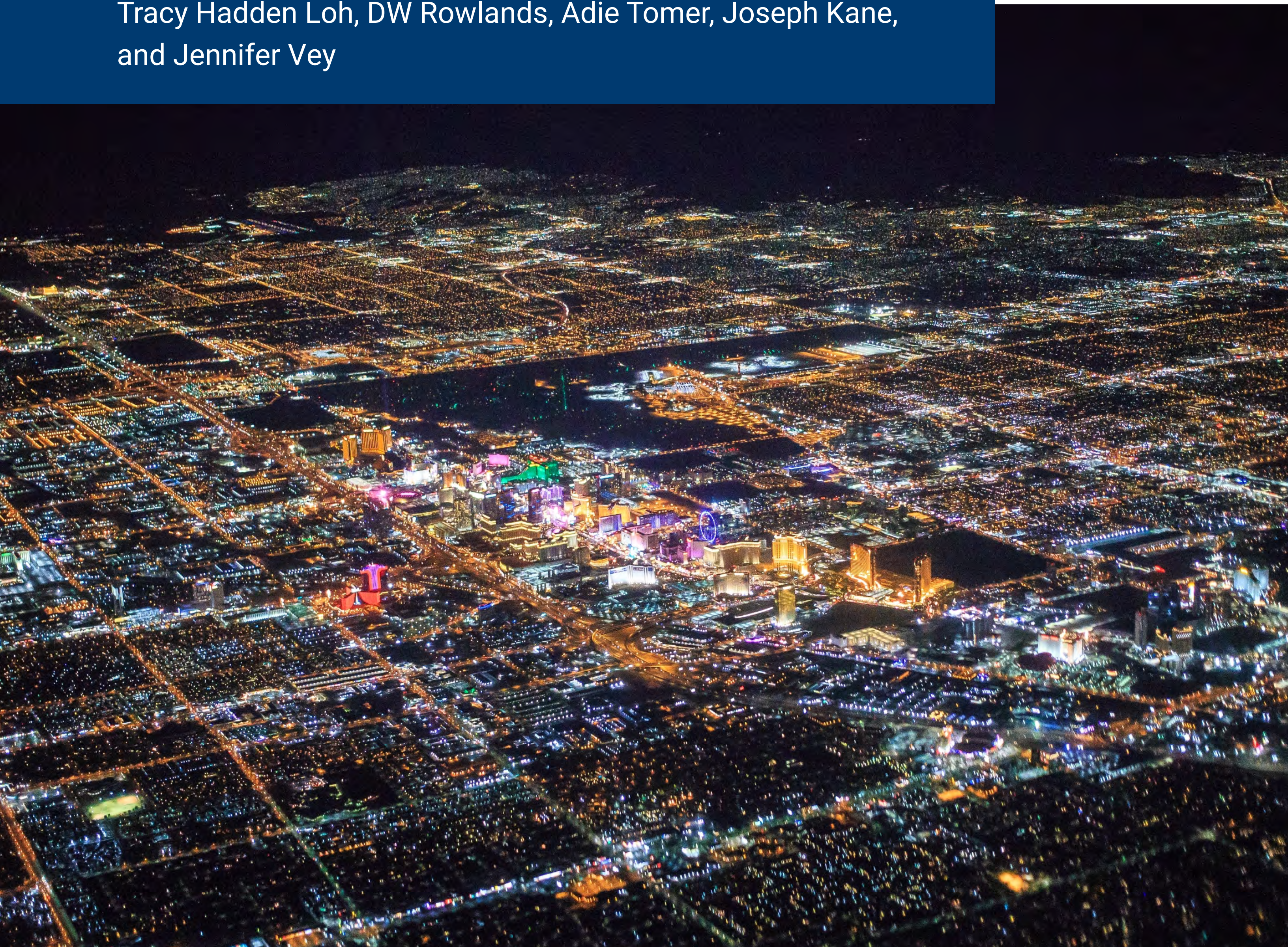
OCTOBER 2022

Brookings Metro

ANNE T. AND ROBERT M. BASS CENTER FOR TRANSFORMATIVE PLACEMAKING

MAPPING AMERICA'S ACTIVITY CENTERS: Methodology Appendix

Tracy Hadden Loh, DW Rowlands, Adie Tomer, Joseph Kane, and Jennifer Vey



CONTENT

COLLECTING DATA FOR ANALYSIS

One important factor in identifying activity centers (or spatial concentrations of any sort) is the scale of the geographic units used to define proximity. As noted by Robert Lang, even in an area with very deconcentrated activity, if you define a large enough boundary, you can enclose enough activity to have something, if not somewhere.[1] For transformative placemaking purposes, a pedestrian-based human scale—the area a person can traverse in a 10- to 30-minute walk, or roughly 120 to 1,200 acres—makes the most sense, since it establishes proximity at a scale where the barriers to travel are smallest.[2]

For an ivory-tower academic analysis, the ideal option would be to establish cells of a fixed size near this scale across each metro area, allowing for equivalent comparisons asset concentration at a human scale. However, such an analysis would likely be of limited practical use to many planners, who must work with Census Bureau geographies due to policy constraints or the need for other data that is only available at the census tract or census block group levels.

