A Better Delineation of U.S. Metropolitan Areas

Jordan Rappaport and McKenzie Humann

Federal Reserve Bank of Kansas City*

April 2025

Abstract

Metropolitan areas are a fundamental unit of economic analysis. Broadly defined, they are unions of built-up locations near each other among which people travel between places of residence, employment, and consumption. Despite the importance of metropolitan areas, metropolitan Core-Based Statistical Areas and other official U.S. delineations considerably stray from this broad definition. We develop a simple algorithm to better match this definition, using commuting flows among U.S. census tracts in 2000. Three judgmental parameters govern the minimum strength of commuting ties between locations to include them in the same metropolitan area, the maximum separating distance between locations, and the minimum density of outlying settlement. A parameterization that balances encompassing commuting flows and excluding sparsely settled land delineates 361 Kernel-Based Metropolitan Areas (KBMAs), in aggregate capturing almost all the population and employment of metropolitan CBSAs in a small fraction of their land area. Additionally, we benchmark KBMAs against two alternative parameterizations, one that prioritizes encompassing commuting flows and one that prioritizes excluding sparsely settled land.

Keywords: delineating metropolitan areas, Core-Based Statistical Areas, commuting flows, city size

JEL Classification Numbers: R12, R14, R23

---

*The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System. Thank you to Chris Cunningham, Giles Duranton, Xavier Gabaix, Todd Gardner, Sanghoon Lee, Albert Saiz, Elisabet Viladecans-Marsal, and several anonymous referees for helpful feedback. Thank you to Caleb Bray, Francis Dillon, and Sai Avinash Sattiraju for excellent research assistance. Supplemental maps, enumerations, and detailed tables are available from the paper's webpage.

1 Introduction

Metropolitan areas are a fundamental unit of economic analysis. We define them broadly as unions of built-up locations near each other with combined population of at least moderate scale and among which a significant share of residents and workers travel on a day-to-day basis between places of residence, places of employment, and places of consumption. By this definition, metropolitan areas correspond to distinct markets for labor and for non-traded goods and services. They are also likely to correspond to the geographic area determining many agglomerative externalities and for sharing many production and consumption amenities. In addition, national populations sort themselves across and within metropolitan areas with respect to numerous characteristics. Conversely, metropolitan areas may foster a sense of shared identity among diverse residents.

Despite this fundamental importance, official delineations considerably stray from our broad definition of metropolitan areas. Metropolitan Core-Based Statistical Areas (CBSAs), delineated by the U.S. Office of Management and Budget, vastly overbound metropolitan land area. For example, the Phoenix CBSA spans more than 14,000 square miles, mostly empty desert, and the Honolulu CBSA includes a Pacific atoll more than 900 miles from its downtown. Even so, many commonly recognized labor markets are split into multiple metropolitan CBSAs, including Raleigh and Durham, Los Angeles and Riverside–San Bernadino, and San Francisco and San Jose. Conversely, other arguably separate metropolitan areas are encompassed by a single metropolitan CBSA. For example, 20 miles of empty desert in 2000 separated the settled portion of the Coachella Valley—Palm Springs and neighboring municipalities in the central portion of the Riverside–San Bernadino CBSA—from the settled western portion of the CBSA.

Two other official delineations of U.S. metropolitan areas in 2000 egregiously stray from our definition. Urbanized Areas (UAs), combinations of census blocks with contiguous population density above a specified threshold, fragment unambiguous metropolitan areas. For example, a slight gap in residential settlement causes a portion of one of Kansas City's main suburban municipalities to be delineated as its own UA, notwithstanding that more than half of its employed residents work in the Kansas City UA and more than one third of its employment is made up of workers living in the Kansas City UA. Commuting Zones (CZs), combinations of U.S. counties delineated to identify rural labor markets, both carve up some unambiguous metropolitan areas and arguably span multiple others. For example, six commuting zones in the vicinity of New York City are connected to Manhattan by commuter rail. Conversely, a commuting zone in Pennsylvania fully encompasses four metropolitan CBSAs (Harrisburg-Carlisle, Lancaster, Lebanon, and York-Hanover).

The failures of official delineations impede decision making, policy implementation, and fundamental understanding. Most immediately, metropolitan delineations affect a wide range of choices, including on infrastructure investment, transportation planning, corporate location, and government-sponsored development. In addition, implementing a slew of legislation depends closely on metropolitan delineations (Congressional Research Service, 2014). For example, the income ceiling in 2024 for a neighborhood to qualify as low income under the Community Reinvestment Act was 13 percent lower in the Durham CBSA than in the Raleigh CBSA.

From a research perspective, Duranton (2021) argues that "meaningful and appropriate" delineations of metropolitan areas are required "to understand anything about fundamental urban questions." Similarly, better delineations are likely to improve our understanding of many other economic processes for which variation across locations underpins empirical study. Such questions include labor supply and demand (e.g., Kennan and Walker, 2011; Autor, Dorn and Hanson, 2013; Monte, Redding and Rossi-Hansberg, 2018), housing supply and demand (e.g., Green, Malpezzi and Mayo, 2005; Saiz, 2010; Van Nieuwerburgh and Weill, 2010, Landvoigt, Piazzesi and Schneider, 2015), non-housing consumption demand (Mian and Sufi, 2012, 2014; Guren et al., 2021), productivity spillovers (Glaeser and Mare, 2001; Moretti, 2004; Combes, Duranton and Gobillon, 2008; Greenstone, Hornbeck and Moretti, 2010), and social mobility (Chetty et al., 2014).

To address these and other questions and purposes, we develop a simple algorithm that delineates metropolitan areas that more closely match our broad definition. It combines elements from the methodologies used to delineate CBSAs, UAs, CZs, and several other alternatives. Similar to the construction of metropolitan CBSAS, UAs serve as cores, anchoring our delineations. The algorithm iteratively joins these cores together to form kernels based on the strength and distance of commuting flows among them and then builds out from the kernels to form Kernel-Based Metropolitan Areas (KBMAs).

Our delineation algorithm, like all others, requires judgment to set the values of parameters. Two such parameters set the minimum strength of commuting ties and the maximum separating distance for locations to be combined in the same metropolitan area; a third parameter sets the minimum density for outlying locations to be included. The algorithm is fully transparent and replicable; unlike CBSAs and UAs, it does not depend on the opinions of local residents.

We think of KBMAs as a baseline parameterization that is likely to be appropriate for most questions and purposes; it balances encompassing commuting flows and excluding locations that arguably are not sufficiently nearby or built up. As alternative benchmarks for other questions and purposes, we delineate more expansive kernel-based metropolitan regions and more compact kernel-based urban areas: The former parameterization more heavily weights encompassing commuting flows and so may better match studying extended connections such as long-distance commuting and regional supply chains. The latter parameterization more heavily weights excluding less built-up and less near locations and so may better match studying narrow geographic spillovers and urban land usage. Our delineation algorithm can also be flexibly parameterized to match numerous other questions and purposes.

The 361 KBMAs we delineate encompass 88 percent of the aggregate population of metropolitan CBSAs and 94 percent of their aggregate employment in just 15 percent of their aggregate land area. Their population ranges from 50,000 to 19 million and their land area from 27 to 6,100 square miles. The population distributions of the kernel-based metropolitan regions and kernel-based urban areas closely match that of KBMAs. Under all three parameterizations, land area expands less than proportionately with population, consistent with centripetal forces, such as centralized employment and amenities, constraining metropolitan expansion.

2 Existing Metropolitan Delineations

Delineating metropolitan areas requires making a number of judgments, either explicit or implicit. What geographic units serve as a building blocks? What makes building blocks sufficiently integrated to join them together in a cluster? Should clusters be anchored by a core? What qualifies a cluster as metropolitan?

2.1 Building Blocks

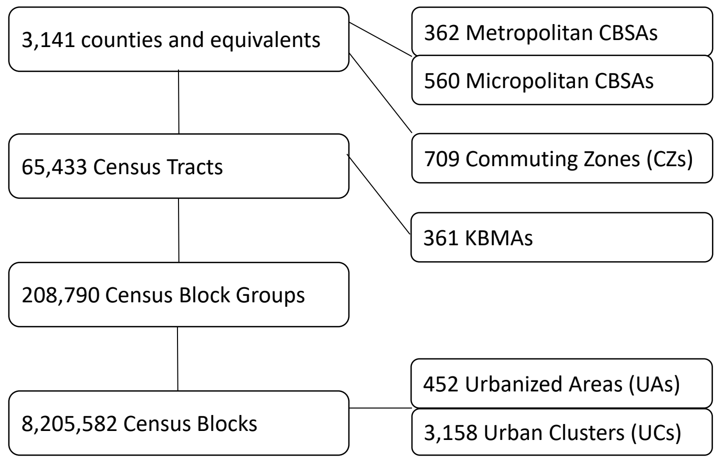
Census blocks are the most granular geographic unit for which the Census Bureau collects data. Blocks are typically quite small, for example the rectangle bounded by one city block on each side, but can also span many square miles in sparsely settled areas. For the 2000 decennial census, the Census Bureau delineated approximately 8.2 million blocks with land area ranging from 0 (all water) to 8,072 square miles (unsettled), with a median of 0.01 square miles. Their population ranged from 0 (more than a third of blocks) to a bit over 23,000. Across census blocks with residents, the median population was 25. Only limited block data is publicly available.

Census block groups combine up to 1,000 blocks and serve mainly for statistical purposes. Specifically, they are the smallest geographic unit for which the Census Bureau reports tabulated data from decennial census sample questions and from the American Community Survey.

Census tracts combine census block groups with several goals, including maintaining relatively stable boundaries over time (U.S. Census Bureau, 1997*b*). Even so, numerous tracts are re-delineated in preparation for each decennial census. For the 2000 census, the Census Bureau targeted tracts to encompass between 1,500 and 8,000 residents. Nevertheless, 2 percent had population in 2000 below 500 and a handful had population above 20,000. Tracts' land area in 2000 ranged from 0.06 square miles at the 1st percentile to 800 at the 99th percentile (median = 1.96 square miles). The endogenous delineation of census tracts to meet the population targets induces a tight negative correlation between land area and population density.

Counties in 2000 combined between 1 and 2,100 tracts, with population ranging from

Figure 1: Metropolitan Building Blocks Using the 2000 Decennial Census



Note: excludes U.S. territories.

1,000 at the 1st percentile to 1.1 million at the 99th percentile (median = 25,000) and land area ranging from 10 square miles at the 1st percentile to 8,100 square miles at the 99th percentile (median = 616). Importantly, their average land area considerably differs across U.S. regions, ranging from a median of 425 square miles in the South Atlantic census division to a median of 2,275 in the Mountain census division; this variation suggests that metropolitan delineations constructed using counties as building blocks are not comparable across different parts of the U.S. On the other hand, county borders have remained relatively stable since 1920, making counties an ideal building block for delineating metropolitan areas with unchanged borders since then. In addition, a wealth of county data is available from government agencies and private organizations.

2.2 Urbanized Areas and Urban Clusters

Urbanized Areas were first delineated by the Census Bureau following the 1950 decennial census, partly in response to rapidly growing population just outside the boundaries of large incorporated municipalities (Ratcliffe, 2015). Beginning with the 2000 decennial census, the Census Bureau used a mostly granular algorithm to delineate UAs and a smaller counterpart, Urban Clusters (UCs), as combinations of census blocks (U.S. Census

Bureau, 2002).[1]

The granular portion of the UA/UC algorithm focuses on residential density and proximity. It starts by identifying clusters of two or more contiguous block groups or blocks with population density of at least 1,000 per square mile. Expanding outward, adjacent block groups and blocks with population density of at least 500 are attached. Each resulting cluster is then joined with any similarly constructed clusters separated from it by no more than 0.5 miles along a connecting road segment, forming *interim cores* of chained clusters. Next, interim cores separated by no more than 2.5 miles along a connecting road segment are joined together, forming chained cores. Additional criteria add in blocks mostly surrounded by the chained cores as well as adjacent blocks containing a major airport. The resulting unions of census blocks constitute *urban area agglomerations*.[2]

All urban area agglomerations with population between 2,500 and 50,000 qualify as UCs under the 2000 algorithm, and most with population above this qualify as UAs. However, some of the larger agglomerations are split into two or more UAs to backwardly align them with metropolitan statistical areas (MSAs) delineated during the 1990s. For example, the Los Angeles UA is contiguous with two other UAs, Riverside–San Bernadino and Thousand Oaks, and so the three constitute a single urban area agglomeration. Splitting it aligns each of the resulting UAs with a legacy MSA. This backward alignment to the 1990s is perpetuated in the algorithms constructing UAs following the 2010 and 2020 decennial census, which split urban area agglomerations to avoid combining UAs previously delineated as separate (U.S. Census Bureau, 2011, 2022).[3]

Backwardly aligning the UAs in 2000 with legacy MSAs breaks an otherwise granular algorithm, contaminating using UAs for some purposes. In addition to the explicit history dependence, doing so introduces ad hoc subjectiveness that was applied idiosyn-

[1] The separate designations were dropped following the 2020 decennial census, in favor of "Urban Area", and the threshold population for inclusion was increased from 2,500 to 5,000 (U.S. Census Bureau, 2022).

[2] The term "urban area agglomeration" was introduced following the 2020 decennial census (U.S. Census Bureau, 2022).

[3] Backwardly aligning UAs has required more extensive splits of urban area agglomerations as urban settlement has spread. For UAs delineated following the 2000 decennial census, we identify 32 pairs with contiguous borders, ranging from a single point to a maximum of 9 miles, implying that splits were made at each. Eight of these contiguous borders exceeded 3 miles in length, the maximum prescribed by the published rule for the 2000 delineations (U.S. Census Bureau, 2002). For the UAs delineated using the 2020 decennial census, we identify 121 pairs with contiguous borders, ranging in length up to 47 miles.

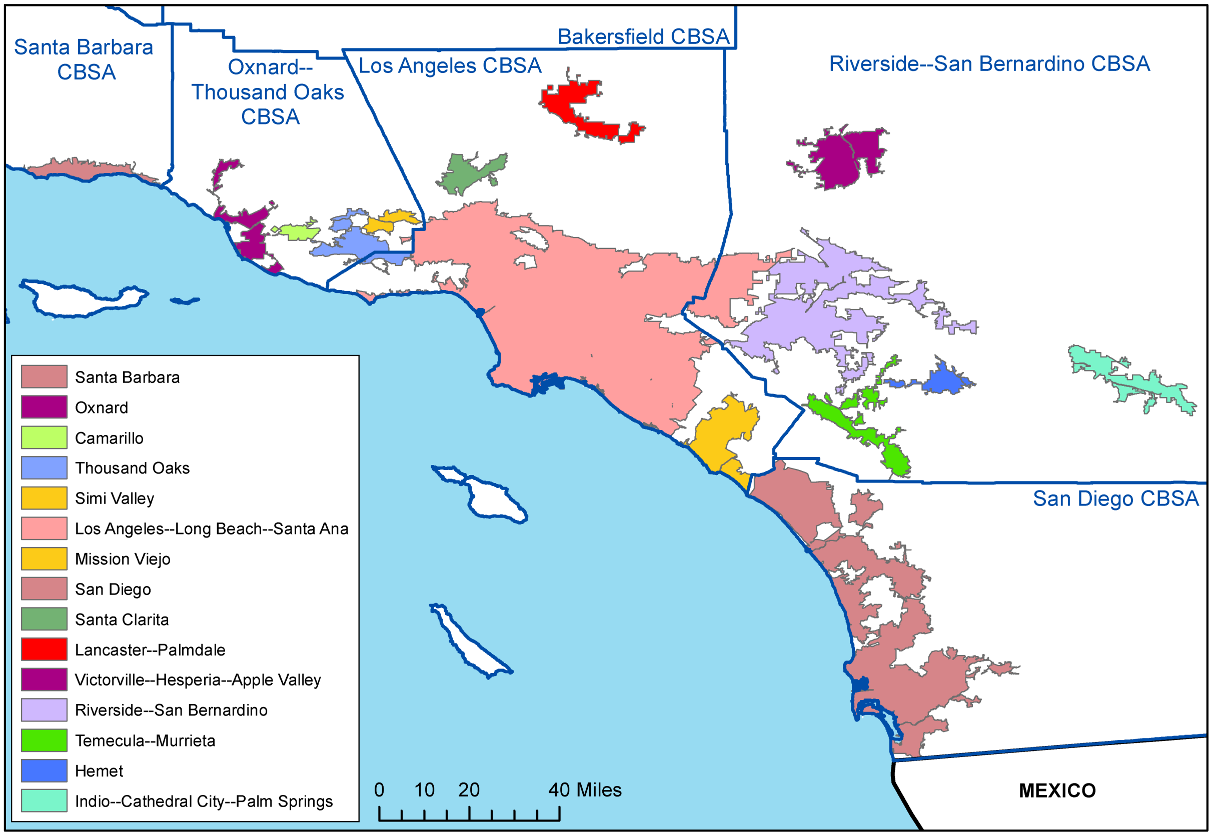
cratically across locations. Specifically, metropolitan delineations during the 1990s relied heavily on local opinion—the views of a wide range of public groups including business and other leaders, chambers of commerce, planning commissions, and local officials—to partition larger MSAs, designated as *consolidated*, into two or more *primary* MSAs (Office of Management and Budget, 1990).

From a pragmatic perspective, a key flaw of UAs and UCs is that they fragment many locations that are arguably integrated. For example, as illustrated in Figure 2, five separate UAs are located along a 25-mile ray in the southern portion of the Oxnard–Thousand Oaks CBSA. Large commuting flows among three of these—Los Angeles, Simi Valley, and Thousand Oaks—far exceed our threshold metric for combining them in a single metropolitan area as do the commuting flows between the Oxnard and Camarillo UAs. Similarly, as described in the introduction, a slight gap in residential settlement causes a portion of one of Kansas City's main suburban municipalities to be delineated as its own UA, notwithstanding that more than half of its residents work in the Kansas City UA and more than one third of its employment is made up of workers living in the Kansas City UA. Analogous suburban UAs, from which at least half of employed residents commute to a much larger UA, were delineated adjacent to the Phoenix, Seattle, Chicago, and New York UAs.[4]

Conversely, the relatively low density threshold for inclusion in urban area agglomerations contributes to some chains that arguably span multiple metropolitan areas. For example, a single urban area agglomeration stretches from northern New Jersey up along the Atlantic coast almost to New London, including a branch extending from New Haven north to Springfield, MA.

[4]The Lee's Summit, MO UA is separated from the Kansas City UA by 0.4 miles (implicitly not along a connecting road segment as otherwise they would not be separate). The Avondale, AZ UA is separated from the Phoenix–Mesa UA by 0.5 miles; the Marysville, WA, UA is separated from the Seattle UA by 0.4 miles; the Round Lake Beach–McHenrey–Grayslake, IL-WI UA is separated from the Chicago UA by 0.4 miles; and the Hightstown, NJ UA is separated from the New York–Newark UA by 0.1 miles.

Figure 2: Urbanized Areas in Southern California



Notes: Borders represent delineations following the 2000 decennial census. Blue lines demarcate metropolitan CBSAs. The Los Angeles UA includes a narrow corridor of census blocks along its northwest coast, some of which are not visible.

2.3 Metropolitan Core-Based Statistical Areas

The predecessor to the Office of Management and Budget (OMB) began delineating metropolitan statistical areas in the late 1940s, motivated by the desire to have geographic units for which government agencies could collect, tabulate, and publish data. The specific criteria and name of the delineated units evolved over time (Berry, Goheen and Goldstein, 1969; Office of Management and Budget, 1998; Gardner, 1999, 2021).

In preparation for the 2000 decennial census, OMB significantly revised its criteria to delineate metropolitan and micropolitan Core-Based Statistical Areas (Office of Management and Budget, 2000). Both are constructed using counties as geographic building blocks, with UAs serving as cores anchoring metropolitan CBSAs and UCs with popula-

Table 1: Official Delineations of U.S. Metropolitan Areas

Delineation	Core	Building Blocks	Integration Criteria	Other Criteria	obs	Population min	Population mdn	Population max	Land Area (sq.mi) min	Land Area (sq.mi) mdn	Land Area (sq.mi) max
Urbanized Areas (UAs)	none	census blocks	proximity	pop density	452	50.1k	118k	17.8m	12	64	3.4k
Urban Clusters (UCs)	none	census blocks	proximity	pop density	3,158	2.5k	5.9k	49.6k	0.04	4	270
Core-Based Statistical Areas (CBSAs)											
metropolitan	UA	counties	commuting, proximity		362	52.5k	223k	18.3m	140	1.6k	27.3k
micropolitan	UC with pop≥10k	counties	commuting, proximity		560	13.0k	43.7k	182k	110	770	21.4k
Commuting Zones (CZs)	none	counties	commuting		709	13.9k	175k	16.4m	620	3.6k	166k

Notes: Delineations are based on the 2000 decennial census and exclude U.S. territories. The CBSA delineations are the version promulgated on June 6, 2003 (Office of Management and Budget, 2003); the Commuting Zone delineations are disseminated by the Economic Research Service (Economic Research Service, 2012).

tion of at least 10,000 serving as cores anchoring micropolitan CBSAs. Each of these UAs and UCs is associated with one or more *central counties*, in which they are substantially encompassed. Nearby counties are designated as *outlying counties* of the CBSA if at least 25 percent of their employed residents work in the central counties or if at least 25 percent of their employment is made up of workers who live in the central counties. The published criteria also allow OMB to combine adjacent groups of counties that are algorithmically delineated as separate CBSAs if local opinion favors doing so.

As summarized in Table 1, OMB delineated 362 metropolitan CBSAs in 2000 with population ranging from 52,500 to 18.3 million and 560 micropolitan CBSAs with population ranging from 13,000 to 182,000. The considerable overlap in population between the two types emphasizes that they are distinguished from each other based on the population of their cores. OMB retained essentially the same criteria for the 2010 and 2020 decennial censuses (Office of Management and Budget, 2010, 2021). It also periodically updates delineations based on intercensal population estimates.

Metropolitan CBSAs deviate from our metropolitan definition in several ways. First, many vastly overbound metropolitan land area, reflecting the use of counties as building blocks. As described in the introduction, the vast majority of the Phoenix CBSA is empty desert. Similarly, the Riverside–San Bernadino and Anchorage CBSAs each span land area more than three times the size of Massachusetts. Even in the New York City–Newark–Edison CBSA, half of the land area in 2000 was accounted for by census tracts with both population and employment density below 500, the threshold for a census block to be included in a UA or UC.

Second, some CBSAs fully encompass what are arguably separate metropolitan areas. For example, the settled portion of the Coachella Valley constitutes its own television broadcasting market, contributing to our judgment that it belongs to a different metropolitan area than the Riverside–San Bernadino UA, which is a part of the Los Angeles television market.

Third, many CBSAs underbound metropolitan areas. For example, the rule splitting urban area agglomerations into multiple UAs causes Los Angeles and Riverside–San Bernadino to be delineated as separate CBSAs as it also does for San Francisco and San Jose. For Raleigh and Durham NC, a slight gap in residential settlement between them in 2000 causes them to be delineated as separate UAs, a necessary condition for them to be included in separate CBSAs. An appendix table enumerates other pairs of metropolitan CBSAs with large cross commuting. Especially egregious, 39 percent of workers residing in the Hinesville–Fort Stewart, GA metropolitan CBSA in 2000 commuted to places of employment in the Savannah, GA metropolitan CBSA.[5]

For research purposes, the CBSA delineations suffer from several methodological flaws. One is that they are history dependent. As described previously, metropolitan CBSA cores, UAs, are split to backwardly align with MSAs delineated during the 1990s. This history dependence deepened with the 2010 and 2020 delineations, reflecting that suburban expansion filled in unsettled gaps between UAs, thereafter requiring more extensive splits. Worse, the backward alignment is to delineations that relied heavily on local

[5]The high commuting outflow rate from Hinesville–Fort Stewart to Savannah does not meet the 25 percent threshold for merging because a significant portion of it originates in an outlying rather than central county. In recognition of overbounding and underbounding, OMB also delineates subsets of CBSAs, Metropolitan Divisions, and supersets of them, Combined Statistical Areas. It also constructs analogs for the six New England states using county subdivisions as geographic building blocks.

opinion, contaminating the CBSA algorithm.[6] Another flaw arises from the considerable variation in counties' average land area across U.S. regions, which implicitly delineates CBSAs differently across regions and so limits comparisons.

2.4 Additional Methodologies

A third set of possible proxies for U.S. metropolitan areas, Commuting Zones (CZs), were delineated by the U.S. Department of Agriculture's Economic Research Service following each of the 1980, 1990, and 2000 decennial censuses with the goal of better understanding rural labor markets (Tolbert and Sizer, 1987, 1996; Foote, Kutzbach and Vilhuber, 2017). Like CBSAs, CZs are constructed using counties as geographic building blocks. The commuting strength between two locations is calculated as the sum of the commuting flows in both directions normalized by the level of employment in the location where it is smaller. The delineation algorithm begins by constructing an N-by-N symmetric matrix of commuting strength between all possible pairs of counties; the pair with the strongest tie are joined, constituting a first cluster. The process is iteratively repeated, calculating a new (N-1)-by-(N-1) symmetric matrix and joining the pair of observations with the strongest tie until it falls below a judgmental threshold. The resulting clusters fully partition the United States.

CZs share many of the same flaws as CBSAs, including over-bounding and under-bounding. A large share of CZs encompass one, two, or three CBSAs together with a handful of adjacent rural counties. Even so, many medium and large CBSAs are split across multiple CZs, with groups of unambiguous suburbs separated from central business districts. As illustrated in an appendix, six CZs in the vicinity of New York City are connected to Manhattan by commuter rail.

