The rise in national industry concentration in the US between 1977 and 2013 is driven by a new industrial revolution in three broad non-traded sectors: services, retail, and wholesale. Top firms in these sectors have grown entirely by expanding into new local markets that are predominantly small and mid-sized U.S. cities. Local concentration has decreased in all U.S. cities but by significantly more in cities that were initially small. Top U.S. firms in the aggregate economy are increasingly specialized in sectors with rising industry concentration, but their aggregate employment share has remained roughly stable. These facts are consistent with the availability of a new set of fixed-cost technologies in non-traded sectors that enable adopters to produce at lower marginal costs in all markets.

*We thank Feng Lin, Harry Li, and Jihoon Sung for extraordinary research assistance. We also thank Rodrigo Adao, Dan Adelman, Audre Bagnall, Jill Golder, Bob Hall, Pete Klenow, Danial Lashkari, Raghuram Rajan, Richard Rogerson, and Chad Syverson for helpful discussions. The data from the US Census has been reviewed by the U.S. Census Bureau to ensure no confidential information is disclosed.
1. Introduction

Modern production relies on scale: The ability to use a technology to produce the same product or service innumerable times. In manufacturing industries, the steam-engine, electricity, and Ford’s assembly line, together with a number of other inventions, allowed firms to scale production in a single location. For many goods, the cost advantages of a larger scale overwhelmed the cost of transporting the goods to their final consumers, leading to great reductions in total average costs. This ability to scale 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 their final consumers. Modern scale production in these sectors had to wait for a different technology, one that allowed firms to replicate the same production process in multiple locations close to consumers.

We argue that new ICT-based technologies and adoption of new management practices have finally made it possible for firms outside of manufacturing to scale production over a large number of locations. The resulting expansion led to an increase in the national market share of top firms in many industries; a central fact about the US economy in the last three or four decades documented by Autor et al. (2017). This fact is the result of a new industrial revolution. One that has taken place in many non-traded service sectors.

Consider Gawande (2012)’s 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, CA – 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.” The cooks in the R&D facility then teach the new recipes to the kitchen managers of each restaurant at a bi-annual meeting in California. The kitchen managers then follow a finely honed procedure to teach the new recipes to the cooks in each restaurant. The roll out 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 that has taken place in sit-down restaurant meals due to companies such as the Cheesecake Factory has taken place in many non-traded sectors. Take hospitals as another example. Four decades ago, about 85% of hospitals were single establishment non-profits. Today, more than 60% of hospitals are owned by for-profit chains or are part of a large network of hospitals owned by an academic institution (such as the University of Chicago Hospitals).\(^1\) 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 6 Catholic hospitals in Boston. In Gawande (2012)’s account, Cerberus’ 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) 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 ICU 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 ICU room and talk directly to the staff on the scene or to the patients themselves.”

Technologies such as the remote ICU has 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 out-patient clinics.\(^2\) By 2019, Steward had expanded from its 6 original hospitals in Boston to 38 hospitals and 271 outpatient clinics located in 10 states and Malta.\(^3\)

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 non-traded services. In this sense, what has happened in non-traded services is akin to the industrial revolution unleashed by Henry

---

\(^1\)The employment-weighted share of multi-establishment hospitals in the Longitudinal Business Database increased from 15% in 1977 to 62% in 2013.

\(^2\)Steward uses software by Meditech. The dominant medical software company is Epic in Madison, Wisconsin.

\(^3\)Steward’s hospitals and out-patient clinics are in Massachusetts, New York, Ohio, Florida, Arkansas, Louisiana, Texas, Arizona, Pennsylvania, New Hampshire, Utah and Malta.
Ford more than a hundred years ago when Ford introduced mass production to a car industry dominated by independent artisans.

We use micro-data from the Longitudinal Business Database from 1977 to 2013 to document the following facts. First, we show that the phenomena of rising concentration documented by Autor et al. (2017) is primarily seen in three broad sectors – services, wholesale, and retail. As Autor et al. (2017) suggest, top firms have become more efficient over time, but our evidence indicates that this is only true for top firms in these three sectors. In manufacturing, for example, concentration has fallen.

Second, rising concentration in these sectors is entirely driven by an increase in the number of local markets served by top firms. Within a typical market served by a top firm in sectors with increasing concentration, employment of top firms is either constant or falling. Specifically, average employment per establishment of top firms falls in sectors with rising concentration. The same is true for employment of top firms in each county they serve.

Third, the new local markets that top firms expanded into tend to be smaller ones. The share of top firms in local employment has grown in smaller and mid-sized U.S. cities. In the very largest U.S. cities, there is no change in the employment share of top firms. The increasing 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 firm in the local market and the Herfindahl-Hirschman index (HHI) in the city declines in all U.S. cities, but the decline is much more pronounced in smaller cities.4

Fourth, we find that total employment rises substantially in industries with rising concentration. This evidence is consistent with our view that increasing concentration is driven by forces such as the adoption of new technologies or management practices that ultimately raise aggregate industry TFP. It is not consistent with the view that concentration is due to declining competition or entry barriers, as suggested by Gutierrez and Philippon (2017) and Furman and Orszag (2018), as these forces will result in a decline in industry employment.

Fifth, we show that the top firms in the economy as a whole have become increasingly specialized in narrow set of sectors, and these are precisely the non-traded sectors

---

4Using a different dataset (the National Establishment Time Series), Rossi-Hansberg et al. (2018) also find that local concentration has fallen significantly.
that have undergone an industrial revolution. At the same time, top firms have exited many sectors. The net effect is that there is essentially no change in concentration by the top firms in the economy as a whole. The “super-star” firms of today’s economy are larger in their chosen sectors and have unleashed productivity growth in these sectors, but they are not any larger as a share of the aggregate economy.

We use a simple theory of firm size and market entry to show that a key ingredient of the industrial revolution in services that we document is a new fixed-cost intensive technology that lowers the marginal cost in all markets served by the firm. The adoption decision of firms involves a trade-off between a proportional reduction in variable costs and an increase in the fixed cost of the firm. With a large enough fixed cost, only the most efficient firms find it profitable to adopt the new technology, which leads to more concentration in the industry. Firms that adopt the new fixed cost technology in an industry expand by serving new markets that are now viable due to their lower marginal cost. Rising input prices due to the expansion of firms that adopt the new technology force multi-product firms to leave sectors where the new technological advances are less relevant or where their relative productivity is low. The net effect on total employment of top firms is ambiguous, but top firms end up increasingly specialized in sectors that exhibit increases in concentration.

The industrial revolution in services has aggregate and local implications that we 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 top firm and the local Herfindahl-Hirschman index 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 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

---

5Our 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).
top national firms into local markets is not measured in official price statistics because current statistical procedures only measures prices from *incumbent* establishments. Following Aghion et al. (2019a), we calculate “missing growth” 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, nor 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 warehouses 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. Information and communication technology (ICT) started in the 60’s 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.\(^6\) There is a vast literature on the effect of these changes on the organization of production.\(^7\) 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 sys-

\(^6\) See Hobijn and Jovanovic (2001) for the diffusion of ICT technologies.

\(^7\) 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 et al., 2002), the span of control of managers (Rajan and Wulf, 2006; Garicano and Rossi-Hansberg, 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.
tems. 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.\footnote{Syverson (2017) argues that if we were mismeasuring the gains from ICT, the high-tech sector would need to be much larger than it is. If ICT is used for fixed costs 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 (Gutierrez and Philippon, 2017; De Loecker et al., 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 et al., 2018; Syverson, 2019; Edmond et al., 2019), 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 preferences and, therefore, fixed markups, in a model with variable markups these same technological changes could generate increases in markups. We chose not to 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 et al. (2018).\footnote{The magnitude of the trend in markups is still controversial. See the discussion in Traina (2018) and Karabarbounis and Neiman (2018).}

The rest of the paper is organized as follows. Section 2 presents our empirical findings organized in five Facts. Section 3 presents the theory and derives the implications of the availability of a menu of new technologies offering combinations of fixed and...
variable costs. Section 4 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 5 concludes.