We chose to use 2020 block groups as our basic geography for identifying activity centers because they are the smallest Census Bureau geographies for which large amounts of data are available. However, because block groups are drawn to have roughly similar populations—usually between 600 and 3,000 residents—their size varies substantially depending on the built environment. In dense urban areas, they can be as small as a few city blocks, while in exurban and rural areas, they can occasionally be hundreds of square miles.

To account for the variation in block group sizes, we converted most of our variables to densities, using developed land area (identified from the 2019 National Land Cover Database) as the denominator. The only exceptions were assets—libraries, major sports stadiums, post offices, and institutional assets—that generally serve areas much larger than a block group, where a count or presence/absence marker were more appropriate. The use of developed land area, rather than all land area, in the denominator of density improves the accuracy of the local densities of the developed parts of suburban and exurban block groups that also contain large swaths of parkland or other undeveloped and uninhabited land.

Although 2020 block groups were our basic unit of analysis, virtually none of our data was native to this geography. A substantial share of our data came in the form of point locations for assets, and most of the rest was available as 2010 block groups or blocks.

SELECTING MAP PROJECTIONS FOR EACH METRO AREA

Transferring data between geographies generally requires converting all data to a well-selected projected coordinate system, to ensure that overlap areas between geographies are proportional to actual land area. A projected coordinate system is also essential for the distance-based analyses we performed on the clustering of activity centers and the demographics of buffers around them. A Universal Transverse Mercator (UTM) zone was identified for each metropolitan area based on the average longitude of the centroids of its 2020 block groups with water areas removed.

We chose this method—rather than the coordinates of the centroid of the metro area as a whole—to better account for the fact that the most developed area of some metro areas is quite far from the center of the metro area. Since block groups are smaller and more numerous in the most developed and populated parts of a metro area, this approach allowed us to select an appropriate UTM zone for the portion of the metro area with the most population. The EPSG codes for the UTM coordinate systems chosen for each metro area can be found in the appendix.

ASSIGNING POINT DATA TO 2020 BLOCK GROUPS

Most of the data variables that went into the activity centers calculations are in the form of point data: national datasets of points with metadata. This data was relatively easy to process, since it did not require converting between 2010 and 2020 census geographies; each feature was simply assigned to the 2020 block group it was located in in the local UTM coordinate system. For most of the datasets, a simple count of points was sufficient but for some, a variable from the metadata was preserved in the combined point dataset to allow a count of, for example, the number of stadium seats in a block group.

It is important to keep in mind that although the raw data is solely in the form of points—which means that each institution or site is assigned to a single block group—some of the assets described with point data, especially colleges and universities, can be quite large and spread out over a number of block groups. The number of parks in each block group was calculated similarly, but since park locations were available as polygons, each park was counted as present in every block group it overlapped with.

ASSIGNING DATA IN 2010 GEOGRAPHIES TO 2020 BLOCK GROUPS

Many of the datasets used to identify and characterize activity centers were aggregated by 2010 census geographies and had to be converted to 2020 block groups. Datasets available for 2010 blocks—Longitudinal Employer-Household Dynamics (LEHD) and 2010 decennial census data—were converted to 2020 blocks using crosswalks provided by the National Historic Geographic Information System (NHGIS), as discussed below. The 2021 EPA Smart Location Database (SLD) is only available at the 2010 block group level, and was converted to 2020 block groups using a crosswalk provided by the Census Bureau. Data from the 2019 American Community Survey (ACS) was used for a buffer analysis that did not require conversion to 2020 block groups; buffers of activity centers were defined as all 2010 block groups with centroids within a given distance of the centroids of the activity centers.

DATA SOURCES

DECENNIAL CENSUS REDISTRICTING POPULATION DATA

Total population, aggregated at the block group level from the 2020 decennial census redistricting population dataset, was used in the process of characterizing activity centers, and is reported in the POP_TOTAL variable. Racial breakdowns were also collected for potential use in characterizing activity center population: the POP_WHITE, POP_BLACK, and POP_ASIAN variables contain the count of people who listed their race as white, Black, or Asian American only and their ethnicity as not Latino or Hispanic. The POP_LATIN variable contains the count of people of any race who reported their ethnicity as Latino or Hispanic on the 2020 census.

The POP_2010 variable—collected to allow the determination of population change in activity centers—contains the total population values from 2010 decennial census redistricting data, tabulated at the census block level. This data is only available for 2010 geographies, so values for 2010 blocks were transferred to 2020 blocks using a 2010-to-2020 block crosswalk produced by the NHGIS and aggregated to 2020 block groups.[3] The Version 1 crosswalk used for this purpose is based on a target-density weighting process as well as area of overlap to determine the populations of 2020 blocks.

AMERICAN COMMUNITY SURVEY

Demographic data for the buffer analysis and for calculating the ratio of per capita income inside to outside of activity centers was drawn from the 2019 American Community Survey (ACS) 5-year estimates. Because this data is tabulated by 2010 block groups, the centroids of the 2010 block groups (rather than 2020 block groups) were used in measuring the distance to the centroids of activity centers for the buffer analysis. Per-capita income for 2020 block groups was calculated using Census Bureau 2010-to-2020 block group crosswalks and weighting the sum of total income in 2010 block groups by land area overlap with each 2020 block group.[4] ACS data Table 1 on the next page shows the ACS variables used.

