The Best Cities for Firms*

Andra Ghent[†] *University of Utah* Eva Steiner[‡] Penn State University

May 2024

Abstract

We examine the spatial distribution of jobs within firms to understand how city characteristics affect employment. Since there is a mechanical relationship between a city's population and employment in firms that produce non-tradables, we propose a new way of identifying firms producing tradables. Our most robust findings are that physical city size and employment density drive employment even after controlling for firm sorting. We find different patterns in the drivers of employment in nontradable firms suggesting that what drives population may differ from what drives productivity. We find no evidence of agglomeration economies from industry diversity or specialization.

JEL: D22, H73, R12, R13.

Keywords: Local employment. Agglomeration economies. Firm location choice.

^{*}Paige Nelson provided excellent research assistance. We thank workshop participants at the Federal Reserve Bank of Dallas, Rutgers University, the University of Alabama, the AREUEA 2023 International Meeting, and the AREUEA Virtual Seminar for feedback on an earlier draft.

⁺SFEBB 8103, Department of Finance, David Eccles School of Business, University of Utah. Email: andra.ghent@eccles.utah.edu.

[‡]331 Business Building, Department of Risk Management, Penn State Smeal College of Business. Email: ems6770@psu.edu.

1 Introduction

Agglomeration economies in production are one of the key rationales for place-based policies (Bartik, 1991). Such place-based policies are costly and their effectiveness is sometimes questionable.¹ Furthermore, it remains unclear whether agglomeration economies operate primarily within industry, in which case policymakers may want to devote significant resources to incentivizing industry clusters, or whether they operate primarily across industries.

In this paper, we explore the role of cities in firm growth by exploiting establishmentlevel data on employment within the same firm over time. To overcome the endogeneity of firms' initial location choices, we focus on large, multi-unit firms that have established operations with employment in multiple cities. We construct an index of the non-tradability of a firm's output based on the spatial dispersion of its production. We then separate firms into those producing more or less tradable output. Conducting our analyses on firms producing more tradable output allows us to mitigate the influence of demand factors from within a city on our results. We exclude small cities from our estimations such that variation in a given firm's local employment does not materially affect aggregate city characteristics. We then identify the determinants of city-level productivity using within-firm heterogeneity in employment across cities over time.

Using data from a large sample of U.S. firms over the 1997–2021 period, we find that all firms have more employment in denser and larger cities. This result corroborates existing findings in the urban economics literature that have not benefitted from the identification enabled by the availability of establishment-level employment data, such as Ciccone and Hall (1996), Rosenthal and Strange (2008), Combes et al. (2010), and Davis et al. (2014).²

¹See discussions of place-based policies and the costs of cities' efforts to attract firms in, e.g., Bartik (1991), Austin et al. (2018), Neumark and Simpson (2015), Bartik (2020), and Slattery and Zidar (2020).

²See Duranton and Puga (2020) for a review of the literature on the benefits and costs of density.

We document different employment patterns in tradable versus non-tradable firms. Tradable firms have more employment in less diverse and less competitive cities indicating that competition keeps tradable firms from fully internalizing the benefits of innovations (cf. Marshall-Arrow-Romer economies). By contrast, non-tradable firms have more employment in more diverse cities, consistent with the rationale proposed in Jacobs (1969). The share of the population with a college degree drives employment only in tradable industries and only in high education industries. The divergent patterns we document in tradable versus non-tradable firms highlight that studies inferring the effects of city characteristics using all firms may incorrectly conclude that characteristics that drive population growth drive productivity growth.

We find no evidence that industry specialization or industry diversity increases firm productivity. The lack of evidence for productivity spillovers across industries suggests little harm to productivity from firms in different industries setting hybrid WFH policies independently of one another since workers in different industries need not be on-site on the same days. Conversely, the absence of benefits from industry specialization suggests limited benefits of local policies incentivizing industry clusters. Rather than incentivizing industrial diversity or specialization, our results indicate that policymakers should incentivize density and city size to increase productivity. Importantly, we study only agglomeration economies that operate at the level of an entire city.³

In addition to the literature on how cities affect the productivity of firms, our findings contribute to an emerging literature on the spatial organization of large firms. Using data on Danish firms, Acosta and Lyngemark (2021) document an increase in the spatial dispersion of firms in recent decades. Hsieh and Rossi-Hansberg (2023) also document an increase in the spatial dispersion of large U.S. service-producing firms, a finding

³Rosenthal and Strange (2020) and Baum-Snow et al. (2024) summarize the literature on agglomeration economies that operate at a much smaller spatial scale than the city. Our results are consistent with the lack of statistically significant industry TFP spillovers in Baum-Snow et al. (2024) who examine the role of firm peers in agglomeration economies at a very small spatial scale.

they attribute to technological improvement in the IT sector and improved management practices. Aarland et al. (2007) study the tendency of firms to have a location for a central administrative office (CAO) that is separate from the location of the rest of firm production. We take the significant spatial dispersion of U.S. firms as given and use it to study how city characteristics affect where firms choose to expand employment.

The rest of the paper proceeds as follows. In the next section, we outline a framework relating an individual firm's employment decisions to a city's productivity and factor prices. In Section 3, we describe our data and selection of firms. We present our results in Section 4. Section 5 offers concluding remarks.

2 Identifying the Drivers of City-Level Employment

The literature on agglomeration economies often focuses on industry-level employment across cities (see, e.g., Glaeser et al., 1992; Henderson et al., 1995). Using industry-level employment as the outcome variable of interest creates an identification challenge. Because local industry employment is a linear combination of two common explanatory variables, local specialization and density, the effects of those variables cannot be independently identified (Combes and Gobillon, 2015).

To address this identification challenge we focus on local firm-level employment. Our choice of dependent variable further alleviates two related endogeneity concerns. First, the employment of an individual firm is unlikely to have a significant impact on city characteristics such as density, reducing the potential for reverse causality. Second, the granularity of our firm-level data allows us to include firm-by-year fixed effects and firmby-CBSA fixed effects. The former set of fixed effects captures unobserved, time-varying firm characteristics that drive firms' employment dynamics over time. In other words, the inclusion of firm-by-year fixed effects allows us to rule out that it is time-varying firm characteristics, rather than city characteristics, that drive our results. We thus identify the effects of interest from cross-city variation in employment within firms, rather than between firms. The latter set of fixed effects captures any potential predilection among firms for certain locations unrelated to those cities' productivity.

Importantly, we only focus on the locations in which firms have an established presence. This approach allows us to account for the possibility that the most productive firms sort into the most productive locations when they first establish their operations. Theoretical literature suggests that the most productive firms indeed sort into the most productive cities (Oberfield et al., 2023) such that the correlation between density and productivity identified in earlier work may be due to firm sorting rather than causal.

In the remainder of this section, we derive the equation that we estimate to identify the drivers of city-level employment. To make ideas concrete, we outline a simple model to illustrate how city-level characteristics influence local employment. Then, we discuss the set of city characteristics whose impact on firms' local employment we assess in the data.

2.1 Estimating Equation

Firms are infinitely lived and operate in multiple cities. Each city potentially has a different productivity. Labor is always city-specific while some types of capital are firm-wide. Firms produce a good or service that can be consumed in any city such that our model describes firms producing tradable output; see Oberfield et al. (2023) and Hsieh and Rossi-Hansberg (2023) for a discussion of the decisions of large firms producing output that must be consumed near its location of production. Specifically, firm *i* located in city *c* produces output $Y_{i,c,t}$ using the technology

$$Y_{i,c,t} = \begin{cases} A_{c,t} \left[\min \left\{ L_{i,c,t}, \frac{K_{i,c,t}}{\theta} \right\} \right]^{\alpha} f_{i,t}^{1-\alpha} & \text{if } m_{i,c,t-j} \ge \overline{m} \\ 0 & \text{if } m_{i,c,t-j} < \overline{m} \end{cases}$$

where $A_{c,t}$ is city-level productivity, $f_{i,t}$ are firm-level productivity characteristics including capital not specific to city c, $L_{i,c,t}$ is the labor the firm uses in city c at time t, and $K_{i,c,t}$ is the real estate the firm uses in city c at time t. The production function with city-specific inputs is Leontief to simplify our analysis. Note, however, that the firm's overall production function allows for substitution between capital and labor because $f_{i,t}$ and city-specific inputs are substitutable.

