and increase their observed fixed costs.

Consider again the case of the Cheesecake Factory. Between 2003 and 2018, employment in the company grew rapidly, from 14,200 to 38,100 employees. This tremendous growth was accompanied by a large expansion in the number of establishments, from 61 to 214, which resulted in a reduction in the average number of employees per establishment, from 233 to 178. The evolution of the company HQ is quite different, though. The number of employees in HQ establishments, a measure of the firm's fixed costs, grew from 140 to 450.<sup>21</sup> We interpret this large expansion in HQ employees as the firm's investment in fixed-cost-intensive technologies.

We can measure a firm's employment in establishments classified as either doing R&D or serving as the firm's HQ. Figure 4 presents a kernel regression of the relationship between changes in log employment in R&D and HQ establishments serving all firms in an industry against changes in log markets per firm in the industry. It shows that this elasticity is, indeed, positive throughout. Namely, industries where firms were motivated to

<sup>21</sup> This information comes from the 2003 and 2018 Form 10-K of The Cheesecake Factory Inc.

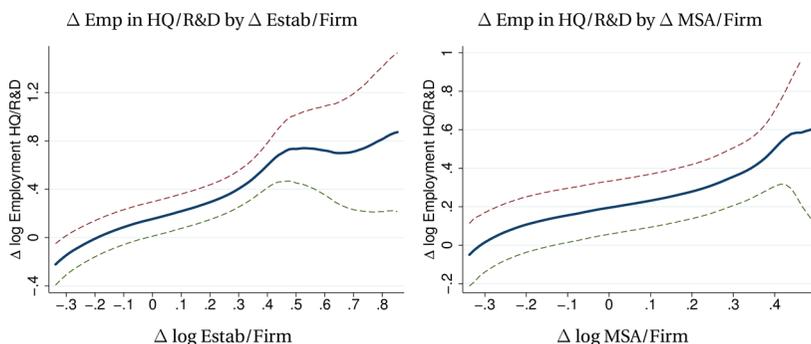


FIG. 4.—Unit of observation is a 4-digit industry ( $N = 445$ ). Figure shows coefficients and 95% confidence intervals of a nonparametric regression of  $\Delta \log$  aggregate employment in HQ and R&D of establishments in an industry from 1977 to 2013 on  $\Delta \log$  establishments (*left*) or MSAs (*right*) per firm of the average firm in the industry from 1977 to 2013. See text for details on how we identify R&D and HQ establishments of firms in each industry.

expand the number of markets rapidly, our service industries, also spent more on these two measures of fixed costs.

Table 3 presents our estimates of the elasticity of fixed costs with respect to markets per firm when we impose the restriction that the elasticity is constant across industries (a reasonable restriction, given the results in fig. 4). The table also presents estimates of elasticities for employment in R&D and HQ establishments separately. The elasticity is about the same for these two types of fixed costs, independently of the definition of a market. In contrast, the market definition is important. The elasticity doubles in magnitude when we use MSAs rather than establishments. Again, this is consistent with larger technological or management innovations in service industries motivating firms to expand their presence across cities and

TABLE 3  
REGRESSION OF  $\Delta \text{LOG}$  EMPLOYMENT IN HQ/R&D ON  $\Delta \text{LOG}$  MARKETS PER FIRM, 1977–2013

	$\Delta \text{Employment,}$ R&D and HQ	$\Delta \text{Employment,}$ R&D Only	$\Delta \text{Employment,}$ HQ Only
$\Delta \text{Log establishments/firm}$	1.491 (.266)	1.595 (.290)	1.772 (.257)
$\Delta \text{Log MSAs/firm}$	3.520 (.667)	3.697 (.728)	4.052 (.648)

NOTE.—Unit of observation is a 4-digit industry ( $N = 445$ ). Table shows coefficient estimates and standard errors from weighted regression of  $\Delta \log$  aggregate employment in R&D and HQ (col. 1), R&D only (col. 2), and HQ only (col. 3) of all firms in the industry on  $\Delta \log$  establishments/firm (row 1) and MSAs/firm (row 2) of the average firm in the industry, all from 1977 to 2013. See text for details on how we identify R&D and HQ establishments of firms in each industry. Weights are Sato-Vartia average of the employment share of the industry in 1977 and 2013.

not only within them. These larger innovations, in turn, motivate firms to invest more in fixed-cost-intensive technologies in order to reduce the cost of operating in the larger number of markets.

*D. Fact 4: Growth in Markets per Firm Is Driven by Top Firms in the Industry*

The increase in the number of markets per firm that we have documented so far has been much more pronounced for the top firms in an industry.<sup>22</sup> Here, we measure top firms by the number of markets in which they operate, but defining top firms by total sales or employment yields similar results. Figure 5 presents the nonparametric relationship between the average change in log markets per firm for the top 1% or 10% of firms in the industry and the average change in log markets per firm for all firms in an industry. As the figure clearly illustrates, the slope of the positive relationship is larger than 1 in all cases (the dashed green line is the 45° line). Namely, in industries where we see a large expansion in the number of markets per firm on average, we see a larger expansion for top-10% firms and an even larger expansion for top-1% firms. This implies that the increase in the average is driven by the top firms. In fact, in appendix D we show that the elasticity of markets per firm is large and significantly positive only for firms in the ninth and tenth deciles of the distribution of markets per firm (for both establishments and MSAs).

Table 4 present results when we estimate a constant elasticity across industries. The elasticity of establishments per firm for top-10% firms to establishments per firm for all firms is 2, and it grows to 2.4 for top-1% firms. For MSAs per firm, the elasticities are much larger, so growth in MSAs per firms is even more skewed toward the largest firms. The elasticity of MSAs per firm for top-1% firms to MSAs per firm for all firms is as large as 4.5.

*E. Fact 5: The Increase in National Industry Concentration Is Driven by the Expansion of Markets per Firm by Top Firms*

It is well known that many industries in the United States have experienced concentration of employment and sales since the late 1970s (as documented by Autor et al. 2017 and Rossi-Hansberg, Sarte, and Trachter 2020, among many others).<sup>23</sup> This increase in concentration is particularly

<sup>22</sup> Figure C1, in app. C, shows the change in markets per firm of top firms in the average industry.

<sup>23</sup> Figure C2, in app. C, shows that the employment share of top-10% firms in each industry (measured by employment) increased by almost 0.1 log points between 1977 and 2013. The employment share of top-1% firms increased by about 0.2 log points over the same period.

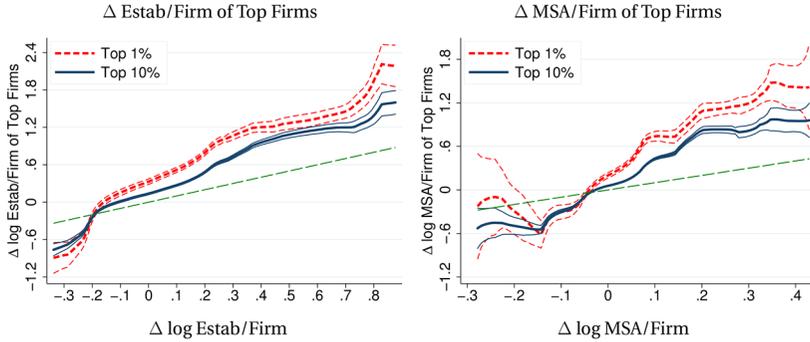


FIG. 5.—Change in log markets per firm: top firms versus all firms. Unit of observation is a 4-digit industry ( $N = 445$ ). Top firms are defined by number of establishments (*left*) or MSAs (*right*) of the firm. Figure shows point estimates and 95% confidence intervals of nonparametric regressions of the change from 1977 to 2013 of log establishments/firm or MSAs/firm of the top 1% and top 10% of firms in the industry on the change in log establishments/firm or MSA/firm of average firm in the industry, also from 1977 to 2013. Green dashed line is the 45° line.

large in service industries where, as we showed in fact 4, top firms have increased rapidly the number of markets in which they operate. We show next that the growth in industry concentration is, in fact, mostly due to the growth in markets per firm of top firms. Table 5 presents our estimates for the relationship between employment concentration and the change in log markets per firm. We present results for three alternative concentration measures (the share of top-1% firms, the share of top-10% firms, and the HHI) as well as measures of concentration using either employment or sales. When we use measures of concentration based on employment,

TABLE 4  
REGRESSION OF  $\Delta \text{LOG MARKETS PER FIRM OF TOP FIRMS}$  ON  $\Delta \text{LOG MARKETS PER FIRM OF ALL FIRMS}$

	$\Delta \text{LOG MARKETS PER FIRM OF TOP FIRMS}$	
	Top 1%	Top 10%
$\Delta \text{Log establishments/firm}$	2.309 (.098)	2.060 (.046)
$\Delta \text{Log MSAs/firm}$	4.343 (.182)	3.331 (.060)

NOTE.—Unit of observation is a 4-digit industry ( $N = 445$ ). Top firms are defined by establishments per firm in row 1 and MSAs per firm in row 2. Entries are coefficient estimates and standard errors from weighted regression of  $\Delta \text{log establishments per firm}$  (row 1) or  $\Delta \text{log MSAs per firm}$  (row 2) of top firms in the industry on  $\Delta \text{log establishments/firm}$  (row 1) or  $\Delta \text{log MSAs/firm}$  (row 2) of the average firm in the industry. Weights are Sato-Vartia averages of the industry's employment share in 1977 and 2013.

