Supplementary Information for: Universal Patterns in the Long-term Growth of Urban Infrastructure in U.S. Cities from 1900 to 2015

1. Generalized city growth scaling

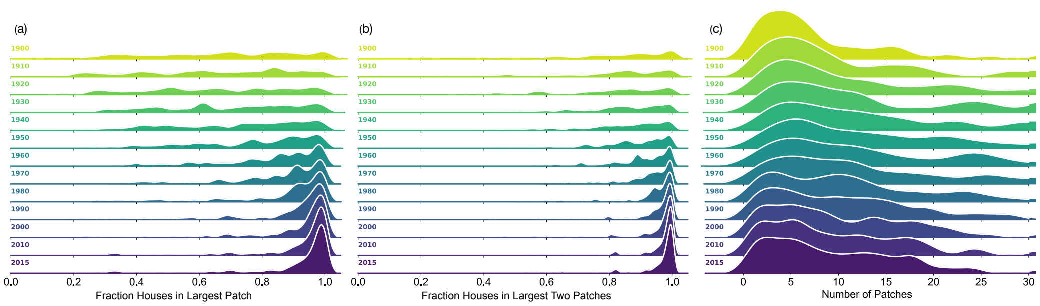

Figure S1: Size of the largest patches. (a) Fraction of houses in the single largest patch within a CBSA. (b) Fraction of houses in the top two largest patches within a CBSA. (c) Patch distribution over time.

City growth need not be strictly looked at from the lens of city scaling, however. Not only are there a number of ways to define scaling laws (6), we cannot be certain that cities strictly follow temporal scaling laws. We therefore want to check the robustness of these correlations without scaling assumptions. To this end, we analyze the correlations between the growth patterns of any pair of cities. We specifically plot the change in various statistics over time, such as new developed area each year and calculate the Spearman correlation between all cities within a small distance window. As

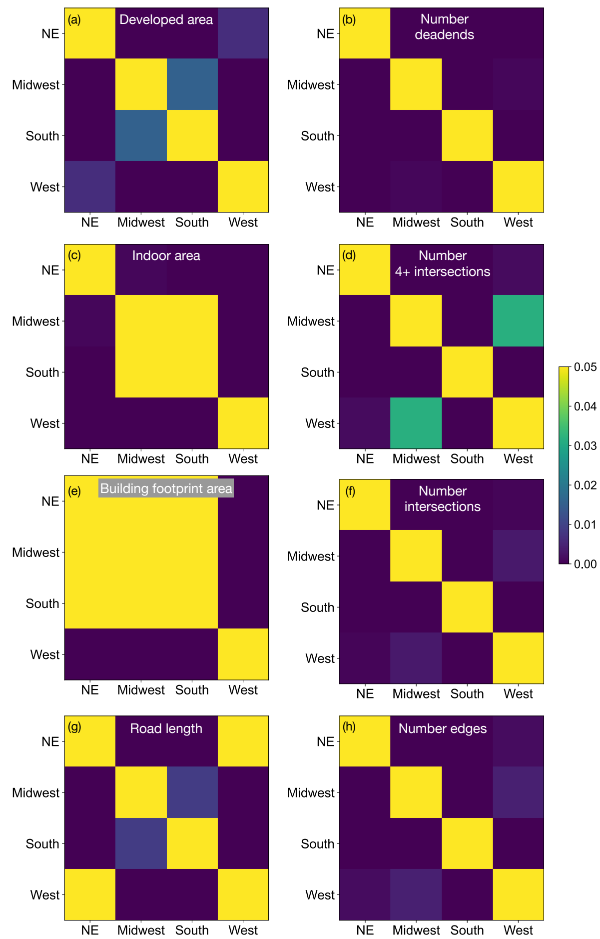

Figure S2: Statistical significance of temporal scaling law distributions within each city size and region for temporal completeness greater than 60% and spatial coverage greater than 40% (see main text Fig. 4 for the distributions). Colors correspond to insignificant differences (dark) and statistically significant differences (light) in these statistics, based on Dunn's test (Dunn (1961)) after rejection by the Kruskal-Wallis test (p-value < 0.05) (Kruskal and Wallis (1952)). Statistically significant differences for (a) developed area, (b) number of deadends, (c) indoor area, (d) number of intersections where four or more edges meet, (e) building footprint area, (f) total number of intersections, (g) road length, and (h) number of edges.