Our specification of $A_{c,t}$, a common component of productivity for all firms operating in city c, is similar to the state-specific productivity assumed by Fajgelbaum et al. (2018) and the local productivity posited by Glaeser et al. (1992). Baum-Snow et al. (2024) also posit a firm-specific component of TFP and a location-specific component although their identification strategy exploits finer geographic data on the location of Canadian firms.

We do not model why a firm operates in some cities and not others. Instead, we conceive of an unobservable historic city-firm characteristic, $m_{i,c,t-j}$, that led the firm to establish a presence in city *c* at time t - j in the past. We then focus on the evolution of labor input in that firm-city pair over time. While a city's productivity may affect the firm's initial decision to establish operations in that location, there may also be historical reasons unrelated to a city's productivity that cause firms to have establishments in some cities and not others.⁴

The challenge for estimating the drivers of city-level productivity $A_{c,t}$ is that we do not observe $Y_{i,c,t}$. We can, however, observe employment in city *c*. This differs from many

⁴See, e.g., Jansen and Winegar (2022) for how amenities influence where entrepreneurs locate.

applications in the industrial organization literature wherein the challenge is that the researcher observes output but not *inputs* (e.g., Orr, 2022). We can observe labor input $L_{i,c,t}$ at the firm-city level, which will allow us to recover the drivers of $A_{c,t}$.

We limit our analyses to tradable firms such that there is no variation across cities in the price of the firm's output. We thus normalize the output price to 1. We also assume the firm is a price-taker in input markets.

Conditional on $m_{i,c,t-j} \ge \overline{m}$, $f_{i,t}$, and quadratic city-specific labor adjustment costs, the Leontief assumption on the location-specific part of the production function implies that the firm's profits each period are

$$\Pi_{i,c,t} = A_{c,t} L_{i,c,t}^{\alpha} f_{i,t}^{1-\alpha} - (w_{c,t} + \theta r_{c,t}) L_{i,c,t} - \frac{\phi}{2} (L_{i,c,t} - L_{i,c,t-1})^2.$$

The firm's problem is thus to choose a sequence of labor to maximize profits, i.e.,

$$\max_{\{L_{i,c,t+j}\}} E_t \sum_{j=0}^{\infty} \beta^j \Pi_{i,c,t+j}.$$

The Bellman equation is

$$V(L_{i,c,t-1}) = \max_{\{L_{i,c,t}\}} \left[\Pi_t + \beta E_t V(L_{i,c,t}) \right].$$

The first order condition is

$$\alpha A_{c,t} L_{i,c,t}^{\alpha-1} f_{i,t}^{1-\alpha} - (w_{c,t} + \theta r_{c,t}) - \phi (L_{i,c,t} - L_{i,c,t-1}) + \beta \phi E_t L_{i,c,t+1} - \beta \phi L_{i,c,t} = 0.$$

We can thus write

$$L_{i,c,t}^* = g(A_{c,t}, f_{i,t}, w_{c,t}, r_{c,t}, L_{i,c,t-1}, E_t L_{i,c,t+1}).$$
(1)

City-level productivity, $A_{c,t}$, is in turn governed by observable productivity characteristics, $X_{c,t}^A$, and an unobservable component according to

$$A_{c,t} = X_{c,t}^A e^{\epsilon_{c,t}}.$$

In equilibrium, factor prices for labor and real estate capital depend on both $A_{c,t}$ and non-productivity related city characteristics. For example, workers may be willing to accept lower wages in cities that offer better amenities (see, e.g., Roback, 1982). Similarly, rents will be higher in cities where it is harder to build more real estate and may be higher in cities that offer workers higher amenities. To capture those relationships, we specify

$$w_{c,t} = (A_{c,t})^{\lambda_A} (X_{c,t}^w)^{\lambda_w}$$
(2)

$$r_{c,t} = (A_{c,t})^{\kappa_A} (X_{c,t}^r)^{\kappa_r}$$
(3)

where $X_{c,t}^w$ and $X_{c,t}^r$ are city characteristics that do not directly affect a city's total factor productivity but may do so indirectly by influencing the prices of factor inputs.⁵

While we do not observe the firm's expected future labor input in city *c*, we assume that the firm forecasts its future expected hiring from its own firm-level productivity and the past realizations of city-level variables that govern city productivity, wages, and rents. Because many city characteristics are slow-moving, which introduces a high degree of collinearity between their current and lagged values, we assume

$$E_{t}L_{i,c,t+1} = h(X_{c,t}^{A}, X_{c,t}^{r}, X_{c,t}^{w}, \Delta X_{c,t}^{A}, \Delta X_{c,t}^{r}, \Delta X_{c,t}^{w}, ..., \Delta X_{c,t-j}^{A}, \Delta X_{c,t-j}^{r}, \Delta X_{c,t-j}^{w})$$
(4)

for j < J.

⁵In the language of Roback (1982), $X_{c,t}^w$ and $X_{c,t}^r$ are city characteristics that don't affect the cost function, i.e., $C_s = 0$.

To recover the full effect of $X_{c,t}^A$ on employment, we plug Equations (2)–(4) into Equation (1). We then estimate unbalanced panel regressions of the form

$$log(L_{i,c,t}) = \gamma_A log(X_{c,t}^A) + \gamma_k log(X_{c,t}^r) + \gamma_w log(X_{c,t}^w) + \gamma_{i,t} + \gamma_{i,c}$$
(5)
+ $\gamma_L log(L_{i,c,t-1}) + \delta_{A,0} \Delta log X_{c,t}^A + \delta_{k,0} \Delta log X_{c,t}^r + \delta_{w,0} \Delta log X_{c,t}^w + \dots + \delta_{A,j} \Delta log X_{c,t-j}^A + \delta_{k,j} \Delta log X_{c,t-j}^r + \delta_{w,j} \Delta log X_{c,t-j}^w + u_{i,c,t}.$

where $\gamma_{i,c}$ are firm-by-CBSA fixed effects to account for unobserved, time-invariant, idiosyncratic preferences among firms for certain locations, such as an entrepreneur's connection to their home town, that the model does not explicitly capture. $u_{i,c,t}$ is the residual.

2.2 City Characteristics

We now outline the city characteristics that we consider in the estimation of Equation (5). Those include direct drivers of city productivity $(X_{c,t}^A)$, as well as drivers of rents $(X_{c,t}^r)$ and wages $(X_{c,t}^w)$.

2.2.1 City Productivity Characteristics

The set of productivity-related city characteristics $(X_{c,t}^A)$ that we examine is informed by the existing literature on the empirics of agglomeration economies. Combes and Gobillon (2015) provide an extensive survey of the local determinants of agglomeration effects.

The first determinant that we consider is the size of the local economy. Following Ciccone and Hall (1996) and Rosenthal and Strange (2008), we measure the size of the local economy as employment density, defined as total employment in a city, scaled by the total land area of that city. We focus on employment density, instead of population density, because employment better captures the magnitude of local economic activity (Combes

and Gobillon, 2015). We use employment density instead of raw local employment to reduce the impacts on our estimations of potential mis-measurement in the scale of the local economy (Briant et al., 2010). We consider the total land area of cities alongside their employment density to distinguish between productivity gains stemming from higher employment (density) and a larger spatial scale of the city (Combes and Gobillon, 2015).