TABLE 5  
REGRESSION OF  $\Delta$ EMPLOYMENT AND SALES CONCENTRATION ON  $\Delta$ LOG MARKETS PER FIRM

	EMPLOYMENT CONCENTRATION			SALES CONCENTRATION		
	Top 1% <sup>a</sup>	Top 10% <sup>a</sup>	HHI	Top 1% <sup>b</sup>	Top 10% <sup>b</sup>	HHI
$\Delta$ Log establishments/firm	.665 (.062)	.326 (.023)	.086 (.014)	.358 (.072)	.244 (.026)	.053 (.022)
$\Delta$ Log MSAs/firm	1.317 (.161)	.592 (.062)	.104 (.036)	.679 (.168)	.477 (.062)	.089 (.051)

NOTE.—Unit of observation is an industry ( $N = 445$ ). Entries are point estimates and standard errors of weighted regressions of the change in employment concentration (log share of top 1%, log share of top 10%, and HHI in cols. 1–3) or sales concentration (log share of top 1%, log share top 10%, and HHI in cols. 4–6) on the change in log markets/firm (establishments in row 1 and MSAs in row 2) of the average firm in each industry. Employment-based concentration regressions use 1977–2013 for all variables. Sales-based concentration regressions use the change from 1977 to 2012 for concentration and that from 1977 to 2013 for growth in markets per firm, except for utilities and transportation and finance, where concentration changes are from 1987 to 2012 and 1992 to 2012 and changes in markets per firm are from 1987 to 2013 and 1992 to 2013, respectively. Weights are Sato-Vartia average of employment share of the industry in 1977 and 2013.

<sup>a</sup> Top 1% and 10% of firms by employment.

<sup>b</sup> Top 1% and 10% of firms by sales.

the relationship is positive and highly significant. As with other facts, it is larger for MSAs per firm, indicating that cross-city expansion reveals a larger underlying industry-level change. The results for sales are also positive and mostly significant, although a bit smaller and more noisy.

The change in the employment share of the top firms in an industry can be decomposed into the contribution of the relative growth in the number of markets per firm of the top firms and the change in the relative average employment size of these markets for top firms. For example, if we define a market as an MSA, the decomposition is given by

$$\Delta \log \frac{L_{\text{top}}}{L} = \Delta \log \frac{\# \text{MSA}_{\text{top}}}{\# \text{MSA}} + \Delta \log \frac{L_{\text{top}} / \# \text{MSA}_{\text{top}}}{L / \# \text{MSA}}. \quad (1)$$

The first term in equation (1) is the contribution from growth in the number of MSAs of the top firms, and the second term is the contribution from changes in employment per MSA of the top firms (both relative to all firms in the industry). The first two columns of table 6 show the results of this decomposition for the relative number of establishments versus employment per establishment (row 1) and the relative number of MSAs versus employment per MSA (row 2).<sup>24</sup> The last two columns show the

<sup>24</sup> Specifically, the first two columns of table 6 show the decomposition of the variance of  $\Delta \log(L_{\text{top}}/L)$  into the variance of  $\Delta \log(\# \text{Estab}_{\text{top}}/\# \text{Estab})$  and the variance of  $\Delta \log((L_{\text{top}}/\# \text{Estab}_{\text{top}})/(L/\# \text{Estab}))$ , where the contribution of the covariance between the last two terms is equally split between the two terms.

TABLE 6  
CHANGE IN SHARE OF TOP 10% FIRMS BY EMPLOYMENT OR SALES:  
NUMBER OF MARKETS VERSUS AVERAGE SIZE

	EMPLOYMENT SHARE		SALES SHARE	
	Markets (1)	Size (2)	Markets (3)	Size (4)
Establishments	1.522 (.092)	-.522 (.092)	1.289 (.094)	-.289 (.094)
MSA	.941 (.072)	.059 (.072)	.789 (.073)	.211 (.073)

NOTE.—Unit of observation is a 4-digit industry ( $N = 445$ ). Column 1 (“Markets”) shows point estimates and standard errors from a regression of  $\Delta \log$  establishments or MSAs of top 10% firms (measured by employment) relative to all firms from 1977 to 2013 on  $\Delta \log$  employment share of top 10% firms from 1977 to 2013. Column 2 (“Size”) shows the results from a regression of  $\Delta \log$  employment per establishment or MSA of top 10% firms relative to all firms from 1977 to 2013 on the same independent variable. Columns 3 and 4 show similar regressions using the  $\Delta \log$  sales share or  $\Delta \log$  sales per establishment or MSA of the top 10% firms measured by sales. The changes in sales are from 1977 to 2012, except for utilities and transportation and finance, where the changes are calculated from 1987 to 2012 and 1992 to 2012, respectively.

same decomposition using sales rather than employment. The first row shows that average employment per establishment of top firms falls by more than 0.5 log points and that average sales per establishment of top firms falls by almost 0.3 log points. Thus, necessarily, more than 100% of concentration growth has to come from the increase in the number of establishments served by the top firms. The second row shows that, for MSAs, most of the growth in concentration also comes from growth in the number of cities served by top firms. Only about 6% of the growth in concentration comes from increased employment per city, and about 21% comes from increased sales per city.

Figure 6 plots the nonparametric relationship between changes in concentration, as measured by the change in the log employment share of top-10% firms, and changes in the log number of markets of top-10% firms relative to all firms (*left*) or changes in the log average size per market of top 10% firms relative to all firms (*right*). The slopes of both curves in the left-hand panel of figure 6 are positive, indicating that in industries where top firms have expanded the most, they have done so by expanding geographically through more establishments or by reaching more MSAs. Note that the slope increases as we adopt narrower definitions of a market. It is the smallest for MSAs and the largest for establishments. The right-hand panel shows that the opposite is true for changes in employment per market. Namely, the relationship with the change in employment of top firms is negative. In sum, these results show that the variation in the change in concentration across industries is entirely driven by variation across industries in the expansion of top firms into new markets.

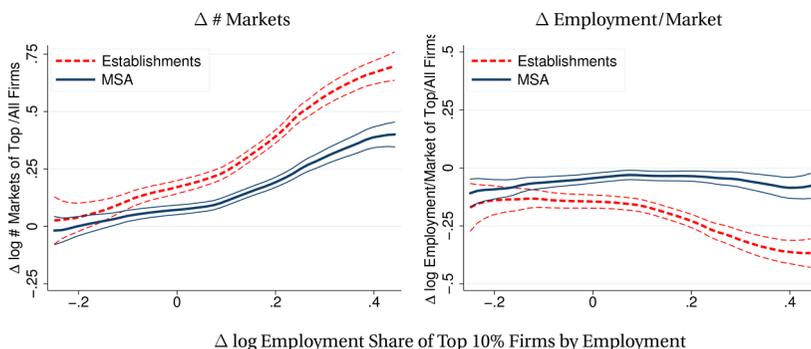


FIG. 6.—Growth of top firms: number of markets versus average size. Unit of observation is a 4-digit industry ( $N = 445$ ). Top firms are defined as the top 10% of firms in an industry, measured by employment. Figure shows point estimates and 95% confidence intervals of nonparametric regressions of  $\Delta \log$  markets of top firms relative to all firms (*left*) and  $\Delta \log$  employment/market of top firms relative to all firms (*right*) against  $\Delta \log$  employment share of top-10% firms in the industry, all from 1977 to 2013. A market is an establishment or an MSA.

