Measuring Two Decades of Urban Spatial Structure: The Evolution of Agglomeration Economies in American Metros*

ELIJAH KNAAP, San Diego State University
SERGIO REY, San Diego State University

In this paper we examine the evolution of urban spatial structure in U.S. metropolitan areas over nearly two decades. Using annual block-level data from the Longitudinal Employment Household Dynamics database, we introduce a technique for identifying regional employment centers that both adheres to urban economic theory and pays homage to classic contributions in local spatial statistics. Centers are defined as local spatial statistical outliers on the network-based job accessibility surface. We proceed by identifying the location and employment makeup of centers for each metropolitan region in the USA from 2002 to 2019 and discuss emergent trends across time and space. Critically, we not only explore empirical patterns, but we discuss the relationship between polycentricity, the evolution of urbanization and localization economies, and regional specialization. We confirm again the pattern of polycentricity in U.S. metros and show that the structure of metropolitan employment is largely stable over time. We also document a continuing trend away from urbanization economies into more specialized subcenters.

Keywords: employment centers, spatial structure, spatial statistics, agglomeration, accessibility

INTRODUCTION

Measuring urban spatial structure is both a classic pursuit in regional science and a burgeoning topic in spatial analysis research. Today, a large and growing body of research documents the presence of urban polycentricity, comprised of an interconnected system of employment subcenters, defined as geographic loci of large concentrations of employment. Yet there remains a great deal of disagreement on (1) how best to define, identify, and study employment centers, and (2) what leads to their configuration and industrial composition. Here we develop a new technique that does not require local tuning parameters and thus scales well to facilitate comparative analysis. We apply the method to nearly 20 years of high-resolution data in the U.S. to examine the prevalence and durability of polycentricity, as well as the industrial composition of centers across the country and their changes over time. Our results show both the stability of spatial structure over time as well as its diversity across space. Monocentrism is the most common form of spatial structure, but polycentrism defines the median. Centers are also becoming moderately more specialised over time.

Although the intellectual tradition of 'spatial structure' extends, arguably to von Thünen (1826), an inflection point in empirical work began in the 1980s with a wave of research on employment centers and polycentric urban form. As concerns over urban sprawl mounted, research on employment subcenters (Anderson & Bogart, 2001; Cervero & Wu, 1997; Cervero & Wu, 1998; Giuliano & Small, 1991; Gordon et al., 1986; Gordon & Richardson, 1996; Griffith, 1981; Levinson & Kumar, 1994; McDonald, 1987) developed alongside theoretical models that explained their emergence (Helsley & Strange, 2007; Helsley & Sullivan, 1991; Straszheim, 1984; White, 1976, 1988; Zhang & Sasaki, 1997).

The motivating problem behind this growing body of early scholarship was understanding the

*This work is supported by NSF-SES Grant 1831615

performance and interrelationship of urban housing and labor markets, and the ways that infrastructure or public policies modified these markets (Anas, 1984; Anas, 1985, 1990; Eberts & McMillen, 1999; Kloosterman & Musterd, 2001). Fundamental to this pursuit are questions about economic development (Dissart, 2003; Kane et al., 2018; Knaap et al., 2016; Niu et al., 2015), transportation performance (Cervero & Wu, 1998; Giuliano & Small, 1991; Giuliano & Small, 1993), and the appropriate provision of housing and jobs (Giuliano, 1991; Knaap et al., 2016; White, 1988), which in turn inform urban planning and policy decisions such as zoning, affordable housing siting, environmental regulations, or transit investments.

Despite maintaining a consistent presence for decades across the literature of urban economics, geography, and planning, research on employment centers and polycentrism has drifted considerably in focus. Today, a large share of employment center research focuses on methodological innovation at the expense of explanation, policy analysis, or theoretical development; methods, for example, have been developed based on employment/housing ratios (Manduca, 2020), diffusion statistics (Hipp et al., 2021), local spatial statistics (Arribas-bel & Sanz-gracia, 2014; Ban et al., 2017; Hajrasouliha & Hamidi, 2017), machine learning (Arribas-Bel et al., 2021; Arribas-Bel & Schmidt, 2013), and remote sensing (Baragwanath et al., 2021), but none of these papers examines the role of subcenters in the larger economy. This intense focus on methods may hinder the ability to study the economic, social, and political forces that lead to the emergence of polycentric structure. Since these forces are most often the objects of social science inquiry (and targets for policy intervention), rather than the polygons or other geometric artifacts that operationalize employment centers, we argue that the former should be a greater focus of study than the latter.

More specifically, apart from the requisite nod to forefathers Alonso, Muth, and Mills, a recent stream of papers on employment center identification and spatial structure eschews much of the context that motivated research on employment centers from its inception, ignoring the urban economic foundations of why centers exist (Hipp et al., 2021; Manduca, 2020). In our view, this trend embodies a classic critique of the field given by Isserman (1995), that "applied regional science should be problem-driven, not method-driven or theory-driven. It must be more than the storehouse of analytical methods that have been regional science's main contribution to date."

In this paper, we consider the problem of employment center identification as one of economic geography and firm location choice. Our goal is to understand the forces that help shape the spatial allocation, competition, concentration, and cooperation of different industries in a metropolitan region, and the ways they shape land-use patterns in metropolitan regions over short and medium time frames. To do so, we first introduce a technique that adheres to both urban economic theory and pays homage to classic contributions in local spatial statistics; this provides a unique opportunity to bridge the gap between approaches rooted in point pattern analysis versus those rooted in threshold criteria discussed by Anas et al. (1998).

We define metropolitan employment centers as local spatial statistical outliers on a transport network-based accessibility surface. This effectively conceptualizes an employment center as a collection of intersections in the transportation network We then use a computational geometry algorithm to bound intersections that belong to the same density cluster, yielding discrete employment centers whose shape is based on the underlying transport network. This conforms to our notion of centers as intra-urban agglomeration economies whose locational return to scale results in the non-random clustering of firms in space (Ahlfeldt, 2011; Anas et al., 1998; Redding, 2023). Our approach is fully reproducible, scalable, and applicable to any metropolitan area on the globe, and we make all developed methods available via open-source software.

We proceed by identifying employment centers in every metropolitan region in the U.S. on an annual basis between 2002 and 2019, and we summarize at a national scale both the pattern of polycentricity that characterizes American metropolitan regions in the modern age and the evolution of these patterns over the two-decade period. Specifically, we examine the patterns of subcenter growth, merging, splitting, and disappearance over time, as well as the interrelationship of these patterns. We then turn to a compositional analysis of the centers, summarizing the evolution of urbanization and localization economies, and emergent trends in regional specialization over time.

The paper proceeds as follows. In the next section, we review the literature focused on metropolitan employment centers and urban spatial structure. Our focus here is the intellectual lineage of employment center identification, the research trajectory over the last three decades, and the different techniques applied in the field, and the discussion trade offs among competing approaches. In the following section, we introduce our new method for identifying centers and subsequently describe results for an application to 379 metropolitan regions in the United States. The final two sections provide an interpretation of our findings, as well as a discussion of the assumptions, global parameters, and potential limitations of our method, before offering a summary and extensions for future research.

POLYCENTRIC URBAN FORM IN ECONOMIC GEOGRAPHY

Volumes of research over the last three decades have documented the emergence of polycentrism as a dominant form of urban spatial structure in the United States and abroad. Given the focus of our empirical application, we limit our primary attention in this paper to the discussion of spatial structure identifying employment centers in the U.S. context. The European literature on polycentrism and employment centering is at least as rich as the American context, however.[1]

As with the formation of cities themselves, employment centers in urban areas result from spillover benefits accruing from the concentration of economic activity in physical space. More directly, employment firms are more productive when they cluster together, which leads to some predictable patterns in urban structure. Classical bid-rent theory describes the logic underlying the general layout of a monocentric city, where workplaces gain more utility from centrally-located

[1] For background, see for example Zhang & Derudder (2019), Riguelle et al. (2007), Heider et al. (2022), Krehl & Siedentop (2019), and Bartosiewicz & Marcinczak (2022) for a survey of recent work. Despite our American focus, the techniques described hereafter are all applicable to any global context. Indeed, as a reviewer pointed out, many countries outside the American context have better-positioned governmental structures for affecting polycentric development, and the strategies developed here may be even more useful in those cases.

land, and thus outbid residences for these areas (Alonso, 1964; von Thünen, 1826). While the pattern of employment concentration is a well documented empirical reality, the mechanisms that lead to this resulting structure remain an active area of research in economic geography (Redding, 2023; Rosenthal & Strange, 2004), particularly as economies restructure and cities are increasingly characterized by polycentric form. Naturally, these trends beg the question of why employment firms cluster together and whether the patterns that drive them such may change (resulting in a new urban layout).

Empirical research on agglomeration emanates from the classic discussion of scale economies given by Marshall (1920). Summarizing the century of work that followed, Duranton & Puga (2004) discern three underlying mechanisms they label sharing, matching, and learning, where *sharing* results from more efficient use of inputs and infrastructure that lowers fixed costs; *matching* increases productivity by increasing the quality or probability of matching (e.g., between producers and consumers, or when "stronger competition helps to save in fixed costs by making the number of firms increase less than proportionately with the labour force" (Duranton & Puga, 2004)); and *learning* results from sharing, generation, or accumulation of knowledge, all of which make firms more efficient.

Empirical work over the last two decades has shown support for each of these mechanisms (Redding, 2023), suggesting there is more to learn about (a) whether different mechanisms are more reliant on space, and (b) whether different industries (alone or in concert) are better equipped to leverage one type of agglomeration mechanism versus another. For example, if knowledge generation is facilitated entirely by digital communication, then we might expect tech clusters to disappear (unless they are also dependent on labor-sharing). Further, we might expect greater industrial mix (i.e. urbanization economies) inside subcenters if matching mechanisms are driving spatial clustering, since greater diversity would provide for greater matching opportunities. This leads to a natural question about the relationship between subcenter formation and the forces that drive agglomeration.

Identifying Employment Centers

There are two major threads in employment center identification research. In regional science, urban planning, and policy analysis, the emphasis is often economic development (as measured by employment growth) and transport efficiency. These papers, of which Giuliano & Small (1991) is the canonical example, focus on the conceptual forces driving patterns away from the classic urban monocentric model. That is, they examine the central role of transportation in defining urban land prices, and ways that processes like land-capital substitution shape the allocation of homes and jobs or the co-location of certain industries (Agarwal et al., 2012; Giuliano et al., 2012; Liu et al., 2020; Meijers et al., 2018).

This thread is also tied closely to the literature on industrial clustering, knowledge spillovers, and new economic geography (Davis & Dingel, 2019; Harris, 2020; Krugman, 1999; Markusen & Porter, 1996; Porter, 1990; Porter, 1998; Rey, 2002). , Well-known examples of applied work in this tradition include Gordon et al. (1986), Gordon & Richardson (1996), Small & Song (1994), Anderson & Bogart (2001), and more recently Niu et al. (2015), Knaap et al. (2016), Craig et al. (2016), and Kane et al.

(2018). Similar work by von Ehrlich & Seidel (2013) stresses the importance of firm heterogeneity in fostering agglomeration, in line with the emergence of multiple polycentric localization economies that succeed the larger urbanization-based monocentric model. Craig et al. (2016) also find similar evidence in an empirical analysis of Houston.

In a growing body of work leveraging data science and machine learning, the emphasis is often on new methodological developments or unique data mining strategies that provide a fresh look at the data compared to classic techniques. Toward these goals, a more recent trend in geographical analysis uses increasingly sophisticated statistical techniques to identify spatial concentrations (or diffusions) of jobs, often without identifying discrete employment center polygons. Unlike the other thread, this literature is focused on the emergence and description spatial patterns via statistical analysis rather than questions about the causes of employment reconfiguration or the performance of various urban social systems or policy measures. Examples include Redfearn (2007), Arribas-bel & Sanz-gracia (2014), Hajrasouliha & Hamidi (2017), Manduca (2020), and Hipp et al. (2021).

While useful from an exploratory spatial analytical perspective, we argue this latter thread has limited utility for understanding the processes of economic development or firm location choice because the techniques eschew, by design, much of the original intent of employment center research (i.e., to examine how intra-regional economies are shaped by spatial agglomerative processes). As such, these techniques provide a thorough description of the *layout* of urban spaces, but largely fail to provide insight into the function and structure thereof. Instead, we view employment centers arising from a (quasi)rational location choices of multiple firms, guided chiefly by transportation costs, access to amenities, and disincentives from externalities. That is, "employment centers are not statistical anomalies-—they are not 'noise' residuals from a correctly specified monocentric model. Employment centers are local agglomerations that have their own fundamentals and may be highly useful as units of analysis." (Agarwal et al., 2012). To study these local agglomerations, it is necessary to first define and identify the discrete units of analysis.

A notable example that straddles these approaches is Arribas-Bel & Schmidt (2013), who define a set of theoretically relevant indicators of the local economy and a spatially-constrained clustering algorithm to delimit employment centers. Despite the unique use of urban theory to select a set of input variables, however, their analysis nevertheless focuses on the location and description of each identified center, rather than an examination of its industrial makeup or its performance within the larger metropolitan system. A second example of integrating spatial machine learning into the identification of employment centers is given by Arribas-Bel et al. (2021) who use building volumes to delineate urban areas, though their dataset does not include employment totals, and thus cannot speak to the composition of agglomeration economies, only the clustering of buildings.

Following Kloosterman & Musterd (2001, p. 631), we agree "more theoretically founded research is needed," and with Agarwal et al. (2012) that "polycentricity—-the dominant urban form for large US metropolitan areas—offers a natural laboratory to reveal much about sub-metropolitan agglomeration economies and their role in determining urban spatial structure... However, surprisingly, little is known about these centers. Much can be learned about the nature of agglomeration economies from studying their characteristics, emergence, growth, and decline."

Here, we attempt such an analysis, leading us to two distinct bodies of inquiry which we examine in turn. Our first question concerns whether the spatial patterning of employment is consistent across time and over space. This comports with the dominant mode of analysis in the literature, except that we consider all metropolitan regions in the U.S. (annually) over a nearly two-decade period to analyze trends at the national scale. This frame provides a look at the *distribution* of general employment patterns, both across the country and within each region over time (i.e., whether cities are typically monocentric or polycentric), whether the number of centers or their spatial footprint is growing or shrinking (e.g., possibly indicating urban sprawl or decline), and whether centers emerge, split, merge, or disappear over time. Together this comprises an analysis of growth, persistence, and land consumption that may be useful for transportation and sustainability research and policymaking.

