City definition affects long-term urban scaling analyses in the United States (1900 - 2015)

Keith Burghardt∗ USC Information Sciences Institute, Marina del Rey, CA 90292, USA

Johannes H. Uhl Institute of Behavioral Science (IBS), University of Colorado Boulder, 483 UCB, Boulder, CO-80309, USA and Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, 216 UCB, Boulder, CO-80309, USA

Kristina Lerman USC Information Sciences Institute, Marina del Rey, CA 90292, USA

Stefan Leyk Institute of Behavioral Science (IBS), University of Colorado Boulder, 483 UCB, Boulder, CO-80309, USA and Department of Geography, University of Colorado Boulder, 260 UCB, Boulder, CO 80309, USA

(Dated: September 28, 2022)

The scaling relations between city attributes and population are emergent and ubiquitous aspects of urban growth. Quantifying these relations and understanding their theoretical foundation, however, is difficult due to the challenge of defining city boundaries and a lack of historical data to study city dynamics over time and space. To address this issue, we analyze scaling between city infrastructure and population across 857 United States metropolitan areas over an unprecedented 115 years using dasymetrically refined historical population estimates, historical urban road network data, and multi-temporal settlement data to define dynamic city boundaries based on settlement density. We demonstrate the clearest evidence that urban scaling exponents can closely match theoretical models over a century if cities are defined as dense settlement patches. In contrast, these results strongly disagree with theoretical models and vary in time if we aggregate cities into administrative-defined metropolitan areas, a method used in previous analyses. Despite the close quantitative agreement with theory, the empirical scaling relations unexpectedly vary across space, a finding that is supported by global scaling analysis. Finally, our analysis of scaling coefficients reveal that a city in 2015 uses more developed land and kilometers of road than a city in 1900 with a similar population, which may imply cities have become increasingly harmful to the local environment. Our results show the strengths and limitations of current theoretical models, and offer a new way to study urban systems based on historical data.

INTRODUCTION

Cities are expanding at a never-before seen pace, with the majority of the world’s population now living in urban environments [1], and a near-majority in each of the world’s continents [2]. Cities encompass highly diverse geographies and populations, both in size and composition, yet seemingly robust urban patterns emerge from this complexity [3][4]. These patterns are important not only because of the significance of cities and because they provide a uniform reference frame for comparing cities, but also because they offer a framework to understand emergent phenomena in complex systems. Among the most widely studied examples of urban patterns are scaling relations between features of a city and its population size [11][13][14], of the form

\[
\text{City property} = \exp(b) \times P^a,
\]

where \(P\) is the city population, \(b\) is the scaling coefficient, and \(a\) is the scaling exponent. We define the constant to be \(\exp(b)\) such that a log-log plot will have the relation \(\log(\text{City property}) = a \times \log(P) + b\). The universality of these patterns has recently been called into question, however, because different definitions of cities [13][15] produce very different scaling relations [11][12][16][17], and, in contrast to theoretical predictions, the scaling exponents appear to vary in time [10][18][21]. This leaves a surprising, if unsettling, problem for researchers: how do we define a city boundary? And what changes, if any, are needed in our modeling framework to encapsulate these findings? In this paper, we argue that both problems are built upon discrepant and static definitions of a city boundary. For example, studies of urban scaling [19][22] were based on static administrative boundaries [11][18][21], but artificial boundaries can lead to incorrect statistical conclusions [23][24]. Moreover, time-varying scaling laws [19][21] could be affected by the city’s expanding boundary, a feature administrative boundaries cannot capture.

In this paper, we address these limitations with an intuitive definition of city boundaries (also known as spatial

* keithab@isi.edu
analitical units [23] based on settlement densities. This definition, however, lacks an administrative boundary usually needed to estimate population size. We therefore develop estimations of city populations based on dasyometric refinement [26,28], in which buildings are used to help estimate populations at resolutions finer than what the United States (US) census data provide. The estimates are used to model populations of cities and their relationships to urban properties over time. Namely, better knowledge of scaling over time means we better understand how urban environments have developed in relationship to their population. This helps us estimate how this environment evolves with regard to, for example, public service, infrastructure, and built development. We apply these two methodological advances to a new dataset that records how cities have evolved since 1900, allowing us to audit several different predicted urban patterns for 857 metropolitan areas within the conterminous US (CONUS). We also compare our results against similar findings for cities across the world in 2013 that are defined by population density. Our results complement some previous work by Arcuate et al., [24] who created population density-based city boundary definitions, although our analysis expands on their work with building-level city boundary resolution and we test the empirical city boundary definition across a century of historical data.

Because the definition of a city, or city core, is an inherently vague concept [29], we use three different definitions of city boundaries: cities aggregated up to their administrative metropolitan boundaries, which is similar to previous city boundary definitions [13], cities disaggregated within administrative boundaries based on their settlement density, and settlement-density based cities that span multiple administrative boundaries. While we show that the first definition poorly represents urban scaling patterns, the settlement density-based disaggregation shows stronger patterns of urban scaling, including more stable scaling exponents, improved agreement with scaling laws based on the $R^2$ fitting metric, and better agreement with theoretical models [14]. We show that the last definition is the most natural one that avoids delineation of cities that span administrative boundaries from being artificially severed. Therefore, in contrast to previous analyses, this work shows that a theoretical model might explain observed empirical findings. That said, we see regional differences in scaling laws, from the CONUS to the world, in disagreement with model predictions. Models therefore need to account for spatial variations in urban constraints to account for these effects. Finally, our scaling analysis reveals a substantial increase in developed land, building footprints, indoor area, and km of road per capita, after controlling for population. Such observations could indicate that cities are using much more otherwise undeveloped land, which may overburden resource management and be harming the surrounding environment [30, 32]. Overall, our work gives new insights into the persistence of urban patterns, while revealing how different geographies, as defined by e.g., topography, may nonetheless impact seemingly universal scaling relations. This adds nuance to theoretical arguments behind and implications of urban scaling [14].