Next, we consider two variables that capture the industrial composition of the local economy, namely, specialization and diversity. The industrial specialization of a city can give rise to localization economies that are based on the accumulation of knowledge from communications between local firms in the same industry (Porter, 1990). Such effects have been termed Marshall-Arrow-Romer (MAR) economies, reflecting their intellectual origins. We define specialization at the city-level following Duranton and Puga (2000) as

$$Specialization_{c,t} = \max\left(\frac{Employment_{c,s,t}}{Employment_{s,t}}\right)$$
(6)

where $Employment_{c,s,t}$ is the share of industry *s* in total employment of city *c* at time *t*, and $Employment_{s,t}$ is the share of industry *s* in total employment nationwide at time *t*. We note that local employment density and specialization may both have a positive impact on city-level productivity as they reflect, respectively, urbanization economies and localization economies (Combes, 2000).

The industrial diversity of a city can give rise to urbanization economies that are driven by the sharing of knowledge and innovations between firms from different local industries (Jacobs, 1969). To capture those effects, we define diversity at the city-level similar to Duranton and Puga (2000) as

$$Diversity_{c,t} = \frac{1}{\sum_{s} (Employment_{c,s,t} - Employment_{s,t})^2}$$
(7)

where $Employment_{c,s,t}$ and $Employment_{s,t}$ are defined as in Equation (6). Combes et al. (2004) argue that measures of diversity which sum over local industries *s* are highly sensitive to the number of locally active industries. To identify the effect of diversity on city-level productivity more accurately, they propose considering the number of locally active industries alongside the diversity measure.

The distribution of local economic activity across the firms in a city may also generate localization economies. Jacobs (1969) and Porter (1990) posit that local competition supports city productivity as it accelerates the adoption of new technology. The MAR theory implies that a local monopoly is more beneficial for city productivity since it allows innovators to internalize the benefits of developing new technologies, fostering innovation and growth. Similar to Glaeser et al. (1992), we define local competition at the city-level as

$$Competition_{c,t} = \frac{Establishments \ per \ Employee_{c,t}}{Establishments \ per \ Employee_t}$$
(8)

where *Establishments per Employee*_{*c*,*t*} is the total number of establishments in city *c* at time *t*, scaled by the corresponding total number of employees. The denominator is *Establishments per Employee*_{*t*}, the total number of establishments nationwide at time *t*, scaled by the corresponding total number of employees.

The educational composition of the local workforce can lead to human capital externalities that facilitate city-level productivity. We measure human capital using the education level of the local workforce (see, e.g., Combes et al., 2008; Moretti, 2004). Specifically, we define education at the city-level as

$$Education_{c,t} = \frac{College-Educated Population_{c,t}}{Total Population Over 25_{c,t}}$$
(9)

where *College-Educated Population*_{c,t} is the total number of persons in city c at time t who hold a bachelor's degree or higher level of educational attainment and *Total Population Over*

 $25_{c,t}$ is the corresponding total population over 25 years of age. We note that *Education*_{c,t} captures the composite effect of human capital externalities and the limited substitutability of more and less educated workers in a city (see, e.g., Combes and Gobillon, 2015; Rossi-Hansberg, Esteban and Sarte, Pierre-Daniel and Schwartzman, Felipe, 2019).

2.2.2 Wage Determinants

The non-productive characteristics that make some locations more attractive to workers $(X_{c,t}^w)$, thus possibly influencing wages, are largely time-invariant. For example, consumption amenities such as weather, proximity to the ocean, and proximity to mountains do not vary over time. We include several natural amenity measures in our regressions to capture the geographic desirability of a location. Following prior literature, we also include in our analyses a point-in-time measure of the number of days with pleasant weather in each city (Rappaport, 2007).

2.2.3 Rent Determinants

As is the case for wages, many of the non-productive characteristics that create higher or lower rents in a city $(X_{c,t}^r)$ are fixed over time. For example, the presence of sloped terrain or bodies of water is fixed. However, many cities have increased man-made land use restrictions over time (Gyourko et al., 2021; Baum-Snow and Han, 2023). Notably, cities with higher productivity may enact more land use restrictions (Davidoff, 2016). To account for those factors, we include in our analyses two point-in-time measures of the restrictiveness of local land use regulation.

3 Data

To estimate Equation (5), we need to construct a data set with a firm-city-year panel of employment and match it to a city-year panel of urban characteristics. In this section, we outline the construction of that data set.

3.1 Employment Data

Our employment data set is based on information from Data Axle. Data Axle is a leading business data and analytics firm that provides annual establishment-level information based on the Infogroup Business Data historical files. The Data Axle database covers all public and private firms in the U.S. since 1997. From that database, we obtain annual data sets for the 1997–2021 period. Those annual data sets contain information on the unique identity of each establishment, its exact location, parent firm, and industry (NAICS code), and the total number of employees per establishment. We define an establishment's location by the Core-Based Statistical Area (CBSA) in which it is situated.

We supplement our Data Axle data with the education attainment required for each occupation with data from the Bureau of Labor Statistics (BLS). We then map occupations to 2-digit NAICS code using the crosswalk provided by the BLS through the Occupational Employment and Wage Statistics program (OEWS) to obtain an industry-level summary of the share of jobs in each industry requiring a bachelor's degree or higher.

From the initial Data Axle sample, we focus on firms that have at least 500 employees in at least one year over the 1997–2021 period. We drop establishments with missing data on the parent firm to which they belong as well as those with missing CBSA codes and those situated in rural locations. We also require firms to have a presence in at least two CBSAs over the 1997–2021 period. We drop firm-years and firm-CBSA-years with zero employment. Lastly, we require firms to have at least three consecutive years of data in a CBSA. Our final sample thus includes 14,541 unique firms.

Figure 1 presents a breakdown of the final sample by industry based on the 2-digit NAICS classification of the sample firms' headquarters. The figure shows that the final sample includes firms from 20 different industries. The figure also indicates that manufacturing, professional services, wholesale trade, finance/insurance, and retail trade represent the top-five largest industries by share of observations.

Figure 1. Breakdown of Industries

This figure depicts the breakdown of the firm-CBSA-year observations in our sample by industry (2-digit NAICS codes). The data used to produce this figure are from Data Axle.



The sample firms are headquartered across 571 unique CBSAs. Figure 2 shows the geographical locations of the sample firms' headquarters by CBSA. The map indicates that the sample firms are headquartered across a wide range of locations in the U.S.

Figure 2. Headquarter Locations of Sample Firms by CBSA

This figure depicts the geographical locations of the sample firms' headquarter locations by CBSA. Darker shading indicates a larger number of sample firm headquarters housed in a CBSA. For readability, the map only depicts the CBSAs located in the continental U.S. The data used to produce this figure are from Data Axle.



3.2 Data on City Characteristics

We match the firm-CBSA-year panel of firms' employment data with information on key drivers of agglomeration economies at the CBSA-year level. The drivers of agglomeration economies that we include in our analyses are outlined in Section 2.2. Here, we provide details on the data used to construct the corresponding variables.

We obtain annual data on the total area of each CBSA in square miles (*Total Area*) from the U.S. Census Bureau's County TIGER/Line Shapefiles. We calculate the variable *Density* by dividing each CBSA's annual total employment, obtained from the U.S. Census Bureau's County Business Patterns Data, by its *Total Area*.

We construct *Specialization*, *Diversity*, and *Competition* directly from the DataAxle employment data. We take the data on educational attainment and the population over 25 years of age that are required to construct the variable *Education* from the U.S. Census Bureau's American Community Survey. We obtain the data for the variable *Regulation* from the Wharton Residential Land Use Regulatory Index (WRLURI).⁶ The WRLURI data are available for 2006 and 2018. We linearly interpolate the data for the missing years and extrapolate for years prior to 2006 and after 2018. The data for the variable *Weather*, proxied by the number of days per year with pleasant weather, are from the National Weather Service and the Federal Aviation Association (NWS/FAA).⁷ The CBSA-level data on the number of days per with pleasant weather represent the average of that number over the 1998–2018 period. We use those CBSA-level average values across all sample

⁶The Wharton Residential Land Use Regulatory Index (WRLURI) data are available from Joseph Gyourko's website at Wharton, see here.