TABLE 1

2019 ACS 5-year Estimates Data Used to Analyze Center Buffers

Measurement	Formula
Total Population	B03002_001
Non-Hispanic White-alone Population	B03002_003
Non-Hispanic Black-alone Population	B03002_004
Non-Hispanic Asian-alone Population	B03002_006
Latino or Hispanic of Any Race	B03002_012
Minority Population	B03002_001 – B03002_003
Total Commuters (workers except work-from-home)	B08301_001 – B08301_021
Number of Driving-Alone Commuters	B08301_003
Aggregate Income	B19313_001
Total Occupied Housing Units	B25003_001
Number of Owner-Occupied Housing Units	B25003_002
Number of Renter-Occupied Housing Units	B25003_001 – B25003_002
Total Housing Units	B25024_001
Number of Detached Single-Family Housing Units	B25024_002
Number of “Missing Middle” Housing Units	B25024_003 + B25024_004
(row houses and buildings of 2-to 9 units)	+ B25024_005 + B25024_006
Number of Housing Units in Buildings of 10+ Units	B25024_007 + B25024_008 + B25024_009
Number of Mobile Home Housing Units	B25024_010 + B25024_011
Total Households	B25044_001
Number of 0-car Households	B25044_003 + B25044_010
Number of 1-car Households	B25044_004 + B25044_011
Number of 2-or-more-car Households	B25044_001
	– B25044_003 – B25044_004
	– B25044_004 – B25044_004

LONGITUDINAL EMPLOYER-HOUSEHOLD DYNAMICS

Longitudinal Employer-Household Dynamics (LEHD) Workplace Area Characteristics (WAC) data from 2018 (except for Alaska, where 2018 data is unavailable so 2016 data was used) was used as a source for private sector employment data. Public sector jobs and public administration jobs (NAICS 92) were excluded from the dataset. The raw data was used to calculate the total number of jobs (JOB_TOTAL), an estimate of tradable jobs (JOB_TRADE), and the number of retail jobs (JOB_RETAIL) in each block group. Tradable jobs were calculated based on estimates of the share of tradable jobs in each two-digit NAICS code provided in Chapter 2 of Global Trade in Services: Fear, Facts, and Offshoring by J. Bradford Jensen of the Peterson Institute for International Economics.[5] All jobs Jensen identified as intermediately or most tradable were treated as tradable jobs, and the number of jobs in each two-digit NAICS code was multiplied by the resulting fraction (shown in Table 2 below) before being summed to produce the JOB_TRADE variable.

In addition, counts of private sector jobs in 2011 (based on 2011 LEHD data), by income level (<$1,250/month, $1,250-$3,333/month, >$3,333/month), education level (no high school diploma, high school but no college, some college, or bachelor's degree or higher), and industry (retail, office, hospitality, or education/medical) were calculated for 2018 (or 2016 for Alaska) and jobs by firm age for 2017 (or 2016 for Alaska) were added for use in characterizing activity centers. Variable names and the values they represent can be found in the data dictionary for the block group level data attachment.

LEHD data is currently only available in 2010 geographies, but is available at the block level, so 2020 block group values were estimated using the Version 0 NHGIS 2010-to-2020 block crosswalks to convert 2010 block data to 2020 blocks, which were then aggregated to 2020 block groups.[6] The Version 0 crosswalks, which are based solely on land area overlap, were used because the population-distribution methods used to produce the Version 1 crosswalks don't make sense for estimating the distribution of jobs.

TABLE 2

Percent Tradable Jobs by Industry

LEHD Data Code	NAICS Code	Industries	Percent Tradeable Jobs
CNS01	11	Agriculture, Forestry, Fishing, and Hunting	100%
CNS02	21	Mining and Petroleum Extraction	100%
CNS03	22	Utilities	19.1%
CNS04	23	Construction	0%
CNS05	31-33	Manufacturing	87.9%
CNS06	42	Wholesale Trade	54.2%
CNS07	44-45	Retail Trade	14.9%
CNS08	48-49	Transportation and Warehousing	78.6%
CNS09	51	Information	66.8%
CNS10	52	Finance and Insurance	68.0%
CNS11	53	Real Estate, Rental, and Leasing	90.9%
CNS12	54	Professional, Scientific, and Technical	86.1%
CNS13	55	Management of Companies and Enterprises	100%
CNS14	56	Administrative, Support, and Waste Management	40.5%
CNS15	61	Educational Services	1.1%
CNS16	62	Healthcare and Social Assistance	2.2%
CNS17	71	Arts, Entertainment, and Recreation	32.7%
CNS18	72	Accommodation and Food Services	18.1%
CNS19	81	Other Services	20.3%

IMLS PUBLIC LIBRARIES SURVEY

The LIB_SQFT variable, measuring the square feet of library space in each block group, was derived from the Institute of Museum and Library Services' (IMLS) Public Libraries Survey (PLS). The results of the FY2019 PLS were imported as a CSV and filtered to remove bookmobiles, book-by-mail services, and closed branches.

This data was supplied as a CSV file with no CRS specified, so we assumed the coordinates given were NAD83 latitude/longitude (EPSG:4269) and converted them to a spatial object with EPSG:4269.