Our second question concerns whether the composition of the revealed employment centers (i.e., the size and share of total employment, and industrial mix contained within the boundary) is consistent over time and space. This frame provides the potential to identify different varieties of agglomeration forces, (like localization versus urbanization) and whether one style dominates the other or leads to faster or more stable growth over time. This frame also provides a view into different potential underlying drivers of agglomeration (like knowledge spillovers or input sharing) by examining whether cooperative or complementary industries co-locate (and which industries, in which places, and whether these patterns uphold over time, etc).

Considering these different frameworks and their interaction allows a more nuanced picture of economic geography than focusing solely on the layout and statistical approaches for identifying employment centers. For example, is urban sprawl fueled by expansion of the tech industry in suburban office parks? Or do all industries gain fewer benefits from concentration today (given the ubiquity of online interaction and transaction), leading to an overall pattern of decentralization? Are urbanized centers more stable than localized ones? Naturally these would have different policy implications

Yet, conducting these inquiries requires a technique for identifying employment centers that is both generalizable to a wide variety of urban contexts and computationally scalable to large-volume, high-resolution employment data. Toward these ends, existing methods have several notable drawbacks, thus an unavoidable third question focuses on developing a technique with desirable conceptual and computational properties for comparative analysis. We review existing strategies briefly below to describe their shortcomings for addressing our research questions.

Conceptual Issues For Employment Center Definitions in Empirical Research

A thorough overview of employment center identification techniques is given by Hajrasouliha & Hamidi (2017). In general, employment center research falls into three primary categories based on the empirical technique used to study the spatial patterning. The classic approach uses minimum thresholds for employment density and total employment to define centers. Well known papers in this vein include Giuliano et al. (2007), Agarwal et al. (2012), Cervero & Wu (1997), Giuliano & Small

(1991), and Cervero & Wu (1998). Another approach uses various flavors of regression, such as Locally Weighted Regression (LWR) to investigate density gradients and examine whether a polycentric model or monocentric model provide a more realistic picture of contemporary cities. Examples include McMillen & McDonald (1997), Schmidt et al. (2020), Garcia-López & Moreno-Monroy (2018), and McMillen & Smith (2003). A final burgeoning body of work uses spatial statistics, sometimes in concert with novel methods that examine diffusion versus centralization. Recently published examples in this tradition include Arribas-bel & Sanz-gracia (2014), Krehl (2015), Manduca (2020), Hipp et al. (2021) and Hajrasouliha & Hamidi (2017)

Density cutoffs are conceptually straightforward but require local knowledge of each study area to set appropriate parameters which may be difficult to posit a-priori and limits potential for comparative analysis since the parameters change with the locale. Regression approaches are designed to test only whether polycentric or monocentric specifications better fit the data, and thus fail to identify the boundaries of any employment center for further analysis. Spatial statistical approaches generally assume that outliers are indicative of "important" concentrations, but some authors like Agarwal et al. (2012) question the justification for this notion. A pocket of employment may draw a disproportionately large share of trips, or provide increasing returns to growth for firms inside the pocket, even if the pocket does not amount to enough employment to be "statistically meaningful".

Furthermore, techniques that focus on dispersion and centralization generally ignore the industrial mix of employment concentrations given their focus on spatial spread, which limits insight into the potential causes of agglomeration. Some approaches such as Guillain et al. (2006) and Manduca (2020) further rely on heuristics like population/employment ratio provide similarly little information about the causes of clustering. We do not expect jobs and housing to be balanced in a single location, because there is "well-documented evidence that economic activity within urban areas is markedly more concentrated and presents different patterns of location to those presented by residential areas" (Arribas-Bel et al., 2021).

In other cases, authors include measures such as distance from the central business district (CBD) as a method for identifying employment centers (Hajrasouliha & Hamidi, 2017). This choice is useful for exploring how structure has evolved from the historic urban core, but it also assumes that the central business district is static, and that the *center* itself is immobile. We argue this is a curious reversal of logic, since the CBD is defined (endogenously) by employment density rather than the reverse. In what follows, we provide a bridge between density cutoffs and spatial statistics, providing a unique method for identifying discrete employment centers based on theoretically-justified criteria that still relies on the best computational methods available.

MEASURING SPATIAL STRUCTURE AND AGGLOMERATION ECONOMIES

Our approach adopts two perspectives; first, that agglomeration economies are best captured by measures of *accessibility* to employment rather than raw aggregates of arbitrary boundaries or grid cells. Among other benefits, this also helps avoid the issue of the modifiable areal unit problem (MAUP) arising in this specific context and discussed by McMillen (2003), Kane et al. (2018), and others who find that the identification of centers is dependent on the zoning scheme (or cell size) used to aggregate data. A corollary to this view is that urban space should be measured in ways that reflect realistic movements through the built environment.

The second perspective holds that employment centers are best used as vehicles for studying urban policy measures and the mechanisms that guide location allocation and generate productive spillovers, rather than objects of study in their own right. That is, identifying employment centers is useful only if the exercise reveals some insight about spatial structure, which is poignant in the case of employment centers, given their widespread use in regional planning efforts both internationally (Derudder et al., 2022) and domestically[2] (Knaap et al., 2016). Adopting this perspective, we attempt to address the critique that "further understanding of urban polycentricity is necessary for understanding how large cities generate gains for their residents" (Craig et al., 2016, p. 26) by developing a method that permits comparative analysis across space and time, and allows compositional analysis of the identified employment centers.

Our approach recognizes travel infrastructure as the skeleton of an urban system; thus, the base-unit of a transport system (a street intersection/network node) serves as the primary unit of analysis. We begin by assigning each census block to its representative intersection in the metropolitan street network. We collect the network from OpenStreetMap (OSM) and it remains fixed in each metropolitan region over time. That is, we assume the infrastructure network is constant, and temporal variation is created by the growth, decline, or relocation of jobs allocated to the network[3]. By using the travel network as the backbone of the analysis, we help ensure that employment center boundaries are constricted to regions served by infrastructure. This is not intended to serve as a perfect depiction of the travel network over time, but to limit the smoothing along actual infrastructure rather than free space (Knaap & Rey, 2023).

We then define employment centers based on accessibility to jobs from each street intersection rather than concentration into an arbitrary administrative polygon. Our technique is proceeds as follows. We first "attach" each census block (with total employment counts for each two-digit category) to their nearest node in the street network) and use network analysis to define employment accessibility at each node. Specifically, we use the contraction hierarchies technique to process each metropolitan street network[4] (Geisberger et al., 2012), then compute the shortest path between each pair of nodes within a distance of two kilometers. At each node i in the network, we compute a gravity-based accessibility measure a_{i_k} as

[2] See, for example https://www.sandag.org/index.asp?classid=16&subclassid=127&projectid=581&fuseaction=projects.detail for the prominent role of employment centers in regional planning at the San Diego Association of Governments (SANDAG)

[3] In general, we believe this is a reasonable assumption, as street network patterns generally change little over time. We recognize, however, that in some metropolitan areas–especially those which are growing rapidly–may have some change in the underlying transport network. In these cases the scale of change is likely to be small and in a single direction (that is, the network generally grows as new streets are added; rarely does the network shrink because streets are destroyed). Because we use the *current* street network, it is likely that we capture the universe of potential employment locations. If there is any bias, it is upward.

[4] The contraction hierarchies technique is a method for graph pre-processing designed specifically for large street networks. It provides a massive computational boost for finding the shortest path through network space by reducing the number of nodes considered by Dijkstra's algorithm

$$a_{i_k} = \sum_{j \in \{j | d_{ij} < 2000\}} exp\left(-1 \times \frac{d_{ij}}{2000}\right) \times t_{j_k} \quad (1)$$

where k is a two-digit employment sector defined by the North American Industrial Classification System (NAICS), d_{ij} is the distance along the shortest network path (in meters) between nodes i and j, and t_{j_k} is the total employment in sector k located at node j [5]. This step accomplishes several tasks simultaneously. First, it helps overcome the sparsity and aggregation tradeoff encountered when using employment data at different scales, namely that observations are too sparse when using address-level data, and variably sized when using polygon-level data as Traffic Analysis Zones (TAZs) or census tracts. Second, it transforms the total employment count into an "attractiveness" measure consistent with theories of agglomeration and location choice (Hansen, 1959a; Hansen, 1959b; Lynch & Rodwin, 1958). Finally, it constricts our analysis to the local transportation network that conditions movement between employment locations, yielding a much more realistic picture of the distance between opportunities (Duschl et al., 2014). We compute all network accessibility measures using the open-source Python package `pandana` (Foti et al., 2012).

These new measures, a_{i_k}, approximate a continuous surface along the transport network's spatial topology, and become the input data in our further analyses. Conceptually, this process is akin to the procedure employed in many other employment center studies Maoh & Kanaroglou (2007), Maoh et al. (2010), Kane et al. (2018), Redfearn (2007), McMillen (2001)], which use a kernel density function through a raster data model, except that our "smoother" uses the transport network rather than Euclidean distances to "kernelize" observations. Our approach overcomes two drawbacks of these two studies (identified by Hajrasouliha & Hamidi (2017)), which are that raster models fail to identify discrete center boundaries, and planar kernel functions ignore urban infrastructure or topological constraints (valleys, mountails, rivers, etc.).

Employment Centers as Nonrandom Concentration(s) of Economic Activity

We define a metropolitan employment center as a collection of street intersections with statistically significant access to a large concentration of jobs (where *large* is measured by a disproportionate share of transportation trips relative to the region). The primary challenges to classic employment center identification, as discussed by McMillen (2003), are determining (1) whether a concentration of employment is large enough to be meaningful, and (2) concentrated enough to comprise a "center". Any 'contiguous' observations meeting these eligibility criteria can be collapsed into a single center. These two criteria map onto the total and density requirements specified by Giuliano & Small (1991), but have difficulty scaling to comparative work, as discussed earlier. To address these drawbacks, we offer two incremental improvements.

To meet the "total threshold" requirement, an observation must be in the top quintile of job ac-

[5] To identify employment centers we examine total employment at each node ($t_{j_k} = \sum^k t_j$), however we include the k subscript in this formulation because it sets the same notation for concentration indices introduced in the next section

cess within the study region (that is, $a_{i_{total}} \geq \pi_{0.8}$). This is a *relative* standard that should apply well in different urban contexts, requires no local knowledge, and may be defined globally for comparative work. Although the 20% metric is subjective, it sets a liberal standard that should identify a large set of candidate sites. We choose 20% given the widespred use of employment centers in practical transportation planning, and that the distribution of Home-Based-Work trips is Pareto-shaped, meaning this definition is likely to capture roughly 80% of all commutes (Arribas-Bel et al., 2021).

To meet the density requirement, the spatial concentration of high job accessibility surrounding each node in an employment center must be large enough to achieve statistical significance at the one percent level, according to an analysis of local spatial association. Specifically, each node must have a G_i^* statistic which is significant at the $\alpha = 0.01$ level, ensuring that jobs are sufficiently concentrated in space to distinguish the employment center from its nearby intersections. Again, while one percent is a subjective threshold, it sets a stringent, widely-accepted statistical standard (whose conservatism narrows down the list of candidate sites).

The G_i^* statistic introduced by Getis & Ord (1992) and extended by Ord & Getis (1995) is a spatial statistical measure of concentration used in the geographical analysis literature to study "hotspot" occurrences of phenomena like disease outbreaks (Getis & Ord, 1992; Kao et al., 2008; Ord & Getis, 1995), traffic accidents (Abdulhafedh, 2017; Songchitruksa & Zeng, 2010), and indeed, employment concentration (Baumont et al., 2004; Guillain et al., 2006; Hajrasouliha & Hamidi, 2017; Scott, 1999). The G_i and G_i^* statistics are uniquely suitable to employment center identification because they deal only with positive numbers and focus exclusively on positive spatial autocorrelation (unlike Moran's I). In many spatial analyses this distinction can be a weakness, but in this case we argue the focus on positive spatial autocorrelation is a strength in favor of G_i^*, as the search for employment centers is similarly focused *only* on discovering large concentrations, not gaps.

The G family of statistics identify clusters of nearby observations with high attribute values (or clusters of nearby observations with low values, which in the case of employment centers is substantively uninteresting). As Hajrasouliha & Hamidi (2017, p. 426) describe, "when the local sum (a feature's value and the values for all of its neighboring features) is much higher than the expected local sum, and that difference is too large to be the result of random chance, there is a statistically good chance that the feature is part of a hot spot." Here, we adopt the G_i^* statistic (which includes the focal observation in the analysis), defined by Ord & Getis (1995) as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij} a_j - \bar{A} \sum_{j=1}^{n} w_{ij}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{ij}^2 - \left(\sum_{j=1}^{n} w_{ij} \right)^2 \right]}{n-1}}} \tag{2}$$

with

$$S = \sqrt{\frac{\sum_{j=1}^{n} a_j}{n} - \bar{A}^2} \tag{3}$$

where $a_j = \sum_k a_{j_k}$ is the sum of industry accessibility measures defined in Equation 1 for center j, $\bar{A}$ is the mean accessibility score for the region, and w_{ij} is a spatial weights matrix indicating the connectivity between units i and j. Here, we use a binary weights matrix based on the 100 nearest neighbors of observation i. Although these statistics have been used in the employment center literature previously, the sparsity of employment data can wreak havoc on an analysis because very small localized pockets of employment can show up as 'significant' simply because they are surrounded by places with little to no development. This is especially true for highly disaggregate data such as the LEHD dataset which is provided at the census block level. The spatial smoothing step we apply by transforming the input data into an accessibility surface skirts this issue.