⁷We obtained the number of days per year with pleasant weather from Brian Brettschneider at the University of Alaska Fairbanks, see here. The data are from the NWS/FAA automated stations and were downloaded from the Iowa State University Mesonet archive site.

years. The data for the variable % *Water*, which captures the percentage share of a CBSA's area that is water, are from the Economic Research Service at the U.S. Department of Agriculture (USDA). The data on a CBSA's topography, which allow us to construct the variable *Mountainous*, are from the same source.

Table 1 presents descriptive statistics on selected characteristics of the top-10 largest CBSAs in the U.S. over the 1997–2021 period. The top-10 largest CBSAs are determined by their mean total employment over that period. The CBSAs are listed in rank order of their mean total employment (from largest to smallest). New York is the top-ranked CBSA by total mean employment, followed by Los Angeles and Chicago. Among the top-10 largest CBSAs, New York has the highest employment density, while Dallas covers the largest land area. The most specialized cities are Washington, D.C. (government and related services) and Houston (oil industry). Chicago and Atlanta have the most diverse industrial composition. All of the top-10 largest CBSAs house nearly 100 distinct industries. San Francisco and Los Angeles are the most competitive CBSAs, based on the local number of establishments per employee. Washington, D.C., San Francisco, and Boston have the most highly educated workforces. Boston, Philadelphia, and San Francisco exhibit the tightest land use regulations relative to the national average. Los Angeles and San Francisco register the highest numbers of days with pleasant weather per year. Coastal cities, namely Los Angeles, New York, Boston, and San Francisco, naturally have the largest percentage shares of water in their areas. Of the CBSAs in this top-10 ranking, only Los Angeles and San Francisco are classified as having a mountainous topography.

We construct the final data set for our main empirical analyses by matching the CBSA-year panel of city characteristics to the firm-CBSA-year panel of employment data by CBSA code and year.

Table 1. Top-10 Cities by Key Characteristics

3.3 Descriptive Statistics

Table 2 presents descriptive statistics on our final sample by firm-year. The statistics reported show that the average firm in the sample operates 174 establishments per year, across nearly 45 CBSAs. The average firm is active in the sample for approximately 18 years. During that time, annual firm-level employment averages approximately 5,500 workers. The annual mean growth rate in total employment (respectively, the total number of establishments) is 22% (12%). The average sample firm opens a new establishment in 66% of sample years and enters a new CBSA in 42% of sample years.

Table 2. Descriptive Statistics on Firm-Year Panel

This table presents descriptive statistics on the firms in the final sample, observed over the 1997–2021 period. # *Establishments* is the number of establishments a firm operates in a given sample year. # *CBSAs* is the number of CBSAs in which a firm operates active establishments in a given sample year. # *Years* is the number of years during which we observe a firm in the sample. *Employment* is the total number of employees that a firm has in a given sample year. *Employment Growth* is the annual growth rate in the number of employees that a firm experiences during the sample period. *Establishment Growth* is the annual growth rate in the number of active establishments that a firm experiences during the sample period. *New Establishment* is an indicator that takes the value of one if a firm opens a new establishment in a given sample year. *New CBSA* is an indicator that takes the value of one if a firm enters a new CBSA, in which it does not previously operate any active establishments, in a given sample year.

	Ν	Mean	Median	Std. Dev.	Min.	Max.
# Establishments	182,216	178.06	29.00	911.34	2.00	66,395.00
# CBSAs	182,216	45.68	15.00	97.39	2.00	933.00
# Years	182,216	18.36	20.00	6.85	3.00	25.00
Employment	182,216	5,572.74	1,217.00	25,077.28	1.00	1,617,586.00
Employment Growth	182,216	0.20	0.00	10.06	-1.00	2,744.71
Establishment Growth	182,216	0.12	0.00	2.41	-1.00	434.67
New Establishment	182,216	0.67	1.00	0.47	0.00	1.00
New CBSA	182,216	0.43	0.00	0.49	0.00	1.00

Table 3 presents descriptive statistics on our final sample by firm-CBSA-year. The statistics reported show that the average firm in the sample operates over four establishments per CBSA each year, and operates those establishments for 16 years on average.

During that time, annual firm-level employment averages approximately 132 workers per CBSA. On average, the CBSA with the most employees hosts approximately 21% of the firms' total workforce. By contrast, the CBSA with the least employees on average hosts less than 1% of the firms' total workforce. The annual mean growth rate in total CBSA-employment (respectively, the total number of establishments) for the sample firms is approximately 30% (6%).

Table 3. Descriptive Statistics on Firm-CBSA-Year Panel

This table presents descriptive statistics on the firm-CBSA-year observations in the final sample over the 1997–2021 period. # *Establishments* is the number of establishments a firm operates in a given CBSA and given sample year. # *Years* is the number of years during which we observe a firm in a given CBSA in the sample. *Employment* is the total number of employees that a firm has in a given CBSA and given sample year. *Max. Employment Share* is the share of total employment that a firm has in the CBSA with the most employees in a given year. *Min. Employment Share* is the share of total employment that a firm has in the CBSA with the least employees in a given year. *Employment Growth* is the annual growth rate in the number of employees that a firm experiences in a given CBSA during the sample period. *Establishment Growth* is the annual growth rate in the number of employees that a firm experiences in a given CBSA during the sample period. *Establishment Growth* is the annual growth rate in the number of employees that a firm experiences in a given CBSA during the sample period.

	Ν	Mean	Median	Std. Dev.	Min.	Max.
# Establishments	2,740,139	7.29	2.00	24.36	1.00	2,843.00
# Years	2,740,139	16.17	16.00	7.13	3.00	25.00
Employment	2,740,139	238.96	37.00	1,098.47	1.00	174,251.00
Max. Employment Share	2,740,139	0.26	0.18	0.21	0.01	1.00
Min. Employment Share	2,740,139	0.00	0.00	0.02	0.00	0.50
Employment Growth	2,740,139	0.44	0.00	21.43	-1.00	19,999.00
Establishment Growth	2,740,139	0.08	0.00	1.19	-1.00	465.00

Table 4 presents descriptive statistics on the firm-CBSA-year observations of employment in the final sample matched to CBSA-year characteristics. Mean firm-CBSA-year employment is approximately 180 employees. The average employment density across the CBSAs in the matched sample is 214 employees per square mile. The average total land area is approximately 4,000 square miles. The mean level of specialization (diversity) across the CBSAs in the matched sample is approximately 1,032 (28). We note that specialization and diversity are not opposites—a city with a specialization in one main industry can, at the same time, have a broad base of other industries (Duranton and Puga, 2000). The average CBSA in the sample houses approximately 95 active industries and has a level of competition consistent with the U.S. as a whole, with a mean value of 0.99. On average, 29% of the CBSA-level population hold a bachelor's degree or higher. The mean value of land use regulation is 0.03, slightly higher than the overall U.S. average (the WRLURI data are standardized across the CBSAs in the U.S. to have a mean of zero, such that a positive value indicates an above-average level of land use regulation). The average number of days per year with pleasant weather in the CBSAs represented in our matched sample is 70. The percentage share of a CBSA that is water averages 7%. Approximately 25% of the observations in the matched sample are from CBSAs that are mountainous.⁸

3.4 Non-Tradability Index

The location choice for firms that produce output not easily consumed outside of the location in which it is produced depends on local consumption rather than how the city affects productive capacity. Further, there is a direct relationship between employment growth in a city and the growth of non-tradable output that does not depend on the agglomeration economies that we are interested in. For example, we would trivially find that McDonald's has operations in nearly every American city. We would also find that employment in the construction industry is growing in successful cities for reasons unrelated to agglomeration economies.

To focus on the drivers of employment growth for firms that have a choice of where to produce because of productivity factors, rather than simply overall population growth

⁸Appendix Table A.1 presents pairwise correlation coefficients for the variables in the matched sample. The statistics reported indicate no serious concerns about multicollinearity between those variables.