## RESULTS

The authors recently developed an approach and data infrastructure to measure the long-term evolution of settlements [35] and road networks for 850 cities within the US since 1900 [34], which adds to the growing literature on reconstructing land use over time [36, 37]. We apply these data to reconstruct the infrastructure of cities as they grow, as shown in Fig. [1]. Because cities are an inherently vague concept, where different people can define the core of a city differently [29], we analyze three city boundary definitions to gain insights into scaling relations over time. First, we treat cities as areas with dense built up settlements within Core-Based Statistical Areas (CBSAs), which we call Density-Based Urban Areas (DBUAs). We separately define cities as DBUAs that are merged together if they cross CBSAs, which we call CBSA-Independent Urban Areas (CIUAs). This last definition belies the heterogeneity of urbanization within the region, and therefore masks where cities may actually be located [38]. These are shown as colored DBUAs in Fig. [1]–b for 1900 and 1950, respectively. Finally, we aggregate all DBUAs within a CBSA to approximate the ways data has been studied in previous work [13], which we call Density-Based Urban Areas (DBUAs). In the rest of the paper, we will focus our analysis on DBUAs, which are similar to city boundaries used in prior work (e.g., [13] [14] [19] [21]), and CIUAs, which are the most natural settlement density-defined cities. We show consistency of our result between CIUAs and DBUAs in the Supplementary Figures.

We calculated city properties within these boundaries, namely their total developed area, indoor area, building footprint area, road length, and population size (see Materials and Methods for details). Developed area is the land area of developments (Fig. [1]), while building footprint area is the acreage all buildings within a city boundary (Fig. [1]), Indoor area (Fig. [1]) meanwhile represents the total building floorspace across all floors rather than the footprint acreage. Road length is defined as the total kilometers of road found within each city boundary (Fig. [1]).

Finally, we reconstructed historical populations within city boundaries at a given point in time using the CBSA-level census population estimates [39, 41] and HISDAC-US building data [35]. We observe a strong correlation between the number of buildings and population within a given county (Supplementary Figure S1). The HISDAC-US data tells us the cumulative number of buildings within a city at a given year, therefore the fraction of buildings within the MSA that existed in a city by a given year is related to the approximate fraction of the
FIG. 1. Different concepts of analytical units for urban scaling analysis and geospatial data used to measure historical city properties for different points in time. (a) Density-Based Urban Areas (DBUAs), (b) Density-Based Urban Patches (DBUPs), and (c) CBSA-Independent Urban Patches (CIUPs) in 1900 and 1950 for the Greater Denver area (Colorado), derived from the historical, generalized urban areas dataset [33]. (d) Developed area, and (e) historical building indoor area estimates within CIUPs (yellow), as a measure of built-up intensity, and (f) historical building footprint estimates, shown in black, within a DBUP in 1900 and 1950. (g) Historical, local road networks reconstructed from [34] from before 1900 through 2000. A modern interstate is shown in white.

total population within that city in that year. The historic population within a city can therefore be estimated based on the MSA population of that year times the fraction of buildings that existed in a city that year. City boundaries based on built up intensity are therefore very similar to those based on population density [24], or similar metrics [6, 16, 42]. While there are a number of other ways to define cities [17, 43, 44], including those based on workforce and consumption, we use the current method because it is a reasonable proxy of these alternative definitions (people often travel near their home [45], making area around a settlement location a good proxy of where people travel). Moreover, our definition is robust enough that we can uncover city boundaries far back in time.
We confirm the accuracy of this method in Supplementary Figures S2 & S3, which shows strong correlations between building-based population estimates and more sophisticated population density estimations [28, 46]. We also show in Supplementary Figure S4 that the different population estimation methods have a minimal effect on scaling laws.

Scaling for different city boundaries

We show urban scaling relations in Figs. 2 & 3. In the first row of each figure, we define a city as all urban regions within the 2010-era DBUAs and therefore aggregate all DBUPs (like those seen in Fig. 1) within the DBUA boundary. The urban scaling laws are near-linear or even superlinear around 1900, implying that larger cities take up more land and have more indoor area per capita, but nearly the same length of roads compared to smaller cities. These exponents, however, drop dramatically to sublinear scaling relations for footprint area and road length as of 2015. This suggests that cities have a strong economy of scale, but the exponents disagree with empirical predictions [13], shown as dashed lines in Fig. 3. Cities also have a nearly linear relation between these areas and population, therefore the area per capita is nearly constant, whether someone lives in a large or small city. However, we observe statistically significant sublinear relations between footprint area and population, suggesting larger cities have smaller building footprints per capita. These results show consistency with previous analysis, which found scaling exponents vary over time [20, 21].

We contrast the previous results with cities based on CIUPs of highly built up land. In this case, the growth patterns are found to be more stable, and are in better agreement with theory as shown in the second row of Figs. 2 & 3. Namely, CIUPs produce a consistently sublinear scaling between developed area or road length and population, and are in close agreement with the 2/3 law for area, and 5/6 law for road length as previously reported [14]. Moreover, we find a consistent linear relation between building footprint area or indoor area and population, meaning that large and small cities have similarly sized homes on average. All results are found to be robust to missing data and CIUP results are consistent with the alternative DBUP city boundaries (see Supplementary Figures S5–S8). Finally, CIUPs have a higher R^2 than DBUAs, demonstrating better agreement with a scaling law (see Supplementary Figures S9–S11). The results we see for CIUPs are consistent with those of DBUPs. The exponent variability seen in DBUAs is therefore due to treating both large and small towns (different sized DBUPs) as though they belong to the largest city, which creates issues related to the modifiable area unit problem [17].

SPATIAL VARIATIONS IN CROSS-SECTIONAL SCALING

We can also split these data into regions within the CONUS, and test how scaling laws vary across both space
and time. We show an example of this in Fig. 3, in which we compare scaling laws based on all cities (“aggregated”) to scaling laws within the Northeast, Midwest, South, and West. Regardless of how city boundaries are defined, scaling law exponents consistently differ between regions; exponents in 2015 are typically statistically significantly different across regions (z-score p-values < 0.001). Surprisingly, we find that the Northeast has consistently higher CIUP-scaling exponents while the South and West have lower exponents. These lower exponents imply the South or West may have a greater economy of scale because their area or road length per capita is smaller for larger cities, thus they better (or more efficiently) utilize the infrastructure they have.