Expanding beyond the United States, the European Union and the OECD delineate Functional Urban Areas in their member countries (Dijkstra and Poelman, 2012; Brezzi

[6] Illustrating this heavy dependence on local opinion, the 1990s algorithm delineated a Consolidated MSA chaining from southern New Jersey up through New York City and then further north into Connecticut and eastern Pennsylvania. OMB relied on local opinion to divide it into 15 Primary MSAs. Among the more idiosyncratic, the Jersey City Primary MSA is made up of a 14-mile strip of land tightly wedged between Newark, NJ and the Hudson River.

et al., 2012; Dijkstra, Poelman and Veneri, 2019). The first of four stages joins adjacent grid cells of 1 square kilometer into the same cluster if each has population density of at least 1,500 persons per square kilometer (3,885 persons per square mile). Resulting clusters that have population of at least 50,000 are considered *urban centres*. These are analogous to U.S. Urbanized Areas except that they have a density threshold almost eight-fold higher and so exclude considerable suburban area. The second stage combines the governmental administrative units that overlap each urban center, forming *core cities*. The third stage joins core cities between which commuting ties exceed a threshold strength, in essence creating kernels. The final stage builds out from the kernels, attaching local government administrative units with which commuting ties exceed a threshold strength.[7]

Duranton (2015) goes a step beyond the examples above, granularly delineating metropolitan areas that are multilevel hierarchies. The directional strength between locations is measured by the commuting outflow rate, calculated as the flow from origin to destination relative to the number of workers residing in the origin. A first round of joins classifies each location as subsidiary to the one with which it has the highest outflow rate, contingent on the rate exceeding a judgmental threshold. The resulting first-round clusters mix hub-and-spoke (B and C each attached to A) and chained configurations (C attached to B and B attached to A). These are then iteratively used to effect a second round of joins, attaching each first-round cluster to the one with which it has the strongest outflow rate. The iteration continues until no outflow rate exceeds the threshold. Applying the algorithm to U.S. counties constructs hierarchical clusters whose population is tightly correlated with that of matched CBSAs (Dingel, Miscio and Davis, 2021).

Reliable commuting data is not available for most nations, and so delineating metropolitan areas must rely on other approaches. One uses satellite images of nighttime lights and land cover. For example, Ch, Martin and Vargas (2021) identify clusters

[7]The European Union, OECD, the United Nations, and The World Bank jointly adopted an urban-rural classification system that complements the EU/OECD delineations (Eurostat et al., 2021; Dijkstra et al., 2021). Countries are gridded into cells of 1 square kilometer, each of which is assigned to one of three categories: an *urban centre*, constructed as described; an *urban cluster*, contiguous grid cells, not in urban centres, that have population density of at least 300 persons per square kilometer (777 persons per square mile) and total population of at least 5,000; and *rural grid cells*, all cells not in an urban centre or urban cluster. These are respectively characterized as "densely populated", "intermediate density", and "thinly populated".

of adjacent pixels with nighttime light intensity above alternative thresholds to delineate between 4,200 and 6,700 metropolitan areas worldwide. Dingel, Miscio and Davis (2021) similarly use nighttime light to cluster pixels, which they match with intersecting administrative units. Using U.S. counties as administrative building blocks, the implied unions have population that is tightly correlated with the population of matched CBSAs across a range of light-intensity thresholds, validating using the methodology for some purposes when commuting data is not available.

Another approach focuses on physical structures. de Bellefon et al. (2021) classify small grid cells in France as urban if the density of buildings in them exceeds a threshold percentile across all grid cells in a country and then cluster the grid cells to construct urban areas. Similarly, Arribas-Bel, Garcia-Lopez and Viladecans-Marsal (2021) apply a spatial clustering algorithm to the precise location of all buildings in Spain to construct urban areas.

A third approach relies more directly on judgment. Galdo, Li and Rama (2021) ask several groups of assessors to use Google Earth and Google Maps to classify a sample of locations in India as either urban or rural. A machine learning algorithm then uses several measurable characteristics of the locations—such as population, population density, land cover, and nighttime luminosity—to predict the urban/rural status of more than 500,000 villages, towns, and cities.

3 Constructing KBMAs

Our metropolitan definition is purposely imprecise, leaving scope for judgment on whether specific delineations are consistent with it. Alternative parameterizations of the KBMA algorithm can match a wide range of judgments on encompassing places of residence and employment that are near each other while excluding lightly settled locations. Like the methodologies described above, no explicit component is included to encompass places of consumption.

3.1 Algorithm

We use census tracts, UAs, and UCs as our building blocks, taking advantage of carefully measured commuting flows among them. As with CBSAs, UAs and UCs with population above 10,000 serve as the cores of KBMAs. The first stage of the algorithm iteratively joins these cores together into kernels, paralleling the iterative construction of CZs. Doing so combines many of the UAs that are split by backward alignments. The second stage builds out from the kernels, attaching outlying tracts tied to them by sufficiently strong commuting flows.

We symmetrically measure the strength of the commuting tie between two locations by the sum of the inflow and outflow rates in both directions. Let $f_{i,j}$ represent the gross commuting flow from location i to location j; let e_i and w_i respectively represent employment in location i and workers residing there. The commuting strength between i and j (and between j and i) is given by,

$$s_{i,j}, s_{j,i} = \frac{f_{i,j}}{w_i} + \frac{f_{i,j}}{e_j} + \frac{f_{j,i}}{w_j} + \frac{f_{j,i}}{e_i} \tag{1}$$

Measuring strength symmetrically abstains from imposing a hierarchy. Normalizing directional flows both by resident workers and employment equally weights the importance of a location serving as a source of labor supply and as a source of labor demand.

Pragmatically, summing the four terms rather than taking the maximum (the CBSA formula) avoids a bias against combining locations of similar size. The normalization of the gross flows yields relatively low values for all four terms when two locations have similar size compared to the maximum value of the four terms when one location is much larger than the other. Correspondingly, we judge that appropriately delineating metropolitan areas when measuring strength by the maximum value would require setting a much lower threshold, σ, for joining two similar locations than for joining two unequal ones.

The initial iteration of the first stage begins by constructing an N-by-N symmetric matrix of the commuting strength between all pairs of cores. The pair with maximum strength, subject to having separating distance no more than δ, is joined together and then a new (N-1)-by-(N-1) matrix is constructed. Iterating continues until the maximum pairwise strength drops below σ.

The second stage attaches outlying tracts to a kernel if the strength of their commuting tie with the kernel weakly exceeds σ, their distance from the kernel does not exceed δ, and either their population or employment density weakly exceeds η.

We rely almost exclusively on one source of data, the Census Transportation Planning Package 2000 (Bureau of Transportation Statistics, 2005). It re-tabulates responses to the journey-to-work questions on the long form of the 2000 decennial census by tract of employment and by origin-destination pairs.[8]

As a prerequisite to implementing our algorithm, we construct tract-based approximations of UAs and UCs with population above 10,000. Most census tracts are either fully overlapped by a UA/UC or else fully disjoint with any UA/UC; even so, many tracts partly intersect with a UA/UC. As described in an appendix, we use a fit criterion to set threshold shares of a tract's population and land area that must intersect with a UA/UC to include the tract in that UA/UC's approximation. Only 1,274 of the 1,374 UAs and eligible UCs are actually approximated, reflecting that 100 of the eligible UCs have no intersecting tract that meets both the population and land-area thresholds. As described in the appendix, this shortfall only negligibly affects delineations.

3.2 Parameterization

Parameterizing the threshold commuting strength and maximum allowed separating distance to join locations and the threshold density of outlying tracts to attach them to kernels rests heavily on judgment. Any set of values will inevitably contradict many people's priors, either because of specific implied delineations or because of judgements of what qualifies as sufficiently integrated, near, and built-up. Alternative parameterizations illuminate how judgments affect delineations.

[8]A question on the 2000 decennial census long form and subsequent American Community Surveys asks respondents to give the address where they primarily worked the previous week. An alternative source of origin-destination commuting flows, the Census Bureau's LEHD Origin-Destination Employment Statistics (LODES), determines workplaces by the address of the business establishment with which employees are associated for payroll reporting. These establishment locations frequently differ from actual work locations. In addition, the LODES dataset does not cover non-payroll employment, such as self-employment.

We judgmentally parameterize KBMAs to balance encompassing commuting flows and excluding sparsely settled land, while resting only lightly on separating distance to constrain joins. As such, KBMAs serve as a baseline likely to match most reasons for using metropolitan delineations. We also delineate more expansive kernel-based metropolitan regions, setting lower values for threshold commuting strength and density and a higher allowed separating distance, and more compact kernel-based urban areas, setting higher vales for threshold strength and density and a lower allowed separating distance.

The kernel portion of the construction depends only on the commuting strength and distance parameters, σ and δ. The composition of many kernels is especially sensitive to the parametrization of σ, which we set to 0.25 based on our judgment of the implied unions. To span a wide range of judgments, we set σ much lower, to 0.10, for kernel-based metropolitan regions and much higher, to 0.40, for kernel-based urban areas. We set the maximum allowed separating distance, δ, to 20 miles, measured between the nearest tract centroids, consistent with our judgment of "near" and sufficiently high that it blocks relatively few joins during the kernel iteration.[9] Our alternative metropolitan region and urban area parameterizations respectively set δ to 40 miles and to 10 miles. Under the KBMA parameterization, iterating ends after 343 joins, leaving 931 kernels.

Prior to iterating, the 1,274 cores form more than 810,000 pairs. Constrained to those separated by no more than 20 miles, the initial iteration joins the Miami and Key Biscayne cores, which have pairwise commuting strength of 1.28 and separating distance of 5.2 miles. Table 2 illustrates some of the later iterations as σ is incrementally lowered from a strength of 0.30 to its KBMA parameterized value of 0.25 and then further down to a strength of 0.20. For example, the 0.30 strength threshold is sufficiently low to join Raleigh and Durham with one additional core, San Francisco and San Jose with nine additional cores, and Los Angeles and Riverside–San Bernadino with 12 additional cores. Lowering σ to 0.25 joins the already combined Manchester and Nashua cores with Boston, Worcester, and two others. Further lowering σ to 0.20 joins this cluster with

[9]The tract centroids we use are the *internal points* reported by the Census Bureau. Some irregular-shaped tracts have geographic centroids that lie outside their physical territory, e.g. crescent shapes and tracts split by a water body. In these cases, the Census Bureau locates the internal point in the tract portion nearest the geographic centroid.

Table 2: Kernels under Alternative Commuting Strength Thresholds

Tighter (σ=0.30)	**Actual** (σ=0.25)	**Looser** (σ=0.20)
Raleigh, NC; Durham, NC; Clayton, NC		Smithfield, NC
San Francisco-Oakland, CA; San Jose, CA; Concord, CA; San Rafael-Novato, CA; Antioch, CA; Vallejo, CA; 5 others	Fairfield, CA; Vacaville, CA; Napa, CA; Dixon, CA	
Los Angeles-Long Beach-Santa Ana, CA; Riverside-San Bernadino, CA; Mission Viejo, CA; Lancaster-Palmdale, CA; Thousand Oaks, CA; Simi Valley, CA; 8 others		Oxnard, CA; Camarillo, CA; Santa Paula, CA
Oxnard, CA; Camarillo, CA; Santa Paula, CA		*included in Los Angeles kernel*
Nashua, NH-MA; Manchester, NH	Boston, MA-NH-RI; Worcester, MA-CT; Leominster-Fitchburg, MA; Athol, MA	Barnstable Town, MA (Cape Cod); Dover-Rochester, NH--ME; Portsmouth, NH--ME; Concord, NH; Franklin, NH
Washington, DC-VA-MD; Frederick, MD; Fredericksburg, VA; St. Charles, MD; 3 others		Baltimore, MD; Aberdeen-Havre De Grace-Bel Air, MD; Westminster, MD, 2 others
Baltimore, MD; Aberdeen-Havre De Grace-Bel Air, MD; Westminster, MD; 2 others		*included in Washington DC, kernel*
Seattle, WA; Marysville, WA; North Bend, WA		Bremerton, WA; Olympia-Lacey, WA; Centralia, WA; Shelton, WA
Detroit, MI; Ann Arbor, MI; South Lyon-Howell-Brighton, MI; Monroue, MI	Lapeer, MI	Port Huron, MI
New York-Newark, NY-NJ-CT; Trenton, NJ; Hightstown, NJ (Princeton); 8 others	Poughkeepsie-Newburgh, NY; Port Jervis, NY--PA; New Paltz, NY	
	Providence, RI-MA; New Bedford, MA	
	Akron, OH; Canton, OH; Alliance, OH	
		Grand Rapids, MI; Holland, MI
		Johnson City, TN; Kingsport, TN-VA

Notes: The left column reports cores joined in selected kernels under a commuting strength threshold an increment above the KBMA parameterized value. The middle and right columns report additional cores joined from lowering the strength threshold to its KBMA value ($\sigma = 0.25$) and then lowering it a further increment. Joined cores must be separated by no more than 20 miles ($\delta = 20$). An online workbook enumerates the sequence of iterative joins under alternative maximum distances.

Barnstable Town (Cape Cod) and three others. Doing so also joins the Washington D.C. and Baltimore clusters, as well as the Los Angeles and Oxnard clusters. It would additionally join the Indio–Cathedral City–Palm Springs core to the Los Angeles cluster except that the separating distance between them is a few tenths above the allowed 20 miles.

We judge that setting σ to a commuting strength of either 0.30 or 0.25 constructs a set of kernels that is consistent with our metropolitan definition. One reason we choose the lower value is to keep the threshold comfortably below the join of Raleigh and Durham, which occurs at a strength slightly above 0.30. We are more skeptical that the set of kernels implied by lowering σ to a strength of 0.20 is consistent with our definition. A number of the incremental joins are with cores we think of as seasonal rather than day-to-day destinations. Specifically, Oxnard, Cape Cod, Olympia, and Port Huron all market themselves as coastal getaways.

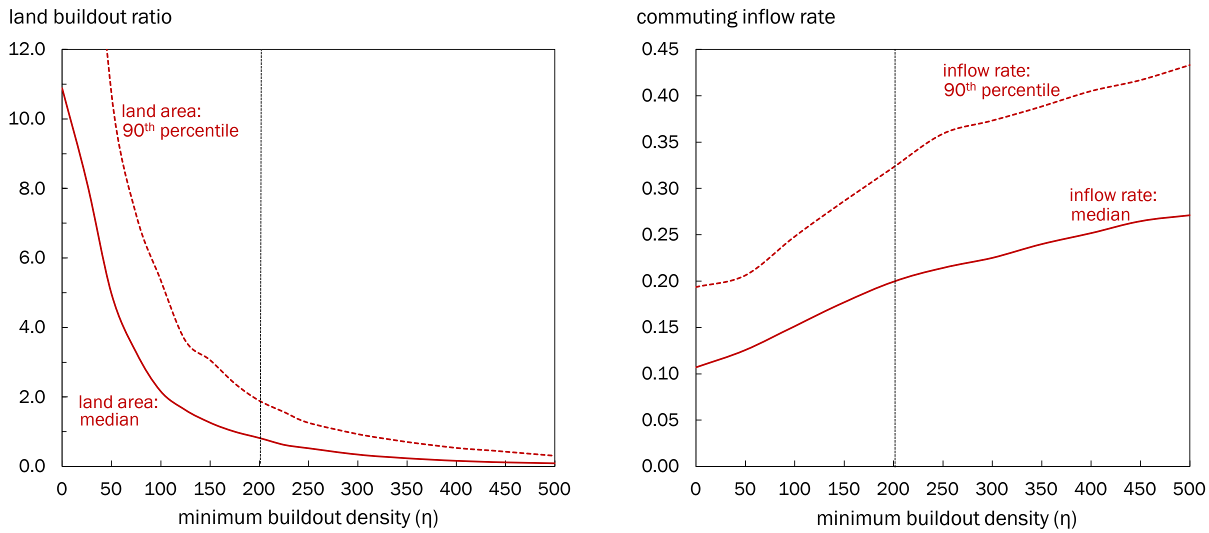
In addition, setting σ to a commuting strength of 0.25 rather than 0.20 implies a more robust buffer against joining the Indio–Cathedral City–Palm Springs core to the Los Angeles kernel, consistent with our prior that it constitutes its own metropolitan area. Otherwise, not joining the two rests fragilely on their separating distance being a tad above the parameterized maximum. More generally, we methodologically prefer commuting strength to serve as the primary determinant of joins, reflecting that strength typically decreases with separating distance and so in part implicitly accounts for distance.

We judge that the kernel iterations appropriately recombine UAs split from each other by the backward alignment with the 1990s MSAs. We are able to identify 32 such splits, 19 of which are recombined at the KBMA parameterization. Several other pairs are recombined if σ is set to a strength of 0.20, including Port Huron and Detroit, Portsmouth and Boston, and Johnson City and Kingsport. All but four pairs are recombined at the parameterization for kernel-based metropolitan regions, which sets σ to a strength of 0.10.

Not rejoining the remaining four pairs of metropolitan regions illustrates the success of the kernel iteration at avoiding constructing long chains of cores. One of the remaining splits of urban area agglomerations divides New York and Bridgeport–Stamford into separate UAs, between which the pairwise commuting strength is 0.19. The iterative sequencing initially joins each to other cores, diluting the pairwise strength between the resulting kernels. Further iterative dilutions delay combining New York and Bridgeport–Stamford in the same kernel until σ falls below a strength of 0.08. As another example, the iterative sequencing delays recombining the split San Diego and Mission Viejo UAs until σ falls below 0.05, at which point the San Diego and Los Angeles kernels join.

In contrast to the high sensitivity for specific kernels, an appendix figure illustrates the

Figure 3: Sensitivity of Built-Out Kernels to the Density Threshold



Notes: Left panel shows the median and 90th-percentile ratios of the land area of the buildout portion of built-out kernels relative to the land area of the kernel portion as η is increased from 0 to 500. The dashed vertical line corresponds to the parameterized KBMA value, $\eta = 200$. The right panel shows the median and 90th percentile rates of commuting inflows. For comparability, kernels are restricted to the 302 with population of at least 50,000. A complementary appendix figure illustrates buildout ratios for population and employment, and the commuting outflow rate.

gradual decline in the number of kernels as σ is lowered. It also illustrates that relatively few iterative joins are constrained by the maximum allowed separating distance, δ, at the parameterizations of KMBAs, kernel-based metropolitan regions, and kernel-based urban areas.

The buildout stage requires additionally parameterizing η, the minimum settlement density for eligible tracts to be attached to a kernel. We judgmentally set it to 200 (residents per square mile or workers per square mile, whichever is higher), balancing encompassing commuting flows and excluding sparsely settled land.

Figure 3 illustrates the tradeoff. The left panel shows the ratio of land in the buildout portion of KBMAs to land in their kernel portion as η is increased from 0 to 500. The median and 90th percentile buildout ratios at 200 are 0.8 and 1.9, respectively. Both begin to blow up as η is lowered below this. The right panel shows the median and 90th percentile of KBMAs' commuting inflow rates. Both rise significantly as η is increased from 0.

Table 3: 50 Largest KBMAs by Population in 2000

rank	KBMA Title	Population	rank	KBMA Title	Population
1	New York--Newark, NY--NJ	19,139,000	26	Orlando, FL	1,657,000
2	Los Angeles--Long Beach--Santa Ana, CA	15,191,000	27	Sacramento, CA	1,614,000
3	Chicago, IL--IN--WI	8,831,000	28	Kansas City, MO--KS	1,498,000
4	San Jose--San Francisco--Oakland, CA	6,257,000	29	Milwaukee, WI	1,471,000
5	Philadelphia, PA--NJ--DE--MD	5,520,000	30	Virginia Beach, VA	1,443,000
6	Boston, MA--NH--CT	5,250,000	31	San Antonio, TX	1,414,000
7	Miami, FL	4,932,000	32	Providence, RI--MA	1,391,000
8	Dallas--Fort Worth--Arlington, TX	4,736,000	33	Charlotte, NC--SC	1,353,000
9	Detroit, MI	4,501,000	34	Salt Lake City, UT	1,339,000
10	Washington, DC--VA--MD--WV	4,428,000	35	Las Vegas, NV	1,322,000
11	Houston, TX	4,278,000	36	Indianapolis, IN	1,304,000
12	Atlanta, GA	3,962,000	37	Columbus, OH	1,295,000
13	Phoenix--Mesa, AZ	2,991,000	38	New Orleans, LA	1,220,000
14	Seattle, WA	2,942,000	39	Buffalo, NY	1,090,000
15	San Diego, CA	2,742,000	40	Hartford, CT	1,018,000
16	Minneapolis--St. Paul, MN--WI	2,640,000	41	Austin, TX	996,000
17	Baltimore, MD--PA	2,504,000	42	Memphis, TN--MS--AR	990,000
18	Denver--Aurora, CO	2,314,000	43	Raleigh--Durham, NC	978,000
19	Tampa--St. Petersburg, FL	2,313,000	44	Nashville-Davidson, TN	970,000
20	St. Louis, MO--IL	2,275,000	45	Akron, OH	928,000
21	Cleveland, OH	2,156,000	46	Louisville, KY--IN	917,000
22	Pittsburgh, PA	2,045,000	47	Jacksonville, FL	907,000
23	Bridgeport--New Haven--Stamford, CT--NY	1,844,000	48	Oklahoma City, OK	871,000
24	Cincinnati, OH--KY--IN	1,744,000	49	Honolulu, HI	864,000
25	Portland, OR--WA	1,729,000	50	Richmond, VA	854,000

Note: An enumeration of all KBMAs, with additional variables, is included as an appendix and online.

We judge that the marginal benefit of encompassing more flows by decreasing η below 200 approximately equals the marginal cost of encompassing additional sparsely settled land area. To span a wide range of judgments, we drop any minimum for kernel-based metropolitan regions, setting η to 0; we set it moderately higher, to 500, for kernel-based urban areas. (500 is the threshold population density for census blocks to be included in a UA or UC.) An analogous appendix figure shows population and employment buildout ratios and the encompassment of commuting outflows, all of which are less sensitive to the parameterization of η.

Lastly, we classify only the 361 built-out kernels that have population above 50,000 as KBMAs, matching the convention used by several of the delineations described in the previous section. Of course, higher or lower population thresholds may better match various purposes. Table 3 lists the 50 KBMAs with largest population.[10] An enumeration of all KBMAs along with detailed data on underlying census tracts, cores, built-out kernels, and the iterative joins constructing kernels are available from the paper's webpage.

4 Realized Delineations

To gauge our success in matching our metropolitan definition, we first show maps of realized KBMAs and describe KBMAs' overlap with metropolitan CBSAs. We then document the extent to which KBMAs encompass commuting inflows and outflows.

4.1 Illustrations

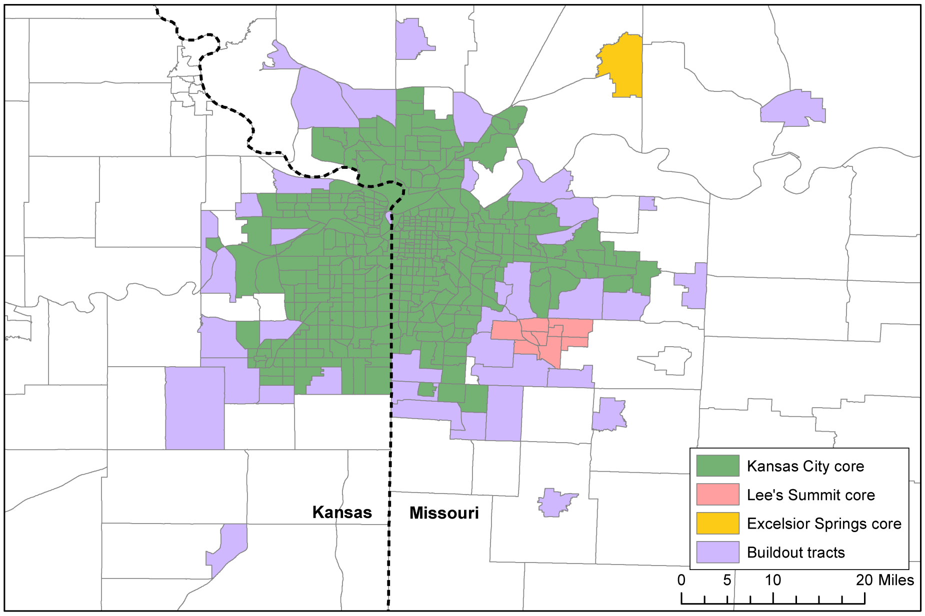
Figure 4 illustrates the composition of the Kansas City MO–KS KBMA. The kernel's three cores correspond to the Kansas City, MO–KS UA; a portion of a major suburb, Lee's Summit; and a very small suburban UA, Excelsior Springs. Most of the buildout tracts are contiguously connected to the kernel; the detached ones correspond to UCs with population below 10,000. Most excluded tracts are expansive, reflecting sparse settlement. The main exception, the cluster of small tracts near the northwest corner, corresponds to the Leavenworth UC, whose commuting strength with the Kansas City kernel falls just short of the threshold to combine with it and whose built-out population on its own is too low to qualify as a KBMA.

Figure 5 zooms out. Kansas City and its three neighboring KBMAs each occupy the most densely settled portion of a corresponding metropolitan CBSA. The commuting strength between the Kansas City and Lawrence kernels is a tick below 0.150, sufficient to combine them under the parameterization for kernel-based metropolitan regions.