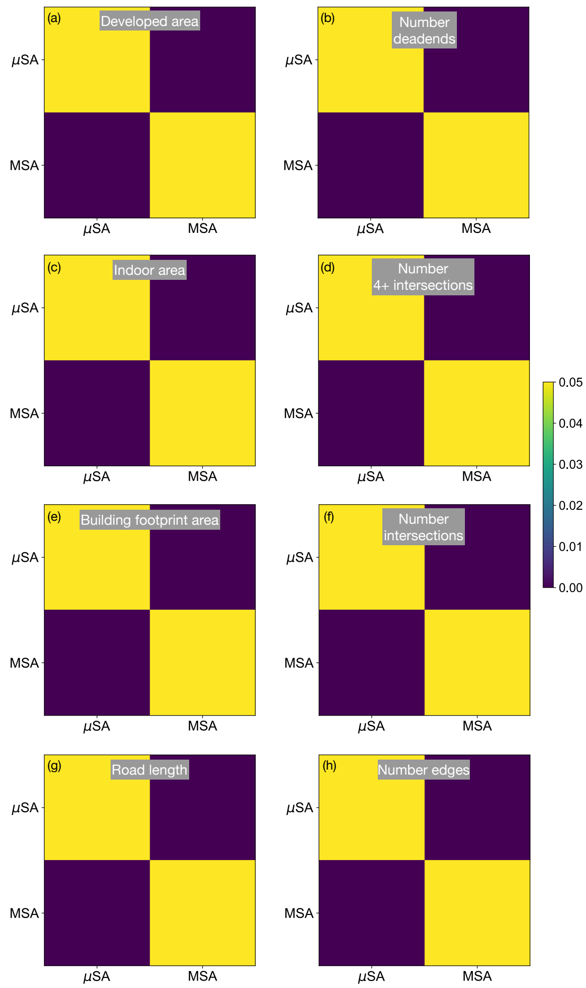

Figure S3: Statistical significance of temporal scaling law distributions within each city size and region for temporal completeness greater than 60% and spatial coverage greater than 40% (see main text Fig. 4 for the distributions). Colors correspond to insignificant differences (dark) and statistically significant differences (light) in these statistics, based on Dunn's test (Dunn (1961)) after rejection by the Kruskal-Wallis test (p-value < 0.05) (Kruskal and Wallis (1952)). Statistically significant differences between micro- and metropolitan statistical areas scaling law exponents for (a) developed area, (b) number of deadends, (c) indoor area, (d) number of intersections where four or more edges meet, (e) building footprint area, (f) total number of intersections, (g) road length, and (h) number of edges.

before, we vary the window to plot how correlations decrease with distance between cities. The statistics we calculate are developed area, indoor area, building footprint area, road length (just as when calculating scaling laws), as well as road network statistics outlined in Table **??**. We calculate the local griddedness (defined in (3), road density (length of road per square kilometer), orientation entropy (variability in road orientations, defined in (2), mean degree (number of edges at an intersection), dead end percentage (the proportion of cul-de-sacs and other roads that do not end at an intersection), and percentage degree 4+ (the proportion of intersections where four or more edges meet). Descriptions of statistics not mentioned in the main text are shown in SI Table S1.

The results are shown in Fig. S4. In Fig. S4a, we see changes in extrinsic variables, namely developed area, building footprint area, indoor area, and road length, all have similarly long-range correlations with city distance. They only reach approximately zero correlation after 1000km, although the decrease in correlation appears faster than scaling law exponents. Figure S4b, meanwhile, shows that changes in intrinsic variables (road network statistics) show a qualitative decrease in correlation with distance, although the correlations are always high and never reach zero. This is probably because all cities, regardless of population, experienced similar road network changes with the advent of cars, as described in prior work (3). All results are found to be robust when splitting by MSAs and μSAs (SI Fig. S12), as well as weaker and stronger data filtering (SI Figs. S5 & S6, respectively). These results broadly match the findings seen for scaling laws, suggesting the growth and evolution of cities has a long-range correlation.

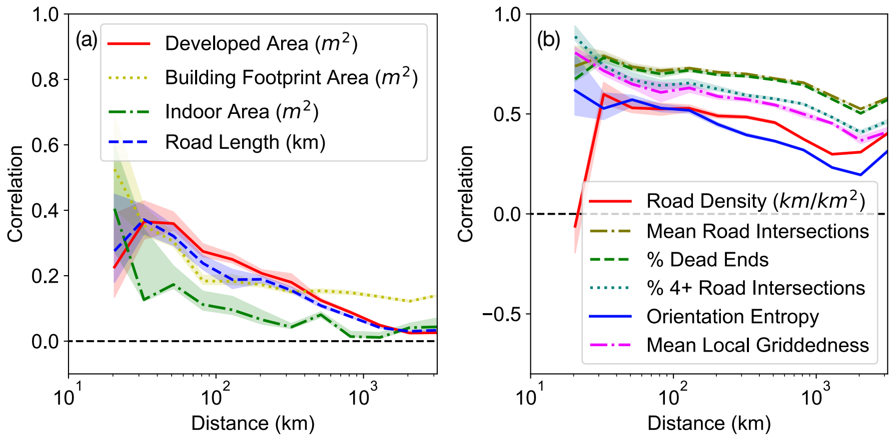

Figure S4: Nearby cities grow similarly. (a) Correlations of extrinsic city properties: developed area growth, building footprint area growth, indoor area growth, and road length growth evolution versus city distance as of 2015. road networks growth between cities (CBSAs). (b) Correlations of road network properties: road density, mean number of intersections, fraction of dead ends, orientation entropy (1), and mean local griddedness (3) of new roads.

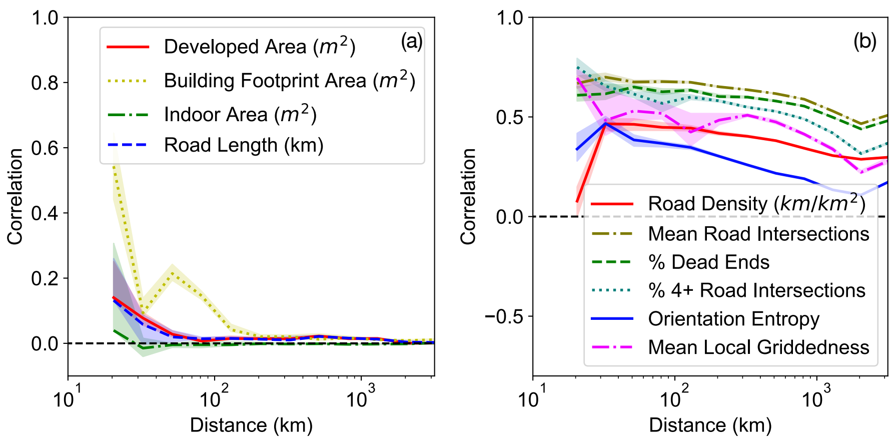

Figure S5: Nearby cities grow similarly. Figures are the same as main text Fig. 6 except we include all CBSAs with temporal completeness and geospatial coverage greater than 0%. (a) Correlations of extrinsic city properties: developed area growth, building footprint area growth, indoor area growth, and road length growth evolution versus city distance as of 2015. road networks growth between cities (CBSAs). (b) Correlations of road network properties: road density, mean number of intersections, fraction of dead ends, orientation entropy (1), and mean local griddedness (3 of new roads.