DHS HOMELAND INFRASTRUCTURE FOUNDATION-LEVEL DATA

The U.S. Department of Homeland Security's Homeland Infrastructure Foundation-Level Data (HIFLD), which contains point location data on a number of public facilities, was used to identify the locations of stadiums, colleges, hospitals, and government buildings. This data allowed the identification of several important types of activity that have a significance beyond their direct impact in terms of jobs. When possible, the sizes of these facilities were measured based on their metadata. Unfortunately, there is no size data available for courthouses, state government buildings, and state capitol buildings, so the counts of these are simply combined into a single variable

HIFLD consists of a number of separate datasets of the locations and additional information about a variety of different public facilities. These datasets were largely downloaded in early 2020. Seven datasets were used:

Major sports stadiums became the STAD_SEAT variable (a count of the total stadium seats in each block group, derived from the dataset's POPULATION variable) and the STAD_CNT variable (a count of the total number of stadiums in each block group). Stadiums listed as "CLOSED" and the roughly one-eighth of stadiums with seating capacities listed as 0 (consisting of golf courses and car races held on city streets) were dropped from the dataset.

Colleges and universities became the COLLEG_CNT variable (a count of the number of colleges and universities in each block group), the COLLEG_STU variable (a count of the total students at colleges and universities in each block group, derived from the TOT_ENROLL variable), and the COLLEG_EMP variable (a count of the total college and university employees in each block group, derived from the dataset's TOT_EMP variable). For-profit schools, non-degree-granting institutions, closed institutions, and institutions with enrollment or staff listed as 0 were excluded.

Courthouses were included in the GOVBLD_CNT variable—a count of courthouses, state government buildings, and state capitol buildings in each block group.

Major state government buildings were included in the GOVBLD_CNT variable—a count of courthouses, state government buildings, and state capitol buildings in each block group.

State capitol buildings were included in the GOVBLD_CNT variable—a count of courthouses, state government buildings, and state capitol buildings in each block group.

GSA office space became the GSA_SQFT variable (a count of the number of occupied square feet of GSA office space in each block group, derived from the difference between the BUILDING_RSF and BLD_VACANT_RSF variables in the dataset) and the GSA_CNT variable (a count of the number of GSA facilities in each block group). Only the roughly 92% of buildings listed as "ACTIVE" and with valid coordinates listed were included.

Hospitals became the HOSP_BED variable (a count of the number of hospital beds in each block group, derived from the BEDS variable in the dataset) and the HOSP_CNT variable (a count of the total hospitals in each block group). "CLOSED" hospitals were excluded, and only general acute care, critical access, military, children's, and women's hospitals were included.

The GSA office space data was supplied as a CSV file with no CRS specified, so we assumed the coordinates given were NAD83 latitude/longitude and converted them to a spatial object with EPSG:4269.

SAFEGRAPH POINTS-OF-INTEREST DATA

While educational facilities and hospitals are represented in the HIFLD data, retail and non-hospital medical offices are not. In addition, while HIFLD does supply a religious institutions dataset, it is woefully incomplete and largely limited to Christian institutions. Proprietary data on the location of “points of interest” in May 2020 from SafeGraph was used to supply this information.

The SafeGraph dataset used contains latitude/longitude coordinates of roughly 6 million facilities, along with their six-digit NAICS industry codes. Since no coordinate reference system is specified in the dataset, the NAD83 ellipsoid, EPSG:4269 was assumed. We filtered the data by NAICS code into nine categories, shown in Table 3, and a separate point object was created for each. Unfortunately, there is no metadata in the SafeGraph dataset that allows us to estimate the size of these assets, so each variable is a count of the number of assets of a type in each block group.

TABLE 3

SafeGraph Data Imported to Block Groups

NAICS Code	Description	Variable Name in Block Group Shapefiles
813110	Religious institutions	RELIG_CNT
712110, 712130	Museums and zoos	MUSE_CNT
721110, 721191	Hotels (except casino hotels), motels, and bed-and-breakfasts	LODGE_CNT
721120, 713210	Casinos, including casino hotels	CASINO_CNT
722400–729999	Restaurants and bars	RESTA_CNT
442000–446999, 448000–453999, 517312, 5221xx, 5322xx, 6244xx, 713940, 8114xx, 812xxx	Retail establishments, except for gas stations and car sales and repairs	RETAIL_CNT
621100–624399	Medical offices	MEDIC_CNT
51213x, 712110, 713120, 713950, 713990	Movie theaters, amusement parks, bowling alleys, and other amusements	AMUSE_CNT
491110	USPS post offices	SHIP_CNT

USGS NATIONAL LAND COVER DATABASE

We used developed land area rather than total land area in the denominator of density calculations to account for suburban block groups that contain both relatively dense development and large tracts of undeveloped land. The developed land area for each 2020 block group was calculated from land cover rasters: the 2019 USGS National Land Cover Database (NLCD) for the contiguous U.S., the 2016 NLCD for Alaska, and 2014-2015 NOAA land cover data for Hawaii. In each case, the most recent available data was used.

The NLCD data consists of 30-m resolution pixels while the NOAA data used for Hawaii consists of 1-m resolution pixels. In both cases, the fractions of the land area in each 2020 census block group covered by pixels coded as "developed open land" (< 20% impervious surface), "low-intensity developed land" (20% to 50% impervious surface), "medium-intensity developed land" (50% to 80% impervious surface), and "high-intensity developed land" (> 80% impervious surface) were counted and approximated as 10%, 35%, 65%, and 90% impervious surface, respectively. These values were then used to calculate the developed land area in square miles, DAREA_MI, defined as the area covered by artificial impervious surfaces.