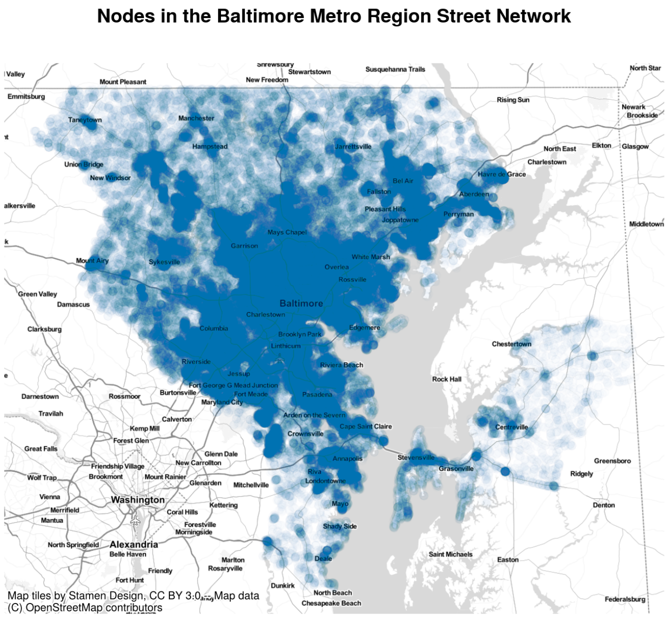
Upon selecting nodes that meet these two criteria, we define nodes as belonging to the same employment center if the are within a one kilometer distance from one another[6], and we use a computational geometry method known as an *alpha shape* to wrap a concave hull polygon around the nodes belonging to each center to define the discrete boundary of each center (Edelsbrunner et al., 1983). These computations are all performed using the open source Python package `libpysal` (Rey et al., 2021). Finally, we discard any polygon whose face is less than one square kilometer in total area. This process yields a distinct (set of) polygon(s) representing the employment center(s) identified in the study region. In fairness, this choice amounts to a third selection criterion, that an employment center must be larger than one square kilometer (half the search distance of the spatial smoothing step). This process is outlined in Figure 1.

Characterizing Agglomeration Economies Inside Subcenters

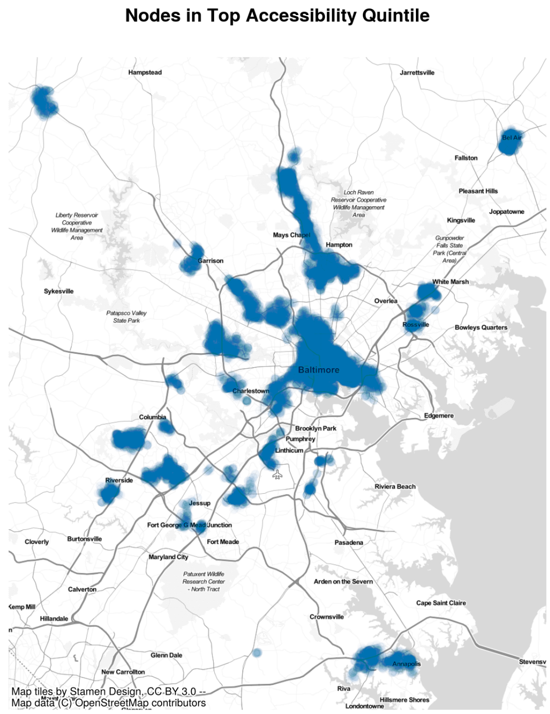
Our object of study is the spatial concentration of employment, as well as its underlying causes and potential consequences. Together this knowledge can facilitate better policy and planning for economic development, transportation congestion, environmental quality, and housing affordability. To understand the benefits of locating inside an employment center, we examine the ways that economies of scale play out differently across centers. The urban economics literature distinguishes between two forms of spatial agglomeration, where localization describes scale economies resulting from similar firms occupying a space, whereas urbanization describes scale economies resulting from dissimilar firms (Henderson, 2003).

We use two metrics to characterize the industrial structure and scale economies in each revealed employment center. Specifically, we use the two-digit NAICS categories to calculate Location Quotient (LQ) and Herfindahl-Hirshman (HHI) indices which together describe the relative strength of different industries inside each center. While the HHI index helps determine whether the location is dominated by a particular industry, the LQ index informs *which* industries maintain a stronger presence in each location. Reviewing our notation, our measures are based on an accessibility score a, computed for each node in the street network i, in each two-digit NAICS category k, in each of the centers c uncovered by our method.

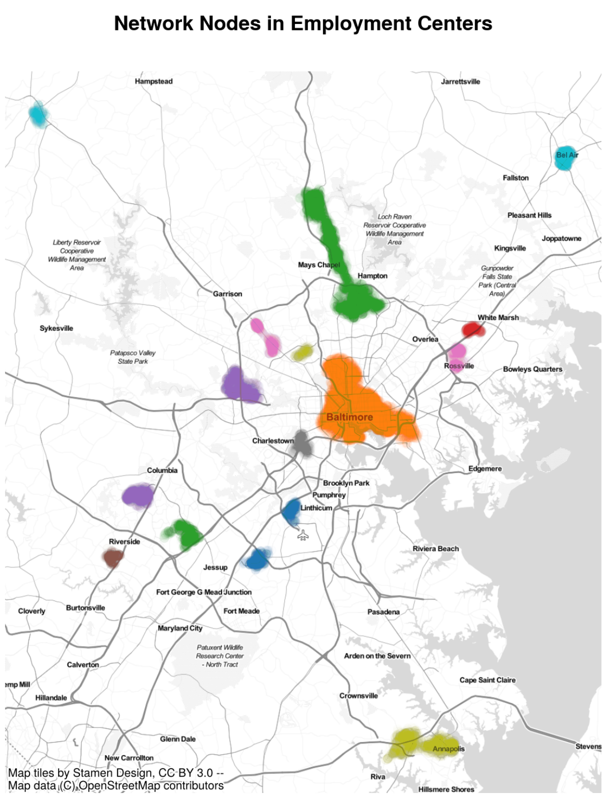
[6] This is identical to the procedure described by McMillen (2003, p. 69), where here the W matrix is defined by a one-kilometer distance-band instead of 1.25 miles



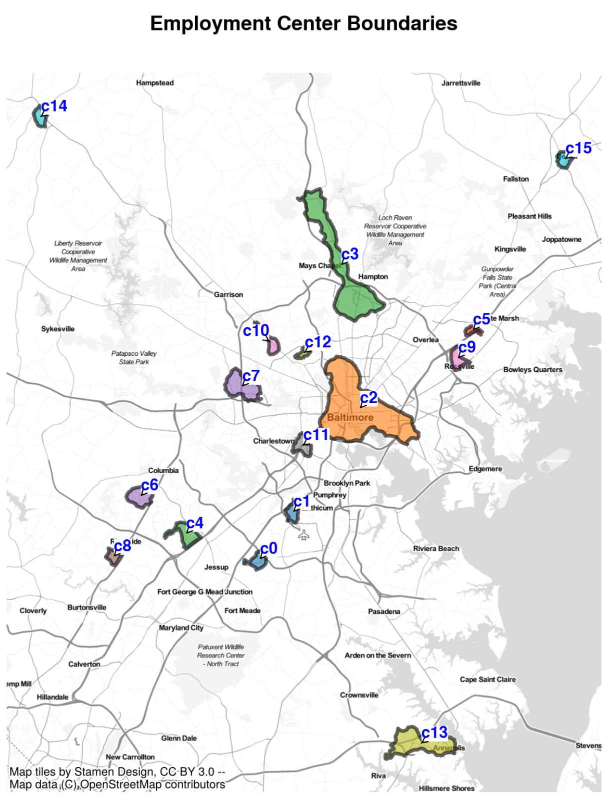
(a) All Nodes in the Baltimore Region



(b) Nodes in top Accessibility Quintile



(c) Nodes in Employment Centers



(d) Employment Center Polygons

Figure 1: Employment Center Selection Process for the Baltimore Region in 2019

Location Quotients

The LQ index quantifies the concentration of an employment center into an employment center, and can be conceived as a measure of local specialization (Isard, 1960). It compares the share of each industry's employment inside an employment center to the share of that industry's total employment in the region, with higher numbers indicating a greater domination by the given industry. Following the notation from Equation 1, we define the spatial location quotient for each industry k in each employment center c as

$$LQ_{kc} = \frac{\left(\frac{\sum^{I} a_{i_{kc}}}{\sum^{K}\sum^{I} a_{i_{kc}}}\right)}{\left(\frac{\sum^{I} a_{ik}}{\sum^{K}\sum^{I} a_{ik}}\right)} \tag{4}$$

where the numerator $\left(\sum^{I} a_{i_{kc}} / \sum^{K}\sum^{I} a_{i_{kc}}\right)$ is the share of accessibility to jobs in industry k belonging to center c, and the denominator $\left(\sum^{I} a_{i_k} / \sum^{K}\sum^{I} a_{ik}\right)$ is the share of access to employment in industry k for all nodes in the region. This ratio-of-ratios describes the concentration of an industry inside each employment center relative to the region's overall share of that industry. A Location Quotient of 2.0 would indicate an employment center has twice the share of jobs in a given industry as its share of total jobs in the region.

Herfindahl-Hirshman Indices

The Herfindahl-Hirshman (HHI) index is a measure of competition and market penetration. The version of the HHI index applied here treats nodes in the transport network as "firms" which provide a share of access to employment. The total employment is the sum of a_i values over all nodes in the network and the contribution of the "firm" is the contribution of the accessibility at each node. The HHI measure is also known in the ecology and segregation literature as Simpson's Diversity index. Following again the notation from Equation 1, we define our HHI index for each employment center c as

$$HHI_c = \sum_{i=1}^{N_{ic}} \left(\frac{a_{i_{kc}}}{\sum a_{i_{kc}}}\right)^2 \tag{5}$$

where $\sum a_{i_{kc}}$ is the sum of accessibility values for nodes i in industry k, located in center c, and $\sum a_{i_{kc}}$ is the sum of accessibility values for all industry categories in nodes located at center c. The HHI index ranges from $1/N$ to unity, and in this case measures the degree to which each employment center is dominated by a single industry. A high value suggests "monopolistic" control of the land area. In other words, high HHI values are indicative of localization economies, since a single industry dominates the spatial market.

Continuing the example shown in Figure 1 from the Baltimore region in 2019, we find strong evidence of localization economies in some subcenters. For example, Center c12, which spans the campus of Sinai Hospital has a very high HHI (0.65) indicating an active, highly specialized localization economy; the largest LQ for that center is NAICS 62 (healthcare and social assistance) with a value of 4.2. Similarly, Center c1, located at the Baltimore-Washington International airport, has a high specialization (HHI=0.35) with a large LQ in manufacturing, and Center c8, anchored at the Maple Lawn applied physics lab, has a high specialization (HHI=0.62) with a large LQ in professional and scientific jobs. By contrast, the region's traditional central business district (CBD) in downtown Baltimore City (c2) has a low HHI of 0.120, with no dominating LQ from a particular industry, indicating evidence of a continued urbanization economy, as expected from the historic urban core.

	c0	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15
HHI	0.296	0.357	0.12	0.085	0.144	0.118	0.121	0.215	0.629	0.237	0.158	0.157	0.651	0.13	0.119	0.126
naics_11	0	0	0.07	3.64	0.35	1.93	0.49	0	0	0.01	0	0	0	1.76	0	1.07
naics_21	0	0	0	1.47	0	0	0	0	0	0	0	0	0	7.36	0	0
naics_22	0	0	1.68	0.08	0	0	0	2.77	0	0.05	0	0	0	0.02	0	0.09
naics_23	0.19	0.31	0.55	1.24	0.9	0.44	0.42	1.17	0.61	1.37	0.75	2.17	0.48	0.59	0.82	0.74
naics_31_33	0.06	22.78	0.57	1.2	2.26	0.76	0.09	1.16	0.02	0.77	0.42	1.96	1.11	0.25	0.2	0.22
naics_42	1.3	0.45	0.43	1.01	4.46	1.04	0.6	0.45	0.22	2.15	0.75	3.7	0.08	0.89	1.98	1.07
naics_44_45	3.13	0.05	0.34	1.19	0.85	1.96	2.13	0.83	0.15	1.24	1.36	0.78	0.2	1.92	1.82	2.07
naics_48_49	0.39	1.67	0.5	2.79	0.55	0.58	0.27	0.09	0.03	0.69	0.13	5.44	0.27	0.43	0.02	0.26
naics_51	0.4	0.55	1.19	1.04	4.22	6.59	0.66	1.02	0.1	0.05	0.74	0.49	0.04	0.62	0.59	0.09
naics_52	0.57	1.45	1	1.66	0.85	1.67	2.38	0.23	0.63	0.1	0.47	0.08	0.03	0.62	0.58	1.49
naics_53	0.61	0.25	1	0.94	1.39	0.21	1.14	1	0.16	0.76	1.85	0.43	0.54	1.18	0.39	0.42
naics_54	0.6	0.6	1	0.98	3.51	0.2	1.91	0.76	8.58	0.14	0.62	0.47	0.04	0.92	0.53	0.75
naics_55	0.05	3.17	1.16	1.49	3.09	2.49	1.74	0.37	0	0	0.07	0.64	0.67	0.2	0.34	0
naics_56	0.41	0.64	1.03	1.11	1.29	2.59	0.96	1.13	0.15	0.46	1.24	1.95	0.38	0.55	0.84	0.19
naics_61	0.04	0.06	1.2	0.77	0.15	0.44	0.69	0.21	0.29	2.09	0.34	0.1	0.34	0.36	1.38	1.05
naics_62	0.14	0.07	1.19	0.74	0.19	0.57	0.67	0.73	0.31	2.06	0.99	1.69	4.22	0.89	0.8	0.4
naics_71	0.83	0.04	1.23	0.67	1.37	0.31	0.8	0.17	0.05	0.36	0.19	0.15	0.77	0.56	0.12	1.44
naics_72	5.46	1.12	0.86	0.76	0.39	1.75	1.57	0.55	0.31	0.37	0.65	0.46	0.14	1.21	1.4	1.68
naics_81	0.62	0.8	0.89	0.73	0.51	0.78	0.77	0.27	0.2	0.79	1.24	1.58	0.35	1.74	1.19	1.08
naics_92	0.38	0.01	1.35	1.12	0.14	0	0.02	3.8	0	0.03	2.75	0	0	2.06	1.32	1.89

Table 1: Locations Quotient & HHI Results for Baltimore Region, 2019

THE DYNAMICS OF SPATIAL EMPLOYMENT STRUCTURE IN U.S.