2. Facts

We use micro-data from the U.S. Census Longitudinal Business Database (LBD). The LBD is based on administrative employment records of every nonfarm private establishment in the U.S. economy. The advantages of the LBD are its broad coverage and quality. The establishment-level variables we use are employment, county, industry (6-digit 2002 NAICS code provided by Fort and Klimek (2018)), 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. We classify each 2002 NAICS code into 450 consistently defined industries from 1977 to 2013. Hereafter, when we refer to an industry we mean one of these 450 industries. We also group counties into metropolitan areas (MSAs) defined consistently over time. We use the firm ID to aggregate employment of establishments to a firm in an industry or in the aggregate economy. 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 concentration by companies for which franchising is an important margin.

We highlights five facts from the LBD data.

Fact 1. Increase in Industry Concentration

Our first fact is the increase in the average industry concentration. Figure 1 shows that the weighted average of the cumulative change in the log employment share of the top 10% of firms in each of our 450 industries increased by .09 log points from 1977 to 2013.10 This fact echoes Autor et al. (2017)’s finding that the share of the top 4, 20, and

\[
\sum_{j=1}^{J} \sum_{k=1}^{L_j} \Delta \log L_{jk}^{j}
\]

where \( J \) denotes the set of industries, \( L_{j} \) denotes employment in industry \( j \), and \( \Delta \) is the change between 1977 and 2013.

\(^{10}\)We use the share of the top 10% because the smallest number of firms in an industry in our sample is 10. The weights are Sato-Vartia weights for each industry \( j \) between 1977 and 2013 defined as \( \sum_{k=1}^{L_k} \Delta \log L_{jk} \)
40 firms has increased in the average sector over a similar time period.\textsuperscript{11}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{\(\Delta\) Employment Share of Top 10\% Firms}
\end{figure}

Note: Figure shows weighted average of cumulative change of the log employment share of top 10\% firms in each four digit industry from 1977 to 2013, weighted by Sato-Vartia employment share of each four-digit industry over these two years.

Table 1 shows the change in the log employment share of the top 10\% firms from 1977 to 2013 for each broad industry. The increase in average concentration is primarily driven by three sectors: wholesale, services, and retail, where the average employment share of the top 10\% firms increased by about .14 log points between 1977 and 2013. In contrast, concentration in manufacturing fell over this time period. These last two facts are also shown in Figure 1, which shows the cumulative change of the top 10\% of firms in the average manufacturing sector and in the average retail, wholesale, and service sector.

**Fact 2. Heterogeneity in the Increase in Industry Concentration**

Table 2 shows the distribution of the change in the employment share from 1977 to 2013 across our 450 industries. The first row shows the overall distribution. The 90-10 gap in the change in the top 10\% share is about .35 log points. The second row shows the dispersion within broad 1-digit sectors. Given the heterogeneity in the mean change in

\textsuperscript{11}Autor et al. (2017) use the micro-data from the Economic Censuses.
Table 1: $\Delta$ Top Firm Share, 1977-2013

<table>
<thead>
<tr>
<th>Industry</th>
<th>$\Delta$ log employment share</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Industries</td>
<td>.088</td>
</tr>
<tr>
<td>Wholesale</td>
<td>.160</td>
</tr>
<tr>
<td>Retail</td>
<td>.145</td>
</tr>
<tr>
<td>Services</td>
<td>.116</td>
</tr>
<tr>
<td>Construction</td>
<td>.075</td>
</tr>
<tr>
<td>Finance</td>
<td>.049</td>
</tr>
<tr>
<td>Utilities and Transportation</td>
<td>.019</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-.018</td>
</tr>
</tbody>
</table>

Note: Weighted average of $\Delta$ log employment share from 1977-2013 of top 10% firms in 4-digit industries within each large sector, weighted by Sato-Vartia average of the employment share of each 4-digit industry in 1977 and 2013.

broad sectors shown in Table 1, it is not surprising that the residual dispersion is smaller than the overall dispersion. Still, the residual dispersion is sizable. The 90-10 gap in the change in residual industry concentration is about .31 log points. Hence, even within sectors such as services, where concentration has increased on average, there is still substantial dispersion in the change in concentration.

Table 2: Distribution of $\Delta$ Top Firm Share, 1977-2013

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>75-25</th>
<th>90-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>.065</td>
<td>.174</td>
<td>.350</td>
</tr>
<tr>
<td>Within 1-Digit Sector</td>
<td>.168</td>
<td>.311</td>
<td></td>
</tr>
</tbody>
</table>

Note: Table shows the median, 90-10 gap, and 75-25 gap in the distribution of $\Delta$ log employment share of top 10% firms in 450 industries from 1977 to 2013. Row 2 shows the dispersion within 1-digit sectors.

Fact 3. Industry Employment and Establishment per Firm Rise with Concentration

Figure 2 plots the non-parametric relationship between the change in the employment share of the industry against the change in the share of the top firms in the industry (both from 1977 to 2013). The left panel, which uses the variation across all 450 industries in our sample, shows a clear relationship between industry growth and growth
of the top firms in the industry. Table 3 (row 1, column 1) shows that an OLS regression of this relationship yields a precisely estimated elasticity of 2.577 (s.e. = 0.296).

Figure 2: ∆ Industry Employment vs. ∆ Share of Top Firms

<table>
<thead>
<tr>
<th>All Sectors</th>
<th>Non-Manufacturing</th>
</tr>
</thead>
</table>

![](image)

Note: Figure shows point estimate and 99% confidence interval of non-parametric regression of ∆ log employment of the industry as a share of aggregate employment in the economy on ∆ log employment of top 10% firms in the industry as a share of total industry employment, both from 1977-2013. Left panel includes all 450 4-digit industries; right panel excludes manufacturing industries.

The right panel in Figure 2 restricts the sample to the non-manufacturing industries. Using only the variation within non-manufacturing, the relationship between industry employment and top firm concentration is weaker but still positive and statistically significant. The elasticity of total employment in a non-manufacturing industry to the change in top firm concentration is 1.097 (s.e. = 0.354).

Figure 3 plots the non-parametric relationship between the change in the log number of establishments per firm (for all firms in the industry) and the change in the log employment share of the top 10% of firms in the industry. Table 3 (row 1, column 2) shows that the OLS coefficient of this relationship is a precisely estimated 0.987 (s.e: 0.085).\(^\text{12}\)

\(^{12}\)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 do not relate the change to measures of industry concentration but show that the increase is more pronounced for larger firms and in the service sector.
Table 3: Elasticity of $\Delta$ Employment and Estab./Firm to $\Delta$ Top Firm Share

<table>
<thead>
<tr>
<th></th>
<th>$\Delta$ Employment</th>
<th>$\Delta$ Establishments/Firm</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Firms in Industry</td>
<td>2.577</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>Bottom 90% Firms</td>
<td>0.193</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.292)</td>
<td>(0.035)</td>
</tr>
</tbody>
</table>

Note: Entries are coefficients and standard errors of regression of $\Delta \log$ total employment (column 1) and $\log$ establishments per firm (column 2) on change in industry concentration from 1977 to 2013. Total employment and establishments per firm are for all firms in the industry (row 1) or only the bottom 90% of firms in the industry (row 2). Industry concentration is log of employment share of top 10% of firms in industry.

Figure 3: $\Delta$ Establishments/Firm vs $\Delta$ Share of Top Firms

Note: Figure shows point estimate and 99% confidence interval of non-parametric regression of $\Delta \log$ # Establishments/Firm of all firms in the industry on $\Delta \log$ employment share of top 10% firms in the industry, both from 1977 to 2013.