The correspondence between KBMAs and metropolitan CBSAs is less clean in Southern California. As illustrated in Figure 6, the Los Angeles KBMA encompasses essentially

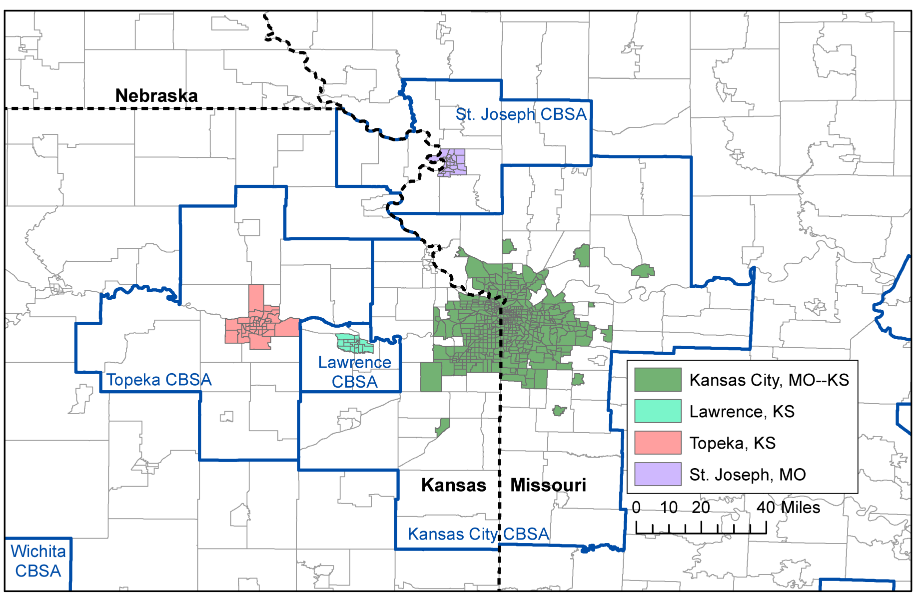
[10] We title KBMAs using the Census Bureau's algorithm for titling UAs and UCs (U.S. Census Bureau, 2002). The first listed name is the incorporated place with highest population. The names of the next two incorporated places by population are included if they have population exceeding 250,000 or if their population is at least two thirds that of the largest incorporated place. State postal abbreviations are ordered to correspond to any places included in the title and then by descending order of each state's population in the KBMA.

Figure 4: Composition of the Kansas City KBMA



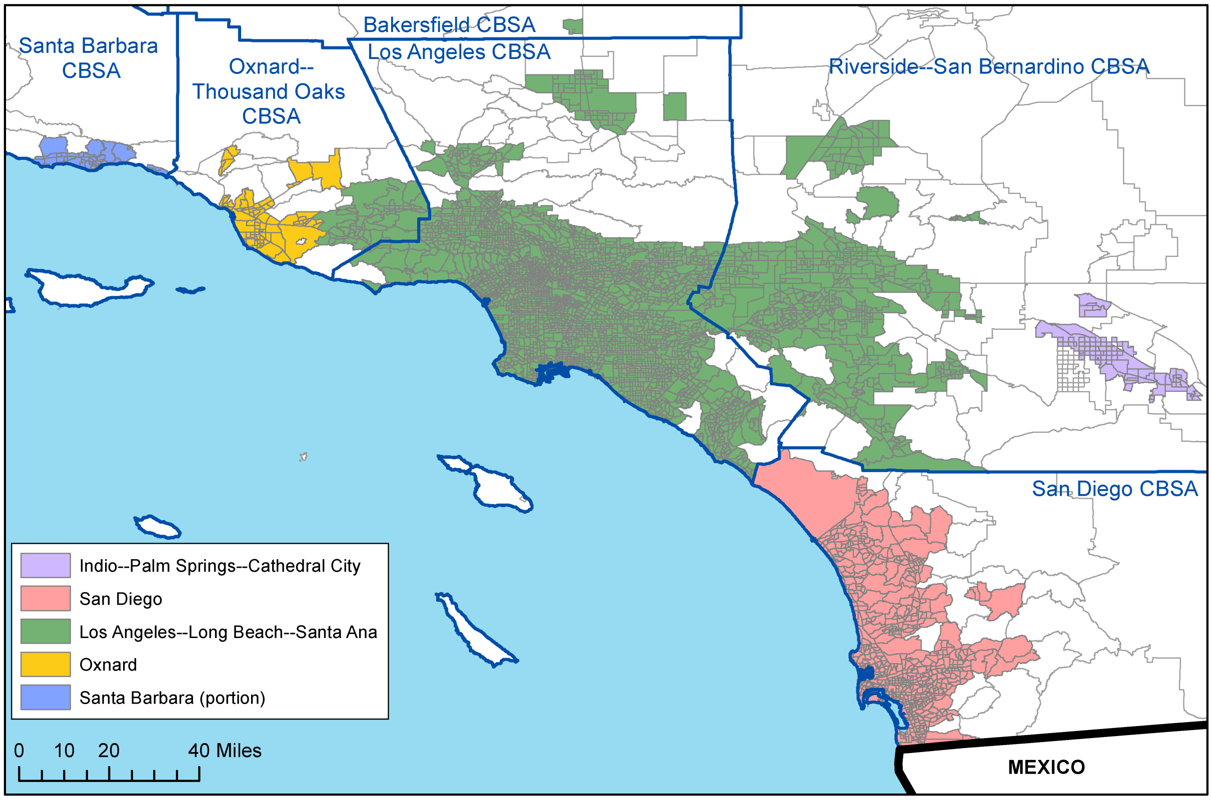
Notes: Census tracts, demarcated in gray, have population density that is inversely correlated with their land area. The displayed area lies entirely within the Kansas City MO-KS metropolitan CBSA.

Figure 5: Kansas City and Neighboring KBMAs.



Notes: Blue lines demarcate the borders of metropolitan CBSAs. Census tracts, demarcated in gray, have population density that is inversely correlated with their land area.

Figure 6: Los Angeles and Neighboring KBMAs



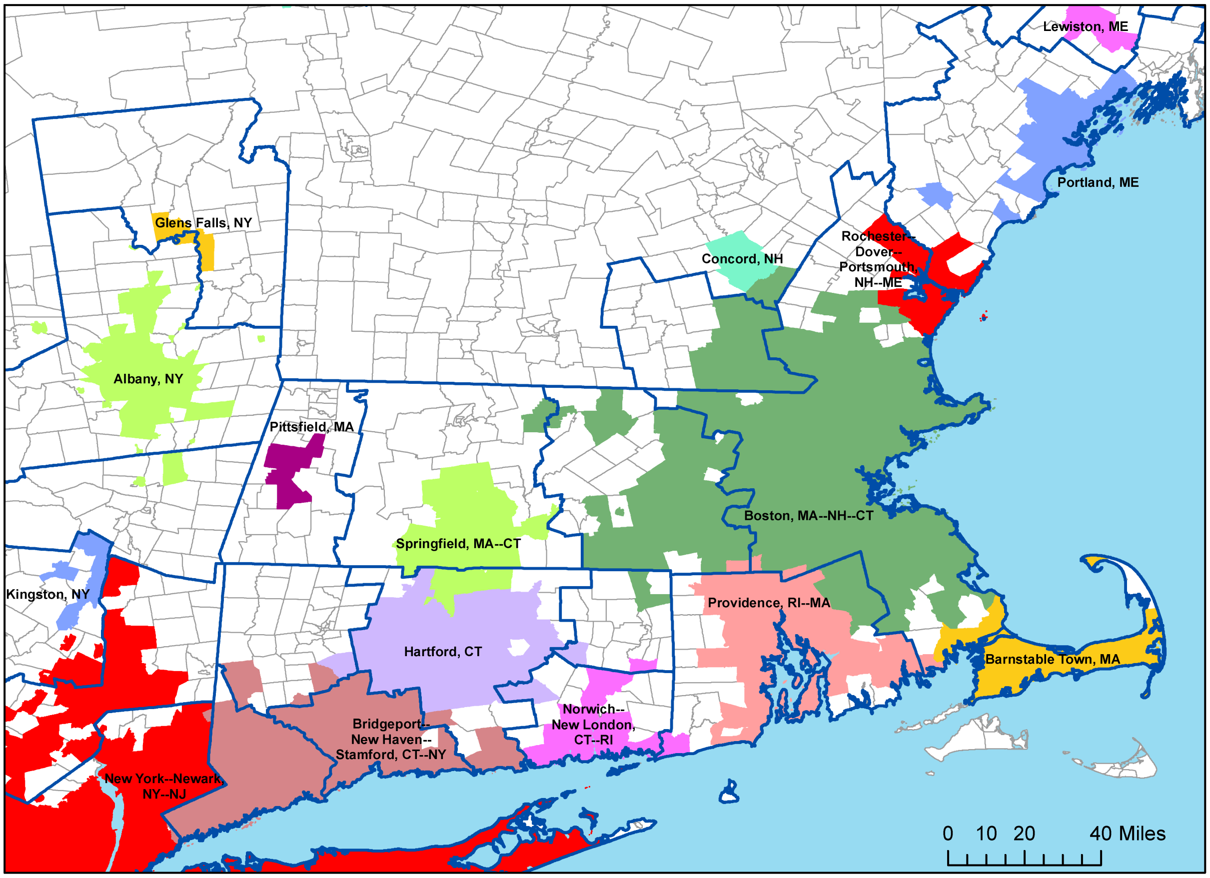
Notes: Blue lines demarcate the borders of metropolitan CBSAs. Census tracts, demarcated in gray, have population density that is inversely correlated with their land area.

the entire population of the Los Angeles CBSA, most of the population of the Riverside-San Bernadino CBSA, and a significant portion of the population of the Oxnard-Thousand Oaks CBSA. The parameterization for kernel-based metropolitan regions combines the Los Angeles, Oxnard, and Indio KBMAs together, closely corresponding to an implicitly crowd-sourced judgement of "Greater Los Angeles", defined by Wikipedia as the combination of the Los Angeles, Riverside-San Bernadino, and Oxnard-Thousand Oaks CBSAs.

KBMAs along the Atlantic Coast of southern New England are tightly packed. For example, as illustrated in Figure 7, the Boston KBMA adjoins four other KBMAs and lies within 30 miles of four more. It also encompasses the most densely settled portions of three metropolitan CBSAs: Boston, Worcester (directly to its west), and Manchester-Nashua (to its northwest).

Figure 8 shows the kernel-based urban areas in the same footprint, which fragment

Figure 7: KBMAs in Southern New England



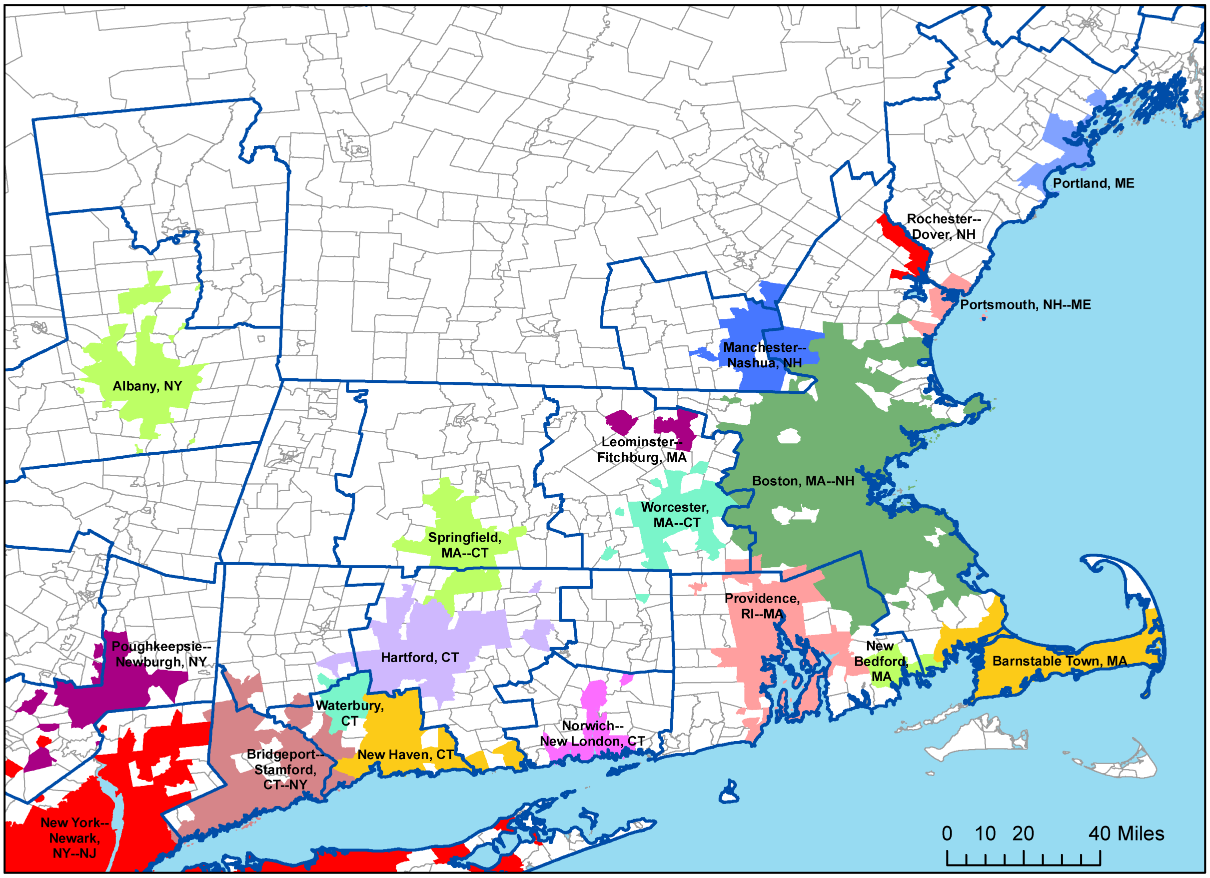
Notes: Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders.

many KBMAs. For example, the Boston KBMA fragments into four separate kernel-based urban areas: Boston, Worcester, Leominster–Fitchburg, and Manchester–Nashua. Other KBMAs "disappear." Specifically, the built-out kernels corresponding to the Kingston, Glens Falls, Pittsfield, Concord, and Lewiston KBMAs no longer have population that exceeds 50,000 and so fail to qualify as kernel-based urban areas.[11] The median land area of the surviving 346 kernel-based urban areas is 34 percent smaller than the median land area of KBMAs.

Conversely, the kernel-based metropolitan regions in the same footprint, shown in Figure 9, consolidate many KBMAs and vastly expand them outward to encompass swathes of rural area. For example, the Boston kernel-based metropolitan region spans five KBMAs:

[11] For each of the three parameterizations, a workbook with detailed variables for all built-out kernels is available from the paper's webpage.

Figure 8: Kernel-Based Urban Areas in Southern New England



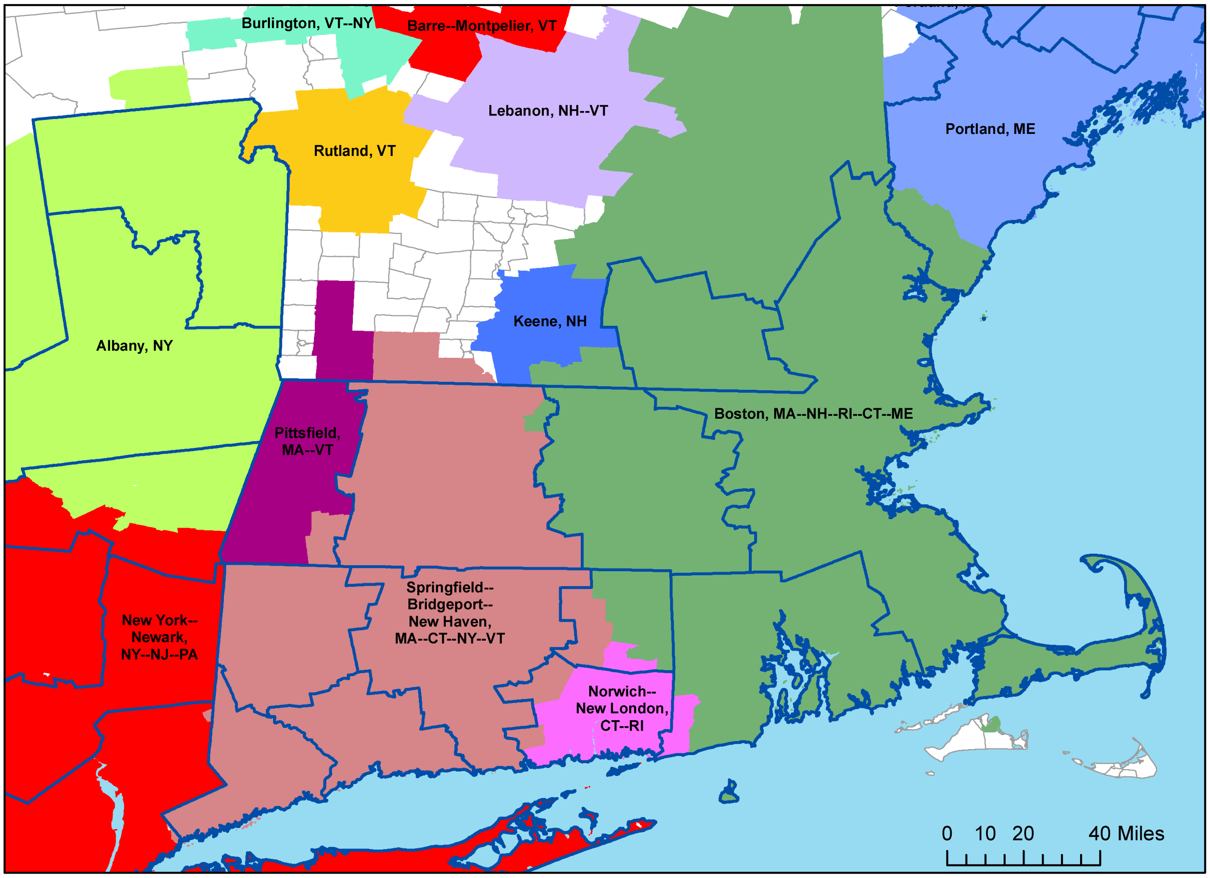
Notes: Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders.

Boston, Providence, Barnstable Town, Rochester-Dover-Portsmouth, and Concord; its land area is more than twice the aggregate of the five KBMAs. The median land area of kernel-based metropolitan regions is more than 16 times that of KBMAs.

4.2 Comparison to Metropolitan CBSAs

Table 4 compares KBMAs with the metropolitan CBSA in which they have their most populous core. This matching is straightforward for the majority of KBMAs but less obvious for others. For example, the criterion implies comparing both the Boston and Rochester–Dover–Portsmouth KBMAs to the Boston–Cambridge–Quincy CBSA. In consequence, some metropolitan CBSAs, such as Worcester and Manchester–Nashua, have no KBMA compared against them. A complementary appendix table reciprocally compares

Figure 9: Kernel-Based Metropolitan Regions in Southern New England



Notes: Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders.

metropolitan CBSAs with the KBMA that has the most populous core in them.

Most KBMAs have population moderately below that of their comparison CBSA, employment modestly below it, and land area far below it. The median ratio of KBMA size to metropolitan CBSA size is 0.72 for population, 0.87 for employment, and 0.12 for land area. But some KBMAs that have cores in multiple metropolitan CBSAs, such as the Boston KBMA, considerably exceed their comparison CBSA in population.

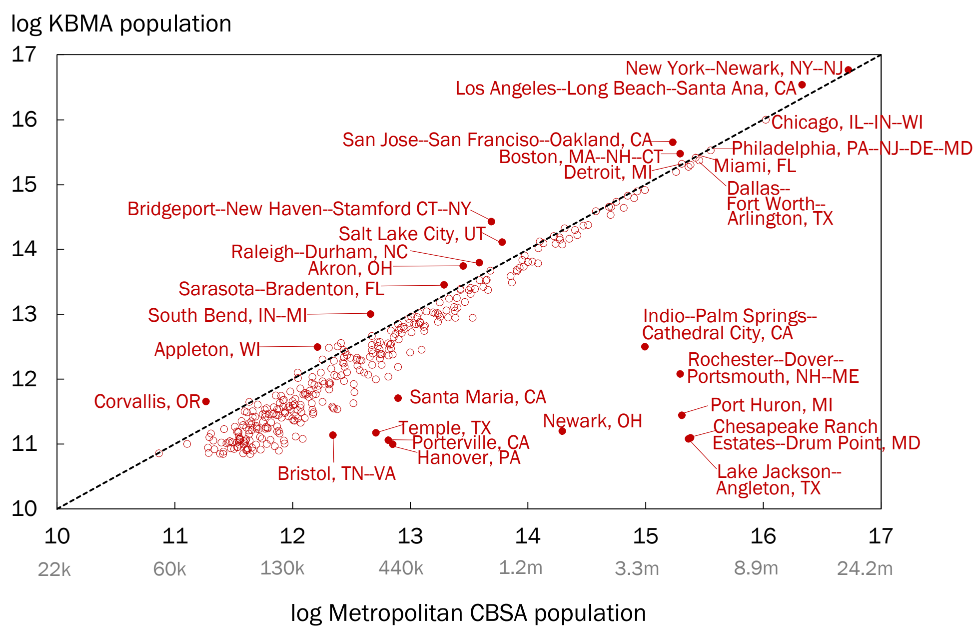
Figure 10 correspondingly plots the population of KBMAs against their comparison CBSAs. All of the labeled KBMAs with solid markers that have population above their comparison CBSA include at least one core that anchors a different CBSA. All of the labeled KBMAs with solid markers that have population below their comparison are paired to a CBSA in which another KBMA also has a core.

Table 4: KBMAs Compared to Matched Metropolitan CBSAs

KBMA	cores	metro CBSAs in which a core	metro CBSAs in which a tract	comparison metropolitan CBSA	size relative to comparison metro CBSA: pop	emp	land	share of KBMA in comparison metro CBSA: pop	emp	land
Kansas City, MO--KS	3	1	1	Kansas City, MO-KS	0.82	0.90	0.14	1.00	1.00	1.00
Los Angeles--Long Beach--Santa Ana, CA	14	4	4	Los Angeles-Long Beach-Santa Ana, CA	1.23	1.17	0.83	0.81	0.85	0.56
Indio--Palm Springs--Cathedral City, CA	2	1	1	Riverside-San Bernardino-Ontario, CA	0.08	0.10	0.01	1.00	1.00	1.00
Oxnard, CA	3	1	1	Oxnard-Thousand Oaks-Ventura, CA	0.55	0.58	0.12	1.00	1.00	1.00
Boston, MA--NH--CT	6	4	5	Boston-Cambridge-Quincy, MA-NH	1.20	1.18	1.14	0.78	0.80	0.62
Rochester--Dover--Portsmouth, NH--ME	2	2	2	Boston-Cambridge-Quincy, MA-NH	0.04	0.04	0.08	0.83	0.87	0.72
Providence, RI--MA	2	3	3	Providence-New Bedford-Fall River, RI-MA	0.88	0.90	0.60	0.98	0.99	0.95
minimum	1	0	0		0.01	0.01	0.00	0.30	0.40	0.20
10th percentile	1	1	1		0.49	0.64	0.03	0.88	0.92	0.82
median	1	1	1		0.72	0.87	0.12	1.00	1.00	1.00
90th percentile	3	2	2		0.97	0.99	0.39	1.00	1.00	1.00
maximum	15	6	6		2.09	1.95	2.26	1.00	1.00	1.00
aggregate (all KBMAs relative to all metropolitan CBSAs)					0.88	0.94	0.15	0.99	0.99	0.96

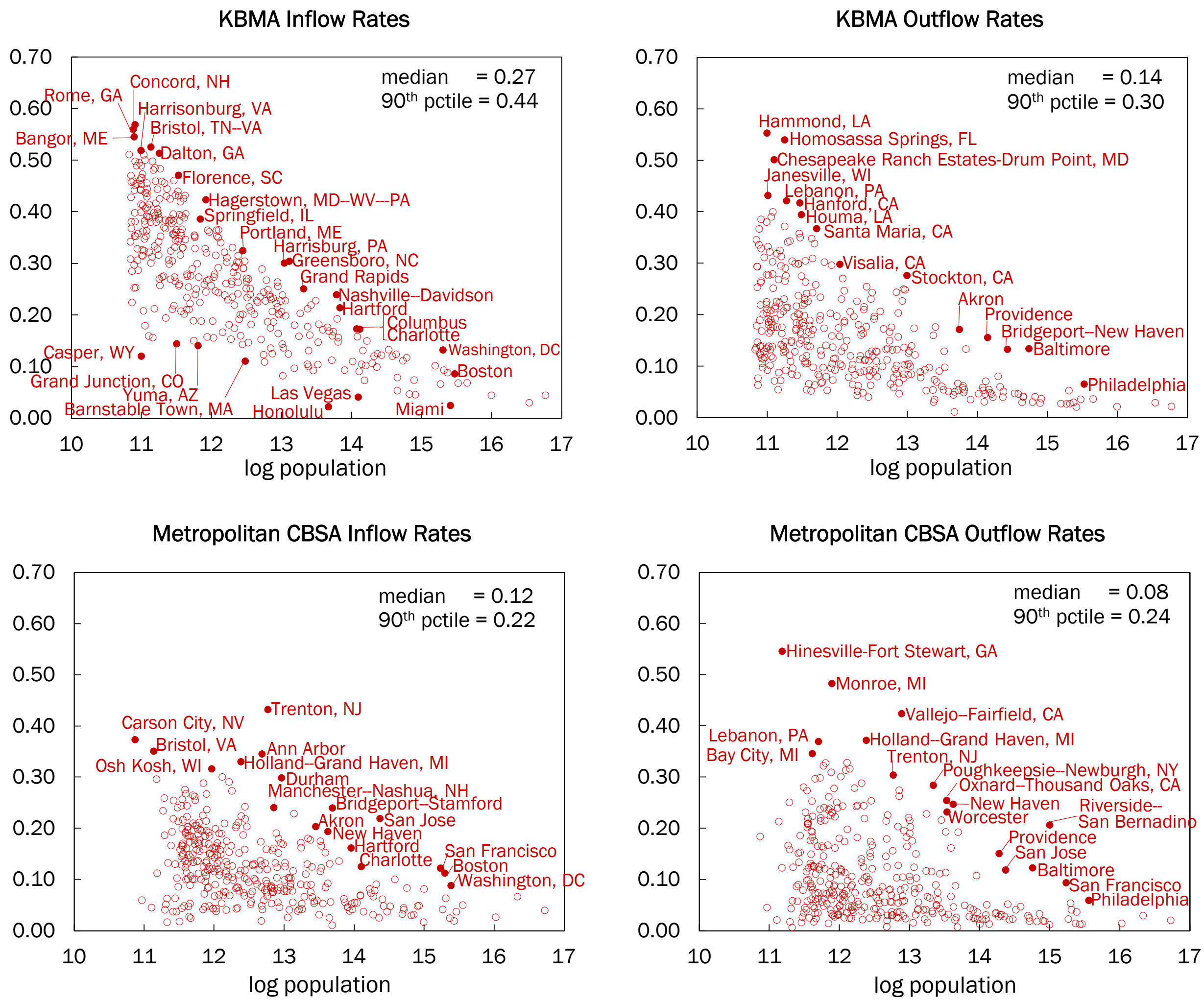
Notes: Each KBMA is compared to the metropolitan CBSA in which it has its most populous core. The summary statistics on relative size and overlap exclude the 19 KBMAs whose largest core is in a micropolitan CBSA. The aggregate statistics compare all KBMAs with all metropolitan CBSAs. A complementary appendix table compares metropolitan CBSAs with the KBMA that has the most populous core in them.