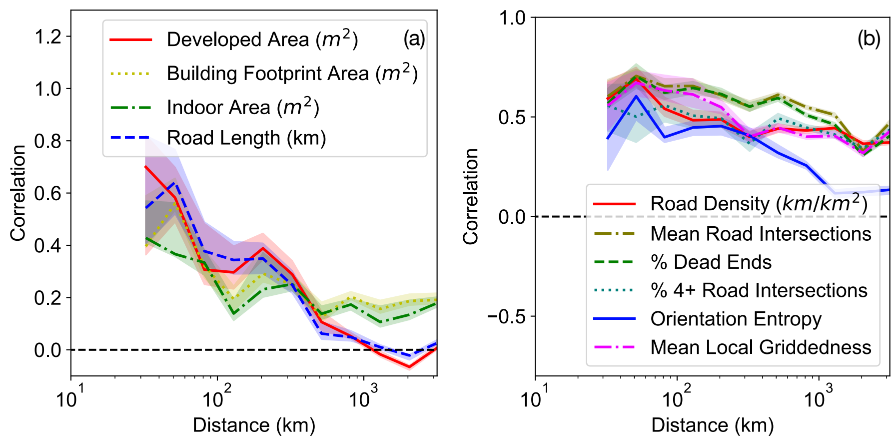

Figure S6: Nearby cities grow similarly. Figures are the same as main text Fig. 6 except we include all CBSAs with temporal completeness and geospatial coverage greater than 80%. (a) Correlations of extrinsic city properties: developed area growth, building footprint area growth, indoor area growth, and road length growth evolution versus city distance as of 2015. road networks growth between cities (CBSAs). (b) Correlations of road network properties: road density, mean number of intersections, fraction of dead ends, orientation entropy (1), and mean local griddedness (3 of new roads.

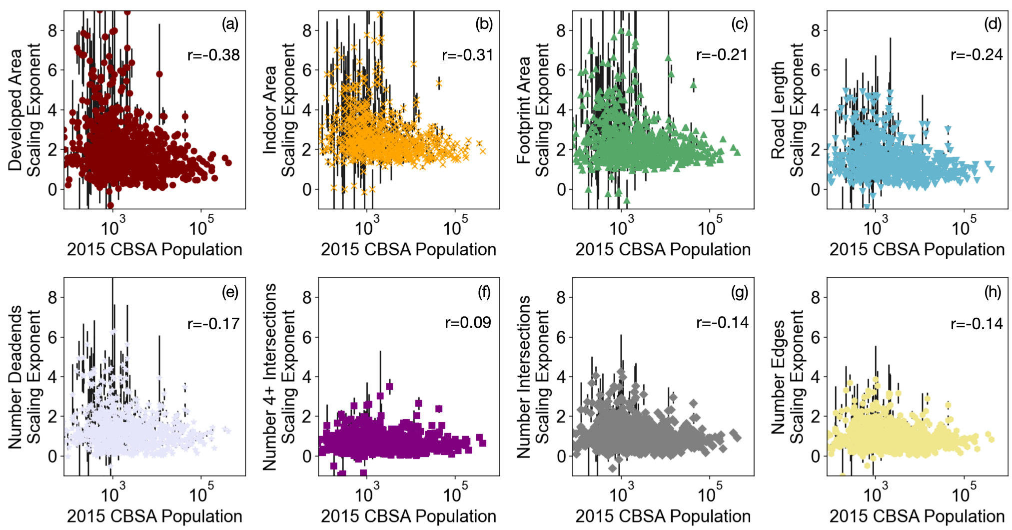

Figure S7: Scaling law exponent versus 2015 population for all CBSAs with temporal completeness and geospatial coverage > 0%. (a) Developed area, (b) indoor area, (c) building footprint area, (d) road length, (e) number of deadends, (f) number of intersections with four or more edges meeting, (g) total number of intersections, and (h) number of edges. Black error bars represent standard errors of scaling law coefficients. Statistically significant Spearman correlations (p-values< 0.05) are shown in each figure, and are otherwise denoted "n.s." (not significant).

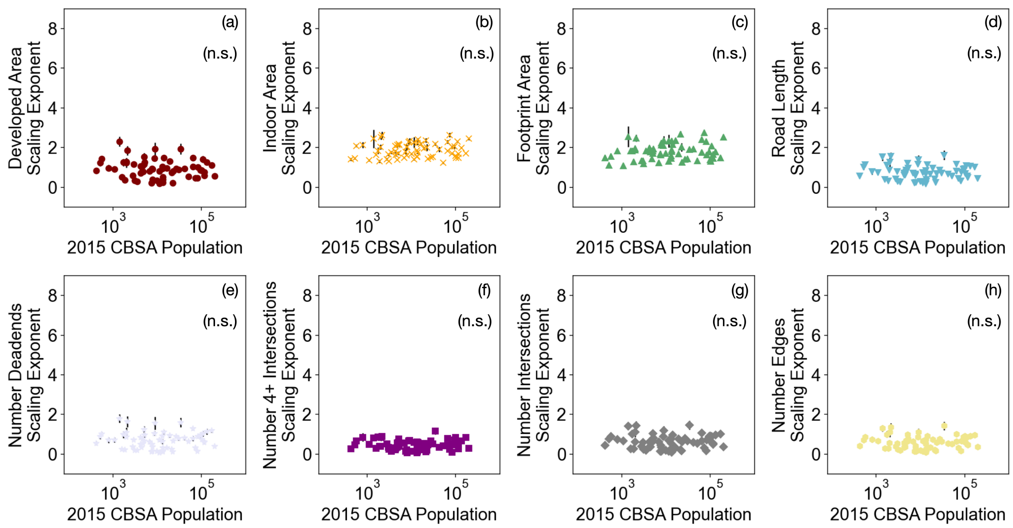

Figure S8: Scaling law exponent versus 2015 population for all CBSAs with temporal completeness and geospatial coverage > 80%. (a) Developed area, (b) indoor area, (c) building footprint area, (d) road length, (e) number of deadends, (f) number of intersections with four or more edges meeting, (g) total number of intersections, and (h) number of edges. Black error bars represent standard errors of scaling law coefficients. Statistically significant Spearman correlations (p-values< 0.05) are shown in each figure, and are otherwise denoted “n.s.” (not significant).