The last row in Table 3 shows the elasticity of the change in total employment (column 1) and establishment per firm (column 2) for the bottom 90% of firms in an industry to the change in the share of the top 10% of firms in the industry. The first column shows
that the elasticity of total employment of the bottom 90% of firms is much smaller than the elasticity of total industry employment and not statistically different from zero. On the other hand, the elasticity of establishments per firm of the bottom 90% of firms with respect to the employment share of the top 10% of firms is positive and precisely estimated, although the elasticity is smaller than that of all firms in the industry.

**Fact 4. Industry Concentration is due to Extensive Margin Growth**

We next show that the growth in industry concentration is mostly due to extensive margin growth by the top firms. The change in the employment share of the top firms in an industry can be decomposed into the contribution of growth on the extensive and intensive margins. For example, if we define a market as an MSA, the decomposition is:

$$\Delta \log \frac{L_{top}}{L} = \Delta \log \frac{\# MSA_{top}}{\# MSA} + \Delta \log \frac{L_{top}}{\# MSA}$$  \hspace{1cm} (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).

Table 4 shows the results of this decomposition for the relative number of MSAs vs employment per MSA (row 1), relative number of counties vs. employment per county (row 2), and relative number of establishments vs. employment per establishment (row 3). The first row shows that 99% of the growth in concentration comes from growth in the number of cities served by top firms, and only about 1% comes from increased employment per city. The next two rows show that average employment per county and per establishment of top firms falls. So necessarily more than 100% of concentration growth has to come from the increase in the number of counties and establishments served by the top firms.

Figure 4 plots the non-parametric relationship between our three measures of extensive vs. intensive margin growth of top firms. The left panel shows the point estimate of a non-parametric regression of our three measures of extensive margin growth ($\#$ of MSAs, counties or establishments) of the top firms vs. overall growth in the employment share of the top firms. The slopes of all curves are positive, indicating that in industries where top firms have expanded the most, they have done so by expanding
Table 4: \( \Delta \) Concentration Due to Extensive vs. Intensive Margin Growth

<table>
<thead>
<tr>
<th></th>
<th>Extensive</th>
<th>Intensive</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSAs</td>
<td>0.993</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Counties</td>
<td>1.261</td>
<td>-0.261</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Establishments</td>
<td>1.538</td>
<td>-0.538</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.095)</td>
</tr>
</tbody>
</table>

Note: Column 1 shows point estimates and standard errors of a regression of \( \Delta \) log # of MSAs/Counties/Establishments of top 10% firms relative to all firms from 1977-2013 on \( \Delta \) log employment share of top 10% firms from 1977-2013. Column 2 shows the same for a regression of \( \Delta \) log employment per MSAs/County/Establishment of top 10% firms relative to all firms from 1977 to 2013 on the same independent variable.

geographically into more establishments, counties, and MSAs. The slope increases as we adopt narrower definitions of a market. It is the smallest for MSAs and the largest for establishments. 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.

The right panel in Figure 4 shows the non-parametric regression of intensive margin growth (employment per MSA, per county, and per establishment of the top firms). The figure shows that employment per establishment declines by more in industries where top firm concentration has increased by more. The same is true for employment per county, although the magnitude of the decline with respect to the change in industry concentration is smaller. And when we look at employment per MSA, the intensive margin growth of top firms is essentially zero.

Table 5 shows the contribution of extensive margin growth to top firm concentration in the three broad sectors where concentration has grown the most: services, retail, and wholesale. Specifically, we decompose the variation in the change in top firm concentration across 4-digit industries within each of these broad sectors, and show the contribution of extensive margin growth in Table 5. For comparison, the first column reproduces the contribution of extensive margin growth across all industries shown earlier in Table 4. The table shows that contribution of extensive margin growth is
The growth in the number of MSAs served by top national firms in services and retail is 30% larger than the growth in their national market share. And the growth in the number of establishments is twice as large as the growth of the national market share of top retail and service firms. So in retail and services, top national firms are much smaller in each local market. This is less true in wholesale, where some of the growth of top national firms has been due to intensive margin growth.

Table 6 probes for evidence that top firms have expanded into smaller more marginal local markets. Specifically, we measure the size of the local market as total employment (in all industries) in the county or 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 6 shows the regression of Δ log size of the local market of a top firm in the industry relative to the size of the local market of an average firm in the industry on Δ log employment share of the top firm in the industry (both are calculated from 1977 to 2013). Table 6 shows that the elasticity of the change in the relative size of the market of top firms particularly large in services and retail. The growth in the number of MSAs served by top national firms in services and retail is 30% larger than the growth in their national market share. And the growth in the number of establishments is twice as large as the growth of the national market share of top retail and service firms. So in retail and services, top national firms are much smaller in each local market. This is less true in wholesale, where some of the growth of top national firms has been due to intensive margin growth.

Table 6 probes for evidence that top firms have expanded into smaller more marginal local markets. Specifically, we measure the size of the local market as total employment (in all industries) in the county or 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 6 shows the regression of Δ log size of the local market of a top firm in the industry relative to the size of the local market of an average firm in the industry on Δ log employment share of the top firm in the industry (both are calculated from 1977 to 2013). Table 6 shows that the elasticity of the change in the relative size of the market of top firms particularly large in services and retail. The growth in the number of MSAs served by top national firms in services and retail is 30% larger than the growth in their national market share. And the growth in the number of establishments is twice as large as the growth of the national market share of top retail and service firms. So in retail and services, top national firms are much smaller in each local market. This is less true in wholesale, where some of the growth of top national firms has been due to intensive margin growth.

Table 6 probes for evidence that top firms have expanded into smaller more marginal local markets. Specifically, we measure the size of the local market as total employment (in all industries) in the county or 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 6 shows the regression of Δ log size of the local market of a top firm in the industry relative to the size of the local market of an average firm in the industry on Δ log employment share of the top firm in the industry (both are calculated from 1977 to 2013). Table 6 shows that the elasticity of the change in the relative size of the market of top firms
Table 5: Extensive Margin Growth in Wholesale, Retail, and Services

<table>
<thead>
<tr>
<th></th>
<th>All Industries</th>
<th>Services</th>
<th>Retail</th>
<th>Wholesale</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSAs</td>
<td>0.993</td>
<td>1.314</td>
<td>1.261</td>
<td>0.896</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.157)</td>
<td>(0.185)</td>
<td>(0.210)</td>
</tr>
<tr>
<td>Counties</td>
<td>1.261</td>
<td>1.605</td>
<td>1.574</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.172)</td>
<td>(0.187)</td>
<td>(0.248)</td>
</tr>
<tr>
<td>Establishments</td>
<td>1.538</td>
<td>1.908</td>
<td>2.009</td>
<td>0.812</td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.186)</td>
<td>(0.208)</td>
<td>(0.286)</td>
</tr>
</tbody>
</table>

Note: Table shows point estimates and standard errors of a regression of $\Delta \log \#$ of MSAs/Counties/Establishments of top 10% firms relative to all firms from 1977-2013 on $\Delta \log$ employment share of top 10% firms from 1977-2013.

Table 6: Elasticity of $\Delta$ Market Size of Top Firms to $\Delta$ Concentration

| $\Delta$ MSA Size of Top Firms/All Firms | -.273 |
|                                         | (.056) |
| $\Delta$ County Size of Top Firms/All Firms | -.425 |
|                                         | (.091) |

Note: Table presents regression of $\Delta \log$ market size (counties or MSAs) of top firms in the industry relative to market size of average firm in the industry on $\Delta \log$ concentration in the industry. Market size is average total employment (in all industries) in a local market; concentration is employment share of top 10% firms in industry. $\Delta$ market size and concentration calculated from 1977 to 2013.