Separately, we analyzed previously collected data from [48, 49] on global urban development, as shown in Fig. 4a–c. In this work, city boundaries are based on coarse-resolution (1 km$^2$) gridded population density estimates, in contrast to the results above, which are based on fine-grained footprint-level settlement density. In Fig. 4d–f, we show that exponents significantly vary (ANOVA p-values < 0.001), which is confirmed by recent analysis of area-scaling of the same data [48], but does not strongly depend on the latest GDP per capita (Spearman rank correlation p-values > 0.3). We also notice, however, that scaling exponents are typically above US-based empirical data (Fig. 3) and theoretical predictions shown as dashed lines. This is likely due to differences in methodology and coarse population resolution [50], which can affect the smallest cities. Namely, small cities may largely lie within coarse-resolution areas used as analytical units.

### SCALING COEFFICIENTS ACROSS SPACE AND TIME

In Fig. 4a–f we also show the scaling coefficients as functions of the latest GDP per capita. Spearman rank correlation is not significant for urban center area (p-value= 0.2), but is 0.62 for built-up area (p-value= 0.006) and 0.60 for road length (p-value= 0.009), where the former two are proxies of developed area (see Supplementary Figure S12). The scaling coefficients differ across the world by roughly 0.5, which suggests that, given a roughly fixed scaling exponent, GDP increases some urban features by up to 65%.

To further understand how scaling coefficients vary, we plot coefficients across space and time in Fig. 5. We notice that, regardless of how cities are defined, the scaling coefficient depends strongly on time, most notably for developed, indoor, and footprint areas. Given the stable scaling exponent for CIUPs, this finding implies the developed and housing area per capita has increased over time (results for DBUPs are consistent with these findings, cf. Supplementary Figure S7–S8). We independently confirm this latter finding using US Census data on newly built homes since 1978 (Supplementary Figure S13), in which we find the average house area has consistently increased until 2010–2015. We also notice that developed area continues to increase over time, with one of the largest jumps occurring between 1950–1960, with a change of ~ 0.5 (implying developed area per capita increased by around 65%). This finding consistent with post-WWII housing boom [51], and white flight [52], which would have contributed to a greater build-
up of housing in suburbs and less-developed areas, after controlling for city population. Finally, using CIUP city boundary definitions, we notice that the road length has recently decreased for cities even as their developed area stagnates. This might imply that cities have more recently become more sparse, but more work needs to confirm these findings. In conclusion, we find that scaling coefficients are not universal, and depend on GDP per capita as well as vary strongly in time. Their increase also implies a greater amount of land used for development than a comparable city from 1900.

**DISCUSSION AND CONCLUSIONS**

In this study, we have tested the robustness of scaling laws across space and time for different city boundary definitions. We find that disaggregating the analytical unit used to define cities into density-based city boundaries produces scaling laws in good agreement with theoretical models and data contains less scatter than using the CBSAs as analytical units. Despite the disagreement on how to define cities, and the inconsistency of scaling laws in different studies, we show that a fundamental limitation of previous work is defining cities from static administrative boundaries. We find agreement with a theoretical model when we define cities by CIUPs (or DBUPs) but not more aggregated DBUA. This difference could help explain why previous research has found exponents vary in time [20, 21] because data was built from administrative-defined city boundaries. These boundaries can create inconsistent patterns because of the modifiable areal unit problem [23, 24], and because different cities might be unintentionally grouped together, therefore statistics could suffer from Simpson’s paradox [25]. An empirical definition of cities accounts for cities expanding and evolving, thus capturing when cities expand beyond their administrative boundaries. We also find that urban scaling laws vary across space, in disagreement with theoretical models. This could be due to topographic constraints such as rivers or mountains, which can affect urban density [25] but also other constraining factors.

Regardless of the city boundary definition, the long time window also allows us to explore an under-appreciated component of scaling: the scaling coefficient shows an increase over time for most properties, as seen in Fig. 3. Therefore, developed area, indoor area, footprint area, and road length per capita are increasing on average. This reduces the amount of otherwise undeveloped land, which complicates resource use and harms the local environment, including water quality and wildlife habitat [26, 27]. While previous work has found the developed land base has recently become denser [28], our work implies that a city 100 years ago was still much more densely developed than the same city would be today. This reduced density is consistent with previous work on the loss of dense urban environments [25], although scaling analysis allows us to correctly control for city population and show consistent results for thousands of settlement density-defined cities. One reasonable hypothesis is that the reduced urban density is a function of increased car use throughout the century [30], but we can only draw associations not cause and effect from these data, and previous work found car ownership can be correlated with increasing urban density [31]. More research is therefore needed to understand this trend.

These results have important consequences for the urban sciences. First, as previously demonstrated, objective definitions of city boundaries are needed to study the evolution of urban areas over long spans of time. Second, and equally important, urban systems need to be studied based on historical data to fully understand their development trajectories. While previous work has made impressive advances in understanding historical urban development [15, 48, 58, 59], there has been a lack of analysis on scaling laws for the same cities over time. Our analysis exploits recently created historical urban data, which allows scaling comparisons to be performed across the same cities over a century. More analysis of this type, especially for superscaling variables like income [13, 20] need to be explored. Finally, regional variations need to be explored. While universal laws are idealized explanations of urban science, regional variations can give insights into region-specific conditions. This can improve the interplay between urban science and policy-makers who focus on region- and city-specific problems.

There are certain limitations in the data we analyze. For example, built years are missing for many records, and historical data is missing in a non-random way [60], leading to survivorship bias. Furthermore, we assume that roads were constructed at approximately the same time as nearby buildings and that population is proportional to the number of buildings within a given DBUP. We evaluated the reliability of such assumptions to the best of our ability, but more research is needed to test and improve upon them. For example, it is important to uncover new ways to approximate DBUP-level population estimations, as well as city infrastructure over time. In addition, our analysis of city-wide properties belies variations within cities that need to be further analyzed and modeled [61]. Finally, there are several other features of cities, including economic indicators or road topology [61, 63], that could be explored in future work.