Our results confirm that metropolitan areas in the U.S. tend toward polycentric urban form, but increasing the scope of analysis to compare form across the country also provides for more nuance. Descriptive statistics for the distribution of employment centers over time is shown in Table 4. Given the difficulty of interpreting succinctly such a large volume of results, we present summaries for both the country as a whole, and broken down by the U.S. economic subregions delineated by the Bureau of Economic Analysis. Table 3 shows the distribution of centers across these zones in 2019, and Figure 3 provides a graphic view. Maps of the centers identified in 2019 in MSAs where other prominent work has taken place are shown in Figure 2.

The maps in Figure 2 help clarify (1) how the identification procedure adapts to diverse contexts in vastly different geographic regions and urban densities and (2) how the underlying network analysis yields an irregular shape for the centers, which conforms to the transport infrastructure. Most centers span along major throughfares or sit at critical intersections. The relative size of centers is also variable across places, e.g. Houston (Figure 2b) has a large geographic core and several small centers that scatter outwards, Detroit (Figure 2a) has several moderately-sized centers scattered throughout the region, Atlanta, while the famously monocentric Chicago (Figure 2d) has a single dominating center with supporting appearances from small concentrations like Northwhestern University in Evanston.

Table 2: Top 25 Most Nucleated MSAs in 2019

name	count
Miami-Fort Lauderdale-Pompano Beach, FL	23
Detroit-Warren-Dearborn, MI	21
Los Angeles-Long Beach-Anaheim, CA	21
Houston-The Woodlands-Sugar Land, TX	21
Fort Smith, AR-OK	20
Riverside-San Bernardino-Ontario, CA	20
Dallas-Fort Worth-Arlington, TX	18
Baltimore-Columbia-Towson, MD	16
Charlotte-Concord-Gastonia, NC-SC	16
Washington-Arlington-Alexandria, DC-VA-MD-WV	15
Virginia Beach-Norfolk-Newport News, VA-NC	14
St. Louis, MO-IL	14
San Diego-Chula Vista-Carlsbad, CA	14
Jacksonville, FL	13
Atlanta-Sandy Springs-Alpharetta, GA	13

Continued on next page

Table 2: Top 25 Most Nucleated MSAs in 2019

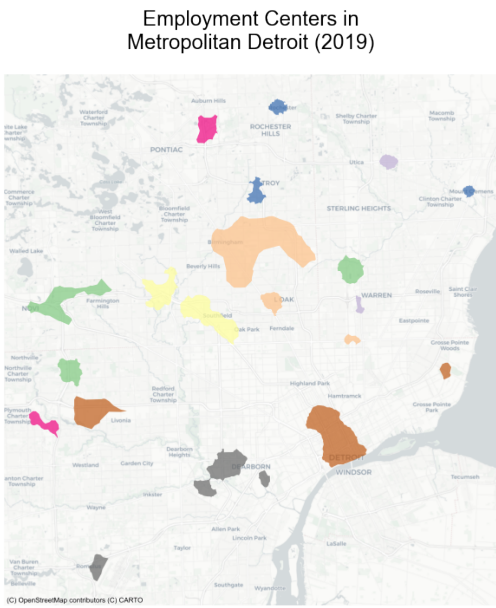
name	count
Phoenix-Mesa-Chandler, AZ	13
Minneapolis-St. Paul-Bloomington, MN-WI	13
Cincinnati, OH-KY-IN	12
Tampa-St. Petersburg-Clearwater, FL	12
Bakersfield, CA	11
Kingsport-Bristol, TN-VA	11
Grand Rapids-Kentwood, MI	11
Raleigh-Cary, NC	11
Orlando-Kissimmee-Sanford, FL	10
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	10

At a national scale, the number of employment centers follows a power distribution (Figure 3a). As Table 4 shows, the median MSA in the country is polycentric with approximately four employment centers. What this statistic masks, however, is that the *modal* MSA is monocentric. Fully 102 of the 379 study areas (~27%) have only a single employment center according to the method developed in this paper. The 'most sprawling' MSAs are shown in Table 2, which lists the top 25 metropolitan regions with the most employment centers. Some of the largest regions, like New York, Chicago, and San Francisco have only a few centers. This latter finding is more in keeping with results from European cities (Ahlfeldt & Wendland, 2013).

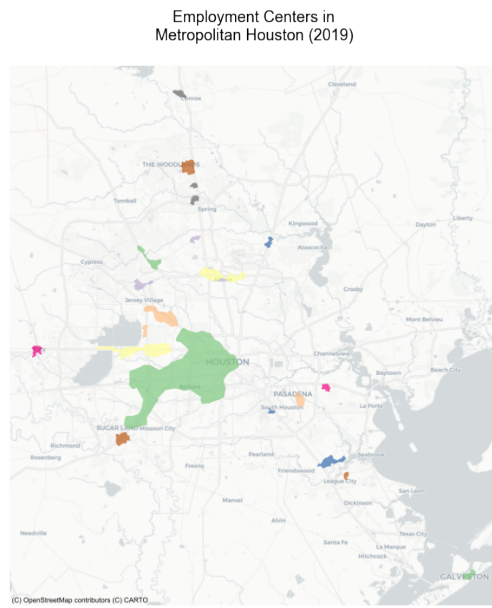
Table 3: Descriptive Statistcs by BEA Region

	count	mean	std	min	Q1	Q2	Q3	max
bea_region								
Far West	49.0	4.53	4.38	1.0	2.0	3.0	6.0	21.0
Great Lakes	58.0	3.26	3.48	1.0	1.0	2.0	4.0	21.0
Mideast	41.0	3.93	3.41	1.0	2.0	3.0	4.0	16.0
New England	15.0	4.73	2.52	1.0	2.5	4.0	6.5	9.0
Plains	33.0	3.06	3.31	1.0	1.0	2.0	4.0	14.0
Rocky Mountain	22.0	2.18	1.84	1.0	1.0	1.0	3.0	8.0
Southeast	112.0	4.50	3.98	1.0	2.0	3.0	6.0	23.0
Southwest	39.0	4.10	4.60	1.0	1.0	2.0	5.0	21.0

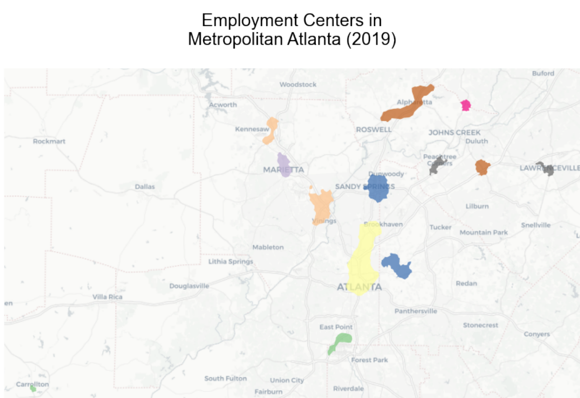
Further, as shown by Figure 3, the distribution is heterogeneous across zones of the country. New England has the highest median number of centers, but the second-smallest maximum, and a very small variance (shown in Figure 3c). The Rocky Mountain region has the smallest number of centers



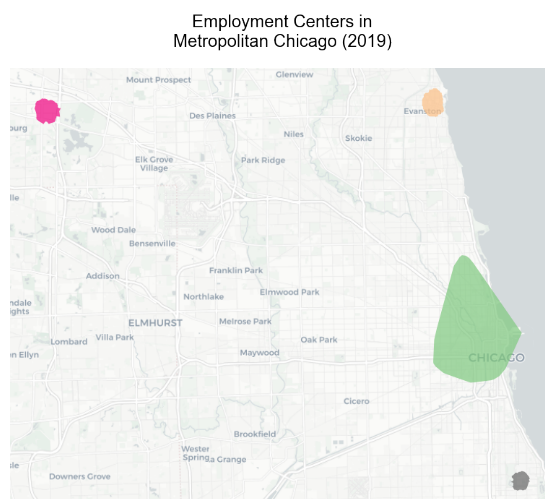
(a) Detroit Centers 2019



(b) Houston Centers 2019



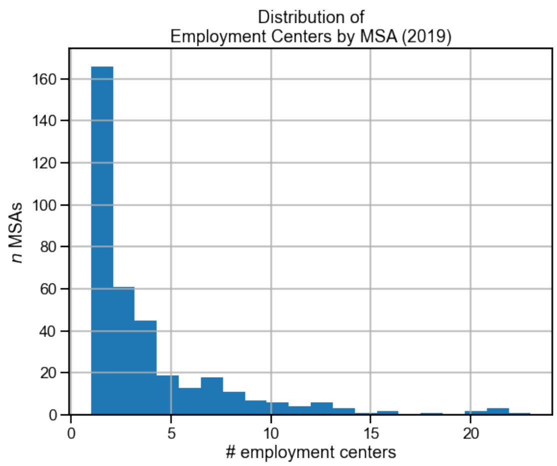
(c) Atlanta Centers 2019



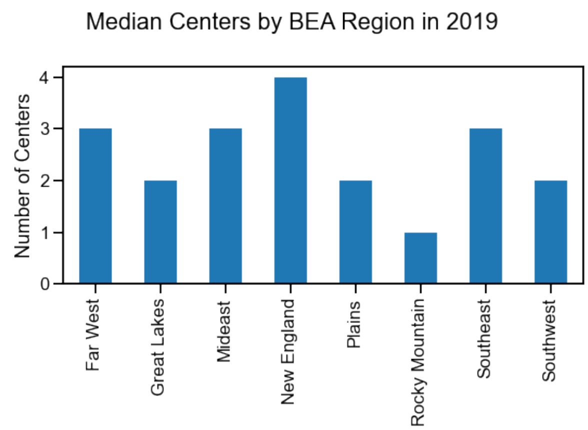
(d) Chicago Centers 2019

Figure 2: Employment Centers in Selected Metropolitan Regions 2019

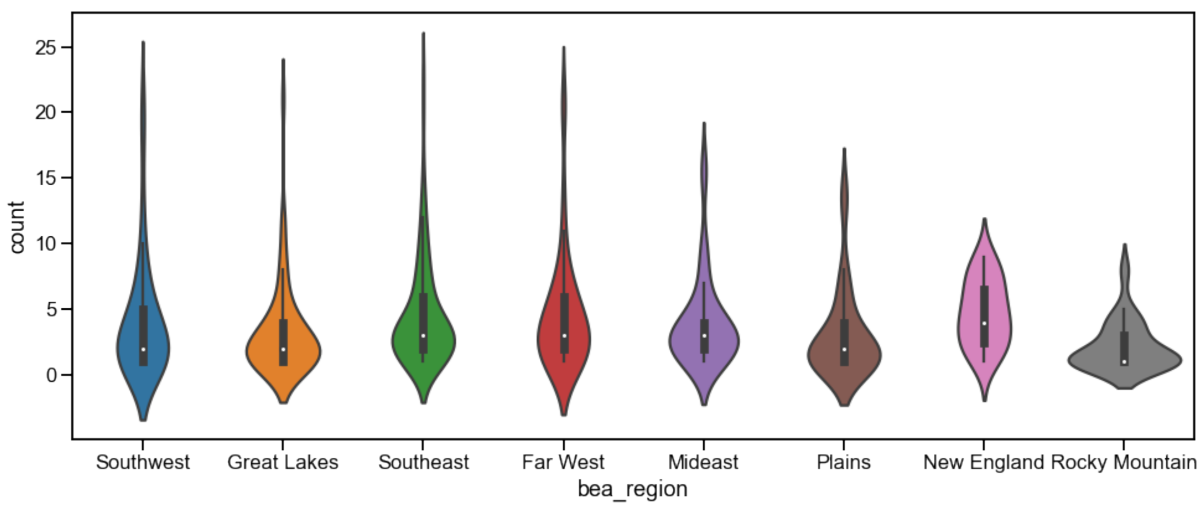
(the median MSA in the region is monocentric) and the smallest variance. The Southwest has a small median but the largest variance. These variations may be due to market forces, land use regulations, or topographical constraints, among other possibilities.



(a) Employment Center Histogram 2019



(b) Employment Centers by BEA Region, 2019

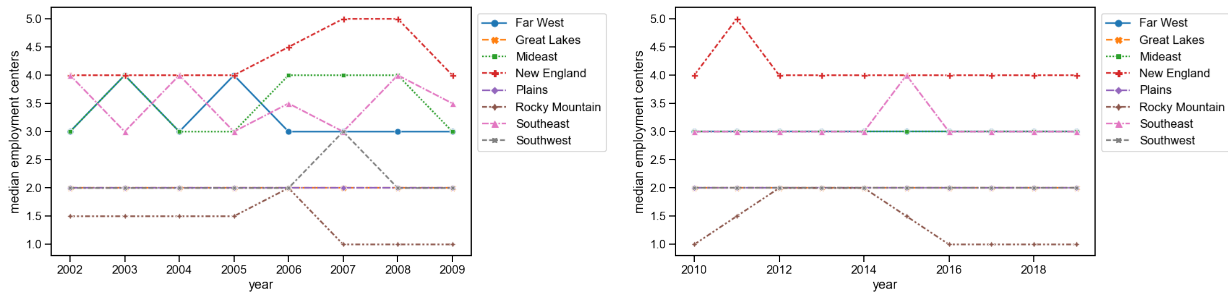


(c) Violin Plot of Center Distributions by BEA Region, 2019

Figure 3: Number of Employment Centers in 2019

Temporal Trends

The total number of employment centers in each MSA is typically quite stable over time, as is the variance and mean number of centers per metropolitan region. One possible explanation for the difference between these results and those from other recent prior work documenting a trend toward decentralization (Hipp et al., 2021) may be growth in subcenters along the perimeter of each MSA, not simply a trend toward generalized dispersal. Another complementary explanation may be the growth of smaller monocentric regions into their more mature polycentric form.