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 local markets. 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 headquarters and larger than its existing markets.

The expansion of top firms through entry in new and, on average, smaller markets implies that their presence in cities that were small in 1977 should have increased. Figure 5 presents the average employment weighted share of top 10% national firms in an industry in each MSA by total employment of the MSA in 1977. The left panel shows this for top national firms in all industries. 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, like Missoula, MT (employment 19 thousand 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. The right panel in Figure 5 focuses on top national firms in manufacturing. As can be seen, the presence of top firms in manufacturing varies little with city size and did not change much between 1977 and 2013.

**Figure 5:** Local Employment Share of Top National Firms

All Industries | Manufacturing Only

![Graph showing local employment share of top national firms in all industries and manufacturing.](image)

Employment in MSA in 1977

Note: Figure shows point estimates and 99% confidence intervals of non-parametric regression of total employment of top national firms in all industries (left panel) or in manufacturing (right panel) as a share of total employment in each MSA in 1977 and 2013 against total employment in the MSA in 1977.

**Fact 5. No Change in Concentration in Aggregate Economy**

All our facts up until now have been about the top firms in an industry. Our last fact is about the top firms in the aggregate economy. The difference between the two is that top firms in the aggregate economy are in multiple industries. Consider General Mills in 1980. This company was one of the 100 largest companies in the US in 1980. As was typical among the largest US companies at the time, General Mills operated in multiple sectors, primarily packaged foods, retail, toy manufacturing, and restaurants.
Its products in 1980 included Cheerios, Wheaties, Talbots, J.Crew, Monopoly, Play-Doh, and Red Lobster. In 1995, after having earlier sold its retail and toy businesses, General Mills was split in two companies, one of which focused on packaged foods and the other on full-service restaurants. The latter is now called Darden Restaurants, and has grown substantially since 1995. Darden is currently the largest operator of full-service restaurants in the US, with over 1,700 restaurants and 180 thousand employees.\textsuperscript{13}

Table 7 shows the employment share of the top 0.1%, 0.01% and 0.001% of firms in the aggregate economy in 1977 and 2013. Although the employment share of top firms such as Darden in their industries has increased substantially, the employment share of the top firms in the aggregate economy has not. The share of a top company such as Darden in the aggregate economy today is not much different from the share of a top company such as General Mills four decades ago. Gutierrez and Philippon (2019) report a similar fact from data on publicly traded firms in Compustat. In our LBD data, the employment share of the top .001% firms increased from 8.3% in 1977 to only 8.5% in 2013.

\begin{table}
\centering
\begin{tabular}{lcc}
\hline
 & 1977 & 2013 \\
\hline
Top 0.1\% & 40.7\% & 41.3\% \\
Top 0.01\% & 21.9\% & 22.5\% \\
Top 0.001\% & 8.3\% & 8.5\% \\
\hline
\end{tabular}
\caption{Employment Share of Top Firms in Aggregate Economy}
\end{table}

Note: Table shows the employment share of top 0.1\%, 0.01\%, and 0.001\% firms in the aggregate economy.

Table 8 reconciles the stable top firm share in Table 7 with increasing concentration at the industry level. The first two columns shows the average number of industries per firm among the top firms relative to the number of industries per firm of all firms in the economy. Top firms produce in more industries than the average firm, but less so in 2013 compared to 1977. The number of industries of a top 0.001\% firm (relative to the average firm) fell from 36.5 in 1977 to 17.4 in 2013. The corresponding number for a top

\textsuperscript{13}Darden’s national chains include Olive Garden, Longhorn Steakhouse, Cheddar’s Scratch Kitchen, Yardhouse, Capital Grille, Seasons 52, Bahama Breeze, Eddie V’s. Darden sold Red Lobster to a private equity company in 2014.
Table 8: Concentration within Top Firms

<table>
<thead>
<tr>
<th></th>
<th># Industries</th>
<th></th>
<th>HHI of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Top/All Firms</td>
<td></td>
<td>Top Firms</td>
</tr>
<tr>
<td>Top 0.1%</td>
<td>7.7 5.0</td>
<td>0.617 0.721</td>
<td></td>
</tr>
<tr>
<td>Top 0.01%</td>
<td>22.6 10.0</td>
<td>0.408 0.658</td>
<td></td>
</tr>
<tr>
<td>Top 0.001%</td>
<td>36.5 17.4</td>
<td>0.374 0.682</td>
<td></td>
</tr>
</tbody>
</table>

Note: Table shows average industries per firm of top firm (in overall economy) relative to the average firm (columns 1-2) and average HHI across the sectors of each top firm.

0.01% firm is 22.6 industries in 1977 and 10 industries in 2013. The last two columns show the Herfindahl-Hirschman Index (HHI) of a top firm. Specifically, we calculate a firm's HHI as the sum of the square of the firm's employment in each industry as a share of the firm's total employment. Not surprisingly the HHI of a top 0.001% firm is larger than that of a top 0.1% firm, but the gap is noticeably smaller in 2013 compared to 1977. Over this time period, the HHI of a top 0.001% firm almost doubled from 0.374 to 0.682.

Going back to the example of General Mills and Darden, General Mills was a diversified conglomerate in 1980 and Darden today only runs restaurants.

Table 9 shows that top firms are also larger in the industries they have chosen to specialize in. Columns 1 and 2 shows average employment per industry of a top firm relative to average employment per industry of all firms. Employment per industry of a top 0.001% firm in 1977 was 230 times larger than that of the average firm. By 2013, the size gap had more than doubled to 494. Columns 3 and 4 shows that the industries that top firms have chosen to specialize in are primarily ones with growing concentration. Specifically, the table shows the employment of top firms in industries with above-median concentration growth as a share of total employment of the top firm. For example, 11% of employment of the top 0.001% firms in the overall economy in 1977 were in industries with growing concentration. By 2013, almost half of employment of the top 0.001% firms were in such industries. In summary, top firms are now more specialized, are larger in the chosen industries, and these are precisely the industries
where concentration has grown.

<table>
<thead>
<tr>
<th>Top Firms in Top Firm Employment</th>
<th>1977</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 0.1%</td>
<td>53.0</td>
<td>82.1</td>
</tr>
<tr>
<td>Top 0.01%</td>
<td>97.4</td>
<td>225.1</td>
</tr>
<tr>
<td>Top 0.001%</td>
<td>229.4</td>
<td>493.8</td>
</tr>
</tbody>
</table>

Note: Columns 1 and 2 show employment per industry of average top firm (in overall economy) relative to the average firm. Columns 3 and 4 show employment of top firms in industries with above median concentration growth between 1977 and 2013 as a share of the top firm's total employment.

3. A 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. The main purpose of the theory is 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 U.S. economy between 1977 and 2013.

3.1. The model

Consider a firm $i$ that produces a good $j$. The firm uses plants to produce in different locations $n$, out of a continuum of locations with mass $N$. The price of good $j$ in location $n$ is given by $p_{jn}$. Assume that the only way to serve market $n$ is to put a plant there. A firm pays a fixed cost $F_j$ (in units of the numeraire) to produce good $j$ and another fixed cost $w_n f_n$ (in units of the numeraire, but with a magnitude indexed to the local wage) to set up an establishment in market $n$. The firm's productivity $A_{ij}$ applies to its establishments in all locations. Labor is the only factor of production, so a firm that
hires $L_{ijn}$ units of labor produces $Y_{ijn} = A_{ij}L_{ijn}$ units of output with local revenues given by $R_{ijn} = p_{jn}A_{ij}L_{ijn}$.