Table 4. Descriptive Statistics on Matched Sample

This table presents descriptive statistics on the firm-CBSA-year observations of employment in the sample matched to CBSA-year characteristics over the 1997–2021 period. Employment is the total number of employees in a given firm-CBSA-year. The CBSA-year characteristics are defined as follows. Density is total employment in a CBSA, scaled by that CBSA's land area in square miles. Total Area is the total land area of a CBSA in square miles. Specialization is the maximum of the employment share of a given industry (defined by 3-digit NAICS code) in a city, scaled by the corresponding share of that industry's employment nationwide. Diversity is the inverse of the sum across all industries active in a CBSA of the squared differences between a given industry's employment share in a given CBSA-year and the corresponding employment share of that industry nationwide. # Industries is the number of active industries (those with non-zero employment) in a given CBSA-year. Competition is the number of establishments per employee in a given CBSA-year, scaled by the corresponding number of establishments per employee nationwide. Education is the share of the total population over 25 years of age with a bachelor's degree or higher level of educational attainment in a given CBSA-year. *Regulation* is the CBSA's assigned value of the Wharton Residential Land Use Regulatory Index (WRLURI). Weather is the number of days with pleasant weather in a CBSA. % Water is the percentage share of a CBSA's total area that is water. Mountainous is an indicator that takes the value of one if the local topography of a given CBSA is mountainous.

	Ν	Mean	Median	Std. Dev.	Min.	Max.
Employment	2,806,328	234.29	35.00	1,087.18	1.00	174,251.00
Density	2,806,328	283.39	198.74	252.03	16.87	1,301.11
Total Area	2,806,328	5,109.08	4,391.90	3,749.20	604.31	27,277.36
Specialization	2,806,328	764.34	493.22	801.16	166.61	24,719.68
Diversity	2,806,328	28.52	28.74	3.16	16.59	37.80
# Industries	2,806,328	96.81	97.00	1.45	89.00	99.00
Competition	2,806,328	0.98	0.97	0.11	0.65	1.50
Education	2,806,328	0.31	0.31	0.07	0.11	0.54
Regulation	2,806,328	0.08	0.10	0.63	-2.63	2.23
Weather	2,806,328	72.87	61.09	30.22	40.78	173.79

that creates demand for the product, we create a measure of the non-tradability of a firm's output.

Larger firms will tend to have more employment in more cities. Further, one of the reasons businesses may have a presence in multiple cities is because they produce output in many industries and concentrate their production of each industry output in one location. To generate a non-tradability index, we therefore regress the total number of cities in which a firm has employment in a given year on deciles of total firm employment and quartiles of the total number of 2-digit NAICS industries in which a firm has employment.

We take the residual from that regression as our non-tradability index. Intuitively, this measure would assign a score close to zero for firms whose geographical dispersion is almost entirely explained by their size and by the number of industries in which they operate. Firms that produce in fewer locations than their size and industry coverage suggest must be producing tradable output since their output is likely consumed across many cities, but only produced in a few. Those firms will have negative values in our non-tradability index. In contrast, firms that produce in a larger number of cities than their size and industry coverage suggest must be producing suggest must be produced in a few. Those firms will have negative values in our non-tradability index. In contrast, firms that produce in a larger number of cities than their size and industry coverage suggest must be producing output consumed locally. Those firms will have positive values in our non-tradability index.

We can aggregate our non-tradability index from the firm-level to the industry-level by computing the average values of the non-tradability index by 4-digit NAICS codes. The resulting measure allows us to assess how non-tradable an industry's output is. Table 5 compares our non-tradability index to the industry classification by Mian and Sufi (2014). Reassuringly, the average and median values of our non-tradability index are highest for industries that Mian and Sufi (2014) classify as non-tradable and lowest for industries that Mian and Sufi (2014) classify as tradable.

Table 5. Descriptive Statistics on Non-Tradable Index by Industry Category

This table presents descriptive statistics on the industry-level Non-Tradability Index by industry categories. The firm-year-level Non-Tradability Index is aggregated to the industry-year level by taking industry-means in each year of the 1997–2021 period. Descriptive statistics are tabulated by the industry categories defined in Mian and Sufi (2014).

	P25	Median	P75	Mean	St. Dev.	Min.	Max.
Other	-12.96	-3.90	9.49	2.00	37.84	-159.83	577.17
Tradable	-22.08	-14.17	-8.02	-15.93	13.73	-153.83	54.40
Non-Tradable	-0.19	20.21	43.53	24.19	34.72	-65.24	183.88
Construction	-10.20	-3.67	4.70	-2.56	14.97	-121.83	109.41
Total	-15.66	-6.71	5.29	-1.55	32.88	-159.83	577.17

Figure 3 depicts the distribution of our non-tradability index. Panel A presents the distribution of that index on the industry-level, broken down by the industry categories defined in Mian and Sufi (2014). The figure shows that the non-tradability index takes a wide range of values, even within the tradable industry category as defined in Mian and Sufi (2014). Those patterns suggest that industries classified as tradable can in fact be populated by firms that produce in many cities. This finding implies that classifications of tradability based solely on where an industry produces as a whole can be misleading. One possible reason for the discrepancy is that the competitive structure of certain industries lends itself to more or less concentration. Our within-firm measure instead directly measures how dispersed firms' production structure is.

Panel B of Figure 3 presents the distribution of the non-tradability index on the firm-level, broken down by industry for the top-five industries in our final sample (see Figure 1). Industries are defined by the 2-digit NAICS code of the firms' headquarters. The figure shows that, even within a given industry, the non-tradability index varies significantly. This finding may be related to the fact that firms operate establishments across multiple industries that can vary by the degree to which their output is tradable.

Figure 3. Distribution of Non-Tradability Index

This figure depicts the distribution of the Non-Tradability Index. Panel A presents the distribution of the Non-Tradability Index on the industry-level, showing the mean Non-Tradability Index values for each industry, broken down by the industry categories defined in Mian and Sufi (2014). Panel B presents the distribution of the Non-Tradability Index on the firm-level, broken down by the firms' industry for the top-five industries in our final sample (see Figure 1). Industries are defined by 2-digit NAICS codes. The values of the Non-Tradability Index are trimmed at the 1st and 99th percentiles. The data used to produce this figure are from Data Axle and Mian and Sufi (2014).



(A) Industry-Level

(B) Firm-Level

4 What Determines Where Firms Employ People?

We now turn to the analyses of the city characteristics that may spur firms' local employment. We estimate Equation (5) on the firms in the final sample over the 1997–2021 period that have an established presence in a given CBSA. We define an established presence as having at least two consecutive years of data prior to a given observation.

4.1 What Types of Agglomeration Economies Matter?

Table 6 presents the results from estimating Equation (5). Columns 1 through 3 present the regression results for all firms. Columns 4 through 6 (columns 7 through 9, respectively) present the corresponding results for tradable firms (non-tradable firms). The results presented in column 1 (columns 4 and 7, respectively) refer to a baseline estimation with lagged employment and the first differences of the city characteristics. In column 2 (columns 5 and 8, respectively), we add the one-period lag of the differenced city characteristics in which we account for up to two lags of the differenced city characteristics.

The estimated coefficients shown in Table 6 are similar across the specifications reported in columns 1 through 3 (4 through 6 and 7 through 9, respectively), suggesting that the regression results are insensitive to the lag length included in the estimations.