NATIONAL REGISTER OF HISTORIC PLACES HISTORIC SITE DATA

The National Register of Historic Places historic site dataset, downloaded October 2020, was used to identify historical sites. Although the dataset contains counts of objects, buildings, and structures at each site, these counts were found to be inconsistent and unreliable, so a count of the number of sites was used for the HIST_CNT variable reflecting the number of National Register historic sites in each block group.

BTS INTERMODAL PASSENGER CONNECTIVITY DATABASE

The U.S. Department of Transportation Bureau of Transportation Statistics' Intermodal Passenger Connectivity Database (IPCD), downloaded in November 2020, was used to identify the locations of intercity rail stations and airports. Bus stations were excluded because it proved impossible to distinguish between actual bus stations and, for example, garages where bus companies store their vehicles or perform maintenance. In addition, the dataset contained numerous duplicates, such as multiple terminals at the same airport listed separately, so only a binary flag, TRANS_BIN, indicating the presence of intercity transportation facilities in a block group, was used.

ESRI PARK POLYGONS DATA

The ESRI parks dataset was downloaded in January 2020 and consists of polygons identified as local, county, regional, state, and national parks. The local/county/regional park distinction is essentially meaningless for comparison between metro areas, as different areas handle park governance entirely differently, but we removed state and national parks from the dataset because these parks tend to be much larger and correspond to wilderness areas rather than urban or community land use. The PARK_CNT variable records the number of park polygons intersecting with a given 2020 block group, meaning that a single park could be recorded as present in more than one block group.

EPA SMART LOCATION DATABASE

The EPA Smart Location Database (SLD) is a national database of transportation-related data, including road network connectivity and walkability, public transportation access, vehicle miles travelled, and transportation greenhouse gas releases, tabulated at the block group level.[7] It also contains National Walkability Index (NWI) scores, a ranking of census block groups according to relative walkability based on density of land uses, proximity to transit stops, and intersection data.

We used the most recent edition (2021), which is tabulated for 2010 block groups, and converted these values to 2020 block groups using Census Bureau block group to block group crosswalks based on land area overlap.[8] Values for the NWI and other SLD data in 2020 block groups were calculated using weighted means, weighted by the amount of land area overlap, except for total greenhouse gas emissions, which were calculated as a sum weighted by land area overlap. Overall walkability scores for metro areas' activity centers were calculated as the averages of the NWI scores for all activity centers in the metro area.

IDENTIFYING ACTIVITY CENTERS

We identified activity centers based on the presence of assets in five categories: community assets, tourism assets, consumption assets, institutional assets, and economic assets. Each asset category was assessed using one or more measures, which are listed below by asset category. Densities per square mile of developed land area (not total land area) are used for most assets. Counts were used for institutional assets, as these assets serve areas much larger than a block group. Binary flags are used for the presence of sports stadiums and libraries (which also serve large areas but where size is less important) and for transport facilities (due to duplicates in the data).

- Community assets
  - Density of population (Brookings analysis of 2020 census data)
  - Binary flag for public library (IMLS FY2019 Public Libraries Survey)
  - Density of places of worship (Brookings analysis of SafeGraph data)
  - Density of historic sites (Brookings Analysis of National Registry of Historic Places data)
  - Density of major and minor parks (Brookings analysis of ESRI data)
- Tourism assets
  - Binary flag for major sports stadium (HIFLD)
  - Density of hotels and motels (Brookings analysis of SafeGraph data)
  - Density of casinos and museums (Brookings analysis of SafeGraph data)
- Consumption assets
  - Density of restaurants (Brookings analysis of SafeGraph data)
  - Density of retail establishments (Brookings analysis of SafeGraph data)
  - Density of medical offices (Brookings analysis of SafeGraph data)
  - Density of amusement parks and theaters (Brookings analysis of SafeGraph data)
  - Binary flag indicating presence of a post office (Brookings analysis of SafeGraph data)
  - Density of retail jobs (Brookings analysis of LEHD data)
- Institutional assets
  - Count of college students, faculty, and staff (HIFLD)[9]
  - Count of hospital beds(HIFLD)5
  - Count of state courthouses, state office buildings, and statehouses (HIFLD)
  - Square feet of GSA office space (HIFLD)
  - Binary flag for intercity rail or airport (BTS Intermodal Passenger Connectivity Database)
- Economic assets
  - Density of tradable private sector jobs (Brookings analysis of LEHD data)

The value for each measure was normalized by dividing it by the average value over all block groups in the metro area. Normalized values for each measure within an asset category were then summed to create an overall location quotient score for the asset category.

We identified activity centers based on their percentile ranking within the metro area for the five asset category location quotient scores. Block groups with land areas of at least 100 square miles were excluded from the calculation of these percentile ranks on the basis that they are too large to reasonably be described as "centers." These block groups, which only occur in very low-density rural areas due to the Census Bureau requirement that block groups have a roughly consistent population, would otherwise sometimes score highly because they combine multiple small communities spread over a large area.

Each block group was assigned a center based on its percentile score location quotients for the five asset categories.

- Primary center: Two or more location quotient scores above 98th percentile
- Secondary center: Two or more CT scores above 95th percentile, but doesn't qualify as primary center
- Monocenter: One CT score above 98th percentile, but doesn't qualify as primary or secondary center
- Non-center: All other block groups

Block groups with land areas over 100 square miles were also classified as non-centers, and were given location quotient scores of zero for all five asset categories.