Figure 10: Population of KBMAs versus Comparison Metropolitan CBSA



Notes: Each KBMA is compared to the metropolitan CBSA in which it has its most populous core. The figure excludes the 19 KBMAs whose largest core is in a micropolitan CBSA.

Figure 11: Commuting Flows: KBMAs and Metropolitan CBSAs

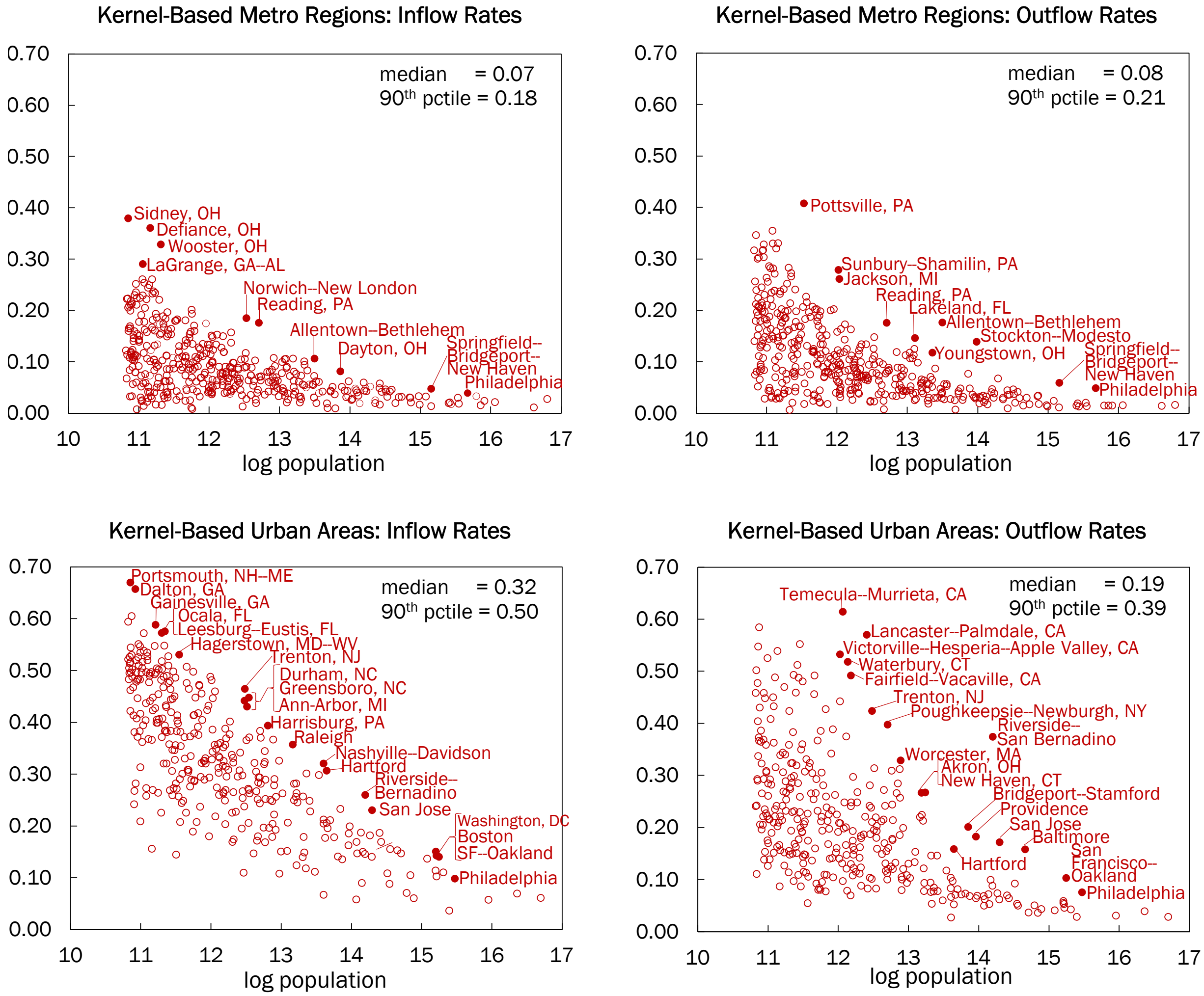


Notes: Inflow rates are measured as a share of employment; outflows rates, as a share of residents who are employed.

4.3 Commuting

Figure 11 shows commuting inflow and outflow rates for KBMAs and metropolitan CBSAs. As illustrated in an appendix figure, the commuting inflows overwhelmingly originate from nearby census tracts that are not attached to a KBMA; commuting outflows split about equally to unattached tracts and to other KBMAs. Measured by the median and 90th percentile rates, KBMAs underperform metropolitan CBSAs in encompassing commuting, unsurprising in the context of the KBMA parameterization's tradeoff to exclude sparsely settled land.

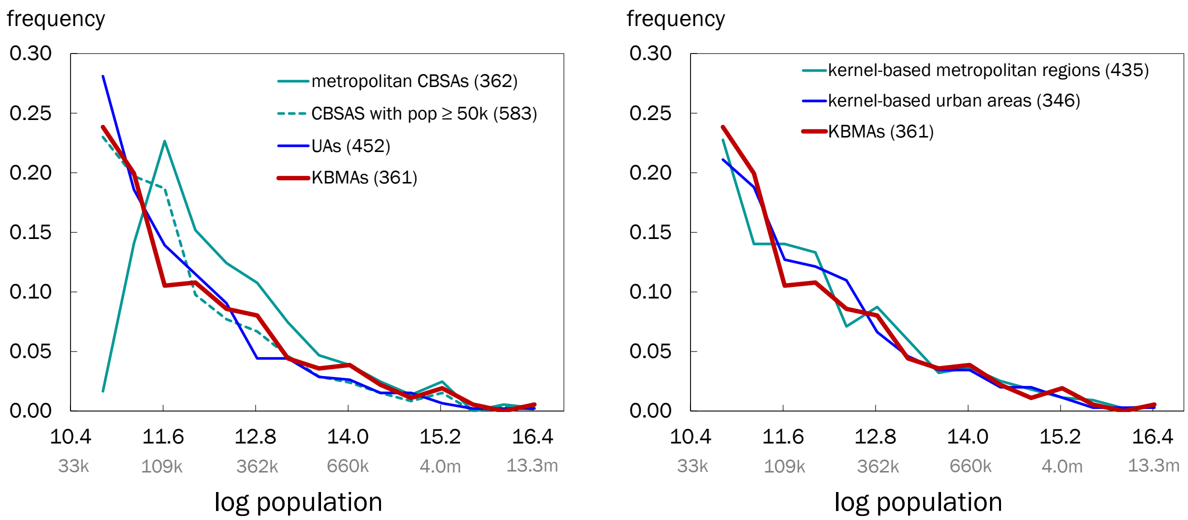
Figure 12: Commuting Flows: Kernel-Based Metropolitan Regions and Kernel-Based Urban Areas



Notes: Inflow rates are measured as a share of employment; outflow rates, as a share of residents who are employed. The Tracy, CA kernel-based urban area, which has log population of 10.9, has a commuting outflow rate of 0.76, above the displayed range.

KBMAs are more competitive by other measures of commuting. The KBMA flow rates have upper envelopes that decline more steeply with population, reflecting that KBMAs more successfully hold down flows for large metropolitan areas. For example, the 90th-percentile inflow rate for KBMAs with population above 500,000 is 0.20, only modestly above the 90th-percentile inflow rate of 0.17 for metropolitan CBSAs with population above 500,000. Even more competitive, the 90th percentile outflow rate for the same set of KBMAs is 0.13, below the outflow rate of 0.17 for the same set of metropolitan CBSAs.

Figure 13: The Distribution of Metropolitan Population



Notes: The horizontal axes measure the logarithm of population at the lower bound of bins with width of 0.4 log points. The lower bound of the leftmost bin corresponds to a population of 50,000 for all histograms. The green-dashed histogram in the left panel describes the set of all CBSAs with population of at least 50,000, both metropolitan and micropolitan.

As illustrated in Figure 12, the kernel-based metropolitan regions considerably hold down commuting flows compared to KBMAs. For example, they achieve median and 90th-percentile inflow rates and a 90th percentile outflow rate below those of metropolitan CBSAs. Commuting flows are considerably higher for kernel-based urban areas.

4.4 Size

Figure 13 shows the histogram of KBMAs' population: on the left compared against those of metropolitan CBSAs and UAs; on the right, against those of kernel-based metropolitan regions and kernel-based urban areas. The KBMA distribution slopes downward across all of the bins, approximately matching the UA distribution. In contrast, the metropolitan CBSA distribution has frequency that increases across the three lowest bins before turning down. The upward-sloping portion reflects that the scale criterion for metropolitan CBSAs applies to their core population rather than total population. To the extent that population buildouts—the ratio of non-core to core population—have a unimodal distribution, relatively few metropolitan CBSAs will have population only moderately above 50,000. In contrast, the more cleanly truncated set of all CBSAs with population above 50,000, both metropolitan and micropolitan, has population frequency that slopes monotonically down, approximately matching the KBMA and UA distributions.

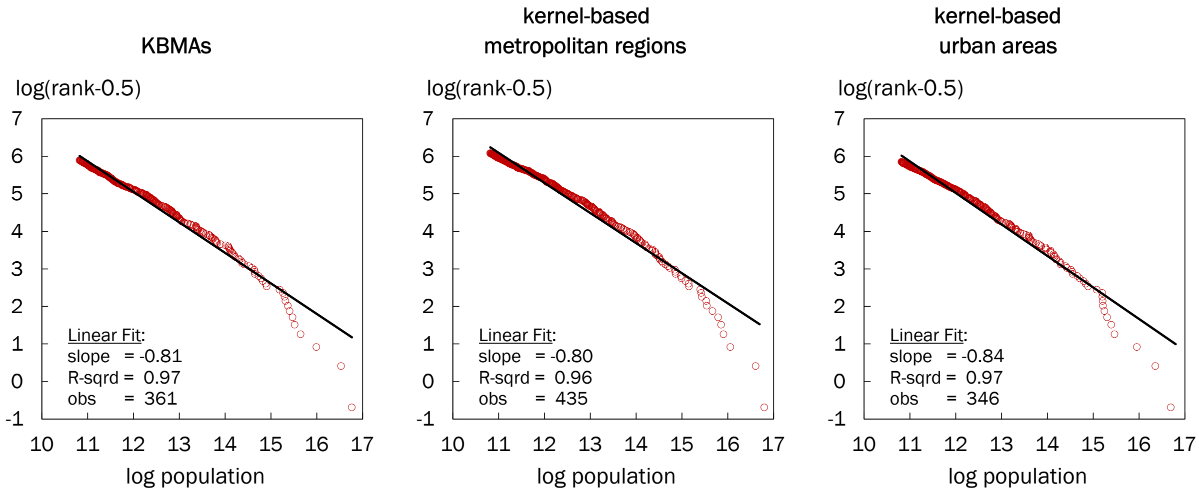
As illustrated in the right panel, the population distributions of the three kernel-based parameterizations approximately match each other. This may seem counterintuitive: During the kernel construction stage, the additional iterations associated with a lower threshold commuting strength and a higher allowed separating distance merge together what otherwise would be separate kernels, shifting the population distribution of built-out kernels rightward. But the shift also pushes more built-out kernels above the 50,000 threshold, contributing to keeping the population distribution approximately unchanged. Equally important, as described previously, the iterative calculation of pairwise commuting strength avoids constructing long chains of large cores, limiting the rightward shift at the top of the population distribution. In aggregate, the 435 kernel-based metropolitan regions have 29 percent more population than the 361 KBMAs; the 346 kernel-based urban areas have 9 percent less population.

Figure 14 plots the logarithm of the population rank of observations for each of the kernel-based delineations against the logarithm of their population, a standard benchmark for local delineations (e.g., Rosen and Resnick, 1980; Gabaix, 1999; Gabaix and Ioannides, 2004; Soo, 2005; Rozenfeld et al., 2011). The rank–size relationship is expected to be approximately linear for Pareto distributions, with a slope close to -1 if the distribution is Zipf, a specialization of Pareto (Gabaix, 1999).[12] Notwithstanding R-squared values very close to 1, the plots are visibly concave.

As emphasized by Eeckhout (2009), the logarithmic rank–size relationship is peculiar, including distorting the visual relationship across the largest observations. Consistent with this, the high R-squared values of linear fits are misleading, among other reasons because the relationship is monotonic by construction. As reported in the appendix, the Monte Carlo simulations prescribed by Gabaix and Ioannides (2004) cannot reject that the

[12]A Pareto distribution has CDF, $\mathrm{F}(S) = 1 - (\widetilde{S}/S)^{\zeta}$ for size $S \geq \widetilde{S}$, $\zeta > 0$. It specializes to a Zipf's distribution when the shape parameter, ζ, equals 1.

Figure 14: Population Rank vs. Size



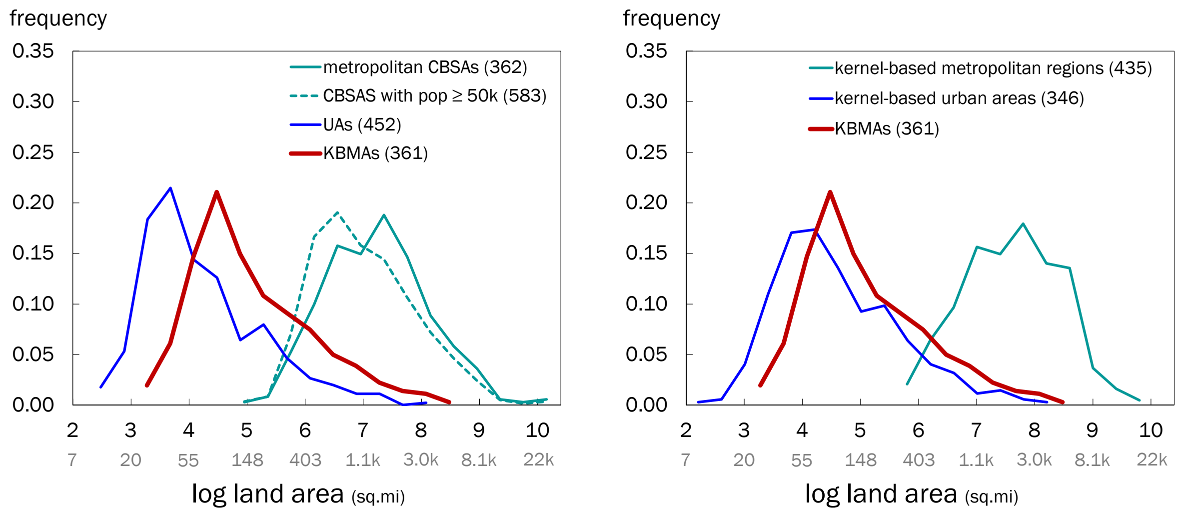
Note: Subtracting 0.5 from the rank improves the fit in the presence of small-sample bias (Gabaix and Ibragimov, 2011).

largest 50 observations of each of the kernel-based delineations are drawn from a Pareto distribution. But they do reject Pareto at the 0.05 level for the largest 75 observations of each. Similarly, they reject Pareto for the largest 75 metropolitan CBSAs and UAs. And they reject it across a much wider range of kernel-based parameterizations, including permutations that set separate values of σ and separate values of δ for the kernel and buildout stages (Rappaport and Humann, 2023).

A benchmark alternative hypothesis is that population distributions are lognormal over their entire range of observations, spanning from very low levels of population to the largest. For example, Eeckhout (2004) shows that the union of municipalities and census designated places in 2000, which includes locations with population as low as 1, is well approximated by a lognormal population distribution with peak frequency at a population close to 1,400.[13] One reason lognormal is important is that it is asymptotically implied by stochastic processes that satisfy Gibrat's law, having growth rates that are uncorrelated with initial levels (Eeckhout, 2004; Lee and Li, 2013).

[13]Census designated places are statistical entities consisting of a closely settled, locally recognized concentration of population that is identified by name. Designation relies heavily on the opinions of local residents, organizations, and officials (U.S. Census Bureau, 1997*a*). This reliance induces a selection bias, specifically that local support for designation is likely to be stronger for locations with larger population.

Figure 15: Distribution of Metropolitan Land Area



Notes: The horizontal axes measure the logarithm of land area of the lower bound of bins with width of 0.4 log points. For each histogram, the lower bound of the leftmost bin corresponds to the observation with minimum land area.

We are skeptical. As illustrated in the appendix, the frequency distribution of the union of UAs and UCs is convex downward sloping from its minimum population of 2,500. The 3,610 UAs and UCs in 2000 together accounted for only 79 percent of the U.S. population, at minimum requiring tens of thousands of smaller clusters to encompass the remaining 59 million residents. Similarly, the union of municipalities and census designated places in Eeckhout (2004) accounted for only 71 percent of the U.S. population (Levy, 2009). Bolstering our skepticism, more recent research finds that population growth rates of various geographical delineations were historically correlated with population levels (Holmes and Lee, 2010; Dittmar, 2021; Michaels, Rauch and Redding, 2012; Desmet and Rappaport, 2017; Rappaport, 2018). For example, Rappaport (2018) documents that intermediate-sized locations grew fastest during recent decades.

Figure 15 compares histograms of land area. Unsurprisingly, the KBMA land distribution is located to the right of the UA distribution and considerably to the left of the metropolitan CBSA distribution. More surprising, the KBMA distribution represents a near parallel rightward shift from the UA distribution, including preserving the latter's pronounced rightward skew. The land distribution of kernel-based urban areas, shown in the right panel, is also a parallel shift rightward from UAs but less far than KBMAs. In aggregate, the 346 kernel-based urban areas have 39 percent less land area than the 361 KBMAs.

The land distribution of kernel-based metropolitan regions compresses that of metropolitan CBSAs, including truncating the latter's right tail. For example, the Dallas–Forth Worth–Arlington kernel-based metropolitan region has the largest land area, 20,500 square miles, comfortably below the land areas of the Riverside-San Bernadino and Anchorage metropolitan CBSAs. The least expansive metropolitan region occupies 334 square miles, more than twice that of the least expansive metropolitan CBSA. In aggregate, the 435 kernel-based metropolitan regions have land area more than 10 times that of the 361 KBMAs.

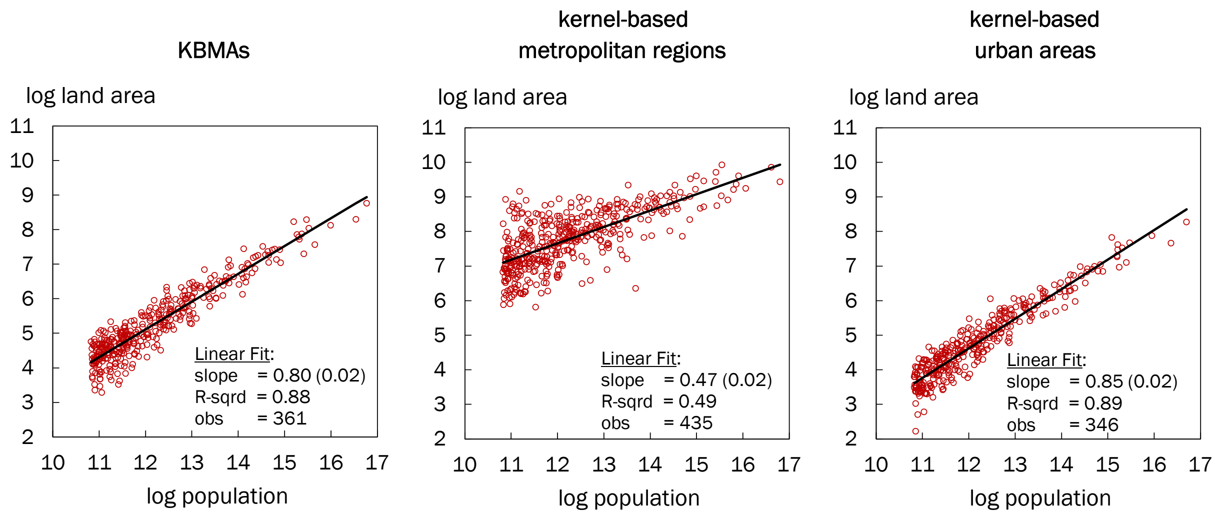
Figure 16 plots the logarithm of land area against the logarithm of population for the three kernel-based parameterizations. The fitted linear relationship for each has a slope less than 1, indicating that higher levels of population are associated with less-than-proportionally higher levels of land area. The slope corresponds to the implicit elasticity of land area with respect to population, estimated to be 0.80 for KBMAs.[14]

The estimated elasticity is lower and the fit less tight for kernel-based metropolitan regions, reflecting that they encompass a higher share of lightly settled land and so can more easily densify to accommodate population growth. Conversely, the estimated elasticity is higher and the fit tighter for kernel-based urban areas, reflecting their lower capacity to densify. As illustrated in the appendix, the estimated implicit elasticities for metropolitan CBSAs and UAs are similar to those of kernel-based metropolitan regions and kernel-based urban areas, respectively.

The tight, less-than-proportional relationship between land area and population intuitively suggests that centripetal forces are pulling residents and establishments to locate near a central location, for example to take advantage of clustered employment (Alonso, 1964; Mills, 1967; Muth, 1969) or amenities. The pull of these centripetal forces may strengthen with distance as the marginal opportunity cost of locating a bit farther away

[14] For comparison, Ahlfeldt and Pietrostefani (2019) estimate that the implicit elasticity of land area to population is 0.57 across 70 U.S. Functional Urban Areas and to range from 0.29 to 0.85 across Functional Urban Areas in each of the additional 13 countries for which they have sufficient data. Combes, Duranton and Gobillon (2019) estimate that the implicit elasticity of land area to population is approximately 0.7 across urban areas in France.

Figure 16: Land Area versus Population



Note: Standard errors for the slope coefficient are reported in parentheses.

bends upward (Rappaport, 2016). Consistent with this, as illustrated in the appendix, the relationship of land to population for the union of UAs and UCs is concave, statistically significant at the 0.01 level. The fitted elasticity declines from 1.01 at the smallest observation to 0.71 at the largest. This concavity probably also reflects topography, which constrains expansion in many large metropolitan areas (Saiz, 2010).

5 Conclusion

Parameterizations of our kernel-based algorithm better match a broad definition of metropolitan areas compared to metropolitan CBSAs and other existing delineations. The KBMA parameterization, which balances encompassing commuting flows and excluding sparsely settled land, is likely to be appropriate for most purposes, including for planning, regulation, and research. The more expansive parameterization for kernel-based metropolitan regions and the more compact parameterization for kernel-based urban areas are likely to be appropriate for other purposes. Datasets for each are available for download. The algorithm's computer code will also be made available for download, allowing for extensive customization, including independently setting the commuting strength parameter and the separating distance parameter for the kernel and buildout stages.

A priority for future research is constructing kernel-based delineations for more recent years. Doing so will realize another advantage of KBMAs compared to CBSAs, the ability to decompose metropolitan growth into intensive and extensive margins (i.e., within and outside starting footprints). CBSAs poorly do so because settlement varies hugely within its county building blocks. Similarly, kernel-based delineations can closely track the geographic expansion of nearby metropolitan areas up against each other, in some cases merging into a single metropolitan area but not in others.

On the other hand, using census tracts as building blocks limits data availability. To be sure, the Census Bureau publishes tabulated tract data for some of its products, including decennial censuses and the American Community Survey. Conversely, anonymity requirements limit the availability of microdata for most counties and metropolitan CBSAs, lessening the disadvantage of using tracts as building blocks for some research purposes.

As argued by Duranton (2021), appropriate metropolitan delineations are required to address fundamental urban questions. KBMAs and other kernel-based parameterizations can contribute to doing so.

References

Ahlfeldt, Gariel M., and Elisabetta Pietrostefani. 2019. "The Economic Effects of Density: a Synthesis." *Journal of Urban Economics*, 111.