Table 6. Determinants of Firm-CBSA-Year Employment

clustered standard errors (by CBSA), are shown in parentheses. Tradable firms are those with a Non-Tradability Index below the 25th percentile. Non-(Emp.), defined as the total number of employees in a given firm-CBSA-year. See Table 4 for independent variable definitions. All variables, except Regulation, are transformed to their natural logarithms. The lag of the dependent variable and (lagged) differences of city characteristics (except Regulation) are included as indicated. Firm-by-year fixed effects and firm-by-CBSA fixed effects are included as indicated. t-statistics, based on tradable firms are those with a Non-Tradability index above the 75th percentile. Significance is indicated as follows: *** p<0.001, ** p<0.05, * p<0.1This table presents output from Equation (5), estimated over the 1997–2021 period. The dependent variable is the natural logarithm of Employment

		All Firms		Tr	adable Firn	SU	Non-	Tradable F	irms
	Emp. (1)	Emp. (2)	Emp. (3)	Emp. (4)	Emp. (5)	Emp. (6)	Emp. (7)	Emp. (8)	Emp. (9)
Density	0.096***	0.103^{***}	0.106^{***}	0.095**	0.105^{**}	0.103^{**}	0.098***	0.105^{***}	0.111^{***}
,	(4.96)	(4.78)	(4.94)	(2.58)	(2.58)	(2.48)	(4.74)	(5.02)	(5.16)
Total Area	0.106^{***}	0.113^{***}	0.113^{***}	0.108^{***}	0.119^{***}	0.110^{***}	0.101^{***}	0.108^{***}	0.108^{***}
	(5.83)	(5.67)	(5.60)	(3.12)	(3.10)	(2.74)	(4.99)	(5.33)	(5.22)
Specialization	-0.004*	-0.004**	-0.005*	-0.001	-0.004	-0.004	-0.003	-0.005*	-0.005
	(-1.94)	(-2.07)	(-1.92)	(-0.27)	(-0.95)	(-0.89)	(-1.35)	(-1.73)	(-1.47)
Diversity	-0.01	-0.013	-0.019	-0.01	-0.017	-0.028	0.016	0.019	0.017
	(-0.84)	(-1.03)	(-1.33)	(-0.38)	(09.0-)	(-0.84)	(1.24)	(1.32)	(1.03)
# Industries	0.043	-0.014	-0.041	-0.162	-0.300	-0.427	0.227^{**}	0.206^{*}	0.180
	(0.41)	(-0.11)	(-0.30)	(-0.66)	(-1.05)	(-1.35)	(2.35)	(1.71)	(1.46)
Competition	-0.021	-0.019	-0.014	-0.127***	-0.136***	-0.160***	0.011	0.008	0.020
	(-1.17)	(-0.92)	(09.0-)	(-3.57)	(-3.50)	(-3.34)	(0.46)	(0.31)	(0.69)
Education	0.026	0.045	0.051^{*}	0.115^{*}	0.162^{**}	0.171^{**}	-0.052	-0.045	-0.046
	(1.03)	(1.60)	(1.72)	(1.75)	(2.26)	(2.33)	(-1.32)	(-1.07)	(-1.04)
Regulation	0.002	0.002	0.001	0.004	0.004	0.004	0.003	0.003	0.002
1	(0.93)	(0.77)	(0.74)	(0.65)	(0.61)	(0.62)	(1.23)	(1.06)	(0.98)
Firm-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-CBSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
L.Employment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
D.City Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LD.City Characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
L2D.City Characteristics	No	No	Yes	No	No	Yes	No	No	Yes
Observations Adjusted R ²	2,059,247 0.94	2,059,247 0.94	2,059,247 0.94	503,459 0.91	503,459 0.91	503,459 0.91	509,014 0.97	509,014 0.97	509,014 0.97

4.1.1 Results for All Firms

The coefficient estimates reported in columns 1 through 3 of Table 6 suggest that higher levels of density in a city are associated with higher employment. This finding is consistent with prior studies assessing the effects of density on productivity in the U.S., such as Ciccone and Hall (1996), Rosenthal and Strange (2008), Combes et al. (2010), and Davis et al. (2014). The results reported across columns 1 through 3 also indicate that firms have significantly higher employment in cities with a larger total area. This result aligns with the estimates presented in Combes et al. (2012) who document, using French data, that firms are more productive in larger cities. In combination with our finding on density, these results suggest that cities experience productivity gains from growth in their density and, simultaneously, from growth in their spatial scale.

Prior work has focused on estimating the effects of local market size on industry-level employment growth in different countries, with mixed results. Combes (2000) documents a positive (negative) impact of market size on local industrial employment growth in manufacturing (service) industries in France. Viladecans-Marsal (2004) shows that, in Spain, market size is insignificant in explaining industrial employment growth across six different industries and has a non-linear effect in three others. We provide firm-level evidence consistent with prior works that identify a positive association between the size of the local economy and industry-level employment growth.

The estimates shown in columns 1 through 3 of Table 6 indicate that higher degrees of industrial specialization in cities have a slightly negative influence on firms' employment in those locations. Henderson et al. (1995) document that MAR economies are associated with higher industrial growth at the industry level in eight manufacturing industries. In contrast, Glaeser et al. (1992) show that industry-level employment grows more slowly in cities where that industry accounts for a large share of total local employment. Combes (2000) also present evidence that, in France, high industrial specialization is associated

with lower employment growth within that industry. Henderson (2003) and Combes et al. (2008) both find larger effects of specialization in service industries, where firms could experience greater technological spillover effects. Instead of looking at the industry-level, we are able to study employment within firms. Our granular firm-CBSA-level data suggest that cities may not benefit from greater specialization.

The coefficient estimates reported indicate that regulation is positively related to local employment (see columns 1 through 3). This finding highlights the importance of accounting for non-productivity related drivers of local factor prices in analyses of city-level determinants of productivity. We find little evidence that diversity, competition, or education have a significant influence on firms' employment choices across cities in any of the specifications reported across columns 1 through 3 of Table 6.

4.1.2 Results for Tradable vs. Non-Tradable Firms

For the purposes of the sub-sample analyses of (non-) tradable firms, we define tradable firms as those with a non-tradability index below the 25th percentile. We define non-tradable firms as those with a non-tradability index above the 75th percentile. The regression results reported for tradable and non-tradable firms show that city density and spatial scale are positively associated with employment for both firm types. Those results are consistent with the findings for all firms discussed above.

Beyond the effects of city density and spatial scale, the regression results indicate two striking differences in the determinants of firm-CBSA-year employment across tradable and non-tradable firms. First, a city's industrial diversity, as proxied by the number of locally active industries, is positively related to employment for non-tradable firms (see columns 7 through 9). For tradable firms, however, the number of locally active industries has a slightly negative or insignificant association with firm-CBSA-year employment (columns 4 through 6). Previous work suggests that the effect of diversity is not robust (Combes and Gobillon, 2015). Our results based on granular establishment-level employment data and newly developed tradability categories reveal evidence that is consistent with the intuition of Jacobs (1969) for non-tradable firms.

Second, the MAR theory implies that higher levels of local competition hinder growth, as a local monopoly allows firms to internalize the benefits of developing innovations. On the other hand, Jacobs (1969) and Porter (1990) posit that local competition is beneficial for growth as it spurs the adoption of new technologies. Our estimates show that competition is negatively associated with local employment for tradable firms (see columns 4 through 6 of Table 6). In contrast, the local employment levels of non-tradable firms appear to be insensitive to the effects of local competition (columns 7 through 9).

We find little evidence that specialization, education, or regulation are significant drivers of tradable firms' local employment. By contrast, we document a positive effect of regulation on the employment levels of non-tradable firms. However, the effect is relatively weak, with a coefficient estimate of 0.004 that is significant at the 10% level.

In sum, we document significant differences in the drivers of local employment choices between tradable and non-tradable firms. Those differences imply that the degree to which a firm's output is tradable can mask nuanced relationships between determinants of cities' productivity and local firm employment.

4.1.3 Results by Education of Parent Establishment

Given the weak effect of the share college in column (4)-(6) of Table 6, it is worth investigating the extent to which the importance of an educated work force varies by industry. In Table 7, we split firms into those operating in industries where a high or low share of jobs are in occupations that require a college degree according to the 2D NAICS code of the establishment of the parent. Indeed, city education only matters in high-education industries.

Table 7. Determinants of Firm-CBSA-Year Employment-By Education

t-statistics, based on clustered standard errors (by CBSA), are shown in parentheses. Significance is indicated as follows: *** p < 0.001, ** p < 0.05, * p < 0.1 Column (1) contains benchmark specification for all firms. Column (2) contains results for the subset of firms for which we can identify the 2D NAICs of the establishment. In column (3), we include only firms whose parent establishment is in a 2D NAICS code with an above median share of occupations that require a bachelor's degree or more. In column (4), we include only firms whose parent establishment is in a 2D NAICS code with a below median share of occupations with a bachelor's degree or more.