(a) Median Employment Centers by BEA Region 2002-2009 (b) Median Employment Centers by BEA Region 2010-2019

Figure 4: Employment Centers in U.S. Metro Areas 2002-2019

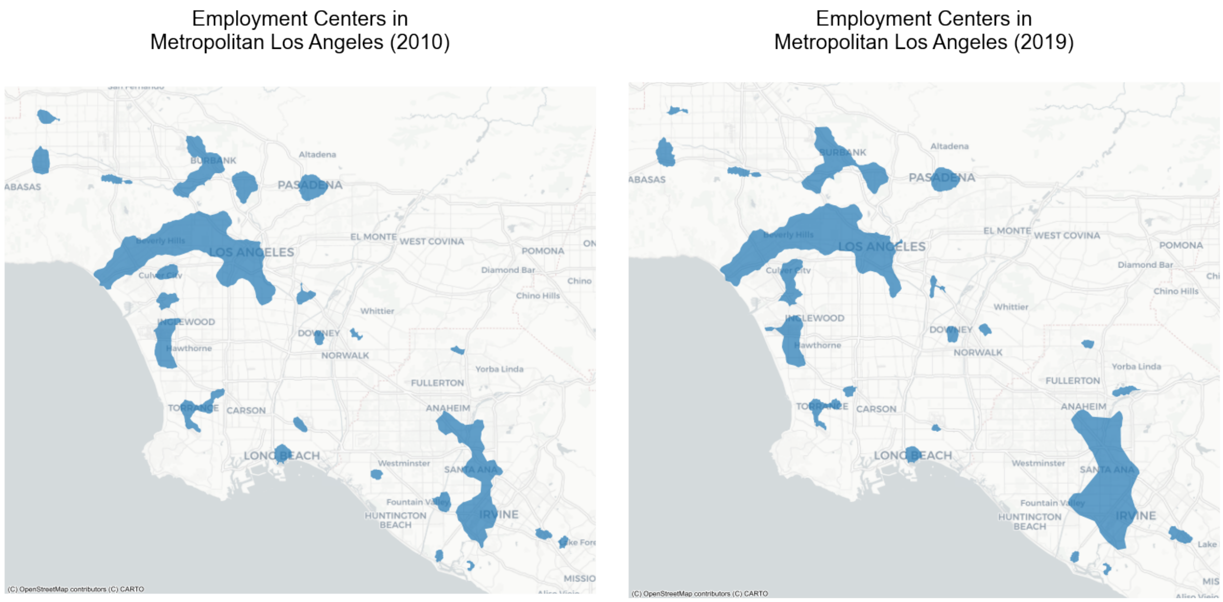
The graphs in Figure 4 show the average number of employment centers in all metropolitan regions in the U.S. from 2002 to 2019, with Figure 4a showing the trend from 2002 to 2009 and Figure 4b showing the trend from 2010 onward. The obvious feature distinguishing between these two is the sharp decline in job centers in 2010, which occurs because prior to 2010, the LEHD data did not include federal employment in its tabulations. As such, the entire timeseries is not interpretable as a single continuous entity, but we believe there is still valuable information in understanding the employment layout of urban areas, even without considering the federal government, so we include those time periods in our analysis as well. One thing this graph makes clear is the centralizing role that federal employment has at the national scale. In other words, when federal employment is omitted in the years prior to 2010, our method discovers more employment centers on average, suggesting that federal employment either locates within existing centers or helps foster agglomeration economies, generating centers of its own.

year	count	mean	std	min	25%	50%	75%	max
2010	1438	4.481	4.767	1	1	3	5	36
2011	1436	4.117	3.806	1	1	3	5	26
2012	1449	4.217	3.926	1	1	3	5	24
2013	1453	4.273	4.009	1	1	3	5	25
2014	1463	4.349	4.136	1	1	3	6	24
2015	1472	4.254	3.992	1	1	3	5	26
2016	1476	4.334	4.062	1	1	3	6	25
2017	1468	4.233	3.922	1	1	3	5	25
2018	1471	4.244	3.899	1	1	3	6	24
2019	1453	4.335	4.048	1	1	3	6	24

Table 4: Descriptive Statistics for Employment Centers, 2010-2019

Spatial Trends

We also examine the spatial configuration of employment centers with respect to births, deaths, mergers and splits over time. A graphic example of these processes is shown in Figure 5, which plots employment center maps for the Los Angeles region in 2010 (Figure 5a) and 2019 (Figure 5b). Here, the centers in Burbank in the North merge into a single center in 2019, and similarly the center at Santa Ana in the Southeast grows in size and absorbs (merges with) another center to its immediate left.



(a) L.A. Centers 2010

(b) L.A. Centers 2019

Figure 5: Employment Centers in Southern California in 2010 and 2019

To quantify these relationships, we adopt a graph-theoretic perspective to examine the configuration inside each metropolitan area between pairs of successive years. We use the python package `networkX`, and treat the employment centers in each time period as a set of nodes that share an edge when they intersect topologically between the two time periods.

Figure 6 displays the graph for the employment centers in the San Diego CBSA. Each node in the graph represents an employment center at a moment in time, with the colorbar indicating the year (0=2010, 9=2019). Each path is formed by a sequence of edges between two centers that overlap in consecutive years. This representation affords the visual identification of several aspects of center dynamics. Ephemeral centers, such as the two singleton nodes in the northwestern portion of the graph, come into exisence and disappear in the same year.[7] Less extreme are paths with lengths greater than 1 but which do not span the entire period. These consist of centers that may have been in existence at the beginning of our sample, but exit before the final year. Alternatively, a center may have come into existence after the first period and continued to exist for multiple years, possibly to the end of the sample.

Other dimensions of the center dynamics include coalesence and splitting. The former obtains when two centers who existed but did not overlap in a given period, grow to intersect in the next pe-

[7] The layout of the graph in Figure 6 is based on a spectral embedding of the graph. The locations of the nodes are not relative to the geographical context of the area.

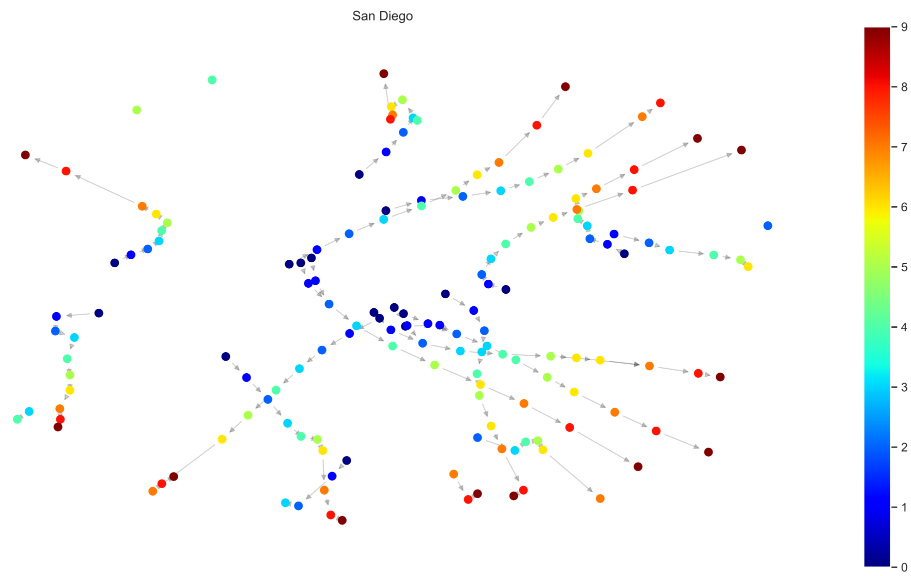


Figure 6: San Diego Employment Center Graph

Table 5: Employment Center Graph Summary Measures.

	coalescence	b_persistence	o_persistence	nc
mean	0.12	0.74	0.84	38.43
std	0.25	0.27	0.25	36.77
min	0.00	0.00	0.00	9.00
25%	0.00	0.56	0.73	13.00
50%	0.00	0.78	1.00	27.50
75%	0.08	1.00	1.00	47.00
max	1.00	1.00	1.00	247.00

riod, thus forming a single center from a pair of parent centers. Coalesence would thus reduce the number of centers between the two periods. Splitting occurs when a center breaks off into two, or more, centers in the next period. This may, or may not, increase the size of the area contained in the children centers relative to the ancestor center. Splitting does increase the number of centers in the area.

In addition to the visual representation of the center dynamics, the adoption of a graph-theoretic perspective provides for a number of summary measures for the evolution of the region's spatial employment centers. We focus on two persistence measures, the first is the proportion of center births that survive, irrespective of year of birth ($b\_persistence$). The second persistence measure considers only the centers that existed in the first year of the sample ($o\_persistence$).

Table 5 reports the summary measures for the employment center graphs from our regions. We find that coalesence is rare, with three quarters of the areas experiencing a coalescence rate of 0.08 or less. This suggests that center growth is dominated by the expansion of individual centers rather than the merging of previous centers. The second finding to note is that the median persistence rate is 0.78. That is, once a center is born, the tendency is to survive to the end of the sample. Interestingly, the persistence rate is higher for centers that were in existence at the beginning of our sample, relative to centers that emerge later in the sample.

The relationship between these different persistence measures is further examined in Figure 10a in the Appendix. The majority of the areas expereince lower birth persistence overall relative to the persistence of the original centers, with more of the observations falling below the red diagonal line, and the slope of the regression line being less than 1. Figure 10b shows a positive association between birth persistence and coalescense. Thus, although coalesence is relatively uncommon, when it occurs it fosters stronger cluster persistence. That is, when centers merge, they are even more likely to persist.

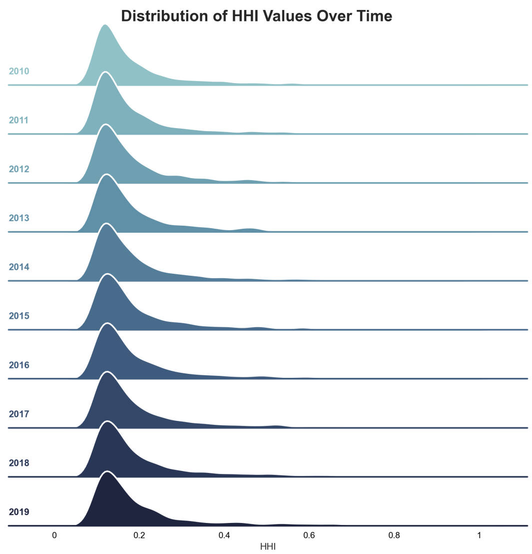
From an economic development perspective, these results may suggest that when an employment center achieves a critical mass inside the region, it provides a spatial anchor for jobs inside the region. When employment centers are located near one another (e.g., along major transport corridors, as many of our centers are) then the variation in total employment at each center causes mergers or splits in fluctuating time periods, leading to variation in the number of centers we discover over time. Because infill development will likely lead to centers merging together, this could bolster the center's staying power even further.

THE COMPOSITION OF METROPOLITAN EMPLOYMENT CENTERS

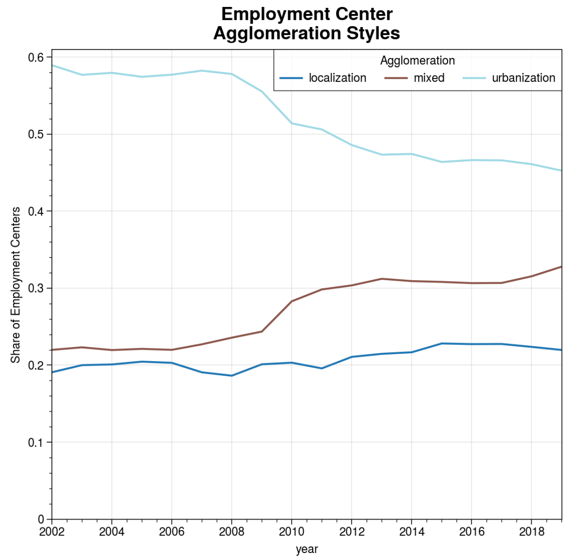
In addition to becoming more polycentric, we find that spatial structure in metropolitan regions is increasingly characterized by more localized economies. There is a trend away from employment centers characterized by urbanization agglomeration, which contain multiple types of jobs in favor of multiple subcenters that specialize in certain industries. We define a typology of agglomeration styles based on the observed HHI values at each subcenter. Following the guidance of Horizontal Merger Guidelines provided by the U.S. Department of Justice[8], we use HHI values to divide employment centers into categories based the degree of 'monopolistic' control by any single industry. We map these rules onto three types of agglomeration economies, with an urbanization economy defined as a center with $HHI < 0.15$, a mixed economy as a center with $0.15 \leq HHI < 0.25$, and a localized economy as a center with $HHI \geq 0.25$

The two graphs in Figure 7 show the change in the industrial composition of employment centers across the United States over time. The graph in Figure 7a shows the distribution of HHI values across centers for each time period, and the graph in Figure 7b shows the relative share of each agglomeration type over time. There is a modest increase in localization, a marked increase in mixed economies, and a dramatic decline in urbanization. Although there tends to be a jump at 2010, when federal employment enters the picture, the trends are still present before and after the critical year.