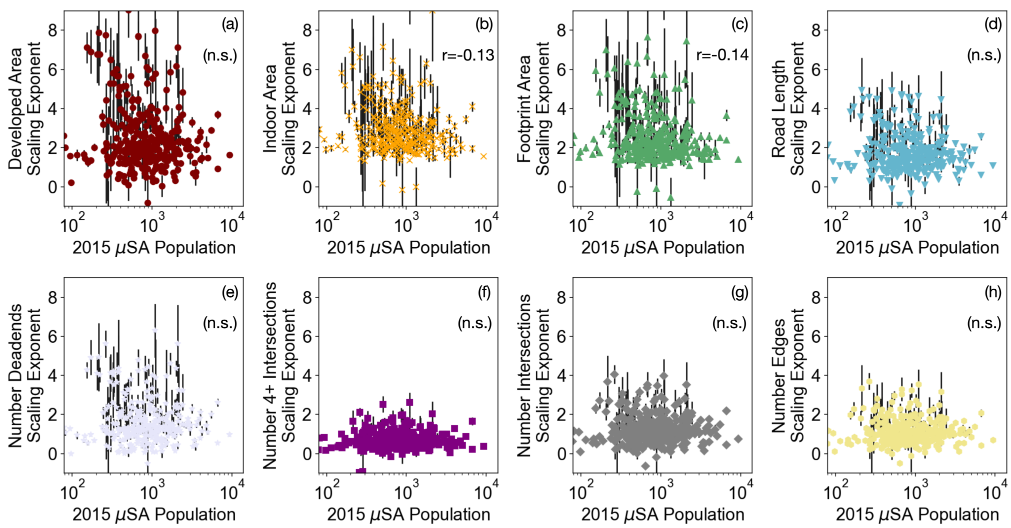

Figure S9: Scaling law exponent versus 2015 population for μSAs with temporal completeness $> 60\%$ and geospatial coverage $> 40\%$ as in the main text. (a) Developed area, (b) indoor area, (c) building footprint area, (d) road length, (e) number of deadends, (f) number of intersections with four or more edges meeting, (g) total number of intersections, and (h) number of edges. Black error bars represent standard errors of scaling law coefficients. Statistically significant Spearman correlations (p-values< 0.05) are shown in each figure, and are otherwise denoted "n.s." (not significant).

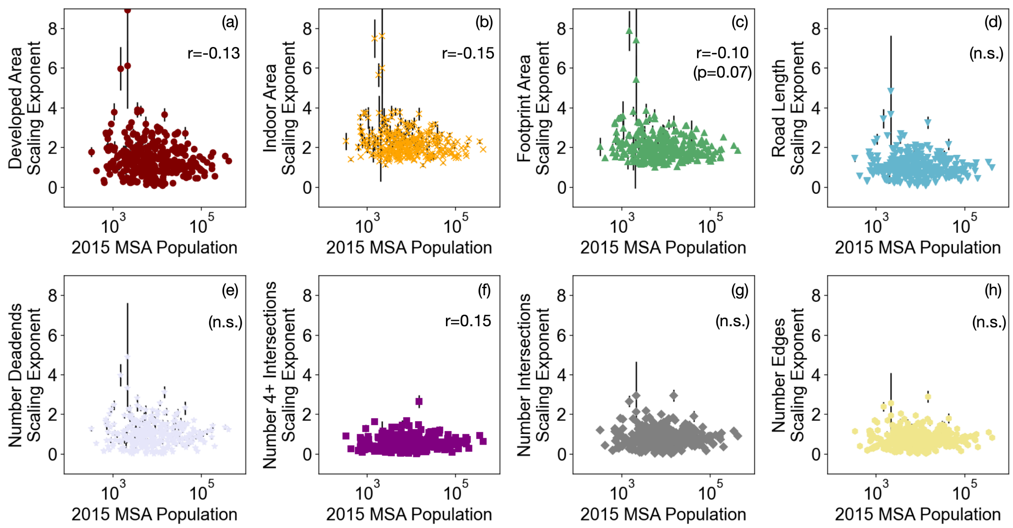

Figure S10: Scaling law exponent versus 2015 population for MSAs with temporal completeness > 60% and geospatial coverage > 40% as in the main text. (a) Developed area, (b) indoor area, (c) building footprint area, (d) road length, (e) number of deadends, (f) number of intersections with four or more edges meeting, (g) total number of intersections, and (h) number of edges. Black error bars represent standard errors of scaling law coefficients. Statistically significant Spearman correlations (p-values< 0.05) are shown in each figure, and are otherwise denoted "n.s." (not significant).