	Emp. (1)	Emp. (2)	Emp. (3)	Emp. (4)
Density	0.10**	0.096**	0.14**	0.059
	(0.04)	(0.05)	(0.06)	(0.07)
Total Area	0.11***	0.16***	0.24***	0.09
	(0.04)	(0.05)	(0.07)	(0.07)
Specialization	-0.0044	0.00077	-0.015	0.0073
	(0.01)	(0.01)	(0.01)	(0.01)
Diversity	-0.028	-0.026	-0.099	0.027
	(0.03)	(0.07)	(0.09)	(0.10)
# Industries	-0.43	0.037	0.59	-0.27
	(0.32)	(0.58)	(0.78)	(0.71)
Competition	-0.16***	-0.1	-0.064	-0.17
	(0.05)	(0.08)	(0.13)	(0.11)
Education	0.17**	0.09	0.34*	-0.17
	(0.07)	(0.13)	(0.19)	(0.18)
Regulation	0.0037	0.019**	0.018*	0.022
	(0.01)	(0.01)	(0.01)	(0.01)
Firm-Year FEs	Yes	Yes	Yes	Yes
Firm-CBSA FEs	Yes	Yes	Yes	Yes
Firm Types	Tradable	Tradable	Tradable	Tradable
Education of 2D NAICS	All	All	High Bach.	Low Bach.
Years	All	All	All	All
Observations	503,459	174,207	77,337	93,127
R^2	0.93	0.93	0.93	0.93

4.1.4 Effects of Time Trends, Sorting, and Idiosyncratic Preferences

In Table 8, we replicate the estimations from our preferred specification in Table 6 for all firms and for tradable firms. Recall that the benchmark specification from Table 6 includes lagged firm-CBSA-level employment along with up to two-period lagged differences in city characteristics. Table 8 presents alternative versions of this specification in which we include, in turn, year fixed effects, firm-by-year fixed effects, and CBSA-by-year fixed effects. This analysis allows us to assess the effects of accounting for time trends, firm sorting, and firms' idiosyncratic preferences on inferences about city productivity.

The results for all firms indicate that only accounting for time trends by including year fixed effects would suggest that diversity and education have a negative association with employment (column 1). Only accounting for firm sorting by including firm-by-year fixed effects still produces results suggesting that education has a negative association with employment (column 2). Lastly, only accounting for firms' idiosyncratic location preferences by including firm-by-CBSA fixed effects would suggest that specialization and competition have a negative association with employment while diversity and regulation would be found to have a positive association with firms' local employment (column 3).

The sub-sample analyses for tradable firms in Table 8 suggest similar conclusions. Notably, estimations without our saturated set of fixed effects would suggest that some city characteristics matter for employment when they are insignificant in our benchmark model. The highly significant coefficient estimates on specialization, diversity, and the number of active local industries in column 7 are an example for this issue. Analyses without our saturated set of fixed effects would also on occasion fail to detect significant associations between city characteristics and employment. For instance, the coefficient estimates on employment density and city size are insignificant in column 7 although the positive effects of those characteristics on employment are among the most robust findings in our benchmark analyses. The estimation results in Table 8 show that accounting for time trends, firm sorting, and firms' location preferences significantly alters inferences about the drivers of city productivity. Failing to account for those effects can confound estimation results.

In sum, our results reveal that firms' sorting and preferences across locations are significant drivers of local employment choices that are unrelated to city-level productivity. We also document stark differences in the drivers of local employment choices between tradable and non-tradable firms. Our results stand in contrast to those from prior work, which did not have the benefit of our identification strategy based on granular, establishment-level data to account for firms' sorting and idiosyncratic preferences.

Table 8. Determinants of Firm-CBSA-Year Employment-Effects of Time Trends, Sorting, and Preferences

This table replicates the results from Table 6. The lag of the dependent variable, the first differences of city characteristics (except Regulation, Weather, % Water, and Mountainous), as well as their one-period lagged and two-period lagged first differences, are included in all specifications. Year fixed effects, firm-by-pear fixed effects, are included as indicated. *t*-statistics, based on clustered standard errors (by CBSA), are shown in parentheses. Significance is indicated as follows: *** p<0.001, ** p<0.05, * p<0.1

		All F	irms			Tradabl	le Firms	
	Emp. (1)	Emp. (2)	Emp. (3)	Emp. (4)	Emp. (5)	Emp. (6)	Emp. (7)	Emp. (8)
Density	0.019***	0.046***	-0.006	0.106***	0.027***	0.040***	-0.077	0.103**
2	(20.95)	(32.14)	(-0.24)	(4.94)	(9.52)	(15.70)	(-1.34)	(2.48)
Total Area	0.020^{***}	0.049^{***}	-0.001	0.113^{***}	0.025^{***}	0.038***	-0.079	0.110^{***}
	(18.54)	(33.43)	(-0.05)	(5.60)	(9.18)	(15.91)	(-1.41)	(2.74)
Specialization	0.000	-0.001	-0.015***	-0.005*	0.002	0.001	-0.022**	-0.004
-	(0.42)	(-0.79)	(-3.03)	(-1.92)	(1.62)	(0.92)	(-2.30)	(-0.89)
Diversity	-0.008*	0.004	0.184^{***}	-0.019	-0.030**	0.002	0.192^{***}	-0.028
2	(-1.70)	(0.66)	(6.77)	(-1.33)	(-2.53)	(0.17)	(3.65)	(-0.84)
# Industries	-0.122**	-0.043	0.612^{**}	-0.041	-0.363***	-0.312**	0.542	-0.427
	(-2.28)	(-0.70)	(2.48)	(-0.30)	(-3.16)	(-2.57)	(1.04)	(-1.35)
Competition	0.004	-0.004	-0.158***	-0.014	-0.014	-0.023	-0.353***	-0.160***
	(0.94)	(-0.68)	(-3.42)	(-0.60)	(-1.06)	(-1.60)	(-3.66)	(-3.34)
Education	-0.004	0.000	0.026	0.051^{*}	-0.008	-0.001	0.087^{*}	0.171**
	(-1.54)	(-0.15)	(1.12)	(1.72)	(-1.27)	(-0.08)	(1.70)	(2.33)
Regulation	0.000	-0.001	0.005	0.001	0.003**	0.002	0.015^{*}	0.004
	(0.10)	(-0.67)	(1.63)	(0.74)	(2.36)	(1.05)	(1.88)	(0.62)
Weather	0.000	0.000			0.000	0.000		
	(-0.05)	(-0.05)			(0.81)	0.00		
% Water	0.000	0.000			0.000	0.000		
	(-0.44)	(0.94)			(60.0-)	(0.49)		
Mountainous	-0.004***	-0.005**			-0.007	-0.005		
	(-2.65)	(-2.58)			(-1.63)	(-1.31)		
Year Fixed Effects	Yes	No	No	No	Yes	No	No	No
Firm-Year Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Firm-CBSA Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
L.Employment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
D.City Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LD.City Characteristics	No	Yes	Yes	Yes	No	Yes	Yes	Yes
L2D.City Characteristics	No	No	Yes	Yes	No	No	Yes	Yes
Observations	2,059,247	2,059,247	2,059,247	2,059,247	503,459	503,459	503,459	503,459
Adjusted R^2	0.91	0.93	0.93	0.94	0.88	0.89	0.9	0.91

5 Conclusions

We corroborate the importance of employment density and city size in productivity. Further, we show that firms' sorting across cities and their idiosyncratic preferences for certain cities can confound analyses of the determinants of city productivity. Based on our firm-level index of tradability, we find different drivers of employment for tradable and non-tradable firms. In particular, we find that local competition hinders employment for tradable firms while not affecting the employment of non-tradable firms. Those differences suggest that the determinants of local population growth—the key determinant of demand for firms producing non-tradable output—may differ from those of city productivity.