Now suppose that demand is CES and firms compete monopolistically, then $p_{jn} = E_nY_{ijn}^{-\frac{1}{\sigma}}$, where $\sigma > 1$ is the elasticity of substitution across varieties within an industry and $E_n$ is a function of local expenditure determined in equilibrium. Conditional on serving market $n$, profit maximizing employment in the local market is given by:

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

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

$$A_{ij} \geq \alpha \left( \sigma, f_n, \frac{E_n}{w_n} \right) \equiv \left( \frac{f_n}{\tilde{\sigma} \left( \frac{E_n}{w_n} \right)^\sigma} \right)^{-\frac{1}{\sigma-1}} \quad (3)$$

where $\tilde{\sigma} \equiv (\sigma - 1)^{\sigma-1}/\sigma^\sigma$. Hence, the firm is more likely to enter a market where the local fixed cost $f_n$ is lower, wages $w_n$ are smaller, and total expenditures $E_n$ are larger.

Suppose that the mass of markets with characteristic $\alpha < \alpha_n$ is given by $\Gamma (\alpha_n)$ with density $\gamma (\alpha_n)$. This distribution $\Gamma (\cdot)$ is determined by parameters, the set of available markets, and the distribution of $E_n/w_n$, which is determined in equilibrium. If all markets are identical and there is free mobility across markets, then the distribution of $E_n/w_n$ would be degenerate and therefore the distribution $\Gamma (\cdot)$ would be degenerate as well. If markets are different in terms of amenities, productivity, housing and other factors, or a variety of other frictions, the distribution of $E_n/w_n$ across markets would not be degenerate, even with free mobility, apart from cases with extreme assumptions on the distribution of the rents of local factors. Limited or frictional mobility would also yield a non-degenerate distribution of $E_n/w_n$. Here, we stop short of specifying a fundamental model of the distribution $\Gamma (\cdot)$ to gain generality and simplify the exposition.

Figure 6 depicts the markets in which the firm will be active given a density $\gamma (\cdot)$ and, only to simplify the illustration, with a constant $f_n$ and $\sigma$ that we omit in the notation. Then, the only relevant local characteristic is the ratio $E_n/w_n$, where a higher ratio means that the market is more profitable. Suppose $\gamma (\cdot)$ is increasing in $\alpha$ and therefore
decreasing in $E_n/w_n$. That is, more desirable markets are more scarce. Firm $i$ will choose to sell product $j$ in all markets for which $A_{ij} > \alpha (E_n/w_n)$, namely $E_n/w_n \geq \alpha^{-1} (A_{ij})$ since $\alpha (\cdot)$ is decreasing in $E_n/w_n$.

**Figure 6: Determination of Active Local Markets**

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,

$$
\int_{n \text{ s.t. } A_{ij} > \alpha_n} [w_n \tilde{\sigma} A_{ij}^{\sigma-1} (\frac{E_n}{w_n})^{\sigma} - w_n f_n] dn - F_j > 0,
$$

where $\alpha_n \equiv \alpha (\sigma, f_n, E_n/w_n)$ is defined in (3). If we further assume that the fixed cost of opening a plant in a local market $w_n f_n$ is constant at $f$, the profits of a firm that enters industry $j$ are given by

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

Denote by $A (F_j, f, \Gamma, \sigma)$ the productivity level such that $\Pi \left( A (F_j, f, \Gamma, \sigma), F_j, f, \Gamma, \sigma \right) = 0$. Since $\Pi$ is increasing in $A_{ij}$, there is a unique such value. Therefore, active firms in industry $j$ are such that $A_{ij} \geq A \left( F_j, f, \Gamma, \sigma \right)$. 

It is useful to illustrate the intuition behind the cutoff productivity with a parametric example. Suppose that the density of markets with characteristic $\alpha$ is $\gamma(\alpha) = \Omega \alpha^a / f$ for some $a > \sigma - 1$. This distribution implies that there are many markets where it is hard to enter, and the more so when $a$ is larger. $\Omega > 0$ is related to the availability of good markets to enter and is determined in general equilibrium by the level of wages and expenditures. Then total profits from product $j$ are given by $\Pi(A_{ij}, F_j, f, \Gamma, \sigma) = \tilde{a} \Omega A_{ij}^{1+a} - F_j$ where $\tilde{a} = \left[ \frac{\sigma - 1}{(2 + a - \sigma)(1 + a)} \right]^{14}$. So active firms in industry $j$ are those with productivity $A_{ij} \geq A(F_j, f, \Gamma, \sigma) \equiv (F_j \tilde{\alpha})^{1/(1-\alpha)}$. The productivity threshold is decreasing in $\Omega$ since having more profitable markets implies a lower entry productivity threshold.

### 3.2. A new technology

Sutton (1991) argues new sunk cost technologies leads to market concentration. We now borrow this idea and examine the effect of a new technology that increases the fixed costs of producing a given good in exchange for a reduction in the variable cost (and leaves the fixed cost of creating plants, $f$, constant). Namely, adopting the new technology results in an increase in fixed costs to $h^\eta F_j$ and an increase in productivity to $h A_{ij}$, for $h > 1$ and $\eta > 0$.

Firms will adopt the technology if $\Pi(hA_{ij}, h^\eta F_j, f, \Gamma, \sigma) \geq \Pi(A_{ij}, F_j, f, \Gamma, \sigma)$. This condition can be rewritten as the sum of the profits of the firm in new markets plus the increased profits in old markets being greater or equal than the increase in the firm-product fixed cost relative to local fixed costs, namely,

$$\int_{A_{ij}}^{hA_{ij}} \left( \frac{h A_{ij}}{\alpha} \right)^{\sigma-1} \Gamma(d\alpha) + \int_0^{A_{ij}} \left( \frac{A_{ij}}{\alpha} \right)^{\sigma-1} (h^{\sigma-1} - 1) \Gamma(d\alpha) \geq (h^\eta - 1) \frac{F_j}{f}. \quad (4)$$

How do the benefits of adopting the new technology, the left-hand side of condition (4), change with firm productivity? Taking the derivative of the left-hand side of condi-

---

\[14\] Note that we defined the level of the distribution $\gamma(\cdot)$ as $\Omega / f$. Other definitions lead to a value of the constant $\tilde{a}$ that depends on $f$. 

tion (4) with respect to $A_{ij}$ yields:

$$\frac{\partial LHS}{\partial A_{ij}} = \int_0^{hA_{ij}} \sigma - 1 \left( \frac{h}{\alpha} \right)^{\sigma-1} A_{ij}^{\sigma-2} \Gamma (d\alpha) - \int_0^{A_{ij}} \sigma - 1 \left( \frac{1}{\alpha} \right)^{\sigma-1} A_{ij}^{\sigma-2} \Gamma (d\alpha)$$

$$= (\sigma - 1) A_{ij}^{\sigma-2} \int_{A_{ij}}^{hA_{ij}} \left( \frac{h}{\alpha} \right)^{\sigma-1} \Gamma (d\alpha)$$

$$+ (\sigma - 1) A_{ij}^{\sigma-2} \int_0^{A_{ij}} \left( \frac{1}{\alpha} \right)^{\sigma-1} \left( h^{\sigma-1} - 1 \right) \Gamma (d\alpha) > 0$$

since, by Leibniz rule, the derivative of $A_{ij}$ in the limit of the integral is equal to zero. Hence, the gains from adopting the new technology increase with a firm’s productivity, while the costs (the right-hand side of condition 4) are fixed. This implies that there exists a threshold $H (F_j, f, \Gamma, \sigma, h, \eta)$ such that if $A_{ij} \geq H (F_j, f, \Gamma, \sigma, h, \eta)$, firm $i$ adopts the new technology. The function $H (\cdot)$ is increasing in $F_j$ and $\eta$, as simple inspection of (4) indicates.