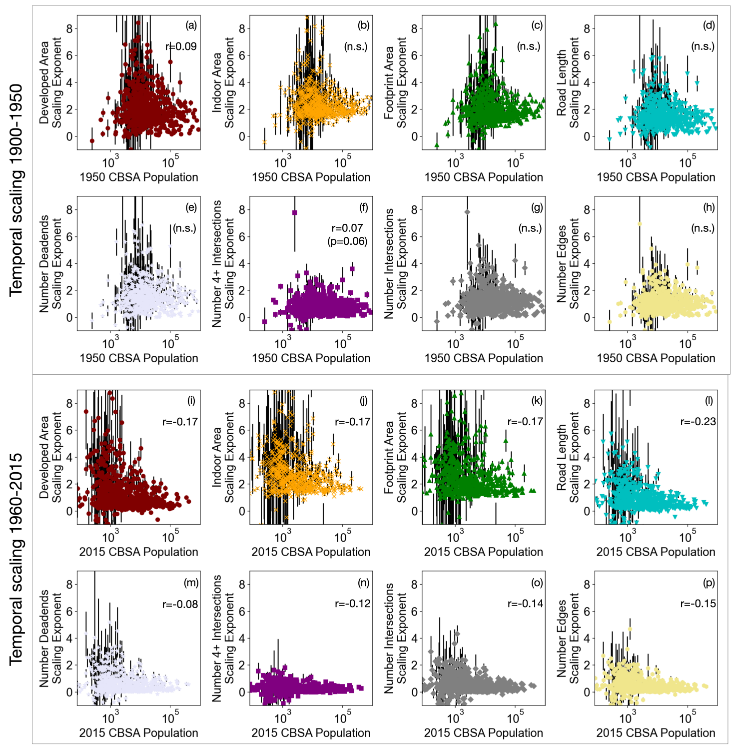

Figure S11: Relationship between scaling exponents and population in 1950 and 2015. (Top panel) scaling law exponent for temporal scaling 1900-1950 versus 1950 population and (bottom panel) scaling law exponent for temporal scaling 1960-2015 versus 2015 population for MSAs with temporal completeness > 60% and geospatial coverage > 40% as in the main text. Top panel: (a,i) Developed area, (b,j) indoor area, (c,k) building footprint area, (d,l) road length, (e,m) number of deadends, (f,n) number of intersections with four or more edges meeting, (g,o) total number of intersections, and (h,p) number of edges. Black error bars represent standard errors of scaling law coefficients. Statistically significant Spearman correlations (p-values< 0.05) are shown in each figure, and are otherwise denoted "n.s." (not significant).

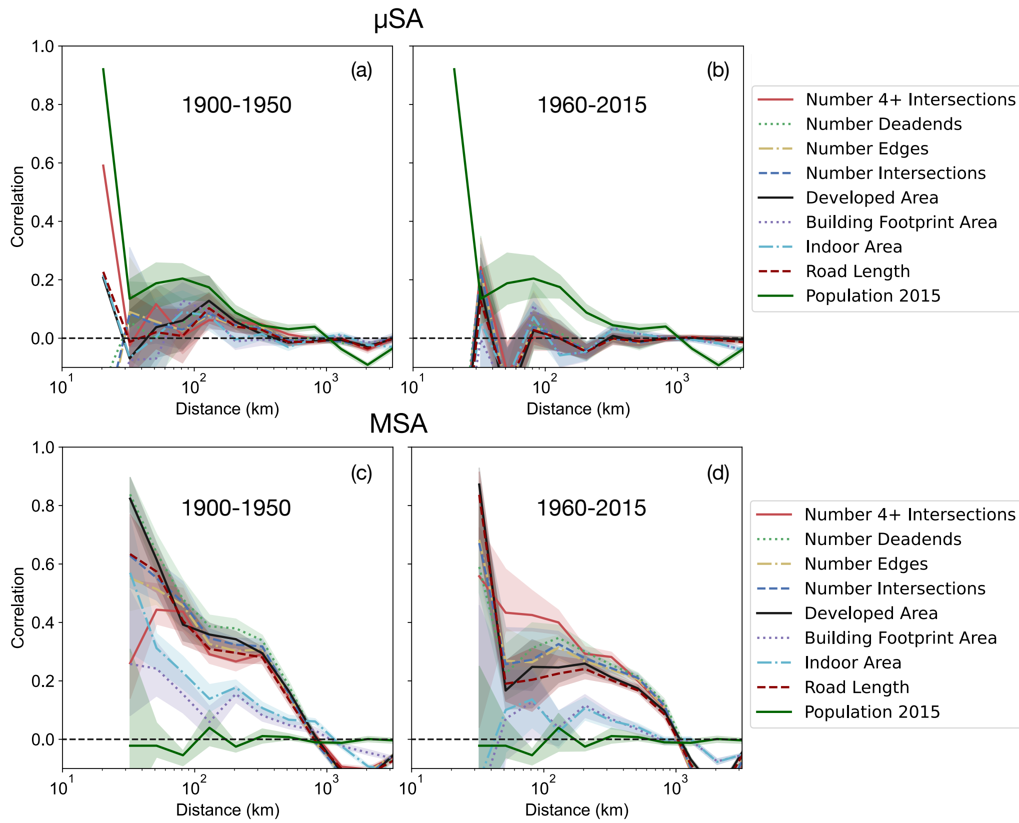

Figure S12: Nearby cities scale similarly. Figures are the same as main text Fig. 6 except we split data by (a–b) MSA and (c–d) μSA with borders defined as of 2010 and we divide the temporal scaling laws to be between (a,c) 1900–1950 and (b,d) 1960–2015. Spearman correlation versus distance for the number of 4+ intersections, number of deadends, number of edges, number of intersections, developed area, building footprint area, indoor area, and road length. Shaded regions represent 68% confidence intervals in the mean correlation.