#### F. Fact 6: Top Firms Have Expanded Their Presence in Small MSAs

We have shown that the number of markets per firm has increased on average and, particularly, in nontraded-service industries. We have also shown that this increase is accompanied by an expansion in industry employment and sales as well as by an expansion in HQ and R&D employment. Furthermore, the expansion is driven by the top firms in the industry and has generated increases in employment and sales concentration. We now ask where these top firms have added new markets.

Table 7 probes for evidence that top firms have expanded into smaller and more marginal local markets. Specifically, we measure the size of the

TABLE 7  
CHANGE IN MARKET SIZE OF TOP-10% FIRMS

	TOP-10% FIRMS	
	By Employment (1)	By Establishments/Firm (2)
$\Delta$ MSA size of top firms/all firms	-.262 (.065)	-.059 (.014)

NOTE.—Unit of observation is a 4-digit industry ( $N = 445$ ). Market size of a firm is total employment (in all industries) in the MSA, averaged across all the MSAs in which the firm operates. Top firms in an industry are defined by employment (col. 1) or number of establishments (col. 2). Column 1 shows a weighted regression of  $\Delta \log$  ratio of the market size of the top 10% of firms to the market size of the average firm in the industry on  $\Delta \log$  employment share of the top 10% of firms. Column 2 shows a weighted regression of  $\Delta \log$  ratio of the market size of the top 10% of firms to the market size of the average firm in the industry on  $\Delta \log$  establishments per firm of the top 10% firms in the industry. All variables are from 1977 to 2013. Weights are Sato-Vartia average of industry employment in 1977 and 2013.

local market as total employment (in all industries) in the MSA. The size of a firm's local market is then the average size of all the local markets in which a given firm has an establishment. Table 7 shows the regression of  $\Delta \log$  ratio of the size of the local market of a top firm in the industry to that of an average firm in the industry on  $\Delta \log$  employment share of the top-10% firms (measured by employment) in the industry (both are calculated from 1977 to 2013). The first column defines a top firm as the top 10% of firms in an industry as measured by their employment; the second column defines a top firm as the top 10% of firms as measured by the number of establishments. Table 7 shows that the elasticity of the change in the relative size of the market of top firms with respect to the change in the market share of top firms is negative and precisely estimated. So top firms, on average, expand by entering into smaller MSAs.<sup>25</sup> Of course, the expansion patterns of specific industries might look different. For example, Holmes (2011) shows that Walmart grew by expanding into new local markets that are typically close to its HQ and larger than its existing markets.

We next directly show the employment share of top national firms in each MSA in 1977 and 2013. Figure 7 plots the average share of employment of the top 10% of national firms in an industry in each MSA by total employment of the MSA in 1977. The left-hand panel defines top firms by their industry employment, while the right-hand panel defines top firms by the number of establishments, as we have done above. In 1977, this share was markedly lower in small cities than in large ones. In contrast, by 2013, the presence of top firms varies significantly less across markets. Small cities in 1977, such as Missoula, Montana (employment 19,000 in 1977), have seen enormous entry of establishments of top firms, while large cities, such as Washington DC (employment 1.2 million in 1977), have seen no significant increase in the share of top firms operating in the city. Clearly, top firms have increased the number of markets per firm by entering much more aggressively into small cities. The observed changes in the share of top firms between 1977 and 2013 are extremely large: the share of top-10% firms has increased by about 15 percentage points for the smallest cities but not at all for the largest ones (independently of the measure of top firms we use).

Put together, these facts paint, we believe, a consistent picture. The industrial revolution in services has affected service industries by providing fixed-cost-intensive technologies that lower the cost of operating in

<sup>25</sup> Table B6 shows that the same pattern holds when a local market is defined as a county or a zip code.

Here, the effect of "aging" works in the opposite direction. In the sample of incumbent firms between 1977 and 2013, the size of the average MSA served by the firm increases by 0.3 log points between 1977 and 2013, and the size of the firm's average establishment increases by 0.28 log points.

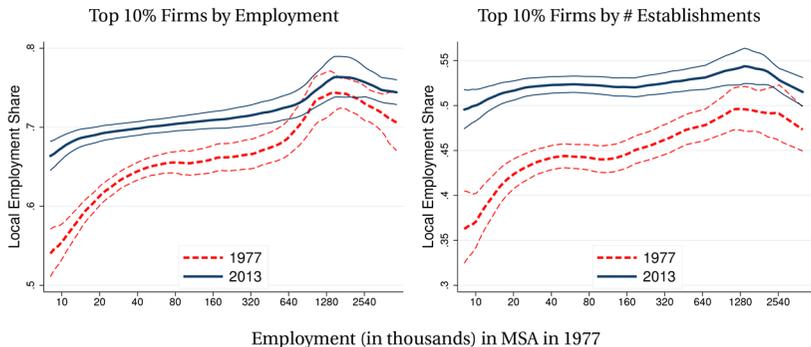


FIG. 7.—Local employment share of top-10% firms. Unit of observation is an MSA ( $N = 329$ ). Top-10% firms are defined by total employment (*left*) or number of establishments (*right*) in the industry (and in all MSAs in the country). Figure shows coefficients and 95% confidence intervals of nonparametric regressions of the weighted average of the employment share of the top 10% of firms in each MSA-industry in 1977 or 2013 on the log of employment in the MSA in 1977. Weights are Sato-Vartia average of the industry in each MSA in 1977 or 2013.

individual geographic markets, particularly for high-productivity firms in the industry. These firms, the top firms, have expanded the number of markets in which they operate. This expansion is accompanied by an overall increase in the size of the industry and an increase in industry national concentration, as well as more investments in fixed costs as measured by employment in R&D and HQ establishments. The expansion of top firms has made them enter more marginal markets in smaller cities. We now propose a formal model that makes more precise this narrative about the nature and implications of the industrial revolution in services.

#### IV. A Simple Model of Firm Size and Market Entry

Our aim in this section is to propose a simple theory of firm production decisions that is rich enough to speak to the facts in the previous section. To do so, we need to propose a theory in which firms can enter multiple markets but where the most productive firms enter more markets and reach more marginal smaller markets. That is, we want to propose a model akin to standard heterogeneous-firms models, but with profits that are supermodular in the productivity of the firm and the number of markets it reaches. The main purpose is then to use the theory to define precisely a form of technological change and trace its implications. This new technology is, we believe, a useful abstract description of the innovations that have driven the large secular changes we have documented in the US economy between 1977 and 2013.

### A. *The Model*

Consider a firm  $i$  that produces a nontraded service  $j$ . The firm uses plants to produce in different locations  $n$ , out of a continuum of locations with mass  $N$ . The price of a variety of service  $j$  produced by firm  $i$  in location  $n$  is given by  $p_{ijn}$ . Assume that the only way to serve market  $n$  is to put an establishment there. A firm pays a fixed cost  $F_j$  (in units of the numeraire, which is the price of a freely traded undifferentiated good) to produce service  $j$  and another fixed cost  $f$  (in units of the numeraire, but indexed by the local wage  $w_n$ ) to set up an establishment in market  $n$ . The firm's productivity  $a_{ijn}A_{ij}$  has two components, one that applies to its establishments in all locations and one that is idiosyncratic to the market,  $a_{ijn}$ , and helps account for firm idiosyncratic entry patterns.<sup>26</sup> Labor is the only factor of production, so a firm that hires  $L_{ijn}$  units of labor produces  $Y_{ijn} = a_{ijn}A_{ij}L_{ijn}$  units of output, with local revenues given by  $R_{ijn} = p_{ijn}a_{ijn}A_{ij}L_{ijn}$ .

Now suppose that demand is CES and firms compete monopolistically. Each variety is produced by a single firm, so the demand function is  $p_{ijn} = E_{jn}Y_{ijn}^{-1/\sigma}$ , where  $\sigma > 1$  is the CES across varieties within an industry and  $E_{jn}$  is a function of local real industry expenditure and the local industry price index determined in spatial equilibrium. Conditional on serving market  $n$ , profit-maximizing employment in the local market is given by

$$L_{ijn} = (a_{ijn}A_{ij})^{\sigma-1} \left[ \left( 1 - \frac{1}{\sigma} \right) \frac{E_{jn}}{w_n} \right]^{\sigma}. \quad (2)$$

The firm will serve market  $n$  if local profits are positive, which is the case when the firm's productivity  $A_{ij}$  is above a threshold  $\alpha_n$  defined by

$$A_{ij} \geq \alpha(a_{ijn}, f, w_n, E_{jn}) \equiv \left( \frac{f}{\tilde{\sigma} a_{ijn}^{\sigma-1} w_n^{1-\sigma} E_{jn}^{\sigma}} \right)^{1/(\sigma-1)}, \quad (3)$$

where  $\tilde{\sigma} \equiv (\sigma - 1)^{\sigma-1} / \sigma^{\sigma}$ . Hence, the firm is more likely to enter a market where its local productivity  $a_{ijn}$  is higher, wages  $w_n$  are smaller, and total real expenditures  $E_{jn}$  are larger. Also, firms enter more markets the smaller the local fixed cost  $f$ .