Alonso, William. 1964. *Location and Land Use.* Cambridge:Harvard University Press.

Arribas-Bel, Daniel, Miquel-Angel Garcia-Lopez, and Elisabet Viladecans-Marsal. 2021. "Building(s and) Cities: Delineating Urban Areas with a Machine Learning Algorithm." *Journal of Urban Economics*, 125.

Autor, David, David Dorn, and Gordon Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103: 2121–2168.

Berry, Brian J.L., Peter G. Goheen, and Harry Goldstein. 1969. "Metropolitan Area Definition: A Re-Evaluation of Concept and Statistical Practice." Bureau of Census.

Brezzi, Monica, Mario Piacenti, Konstantin Rosina, and Daniel Sanchez-Serra. 2012. "Redefining Urban Areas in OECD Countries." In *Redefining "Urban": A New Way to Measure Metropolitan Areas*. Chapter 1. Paris:OECD Publishing.

Bureau of Transportation Statistics. 2005. "Census Transportation Planning Package (CTPP) 2000."

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where it the land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*, 129: 1553–1623.

Ch, Rafael, Diego A. Martin, and Juan F. Vargas. 2021. "Measuring the Size and Growth of Cities Using Nighttime Light." *Journal of Urban Economics*, 125.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics*, 62: 723–742.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2019. "The Costs of Agglomeration: House and Land Prices in French Cities." *Review of Economic Studies*, 86: 1556–1589.

Congressional Research Service. 2014. "Metropolitan Area Designations by OMB: History, 2010 Standards, and Uses."

de Bellefon, Marie-Pierre, Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, and Clément Gorin. 2021. "Delineating Urban Areas Using Building Density." *Journal of Urban Economics*, 125.

Desmet, Klaus, and Jordan Rappaport. 2017. "The Settlement of the United States, 1800-2000: The Long Transition towards Gibrat's Law." *Journal of Urban Economics*, 98: 50–68.

Dijkstra, Lewis, and Hugo Poelman. 2012. "Cities in Europe: The New OECD-EC Definition." European Commision, Directorate-General for Regional and Urban Policy.

Dijkstra, Lewis, Aneta J. Florczyk, Sergio Freire, Thomas Kemper, Michele Melchiorri, Martino Pesaresi, and Marcello Schiavina. 2021. "Applying the Degree of Urbanisation to the Globe: a New Harmonised Definition Reveals a Different Picture of Global Urbanisation." *Journal of Urban Economics*, 125.

Dijkstra, Lewis, Hugo Poelman, and Paolo Veneri. 2019. "The EU-OECD Definition of a Functional Urban Area." OECD.

Dingel, Jonathan I., Antonio Miscio, and Donald R. Davis. 2021. "Cities, Lights, and Skills in Developing Economies." *Journal of Urban Economics*, 125.

Dittmar, Jeremiah. 2021. "Cities, Markets, and Growth: The Emergence of Zipf's Law." Institute for Advanced Study.

Duranton, Gilles. 2015. "Delineating Metropolitan Areas: Measuring Spatial Labour Market Networks Through Commuting Patterns." In *The Economics of Interfirm Networks.* , ed. Iichiro Uesugi Tsutomu Watanabe and Arito Ono, 107–133. Tokyo:Springer.

Duranton, Gilles. 2021. "Classifying Locations and Delineating Space: An Introduction." *Journal of Urban Economics*, 125.

Economic Research Service. 2012. "2000 Commuting Zones." data retrieved on 3/24/2020 from https://www.ers.usda.gov/webdocs/DataFiles/48457/cz00_eqv_v1.xls.

Eeckhout, Jan. 2004. "Gibrat's Law for (All) Cities." *American Economic Review*, 94: 1429–1451.

Eeckhout, Jan. 2009. "Gibrat's Law for (All) Cities: Reply." *American Economic Review*, 99: 1676–1683.

Eurostat, OECD, United Nations, and The World Bank. 2021. *Applying the Degree of Urbanisation: A Methodological Manual to Define Cities, Towns and Rural Areas for International Comparisons.* Paris and Brussels:OECD Publishing.

Foote, Andrew, Mark J. Kutzbach, and Lars Vilhuber. 2017. "Recalculating...: How Uncertainty in Local Labor Market Definitions Affects Empirical Findings." U.S. Census Bureau, Center for Economic Studies.

Gabaix, Xavier. 1999. "Zipf's Law for Cities: an Explanation." *Quarterly Journal of Economics*, 114.

Gabaix, Xavier, and Rostam Ibragimov. 2011. "Rank - 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents." *Journal of Business and Economic Statistics.*

Gabaix, Xavier, and Yannis M. Ioannides. 2004. "The Evolution of City Size Distributions." In *Handbook of Regional and Urban Economics.* , ed. J.V Henderson and J.F. Thisse, 173–204. Amsterdam:North Holland.

Galdo, Virgilio, Yue Li, and Martin Rama. 2021. "Identifying Urban Areas by Combining Human Judgement and Machine Learning: An Application to India." *Journal of Urban Economics*, 125.

Gardner, Todd. 1999. "Metropolitan Classification for Census Years before World War II." *Historial Methods*, 32.

Gardner, Todd. 2021. "Changes in Metropolitan Area Definition, 1910-2010." U.S. Census Bureau.

Glaeser, Edward L., and David C. Mare. 2001. "Cities and Skills." *Journal of Labor Economics*, 19: 316–342.

Green, Richard K., Stephen Malpezzi, and Stephen K. Mayo. 2005. "Metropolitan-Specific Estimates of the Price Elasticity of the Supply of Housing and Their Sources." *American Economic Review*, 95: 334–339.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy*, 118: 536–598.

Guren, Adam M., Alisdair McKay, Emi Nakamura, and Jón Steinson. 2021. "Housing Wealth Effects." *The Review of Economic Studies*, 88: 669–707.

Holmes, Thomas J., and Sanghoon Lee. 2010. "Cities as Six-by-Six-Mile Squares: Zipf's Law?" In *Agglomeration Economics.* , ed. Edward L. Glaeser, 105–131. Chicago and London:University of Chicago Press.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79: 211–251.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider. 2015. "The Housing Market(s) of San Diego." *American Economic Review*, 105: 1371–1407.

Lee, Sanghoon, and Qiangi Li. 2013. "Uneven Landscapes and City Size Distributions." *Journal of Urban Economics*, 78: 19–29.

Levy, Moshe. 2009. "Gibrat's Law for (All) Cities: Comment." *American Economic Review*, 99: 1672–1675.

Mian, Atif, and Amir Sufi. 2012. "The Effects of Fiscal Stimulus: Evidence from the Cash for Clunkers Program." *Quarterly Journal of Economics*, 127: 1107–1142.

Mian, Atif, and Amir Sufi. 2014. "What Explains the 2007-2009 Drop in Employment?" *Econometrica*, 82: 2197–2223.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding. 2012. "Urbanization and Structural Transformation." *Quarterly Journal of Economics*, 127: 535–586.

Mills, Edwin S. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *American Economic Review*, 57: 197–210.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Densities." *Amerian Economic Review*, 108: 3855–3890.

Moretti, Enrico. 2004. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review*, 94: 656–690.

Muth, Richard F. 1969. *Cities and Housing.* Chicago:University of Chicago Press.

Office of Management and Budget. 1990. "Revised Standards for Defining Metropolitan Areas in the 1990's; Notice." *Federal Register: The Daily Journal of the United States*, 55: 12154–12160.

Office of Management and Budget. 1998. "Alternative Approaches to Defining Metropolitan and Nonmetropolitan Areas; Notice." *Federal Register: The Daily Journal of the United States*, 63.

Office of Management and Budget. 2000. "Standards for Defining Metropolitan and Micropolitan Statistical Areas; Notice." *Federal Register: The Daily Journal of the United States*, 65: 82228–82238.

Office of Management and Budget. 2003. "Metroplitan Statistical Areas, Micropolitan Statistical Areas, Combined Statistical Areas, New England City and Town Areas, Combined New England City and Town Areas, 2003." *OMB Bulletin No. 03-04*, 82228–82238.

Office of Management and Budget. 2010. "2010 Standards for Delineating Metropolitan and Micropolitan Statistical Areas; Notice." *Federal Register: The Daily Journal of the United States*, 75.

Office of Management and Budget. 2021. "2020 Standards for Delineating Metropolitan and Micropolitan Statistical Areas." *Federal Register: The Daily Journal of the United States*, 86.

Rappaport, Jordan. 2016. "Productivity, Congested Commuting, and Metro Size." Federal Reserve Bank of Kansas City Working Paper.

Rappaport, Jordan. 2018. "The Faster Growth of Larger, Less Crowded Locations." *Economic Review*, 105. Federal Reserve Bank of Kansas City.

Rappaport, Jordan, and McKenzie Humann. 2023. "The Size of U.S. Metropolitan Areas." Federal Reserve Bank of Kansas City Working Paper.

Ratcliffe, Michael. 2015. "A Century of Delineating a Changing Landscape: The Census Bureau'S Urban and Rural Classification, 1910 To 2010." accessed 2/16/2022 from http://www2.census.gov/geo/pdfs/reference/ua/Century_of_Defining_Urban.pdf.

Rosen, Kenneth T., and Mitchel Resnick. 1980. "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy." *Journal of Urban Economics*.

Rozenfeld, Hernan D., Diego Rybski, Xavier Gabaix, and Hernan A. Makse. 2011. "The Area and Population of Cities: New Insights from a Different Perspective on Cities." *American Economic Review*, 101: 2205–2225.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics*, 125: 1253–1296.

Soo, Kwok Tong. 2005. "Zipf's Law for Cities: A Cross Country Investigation." *Regional Science and Urban Economics*, 35: 239–263.

Tolbert, Charles M., and Molly Sizer. 1987. "US Commuting Zones and Labor Market Areas." United States Department of Agriculture.

Tolbert, Charles M., and Molly Sizer. 1996. "US Commuting Zones and Labor Market Areas: a 1990 update." United States Department of Agriculture.

U.S. Census Bureau. 1997*a*. "Census Designated Place (CDP), Program for Census 2000–Final Criteria; Notice." *Federal Register: The Daily Journal of the United States*, 62: 44638–44641.

U.S. Census Bureau. 1997*b*. "Census Tract Program for Census 2000–Final Criteria; Notice." *Federal Register: The Daily Journal of the United States*, 62: 55778–55781.

U.S. Census Bureau. 2002. "Urban Area Criteria for Census 2000; Notice." *Federal Register: The Daily Journal of the United States*, 67: 11663–11670.

U.S. Census Bureau. 2011. "Urban Area Criteria for the 2010 Census; Notice." *Federal Register: The Daily Journal of the United States*, 76: 53030–53043.

U.S. Census Bureau. 2022. "Urban Area Criteria for the 2020 Census–Final Criteria." *Federal Register: The Daily Journal of the United States*, 87: 16706–16715.

Van Nieuwerburgh, Stijn, and Piere-David Weill. 2010. "Why Has Housing Dispersion Gone Up?" *The Review of Economic Studies*, 77: 1567–1606.

Appendices

A. Supplemental Figures

1. Tract-Based Approximations of UAs and UCs
2. Sensitivity of Kernel Construction to the Strength and Distance Parameters
3. Sensitivity of Population, Employment, and Commuting Outflows to the Density Parameter
4. The Population and Land Area of UAs and UCs
5. Population Rank versus Size for Official Delineations
6. Land Area versus Population for Official Delineations
7. Commuting Flows for UAs and Commuting Zones
8. Decomposition of KBMA Commuting Flows

B. Supplemental Tables

1. Monte Carlo Tests for Pareto Population Distributions
2. Comparison of Metropolitan CBSAs with their Corresponding KBMA
3. Metropolitan CBSA Pairs with Strong Cross Commuting

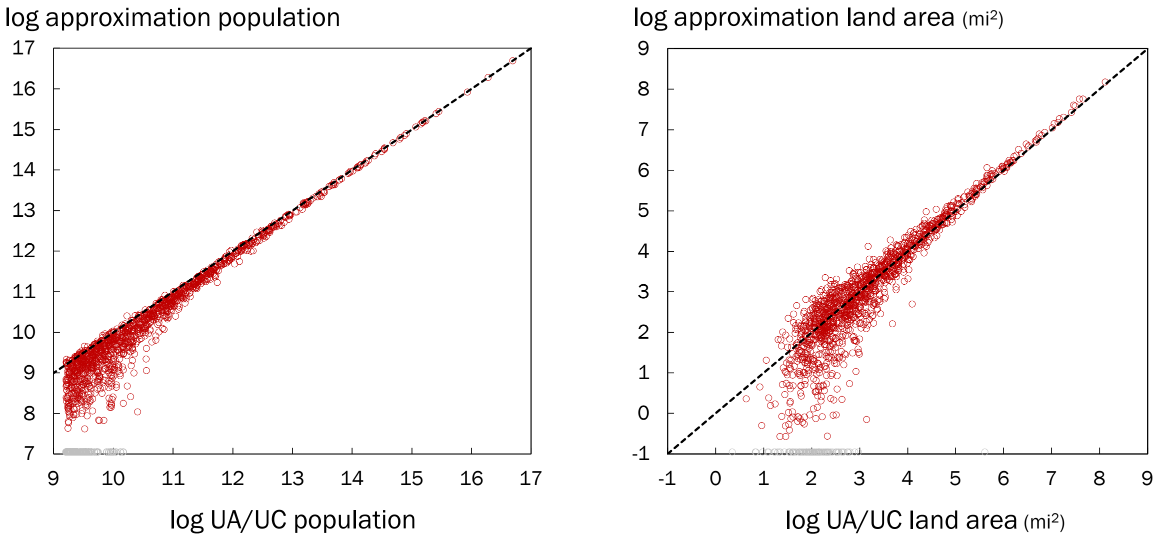
C. Supplemental Maps

1. Commuting Zones Near New York City
2. KBMAs from New York City to Virginia Beach
3. KBMAs in Northern California
4. KBMAs in Florida
5. KBMAs in Texas

D. Online: The files directly linked below are also available from the paper's webpage.

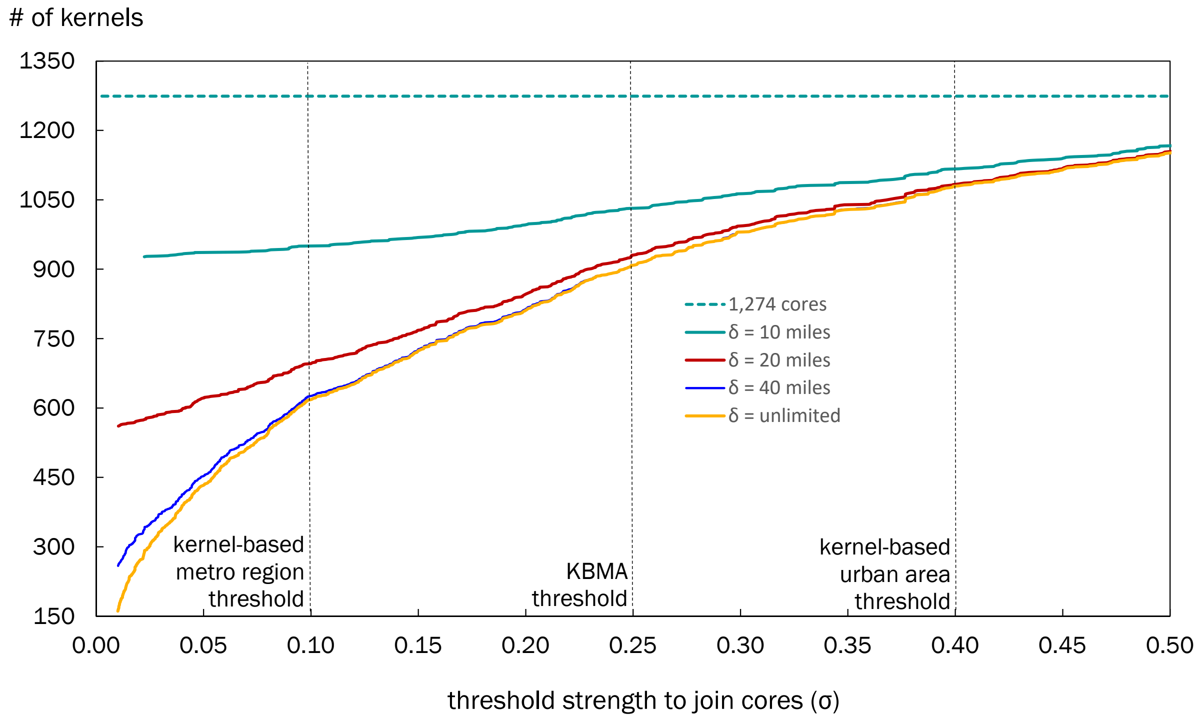
1. Enumerations of
   KBMAs (kernel-based metropolitan areas), KBUAs (kernel-based urban areas), and KBMRs (kernel-based metropolitan regions)
2. Maps of KBMAs, KBUAs, and KBMRs
3. Maps comparing KBMAs, KBUAs, and KBMRs
4. Detailed tables (additional variables, constituent census tracts, cores, built-out kernels, and kernel iterations) for KBMAs, KBUAs, and KBMRs
5. Shape Files for KBMAs, KBUAs, and KBMRs
6. Pairwise flows between cores (prior to kernel construction)
7. Iterative kernel joins as commuting strength drops to 0.01. (underlying data for Supplemental Figure A.2)

Figure A.1: Tract-Based Approximations of UAs and UCs



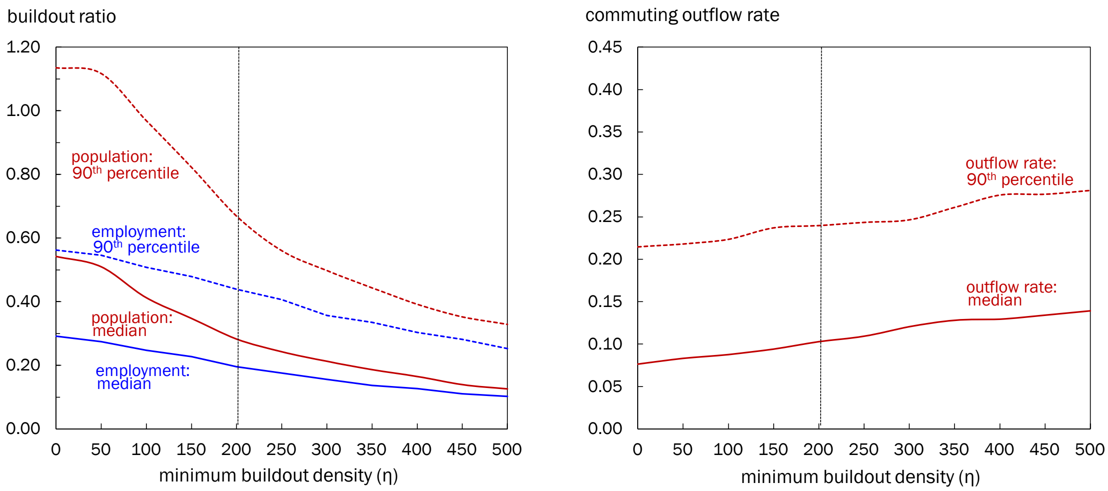
Notes: Census tracts are assigned to the approximation of a UA or a UC with population of at least 10,000 if the shares of their population and land in the UA/UC are at least 0.55 and 0.30, respectively. These thresholds minimize the sum of squared percentage differences of approximation population and land area from their actual values. Dashed lines are drawn where approximation size equals actual. Only 1,274 of the 1,374 qualifying UAs/UCs have at least one tract that meets both thresholds; the gray dots represent the 100 UCs with no tract that meets both. These null approximations only negligibly affect the delineated set of KBMAs. As illustrated by the Kansas City KBMA (Figure 4), many tracts overlapping unapproximated UCs join to a nearby kernel during the buildout. The built-out kernels of most remaining unapproximated UCs would likely fail to meet the 50,000 population threshold to qualify as a KBMA, similar to the 570 actual built-out kernels that fail to qualify.

Figure A.2: Sensitivity of Kernel Construction to the Strength and Distance Parameters



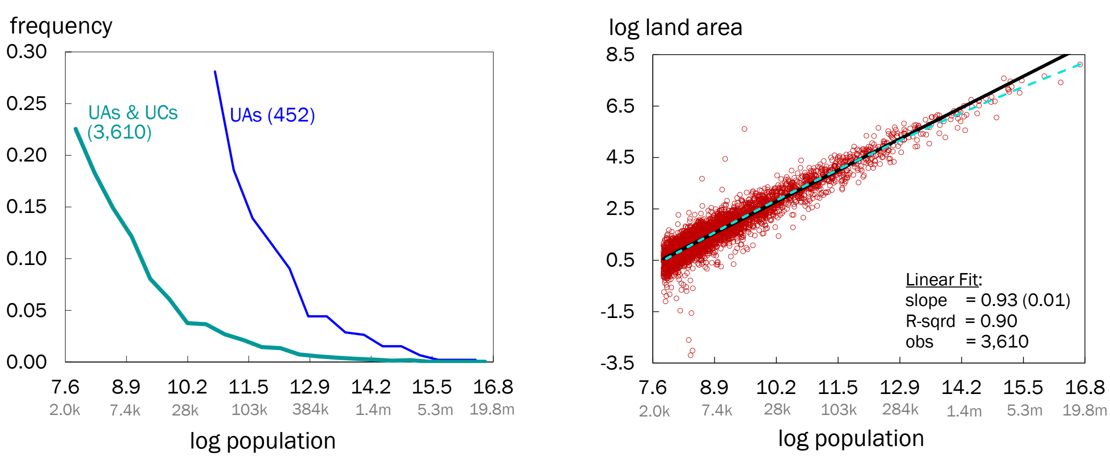
Notes: The kernel-iteration stage of the KBMA algorithm starts with 1,274 separate cores (green dashed line), corresponding to tract-based approximations of UAs and UCs with population above 10,000. Moving from right to left, the figure shows the number of surviving kernels as σ is decreased from 0.50 to 0.01 for each of three distance maximums (green, red, and blue lines) and for unconstrained distance (yellow line). At the KBMA strength threshold, σ=0.25, the 20-mile maximum distance binds for 22 latent joins (vertical distance between the red and yellow lines). At the kernel-based urban area strength threshold, σ=0.40, the 10-mile maximum distance binds for 37 latent joins (vertical distance between the green and yellow lines). At the kernel-based metropolitan region strength threshold, σ=0.10, the 40-mile maximum distance binds for 8 latent joins (distance between the blue and yellow lines). An online workbook documents each of the iterative joins for each of the distance maximums.

Figure A.3: Sensitivity of Population, Employment, and Commuting Outflows to the Density Parameter



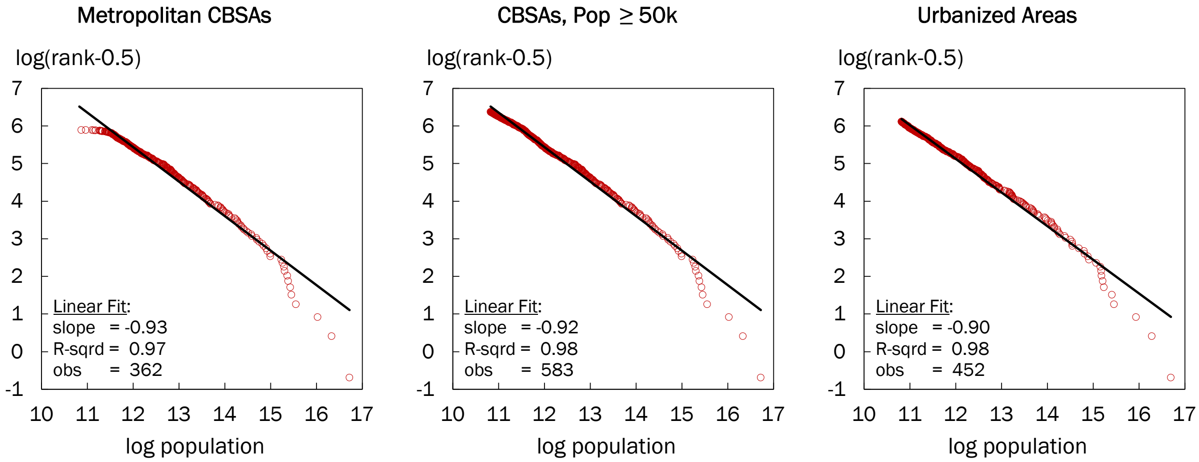
Notes: These complement the land buildout ratios and commuting inflow rates in Figure 3. Left panel shows the median and 90th-percentile ratios of the population and employment in the buildout portion of built-out kernels relative to the kernel portion as η is increased from 0 to 500. The dashed vertical line corresponds to the parameterized KBMA value, $\eta = 200$. The vertical scale is one-tenth that used for the land buildout ratio in Figure 3, reflecting that population and employment are considerably less sensitive to η. The right panel shows the median and 90th percentile rates of commuting outflows, which are less sensitive to η than the corresponding inflow rates shown in Figure 3. For comparability, kernels are restricted to the 302 with population of at least 50,000.