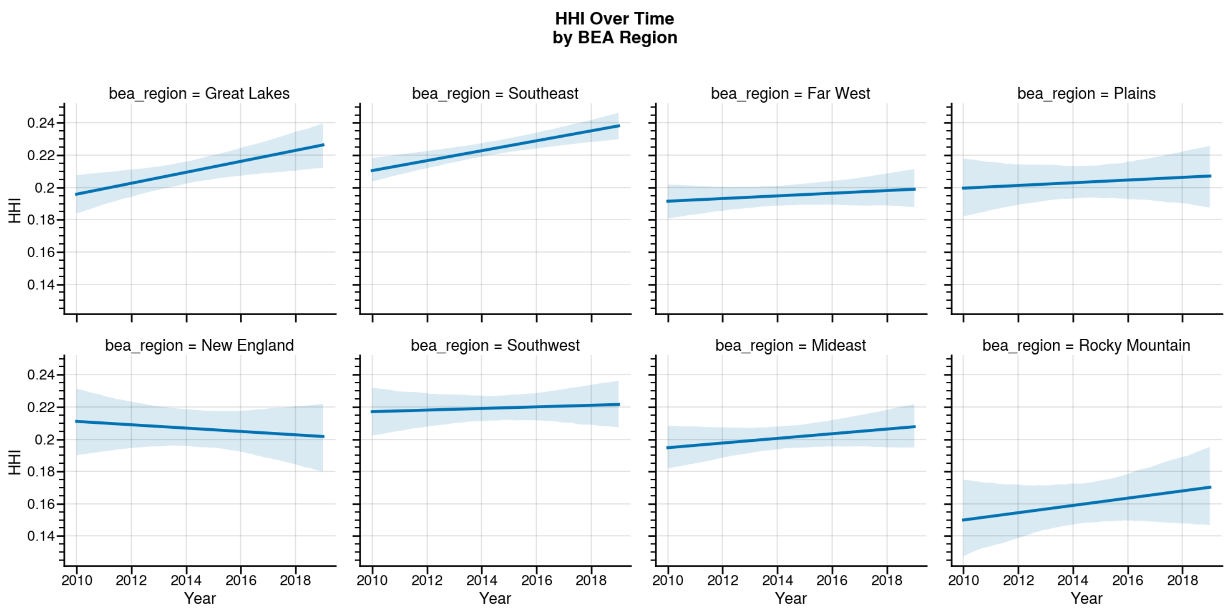
[8] https://www.justice.gov/atr/horizontal-merger-guidelines-08192010



(a) Distribution of HHI Values by Year



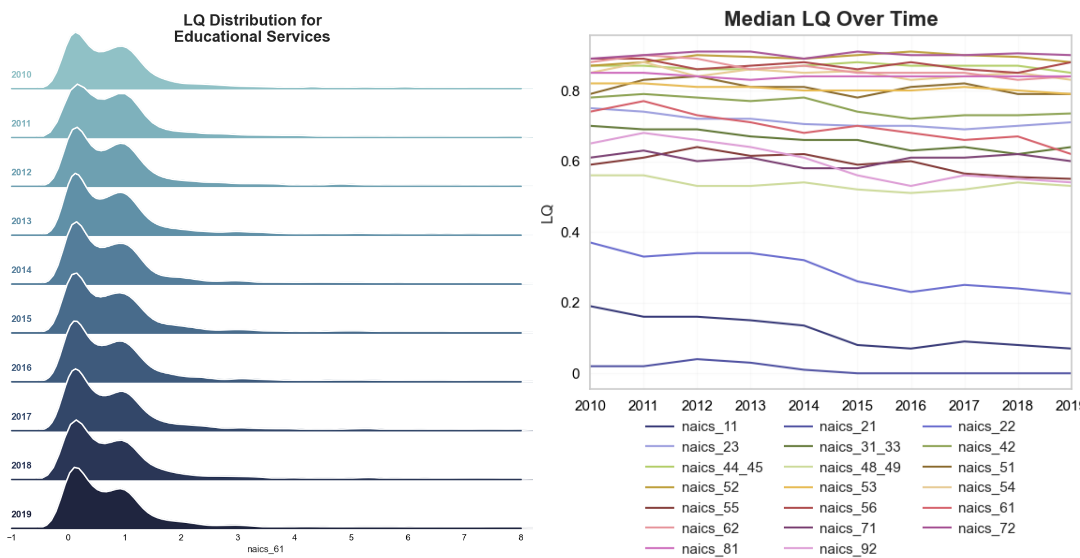
(b) Agglomeration Styles Over Time



(c) HHI Over Time by Region

Figure 7: Employment Center Agglomeration Over Time

In the continuous view of the data, shown in Figure 7a, this trend is displayed by the modest flattening of the largest peak, and the increasing density around the 0.2 level as time moves forward. In the discrete view of the data shown in Figure 7b, the trend is depicted easily by the falling line representing urbanization agglomeration. This trend is not uniform across the country, as Figure 7c shows that some regions, such as the Southwest and Great Lakes are developing more specialized subcenters faster than other regions, as shown by the steeper slope over time.



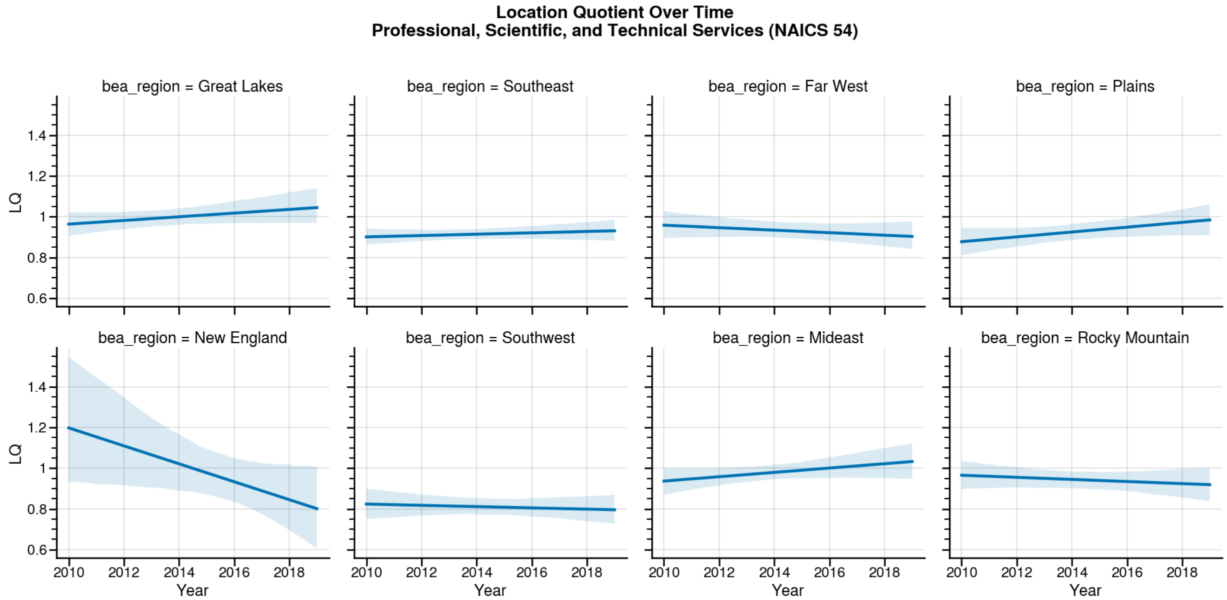
(a) LQs for NAICS 61

(b) Location Quotient by Industry Category Over Time

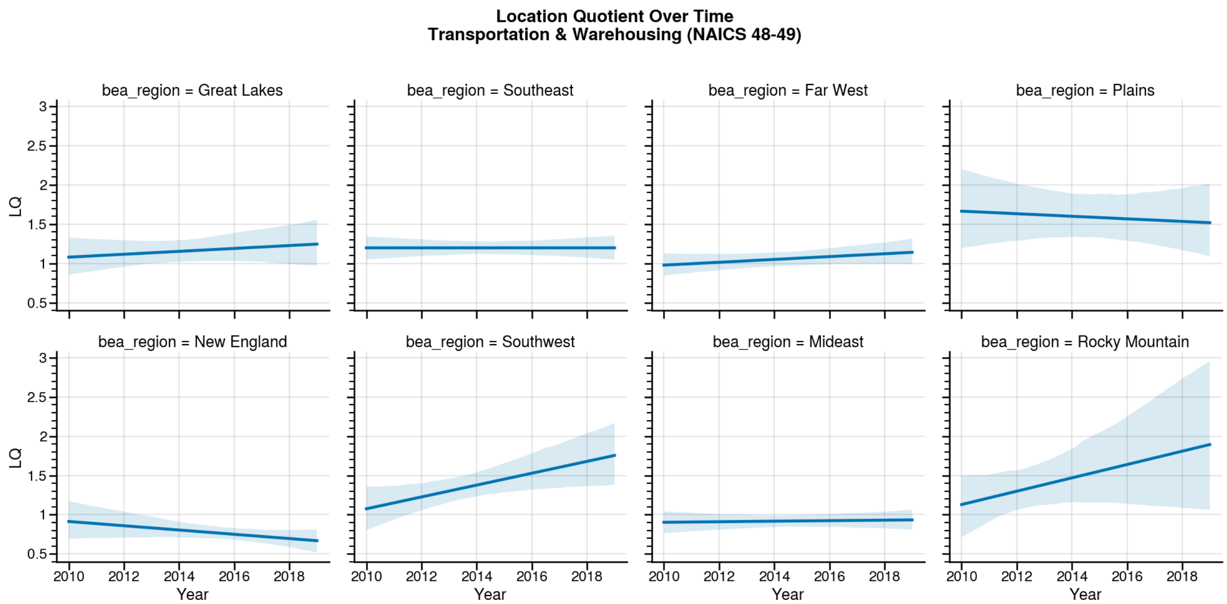
Figure 8: Location Quotients Over Time

Examining the distribution of Location Quotients across centers over time, we find that certain industries are represented in employment centers more often. According to the median LQ in 2019, the most prevalent industries in centers, in descending order, include NAICS categories 72 (Accommodation and Food), 52 (Finance & Insurance), 56 (Management of Companies and Enterprises), 44-45 (Retail), 62 (Healthcare and Social Services), 81 (Other Services) and 54 (Professional, Scientific, and Technical), all of which have a median LQ greater than 0.8 during the entire timeseries, which is shown in Figure 8b.

Distributions of each industry over time are provided in the technical appendix, and demonstrate that there is considerable variation in the employment center composition across the country. Many LQ distributions are bimodal, suggesting that these industries are either dominant players in their subcenter, or do not co-locate with other industries. Examples include industries such as healthcare, or education (shown in Figure 8a), which typically operate on large campuses such as hospitals or universities, both of which require considerable labor forces. The very long righthand tails on these distributions show that there is at least one subcenter in the country each year that specialized in each industry (note that the plots are artificially bounded at an upper limit of 8.0). The seven industries listed above help provide further evidence of knowledge spillovers, as their high LQ values show that



(a) Scientific & Professional by Region



(b) Transport & Warehousing by Region

Figure 9: Location Quotients By Region

these jobs appear in most centers, and Figure 8b shows little evidence of these trends changing over time. By contrast, NAICS 11 (Agriculture), 21 (Mining) and 22 (Utilities) all have LQs lower than 0.4 and declining over time.

Our results also show that the composition of metropolitan employment centers varies by region. Figure 9 displays the linear relationship between LQ measures in the Professional and Scientific (Figure 9a) Transportation and Warehousing (Figure 9b) sectors and time, with separate plots for each of the eight regions defined by the Bureau of Economic Analysis. The figure shows the rising prominence of transportation and warehousing in the Southwest and Rocky Mountain regions, as well as the increasing importance of the Scientific and Professional sector in the Mideast. Taken together, these results could begin to help shape a regional economic development strategy focused on leveraging the existing industrial makeup in a given place, while leveraging spillovers from targeted polycentric development.

DISCUSSION

Our method for identifying employment centers works well at the metropolitan scale for both large and small metros, yielding comparable datasets over time and across space. Nevertheless, we rely on a handful of global tuning parameters that perform well in our analysis but could be subject to greater research. First, we compute our accessibility measure A_i using a two-kilometer threshold, which serves as a spatial smoother and affects the granularity of the resulting employment centers. With smaller values, employment centers are more localized in space, whereas with larger distances, otherwise distinct employment centers get merged together. In this analysis we choose a single threshold and hold it constant to facilitate temporal comparisons. The two kilometer threshold comports with recent work on the geographic scale of knowledge spillovers (Barbieri et al., 2022), though it may be valuable to examine other distances as well. Setting a larger distance threshold would likely reduce the number of centers overall (as more mergers are likely), whereas reducing the smoothing by lowering the distance would result in a greater number of centers.

Second, we use an employment volume criterion that each node within an employment center must be in the top quintile of accessibility values in the region. This seems like a reasonable choice given that employment and transportation trips have long been shown to be distributed according to a Power law, though the twenty percent threshold is nonetheless arbitrary. Increasing this threshold (i.e., raising it to 10%) would reduce the number of candidate sites eligible to be considered employment centers, and would likely reduce the overall number of centers in each MSA.

To capture spatial concentration, our method relies on the G_i^* local spatial statistic, and we use a spatial weights matrix of the 100 nearest neighbors for each street intersection. We make this choice for computational convenience and because it satisfies the recommendation by Ord & Getis (1995) that the weights matrix for the statistic be neither too large nor too small. In a metropolitan region with hundreds of thousands of nodes, this seems like another reasonable choice, but other weights matrices could be considered. Finally, we require that all employment centers be greater than one kilometer in size, to avoid small ephemeral centers that may occur from a single employer, but this choice, too, could be examined. Because we implement all our methods in open-source tooling and release them as a software package with this paper, other researchers in the field can test all these assumptions easily and thoroughly.

The LEHD data used in this study provide a range of benefits including very high geographic resolution and annual tabulations of employment by industry, which allow us to create detailed measures over fine time scales. But these data, which are built from private unemployment insurance records (and combined with other federal sources later) also have drawbacks. Not all employers report employment perfectly, and the address reported may be different from the actual working address. In addition, for some companies, all employment is reported at the headquarter location or at a P.O. Box., rather than the local office where work takes place, Also, the data exclude military employment and thus miss a large picture of the labor force in some regions.

In the results presented above we develop a new technique for identifying employment centers and provide a descriptive and exploratory analysis of the revealed centers over time. With centers identified, this work scratches the surface of additional analyses to address Agarwal et al. (2012)'s call for a better understanding of the relationship between polycentricity and urban agglomeration. In future research, we plan to combine our center identification approach with models that help explain why centers emerge over time and which industries tend to locate together and benefit the most from doing so.

Additional work could also compare the employment centers discovered using our method to those discovered by Arribas-Bel et al. (2021) who use a modification of the DBSCAN algorithm applied to building data, or Baragwanath et al. (2021) who use satellite imagery and light intensity data to define centers. Since those techniques can also be applied to publicly-available data and combined with the LEHD data used here, it would be possible to compare the robustness of our composition results to other delimitation methods. Finally, it would also be useful to conduct a case study on a single or a small set of metropolitan areas to examine the evolving patterns of employment centers in greater detail. At the national scale, our results reveal new insights about macro trends, but do little to describe how an economy in the northeast compares to a similarly sized metropolitan region in the southwest. Further work could exploit these interregional comparisons for a more complete picture.

CONCLUSION

In this paper we develop a new technique for identifying metropolitan employment centers. We apply our method to 18 years of data from the Bureau of Labor Statistics in over 350 Core Based Statistical Areas to examine the configuration and makeup of employment centers in the United States. Our results demonstrate that our technique is well suited to identify employment centers at regional scales both large and small. The methods we employ rely on theoretically-sound parameters which are relative to the study region and need not be calibrated specifically to account for size and density characteristics of a particular location. We use the metropolitan transportation network to encode urban space and ensure that the resulting employment centers represent the underlying structure accurately. By leveraging computational spatial statistics, we avoid the need to specify local parameters to suit different regions.