Suppose that the distribution of a firm's  $\alpha_{ijn}$  is given by a CDF  $\Gamma(\cdot)$  with density  $\gamma(\cdot)$ . This distribution  $\Gamma(\cdot)$  is determined by parameters; the set of available markets; the distribution of idiosyncratic local productivity  $a_{ijn}$ , which we assume is independently and identically distributed across firms; and the joint distribution of  $E_{jn}$  and  $w_n$ , which is determined in equilibrium. The latter distribution is determined by the distribution of amenities, productivity, housing, and other geographic factors, as well as a variety

<sup>26</sup> Alternatively, we could make the firm's local fixed costs idiosyncratic to explain idiosyncratic entry across markets.

of other mobility and trade frictions. Here, we stop short of specifying a fundamental model of the distribution  $\Gamma(\cdot)$  to gain generality and simplify the exposition. For concreteness, appendix K7 provides a parametric example.

Now consider the decision of the firm to enter industry  $j$ . The firm will enter if total profits from industry  $j$  are greater than zero, namely,

$$\int_{n \text{ s.t. } A_{ij} > \alpha_{ijn}} \left[ \tilde{\sigma} (a_{ijn} A_{ij})^{\sigma-1} w_n^{1-\sigma} E_{jn}^{\sigma} - f \right] dn - F_j > 0,$$

where s.t. means “such that” and  $\alpha_{ijn} \equiv \alpha(a_{ijn}, f, w_n, E_{jn})$  is defined in equation (3). The profits of a firm that enters industry  $j$  are given by

$$\Pi(A_{ij}, F_j, f, \Gamma) = \int_0^{A_{ij}} \left( \left( \frac{A_{ij}}{\alpha} \right)^{\sigma-1} - 1 \right) f \Gamma(d\alpha; f) - F_j, \quad (4)$$

which is increasing in firm productivity,  $A_{ij}$ , and decreasing in industry fixed costs,  $F_j$  (and in local fixed costs,  $f$ , through their effect on  $\Gamma(\cdot)$ ). Thus, denote by  $\underline{A}(F_j, f, \Gamma)$  the unique productivity level such that  $\Pi(\underline{A}(F_j, f, \Gamma), F_j, f, \Gamma) = 0$ . Active firms in industry  $j$  are such that  $A_{ij} \geq \underline{A}(F_j, f, \Gamma)$ .

### B. A Menu of New Technologies

Sutton (1991) argues that new sunk-cost-intensive technologies lead to market concentration. We now borrow this idea and examine the effect of a menu of new technologies that increase the fixed cost of producing a given service in exchange for a reduction in the variable cost (and, for now, leaves the fixed cost of creating plants,  $f$ , constant). We consider a menu of new technologies indexed by  $h$ , where adopting the new technology  $h$  results in an increase in fixed costs to  $h^n F_j$  and an increase in productivity to  $h A_{ij}$ , for  $h \geq 1$  and  $\eta > 0$ . The old technology is given by  $h = 1$ . Hence, we assume that the economy starts in an equilibrium with  $h = 1$  and that, at some point, the new menu of technologies with  $h > 1$  becomes available. This is the only source of dynamics in our model.

We start by showing that, for a given  $h$ , the most productive firms are the ones that adopt the new technology and expand by entering new markets. All proofs are relegated to appendix K.

**PROPOSITION 1.** Given the distribution  $\Gamma$ , there exists a threshold  $H(F_j, f, \Gamma, h, \eta) > 0$  such that if  $A_{ij} \geq H(F_j, f, \Gamma, h, \eta)$ , then firm  $i$  adopts the new technology. Thus, in equilibrium the highest-productivity firms use the new technology and the lowest-productivity ones (if active) use the old technology. Firms that adopt the new technology are larger in employment and sales and enter more markets.

Now consider the case when firms can choose the level of  $h \geq 1$ . Assume that  $\eta > \sigma - 1$ , so the profit function is concave in  $h$ . It is easy to show that more productive firms will choose technologies with higher  $h$ . They also adopt more the more useful is the technology, parameterized by a lower  $\eta$ .

**PROPOSITION 2.** Given the distribution  $\Gamma$ , if a firm with productivity  $A$  chooses a technology  $h(A)$ , then firms in the same sector with technology  $A' \leq A$  choose technology  $h(A') \leq h(A)$ . That is,  $h(\cdot)$  is a weakly increasing function. Furthermore, there exists a threshold  $\eta_0$  such that if  $\eta < \eta_0$ , then  $h(A) > 1$  and is strictly increasing in  $\eta$  for all  $A$ .

We can also show that the new menu of technologies results in more industry concentration, with a relative expansion of top firms into new markets relative to the average firm in the industry. Furthermore, these effects will be heterogeneous across industries with different  $\eta$ s.

**PROPOSITION 3.** Given the distribution  $\Gamma$ , the menu of new technologies increases industry concentration and the number of markets of the most productive firms relative to average firms. It also increases average employment per market. The effects are more pronounced for small values of  $\eta$ , with no effect if  $\eta$  is sufficiently large.

### C. *A Technology That Reduces Local Fixed Costs Too*

The model above implies that the advent of the new technology results in increases in an adopter firm's average employment in a market. This prediction is consistent with the evidence if we interpret a market as a city (MSA). However, it is counterfactual if we interpret a market as a single establishment, where employment and sales of the average establishment of an adopter firm fall. To generate declines in the average size of adopters, we need to allow the new menu of technologies to reduce local fixed costs as well.

Suppose that the new menu of technologies is as before but, in addition, local fixed costs are now given by  $fh^{-\varphi}$ . The exponent  $\varphi > 0$  determines the extent to which fixed costs decline with the new chosen technology  $h$ . The exponent should depend on the definition of a market. For a large geographic area, we might think that the cost does not change much beyond the overall firm fixed costs, and so  $\varphi = 0$ . For a smaller area, such as the one covered by a single establishment,  $\varphi > 0$ , because of the ease in replicating standardized establishments (as exemplified by companies such as Starbucks). The next proposition shows that if local fixed costs fall sufficiently, the minimum and average establishment sizes decline.

**PROPOSITION 4.** Given the distribution of markets  $\Gamma$ , if the new technology also reduces local fixed costs to  $fh^{-\varphi}$ , the minimum employment size of the firm's establishments falls, and average establishment size falls if, conditional on  $\eta$ ,  $\varphi$  is large enough.

#### D. *Nontraded Services versus Traded Goods*

The fundamental difference between firms producing nontraded services and firms producing traded goods (as in many manufacturing industries) is that the former have to deliver their services locally, while the latter can ship goods at a relatively low cost from a distance. This allows traded-good producers to concentrate in one (or a few) large plants that supply many locations. In the extreme case when transporting manufacturing goods is free, the firm will produce in only one location. The model above then applies to traded-good industries but with firms that produce in a single location and so pay local fixed costs only there. Namely, the profits of a firm  $i$  in a traded-good industry  $j$  are given by

$$\max_m \int_n \left[ (A_{ij} a_{ijm})^{\sigma-1} \tilde{\sigma} w_m^{1-\sigma} E_{jn}^\sigma \right] dn - f - F_j > 0,$$

which can be written in terms of the distribution  $\Gamma(\cdot)$  as

$$\max_m \int \left( \frac{A_{ij}}{\alpha} \right)^{\sigma-1} \left( \frac{a_{ijm} w_n}{w_m a_{ijn}} \right)^{\sigma-1} f \Gamma(d\alpha; f) - f - F_j.$$

Firms choose their location optimally, so  $a_{ijm}/w_m \geq w_n/a_{ijn}$ , with equality when  $n$  is the preferred location. Hence,  $[(a_{ijm}/w_m)(w_n/a_{ijn})]^{\sigma-1} > 1$ . Perhaps not surprisingly, since firms in traded industries do not expand by adding new plants, the menu of new fixed-cost-intensive technologies is less relevant for them. The new technologies encourage productive firms in nontraded-service industries to reach more customers by adding additional establishments, a margin that is not present for firms producing traded goods, since they can already reach all consumers. Thus, conditional on their initial sales and fixed costs, service firms invest in the new technologies more intensively, compared to traded-goods firms.