Figure A.4: The Population and Land Area of UAs and UCs



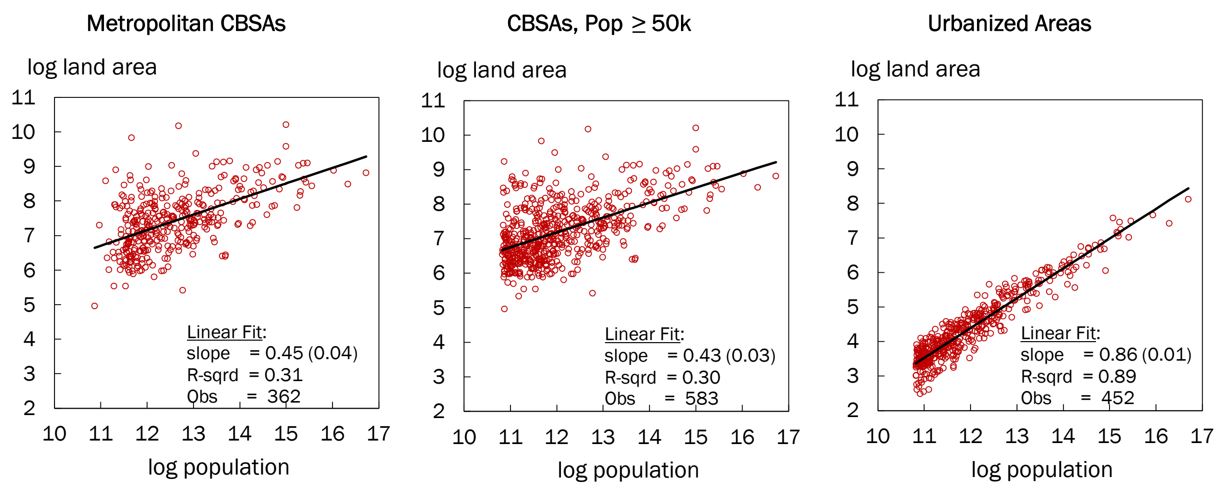
Notes: The left panel shows the population distributions of UAs and the union of UAs and UCs. Its horizontal axis measures the log of population at the lower bound of bins with width of 0.4 log points. The right panel plots the land area of UAs and UCs against their population. The solid and dashed lines respectively represent linear and quadratic fits. The negative estimated quadratic coefficient for the quadratic fit statistically differs from 0 at the 0.01 level. The slope of the quadratic fit declines from 1.01 at the population of the smallest UA/UC to 0.71 at the population of the largest.

Figure A.5: Population Rank versus Size for Official Delineations



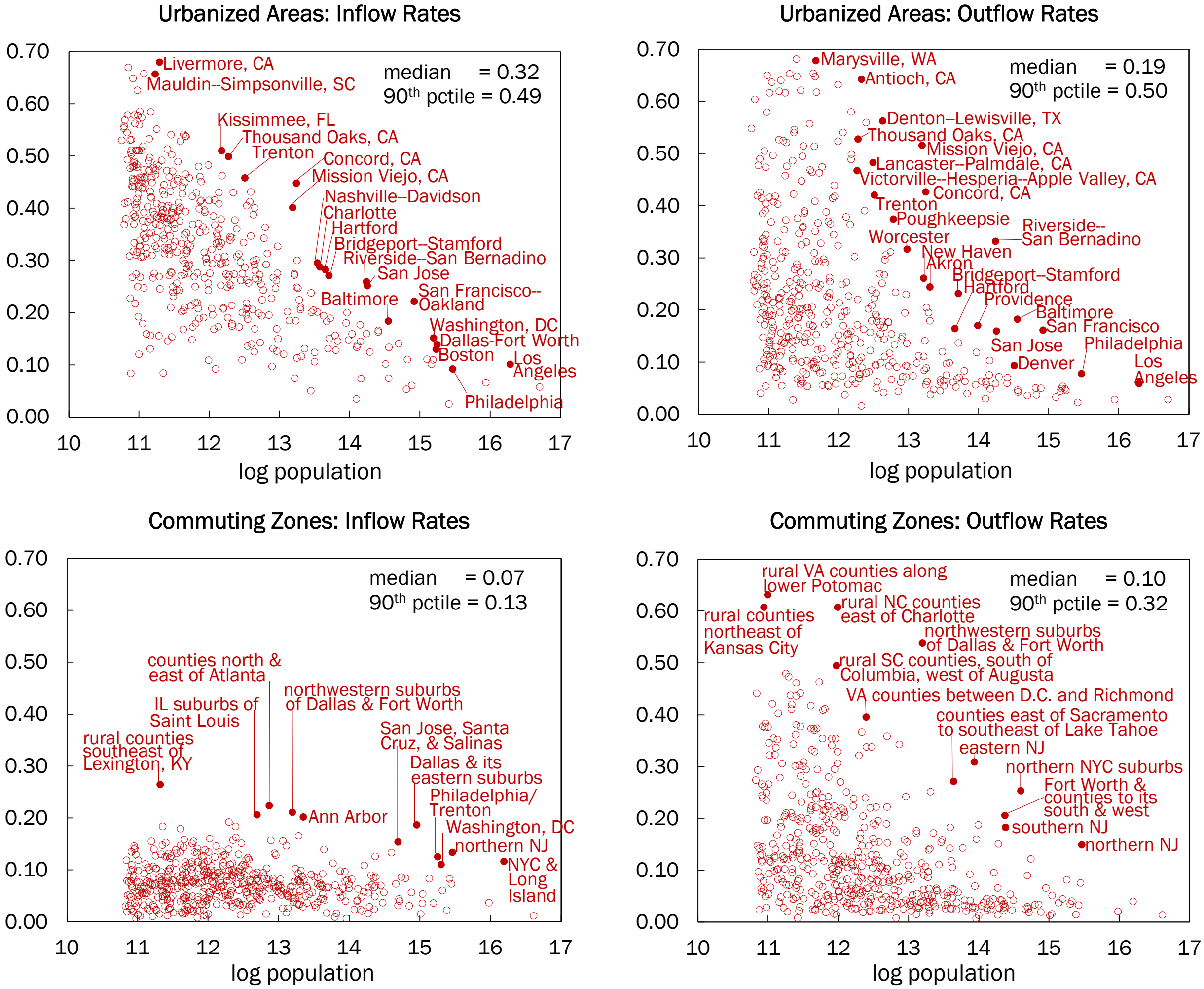
Notes: Subtracting 0.5 from the rank improves the fit in the presence of small-sample bias (Gabaix and Ibragimov, 2011). In the left panel, the sharp concavity of the rank–size scatter at the bottom of the distribution reflects that metropolitan CBSAs are truncated based on the population of their cores rather than their total population. As illustrated in the middle panel, the rank-size scatter is approximately linear at the bottom of the distribution for the set of all CBSAs, metropolitan and micropolitan, cleanly truncated at a total population of 50,000.

Figure A.6: Land Area versus Population for Official Delineations



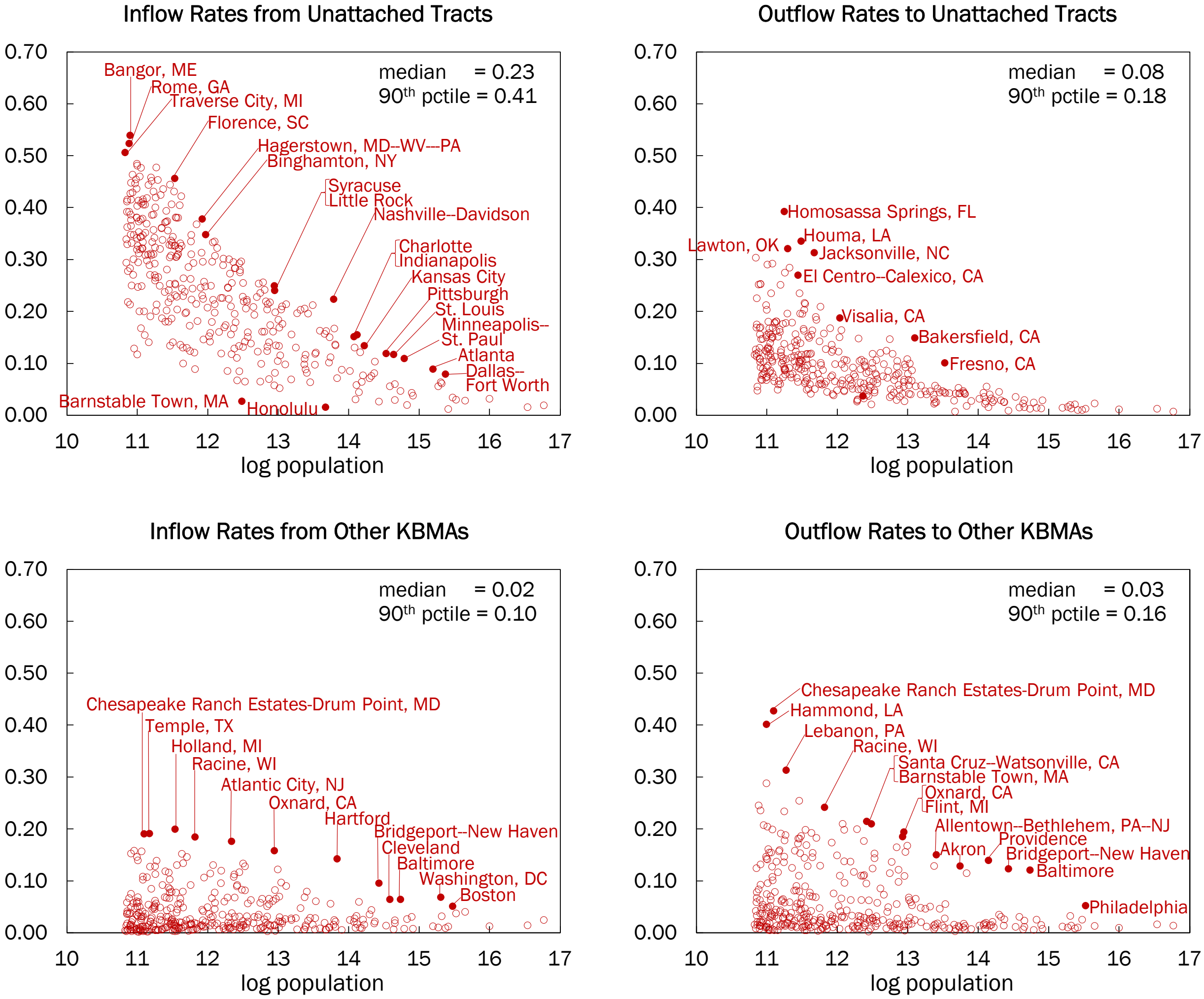
Note: Standard errors for the slope coefficient are reported in parentheses.

Figure A.7: Commuting Flows for UAs and Commuting Zones



Notes: Commuting inflow rates are measured as a share of employment. Commuting outflow rates are measured as a share of residents who are employed. Delineations and commuting flows are based on the 2000 decennial census. The Hightstown, NJ UA, has the highest commuting inflow and outflow rates, 0.74 and 0.81, respectively (above the displayed range). One additional UA has an inflow rate above the displayed range and nine additional ones have outflow rates above the displayed range. We use tract-based approximations of UAs to access commuting data, constructed with more relaxed threshold allocation factors than those for the kernel-based delineations. Specifically, we assign a tract to the tract-based approximation of a UA if at least 50 percent of its population is located in the UA, regardless of how much of its land area is in the UA. Doing so effects UA approximations that have population closer to the actual UA value. The associated inclusion of more unsettled land does not directly affect the measured commuting rates. The Commuting Zone delineations are disseminated by the Economic Research Service (Economic Research Service, 2012).

Figure A.8: Decomposition of KBMA Commuting Flows



Notes: The top panels show inflow rates to KBMAs and outflow rates from them with respective origins and destinations in tracts that are not part of a KBMA. The bottom panels show inflow rates to KBMAs and outflow rates from them with respective origins and destinations in another KBMA. Commuting inflows overwhelmingly originate from unattached tracts. Commuting outflows split about equally to unattached tracts and to other KBMAs.

Table B.1: Metropolitan CBSAs Compared to KBMAs

metropolitan CBSA	KBMA cores in CBSA	KBMAs with a core in CBSA	KBMAs with a tract in CBSA	comparison KBMA	size relative to comparison KBMA			share of CBSA in comparison KBMA		
					pop	emp	land	pop	emp	land
Kansas City, MO-KS	3	1	1	Kansas City, MO--KS	1.23	1.11	7.33	0.82	0.90	0.14
Los Angeles-Long Beach-Santa Ana, CA	6	1	1	Los Angeles--Long Beach--Santa Ana, CA	0.81	0.85	1.21	0.99	0.99	0.46
Riverside-San Bernardino-Ontario, CA	9	2	2	Los Angeles--Long Beach--Santa Ana, CA	0.21	0.17	6.79	0.80	0.79	0.06
Oxnard-Thousand Oaks-Ventura, CA	5	2	2	Oxnard, CA	1.80	1.73	8.20	0.55	0.58	0.12
Boston-Cambridge-Quincy, MA-NH	7	4	4	Boston, MA--NH--CT	0.84	0.85	0.88	0.93	0.95	0.70
Worcester, MA	5	2	2	Boston, MA--NH--CT	0.14	0.12	0.38	0.89	0.94	0.59
Manchester-Nashua, NH	3	1	1	Boston, MA--NH--CT	0.07	0.07	0.22	0.87	0.91	0.40
Providence-New Bedford-Fall River, RI-MA	4	3	3	Providence, RI--MA	1.14	1.11	1.66	0.86	0.89	0.57
Raleigh-Cary, NC	2	1	1	Raleigh--Durham, NC	0.81	0.70	1.70	0.84	0.91	0.44
Durham, NC	1	1	2	Raleigh--Durham, NC	0.44	0.42	1.42	0.69	0.87	0.15
minimum	0	0	0		0.02	0.02	0.04	0.26	0.19	0.00
10th percentile	1	1	1		0.90	0.88	1.84	0.50	0.69	0.03
median	1	1	1		1.34	1.14	7.09	0.72	0.86	0.12
90th percentile	4	2	2		1.92	1.44	31.6	0.91	0.95	0.39
maximum	9	5	5		2.75	2.07	276	1.00	1.00	1.00
aggregate (all metropolitan CBSAs relative to all KBMAs)					1.14	1.07	6.57	0.87	0.93	0.15

Notes: This table complements Table 4 in the main text. Each metropolitan CBSA is compared to the KBMA that has the most populous core in it. The summary statistics on relative size and overlap exclude 8 metropolitan CBSAs that have no comparison KBMA, reflecting that the UA/UC cores located in them belong to built-out kernels with population below 50,000. The aggregate statistics compare all metropolitan CBSAs with all KBMAs.

Table B.2: Monte Carlo Tests for Pareto Population Distributions

	Monte Carlo Simulations							
	F(t) = 0.050	F(t) = 0.025	F(t) = 0.005	KBMAs	kernel-based metro regions	kernel-based urban areas	Metro CBSAs	UAs
25 obs	-7.32	-8.77	-11.84	-3.27	-5.96	-3.16	-2.63	-3.42
50 obs	-9.90	-11.86	-15.90	-8.82	-10.02*	-8.94	-9.23	-10.34*
75 obs	-11.78	-14.00	-19.21	-17.09**	-17.27**	-16.31**	-16.83**	-16.90**

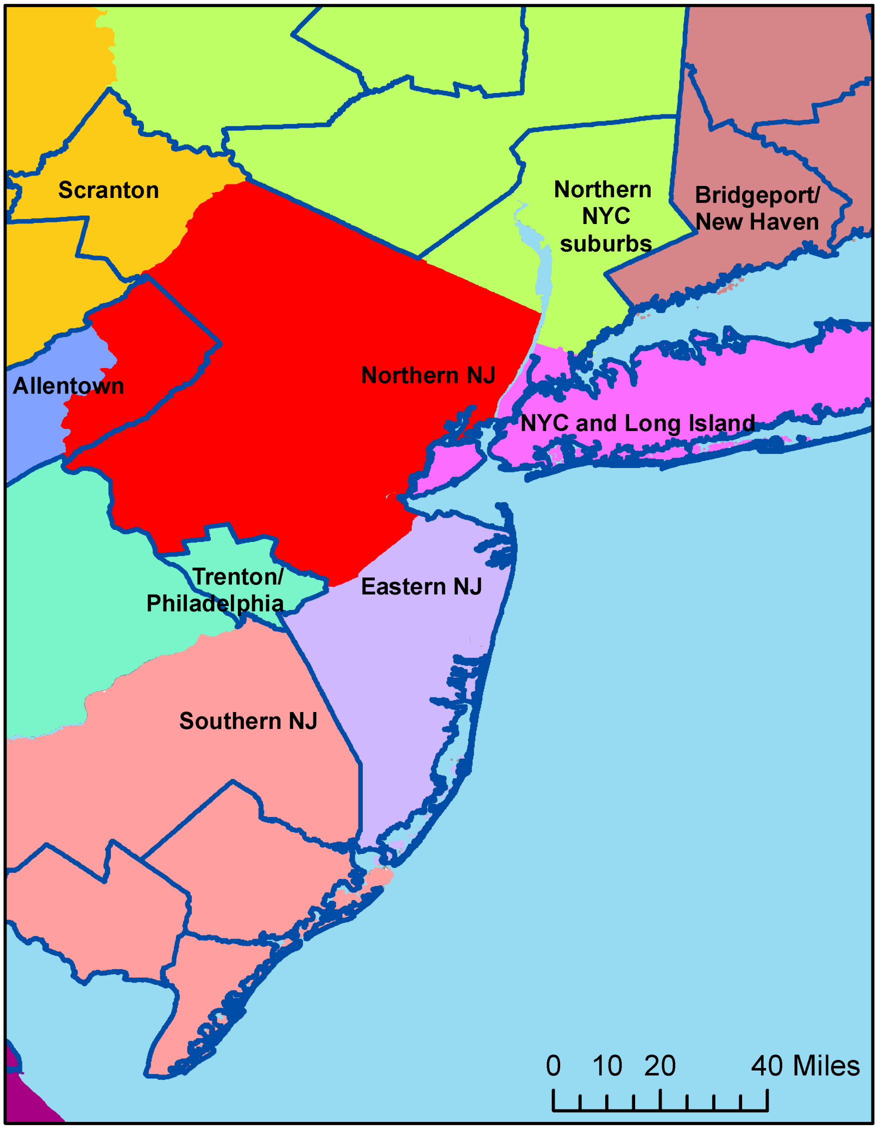
Notes: Table reports results from regressing log(rank - 0.5) on linear and quadratic log population using the largest 25, 50, and 75 observations. Following the methodology described in Gabaix and Ioannides (2004), the left three columns report the value of the t-statistic on the quadratic term at which the cumulative distribution from Monte Carlo simulations equals the specified value. The remaining columns report the t-statistic on the quadratic coefficient for each of the delineations. The Monte Carlo simulations draw the specified number of observations from a Pareto distribution 100,000 times; the t-statistic is independent of the shape parameter, reflecting that the shape parameter linearly affects both the coefficient and standard error. * denotes that we can reject that the observations are drawn from a Pareto distribution at the 0.10 level based on a two-tail test; ** that we can reject at the 0.05 level.

Table B.3: Metropolitan CBSA Pairs with Strong Cross Commuting

Rank	Sum of Flow Rates	Smaller Metropolitan CBSA				Larger Metropolitan CBSA			
		Title	Population	Outflow Rate	Inflow Rate	Title	Population	Outflow Rate	Inflow Rate
1	0.705	Oshkosh-Neenah, WI	157,000	0.183	0.211	Appleton, WI	202,000	0.175	0.136
2	0.507	Bristol, VA	68,000	0.160	0.207	Kingsport-Bristol, TN-VA	230,000	0.084	0.056
3	0.496	Madera, CA	123,000	0.240	0.197	Fresno, CA	799,000	0.026	0.033
4	0.489	Durham, NC	426,000	0.106	0.207	Raleigh-Cary, NC	797,000	0.122	0.053
5	0.479	Johnson City, TN	182,000	0.135	0.127	Kingsport-Bristol, TN-VA	230,000	0.101	0.117
6	0.475	Holland-Grand Haven, MI	238,000	0.249	0.111	Grand Rapids-Wyoming, MI	740,000	0.036	0.079
7	0.458	Hinesville-Fort Stewart, GA	72,000	0.388	0.040	Savannah, GA	293,000	0.002	0.028
8	0.454	Pascagoula, MS	151,000	0.222	0.075	Gulfport-Biloxi, MS	246,000	0.039	0.117
9	0.443	Anderson, SC	166,000	0.242	0.110	Greenville, SC	560,000	0.027	0.064
10	0.442	Ann Arbor, MI	323,000	0.192	0.213	Detroit-Warren-Livonia, MI	4,453,000	0.022	0.016
11	0.439	Trenton-Ewing, NJ	351,000	0.231	0.198	New York-Newark-Edison, NY-NJ-PA	18,323,000	0.005	0.005
12	0.428	Bay City, MI	110,000	0.204	0.083	Saginaw-Saginaw Township North, MI	210,000	0.040	0.101
13	0.427	Warner Robins, GA	111,000	0.166	0.100	Macon, GA	222,000	0.062	0.099
14	0.417	Kingston, NY	178,000	0.238	0.081	Poughkeepsie-Newburgh-Middletown, NY	622,000	0.019	0.078
15	0.400	Vallejo-Fairfield, CA	395,000	0.295	0.076	San Francisco-Oakland-Fremont, CA	4,124,000	0.005	0.025
16	0.387	Glens Falls, NY	124,000	0.183	0.157	Albany-Schenectady-Troy, NY	826,000	0.022	0.025
17	0.384	Manchester-Nashua, NH	381,000	0.209	0.144	Boston-Cambridge-Quincy, MA-NH	4,391,000	0.012	0.018
18	0.384	Midland, TX	116,000	0.088	0.099	Odessa, TX	121,000	0.106	0.090
19	0.378	San Jose-Sunnyvale-Santa Clara, CA	1,736,000	0.104	0.157	San Francisco-Oakland-Fremont, CA	4,124,000	0.075	0.042
20	0.374	Elkhart-Goshen, IN	183,000	0.056	0.156	South Bend-Mishawaka, IN-MI	317,000	0.125	0.037
21	0.370	Muskegon-Norton Shores, MI	170,000	0.158	0.069	Holland-Grand Haven, MI	238,000	0.039	0.104
22	0.369	Springfield, OH	145,000	0.219	0.101	Dayton, OH	848,000	0.015	0.034
23	0.366	Lebanon, PA	120,000	0.241	0.064	Harrisburg-Carlisle, PA	509,000	0.012	0.049
24	0.362	Ogden-Clearfield, UT	443,000	0.197	0.061	Salt Lake City, UT	969,000	0.024	0.080
25	0.359	Worcester, MA	751,000	0.216	0.095	Boston-Cambridge-Quincy, MA-NH	4,391,000	0.014	0.033
26	0.356	Monroe, MI	146,000	0.198	0.101	Toledo, OH	659,000	0.016	0.042
27	0.352	New Haven-Milford, CT	824,000	0.130	0.058	Bridgeport-Stamford-Norwalk, CT	883,000	0.051	0.113
28	0.345	Decatur, AL	146,000	0.172	0.082	Huntsville, AL	342,000	0.031	0.060
29	0.343	Spartanburg, SC	254,000	0.132	0.108	Greenville, SC	560,000	0.050	0.054
30	0.343	Auburn-Opelika, AL	115,000	0.205	0.037	Columbus, GA-AL	282,000	0.014	0.086
31	0.340	Anderson, IN	133,000	0.240	0.076	Indianapolis, IN	1,525,000	0.005	0.018
32	0.339	Akron, OH	695,000	0.166	0.089	Cleveland-Elyria-Mentor, OH	2,148,000	0.030	0.054
33	0.332	Ocean City, NJ	102,000	0.169	0.078	Atlantic City, NJ	253,000	0.026	0.059
34	0.328	Gainesville, GA	139,000	0.214	0.104	Atlanta-Sandy Springs-Marietta, GA	4,248,000	0.003	0.006
35	0.326	Greeley, CO	181,000	0.221	0.085	Denver-Aurora-Broomfield, CO & Boulder, CO	2,449,000	0.005	0.015
36	0.321	Oxnard-Thousand Oaks-Ventura, CA	753,000	0.197	0.105	Los Angeles-Long Beach-Santa Ana, CA	12,366,000	0.006	0.013
37	0.311	Lewiston-Auburn, ME	104,000	0.175	0.084	Portland-South Portland, ME	488,000	0.017	0.035
38	0.309	Trenton-Ewing, NJ	351,000	0.066	0.222	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	5,687,000	0.017	0.004
39	0.307	Monroe, MI	146,000	0.200	0.098	Detroit-Warren-Livonia, MI	4,453,000	0.002	0.007
40	0.306	Racine, WI	189,000	0.198	0.077	Milwaukee-Waukesha-West Allis, WI	1,501,000	0.008	0.024
41	0.301	Riverside-San Bernardino-Ontario, CA	3,255,000	0.181	0.064	Los Angeles-Long Beach-Santa Ana, CA	12,366,000	0.014	0.043
42	0.298	Poughkeepsie-Newburgh-Middletown, NY	622,000	0.239	0.049	New York-Newark-Edison, NY-NJ-PA	18,323,000	0.002	0.008
43	0.298	Greeley, CO	181,000	0.098	0.086	Fort Collins-Loveland, CO	251,000	0.046	0.067
44	0.295	Michigan City-La Porte, IN	110,000	0.151	0.140	Chicago-Naperville-Joliet, IL-IN-WI	9,098,000	0.002	0.002
45	0.295	Olympia, WA	207,000	0.200	0.077	Seattle-Tacoma-Bellevue, WA	3,044,000	0.005	0.013
46	0.294	Las Cruces, NM	175,000	0.158	0.075	El Paso, TX	680,000	0.019	0.042
47	0.285	Merced, CA	211,000	0.121	0.078	Modesto, CA	447,000	0.029	0.057
48	0.284	Carson City, NV	52,000	0.143	0.105	Reno-Sparks, NV	343,000	0.018	0.018
49	0.283	Naples-Marco Island, FL	251,000	0.049	0.126	Cape Coral-Fort Myers, FL	441,000	0.079	0.029
50	0.278	Napa, CA	124,000	0.064	0.136	Vallejo-Fairfield, CA	395,000	0.047	0.030
51	0.275	Flint, MI	436,000	0.172	0.080	Detroit-Warren-Livonia, MI	4,453,000	0.007	0.016
52	0.275	Punta Gorda, FL	142,000	0.139	0.092	Sarasota-Bradenton-Venice, FL	590,000	0.017	0.028
53	0.274	Elizabethtown, KY	108,000	0.148	0.103	Louisville, KY-IN	1,162,000	0.009	0.013
54	0.273	Bremerton-Silverdale, WA	232,000	0.197	0.059	Seattle-Tacoma-Bellevue, WA	3,044,000	0.004	0.013
55	0.270	Canton-Massillon, OH	407,000	0.123	0.050	Akron, OH	695,000	0.027	0.071
56	0.269	Coeur d'Alene, ID	109,000	0.169	0.049	Spokane, WA	418,000	0.011	0.041
57	0.269	Vineland-Millville-Bridgeton, NJ	146,000	0.115	0.147	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	5,687,000	0.003	0.003
58	0.265	Niles-Benton Harbor, MI	162,000	0.100	0.074	South Bend-Mishawaka, IN-MI	317,000	0.037	0.053
59	0.264	Kankakee-Bradley, IL	104,000	0.196	0.065	Chicago-Naperville-Joliet, IL-IN-WI	9,098,000	0.001	0.002
60	0.264	New Haven-Milford, CT	824,000	0.076	0.083	Hartford-West Hartford-East Hartford, CT	1,149,000	0.054	0.050
61	0.261	Winchester, VA-WV	103,000	0.169	0.087	Washington-Arlington-Alexandria, DC-VA-MD-WV	4,796,000	0.002	0.003
62	0.261	Longview-Kelso, WA	93,000	0.157	0.097	Portland-Vancouver-Beaverton, OR-WA	1,928,000	0.003	0.004
63	0.257	Lawrence, KS	100,000	0.164	0.079	Kansas City, MO-KS	1,836,000	0.004	0.010
64	0.250	York-Hanover, PA	382,000	0.111	0.038	Harrisburg-Carlisle, PA	509,000	0.027	0.074