**MATERIALS AND METHODS**

**Historical urban boundary modeling.** We modeled historical urban boundaries based on the historical, generalized built-up areas (GBUA; [33]), which are based on the Historical Settlement Data Compilation for the US (HISDAC-US; [55]). Specifically, the GBUA dataset has been created by generating focal built-up density layers using gridded data on built-up areas [64] available in a grid of cells of 250 m x 250 m, in 5-year in-
FIG. 4. Scaling laws across the world. Scaling laws between (a) developed area, (b) built-up area, and (c) road length, versus population for world sub-regions. Also shown are scaling exponents and coefficients versus sub-region GDP for (d) developed area, (e) built-up area, and (f) road length distance. Data from [48, 49]. Error bars represent standard errors.

FIG. 5. Scaling coefficients over time. Rows represent the city boundaries, either (first row) DBUA or (second row) CIUPs. Columns are coefficients for models fitted to (a,e) developed area, (b,f) indoor area, (c,g) footprint area, and (d,h) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West. Error bars represent standard errors.

tervals from 1900 to 2010, using a circular focal window of 1km radius. From these built-up density layers, only areas of ≥ 5% built-up density were retained, to exclude low-density, scattered rural settlements.
remaining grid cells (delineating higher density settlements) were then segmented to obtain vector objects for each group of contiguous higher density grid cells (called “patches”). For each patch, we calculated the total number of buildings contained in it (using historical building counts from HISDAC-US i.e., the Built-Up Property Location (BUPL) layers, [65]) using a GIS-based zonal statistics tool, and ranked the patches by the number of buildings. Finally, we retained the upper 10% of these ranked patches, representing the high-density patches only. This process was done separately within each 2010 CBSA polygon [66]. Each of these patches represents a density-based urban patch (DBUP). We then dissolved these polygon objects into one multi-part polygon object per CBSA, representing the density-based urban areas (DBUAs). Lastly, the DBUPs in all CBSAs were dissolved based on adjacency only, creating natural, contiguous built-up patches i.e., potentially extending across CBSA boundaries (i.e., the CIUPs). This process was done for each point in time, resulting in three representations of urban boundaries per year: DBUAs, DBUPs and CIUPs. Importantly, we recorded the spatial relationships underlying these spatial dissolve operations (i.e., the containment relationships between CIUP and DBUA polygons, and the adjacency relationships between CIUPs, for each year). We show in Supplementary Figure S14 that DBUP sizes are highly skewed, with a small number of DBUPs representing most settlements within DBUAs. Nonetheless, the smaller settlements are important in order for DBUP and CIUP-based scaling laws to differ (and presumably create more consistent scaling law exponents) compared to DBUA-based scaling laws.

**Historical population modeling.** We acquired historical population counts at the US county level (boundaries of 2010, NHGIS) and aggregated the county populations by CBSA to obtain population time series for each 2010 CBSA boundary, for each decade from 1900 to 2010 (from the decennial censuses) and for 2015 (from the American Community Survey) [39][41]. We spatially refined these CBSA-level population counts, for each point in time. To do so, we carried out a dasymetric refinement [26] by constraining the population estimates reported for the CBSAs to the generalized built-up areas of the respective year, within each CBSA, and assigning population counts to each DBUP proportionally to the number of built-up property records [67] within that DBUP. The BUPR variable counts the number of built-up cadastral properties within each grid cell, and thus, approximates the spatial distribution of built-up structures, or cadastral parcels, or address points, which have been used for dasymetric modeling in previous work (e.g., [26][28][68][70]). Moreover, previous work has revealed strongly linear correlations between the number of houses within a spatial unit and its population over long time periods [71]; see also Supplementary Figure S1). Our historical, dasymetrically refined population estimates per DBUP represent unique, long-term depictions of urban populations at fine spatial grain, and are in strong agreement with more sophisticated population estimations, shown in Supplementary Figures S2 & S3. Based on the previously established spatial relationships (i.e., containment of DBUPs within CBSA boundaries and spatial contiguity of DBUPs) we then calculated the refined population estimates for the DBUAs and the CIUPs. The correlations between population estimates across different city boundary definitions are shown in Supplementary Figure S15.

**Historical built environment characteristics.** We used three variables to historically characterize the built environment within each urban area representation: (a) developed area (measured by the number of built-up grid cells from the HISDAC-US Built-Up Area (BUA) layers, per CBSA polygon, for each year; (b) indoor area (measured by the built-up intensity layer from HISDAC-US: Built-Up Intensity (BUI) [72], which reports the total indoor area of all buildings within a grid cell, per year); and (c) building footprint area per grid cell and year [73]. To obtain the latter, we spatially integrated cadastral data (containing construction year information) available at discrete, point-based locations from the Zillow Transaction and Assessment Dataset (ZTRAX, [74]), with vector data representing each building footprint in the US from Microsoft [75]. This integration assigns the thematic information from each ZTRAX record to the closest building footprint and thus, allowed us to merge the attribute richness of ZTRAX with the fine spatial detail of the US-BuildingFootprints dataset. Using these integrated spatial data, we were able to stratify the building footprint data by their construction year and calculated the aggregated building footprint area within each grid cell as defined by the HISDAC-US 250 m x 250 m grid. For each of the gridded surfaces measuring developed area, indoor area, and building footprint area, for each year, we calculated zonal sums in a GIS, within each DBUP of the respective year. These zonal sums were then further aggregated to the DBUAs and CIUPs based on the recorded spatial relationships.