A limitation of our analysis is that the city-level determinants of firm births may differ from those of growth within existing firms. Suggestive of such a result, Combes et al. (2004) find different determinants of growth within a plant versus growth in the number of plants in France. Furthermore, we do not identify the city-level productivity factors that influence firms' choices of whether and when to enter a new city. Nevertheless, for the large firms we study here, that many cities try to court, our analysis suggests that cities can increase employment by taking steps to allow for additional density and a larger spatial scale.

References

- Aarland, Kristin, James Davis, Vernon Henderson, and Yukako Ono, 2007, Spatial Organization of Firms: The Decision to Split Production and Administration, <u>RAND Journal</u> of Economics 38, 480–494.
- Acosta, Camilo, and Ditte Håkonsson Lyngemark, 2021, The Internal Spatial Organization of Firms: Evidence from Denmark, Journal of Urban Economics 124, 103366.
- Austin, Benjamin, Edward Glaeser, and Lawrence Summers, 2018, Jobs for the Heartland: Place-Based Policies in 21st Century America, <u>Brookings Papers on Economic Activity</u> Spring, 151–232.
- Bartik, Timothy, 1991, Who Benefits from State and Local Economic Development Policies? (W.E. Upjohn Institute for Employment Research).
- Bartik, Timothy, 2020, Using Place-Based Jobs Policies to Help Distressed Communities, Journal of Economic Perspectives 34, 99–127.
- Baum-Snow, Nathaniel, Nicolas Gendron-Carrier, and Ronni Pavan, 2024, Local Productivity Spillovers, <u>American Economic Review</u> 114, 1030–69.
- Baum-Snow, Nathaniel, and Lu Han, 2023, The Microgeography of Housing Supply, Journal of Political Economy Forthcoming.
- Briant, Anthony, Pierre-Philippe Combes, and Miren Lafourcade, 2010, Dots to Boxes: Do the Size and Shape of Spatial Units Jeopardize Economic Geography Estimations?, Journal of Urban Economics 67, 287–302.

- Ciccone, Antonio, and Robert Hall, 1996, Productivity and the Density of Economic Activity, American Economic Review 86, 54–70.
- Combes, Pierre-Philippe, 2000, Economic Structure and Local Growth: France, 1984–1993, Journal of Urban Economics 47, 329 – 355.
- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, 2008, Spatial Wage Disparities: Sorting Matters!, Journal of Urban Economics 63, 723–742.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux, 2012, The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection, Econometrica 80, 2543–2594.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux, 2010, Estimating Agglomeration Economies with History, Geology, and Worker Effects, in Agglomeration Economics, 15–66 (National Bureau of Economic Research).
- Combes, Pierre-Philippe, and Laurent Gobillon, 2015, The Empirics of Agglomeration Economies, in Handbook of Urban and Regional Economics, volume 5 (Elsevier).
- Combes, Pierre-Philippe, Thierry Magnac, and Jean-Marc Robin, 2004, The Dynamics of Local Employment in France, Journal of Urban Economics 56, 217–243.
- Davidoff, Thomas, 2016, Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Many Demand Factors, <u>Critical Finance</u> Review 5, 177–206.

- Davis, Morris, Jonas Fisher, and Toni Whited, 2014, Macroeconomic Implications of Agglomeration, Econometrica 82, 731–764.
- Duranton, Gilles, and Diego Puga, 2000, Diversity and Specialization in Cities: Why, Where and When Does It Matter?, Urban Studies 37, 533–555.
- Duranton, Gilles, and Diego Puga, 2020, The Economics of Urban Density, Journal of Economic Perspectives 34, 3–26.
- Fajgelbaum, Pablo, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar, 2018, State Taxes and Spatial Misallocation, Review of Economic Studies 86, 333–376.
- Glaeser, Edward, Heidi Kallal, José Scheinkman, and Andrei Shleifer, 1992, Growth in Cities, Journal of Political Economy 100, 1126–1152.
- Gyourko, Joseph, Jonathan Hartley, and Jacob Krimmel, 2021, The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index, Journal of Urban Economics 124, 103337.
- Henderson, Vernon, 2003, Marshall's Scale Economies, Journal of Urban Economics 53, 1–28.
- Henderson, Vernon, Ari Kuncoro, and Matthew Turner, 1995, Industrial Development in Cities, Journal of Political Economy 103, 1067–1090.
- Hsieh, Chang-Tai, and Esteban Rossi-Hansberg, 2023, The Industrial Revolution in Services, Journal of Political Economy Macroeconomics 1, 3–42.

Jacobs, Jane, 1969, The Economy of Cities (Vintage).

- Jansen, Mark, and Adam Winegar, 2022, Non-Pecuniary Benefits: Evidence from the Location of Private Company Sales, Working Paper, University of Utah.
- Mian, Atif, and Amir Sufi, 2014, What Explains the 2007–2009 Drop in Employment?, Econometrica 82, 2197–2223.
- Moretti, Enrico, 2004, Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions, American Economic Review 94, 656–690.
- Neumark, David, and Helen Simpson, 2015, Place-Based Policies, in <u>Handbook of</u> Regional and Urban Economics, volume 5b (Elsevier).
- Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre Daniel-Sarte, and Nicholas Trachter, 2023, Plants in Space, Journal of Political Economy Forthcoming.
- Orr, Scott, 2022, Within-Firm Productivity Dispersion: Estimates and Implications, Journal of Political Economy 130, 2771–2828.
- Porter, Michael, 1990, The Competitive Advantage of Nations, <u>Harvard Business Review</u> 68, 73–93.
- Rappaport, Jordan, 2007, Moving to Nice Weather, <u>Regional Science and Urban</u> Economics 37, 375–398.
- Roback, Jennifer, 1982, Wages, Rents, and the Quality of Life, <u>Journal of Political Economy</u> 90, 1257–1278.
- Rosenthal, Stuart, and William Strange, 2008, The Attenuation of Human Capital Spillovers, Journal of Urban Economics 64, 373–389.

- Rosenthal, Stuart S., and William C. Strange, 2020, How close is close? the spatial reach of agglomeration economies, Journal of Economic Perspectives 34, 27–49.
- Rossi-Hansberg, Esteban and Sarte, Pierre-Daniel and Schwartzman, Felipe, 2019, Cognitive Hubs and Spatial Redistribution, Working Paper 26267, National Bureau of Economic Research.
- Slattery, Cailin, and Owen Zidar, 2020, Evaluating State and Local Business Incentives, Journal of Economic Perspectives 34, 90–118.
- Viladecans-Marsal, Elisabet, 2004, Agglomeration Economies and Industrial Location: City-Level Evidence, Journal of Economic Geography 4, 565–582.

APPENDIX

Sample
Matched
Coefficients in
Correlation
Pairwise
Table A.1.

This table presents pairwise correlation coefficients between the variables included in the matched sample over the 1997–2021 period. See Table 4 for variable definitions.

	(1)	(2)	(3)	(4)	(5)	(9)	(2)	(8)	(6)	(10)	(11)	(12)
(1) Employment	1.000											
(2) Density	0.161	1.000										
(3) Total Area	0.109	-0.175	1.000									
(4) Specialization	-0.037	-0.157	-0.049	1.000								
(5) Diversity	0.007	-0.016	0.119	0.010	1.000							
(6) # Industries	0.145	0.425	0.458	0.017	0.000	1.000						
(7) Competition	0.011	0.121	-0.038	-0.034	-0.050	-0.023	1.000					
(8) Education	0.101	0.454	0.120	-0.032	-0.243	0.442	-0.063	1.000				
(9) Regulation	0.029	0.112	0.079	0.012	-0.066	0.042	0.284	0.190	1.000			
(10) Weather	0.013	0.098	-0.010	0.124	0.002	-0.085	0.394	0.056	0.264	1.000		
(11) % Water	0.039	0.407	-0.252	0.082	0.014	0.141	0.067	0.087	0.172	0.008	1.000	
(12) Mountainous	-0.031	-0.211	0.096	0.102	-0.054	-0.200	0.293	-0.069	0.353	0.674	-0.073	1.000