Some firms will decide not to adopt the new technology as long as $A (h^n F_j, f, \Gamma, \sigma) / h > A (F_j, f, \Gamma, \sigma)$. This is not always the case. In our parametric example with $\gamma (\alpha) = \Omega \alpha^a / f$, it requires $\eta > 1 + a$. This is because $A (F_j, f, \Gamma, \sigma) = \left( \frac{F_j}{\Omega} \right)^{\frac{1}{1+a}}$ and so

$$A (h^n F_j, f, \Gamma, \sigma) / h = h^{\frac{n}{1+a} - 1} \left( \frac{F_j}{\Omega} \right)^{\frac{1}{1+a}} > A (F_j, f, \Gamma, \sigma),$$

if $\eta > 1 + a$, since $h > 1$. Namely, some firms do not adopt if the elasticity of fixed cost to $h$ is larger than one plus the elasticity of the density of tougher markets (higher $1 + \alpha$).

More generally, we need the increase in fixed costs in the new technology to be large enough.

We have proven the following proposition:

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

The introduction of a new technology $h$ in the model above is consistent with several facts outlined in the previous section. In particular, it is consistent with Fact 1, since it leads to top firms adopting the new technology. It is also consistent with Fact 2 if the new technology is not available in all sectors, and partially with Fact 3, since adopting firms expand the number of establishments per firm. However, because there is only one technology level $h$, in the model above firms only have a binary decision: adopt or not. Fact 3 shows us that in industries where concentration has increased, which we interpret as industries where a new technology $h$ is available, the number of establishments of the bottom 90% of firms also increased. It seems that in these industries, firms can choose to adopt the new technology with different degrees of intensity. We introduce this margin in our model next.

### 3.2.1. A menu of new technologies

Consider the case where firms can choose among a continuum of technologies indexed by $h \geq 1$. If they choose $h = 1$, they maintain the old technology, while if they choose $h > 1$, they choose a technology with the characteristics described above. As before, firm fixed costs are given by $h^n F_j$ for $n > 0$.

Proposition 1 shows that, given a single $h$, high productivity firms will adopt while others might not. Here we show that the gains from adopting a new technology $h$ are not only increasing in firm productivity, but that the increase with firm productivity grows with $h$. That is, the cross-derivative of the left-hand side of condition (4) is positive. Since the right-hand side of condition (4) does not depend on $h$, this implies that more productive firms choose technologies with higher $h$. Namely, if we denote by $\phi (A)$ the technology adopted by firms with productivity $A$, $\phi (A) \geq \phi (A')$ for $A \geq A'$.

The cross-derivative of the left-hand side of (4) is given by

$$\frac{\partial^2 LHS}{\partial A_{ij} \partial h} = (\sigma - 1) \gamma (h A_{ij}) + (\sigma - 1)^2 A_{ij}^{\sigma - 2} \int_0^{h A_{ij}} h^{\sigma - 2} \left( \frac{1}{\alpha} \right)^{\sigma - 1} \Gamma (d\alpha) > 0.$$  

Note also that the right-hand side of equation (4), $(h^n - 1) F_j / f$, is increasing in $h$ for
\( \eta > 0 \) and the slope grows with \( \eta \) for \( h > 1 \).\(^{15}\) Furthermore, as \( \eta \to 0 \), \( h^\eta - 1 \to 0 \), and the derivatives above converge to zero. Hence, since the left-hand side of (4) is strictly positive for \( h > 1 \), there exists a threshold \( \eta_0 \) such that if \( \eta < \eta_0 \), low productivity firms also adopt a new technology, although with weakly lower \( h \). In our example, this threshold is such that \( \eta_0 = 1 + a \), as proven above. Hence we have proven the following proposition:

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

The results above allows us to match the fact, described as part of Fact 3 above, that in industries where concentration by top firms has increased, even the smaller firms in the industry increase their number of establishments. We now elaborate on these results when we keep the distribution of local markets constant.

### 3.3. Equilibrium implications for a fixed distribution of local markets

Absent general equilibrium effects that determine the distribution \( \Gamma \), firms that adopt a better technology \( h \), which are the more productive firms, enter more locations. This is simply implied by \( \Gamma(hA_{ij}) \) increasing in \( hA_{ij} \). Furthermore, when the new technology is available, the difference between the number of markets of productive and unproductive firms increases. Namely,

\[
\frac{\partial \Gamma (\phi(A) A)}{\partial A} \bigg|_{\phi(A)=1} = \gamma (A) [\phi' (A) A + 1] > \gamma (A) = \frac{\partial \Gamma (A)}{\partial A}
\]

since \( \phi' (A) > 0 \) by Proposition 2.

Note also that, absent general equilibrium effects, firms that adopt a better technology have larger establishments, since \( L_{ijn} = (\sigma - 1) \int \left( \frac{hA_{ij}}{\alpha n} \right)^{\sigma - 1} \). Furthermore, even

\(^{15}\)The derivatives are given by \( \frac{\partial (h^\eta - 1)}{\partial h} = \eta h^{\eta - 1} F_i > 0 \), and \( \frac{\partial (h^\eta - 1)}{\partial h \partial \eta} = [\eta h^{\eta - 1} + \eta h^{\eta - 1} \log h] F_i > 0 \) for \( h > 1 \).
though the new markets where the firm enters are less profitable, the increase in productivity due to the new technology implies that the marginal market has constant employment size. Employment size in the firm’s marginal market when $h A_{ij} = \alpha_n$ is $(\sigma - 1) f$, which does not depend on $h$. This is illustrated in Figure 7. Since firms with better technology enter more markets and have more employees per establishment, their employment share necessarily increases when $h > 1$.

Figure 7: The Effect of $h > 1$ on Employment and Market Entry

Finally, note that in industries where the new technology is very good ($\eta$ is low) there is more concentration since top firms will adopt a larger $h$. In those industries, less productive firms will also adopt a better technology, although with a lower $h$ (see Proposition 2). Hence, in a multi-industry economy with elasticities of substitution across industries greater than one, employment of the whole industry will increase, including employment of the bottom firms.

The arguments above establish the following result for a single industry in partial equilibrium:

**Proposition 3** Given the distribution of markets $\Gamma$, the new menu of technologies results in more concentration of employment in more productive firms. Firms, and more productive firms in particular, enter more markets and are larger in each of them. The effects
are more pronounced for small values of $\eta$, with no effect if $\eta$ is very large (since there is no adoption).

Hence, our model can match Facts 1 to 4 if we identify a market $n$ in the theory as a city (MSA) in the data. In that case, national concentration in the industry rises because top firms enter more markets with a larger scale. As Table 4 shows, this is the case for MSA's but not for counties or establishments. Matching Fact 4 for these narrower geographic units, requires us to tweak the effects of the new technology further. We do so in the next section.

3.4. A technology that reduces local fixed costs too

The model above implies that the advent of the new technology brings 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 county or a single establishment. To generate declines in the average employment size of adopters we need to allow the new technology 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 $f h^{-\varphi}$. The exponent $\varphi > 0$ determines the extent to which fixed costs decline with the new technology. The exponent should depend on the definition of a market. For a large geographic area we might think that the cost did not change much beyond the overall firm fixed costs, and so $\varphi = 0$. For a smaller area, like a county or a single establishment, $\varphi > 0$, due to the ease in replicating standardized establishments (as exemplified by companies like Starbucks).

Note from equation (2) that, given $E_n/w_n$, the fixed cost does not affect establishment sizes directly. It does, however, determine entry into marginal markets and the size of the smallest establishment of the firm, which is given by $(\sigma - 1) f h^{-\varphi}$. For firms that choose $h > 1$, this implies that the smallest establishment size of the firm falls. Hence, with the new technology, average establishment sizes of existing firms fall for $\varphi$ large enough. This is illustrated in Figure 8.