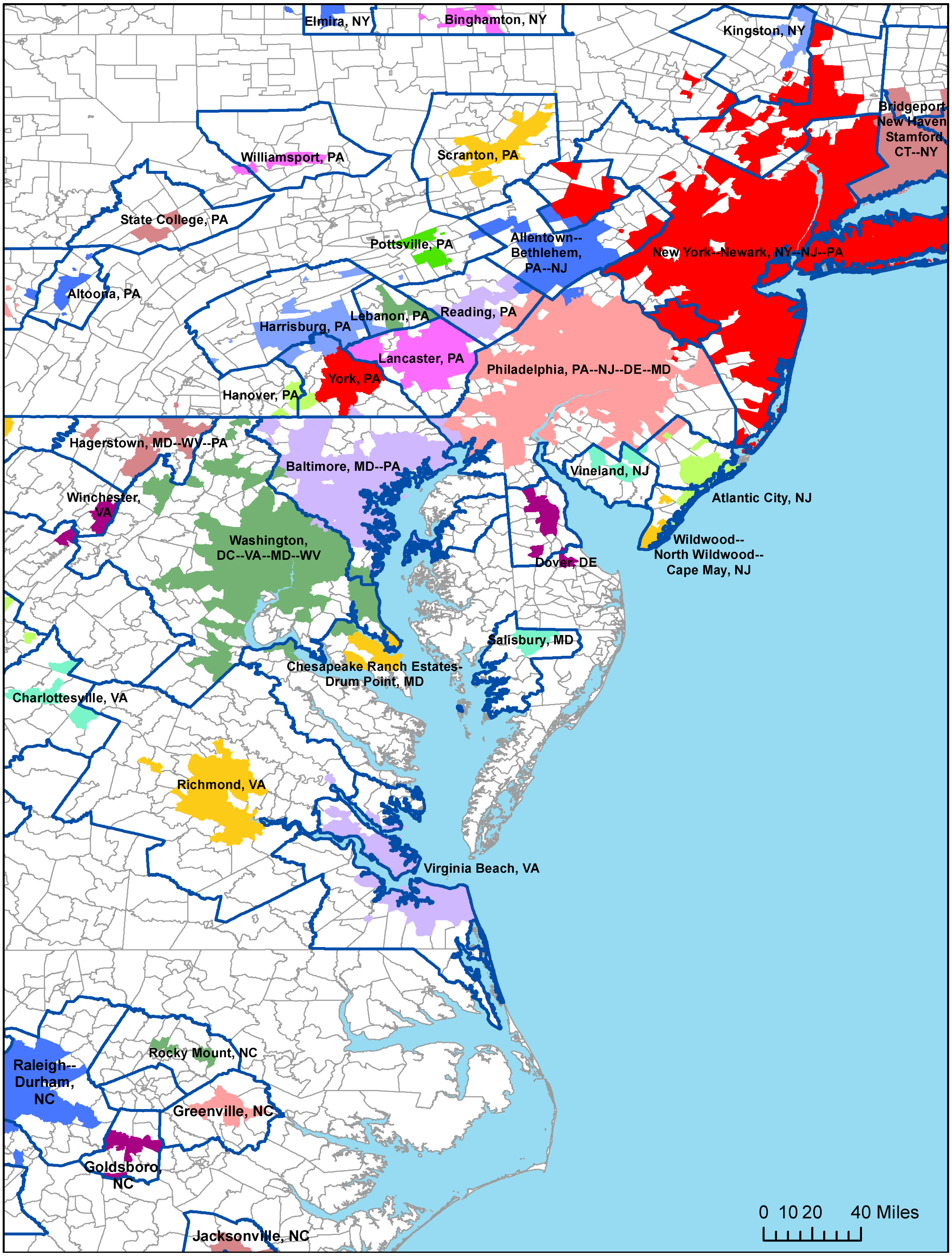
Notes: Outflow rates are measured by workers commuting to the paired CBSA as a share of employed residents. Inflow rates are measured by workers commuting from the paired CBSA as a share of employment. The 64 pairs are all those for which the sum of the four inflow and outflow rates exceeds 0.25, the threshold value used for the KBMA parameterization. Delineations and flows are based on the 2000 decennial census.

Figure C.1: Commuting Zones in the Vicinity of New York City



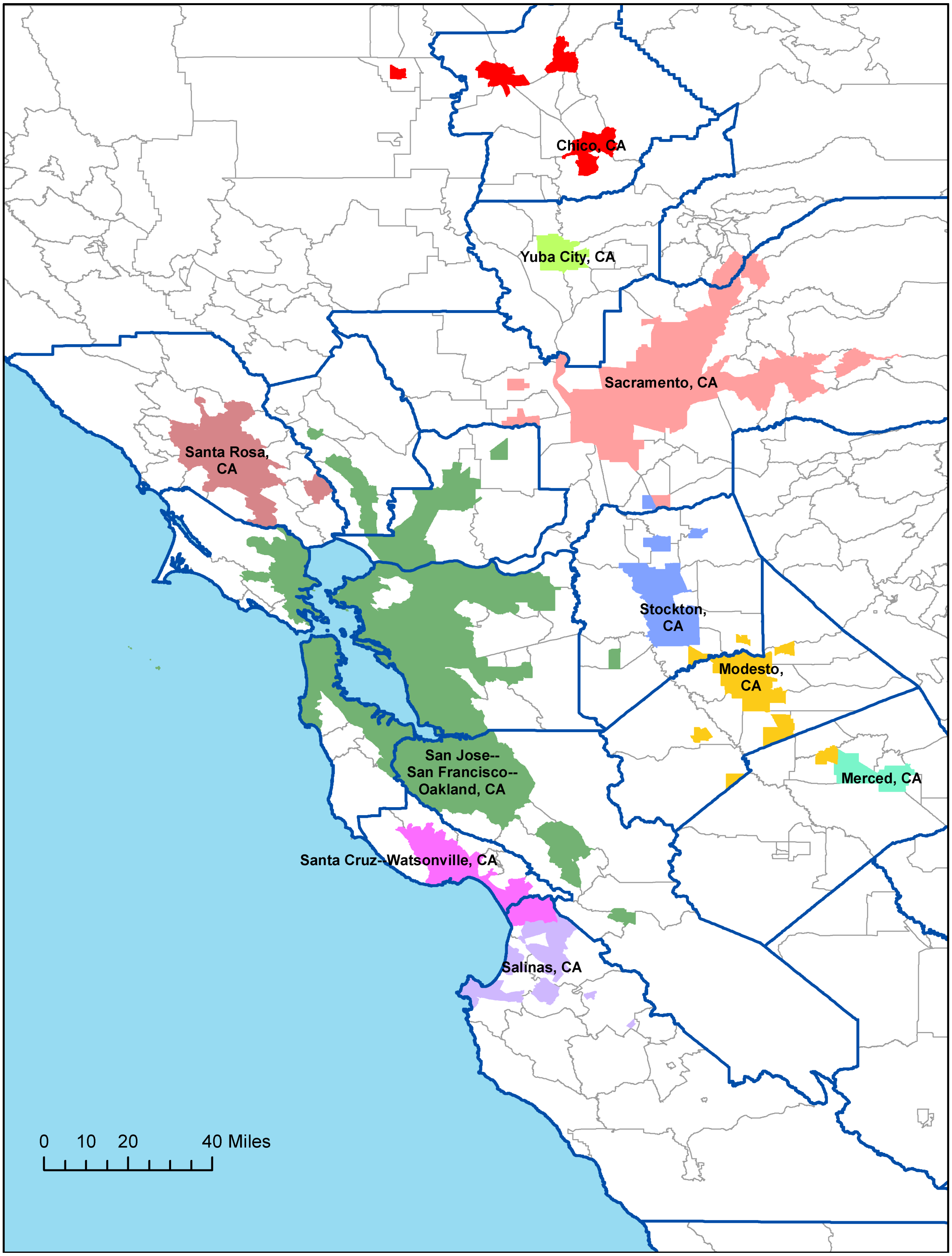
Blue lines demarcate the borders of metropolitan CBSAs. The commuting zone delineations are based on commuting patterns in 2000 (Economic Research Service, 2012).

Figure C.2: KBMAs from NYC to Virginia Beach



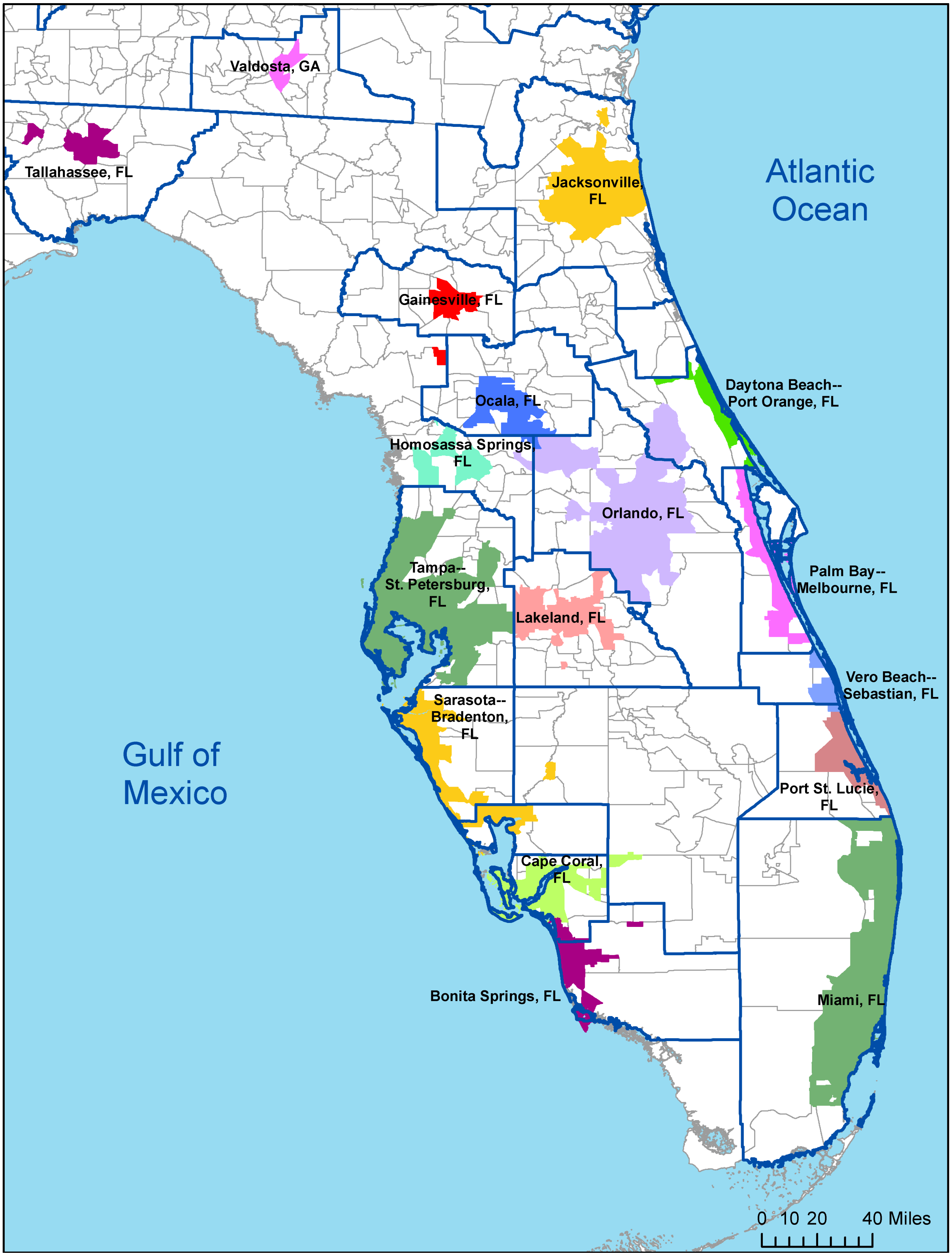
Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders. Additional maps of KBMAs, kernel-based urban areas, and kernel-based metropolitan regions are available from the paper's webpage.

Figure C.3: KBMAs in Northern California



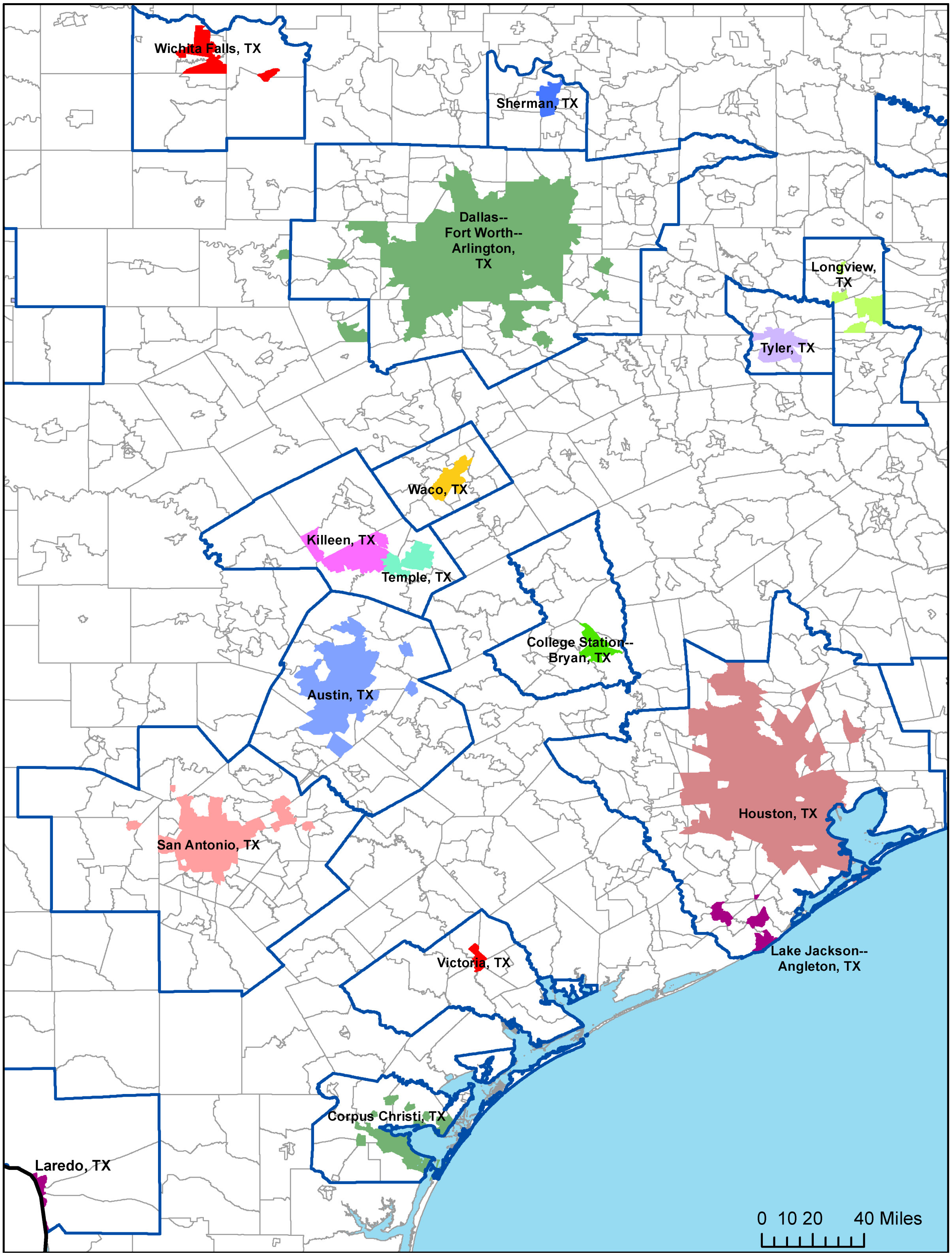
Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders. Additional maps of KBMAs, kernel-based urban areas, and kernel-based metropolitan regions are available from the paper's webpage.

Figure C.4: KBMAs in Florida



Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders. Additional maps of KBMAs, kernel-based urban areas, and kernel-based metropolitan regions are available from the paper's webpage.

Figure C.5: KBMAs in Texas



Blue lines demarcate the borders of metropolitan CBSAs. Gray lines demarcate tract borders. Additional maps of KBMAs, kernel-based urban areas, and kernel-based metropolitan regions are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 1 to 50 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
1	New York--Newark, NY--NJ--PA	14	19,138,832	8,293,650	6,391	4,635	229	0.045	0.022
2	Los Angeles--Long Beach--Santa Ana, CA	14	15,190,744	6,140,949	4,014	3,099	145	0.030	0.029
3	Chicago, IL--IN--WI	7	8,830,933	4,075,080	3,381	1,994	120	0.045	0.021
4	San Jose--San Francisco--Oakland, CA	15	6,257,164	3,092,204	1,940	1,292	99	0.069	0.036
5	Philadelphia, PA--NJ--DE--MD	4	5,519,614	2,464,841	3,106	1,401	106	0.068	0.066
6	Boston, MA--NH--CT	6	5,250,008	2,710,238	3,987	1,095	118	0.086	0.036
7	Miami, FL	2	4,932,226	2,076,669	1,484	871	36	0.025	0.021
8	Dallas--Fort Worth--Arlington, TX	8	4,736,484	2,409,101	2,634	967	127	0.089	0.030
9	Detroit, MI	5	4,501,066	2,117,786	2,261	1,299	108	0.093	0.027
10	Washington, DC--VA--MD--WV	6	4,428,188	2,464,935	2,411	931	95	0.133	0.052
11	Houston, TX	7	4,277,957	1,960,051	2,588	796	73	0.066	0.027
12	Atlanta, GA	7	3,962,469	2,101,116	3,740	631	107	0.102	0.027
13	Phoenix--Mesa, AZ	3	2,990,732	1,364,523	1,141	637	76	0.046	0.037
14	Seattle, WA	3	2,942,476	1,540,236	1,666	640	68	0.074	0.029
15	San Diego, CA	2	2,742,261	1,257,645	1,254	585	45	0.047	0.041
16	Minneapolis--St. Paul, MN--WI	7	2,639,647	1,525,513	1,694	671	92	0.117	0.031
17	Baltimore, MD--PA	6	2,504,099	1,146,193	1,652	613	77	0.111	0.134
18	Denver--Aurora, CO	7	2,314,392	1,235,379	1,153	552	57	0.079	0.041
19	Tampa--St. Petersburg, FL	4	2,313,087	1,034,051	1,418	521	53	0.054	0.040
20	St. Louis, MO--IL	4	2,275,085	1,174,114	1,584	467	101	0.124	0.030
21	Cleveland, OH	3	2,156,397	1,055,456	1,443	682	96	0.121	0.046
22	Pittsburgh, PA	5	2,044,606	981,499	1,830	624	80	0.133	0.047
23	Bridgeport--New Haven--Stamford, CT--NY	4	1,843,931	848,378	1,414	422	91	0.122	0.133
24	Cincinnati, OH--KY--IN	5	1,743,906	893,751	1,364	426	64	0.125	0.051
25	Portland, OR--WA	4	1,728,986	905,121	960	378	61	0.099	0.040
26	Orlando, FL	5	1,657,169	836,996	1,318	327	59	0.120	0.060
27	Sacramento, CA	6	1,613,617	725,772	853	348	83	0.107	0.080
28	Kansas City, MO--KS	3	1,498,432	822,433	1,072	432	65	0.150	0.040
29	Milwaukee, WI	5	1,471,013	738,826	1,124	411	46	0.102	0.046
30	Virginia Beach, VA	2	1,442,566	704,310	759	338	61	0.074	0.048
31	San Antonio, TX	3	1,413,845	662,617	654	278	69	0.131	0.055
32	Providence, RI--MA	2	1,390,708	598,270	963	313	62	0.105	0.156
33	Charlotte, NC--SC	8	1,353,014	751,632	1,717	272	71	0.173	0.071
34	Salt Lake City, UT	4	1,339,304	653,894	734	285	67	0.091	0.046
35	Las Vegas, NV	2	1,321,981	608,039	458	319	36	0.041	0.031
36	Indianapolis, IN	5	1,304,430	724,869	1,000	273	74	0.159	0.053
37	Columbus, OH	6	1,294,551	736,730	825	319	88	0.173	0.059
38	New Orleans, LA	5	1,219,997	545,471	610	361	67	0.110	0.047
39	Buffalo, NY	2	1,090,429	506,077	760	279	47	0.094	0.035
40	Hartford, CT	3	1,017,991	528,261	1,048	252	64	0.214	0.134
41	Austin, TX	4	995,870	584,588	738	215	42	0.157	0.054
42	Memphis, TN--MS--AR	1	989,579	489,539	638	235	39	0.148	0.049
43	Raleigh--Durham, NC	3	978,319	588,087	1,243	175	61	0.192	0.062
44	Nashville-Davidson, TN	3	970,068	599,566	1,005	206	77	0.240	0.058
45	Akron, OH	3	928,201	431,173	770	214	53	0.168	0.172
46	Louisville, KY--IN	2	917,392	489,877	677	217	51	0.160	0.053
47	Jacksonville, FL	1	906,595	452,602	680	162	32	0.113	0.067
48	Oklahoma City, OK	3	870,650	456,485	594	281	48	0.169	0.050
49	Honolulu, HI	3	863,696	406,551	444	210	39	0.022	0.011
50	Richmond, VA	1	853,514	472,891	665	219	43	0.177	0.044