CHARACTERIZING ACTIVITY CENTERS

CALCULATING ACTIVITY CENTER JOB DENSITY AND JOB SHARE

Although there are a number of different components to activity center strength, we used two primary measures: overall job density and overall job share. The presence of jobs across all sectors was used as a rough, one-dimensional measure of the amount of activity in centers for the purpose of comparing the "strength" or overall economic and social significance of centers to each metro area. The total private sector jobs counts from the 2018 LEHD (2016 LEGD for Alaska) dataset served as the basis for these calculations. However, since LEHD public sector jobs data is not reliable at the block group level, a way to measure government jobs was needed as well.

Although it did not prove possible to take state and local jobs into account, federal government jobs were approximated using the floor space values listed in the HIFLD dataset of federal buildings operated by the General Services Administration (GSA). The number of federal workers in each block group was estimated as the usable square footage of GSA space divided by 190, based on the GSA's 2012 recommendations for the amount of office space per employee in federal facilities.[10] Total job count was then figured as the sum of private sector jobs from the LEHD data and estimated federal jobs.

The job density measure was calculated with developed land area (as estimated from NLCD data) rather than total land area in the denominator to better account for block groups that contain both developed areas and parkland or other open space—something that is relatively common, especially in suburban job clusters.

A job-weighted median, rather than an average, was used to measure the density that the median worker experiences and avoid the overweighting of relatively large but low-density activity centers on the outskirts of metro areas. In addition, using a *weighted* median is particularly important with jobs, because the variation in the number of jobs per block group is much greater than the variation in the population of block groups, since block groups are constructed to be of roughly equal population.

Like activity center job density, the share of metro area jobs located in activity centers was calculated based on the sum of private sector jobs from LEHD data and an estimate of one federal job per 190 square feet of GSA office space.

ANALYZING COMMERCIAL AND RESIDENTIAL REAL ESTATE VALUE

Our analysis of the share of commercial and residential real estate value in activity centers is based on the Zillow ZTRAX database of real estate. Calculations are based on total (structure plus land) tax-assessed values for commercial-office and commercial-retail properties (for commercial real estate) and on total tax-assessed values divided by structure square feet for residential and income-generating residential properties (for residential real estate). We did not evaluate industrial land values because location data was missing for too many industrial properties, and we did not evaluate agricultural land values because agriculture is a fundamentally non-urban land use that is unlikely to benefit from proximity to activity centers or density.

The Zillow analysis was limited to 45 metro areas for commercial property and 44 metro areas for residential property (indicated in the table in the Appendix) out of the 110 largest metro areas used in the rest of our analysis. These metro areas were selected because 90% or more of both residential and commercial plots had values listed and plots making up at least 80% of both the residential and the commercial real estate value had locations listed. Since most counties in the dataset had locations listed for less than 100% of commercial and residential plots with values listed, the commercial and residential values of plots in each county with locations listed were scaled up to sum to the total commercial and residential land values in the county. The Columbia, S.C. metro area was excluded from the residential property analysis because structure floor area values for much of the metro area were not reliable.

FITTING ACTIVITY CENTER JOB DENSITY TO PRODUCTIVITY

The relationship between primary center strength and productivity at the MSA level was determined using 2019 MSA-level gross metropolitan product (GMP) per job values from Emsi as our measure of productivity. We tested a substantial number of potential independent variables, including activity center job density and job share, primary center job density and job share, the share of jobs in technology sectors (NAICS codes 51 and 54), and share of the population with college degrees, but found that the best fit was to a single independent variable: activity center job density. The results of that regression are shown in Table 4.

TABLE 4

Regression Coefficients for GMP per Job

Variable	Coefficient	Std. Error	p-Value	Fit
(Intercept)	9.907 × 104	3.238 × 103	< 2 × 10−16	
Activity Center Density (Jobs / Sq. Mi.)	1.723 × 100	2.081 × 10−1	5.32 × 10−13	
				R^2 = 0.404
				F-Statistic = 68.57
				Degrees of Freedom = 101

MEASURING ACTIVITY CENTER CENTRALITY

We measured the centrality of activity centers within regions to characterize their accessibility to each other and the region as a whole. For this, we used two different measures: their average distances from regional cores and the number of centers within 3 miles of other centers. In addition, we calculated the shares of several assets found in activity centers within the largest cities in each metro area. All distance-dependent calculations were performed using the appropriate UTM coordinate system, as indicated in the Appendix.

Due to the lack of reliable assignments of central business districts (CBDs) for all metro areas in our study (the Census Bureau last identified CBDs in 1982, and some of the newer metro areas we studied, especially those in Florida, do not contain Census Bureau-identified CBDs), we calculated our own approximations of regional cores. For each metro area, we approximated the location of the regional core as the centroid of jobs in primary centers. A simple average of distances was then used to determine the average distance of activity centers of each type from the regional core and the share within 3 miles from the core.

Activity center clustering was measured by calculating the average number of activity centers of each type and overall within 3 miles over the entire set of 110 metro areas. Because there was no intermediate step of calculating averages within each metro area, metro areas with more activity centers have a large impact on the overall values.