**Historical road network modeling.** In order to model historical road network densities, we followed a method proposed in Burghardt et al. [34]. Specifically, we clipped contemporary road network vector data from the National Transportation Dataset [76] to the historical urban delineations from the generalized urban area dataset for each year, and calculated the sum of the road network segment lengths per DBUP and year. We then aggregated these road network length estimates to the DBUAs and CIUPs based on the recorded spatial relationships. We show correlations between built environment and road network related characteristics in Supplementary Figures S16–S18, which demonstrates how these city metrics relate to each other, but also diverge strongly from each other. As can be seen, there is significant scatter, although footprint area and indoor area are notably similar with a Spearman rank correlation of 0.93.

**Multi-temporal, cross-sectional scaling analy-**
sis. To find scaling laws between population estimates and the metrics describing the historical urban built environment and road network, we took the log of the property versus the estimated population within the DBUA or DBUP and computed the least-squares fit of these trends, for each point of time. The slope of this line corresponds to the exponent of the scaling law. Namely, let $Y$ be the city property (e.g., developed area), and $P$ be the population. We fitted $Y = A \times P^B$ as the linear fit of $\log(Y) = \log(A) + B \times \log(P)$. The $y$-intercept is $\log(A)$, while the exponent, $B$ describes the slope. For regional stratification of our scaling results, we used geospatial data on Census Regions in the US [77].

Uncertainty quantification, cross-comparison and validation. Data Coverage and quality. Approximately 10% of US counties have low geographic coverage, or low completeness of construction date information (i.e., temporal coverage) in HISDAC-US and the underlying ZTRAX data. These completeness issues have been extensively studied in [34, 35, 60]). We excluded CBSAs affected by these incompleteness issues from our analyses using a geospatial completeness threshold of 40% and temporal coverage of 60%, leaving 647 out of 857 CBSAs for analysis. For CIUP analysis, however, large CIUPs could span CBSAs with both high and low temporal coverage or geospatial completeness, therefore we created a weighted average estimate of the temporal coverage or geospatial completeness for each CIUP. The weight was given by the number of DBUPs within each CBSA that make up a given CIUP. We then thresholded these weighted values to have temporal coverage above 60% and geospatial completeness above 40%.

We conducted a sensitivity analysis showing that our results are largely robust to the choice of this exclusion threshold (Supplementary Figures S5 and S6), and we find that larger cities tend to have higher levels of data completeness (Supplementary Figure S19). Given the robustness of our results we do not believe this strongly affects our conclusions. Moreover, the BUPR, BUPL, BUA, and BUI layers have been validated extensively in prior work [78] and show high levels of coherence to other historical measurements of population and building characteristics. Effects of survivorship bias (i.e., the limitation of the HISDAC-US to measure urban shrinkage and historical changes in the built environment) are assumed to be of minor nature (e.g., [34, 79–81]). Historical road network models have been cross-compared to other models and datasets in [34] reporting that there is sufficient consistency between these products.

International scaling laws. To create scaling laws seen in Fig. 4 and Supplementary Figure S12, we utilize data available from [49]. This dataset consists of five features we use in these plots: world_region, world_subregion, resident_pop (population), area (developed area), built_up_area (built up area), and length_total (km of road), with definitions of these features in Boeing’s original work. This work compliments the results from [49], where he plots total street length versus population across all regions.

To create correlations between scaling coefficients and exponents and GDP per capita, we extract the latest GDP [82] and population [83] from the World Bank, which is then aggregated to regions or subregions [49]. Code for all our analysis can be found at [https://github.com/KeithBurghardt/urbanscaling](https://github.com/KeithBurghardt/urbanscaling).