**PROPOSITION 5.** Conditional on total sales and fixed costs, a traded-goods-producing firm  $i$  adopts the new technologies less intensively than a nontraded-service firm  $i'$ , namely,  $h_{iM} < h_{iS}$ .

#### E. *New Technologies and Industry Expansion*

Consider now industries that vary in the level of fixed costs needed to implement the new technologies, namely,  $\eta_j$ . We assume that agents have nested CES preferences with elasticity of substitution across industry consumption bundles given by  $\rho > 1$ . Because of CES preferences across industries with elasticity of substitution greater than 1, given the distribution of wages, the industry price index will fall with the adoption of the new technology, and aggregate industry quantities and sales will increase more than proportionally. This implies, for  $\eta_j$  small and given the distribution of wages and local price indexes, that industry expenditures in all

markets increase and so does total industry employment. In contrast, when  $\eta_j$  is large, firms do not invest and the industry does not change. The implication is that the advent of the menu of new technologies increases local and aggregate employment and sales in industries with low  $\eta_j$ .

**PROPOSITION 6.** Given the distribution of local wages,  $w_n$ , and price indexes,  $P_n$ , if  $\eta_j$  is sufficiently low that  $h_j(A_{ij}) > 1$  for some  $i$ , and  $\rho > 1$ , then local and aggregate industry sales and employment are decreasing in  $\eta_j$ .

Put together, these results show that a menu of new fixed-cost-intensive technologies naturally generates firm behavior consistent with the facts associated with the new industrial revolution in services that we documented in the previous section.<sup>27</sup> We now turn to study some of the implications on local markets.

## V. Implications and Evidence for Local Markets

In the previous section, we showed that top firms that take advantage of new technologies for delivering nontraded services grow by expanding into new local markets. Furthermore, these new local markets are typically smaller. In this section, we examine the effect on local markets of the entry of top national firms.

### A. Top-Firm Entry and Local Market Concentration

The increasing presence of top firms, particularly in the smallest cities, allows local residents to access new varieties of good and services. In the model we presented in section IV, the local employment and sales share of firm  $i$  producing product  $j$  in market  $n$  is given by

$$s_{ijn} = \frac{a_{ijn}A_{ij}^{\sigma-1}}{\int_{i \in \mathcal{I}_{jn}} (a_{ijn}A_{ij})^{\sigma-1} di},$$

where  $\mathcal{I}_{jn}$  is the set of producers of good  $j$  in market  $n$ . Employment or sales shares depend directly on the relative productivity of firms in a market. Top firms gain large market shares when they enter, since they tend to be more productive than local incumbents. Consistently, figure 7 shows that the share of top firms increased significantly in small and mid-sized local markets.

<sup>27</sup> In app. K8, we show that an explanation based on increases in the demand for services has implications on firm behavior that are isomorphic to changes in firm productivity. Because those explanations do not affect industry fixed costs,  $F_j$ , the resulting changes are not biased in favor of top firms.

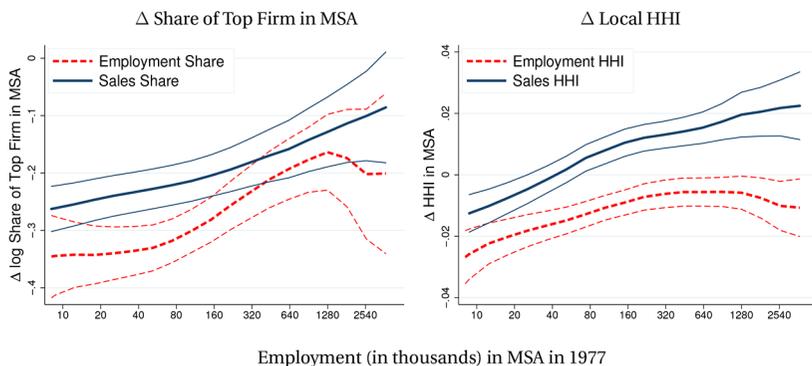


FIG. 8.—Change in local concentration from 1977 to 2013 by MSA size in 1977. Unit of observation is an MSA ( $N = 329$ ). Figure shows point estimates and 95% confidence intervals of nonparametric regressions of the weighted average of the change in the log share of employment or sales of the top firm (*left*) or the HHI (also employment or sales; *right*) in each industry-MSA from 1976 to 2013 on the log of total employment in the MSA in 1977. Weights are the Sato-Vartia average of the employment share of each industry-MSA in 1976 and 2013.

Within each market, however, we also see the share of the largest firm in each industry-city falling everywhere and particularly in cities that were small in 1977. Specifically, we calculate the change in the log employment and sales share of the top firm in each industry and city from 1977 to 2013 and then take the weighted average across all the industries in a city.<sup>28</sup> The left-hand panel in figure 8 plots the change in the average employment or sales share of the top firm in the MSA-industry against the size (total employment) of the city in 1977. The log employment share of the top firm in each industry-city declined by about 20% between 1977 and 2013 in the largest cities and by about 35% in the smallest ones. For sales, the decline was 25% in the smallest cities and about 10% in the largest ones. The figure suggests that top firms entering new markets gained market share by competing with local providers that had very large market shares themselves. Rather than seeing new top firms monopolizing the new markets where they enter, we see top firms taking away some of the market share of local monopolists (or oligopolists).

The right-hand panel in figure 8 shows the average change in the HHI in each MSA between 1977 and 2013. Here, we calculate the HHI for each industry in the MSA as the sum of the squares of the employment or sales share of each firm in the industry in the MSA, and we take the weighted average using Sato-Varia weights of the change in this index between 1977 and 2013 across all industries in the MSA present in the two years. Local

<sup>28</sup> We use Sato-Vartia employment weights for each industry-city in 1977 and 2013 to aggregate across all the industries in a given MSA that exist in the two years.

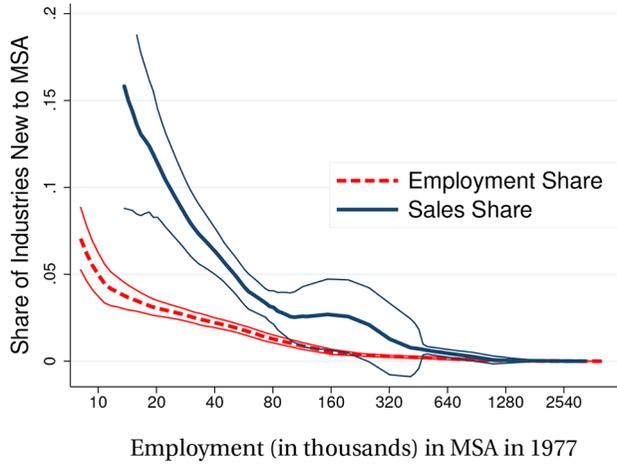


FIG. 9.—Employment and sales share of new industries. Unit of observation is a MSA ( $N = 329$ ). Figure shows coefficients and 95% confidence intervals from nonparametric regressions of employment or sales of top firms in industries new to the MSA in 2013 (employment) or 2012 (sales) as a share of total MSA employment or sales on total MSA employment in 1977. New industries in MSA are industries not in the MSA in 1977. Top firms are the top 10% of firms in each industry in 2013 (employment) or 2012 (sales). Sales share excludes finance and utilities and transportation.

employment concentration has fallen across MSAs of all sizes. This is consistent with the evidence in Rossi-Hansberg, Sarte, and Trachter (2020) that shows the diverging trends between increasing national and decreasing local product market concentration. Furthermore, as is the case with the share of the top local firm, the fall in the local HHI is particularly pronounced in smaller cities where top firms have entered more. For sales, we also see a decline in MSA HHI for the smallest cities but an increase for the largest ones. The pattern across cities is similar, but the level of the change is significantly larger.