In addition to measuring clustering and distance from metro area cores, we assessed the centrality of activity centers by calculating the share of developed land area, jobs, and commercial land value in centers in the largest principal city in each metro area using 2020 TigerLine shapefiles for Census Bureau places. The largest place by population (whether incorporated or not) listed by the Census Bureau as a principal city or principal place was used for each metro area, and all centers that intersected with a principal city were counted as being within a principal city. For each measure and center type, both the share of the measure in centers of that type and in the metro area as a whole located in centers of that type in the largest city were calculated.

MEASURING ACTIVITY CENTER BUFFER DEMOGRAPHICS

For many purposes, the demographics of areas close to centers are important as those of the centers themselves, since proximity to a center increases the accessibility of jobs and amenities located there. One-mile (to approximate easy access by foot or transit) and 3-mile (to approximate easy access by car) buffers around the activity centers and the primary centers in each metro area were defined as the 2010 block groups with centroids within 1 or 3 miles of the centroids of the activity or primary centers. These buffers were defined in terms of 2010 block groups because the ACS data used for the buffer demographic measurements was only available in 2010 geographies.

METRO AREAS ANALYZED

All metropolitan statistical areas with populations of at least 500,000 residents according to the 2020 census—the 110 largest—were included. Smaller MSAs have too few block groups for using percentile scores of block groups to make sense. Spatial analysis for each metro area was done in the Universal Transverse Mercator (UTM) zone in which the centroid of the metro area's block groups is located.

FIPS CODE	Metropolitan Statistical Area	2020 Population	UTM EPSG Code	Included in Zillow Analysis?
10420	Akron, OH	702,219	32617	YES
10580	Albany-Schenectady-Troy, NY	899,262	32618	
10740	Albuquerque, NM	916,528	32613	
10900	Allentown-Bethlehem-Easton, PA-NJ	861,889	32618	
12060	Atlanta-Sandy Springs-Alpharetta, GA	6,089,815	32616	
12260	Augusta-Richmond County, GA-SC	611,000	32617	
12420	Austin-Round Rock-Georgetown, TX	2,283,371	32614	
12540	Bakersfield, CA	909,235	32611	
12580	Baltimore-Columbia-Towson, MD	2,844,510	32618	
12940	Baton Rouge, LA	870,569	32615	
13820	Birmingham-Hoover, AL	1,115,289	32616	
14260	Boise City, ID	764,718	32611	
14460	Boston-Cambridge-Newton, MA-NH	4,941,632	32619	YES
14860	Bridgeport-Stamford-Norwalk, CT	957,419	32618	YES
15380	Buffalo-Cheektowaga, NY	1,166,902	32617	YES
15980	Cape Coral-Fort Myers, FL	760,822	32617	YES
16700	Charleston-North Charleston, SC	799,636	32617	
16740	Charlotte-Concord-Gastonia, NC-SC	2,660,329	32617	
16860	Chattanooga, TN-GA	562,647	32616	
16980	Chicago-Naperville-Elgin, IL-IN-WI	9,618,502	32616	YES
17140	Cincinnati, OH-KY-IN	2,256,884	32616	
17460	Cleveland-Elyria, OH	2,088,251	32617	YES
17820	Colorado Springs, CO	755,105	32613	YES
17900	Columbia, SC	829,470	32617	YES[xi]
18140	Columbus, OH	2,138,926	32617	
19100	Dallas-Fort Worth-Arlington, TX	7,637,387	32614	YES
19430	Dayton-Kettering, OH	814,049	32616	
19660	Deltona-Daytona Beach-Ormond Beach, FL	668,921	32617	YES
19740	Denver-Aurora-Lakewood, CO	2,963,821	32613	
19780	Des Moines-West Des Moines, IA	709,466	32615	YES
19820	Detroit-Warren-Dearborn, MI	4,392,041	32617	
20500	Durham-Chapel Hill, NC	649,903	32617	
21340	El Paso, TX	868,859	32613	
22180	Fayetteville, NC	520,378	32617	
22220	Fayetteville-Springdale-Rogers, AR	546,725	32615	
23420	Fresno, CA	1,008,654	32611	YES
24340	Grand Rapids-Kentwood, MI	1,087,592	32616	
24660	Greensboro-High Point, NC	776,566	32617	YES
24860	Greenville-Anderson, SC	928,195	32617	
25420	Harrisburg-Carlisle, PA	591,712	32618	