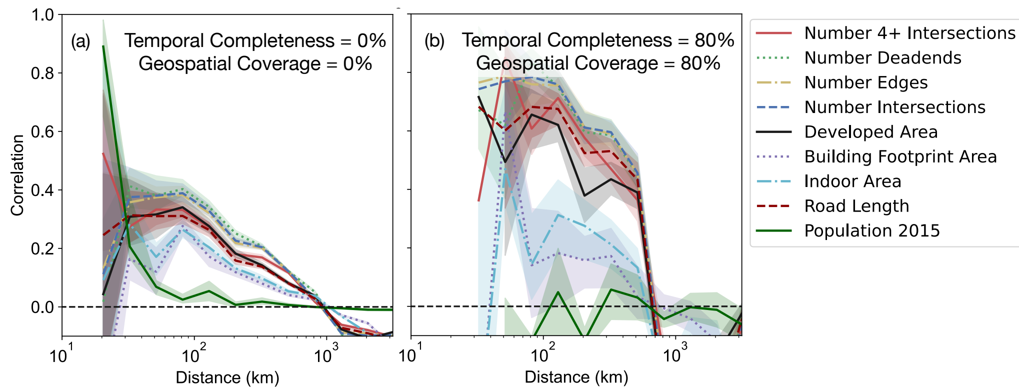

Figure S13: Nearby cities scale similarly. Figures are the same as main text Fig. 6 except CBSAs have temporal completeness and spatial coverage greater than (a) 0% and (b) 80%. Spearman correlation versus distance for the number of 4+ intersections, number of deadends, number of edges, number of intersections, developed area, building footprint area, indoor area, and road length. Shaded regions represent 68% confidence intervals in the mean correlation.

References

Boeing, G., 2019. Urban spatial order: street network orientation, configuration, and entropy. Applied Network Science 4, 67. URL: https://doi.org/10.1007/s41109-019-0189-1, doi:10.1007/s41109-019-0189-1.

Boeing, G., 2020. Off the grid...and back again? the recent evolution of american street network planning and design. Journal of the American Planning Association 87, 1–15.

Burghardt, K., Uhl, J.H., Lerman, K., Leyk, S., 2022. Road network evolution in the urban and rural united states since 1900. Computers, Environment and Urban Systems 95, 101803. URL: https://

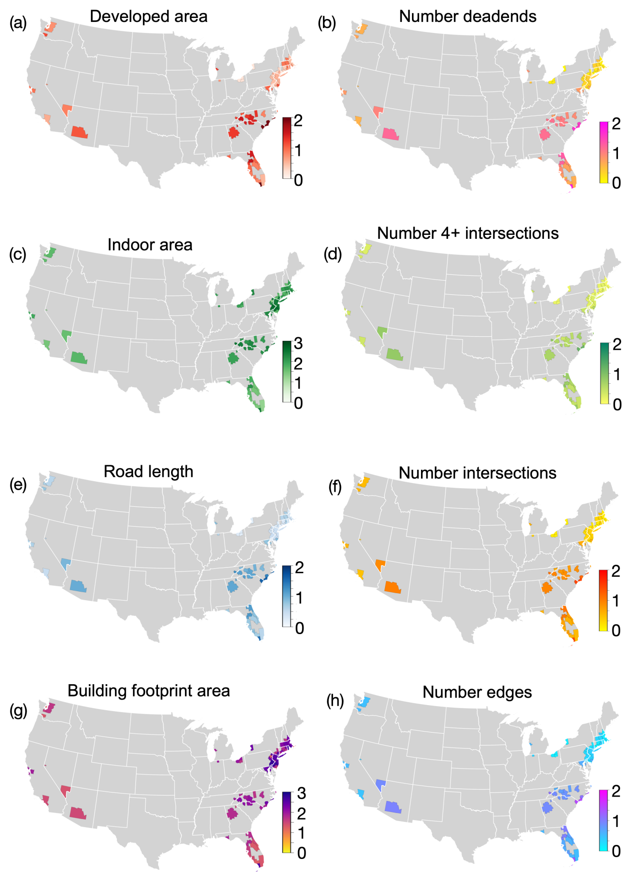

Figure S14: Scaling law exponents across the US for CBSAs with temporal completeness and geospatial coverage > 80%. Map of temporal scaling exponents for (a) developed area, (b) number of deadends, (c) building footprint, (d) number of 4+ intersections, (e) indoor area, (f) total number of intersections, (g) road length, and (h) number of edges.

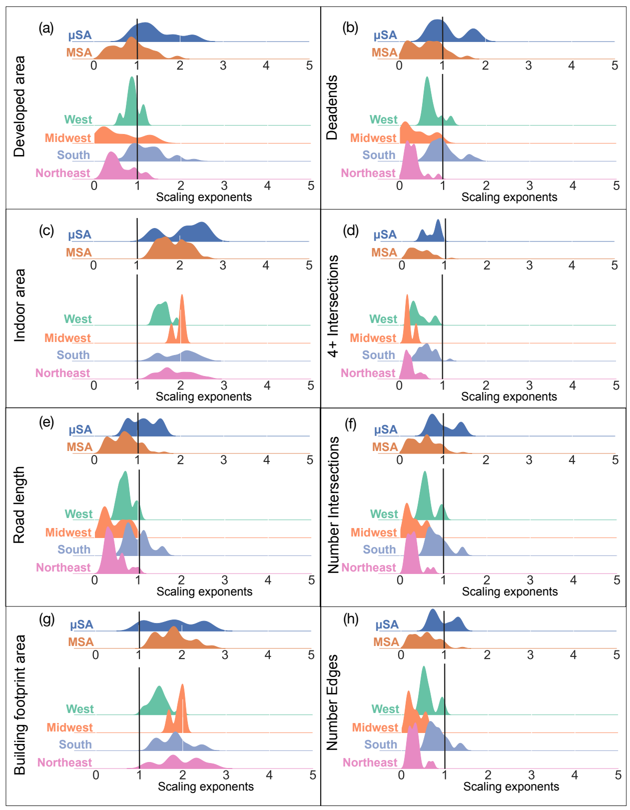

Figure S15: Scaling law exponents split by city size and region for CBSAs with temporal completeness and geospatial coverage > 80%. Distribution of temporal scaling exponents for (a) developed area, (b) number of deadends, (c) building footprint, (d) number of 4+ intersections, (e) indoor area, (f) total number of intersections, (g) road length, and (h) number of edges. Distributions are split by major (MSA) and minor (μSA) cities as well as OMB-defined regions. Black vertical lines represent linear scaling.