### B. *Introducing New Products to Local Markets*

Top firms can come to new markets to compete with local producers, as we showed above, but they also introduce new products into these markets. Figure 9 shows the local share of employment and sales in 2013 of top firms in industries that were not present in 1977.<sup>29</sup> That is, it measures the extent to which top firms are responsible for bringing new industries to particular cities. The employment share of new industries in 2013 is as large as 7% for the cities that were among the smallest in 1977

<sup>29</sup> The sales data exclude finance and utilities and transportation, as the microdata for these industries are not available in the 1977 economic censuses.

but is negligible for the cities among the largest in 1977. The sales share of new industries in 2012 in the smallest cities in 1977 is even larger, at almost 15%, while again it is zero for the largest cities in 2012. Hence, not only are top firms changing the distribution of market shares,  $s_{ijm}$ , by changing the local distribution of productivities and potentially adding new varieties, but they are also changing the set of industries available in a market. Of course, in our model, both margins increase consumer welfare, since agents exhibit “taste for variety” modulated by the parameter  $\sigma$  for varieties within an industries and  $\rho$  for products across industries.

### C. Local Market Employment Implications

The entry of top firms, particularly in small cities, can generate new employment in those locations or mostly replace current employment by simply redistributing existing workers to the top firms. In our framework, an additional top firm can never reduce total employment in an industry-city, since we are assuming an elasticity of substitution between varieties greater than 1,  $\sigma > 1$ . The extent to which employment in the aggregate increases as a result of the entry of top firms depends on the elasticity of local population to local real wages. In turn, this depends on mobility costs, preference heterogeneity, and other characteristics of the moving behavior of agents that we have not fully specified. In any case, our hypothesis and model suggest that small cities that have seen the bulk of the increase in top-firm establishment entry should have grown faster than larger ones since the late 1970s. Figure 10 shows that this is indeed the case. On average, the smallest cities in 1977, such as Missoula, Montana (employment of 19,000 in 1977), doubled their size between 1977 and 2013, while the largest cities, such as New York, increased by only 35%. The documented scale dependence in employment at the MSA level over this long period is a violation of Gibrat’s law, which states that city growth is independent of city size.<sup>30</sup> The secular changes that resulted from the industrial revolution in services are a likely culprit.

The entry of top firms can potentially have negative implications for some local residents if it leads to job destruction and exit by incumbents. In turn, these forces can potentially be compensated for or overwhelmed by overall local employment growth, particularly in small cities where employment growth was faster, as documented in figure 10. The implications of the industrial revolution in services for job destruction and its variation across cities of different sizes is, therefore, ambiguous. Figure 11 plots the average 5-year job destruction rate between 1977 and 2013 as a function of

<sup>30</sup> In the literature, Gibrat’s law has been established using population data, not employment as in fig. 10. See, e.g., Gabaix and Ioannides (2004).

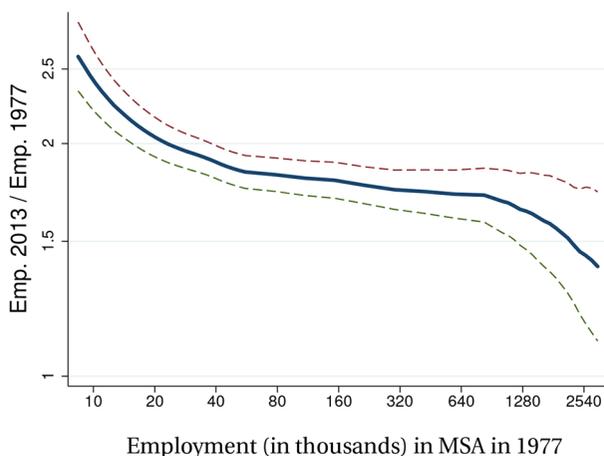


FIG. 10.—Local employment growth by city size. Unit of observation is a MSA ( $N = 329$ ). Figure shows coefficients and 95% confidence intervals from a nonparametric regression of ratio of total employment in the MSA in 2013 to employment in the MSA in 1977 (y-axis) against total employment in the MSA in 1977 (x-axis).

city size at the beginning of the sample.<sup>31</sup> The left-hand panel plots job destruction due to firms that exit the MSA, while the right-hand panel plots job destruction due to shrinking employment in incumbent firms in the MSA. Perhaps surprisingly, job destruction does not seem to vary much by initial city size, and, if anything, the relationship is positive when we look at MSAs with more than 20,000 jobs in 1977. That is, there is more job destruction due to exit and incumbent downsizing in large than in small cities.

#### D. *Missing Growth and the Industrial Revolution in Services*

We now estimate the implications of the technological revolution in the service sector for TFP growth. Aggregate TFP in locality  $n$  is defined as  $TFP_n = Y_n/L_n = R_n/(P_n L_n)$ . In the data, employment  $L_n$  and nominal expenditures  $R_n$  can be easily measured, but measuring prices  $P_n$  is complicated, since it requires the prices per unit of quality of goods and services sold in each market. These complications are particularly salient

<sup>31</sup> Following Davis and Haltiwanger (1992), we measure the job destruction rate by exiting firms between  $t$  and  $t + 1$  as the ratio of employment at time  $t$  of firms that exit the MSA by  $t + 1$  to the average of total employment in the MSA in  $t$  and  $t + 1$  and the job destruction rate of incumbent firms between  $t$  and  $t + 1$  as the ratio of employment losses of firms in the MSA that shrink between  $t$  and  $t + 1$  to the average of total employment in the MSA in the two years.

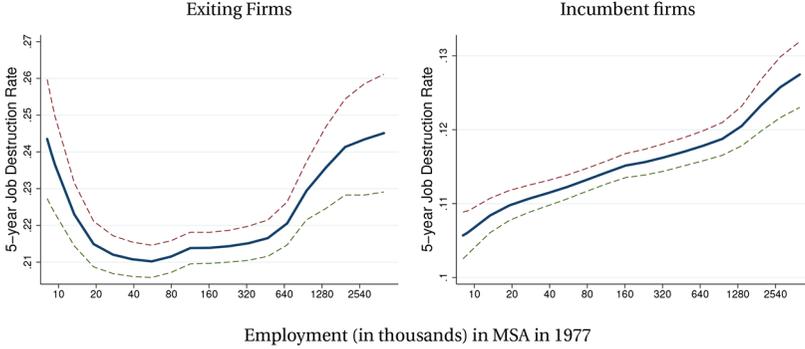


FIG. 11.—Job destruction rate by MSA size. Unit of observation is an MSA ( $N = 329$ ). Figure shows point estimates and 95% confidence intervals of nonparametric regressions of job destruction rate in each MSA calculated over 5-year periods from 1977 to 2013 from exiting firms (*left*) and incumbent firms that shrink (*right*) in each 5-year period on total employment in the MSA in 1977. The job destruction rate is calculated as total jobs lost from exiting firms or shrinking incumbent firms as a share of the average of total employment in the MSA in the beginning and end of each 5-year period.

for the service industries, where quality-adjusted prices are notoriously hard to measure. In the service sector, the BLS (Bureau of Labor Statistics) measures the price of real output as the price of a well-defined service in the same establishment. However, we have shown that the growth of top firms in the service sector is entirely driven by entry of top firms into new markets. As argued by Aghion et al. (2019a), quality growth due to firm entry into new markets is not measured by the BLS.

We use Aghion et al.’s (2019a) procedure to measure the growth not captured by the BLS because of entry of new establishments. We differ in that we measure missing growth in each location and then aggregate missing growth across all locations. As we mentioned in section IV, utility of the representative consumer in  $n$  is given by

$$U_n \equiv \left( \int_{j \in \mathcal{J}_n} Q_{jn}^{(\rho-1)/\rho} dj \right)^{\rho/(\rho-1)},$$

where  $\mathcal{J}_n$  is the set of industries present in location  $n$ , which can change over time;  $Q_{jn}$  is consumption of varieties of industry  $j$ , which are aggregated according to

$$Q_{jn} \equiv \left( \int_{i \in \mathcal{I}_{jn}} Y_{ijn}^{(\sigma-1)/\sigma} di \right)^{\sigma/(\sigma-1)},$$

where  $\mathcal{I}_{jn}$  is the set of firms in industry  $j$  in location  $n$ , which again can change over time.