Our analysis blends classic concepts from regional science including locations quotients, indices of industrial composition, and gravity-based accessibility measures, along with G_i^* a foundational contribution in local spatial statistics, and fast algorithms from computational geometry and data science. In doing so, we devise a technique that builds a bridge across disparate disciplines to provide policy-relevant insight into urban spatial structure. Our empirical analysis spans nearly 20 years of high-resolution data for every Core-Based Statistical Area in the United States (for which LEHD data is available) comprising a total set of observations exceeding several billion. The entire analysis is conducted using open-source software on commercial hardware, and is reproducible from start to finish, with included data and final results for the entire country.

Similar to prior work, we confirm the dominant pattern of polycentricity in modern urban spatial structure. Unlike some recent work, however, we do not find evidence of continued dispersal. While some employment sectors may be decentralizing, our results show that agglomeration economies are still a strong factor shaping the location decisions of employment firms, although urbanization seems to be falling as the dominant mode for leveraging scale economies. Instead, polycentric development appears to be driven by moderately-specialized employment centers, particularly in industries such as Accommodation and Food, Finance and Insurance, and Management of Companies. The bimodal distributions and long tails associated with LQ values for knowledge industries such as Information, Education, and Healthcare also support prior work on the spillovers provided by these sectors.

The polycentric structure of employment centers inside most metropolitan regions generally remains stable over time, even in the presence of job growth (Huang et al., 2021), the median number of employment centers for metropolitan regions in the United States is three, and the mean is slightly above four. This pattern has remained essentially unchanged for a decade. Like prior work, we confirm that the number of employment centers follows a Power distribution and that the rank-size rule holds, even at an intra-metropolitan level (distributions shown in the appendix).

These results show important relevance for public policy and urban planning efforts to promote sustainable, equitable and prosperous economic growth. First, the stability of employment centers shows that geographic targeting may serve as a sound economic development strategy (Derudder et al., 2022; Knaap et al., 2016). The graph-theoretic measures of persistence we develop provide strong evidence of the temporal staying-power of employment centers. The dominant tendency is for centers to survive rather than exit during our sample. When centers merge, it increases the likelihood of center survival even further. The persistence of employment centers over time suggests that urban planning efforts that encourage infill development along transport corridors linking existing centers are likely to be successful in curbing urban sprawl.

Second, we show that polycentrism is accompanied by increasing industrial specialization within the metropolitan area. Unique to the present study, our examination of the composition of each employment center shows a clear trend over the study period. This reveals that U.S. metropolitan areas are moving increasingly toward a greater number of specialized subcenters and away from a monolithic urbanization economy at the core. On average, the HHI values for employment centers are increasing over time, in a continuing pattern of increased localization. When centers merge, they often result in greater urbanization, but their staying power increases even further. Combined with the results described above, this is likely to increase the efficacy of infill development that links centers.

Finally, we document that industrial specialization is not happening uniformly throughout the country, and that employment centers in different regions of the country show evidence of heterogeneous restructuring. For example, the Rocky Mountain Region and the Southwest are becoming increasingly specialized in transportation and warehousing, whereas the Information sector is growing faster in the Mideast, and specialization in the professional/scientific sector is growing fastest in the Great Lakes region. Using both these emerging trends and further analyses leveraging the techniques we develop here, the dynamics of spatial structure of U.S. metropolitan areas can yield important insight into regional economic development.

REFERENCES

Abdulhafedh, A. (2017). A Novel Hybrid Method for Measuring the Spatial Autocorrelation of Vehicular Crashes: Combining Moran's Index and Getis-Ord Gi* Statistic. *Open Journal of Civil Engineering*, *07*(02), 208–221. https://doi.org/10.4236/ojce.2017.72013

Agarwal, A., Giuliano, G., & Redfearn, C. L. (2012). Strangers in our midst: The usefulness of exploring polycentricity. *Annals of Regional Science*, *48*(2), 433–450. https://doi.org/10.1007/s00168-012-0497-1

Ahlfeldt, G. (2011). If Alonso Was Right: Modeling Accessibility and Explaining the Residential Land Gradient. *Journal of Regional Science*, *51*(2), 318–338. https://doi.org/10.1111/j.1467-9787.2010.00694.x

Ahlfeldt, G. M., & Wendland, N. (2013). How polycentric is a monocentric city? Centers, spillovers and hysteresis. *Journal of Economic Geography*, *13*(1), 53–83. https://doi.org/10.1093/jeg/lbs013

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press.

Anas, A. (1984). Discrete Choice Theory and the General Equilibrium of Employment, Housing, and Travel Networks in a Lowry-Type Model of the Urban Economy. *Environment and Planning A: Economy and Space*, *16*(11), 1489–1502. https://doi.org/10.1068/a161489

Anas, A. (1985). The combined equilibrium of travel networks and residential location markets. *Regional Science and Urban Economics*, *15*(1), 1–21. https://doi.org/10.1016/0166-0462(85)90029-8

Anas, A. (1990). Taste heterogeneity and urban spatial structure: The logit model and monocentric theory reconciled. *Journal of Urban Economics*, *28*(3), 318–335. https://doi.org/10.1016/0094-1190(90)90031-H

Anas, A., Arnott, R., & Small, K. A. (1998). Urban Spatial Structure. *Journal of Economic Literature*, *36*(3), 1426–1464. http://links.jstor.org/sici?sici=0022-0515%28199809%2936%3A3%3C1426%3AUSS%3E2.0.CO%3B2-Z

Anderson, N. B., & Bogart, W. T. (2001). The Structure of Sprawl: Identifying and Characterizing Employment Centers in Polycentric Metropolitan Areas. *American Journal of Economics and Sociology*, *60*(1), 147–169. https://doi.org/10.1111/1536-7150.00058

Arribas-Bel, D., Garcia-López, M.-À., & Viladecans-Marsal, E. (2021). Building(s and) cities: Delineating urban areas with a machine learning algorithm. *Journal of Urban Economics*, *125*, 103217. https://doi.org/10.1016/j.jue.2019.103217

Arribas-bel, D., & Sanz-gracia, F. (2014). The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990 , 2000 and 2010. *Urban Geography*, *35*(7), 980–997. https:

//doi.org/10.1080/02723638.2014.940693
Arribas-Bel, D., & Schmidt, C. R. (2013). Self-Organizing Maps and the US Urban Spatial Structure. *Environment and Planning B: Planning and Design*, *40*(2), 362–371. https://doi.org/10.1068/b37014
Ban, J., Arnott, R., & Macdonald, J. (2017). Identifying Employment Subcenters: The Method of Exponentially Declining Cutoffs. *Land*, *6*(1), 17. https://doi.org/10.3390/land6010017
Baragwanath, K., Goldblatt, R., Hanson, G., & Khandelwal, A. K. (2021). Detecting urban markets with satellite imagery: An application to India. *Journal of Urban Economics*, *125*, 103173. https://doi.org/10.1016/j.jue.2019.05.004
Barbieri, N., Ramaciotti, L., & Rizzo, U. (2022). The relationship between R&D knowledge spillovers and employment entry. *The Annals of Regional Science*. https://doi.org/10.1007/s00168-022-01182-2
Bartosiewicz, B., & Marcinczak, S. (2022). Urban structure in transition: Evidence from Poland, 1983–2011. *Regional Studies*, *56*(1), 36–47. https://doi.org/10.1080/00343404.2021.1878125
Baumont, C., Ertur, C., & Gallo, J. (2004). Spatial Analysis of Employment and Population Density: The Case of the Agglomeration of Dijon 1999. *Geographical Analysis*, *36*(2), 146–176. https://doi.org/10.1111/j.1538-4632.2004.tb01130.x
Cervero, R., & Wu, K.-L. (1997). Polycentrism, Commuting, and Residential Location in the San Francisco Bay Area. *Environment and Planning A: Economy and Space*, *29*(5), 865–886. https://doi.org/10.1068/a290865
Cervero, R., & Wu, K.-L. (1998). Sub-centring and Commuting: Evidence from the San Francisco Bay Area, 1980-90. *Urban Studies*, *35*(7), 1059–1076. https://doi.org/10.1080/0042098984484
Craig, S. G., Kohlhase, J. E., & Perdue, A. W. (2016). Empirical polycentricity: The complex relationship between employment centers. *Journal of Regional Science*, *56*(1), 25–52. https://doi.org/10.1111/jors.12208
Davis, D. R., & Dingel, J. I. (2019). A Spatial Knowledge Economy. *American Economic Review*, *109*(1), 153–170. https://doi.org/10.1257/aer.20130249
Derudder, B., Meijers, E., Harrison, J., Hoyler, M., & Liu, X. (2022). Polycentric urban regions: Conceptualization, identification and implications. *Regional Studies*, *56*(1), 1–6. https://doi.org/10.1080/00343404.2021.1982134
Dissart, J. C. (2003). Regional economic diversity and regional economic stability: Research results and agenda. *International Regional Science Review*, *26*(4), 423–446. https://doi.org/10.1177/0160017603259083
Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063–2117). Elsevier. https://doi.org/10.1016/S1574-0080(04)80005-1
Duschl, M., Schimke, A., Brenner, T., & Luxen, D. (2014). Firm Growth and the Spatial Impact of Geolocated External Factors. *Jahrbücher Für Nationalökonomie Und Statistik*, *234*(2-3), 234–256. https://doi.org/10.1515/jbnst-2014-2-308
Eberts, R. W., & McMillen, D. P. (1999). Agglomeration economies and urban public infrastructure. In P. Nijkamp, E. S. Mills, & P. C. Cheshire (Eds.), *Handbook of Regional and Urban Economics* (Vol. 3, pp. 1455–1495). North-Holland. https://doi.org/10.1016/S1574-0080(99)80007-8
Edelsbrunner, H., Kirkpatrick, D., & Seidel, R. (1983). On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, *29*(4), 551–559. https://doi.org/10.1109/TIT.1983.1056714
Foti, F., Waddell, P., & Luxen, D. (2012). A Generalized Computational Framework for Accessibility : From the Pedestrian to the Metropolitan Scale. *4th Transportation Research Board Conference on Innovations in Travel Modeling (ITM)*, 1–14.
Garcia-López, M.-À., & Moreno-Monroy, A. I. (2018). Income segregation in monocentric and polycentric cities: Does urban form really matter? *Regional Science and Urban Economics*, *71*, 62–79. https://doi.org/10.1016/j.regsciurbeco.2018.05.003
Geisberger, R., Sanders, P., Schultes, D., & Vetter, C. (2012). Exact routing in large road networks using contraction hierarchies. *Transportation Science*, *46*(3), 388–404. https://doi.org/10.1287/trsc.1110.0401
Getis, A., & Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, *24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Giuliano, G. (1991). Is Jobs-Housing Balance a Transportation Issue? *Transportation Research Record*, *1305*, 305–312. https://www.youtube.com/watch?v=PCztXEfNJLM

Giuliano, G., Redfearn, C., Agarwal, A., & He, S. (2012). Network Accessibility and Employment Centres. *Urban Studies*, *49*(7), 77–95. https://doi.org/10.1177/0042098011411948

Giuliano, G., Redfearn, C., Agarwal, A., Li, C., & Zhuang, D. (2007). Employment concentrations in Los Angeles, 1980-2000. *Environment and Planning A*, *39*(12), 2935–2957. https://doi.org/10.1068/a393

Giuliano, G., & Small, K. A. (1991). Subcenters in the Los Angeles region. *Regional Science and Urban Economics*, *21*(2), 163–182. https://doi.org/10.1016/0166-0462(91)90032-I

Giuliano, G., & Small, K. A. (1993). Is the Journey to Work Explained by Urban Structure? *Urban Studies*, *30*(9), 1485–1500. https://doi.org/10.1080/00420989320081461

Gordon, P., & Richardson, H. W. (1996). Employment decentralization in US metropolitan areas: Is Los Angeles an outlier or the norm? *Environment and Planning A*, *28*(10), 1727–1743. https://doi.org/10.1068/a281727

Gordon, P., Richardson, H. W., & Wong, H. L. (1986). The Distribution of Population and Employment in a Polycentric City: The Case of Los Angeles. *Environment and Planning A: Economy and Space*, *18*(2), 161–173. https://doi.org/10.1068/a180161

Griffith, D. A. (1981). Evaluating the transformation from a monocentric to a polycentric city. *Professional Geographer*, *33*(2), 189–196. https://doi.org/10.1111/j.0033-0124.1981.00189.x

Guillain, R., Le Gallo, J., & Boiteux-Orain, C. (2006). Changes in Spatial and Sectoral Patterns of Employment in Ile-de-France, 1978-97. *Urban Studies*, *43*(11), 2075–2098. https://doi.org/10.1080/00420980600945203

Hajrasouliha, A. H., & Hamidi, S. (2017). The typology of the American metropolis: Monocentricity, polycentricity, or generalized dispersion? *Urban Geography*, *38*(3), 420–444. https://doi.org/10.1080/02723638.2016.1165386

Hansen, W. G. (1959a). Accessibility and Residential Growth. *Massachusetts Institute of Technology*.

Hansen, W. G. (1959b). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, *25*(2), 73–76. https://doi.org/10.1080/01944365908978307

Harris, J. L. (2020). Rethinking cluster evolution: Actors, institutional configurations, and new path development. *Progress in Human Geography*, 030913252092658. https://doi.org/10.1177/0309132520926587

Heider, B., Mast, J., Roth, D., Standfuß, I., Siedentop, S., & Taubenböck, H. (2022). Dynamics of intraurban employment geographies: A comparative study of U.S. And German metropolitan areas. *Journal of Urban Affairs*, *0*(0), 1–21. https://doi.org/10.1080/07352166.2022.2122833

Helsley, R. W., & Strange, W. C. (2007). Urban interactions and spatial structure. *Journal of Economic Geography*, *7*(2), 119–138. https://doi.org/10.1093/jeg/lbl027

Helsley, R. W., & Sullivan, A. M. (1991). Urban subcenter formation. *Regional Science and Urban Economics*, *21*(2), 255–275. https://doi.org/10.1016/0166-0462(91)90036-M

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics*, *53*(1), 1–28. https://doi.org/10.1016/S0094-1190(02)00505-3

Hipp, J. R., Kim, J. H., & Forthun, B. (2021). Proposing new measures of employment deconcentration and spatial dispersion across metropolitan areas in the US. *Papers in Regional Science*, *100*(3), 815–841. https://doi.org/10.1111/pirs.12593

Huang, D., Liu, T., Kong, F., & Guang, R. (2021). Employment centers change faster than expected: An integrated identification method and application to Beijing. *Cities*, *115*, 103224. https://doi.org/10.1016/j.cities.2021.103224

Isard, W. (1960). *Methods of regional analysis: An introduction to regional science*.