ACKNOWLEDGEMENTS

K.B., J.U., K.L., and S.L. designed research; K.B. and J.U. performed analysis; K.B., J.U., K.L., and S.L. wrote the paper. The authors declare no competing interests. Partial funding for this work was provided through the National Science Foundation (awards 1924670 and 2121976) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under award number P2CHD066613. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Moreover, Safe Software, Inc., is acknowledged for providing a Feature Manipulation Engine (FME) Desktop license used for data processing.
FIG. S1. Correlations between number of building footprints and CBSA population over time. Distributions are for each CBSA with temporal completeness over 60% and geospatial completeness over 40% from 1900 and 2015 (647 CBSAs). (a) Pearson correlation, (b) Kendall’s Tau, and (c) Spearman rank correlation.
FIG. S2. Comparisons between our population estimation method and competing methods for 2010-2015 populations with CBSAs containing temporal completeness and geospatial completeness > 0%. (a) Baynes et al. 2010 DBUA population estimates, (b) Huang et al. 2015 DBUA population estimates, (c) building footprint-based 2015 DBUA population estimates, (d) Baynes et al. 2010 DBUP population estimates, (e) Huang et al. 2015 DBUP population estimates, (f) building footprint-based 2015 DBUP population estimates, (g) Baynes et al. CIUP population estimates, (h) Huang et al. 2015 CIUP population estimates, and (i) building footprint-based 2015 CIUP population estimates versus our equivalent 2015 population estimates. Building footprint-based 2015 DBUA population estimates are an alternative to our current method in which we estimate population as the proportion of building footprint areas within a DBUA, DBUP, or CIUP rather than proportion of buildings. This method is found to correlate slightly better with the alternative population estimates above (correlations are 1.0 for DBUAs and 0.96 and 0.95 for Baynes et al. and Huang et al., respectively, for DBUPs), however it will show a more trivial linear correlation with building footprint area (although it need not be strictly linear), and could similarly bias developed area estimates.
FIG. S3. Comparisons between our population estimation method and competing methods for 2010-2015 populations with CBSAs containing temporal completeness > 60% and geospatial completeness > 40%. (a) Baynes et al. [46] 2010 DBUA population estimates, (b) Huang et al. [28] 2015 DBUA population estimates, (c) building footprint-based 2015 DBUA population estimates, (d) Baynes et al. [46] 2010 DBUP population estimates, (e) Huang et al. [28] 2015 DBUP population estimates, (f) building footprint-based 2015 DBUP population estimates, (g) Baynes et al. [46] CIUP population estimates, (h) Huang et al. [28] 2015 CIUP population estimates, and (i) building footprint-based 2015 CIUP population estimates versus our equivalent 2015 population estimates. Building footprint-based 2015 DBUA population estimates are an alternative to our current method in which we estimate population as the proportion of building footprint areas within a DBUA, DBUP, or CIUP rather than proportion of buildings. This method is found to correlate slightly better with the alternative population estimates above (correlations are 1.0 for DBUAs and 0.96 and 0.95 for Baynes et al. and Huang et al., respectively, for DBUPs), however it will show a more trivial linear correlation with building footprint area (although it need not be strictly linear), and could similarly bias developed area estimates.
FIG. S4. Scaling relation between developed area and population as of 2015. Comparisons with alternative estimates of populations within build-up areas: Huang et al. [28], Baynes et al. [46] (estimated in 2010), population estimated as the fraction of CBSA building areas within a DBUP [75], and the current method. (a) DBUA, (b) DBUP, and (c) CIUP population estimates. Lines represent linear regression on the log-log plot for each population estimation method. DBUA exponents are 0.48 $\pm$ 0.02 for all methods except Baynes et al., which is 0.49 $\pm$ 0.02. DBUP exponents are 0.685 $\pm$ 0.007, 0.627 $\pm$ 0.006, 0.646 $\pm$ 0.006, and 0.718 $\pm$ 0.007 for our method, Baynes et al., Huang et al., and building area-based population estimates. CIUP exponents are 0.622 $\pm$ 0.007, 0.540 $\pm$ 0.007, 0.586 $\pm$ 0.007, and 0.664 $\pm$ 0.007 for our method, Baynes et al., Huang et al., and building area-based population estimates. These scaling results are representative of the broader agreement between scaling analyses of indoor area, building footprint area, and road length. CIUP city populations are highly correlated with DBUP city populations (see Fig. S15) and are therefore not shown for brevity.
FIG. S5. Scaling exponents over time. Plotted are exponents over time among all DBUAs with CBSA temporal completeness and spatial coverage greater than 0%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are exponents for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West. Error bars represent standard errors.
FIG. S6. Scaling exponents over time. Plotted are exponents over time among all DBUAs with CBSA temporal completeness and spatial coverage greater than 80%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are exponents for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West. CIUPs were too infrequent for sufficient statistics in the Midwest past 1950. Error bars represent standard errors.
FIG. S7. Scaling coefficients over time. Plotted are coefficients over time among all DBUAs with CBSA temporal completeness and spatial coverage greater than 0%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are coefficients for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West. Error bars represent standard errors.
FIG. S8. Scaling coefficients over time. Plotted are coefficients over time among all DBUAs with CBSA temporal completeness and spatial coverage greater than 80%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are coefficients for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West. CIUPs were too infrequent for sufficient statistics in the Midwest past 1950. Error bars represent standard errors.
FIG. S9. Quality of scaling law fit to data. Plotted are $R^2$ values for data among all DBUAs with CBSA temporal completeness and spatial coverage greater than 0%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are $R^2$ values for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West.
FIG. S10. Quality of scaling law fit to data. Plotted are $R^2$ values for data among all DBUAs with CBSA temporal completeness greater than 60% and spatial coverage greater than 40%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are $R^2$ values for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West.
FIG. S11. Quality of scaling law fit to data. Plotted are $R^2$ values for data among all DBUAs with CBSA temporal completeness and spatial coverage greater than 80%. Columns represent the city boundary: DBUA, DBUP, and CIUP. Rows are $R^2$ values for models fitted to (a-c) developed area, (d-f) indoor area, (g-i) footprint area, and (k-l) road length. Data is separately analyzed by CONUS region: Northeast, Midwest, South, and West. CIUPs were too infrequent for sufficient statistics in the Midwest past 1950.

FIG. S12. Scaling laws across the world, split by world region. Scaling laws between (a) developed area, (b) built up area, and (c) road length, versus population for world sub-regions. Data from [48, 49].
FIG. S13. Average house indoor area and number of houses sold between 1978 and 2020. We discover that single family houses (the vast majority of houses sold) show substantial increases in their average indoor area in the past 40 years. The dip in houses sold around 2008 corresponds to a recession in the US. Data from [84].

FIG. S14. Size of the largest DBUPs within CBSAs. (a) Fraction of houses in the single largest DBUP within a CBSA. (b) Fraction of houses in the top two largest DBUPs within a CBSA. (c) DBUP distribution over time.
FIG. S15. Quantile-quantile plots between (a) DBUP and DBUA data and (b) DBUP and CIUPs. We see that the largest DBUPs have populations similar to the DBUA cities. Population estimates within DBUPs and CIUPs are similar.
FIG. S16. Relationship between different DBUP city properties as of 2015. (a) Footprint area versus developed area, (b) footprint area versus indoor area, (c) footprint area versus road length, (d) road length versus developed area, (e) road length versus indoor area, (f) indoor area versus developed area. Legends are the Spearman rank correlation.
FIG. S17. Relationship between different DBUA city properties as of 2015. (a) Footprint area versus developed area, (b) footprint area versus indoor area, (c) footprint area versus road length, (d) road length versus developed area, (e) road length versus indoor area, (f) indoor area versus developed area. Legends are the Spearman rank correlation.
FIG. S18. Ratios of various DBUA city properties as of 2015. (a) Indoor area divided by footprint area, indoor area divided by road length, (c) indoor area divided by development area, (d) footprint area divided by development area, footprint area divided by road length, (f) development area divided by road length. Colors show CBSAs whose ratios are higher or lower than surrounding areas, but ratios are not expected to be near 1.0.