Following Feenstra (1994), the growth rate (denoted by a hat accent) of the ideal price of  $Q_{jn}$  between  $t$  and  $t + 1$  is

$$\hat{P}_{jn,t} = \hat{P}_{jn,t|i \in \mathcal{I}_{jn,t}} - \frac{1}{\sigma - 1} \hat{s}_{jn,t|i \in \mathcal{I}_{jn,t}}, \quad (5)$$

where  $\mathcal{I}_{jn,t}$  denotes the set of incumbent firms in industry  $j$  in locality  $n$  between  $t$  and  $t + 1$  and  $s_{jn,t|i \in \mathcal{I}_{jn,t}} \equiv \int_{i \in \mathcal{I}_{jn,t}} s_{ijn} di$  is the sales share of the incumbent firms in industry  $j$  in locality  $n$ . The first term in equation (5) is the growth in the ideal price of the varieties produced by incumbent firms in the locality. Since the BLS collects prices from incumbent firms in a location, this term, in theory, is captured by official price statistics. The second term in equation (5) is the change in the price index due to the entry of new establishments into the local market, which is not measured by the BLS. The resulting bias is given by the change in the nominal sales share of the incumbent firms in sector  $j$  in location  $i$ , multiplied by  $1/(\sigma - 1)$ . Since they did not have sales data, Aghion et al. (2019a) proxy the sales share by the employment share. In our case, we use the sales data from the economic censuses to measure the change in the sales share of incumbent firms,  $\hat{s}_{jn,t|i \in \mathcal{I}_{jn,t}}$ . Specifically, for each industry in a location over each 5-year period, we measure “missing growth” from firm entry in a location-industry as the product of  $1/(\sigma - 1)$  and the change in the log sales share of incumbent firms in each industry in the city over the 5-year period.

Aggregating across all the products in a city and using equation (5), the growth rate of the aggregate local price in  $n$  between  $t$  and  $t + 1$  is then given by

$$\hat{P}_{n,t} = \underbrace{\int_{j \in \mathcal{J}_{n,t}} \beta_{jn,t} \hat{P}_{jn,t|i \in \mathcal{I}_{jn,t}} dj}_{\text{measured by BLS}} - \underbrace{\frac{1}{\sigma - 1} \int_{j \in \mathcal{J}_{n,t}} \beta_{jn,t} \hat{s}_{jn,t|i \in \mathcal{I}_{jn,t}} dj}_{\text{missing growth new varieties}} - \underbrace{\frac{1}{\rho - 1} \hat{s}_{n|j \in \mathcal{J}_{n,t}}}_{\text{missing growth new industries}}, \quad (6)$$

where  $\beta_{jn,t}$  is the Sato-Vartia weight of industry  $j$  and  $\mathcal{J}_{n,t}$  is the set of incumbent industries in location  $n$  in  $t$  and  $t + 1$ . The first term in equation (6) is the Sato-Vartia weighted average of the growth in prices by incumbent establishments in the incumbent products (the first term in eq. [5]). The last two terms are not measured by the BLS and capture the effect of entry into the local market on the price index. The second term is the Sato-Vartia weighted average of missing growth due to the entry of new varieties for the incumbent industries (the second term in eq. [5]). The third term is the missing-growth term from entry of new industries into the local market. As shown in figure 9, top firms also create brand new service industries in the cities they enter, and this effect is larger in smaller cities. This effect is the last term in equation (6), which

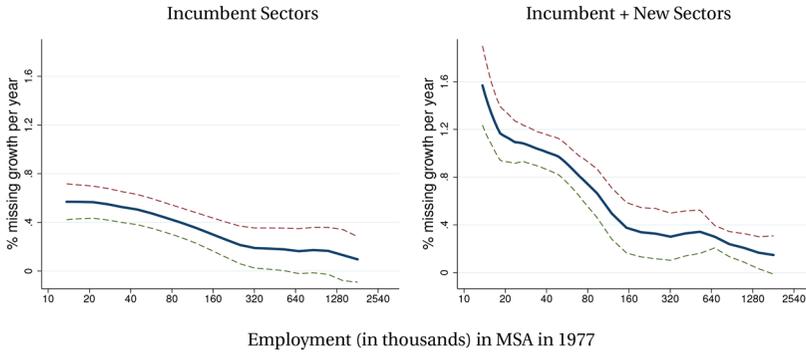


FIG. 12.—Local missing growth by MSA size. Unit of observation is an MSA ( $N = 329$ ). Figure shows the coefficients and 95% confidence intervals from nonparametric regressions of the average annual missing growth from 1977 to 2012 in the MSA due to entry of top firms into incumbent sectors in the MSA (*left*) and all sectors in the MSA (*right*) on total MSA employment in 1977. Missing growth is calculated for each 5-year period from sales data using equation (6). Incumbent sectors are present in the MSA at the beginning and end of each 5-year period. New sectors are in the MSA at the end but not at the beginning of each 5-year period. Establishments in finance and utilities and transportation are excluded.

is the change in the sales share of the incumbent industries multiplied by  $1/(\rho - 1)$ . Total missing growth in a local market is the sum of missing growth due to entry by top firms into incumbent industries and that due to their entry into industries that are new to the city.

Figure 12 presents the resulting estimates of missing growth for each MSA based on sales data from the economic censuses every 5 years between 1977 and 2012.<sup>32</sup> We limit the industries to those with microdata on sales over this entire period.<sup>33</sup> The unit on the y-axis is the average annual growth rate per year in each MSA between 1977 and 2012 missed by BLS. The left-hand panel displays missing growth only from industries that were present in the MSA throughout each 5-year period (the second term in eq. [6]). Missing growth due to entry of top firms into local markets in incumbent industries is 0.6% per year in small cities but only 0.1% in the largest ones. The right-hand panel in figure 12 adds to the missing growth in the incumbent industries the contribution of missing growth due to the local entry of establishments in new industries (third term in eq. [6]). According to this calculation, the BLS's procedures understate TFP growth by 1.6% per year in the smallest US cities and by a more modest 0.2% per year in the largest cities. Top firms have not brought new

<sup>32</sup> We assume that  $\sigma = 3$  and  $\rho = 2$ .

<sup>33</sup> We omit finance and utilities and transportation. The microdata for the former are available only starting in 1992 and those for the latter starting in 1987. Figure J1, in app. K, shows missing growth for all industries (including finance and utilities and transportation), on the basis of employment data in the LBD.

industries to the largest cities; they have always been there, so missing growth in large US cities is all due to entry into incumbent industries.

Finally, after we aggregate missing growth across all MSAs, aggregate missing growth due to the entry of top firms into local markets averaged 0.5% per year from 1977 to 2012. To be clear, our estimates of “missing growth” capture only the effect of entry of top firms in a locality in a 5-year period. New establishments of top firms could have grown after entry, and this growth, in theory, is measured by the BLS. But of course, it is an open question whether the BLS’s procedures capture quality growth in incumbent service-sector establishments.

### *E. A Discussion on Markups*

Our data do not allow us to measure markups or profits of top firms that enter into new local markets. In the model in section IV, the markup of all firms is constant at  $\sigma/(\sigma - 1)$ . Of course, the number of entrants and the scale of production vary so that total firm profits cover establishment- and firm-level fixed costs in each industry where the firm is active. Hence, if firm-industry fixed costs have risen and firms are paying more local fixed costs to open establishments in more locations, the total fixed costs paid by top firms must have risen as well. These fixed costs could take the form of investments in intangibles such as marketing, information technology, and worker training. This is consistent with the evidence in Haskel and Westlake (2017) that investment in intangibles has risen in the United States in the period we study. Furthermore, in fact 3 in section III we showed that employment in HQ and R&D has grown in industries where firms have expanded the number of establishments per firm. Our mechanism also implies that profits by top firms must also have increased to pay for these fixed costs, which is consistent with the evidence in Barkai (2020). In short, an integral part of our hypothesis is that the industrial revolution in services leads to rising investments in fixed costs, some of which could be intangibles, and rising profits by top firms.

Of course, our monopolistic competition model with fixed markups could be extended to incorporate firms with variable markups. In such models, dominant firms in a market could take advantage of local consumers by raising prices, particularly if other competitors have exited or cannot produce similar products. However, in most models with variable markups, profits would fall in markets where the top firm has a smaller employment share and market concentration in terms of the HHI index has fallen. Vogel (2008) presents a model of a local market where firms can position their product (by, e.g., choosing their location) and choose their price. It shows that, if the dispersion in firm productivities is not too large, the unique subgame perfect Nash equilibrium exhibits firm profits that are proportional to local population size and quadratic in market share. The result

is that total local profits are proportional to the HHI index, which we have shown has fallen, especially in small cities. Most models of variable mark-ups produce similar results.

## VI. Conclusion

Over the past 4 decades, the US economy has experienced a new industrial revolution that has enabled firms to scale up production over a large number of establishments dispersed across space. The adoption of these technologies has particularly favored productive firms in nontraded-service industries.