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 51 to 100 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
51	Rochester, NY	3	796,657	424,170	678	199	77	0.172	0.051
52	Dayton, OH	1	752,763	385,183	614	185	36	0.182	0.098
53	Fresno, CA	5	746,386	278,217	414	144	67	0.147	0.129
54	Tucson, AZ	2	742,531	346,140	405	172	27	0.104	0.055
55	Birmingham, AL	2	720,609	387,230	797	160	63	0.228	0.061
56	Sarasota--Bradenton, FL	3	696,300	273,498	506	159	60	0.093	0.090
57	Albany, NY	3	673,220	371,875	544	173	59	0.219	0.070
58	Allentown--Bethlehem, PA--NJ	2	661,446	288,458	730	150	59	0.161	0.176
59	El Paso, TX--NM	1	652,316	238,512	325	120	33	0.100	0.074
60	Omaha, NE--IA	1	651,884	364,687	402	200	36	0.131	0.035
61	Tulsa, OK	3	642,057	366,505	588	203	57	0.212	0.030
62	Springfield, MA--CT	2	641,112	297,119	608	136	39	0.177	0.160
63	Albuquerque, NM	3	638,220	309,789	429	152	33	0.105	0.054
64	Grand Rapids, MI	1	605,501	357,768	535	127	51	0.251	0.097
65	Knoxville, TN	4	578,677	297,608	941	122	58	0.203	0.053
66	Toledo, OH--MI	2	567,085	279,105	417	148	42	0.159	0.093
67	Baton Rouge, LA	2	533,409	259,917	577	109	41	0.181	0.078
68	McAllen, TX	1	530,724	165,609	549	73	38	0.097	0.089
69	Greensboro, NC	4	493,887	306,826	598	113	32	0.304	0.122
70	Bakersfield, CA	6	487,473	159,284	229	100	42	0.125	0.174
71	Youngstown, OH--PA	2	469,306	210,122	482	135	37	0.198	0.139
72	Columbia, SC	1	466,756	265,617	588	109	53	0.228	0.085
73	Harrisburg, PA	3	460,730	286,055	543	102	79	0.301	0.100
74	Colorado Springs, CO	2	454,418	228,366	265	102	26	0.137	0.130
75	Palm Bay--Melbourne, FL	2	450,774	187,411	334	86	25	0.083	0.105
76	Greenville, SC	2	448,979	249,544	655	104	36	0.239	0.126
77	Charleston--North Charleston, SC	1	446,392	227,356	431	95	32	0.144	0.084
78	South Bend, IN--MI	2	441,799	236,334	411	100	39	0.207	0.103
79	Stockton, CA	3	439,187	149,285	195	99	19	0.228	0.276
80	Modesto, CA	5	426,915	149,714	200	83	32	0.201	0.248
81	Wichita, KS	2	423,884	224,801	249	114	50	0.203	0.103
82	Lancaster, PA	2	421,674	212,231	632	84	48	0.165	0.137
83	Little Rock, AR	1	419,556	243,623	522	94	56	0.256	0.085
84	Syracuse, NY	1	418,085	235,702	278	130	40	0.264	0.077
85	Oxnard, CA	3	417,622	171,988	225	82	25	0.216	0.250
86	Flint, MI	1	410,888	170,376	447	123	34	0.222	0.227
87	Des Moines, IA	2	393,771	242,217	362	84	25	0.187	0.047
88	Scranton, PA	1	392,935	198,619	320	121	48	0.233	0.068
89	Cape Coral, FL	2	387,276	149,939	409	93	47	0.125	0.157
90	Lakeland, FL	3	387,256	163,430	399	87	34	0.207	0.203
91	Mobile, AL	2	380,918	171,952	443	110	39	0.231	0.132
92	Boise City, ID	2	379,646	194,914	314	62	37	0.136	0.101
93	Santa Rosa, CA	3	377,751	170,118	317	70	32	0.175	0.244
94	Chattanooga, TN--GA	1	377,499	201,171	508	80	25	0.235	0.085
95	Madison, WI	2	370,462	249,151	305	79	49	0.240	0.094
96	Jackson, MS	3	363,880	198,350	441	89	26	0.249	0.077
97	Spokane, WA	2	363,152	189,018	267	92	42	0.187	0.059
98	Lexington-Fayette, KY	6	357,119	212,806	350	77	49	0.268	0.132
99	Winston-Salem, NC	1	352,235	177,455	558	80	27	0.227	0.176
100	Corpus Christi, TX	3	347,449	150,678	242	68	42	0.109	0.065

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 101 to 150 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
101	Provo--Orem, UT	2	342,579	145,715	218	76	20	0.123	0.150
102	Augusta-Richmond County, GA--SC	1	342,351	161,255	371	65	55	0.244	0.167
103	Lansing, MI	2	334,492	201,317	257	91	38	0.289	0.135
104	Pensacola, FL	1	331,060	151,274	302	63	27	0.153	0.112
105	Fort Wayne, IN	2	326,473	181,261	302	84	32	0.236	0.133
106	Rockford, IL--WI	2	315,357	155,625	228	85	30	0.225	0.173
107	Beaumont, TX	3	310,129	135,200	396	84	40	0.228	0.100
108	Salinas, CA	2	299,572	121,152	166	64	39	0.202	0.207
109	Reno, NV	1	287,283	158,204	184	54	15	0.182	0.080
110	Brownsville, TX	3	286,227	98,562	276	73	45	0.208	0.128
111	Gulfport--Biloxi, MS	2	283,228	150,183	333	65	62	0.251	0.110
112	Peoria, IL	2	282,484	142,913	295	71	51	0.240	0.129
113	Davenport, IA--IL	1	279,155	152,579	221	80	34	0.216	0.094
114	Reading, PA	1	275,515	137,717	304	66	28	0.270	0.197
115	Port St. Lucie, FL	1	275,248	103,174	275	52	20	0.147	0.192
116	Fayetteville, NC	1	270,980	144,961	259	46	27	0.234	0.089
117	Lafayette, LA	3	268,586	125,431	399	55	31	0.271	0.120
118	Indio--Palm Springs--Cathedral City, CA	2	267,989	103,579	182	62	28	0.266	0.243
119	Appleton--Oshkosh, WI	2	266,858	159,702	184	63	47	0.252	0.130
120	Barnstable Town, MA	1	263,125	101,240	437	56	55	0.111	0.231
121	Shreveport, LA	1	261,299	146,747	232	67	21	0.305	0.063
122	Asheville, NC	2	256,958	110,282	511	56	47	0.291	0.081
123	Hickory, NC	1	256,689	156,120	564	51	59	0.290	0.093
124	Portland, ME	2	254,975	167,390	382	59	49	0.325	0.124
125	Bonita Springs, FL	3	254,734	116,102	256	54	32	0.202	0.110
126	Anchorage, AK	3	253,897	133,795	163	52	65	0.143	0.040
127	Santa Cruz--Watsonville, CA	2	246,724	104,349	207	50	32	0.160	0.255
128	Daytona Beach--Port Orange, FL	1	245,344	105,951	187	50	20	0.201	0.177
129	Eugene, OR	1	241,053	126,448	168	56	30	0.192	0.113
130	Montgomery, AL	2	240,600	130,055	202	58	23	0.302	0.111
131	Huntsville, AL	1	238,407	150,505	373	62	25	0.315	0.091
132	Salem, OR	3	238,162	111,079	190	42	39	0.264	0.216
133	York, PA	1	237,138	121,574	324	55	23	0.241	0.210
134	Norwich--New London, CT--RI	2	235,237	132,043	339	56	38	0.276	0.163
135	Springfield, MO	1	233,800	154,184	231	60	20	0.302	0.049
136	Kalamazoo, MI	2	229,118	132,070	294	56	34	0.293	0.171
137	Atlantic City, NJ	1	228,081	119,972	223	58	30	0.294	0.161
138	Columbus, GA--AL	1	225,943	104,709	189	63	16	0.220	0.120
139	Lincoln, NE	1	225,386	135,841	91	53	12	0.174	0.088
140	Evansville, IN--KY	1	222,939	128,775	179	60	44	0.283	0.113
141	Saginaw, MI	2	220,350	112,665	226	57	22	0.331	0.180
142	Erie, PA	1	218,969	117,659	149	58	37	0.235	0.063
143	Roanoke, VA	1	217,319	128,312	249	45	21	0.265	0.051
144	Savannah, GA	1	216,182	122,482	275	63	27	0.260	0.043
145	Bremerton, WA	1	215,064	89,150	260	47	22	0.153	0.227
146	Lubbock, TX	1	209,552	106,482	106	55	25	0.181	0.083
147	Fort Collins, CO	1	206,170	113,703	128	48	16	0.230	0.197
148	Fayetteville--Springdale, AR	1	202,719	122,977	308	38	15	0.311	0.103
149	Tallahassee, FL	1	202,312	129,147	158	43	28	0.303	0.065
150	Green Bay, WI	1	200,209	127,086	201	44	15	0.261	0.094

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: **Enumeration of KBMAs** (ranks 151 to 200 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
151	Santa Barbara, CA	1	196,400	110,976	99	40	45	0.239	0.120
152	Spartanburg, SC	1	195,653	104,350	384	41	26	0.322	0.206
153	Amarillo, TX	2	190,748	92,560	134	53	18	0.167	0.079
154	Huntington, WV--KY--OH	1	186,027	92,237	255	53	43	0.341	0.150
155	Ocala, FL	2	180,096	79,850	314	30	20	0.365	0.188
156	Muskegon, MI	1	179,598	86,452	231	43	19	0.244	0.185
157	Utica, NY	3	179,522	92,305	178	61	43	0.350	0.137
158	Olympia--Lacey, WA	2	179,190	87,931	215	32	38	0.292	0.247
159	Rochester--Dover--Portsmouth, NH--ME	2	175,589	102,515	287	37	27	0.353	0.271
160	Charleston, WV	1	174,280	101,607	243	45	29	0.365	0.100
161	Gainesville, FL	1	172,789	99,386	147	33	14	0.316	0.129
162	San Luis Obispo, CA	4	172,561	79,552	131	32	38	0.247	0.231
163	Cedar Rapids, IA	1	169,855	109,305	142	37	51	0.279	0.104
164	Visalia, CA	2	168,499	60,961	107	36	20	0.274	0.298
165	Laredo, TX	1	168,032	51,985	75	29	4	0.156	0.158
166	Fort Walton Beach--Crestview, FL	2	165,360	79,787	145	31	34	0.188	0.174
167	Killeen, TX	1	161,727	85,412	246	27	23	0.257	0.198
168	Binghamton, NY	1	157,523	94,574	134	44	30	0.361	0.116
169	Champaign, IL	2	156,974	95,359	135	38	33	0.269	0.093
170	Macon, GA	1	154,157	83,327	243	41	18	0.364	0.115
171	Kennewick--Richland, WA	1	152,619	66,694	194	28	22	0.158	0.137
172	Topeka, KS	1	152,250	92,329	178	37	15	0.267	0.073
173	Waco, TX	1	151,254	81,673	103	35	11	0.321	0.105
174	Wilmington, NC	1	150,256	82,467	143	30	11	0.245	0.159
175	Hagerstown, MD--WV--PA	1	150,012	69,045	249	33	46	0.423	0.296
176	Myrtle Beach, SC--NC	1	148,692	91,122	271	33	39	0.329	0.096
177	Chico, CA	3	147,054	62,168	119	30	50	0.259	0.203
178	Fargo, ND--MN	1	140,463	90,288	76	30	15	0.189	0.053
179	Springfield, IL	1	138,792	102,407	101	41	14	0.386	0.080
180	Yakima, WA	2	138,277	60,111	163	21	26	0.224	0.140
181	Racine, WI	1	135,472	58,949	89	28	17	0.274	0.303
182	Yuma, AZ	1	134,671	41,906	144	25	26	0.141	0.174
183	Duluth, MN--WI	2	132,613	79,262	154	49	31	0.300	0.097
184	Medford, OR	1	131,096	70,750	131	24	18	0.250	0.069
185	College Station--Bryan, TX	1	129,426	66,338	80	26	8	0.240	0.122
186	Burlington, VT	1	128,608	97,541	190	27	18	0.353	0.067
187	Lafayette, IN	1	128,416	77,389	82	33	39	0.329	0.130
188	Lake Charles, LA	1	127,810	62,462	109	31	39	0.351	0.234
189	Johnson City, TN	1	127,399	62,307	214	29	21	0.355	0.198
190	Pueblo, CO	1	126,305	47,446	113	44	16	0.150	0.190
191	Panama City, FL	1	125,875	64,614	119	24	22	0.189	0.080
192	Sioux Falls, SD	1	123,538	87,412	61	25	15	0.290	0.084
193	Santa Maria, CA	1	120,985	41,614	54	25	8	0.276	0.367
194	Bellingham, WA	1	119,656	55,423	211	18	28	0.223	0.258
195	Wichita Falls, TX	2	118,074	63,216	142	31	33	0.211	0.064
196	Jacksonville, NC	1	117,095	53,407	160	22	32	0.255	0.339
197	Redding, CA	1	115,325	55,634	143	22	14	0.275	0.098
198	Corvallis--Albany, OR	3	114,723	53,318	115	25	32	0.279	0.171
199	Vero Beach--Sebastian, FL	1	114,468	43,439	118	24	13	0.234	0.243
200	Merced, CA	1	114,190	34,420	77	28	13	0.210	0.310

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 201 to 250 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
201	Vineland, NJ	2	112,858	47,394	125	25	21	0.396	0.348
202	Tuscaloosa, AL	1	112,243	62,077	107	31	9	0.347	0.162
203	Fort Smith, AR--OK	1	111,760	80,660	145	24	15	0.432	0.052
204	Athens-Clarke County, GA	1	110,047	69,539	135	30	14	0.389	0.138
205	Odessa, TX	1	109,987	44,056	95	26	13	0.202	0.189
206	Tyler, TX	1	109,153	67,260	143	26	11	0.426	0.180
207	Rochester, MN	1	108,417	77,469	100	30	40	0.331	0.066
208	Waterloo, IA	1	107,881	64,160	104	33	14	0.330	0.167
209	Jackson, MI	1	106,823	54,573	159	27	17	0.367	0.226
210	Lynchburg, VA	1	106,579	65,117	132	29	14	0.392	0.140
211	Bloomington--Normal, IL	1	105,888	76,907	46	28	7	0.354	0.105
212	Billings, MT	1	105,228	61,389	64	23	24	0.208	0.055
213	Kingsport, TN	1	104,993	46,560	202	24	21	0.360	0.242
214	Charlottesville, VA	1	104,560	68,703	175	25	27	0.420	0.149
215	Joplin, MO--KS	2	104,542	59,234	182	21	30	0.316	0.120
216	Burlington, NC	1	103,399	58,328	144	18	23	0.383	0.264
217	Clarksville, TN--KY	1	103,031	56,568	109	21	13	0.395	0.305
218	Abilene, TX	1	102,810	55,300	57	31	15	0.260	0.105
219	Columbia, MO	1	102,805	71,163	112	23	34	0.348	0.110
220	Holland, MI	1	102,228	69,097	129	15	15	0.431	0.229
221	Sioux City, IA--NE	1	101,948	51,202	83	25	10	0.257	0.222
222	Florence, SC	3	101,770	61,189	174	24	24	0.471	0.232
223	Mansfield, OH	2	101,340	53,544	155	27	22	0.344	0.199
224	Monroe, LA	1	100,888	57,475	101	30	32	0.440	0.179
225	Lima, OH	2	99,144	52,596	127	30	31	0.365	0.218
226	Grand Junction, CO	1	99,059	46,071	90	23	24	0.145	0.116
227	Wheeling, WV--OH	1	98,254	49,208	154	37	27	0.360	0.184
228	Houma, LA	1	97,460	39,803	115	22	21	0.432	0.395
229	Warner Robins, GA	1	97,411	45,213	124	16	8	0.265	0.275
230	Bloomington, IN	1	97,296	59,314	83	24	19	0.345	0.151
231	Johnstown, PA	1	96,386	45,831	139	35	19	0.331	0.162
232	Greenville, NC	1	96,302	57,348	152	14	20	0.354	0.167
233	Yuba City, CA	1	95,823	31,288	58	19	9	0.274	0.350
234	Hanford, CA	3	95,330	14,487	71	18	26	0.218	0.417
235	Longview, TX	2	95,027	58,148	113	21	17	0.461	0.144
236	Battle Creek, MI	1	94,645	54,556	139	27	22	0.383	0.182
237	Anderson, IN	1	94,601	40,831	111	25	12	0.357	0.358
238	Dover, DE	3	94,303	54,768	151	22	16	0.396	0.201
239	Springfield, OH	1	93,518	43,650	87	28	9	0.358	0.298
240	El Centro--Calexico, CA	3	93,250	18,210	37	17	14	0.211	0.299
241	Port Huron, MI	1	93,160	46,340	107	28	8	0.305	0.237
242	St. Cloud, MN	1	92,449	64,903	86	17	12	0.378	0.202
243	Albany, GA	1	91,606	51,627	110	26	13	0.386	0.082
244	Muncie, IN	1	90,874	49,160	98	25	9	0.324	0.185
245	Midland, TX	1	90,783	41,775	34	23	7	0.269	0.217
246	Altoona, PA	1	90,418	53,742	82	26	24	0.425	0.173
247	Eau Claire, WI	1	89,750	61,121	74	20	20	0.383	0.170
248	Greeley, CO	1	89,021	41,325	46	20	6	0.299	0.300
249	Terre Haute, IN	1	88,709	53,025	128	26	11	0.382	0.128
250	Benton Harbor--St. Joseph, MI	1	88,157	44,909	159	25	27	0.307	0.196

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 251 to 300 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
251	Parkersburg, WV--OH	1	86,149	53,077	88	29	19	0.444	0.096
252	Anderson, SC	1	85,385	46,385	172	19	20	0.431	0.304
253	San Angelo, TX	1	84,914	43,014	63	19	6	0.194	0.085
254	Decatur, IL	1	84,528	49,034	56	28	7	0.384	0.150
255	La Crosse, WI--MN	1	84,444	56,039	61	20	29	0.368	0.147
256	State College, PA	2	83,249	61,360	81	17	32	0.456	0.109
257	Las Cruces, NM	1	82,770	34,879	82	16	10	0.370	0.343
258	Sheboygan, WI	1	82,184	51,378	118	15	19	0.275	0.121
259	Lawton, OK	1	80,316	29,753	80	23	11	0.274	0.341
260	Jackson, TN	1	79,957	51,127	132	21	10	0.459	0.214
261	Lebanon, PA	1	78,587	33,179	124	19	19	0.353	0.422
262	Florence, AL	1	78,443	44,697	88	20	9	0.484	0.213
263	Lawrence, KS	1	77,974	41,515	38	16	8	0.258	0.276
264	Texarkana, TX--AR	1	77,560	46,288	100	18	11	0.425	0.092
265	Kahului--Kihei, HI	3	77,316	23,014	68	15	24	0.270	0.291
266	Dalton, GA	2	77,286	62,645	166	16	18	0.513	0.129
267	Michigan City, IN--MI	1	76,983	41,926	124	22	15	0.388	0.265
268	Iowa City, IA	1	76,965	60,807	48	17	7	0.385	0.148
269	Pottsville, PA	2	76,570	15,586	98	18	17	0.316	0.372
270	Homosassa Springs, FL	2	76,561	16,169	207	10	20	0.342	0.540
271	Weirton--Steubenville, WV--OH	1	76,143	38,194	103	24	12	0.395	0.230
272	Williamsport, PA	1	75,406	44,494	94	17	31	0.399	0.134
273	Santa Fe, NM	1	75,377	54,285	42	22	10	0.406	0.127
274	Elmira, NY	2	75,239	39,288	58	22	19	0.444	0.207
275	St. Joseph, MO	1	75,154	40,497	52	23	6	0.328	0.139
276	Auburn, AL	2	73,839	38,135	101	19	24	0.380	0.163
277	Anniston, AL	1	73,743	41,330	119	19	10	0.417	0.176
278	Alexandria, LA	1	73,142	43,239	79	22	9	0.459	0.142
279	Newark, OH	1	72,898	37,941	96	17	10	0.394	0.313
280	Rocky Mount, NC	2	72,649	31,885	92	16	24	0.486	0.295
281	Idaho Falls, ID	1	72,335	42,313	86	17	12	0.334	0.073
282	Logan, UT	1	72,054	34,050	72	15	6	0.213	0.201
283	Mount Vernon, WA	2	71,823	34,472	92	17	31	0.398	0.325
284	Temple, TX	1	71,055	44,670	120	19	10	0.481	0.224
285	Cheyenne, WY	1	70,356	35,678	51	16	7	0.156	0.087
286	Decatur, AL	2	70,222	45,350	107	17	10	0.498	0.269
287	Rapid City, SD	1	69,809	46,424	104	17	14	0.313	0.055
288	Eureka, CA	3	68,700	24,921	71	14	16	0.301	0.176
289	Bristol, TN--VA	2	68,590	41,663	108	15	23	0.526	0.319
290	Beckley, WV	2	68,192	25,200	151	15	25	0.380	0.226
291	Prescott, AZ	1	67,131	30,380	123	11	15	0.322	0.216
292	Gadsden, AL	1	66,979	34,188	132	20	12	0.389	0.192
293	Hattiesburg, MS	1	66,824	39,025	92	15	12	0.414	0.212
294	Great Falls, MT	1	66,363	33,825	78	19	12	0.159	0.071
295	Chesapeake Ranch Estates-Drum Point, MD	1	65,834	13,297	158	11	17	0.391	0.501
296	Lewiston, ME	1	64,998	39,935	94	18	15	0.444	0.252
297	Wausau, WI	1	64,961	45,492	95	15	10	0.409	0.176
298	Lake Jackson--Angleton, TX	1	64,770	28,849	96	14	20	0.456	0.400
299	Bowling Green, KY	1	64,724	48,322	99	14	8	0.429	0.116
300	Kingston, NY	1	64,641	35,704	107	20	12	0.436	0.303

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 301 to 350 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
301	Kankakee, IL	1	64,504	32,516	73	16	7	0.363	0.280
302	Owensboro, KY	1	64,337	38,504	66	17	9	0.391	0.138
303	Pittsfield, MA	1	63,747	37,180	110	15	10	0.353	0.185
304	Porterville, CA	2	63,506	21,719	27	13	10	0.406	0.390
305	Sumter, SC	1	63,220	38,753	102	13	11	0.437	0.175
306	Goldsboro, NC	1	63,025	40,561	126	14	14	0.501	0.208
307	Pocatello, ID	1	62,999	30,203	49	17	6	0.179	0.119
308	Blacksburg, VA	1	62,914	35,754	80	11	6	0.376	0.190
309	Salisbury, MD	1	62,878	24,181	98	13	7	0.383	0.169
310	Missoula, MT	1	62,485	42,744	75	12	6	0.347	0.133
311	Jonesboro, AR	1	62,289	39,315	161	9	9	0.325	0.102
312	Winchester, VA	1	62,004	39,276	101	10	17	0.510	0.322
313	Grand Forks, ND--MN	1	61,518	34,065	40	16	24	0.208	0.165
314	Dubuque, IA	1	61,132	40,704	32	17	6	0.362	0.139
315	Bismarck, ND	1	61,083	37,979	71	13	10	0.314	0.186
316	Coeur d'Alene, ID	1	60,994	33,588	48	12	12	0.413	0.277
317	Kokomo, IN	1	60,951	41,432	40	16	5	0.474	0.181
318	Longview, WA	1	60,600	20,064	45	16	8	0.326	0.192
319	Ashtabula, OH	2	60,417	25,389	110	15	28	0.316	0.332
320	Janesville, WI	1	60,298	31,816	86	14	6	0.466	0.432
321	Casper, WY	1	59,770	28,119	61	15	10	0.120	0.135
322	Dothan, AL	1	59,766	43,226	88	14	7	0.489	0.093
323	Findlay, OH	2	59,644	37,975	51	12	19	0.454	0.278
324	Staunton--Waynesboro, VA	2	59,598	23,863	107	13	15	0.415	0.211
325	Hammond, LA	1	59,521	9,609	152	10	5	0.399	0.553
326	Harrisonburg, VA	1	59,450	43,800	70	10	24	0.519	0.203
327	Hanover, PA	1	59,417	34,685	87	14	13	0.465	0.380
328	Valdosta, GA	1	59,297	34,600	75	17	6	0.443	0.225
329	Cleveland, TN	1	59,151	36,975	85	13	8	0.442	0.220
330	Paducah, KY--IL	1	58,711	37,480	81	15	15	0.461	0.135
331	Jefferson City, MO	1	57,837	48,414	78	11	30	0.495	0.117
332	Wildwood--North Wildwood--Cape May, NJ	1	57,213	21,973	67	13	11	0.287	0.311
333	Cumberland, MD--WV	2	56,868	28,816	108	18	13	0.358	0.129
334	Bend, OR	1	56,576	20,891	73	10	6	0.306	0.114
335	Sandusky, OH	1	56,529	30,678	90	13	16	0.328	0.201
336	Ithaca, NY	1	55,942	46,698	64	14	5	0.489	0.097
337	Morgantown, WV	1	55,807	40,100	62	15	8	0.455	0.106
338	Wenatchee, WA	1	55,229	17,581	47	10	14	0.253	0.168
339	Danville, IL	1	55,211	28,900	99	17	6	0.353	0.116
340	Victoria, TX	1	54,796	27,042	29	14	3	0.381	0.266
341	Glens Falls, NY	1	54,754	28,405	72	13	10	0.452	0.385
342	Concord, NH	1	54,722	42,092	115	11	10	0.569	0.329
343	Sherman, TX	1	54,594	34,218	67	14	6	0.498	0.263
344	Mankato, MN	2	54,421	27,396	34	13	5	0.354	0.140
345	Cape Girardeau, MO	1	54,414	36,784	104	13	11	0.398	0.122
346	Danville, VA	1	54,213	30,440	59	15	9	0.442	0.219
347	Bangor, ME	1	53,969	47,663	61	16	12	0.545	0.138
348	Beaufort, SC	1	53,884	29,354	125	9	11	0.253	0.189
349	Flagstaff, AZ	1	53,861	35,617	68	12	8	0.297	0.124
350	Adrian, MI	1	53,375	28,106	104	11	14	0.460	0.367

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.

Table D.1: Enumeration of KBMAs (ranks 351 to 361 of 361)

Rank	KBMA Title	Cores	Population	Employment	Land Area (sq.mi)	Tracts	Max Dist (mi)	Inflow Rate	Outflow Rate
351	Rome, GA	1	53,148	29,886	74	13	8	0.560	0.336
352	Hilo, HI	1	52,396	17,722	108	10	11	0.285	0.077
353	Greenville, MS	1	52,372	11,346	61	15	15	0.224	0.063
354	Ames, IA	1	51,944	33,734	40	14	16	0.368	0.253
355	Pine Bluff, AR	1	51,863	24,860	49	15	8	0.457	0.294
356	Richmond, IN	1	51,851	31,920	62	11	22	0.381	0.147
357	Carson City, NV	2	51,628	26,472	42	10	11	0.470	0.294
358	Hot Springs, AR	1	51,537	29,950	71	13	8	0.428	0.179
359	St. George, UT	1	51,194	12,551	79	11	8	0.308	0.346
360	Farmington, NM	1	51,115	18,870	95	11	13	0.342	0.136
361	Traverse City, MI	1	50,313	42,206	116	11	22	0.511	0.132

"Max Dist" is the distance between the tract centroids within a KBMA that are farthest from each other. Tables with more detailed variables for KBMAs and analogous tables for kernel-based metropolitan regions and kernel-based urban areas are available from the paper's webpage.