Finally it is useful to realize that if $\eta$ and $\varphi$ are high enough, average firm size necessarily falls, since firms choose a small $h$ and so the establishments in the best markets
only increase marginally in size. In contrast, for $\varphi$ large, the firm adds many new markets with smaller establishments. This is consistent with Fact 4 in the previous section, if we interpret a market as a county or the area served by a single establishment. We summarize these results in the following proposition.

**Proposition 4** Given the distribution of markets $\Gamma$, if the new technology also reduces fixed cost to $fh^{-\varphi}$, the minimum employment size of the firm’s establishments falls, and average establishment size falls if $\eta$ and $\varphi$ are large enough.

### 3.5. Multi-product firms and general equilibrium

Consider a firm that owns a collection of $J$ technologies in a variety of industries $j$ with productivity $A_{ij}$. Each variety requires the firm to pay a fixed cost $F_j$. When the firm obtains access to the new technology in each sector (with potentially different parameters $\eta_j$ and $\varphi_j$ in each sector), it will make a set of upgrading decision $\phi_j (A_{ij})$. In some of its industries the firm might decide to set $\phi_j (A_{ij}) = 1$ (e.g. if $\eta_j$ is very high or its productivity is too low), or exit.

Denote total industry employment by $L_j$, and economy wide employment, which we assume fixed, by $\bar{L} = \sum_{j=1}^{J} L_j$. Then if, for example, there is free mobility, the labor
market clearing condition is given by

\[
\bar{L} = \sum_{j=1}^{J} \int_{\Delta(F, f)}^{\infty} \int_{0}^{A} \left( (\sigma - 1) f \left( \frac{\phi_j(A)}{\alpha} \right)^{\sigma-1} \right) \Gamma_j(d\alpha) \Phi(dA), \tag{5}
\]

where

\[
L_{ij} = \int_{0}^{A_{ij}} \left( (\sigma - 1) f \left( \frac{\phi_j(A_{ij})}{\alpha} \right)^{\sigma-1} \right) \Gamma_j(d\alpha),
\]

denotes firm \( i \)'s employment in industry \( j \),

\[
L_j = \int_{\Delta(F, f)}^{\infty} \int_{0}^{A} \left( (\sigma - 1) f \left( \frac{\phi_j(A)}{\alpha} \right)^{\sigma-1} \right) \Gamma_j(d\alpha) \Phi(dA),
\]

denotes total employment in industry \( j \), and

\[
\sum_{j=1}^{J} L_{ij} = \sum_{j=1}^{J_i} \int_{0}^{A_{ij}} \left( (\sigma - 1) f \left( \frac{\phi_j(A_{ij})}{\alpha} \right)^{\sigma-1} \right) \Gamma_j(d\alpha),
\]

denotes total firm employment. Note that we have recognized above that the distribution \( \Gamma_j \) varies by industry since it depends on expenditure in the industry, \( E_{jn} \).

Overall, as long as \( (\eta_j, \phi_j) \) is such that when the new technology is available some firms adopt, \( \phi_j(A_{ij}) > 1 \), in some industries, those firms will hire more employees for a given distribution of locations \( \Gamma_j \), as described in Proposition 3. Hence, in equilibrium, the distribution \( \Gamma_j \) has to shift out to satisfy the labor market equilibrium condition, as apparent from condition (5). The shift happens through an increase in the level of wages, which in the specific example of \( \Gamma_j \) used above is represented by \( \Omega \). Hence, \( \Omega \) decreases which selects some firms out of particular industries. As discussed above, \( A(F, f, \Gamma, \Omega, \sigma) \) is decreasing in \( \Omega \), if it represents the availability of good markets, as in our example.\(^{16}\)

The implication is that if the bottom firms in the industry do not adopt, namely

\(^{16}\)This argument assumes that the distribution \( \Gamma_j \) is invariant to the new technology apart from its level \( \Omega \). This will be the case with free mobility (or heterogeneous preferences) and constant proportional differences in amenities or productivity. Constant factor shares and constant differences in local factor availability will preserve the distribution under free mobility as well. As stated before, we stop short of developing the fully parameterized fundamental model that gives rise to the distribution \( \Gamma_j \) to gain generality and avoid some distracting notational details. A model with autarkic markets where workers cannot move at all would not be consistent with this assumption.
\( \eta_j > \eta_0 \), then some firms will exit. If, on the contrary, everyone adopts, then all outcomes are possible and depend on the total employment expansion of industries, which in turn depends on the elasticity of substitution in consumption between industry aggregates. In what follows we assume that the bottom firms do not adopt (or adopt only marginally), namely, \( \eta_j > \eta_0 \).

As long as agents have CES preferences with elasticity of substitution in consumption greater than one across industry aggregates, this implications translate into overall industry employment. Hence, firms gain employment in industries where they are productive and the new technology is good (e.g. low \( \eta_j \)), and lose employment or exit industries in which they are not as productive and the technology change is also large (low \( \eta \)). The latter implication obtains since the top firm in that industry then upgrade more significantly. In industries with high \( \eta \), nothing happens directly since there is no adoption. The general equilibrium effect makes employment and establishments smaller, makes firms exit local markets, and makes some firms exit the industry.

In sum, the advent of the new technology implies that firms specialize by leaving some markets and investing in others. The total employment size of the firm grows because of the new investments in its main industries, but declines due to tougher competition in its marginal industries and the decline in \( J_i \). The overall implication on employment size is ambiguous. We summarize the findings in the following proposition:

**Proposition 5.** The availability of the new technology makes firms specialize, invest, and grow their employment in industries where they are most productive, and reduce employment or exit industries where they are less productive (if \( \eta_j > \eta_0 \)). The overall effect on firm employment is ambiguous.

Consider a distribution of new technologies \( \eta_j \) across sectors \( J \). Assume that in a group of sectors \( \eta_j \) is large enough so that, as described in Proposition 3, firms do not adopt the new technology given the distribution \( \Gamma_j \). Suppose also, that in other sectors, \( \eta \) is low enough such that some firms adopt. Because the marginal cost of these firms declines \( (w_n/(hA_{ij})) \), the price \( p_{ijn} = (\alpha/(\alpha + 1)) (w_n/(hA_{ij})) \) declines as well since with CES preferences and monopolistic competition markups are constant. Hence, the industry ideal CES price index, \( P_{jn} \), in those industries falls relative to the price index
of industries with higher $\eta_j$ where firms choose a smaller $\phi_j(A_{ij})$. The result is that consumption expands in industries with low $\eta_j$ and contracts in industries with high $\eta_j$. Note that if the elasticity of substitution in consumption across sectors is greater than one, this implies that output and employment shares in industries with low $\eta_j$ expand relative to industries with high $\eta_j$, as evident from equation (5) and the fact that $\Gamma_j$ is a decreasing function of expenditure in the industry, $E_{jn}$. We summarize the result in the following proposition:

**Proposition 6**  *If the technological innovation $\eta_j$ is sufficiently heterogeneous across industries so that $\phi_j(\cdot)$ is not identical for all $j \in J$, and the elasticity of substitution of consumption across industries is greater than one, industries with low $\eta_j$ increase their employment share, while those with high $\eta_j$ contract.*

The results above show that in general equilibrium, our model is consistent with all the facts outlined in the previous section. We do not show that the total employment of top firms does not grow, but the presence of effects in opposite directions implies that the total effect on firm size should be smaller than the effect on overall industrial concentration. Namely, firms get larger in their main sectors, but specialize and drop more marginal ones. The last proposition also illustrates the key parameter leading to the heterogeneity across industries discussed in Fact 2, namely, $\eta_j$.