Isserman, A. M. (1995). The History, Status, and Future of Regional Science: An American Perspective. *International Regional Science Review*, *17*(3), 249–296. https://doi.org/10.1177/016001769501700301

Kane, K., Hipp, J. R., & Kim, J. H. (2018). Los Angeles employment concentration in the 21st century. *Urban Studies*, *55*(4), 844–869. https://doi.org/10.1177/0042098016678341

Kao, A. S., Getis, A., Brodine, S., & Burns, J. C. (2008). Spatial and temporal clustering of kawasaki syndrome cases. *Pediatric Infectious Disease Journal*, *27*(11), 981–985. https://doi.org/10.1097/INF.

0b013e31817acf4f
Kloosterman, R. C., & Musterd, S. (2001). The Polycentric Urban Region: Towards a Research Agenda. *Urban Studies*, *38*(4), 623–633. https://doi.org/10.1080/00420980120035259
Knaap, E., Ding, C., Niu, Y., & Mishra, S. (2016). Polycentrism as a sustainable development strategy: Empirical analysis from the state of Maryland. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *9*(1), 73–92. https://doi.org/10.1080/17549175.2015.1029509
Knaap, E., & Rey, S. (2023). Segregated by design? Street network topological structure and the measurement of urban segregation. *Environment and Planning B: Urban Analytics and City Science*, 23998083231197956. https://doi.org/10.1177/23998083231197956
Krehl, A. (2015). Urban spatial structure: An interaction between employment and built-up volumes. *Regional Studies, Regional Science*, *2*(1), 290–308. https://doi.org/10.1080/21681376.2015.1034293
Krehl, A., & Siedentop, S. (2019). Towards a typology of urban centers and subcenters – evidence from German city regions. *Urban Geography*, *40*(1), 58–82. https://doi.org/10.1080/02723638.2018.1500245
Krugman, P. (1999). The role of geography in development. *International Regional Science Review*, *22*(2), 142–161. https://doi.org/10.1177/016001799761012307
Levinson, D. M., & Kumar, A. (1994). The Rational Locator: Why Travel Times Have Remained Stable. *Journal of the American Planning Association*, *60*(3), 319–332. https://doi.org/10.1080/01944369408975590
Liu, C. H., Rosenthal, S. S., & Strange, W. C. (2020). Employment density and agglomeration economies in tall buildings. *Regional Science and Urban Economics*, *84*, 103555. https://doi.org/10.1016/j.regsciurbeco.2020.103555
Lynch, K., & Rodwin, L. (1958). A Theory of Urban Form. *Journal of the American Institute of Planners*, *24*(4), 201–214. https://doi.org/10.1080/01944365808978281
Manduca, R. (2020). The spatial structure of US metropolitan employment: New insights from administrative data. *Environment and Planning B: Urban Analytics and City Science*, *0*(0), 1–16. https://doi.org/10.1177/2399808320934821
Maoh, H. F., Koronios, M., & Kanaroglou, P. S. (2010). Exploring the land development process and its impact on urban form in Hamilton, Ontario: Land development process and urban form. *The Canadian Geographer / Le Géographe Canadien*, *54*(1), 68–86. https://doi.org/10.1111/j.1541-0064.2009.00303.x
Maoh, H., & Kanaroglou, P. (2007). Business establishment mobility behavior in urban areas: A microanalytical model for the City of Hamilton in Ontario, Canada. *Journal of Geographical Systems*, *9*(3), 229–252. https://doi.org/10.1007/s10109-007-0043-3
Markusen, A., & Porter, M. E. (1996). Competitive {Advantage}, {Agglomeration} {Economies}, and {Regional} {Policy}. *International Regional Science Review*, *19*, 85–90. https://doi.org/10.1177/016001769601900208
Marshall, A. (1920). *Principles of Economics (8th ed.)*. The Macmillan Press. https://oll.libertyfund.org/titles/marshall-principles-of-economics-8th-ed
McDonald, J. F. (1987). The identification of urban employment subcenters. *Journal of Urban Economics*, *21*(2), 242–258. https://doi.org/10.1016/0094-1190(87)90017-9
McMillen, D. P. (2001). Nonparametric Employment Subcenter Identification. *Journal of Urban Economics*, *50*(3), 448–473. https://doi.org/10.1006/juec.2001.2228
McMillen, D. P. (2003). Identifying Sub-centres Using Contiguity Matrices. *Urban Studies*, *40*(1), 57–69. https://doi.org/10.1080/00420980220080161
McMillen, D. P., & McDonald, J. F. (1997). A nonparametric analysis of employment density in a polycentric city. *Journal of Regional Science*, *37*(4), 591–612. https://doi.org/10.1111/0022-4146.00071
McMillen, D. P., & Smith, S. C. (2003). The number of subcenters in large urban areas. *Journal of Urban Economics*, *53*(3), 321–338. https://doi.org/10.1016/S0094-1190(03)00026-3
Meijers, E., Hoogerbrugge, M., & Cardoso, R. (2018). Beyond Polycentricity: Does Stronger Integration Between Cities in Polycentric Urban Regions Improve Performance?: BEYOND POLYCENTRICITY. *Tijdschrift Voor Economische En Sociale Geografie*, *109*(1), 1–21. https://doi.org/10.1111/tesg.12292
Niu, Y., Ding, C., & Knaap, G.-J. (2015). Employment Centers and Agglomeration Economies. *Economic Development Quarterly*, *29*(1), 14–22. https://doi.org/10.1177/0891242414560813

Ord, J. K., & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, *27*(4), 286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

Porter, M. E. (1990). *The {Competitive} {Advantage} of {Nations}*. Basic Books.

Porter, M. E. (1998). Clusters and the new economics of competition. *Harvard Business Review*, *76*(6), 77–90.

Redding, S. J. (2023). Quantitative Urban Models: From Theory to Data. *Journal of Economic Perspectives*, *37*(2), 75–98. https://doi.org/10.1257/jep.37.2.75

Redfearn, C. L. (2007). The topography of metropolitan employment: Identifying centers of employment in a polycentric urban area. *Journal of Urban Economics*, *61*(3), 519–541. https://doi.org/10.1016/j.jue.2006.08.009

Rey, S. J. (2002). *Identifying regional industrial clusters in {Imperial} {County} {California}*. Technical

Rey, S. J., Anselin, L., Amaral, P., Arribas-Bel, D., Cortes, R. X., Gaboardi, J. D., Kang, W., Knaap, E., Li, Z., Lumnitz, S., Oshan, T. M., Shao, H., & Wolf, L. J. (2021). The PySAL Ecosystem: Philosophy and Implementation. *Geographical Analysis*, gean.12276. https://doi.org/10.1111/gean.12276

Riguelle, F., Thomas, I., & Verhetsel, A. (2007). Measuring urban polycentrism: A European case study and its implications. *Journal of Economic Geography*, *7*(2), 193–215. https://doi.org/10.1093/jeg/lbl025

Rosenthal, S. S., & Strange, W. C. (2004). Chapter 49 Evidence on the nature and sources of agglomeration economies. In *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2119–2171). Elsevier Inc. https://doi.org/10.1016/S1574-0080(04)80006-3

Schmidt, S., Krehl, A., Fina, S., & Siedentop, S. (2020). Does the monocentric model work in a polycentric urban system? An examination of German metropolitan regions. *Urban Studies*, 004209802091298. https://doi.org/10.1177/0042098020912980

Scott, L. M. (1999). *The Accessible City: Employment Opportunities in Time and Space* [PhD thesis, San Diego State University]. https://escholarship.org/uc/item/51h6f8qx

Small, K. A., & Song, S. (1994). Population and Employment Densities: Structure and Change. *Journal of Urban Economics*, *36*, 292–313. https://doi.org/10.1006/juec.1994.1037

Songchitruksa, P., & Zeng, X. (2010). Getis-ord spatial statistics to identify hot spots by using incident management data. *Transportation Research Record*, *2165*(2165), 42–51. https://doi.org/10.3141/2165-05

Straszheim, M. R. (1984). Urban agglomeration effects and employment and wage gradients. *Journal of Urban Economics*, *16*(2), 187–207. https://doi.org/10.1016/0094-1190(84)90041-X

von Ehrlich, M., & Seidel, T. (2013). More similar firms - More similar regions? On the role of firm heterogeneity for agglomeration. *Regional Science and Urban Economics*, *43*(3), 539–548. https://doi.org/10.1016/j.regsciurbeco.2013.02.007

von Thünen, J. H. (1826). *Der Isolierte Staat*. Hamburg.

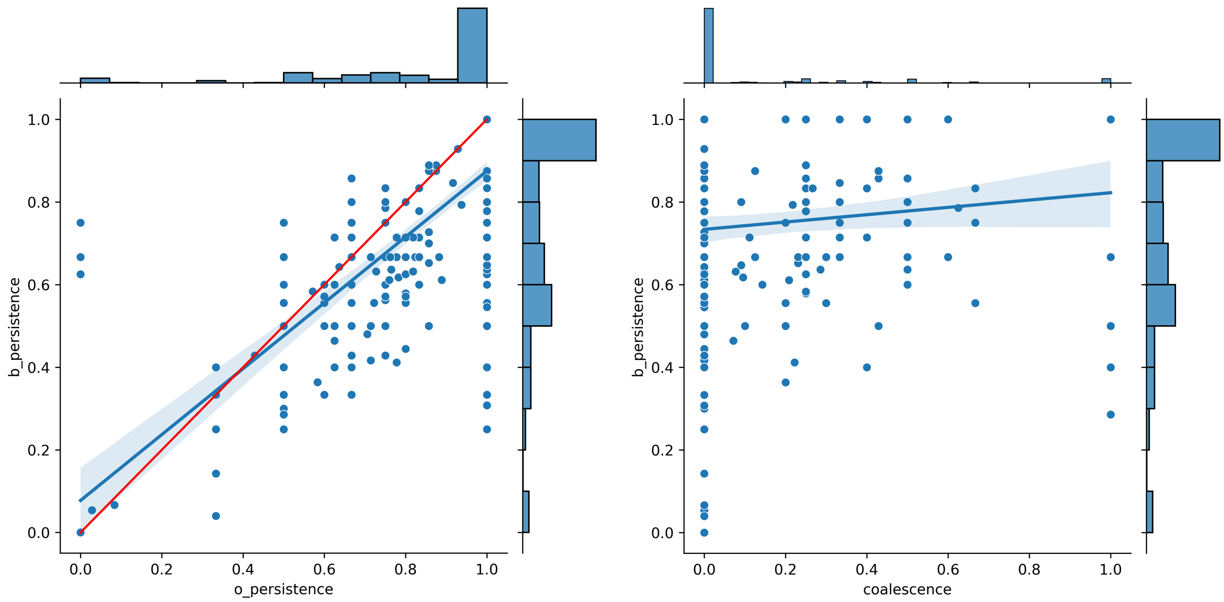
White, M. J. (1976). Firm suburbanization and urban subcenters. *Journal of Urban Economics*, *3*(4), 323–343. https://doi.org/10.1016/0094-1190(76)90033-4

White, M. J. (1988). Location choice and commuting behavior in cities with decentralized employment. *Journal of Urban Economics*, *24*(2), 129–152. https://doi.org/10.1016/0094-1190(88)90035-6

Zhang, W., & Derudder, B. (2019). How sensitive are measures of polycentricity to the choice of "centres"? A methodological and empirical exploration. *Urban Studies*, *56*(16), 3339–3357. https://doi.org/10.1177/0042098019843061

Zhang, Y., & Sasaki, K. (1997). Effects of subcenter formation on urban spatial structure. *Regional Science and Urban Economics*, *27*(3), 297–324. https://doi.org/10.1016/s0166-0462(96)02164-3

APPENDIX



(a) Center Persistence Measures

(b) Birth Perstence and Coalescence

Figure 10: Center Persistence and Coalescence

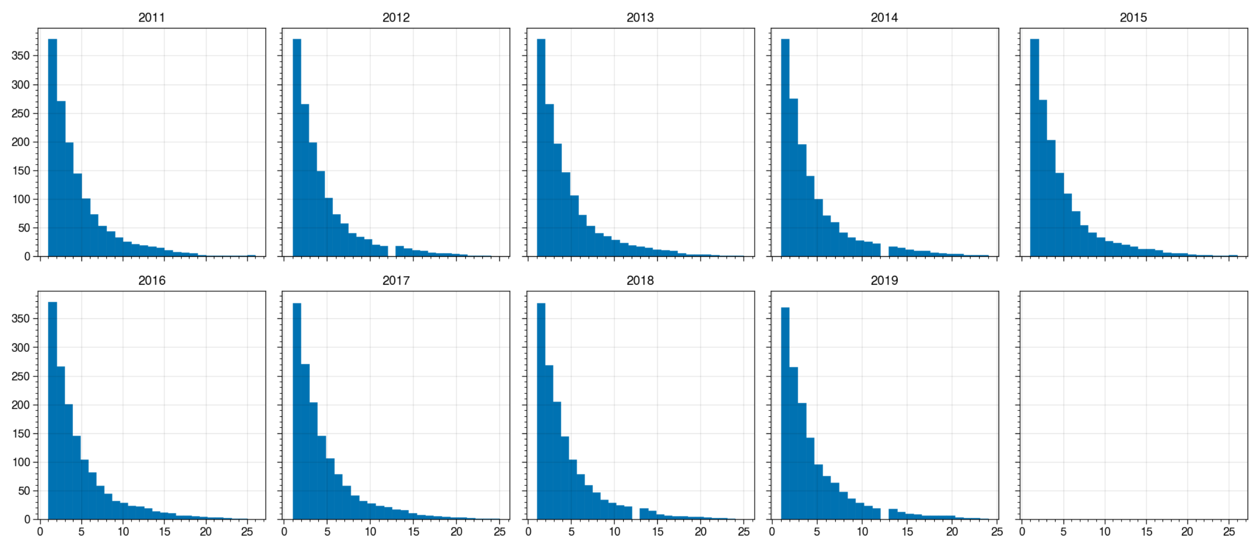


Figure 11: Histograms of Employment Center Counts by Year