The industrial revolution in services has had its largest effect in smaller and mid-sized local markets. Top nontraded-service firms have expanded into small local markets but have always been present in the largest US cities. Over the past 4 decades, small and mid-sized US cities saw the largest declines in local concentration and the highest growth rate of employment. The gain to local consumers from access to more, better, and novel varieties of local services from the entry of top firms into local markets is not captured by the BLS. We estimate that such “missing growth” is as large as 1.6% in the smallest markets and averages 0.5% per year from 1977 to 2013 across all US cities. Although quite large, this number is not an estimate of the full effect of the industrial revolution in services on aggregate TFP. To provide one we would also need to estimate productivity growth of top firms after they enter into each local market as well as estimate the effect of entry of top firms on competition and markups in each local market.

We leave three important questions for future work. First, although we interpret all the facts we have uncovered with a consistent explanation based on the industrial revolution in services, some of these facts might be partly the result of other concurrent phenomena. In particular, rising incomes have probably led to an increase in the demand for services due to non-homotheticities in preferences. This could explain the relative expansion of service industries, some of which naturally would happen across markets. However, such a change is not obviously biased toward the most productive firms. Similarly, some papers have argued that changes in the distribution of firm age can explain some of the observed firm and market dynamics over recent decades (e.g., Hopenhayn, Neira, and Singhania 2018). We show that our facts are not mainly driven by age differences, but changes in firm entry can account for a fraction of some of them. More generally, future work should aim at distinguishing quantitatively the role of these and other related forces.

Second, it is important to establish more precisely what is the new technology behind the industrial revolution in services. We know that it has

been implemented by hiring more workers in HQ and R&D establishments. The timing of these trends also suggests that general-purpose innovations in ICT have probably facilitated these fixed-cost-based sectoral innovations. We gave some examples in our narrative in section I about the Cheesecake Factory and the Steward Health Care Group, but that only scratches the surface. We believe that a blend of quantitative and narrative accounts of this new industrial revolution, in the style of Chandler's (1993) seminal work on the history of the industrial revolution in US manufacturing, would be very useful.

Third, the industrial revolution in services has implications for the employment of workers of different skills across locations. If labor markets are industry specific and local, the decline in local concentration of employment caused by the entry of top firms should reduce the monopsony power of employers in small markets. However, as we have argued, the revolution in services implies a relative shift from employment of workers in local establishments to workers needed for the firm-wide fixed-cost investments. The fixed costs are likely to be skilled-worker intensive and can be located anywhere. Hence, top firms may choose to hire workers performing them in large and skill-abundant cities. Drawing out some of these implications more fully seems potentially fruitful.

### Data Availability

This paper uses the confidential microdata of the US economic censuses and LBD. Code to replicate the tables and figures in this paper can be found in Hsieh and Rossi-Hansberg (2022) in the Harvard Dataverse, <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/TNO1AH>.

### References

- Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter Klenow, and Huiyu Li. 2019a. "Missing Growth from Creative Destruction." *A.E.R.* 109 (1): 2795–822.
- . 2019b. "A Theory of Falling Growth and Rising Rents." Working Paper no. 26448 (November), NBER, Cambridge, MA.
- Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. "The Fall of the Labor Share and the Rise of Superstar Firms." Working Paper no. 23396 (May), NBER, Cambridge, MA.
- Barkai, Simcha. 2020. "Declining Labor and Capital Shares." *J. Finance* 75 (5): 2421–63.
- Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Q.J.E.* 117 (1): 339–76.

- Cao, Dan, Henry R. Hyatt, Toshihiko Mukoyama, and Erick Sager. 2019. "Firm Growth through New Establishments." Working Paper. <https://doi.org/10.2139/ssrn.3361451>.
- Caroli, Eve, and John Van Reenen. 2001. "Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments." *Q.J.E.* 116 (4): 1449–92.
- Chandler, Alfred D., Jr. 1993. *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, MA: Harvard Univ. Press.
- Crouzet, Nicolas, and Janice Eberly. 2018. "Intangibles, Investment, and Efficiency." *AEA Papers and Proc.* 108:426–31.
- Davis, Steven J., and John Haltiwanger. 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation." *Q.J.E.* 107 (3): 819–63.
- Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2017. "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown." *A.E.R.* 107 (5): 322–26.
- De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2018. "The Rise of Market Power and the Macroeconomic Implications." Working Paper no. 23687 (August), NBER, Cambridge, MA.
- Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2023. "How Costly Are Markups?" *J.P.E.*, forthcoming.
- Feenstra, Robert C. 1994. "New Product Varieties and the Measurement of International Prices." *A.E.R.* 84 (1): 157–77.
- Fort, Teresa, and Sean Klimek. 2018. "The Effects of Industry Classification on US Employment Composition." Working Paper no. CES-18-28, US Census Center for Economic Studies, Washington, DC.
- Gabaix, Xavier, and Yannis Ioannides. 2004. "The Evolution of City Size Distributions." In *Handbook of Regional and Urban Economics*, vol. 4, edited by J. Vernon Henderson and Jacques-François Thisse, 2341–78. Amsterdam: North-Holland.
- Ganapati, Sharat. 2018. "The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies." Working Paper CES-18-49, US Census Center for Economic Studies, Washington, DC.
- Garicano, Luis, and Esteban Rossi-Hansberg. 2004. "Inequality and the Organization of Knowledge." *A.E.R.* 94 (2): 197–202.
- . 2006. "Organization and Inequality in a Knowledge Economy." *Q.J.E.* 121 (4): 1383–435.
- Gaubert, Cecile. 2018. "Firm Sorting and Agglomeration." *A.E.R.* 108 (11): 3117–53.
- Gawande, Atul. 2012. "Big Med." *New Yorker*, August 6, 52–63.
- Gutiérrez, Germán, and Thomas Philippon. 2017. "Declining Competition and Investment in the U.S." Working Paper no. 23583 (July), NBER, Cambridge, MA.
- Haskel, Jonathan, and Stian Westlake. 2017. *Capitalism without Capital: The Rise of the Intangible Economy*. Princeton, NJ: Princeton Univ. Press.
- Hobijn, Bart, and Boyan Jovanovic. 2001. "The Information-Technology Revolution and the Stock Market: Evidence." *A.E.R.* 91 (5): 1203–20.
- Holmes, Thomas J. 2011. "The Diffusion of Wal-Mart and Economies of Density." *Econometrica* 79 (1): 253–302.
- Hopenhayn, Hugo, Julian Neira, and Rish Singhania. 2018. "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share." Working Paper no. 25382 (December), NBER, Cambridge, MA.
- Hortaçsu, Ali, and Chad Syverson. 2015. "The Ongoing Evolution of US Retail: A Format Tug-of-War." *J. Econ. Perspectives* 29 (4): 89–112.

- Hsieh, Chang-Tai and Esteban Rossi-Hansberg. 2022. "Replication Code for 'The Industrial Revolution in Services.'" Harvard Dataverse. <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/TNO1AH>.
- Karabarbounis, Loukas, and Brent Neiman. 2018. "Accounting for Factorless Income." Working Paper No. 24404 (March), NBER, Cambridge, MA.
- Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter. 2020. "Plants in Space." Working Paper no. 27303 (June), NBER, Cambridge, MA.
- Rajan, Raghuram, and Julie Wulf. 2006. "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies." *Rev. Econ. and Statis.* 88 (4): 759–73.
- Ramondo, Natalia. 2014. "A Quantitative Approach to Multinational Production." *J. Internat. Econ.* 93 (1): 108–22.
- Rosenberg, Nathan, and Manuel Trajtenberg. 2004. "A General-Purpose Technology at Work: The Corliss Steam Engine in the Late-Nineteenth-Century United States." *J. Econ. Hist.* 64 (1): 61–99.
- Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2020. "Diverging Trends in National and Local Concentration." *NBER Macroeconomics Ann.* 35:115–150.
- Sutton, John. 1991. *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. Cambridge, MA: MIT Press.
- Syverson, Chad. 2017. "Challenges to Mismeasurement Explanations for the US Productivity Slowdown." *J. Econ. Perspectives* 31 (2): 165–86.
- . 2019. "Macroeconomics and Market Power: Context, Implications, and Open Questions." *J. Econ. Perspectives* 33 (3): 23–43.
- Traina, James. 2018. "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements." New Working Paper Series no. 17, Stigler Center, Univ. Chicago.
- Vogel, Jonathan. 2008. "Spatial Competition with Heterogeneous Firms." *J.P.E.* 116 (3): 423–66.