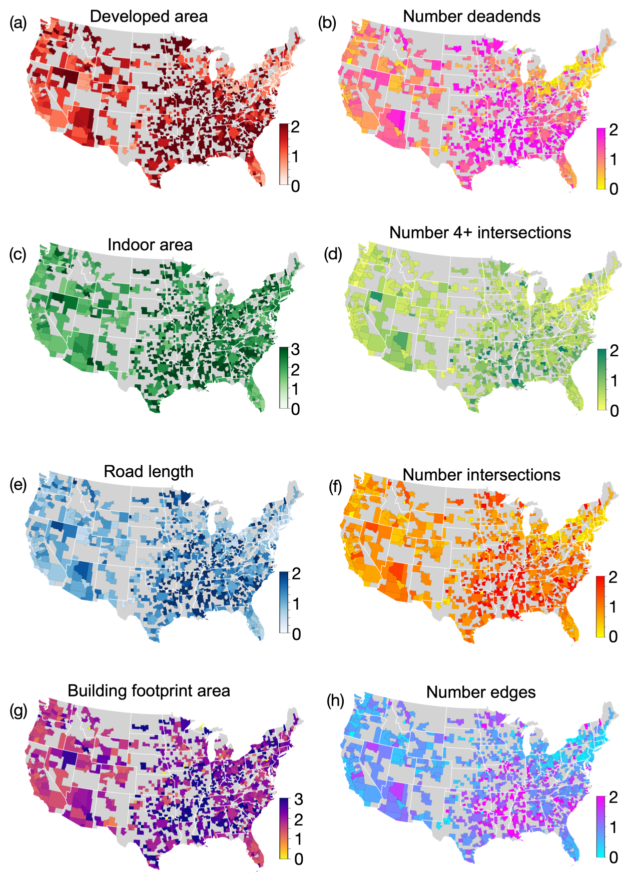

Figure S16: Scaling law exponents across the US for CBSAs with temporal completeness and geospatial coverage > 0%. Map of temporal scaling exponents for (a) developed area, (b) number of deadends, (c) building footprint, (d) number of 4+ intersections, (e) indoor area, (f) total number of intersections, (g) road length, and (h) number of edges.

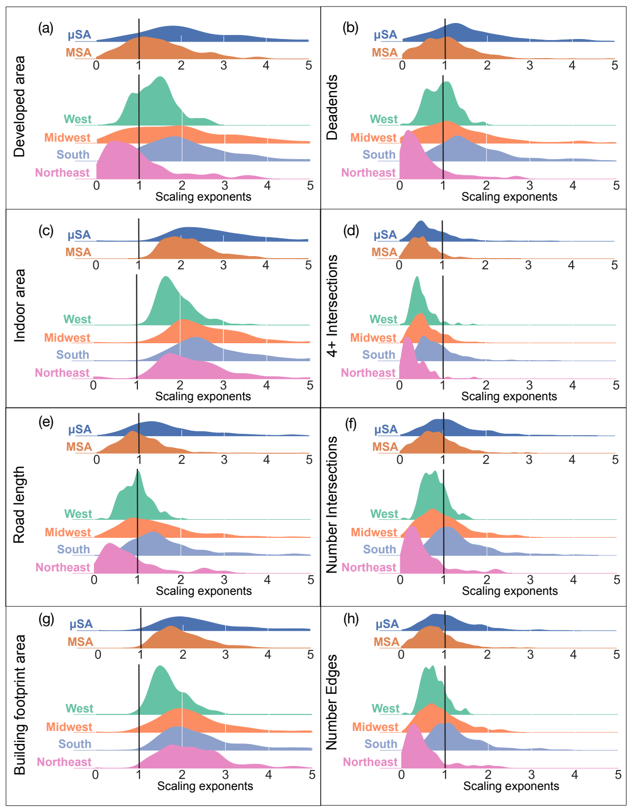

Figure S17: Scaling laws split by city size and region for CBSAs with temporal completeness and geospatial coverage > 0%. Distribution of temporal scaling laws for (a) developed area, (b) number of deadends, (c) building footprint, (d) number of 4+ intersections, (e) indoor area, (f) total number of intersections, (g) road length, and (h) number of edges. Distributions are split by major (MSA) and minor (μSA) cities as well as OMB-defined regions. Black vertical lines represent linear scaling.