FIG. S19. Geographic coverage and temporal completeness of CBSAs versus DBUA population. Results in the paper are for geographic coverage greater than 40% and temporal completeness greater than 60%. We see a moderate Spearman rank correlation of 0.58 and 0.18 for the respective curves (p-value < $10^{-6}$). This implies larger cities tend to have somewhat better geographic coverage or temporal completeness, although, after cleaning, we made a regression of this relation, and find that increasing a city population by an order of magnitude increases geographic coverage by 18% and temporal coverage by only 7%, suggesting this effect, while statistically significant, does not significantly affect our results. We demonstrate the robustness of our results in Supplementary Figs. S5 & S6.
[1] D. of Economic and P. D. Social Affairs, World urbanization prospects, the 2018 revision (2019).
[2] M. R. Montgomery, The urban transformation of the developing world, Science 319, 761 (2008).
[3] G. K. Zipf, Human behavior and the principle of least effort (Addison-Wesley Press, 1949).
[4] B. Jiang, J. Yin, and Q. Liu, Zipf’s law for all the natural cities around the world, International Journal of Geographical Information Science 29, 498 (2015).
[5] J. Eeckhout, Girait’s law for (all) cities, American Economic Review 94, 1429 (2004).
[6] B. Jiang and T. Jia, Zipf’s law for all the natural cities in the united states: a geospatial perspective, International Journal of Geographical Information Science 25, 1269 (2011).
[7] H. D. Rozenfeld, D. Rybski, J. S. Andrade, M. Batty, H. E. Stanley, and H. A. Makse, Laws of population growth, Proceedings of the National Academy of Sciences 105, 18702 (2008).
[8] H. D. Rozenfeld, D. Rybski, X. Gabaix, and H. A. Makse, The area and population of cities: New insights from a different perspective on cities, American Economic Review 101, 2205–2225 (2011).
[9] X. Gabaix and Y. M. Ioannides, Chapter 53 - the evolution of city size distributions, in Cities and Geography, Handbook of Regional and Urban Economics, Vol. 4, edited by J. V. Henderson and J.-F. Thisse (Elsevier, 2004) pp. 2341–2378.
[10] M. Batty, Rank clocks, Nature 444, 592 (2006).
[11] J. Eeckhout, Girait’s law for (all) cities, American Economic Review 94, 1429 (2004).
[12] L. M. A. Bettencourt, V. C. Yang, J. Lobo, C. P. Kempes, D. Rybski, and M. J. Hamilton, The interpretation of urban scaling analysis in time, Journal of The Royal Society Interface 17, 20190846 (2020).
[13] M. Keuschmigg, Scaling trajectories of cities, Proceedings of the National Academy of Sciences 116, 13759 (2019).
[14] J. Depersin and M. Barthelemy, From global scaling to the dynamics of individual cities, Proceedings of the National Academy of Sciences 115, 2317 (2018).
[15] L. M. A. Bettencourt and D. Zünd, Demography and the emergence of universal patterns in urban systems, Nature Communications 11, 4584 (2020).
[16] C. E. Gehrke and K. Bielli, Certain effects of grouping upon the size of the correlation coefficient in census tract material, Journal of the American Statistical Association 29, 169 (1934).
[17] E. Arcaute, E. Hatna, P. Ferguson, H. Youn, A. Johansson, and M. Batty, Constructing cities, deconstructing scaling laws, Journal of The Royal Society Interface 12, 20140745 (2015).
[18] C. Bartesaghi-Koc, P. Osmond, and A. Peters, Innovative use of spatial regression models to predict the effects of green infrastructure on land surface temperatures, Energy and Buildings 254, 111564 (2022).
[19] J. A. Maantay, A. R. Maroko, and C. Herrmann, Mapping population distribution in the urban environment: The cadastral-based expert dasymetric system (ceds), Cartography and Geographic Information Science 34, 77 (2007).
[20] C. Wei, H. Taubenböck, and T. Blaschke, Measuring urban agglomeration using a city-scale dasymetric population map: A study in the pearl river delta, china, Habitat International 59, 32 (2017).
[21] E. L. Glaeser and M. E. Kahn, Sprawl and urban growth, in Handbook of regional and urban economics, Vol. 4 (Elsevier, 2004) pp. 2481–2527.
[22] J. J. Lawler, D. J. Lewis, E. Nelson, A. J. Plantinga, S. Polasky, J. C. Withey, D. P. Helmers, S. Martinuzzi, D. Pennington, and V. C. Radloff, Projected land-use change impacts on ecosystem services in the united states, Proceedings of the National Academy of Sciences 111, 7492 (2014).
[23] G. S. Cumming, A. Buerkert, E. M. Hoffmann, E. Schlech, S. von Cramon-Taubadel, and T. Tscharntke, Implications of agricultural transitions and urbanization for ecosystem services, Nature 515, 50 (2014).
[34] K. Burghardt, J. H. Uhl, K. Lerman, and S. Leyk, Road network evolution in the urban and rural united states since 1900, Computers, Environment and Urban Systems 95, 101803 (2022).

[35] S. Leyk and J. H. Uhl, Hisdac-us, historical settlement data compilation for the conterminous united states over 200 years, Scientific data 5, 180175 (2018).

[36] X. Li, H. Tian, S. Pan, and C. Lu, Four-century history of land transformation by humans in the united states: 1630-2020, Earth System Science Data Discussions 2022, 1 (2022).

[37] H. El Gouj, C. Rincón-Acosta, and C. Lagesse, Urban morphogenesis analysis based on geohistorical road data, Applied Network Science 7, 6 (2022).

[38] S. Leyk, D. Balk, B. Jones, M. R. Montgomery, and H. Engin, The heterogeneity and change in the urban structure of metropolitan areas in the united states, 1990–2010, Scientific Data 6, 321 (2019).

[39] U. C. Bureau, County population totals: 2010-2019, https://www2.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html (2020).

[40] U. C. Bureau, cenpop2000, https://www2.census.gov/geo/docs/reference/cenpop2000/county/ (2020).

[41] U. C. Bureau, Census-us decennial county population data, 1900-1990, https://www.nber.org/research/data/census-us-decennial-county-population-data-1900-1990 (2020).

[42] B. Jiang and Y. Miao, The evolution of natural cities from the perspective of location-based social media, The Professional Geographer 67, 295 (2015).

[43] J. B. Parr, Spatial definitions of the city: four perspectives, Urban studies 44, 381 (2007).