**4. Implications for Local Markets**

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

The increasing presence of top firms, particularly in the smallest cities, allows local residents to access new varieties of goods and services. In the model we presented in Section 3, the local employment share of a firm $i$, producing product $j$ in market $n$, is given by

$$s_{ijn} = \frac{A_{ij}^{\sigma-1}}{\sum_{i \in I_{jn}} (A_{ij})^{\sigma-1}}$$
where $I_{jn}$ is the set of producers of good $j$ in market $n$. Thus, employment 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.

Nevertheless, within each market, 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 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. Figure 9 (left panel) plots this number 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 almost double that amount (about 40%) in the smallest 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 panel in Figure 9 shows the average change in the Herfindahl–Hirschman Index (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 share of each firm in the industry in the MSA, and 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 concentration has fallen across MSAs of all sizes. This is consistent with the evidence in Rossi-Hansberg et al. (2018) that has emphasized 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.

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 10 shows the local share of employment in 2013 of top firms in industries that were not present in 1977. That is, it measures the extent to which top firms are responsible for bringing new industries to particular cities. The share is as large as 4.5% for the smallest cities in 1977,

---

17 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.
but negligible for the largest cities. Hence, not only are top firms changing the distribution of market shares, $s_{ijn}$, by changing the local distribution of productivities and potentially adding new varieties, but they are also changing the set of industries available. Of course, in our model both margins increases consumer welfare since agents exhibit “taste for variety” modulated by the parameter $\sigma$ for varieties within an industries and $\rho$ for products across industries.

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 one, $\sigma > 1$. 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 70’s. Figure 11 shows that this is indeed the case. On average, the smallest cities in 1977 such as Missoula, MT (employment of 19 thousand 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, that states that city growth is independent of city size.\footnote{See for example Gabaix and Ioannides (2004).} 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 or overwhelmed by overall local employment growth, particularly in small cities were employment growth was faster, as documented in Figure 11. The implications of the industrial revolution in services for job destruction and its variation across cities of different sizes is, therefore, ambiguous. Figure 12 plots the average 5-year job destruction rate between 1977 and 2013 as a function of city sizes.
Figure 11: Local Employment Growth by City Size

![Graph showing local employment growth by city size.](image)

Note: Figure shows ratio of total employment in the MSA in 2013 to employment in the MSA in 1977 (on y-axis) against total employment in the MSA in 1977 (on x-axis).

at the beginning of the sample. The left panel plots job destruction due to firms that exit the MSA, while the right 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 thousand jobs in 1977. That is, there is more job destruction due to exit and incumbent downsizing in large rather than in small cities.

We now estimate the implications of the technological revolution in the service sector for aggregate and local TFP growth. The aggregate measure of TFP in an industry is defined as $TPF_{jt} = Y_{jt}/L_{jt} = R_{jt}/(P_{jt}L_{jt})$. In the data, $L_{jt}$ and $R_{jt}$ can be easily measured, but measuring $P_{jt}$ is complicated since it requires the prices per unit of quality of goods and services sold in each market, plus a methodology to aggregate them across locations. These complications are particularly salient for the service industries, where quality adjusted prices are notoriously hard to measure. In the service sector,

---

19Following 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 $t + 1$. 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.
**Figure 12: Job Destruction Rate by MSA Size**

<table>
<thead>
<tr>
<th>Exiting Firms</th>
<th>Incumbent firms</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
</tbody>
</table>

**Note:** Figure shows the average job destruction rate in each MSA calculated over five-year periods from 1977 to 2013 from exiting firms (left panel) and incumbent firms that shrink (right panel) in each five-year period. Job destruction rate 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 five year period.

the BLS 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 sectors 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 will use Aghion et al. (2019a)’s procedure to measure the growth not captured by the BLS due to entry of new establishments. We differ in that we measure missing growth in each locality, and then aggregate missing growth across all the localities. Specifically, for the set of industries present in a given location in each five year period from 1976 to 2013 (“incumbent industries”), we measure “missing growth” in each industry and location with Aghion et al. (2019a)’s formula. Specifically, for the incumbent industries in a city, “missing growth” due to firm entry can be measured as the weighted average of the product of $1/(\sigma - 1)$ and the change in the log employment share of incumbent establishments in each industry in the city over each 5 year period, where $\sigma$ is the elasticity of substitution across varieties in an industry and the weights are the
Sato-Vartia weights of the industry-city in each five-year time period.

As shown in Figure 10, top firms also create brand new service industries in the cities it enters, and this effect is larger in smaller cities. We calculate missing growth from new industries in a city as the product of $1/(\rho - 1)$ and the inverse of 1 minus the employment share of all new industries in that city, where $\rho$ is the elasticity of substitution across industries. Total missing growth in the given location is the sum of missing growth due to entry by top firms into incumbent industries and into industries that are new to the MSA.

Figure 13 presents the resulting estimates of missing growth for each MSA. The unit on the y-axis is the average annual growth rate per year in each MSA between 1977 and 2013 missed by BLS. The left panel displays missing growth only from industries that were present in the MSA throughout each five year period. Missing growth due to entry of top firms into local markets in incumbent industries is 0.76% per year in small cities but only 0.2% in large ones. The right panel in Figure 13 adds the contribution of missing growth due to the local entry of establishments in new industries. According to this calculation, the BLS’ procedures understate TFP growth by 1.22% per year in the smallest U.S. cities and by a more modest 0.2% per year in the largest cities. Top firms have not brought new industries to the largest cities; they have always been there, so missing growth in large U.S. 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 1977-2013. To be clear, our estimates of “missing growth” only capture the effect of entry of top firms in a locality in a 5 year period. New establishments of top firms could have grown post-entry, and this growth in theory is measured by the BLS. But, of course, it is an open question whether the BLS’ procedures capture quality growth in incumbent service sector establishments.

Our data does not allow us to measure markups or profits of top firms that enter into new local markets. In the model in Section 3, 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, total fixed cost 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 U.S. in the period we study. 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 (forthcoming). In short, an integral part of our hypothesis is the industrial revolution in services leads to rising investments in fixed costs, much 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 rising 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, for example, choosing their location) and choose their price. It shows that, if the dispersion in firm productivities is not too large, the unique sub-game 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 markups produce similar results.

5. Conclusion

Over the last four decades, there has been an increase in industry concentration, a decline in local concentration, and no change in aggregate concentration. We argue that all three facts can be driven by the emergence of new technologies that have enabled firms that adopt them to scale production over a large number of establishments dispersed across space.

Industry concentration rises because firms that learn how to scale the delivery of non-traded services over a large number of locations expand into new local markets and thus capture a larger share of the national market in the industry. Local concentration falls because local markets were typically dominated by local firms that had a large local share but only operated in that one market. Thus, entry of top national firms into new local markets made new services available to local consumers and lowered the market share of previously dominant local service sector firms. Finally, the competitive forces unleashed by the industrial revolution services forces multi-product firms to exit industries where they are less productive. Empirically we see that top firms in the overall economy are more focused and have larger market shares in their chosen sectors, but their size as a share of employment in the overall economy has not changed.

The industrial revolution in services had the largest effect in smaller and mid-sized local markets. Top firms expanded in small local markets, but not in the largest US cities. Over the last four 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 and possibly better 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 the largest in small and mid-sized U.S. cities, and averages 0.5% per year from 1977 to 2013 across all U.S. cities.

We leave three important questions for future work. First, it is important to say more precisely what this new technology is. The timing of these trends suggests that general purpose innovations in information and communication technologies have probably facilitated these fixed-cost based sectoral innovations. We also gave some hints in our narrative in the introduction 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 (1993)’s seminal work on the history of the industrial revolution in U.S. manufacturing, would be very useful.

Second, we provided a back of the envelope calculation of “missing growth” due to the expansion of top firms into new local markets. To be clear, 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 markups in each local market. We do not do this in this paper, but our hope is that such estimates will be forthcoming in the future.

Finally, the industrial revolution in services has implications on 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.
References