25540	Hartford-East Hartford-Middletown, CT	1,213,531	32618	YES
26420	Houston-The Woodlands-Sugar Land, TX	7,122,240	32615	YES
26900	Indianapolis-Carmel-Anderson, IN	2,111,040	32616	
27140	Jackson, MS	591,978	32615	
27260	Jacksonville, FL	1,605,848	32617	YES
28140	Kansas City, MO-KS	2,192,035	32615	
28940	Knoxville, TN	879,773	32617	
29460	Lakeland-Winter Haven, FL	725,046	32617	YES
29540	Lancaster, PA	552,984	32618	
29620	Lansing-East Lansing, MI	541,297	32616	
29820	Las Vegas-Henderson-Paradise, NV	2,265,461	32611	YES
30460	Lexington-Fayette, KY	516,811	32616	YES
30780	Little Rock-North Little Rock-Conway, AR	748,031	32615	
31080	Los Angeles-Long Beach-Anaheim, CA	13,200,998	32611	YES
31140	Louisville-Jefferson County, KY-IN	1,285,439	32616	
31540	Madison, WI	680,796	32616	
32580	McAllen-Edinburg-Mission, TX	870,781	32614	
32820	Memphis, TN-MS-AR	1,337,779	32616	
33100	Miami-Fort Lauderdale-Pompano Beach, FL	6,138,333	32617	YES
33340	Milwaukee-Waukesha, WI	1,574,731	32616	YES
33460	Minneapolis-St. Paul-Bloomington, MN-WI	3,690,261	32615	
33700	Modesto, CA	552,878	32610	YES
34980	Nashville-Davidson--Murfreesboro--Franklin, TN	1,989,519	32616	
35300	New Haven-Milford, CT	864,835	32618	YES
35380	New Orleans-Metairie, LA	1,271,845	32615	
35620	New York-Newark-Jersey City, NY-NJ-PA	20,140,470	32618	
35840	North Port-Sarasota-Bradenton, FL	833,716	32617	YES
36260	Ogden-Clearfield, UT	694,863	32612	
36420	Oklahoma City, OK	1,425,695	32614	
36540	Omaha-Council Bluffs, NE-IA	967,604	32614	
36740	Orlando-Kissimmee-Sanford, FL	2,673,376	32617	
37100	Oxnard-Thousand Oaks-Ventura, CA	843,843	32611	YES
37340	Palm Bay-Melbourne-Titusville, FL	606,612	32617	YES
37860	Pensacola-Ferry Pass-Brent, FL	509,905	32616	
37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	6,245,051	32618	
38060	Phoenix-Mesa-Chandler, AZ	4,845,832	32612	YES
38300	Pittsburgh, PA	2,370,930	32617	
38860	Portland-South Portland, ME	551,740	32619	YES
38900	Portland-Vancouver-Hillsboro, OR-WA	2,512,859	32610	
39100	Poughkeepsie-Newburgh-Middletown, NY	697,221	32618	YES
39300	Providence-Warwick, RI-MA	1,676,579	32619	YES
39340	Provo-Orem, UT	671,185	32612	
39580	Raleigh-Cary, NC	1,413,982	32617	YES

40060	Richmond, VA	1,314,434	32618	
40140	Riverside-San Bernardino-Ontario, CA	4,599,839	32611	YES
40380	Rochester, NY	1,090,135	32618	
40900	Sacramento-Roseville-Folsom, CA	2,397,382	32610	
41180	St. Louis, MO-IL	2,820,253	32615	
41620	Salt Lake City, UT	1,257,936	32612	
41700	San Antonio-New Braunfels, TX	2,558,143	32614	
41740	San Diego-Chula Vista-Carlsbad, CA	3,298,634	32611	YES
41860	San Francisco-Oakland-Berkeley, CA	4,749,008	32610	YES
41940	San Jose-Sunnyvale-Santa Clara, CA	2,000,468	32610	YES
42540	Scranton--Wilkes-Barre, PA	567,559	32618	
42660	Seattle-Tacoma-Bellevue, WA	4,018,762	32610	
44060	Spokane-Spokane Valley, WA	585,784	32611	YES
44140	Springfield, MA	699,162	32618	YES
44700	Stockton, CA	779,233	32610	YES
45060	Syracuse, NY	662,057	32618	
45300	Tampa-St. Petersburg-Clearwater, FL	3,175,275	32617	YES
45780	Toledo, OH	646,604	32617	
46060	Tucson, AZ	1,043,433	32612	YES
46140	Tulsa, OK	1,015,331	32615	
46520	Urban Honolulu, HI	1,016,508	32604	YES
47260	Virginia Beach-Norfolk-Newport News, VA-NC	1,799,674	32618	
47900	Washington-Arlington-Alexandria, DC-VA-MD-WV	6,385,162	32618	YES
48620	Wichita, KS	647,610	32614	
49180	Winston-Salem, NC	675,966	32617	
49340	Worcester, MA-CT	978,529	32619	YES
49660	Youngstown-Warren-Boardman, OH-PA	541,243	32617	

END NOTES

1 Edgeless Cities by Robert Lang
2 Defining the 15-minute city | CNU
3 Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. 2021. http://doi.org/10.18128/D050.V16.0 and https://www.nhgis.org/geographic-crosswalks
4 https://www.census.gov/geographies/reference-files/time-series/geo/relationship-files.html#blkgrp
5 Citation for Chapter 2 of Global Trade in Services: Fear, Facts, and Offshoring by J. Bradford Jensen of the Peterson Institute for International Economics
6 Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. 2021. http://doi.org/10.18128/D050.V16.0 and https://www.nhgis.org/geographic-crosswalks
7 https://www.epa.gov/smartgrowth/smart-location-mapping#SLD
8 https://www.census.gov/geographies/reference-files/time-series/geo/relationship-files.html#blkgrp
9 Our college and hospital data comes in the form of point objects, but these facilities are often large and cross block group borders. As a result, large colleges and hospitals may be represented by a single block group despite actually having portions in several.
10 https://www.gsa.gov/cdnstatic/Workspace_Utilization_Banchmark_July_2012_%281%29.pdf

Brookings Metro

ANNE T. AND ROBERT M. BASS CENTER FOR
TRANSFORMATIVE PLACEMAKING

1775 Massachusetts Ave NW,
Washington, DC 20036
(202) 797-6000
www.brookings.edu