[44] C. Small, F. Pozzi, and C. Elvidge, Spatial analysis of global urban extent from dmsp-ols night lights, Remote Sensing of Environment 96, 277 (2005).

[45] M. C. González, C. A. Hidalgo, and A.-L. Barabási, Understanding individual human mobility patterns, Nature 453, 779 (2008).

[46] J. Baynes, A. Neale, and T. Hultgren, Improving intelligent dasymetric mapping population density estimates at 30 m resolution for the conterminous united states by excluding uninhabited areas, Earth System Science Data 14, 2833 (2022).

[47] S. Openshaw and P. J. Taylor, A million or so correlation coefficients: Three experiments on the modifiable areal unit problem, in Statistical Applications in the Spatial Sciences, edited by N. Wrigley (Pion, London, 1979) pp. 127–144.

[48] S. G. Ortman, J. Lobo, and M. E. Smith, Cities: Complexity, theory and history, PLOS ONE 15, 1 (2020).

[49] G. Boeing, Street network models and indicators for every urban area in the world, Geographical Analysis 54, 519 (2022), https://onlinelibrary.wiley.com/doi/pdf/10.1111/gean.12281

[50] D. Balk, S. Leyk, B. Jones, M. R. Montgomery, and A. Clark, Understanding urbanization: A study of census and satellite-derived urban classes in the united states, 1990-2010, PLOS ONE 13, 1 (2018).

[51] M. Moffatt, What caused the post-war economic housing boom after wwi?, ThoughtCo (2021).

[52] W. H. Frey, Central city white flight: Racial and nonracial causes, American Sociological Review 44, 425 (1979).

[53] S. Dark and D. Bram, The modifiable areal unit problem (maup) in physical geography, Progress in Physical Geography: Earth and Environment 31, 471 (2007).

[54] E. H. Simpson, The interpretation of interaction in contingency tables, Journal of the Royal Statistical Society. Series B (Methodological) 13, 238 (1951).

[55] S. Angel, J. Parent, D. L. Civco, and A. Blei, The persistent decline in urban densities: Global and historical evidence of 'sprawl' (2017).

[56] D. P. Bigelow, D. J. Lewis, and C. Mihia, A major shift in u.s. land development avoids significant losses in forest and agricultural land, Environmental Research Letters 17, 024007 (2022).

[57] A. W. Schäfer, Long-term trends in domestic us passenger travel: the past 110 years and the next 90, Transportation 44, 293 (2017).

[58] J. Lobo, T. Whitelaw, L. M. A. Bettencourt, P. Wiessner, M. E. Smith, and S. Ortman, Scaling of hunter-gatherer camp size and human sociality, Current Anthropology 63, 68 (2022).

[59] M. E. Smith, S. G. Ortman, J. Lobo, C. E. Ebert, A. E. Thompson, K. M. Pruer, R. Liendo Stuardo, and R. M. Rosenswig, The low-density urban systems of the classic period maya and izapa: Insights from settlement scaling theory, Latin American Antiquity 32, 120-137 (2021).

[60] G. Boeing, A multi-scale analysis of 27,000 urban street networks: Every us city, town, urbanized area, and zillow neighborhood, Environment and Planning B: Urban Analytics and City Science 47, 590 (2020).

[61] J. Xue, N. Jiang, S. Liang, Q. Pang, T. Yabe, S. V. Ukkusuri, and J. Ma, Quantifying the spatial homogeneity of urban road networks via graph neural networks, Nature Machine Intelligence 4, 246 (2022).

[62] B. C. Scheer, The anatomy of sprawl, Places 14 (2001).

[63] B. C. Geiger, G. U. Jang, J. C. Joo, and J. Park, Capturing the signature of topological evolution from the snapshots of road networks, Complexity 2020, 8054316 (2020).

[64] J. H. Uhl and S. Leyk, Historical built-up areas (BUA) - gridded surfaces for the U.S. from 1810 to 2015 (2020).

[65] J. H. Uhl and S. Leyk, Historical built-up property locations (BUPL) - gridded surfaces for the U.S. from 1810 to 2015 (2020).

[66] U. C. Bureau, t1_2015_us_cbsa, https://www2.census.gov/geo/tiger/TIGER2015/CBSA/ (2015).

[67] J. H. Uhl and S. Leyk, Historical built-up property records (bupr) - gridded surfaces for the u.s. from 1810 to 2015 (2020).

[68] P. A. Zandbergen, Dasymetric mapping using high resolution address point datasets, Transactions in GIS 15, 5 (2011).

[69] A. F. Tapp, Areal interpolation and dasymetric mapping methods using local ancillary data sources, Cartography and Geographic Information Science 37, 215 (2010).

[70] C. Yin, Y. Shi, H. Wang, and J. Wu, Disaggregation of an urban population with m_idw interpolation and building information, Journal of Urban Planning and Development 141, 04014012 (2015).

[71] S. Leyk, J. H. Uhl, D. S. Connor, A. E. Braswell, N. Mietkiewicz, J. K. Balch, and M. Gutmann, Two centuries of settlement and urban development in the united states, Science Advances 6, 10.1126/sciadv.aba2937 (2020)
datasets measuring 200 years of land development in the United States, Earth System Science Data 13, 119 (2021).

[79] G. Boeing, Off the grid...and back again? The recent evolution of American street network planning and design, Journal of the American Planning Association 87, 1 (2020).

[80] C. Barrington-Leigh and A. Millard-Ball, A century of sprawl in the United States, Proceedings of the National Academy of Sciences 112, 8244 (2015).

[81] J. R. Meijer, M. A. J. Huibregts, K. C. G. J. Schotten, and A. M. Schipper, Global patterns of current and future road infrastructure, Environmental Research Letters 13, 064006 (2018).

[82] World Bank, GDP (current US$), https://data.worldbank.org/indicator/NY.GDP.MKTP.CD (2022).

[83] World Bank, Population, total, https://data.worldbank.org/indicator/SP.POP.TOTL (2022).

[84] Census Bureau, Characteristics of new housing, https://www.census.gov/construction/chars/ (2021).