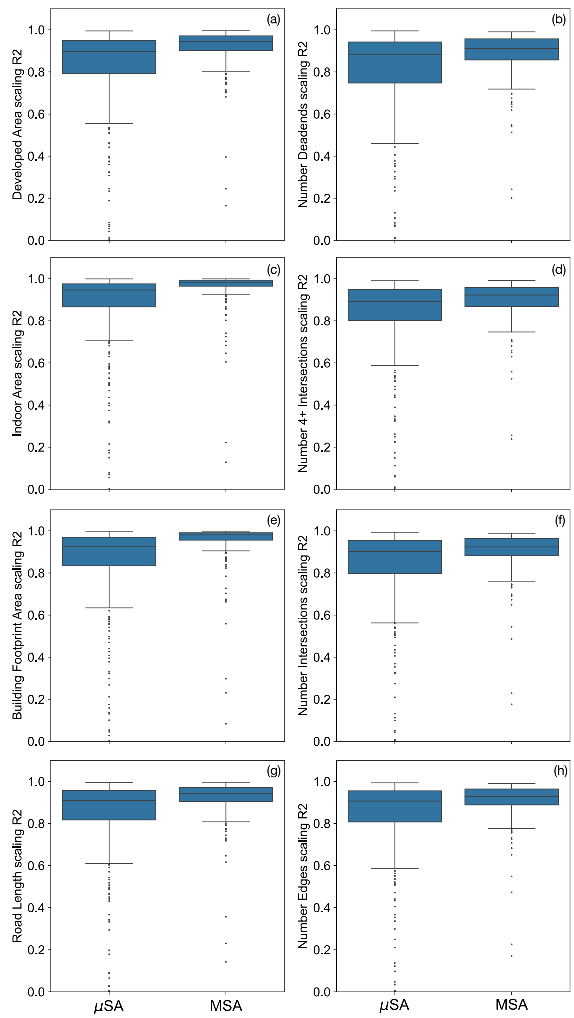

Figure S18: Agreement of temporal relations with power law for cities split by CBSA type (MSA or μSA) whose temporal completeness is greater than 60% and spatial coverage is greater than 40%, as in the main text. Powerlaw fit R^2 for (a) developed area, (b) building footprint, (c) indoor area, (d) road length, (e) number of deadends, (f) number of intersections where four or more edges meet, (g) total number of intersections, and (h) number of edges. Values closer to one correspond to better agreement with power law scaling relation. We find both MSAs and μSAs have strong fits to the scaling relation, but MSA fits are typically better (Mann-Whitney U test p-value < 0.001 for each statistic).

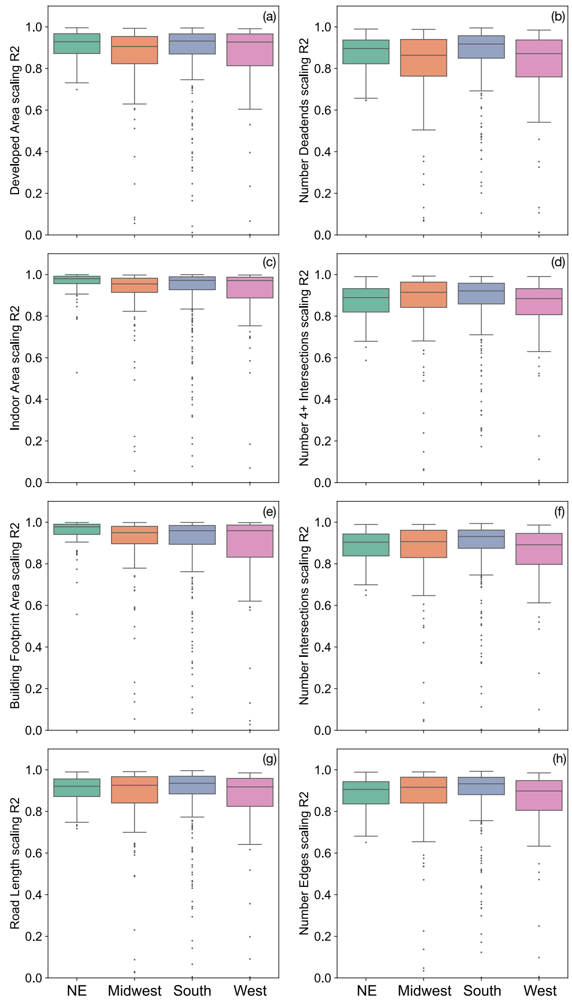

Figure S19: Agreement of temporal relations with power law for cities split by OMB-defined regions whose temporal completeness is greater than 60% and spatial coverage is greater than 40%, as in the main text. Powerlaw fit R^2 for (a) developed area, (b) building footprint, (c) indoor area, (d) road length, (e) number of deadends, (f) number of intersections where four or more edges meet, (g) total number of intersections, and (h) number of edges. Values closer to one correspond to better agreement with power law scaling relation. We find strong fits to the scaling relation across regions.

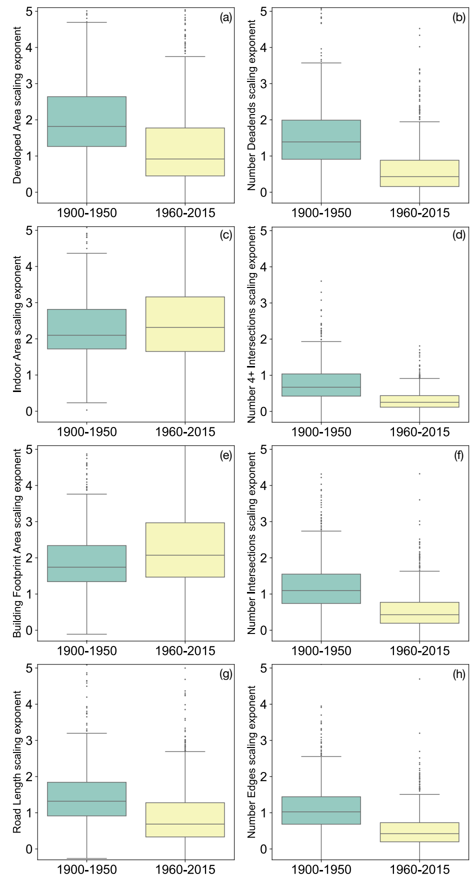

Figure S20: Change in scaling exponents from 1900–1950 and from 1960–2015. Scaling exponents for (a) developed area, (b) building footprint area, (c) indoor area, (d) road length, (e) number of deadends, (f) number of intersections where four or more edges meet, (g) total number of intersections, and (h) number of edges. In agreement with main text Fig. 4, we see that scaling laws are higher for MSAs (Mann-Whitney U test p-values < 0.001 except for indoor area scaling exponents, p-value$= 0.2$). In this analysis, we filter cities to have geographic coverage greater than 40% and temporal completeness greater than 60% as in the main text.

www.sciencedirect.com/science/article/pii/S0198971522000473, doi:10.1016/j.compenvurbsys.2022.101803.

Dunn, O.J., 1961. Multiple comparisons among means. Journal of the American Statistical Association 56, 52–64.

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association 47, 583–621.

Lemoy, R., Caruso, G., 2021. Radial analysis and scaling of urban land use. Scientific Reports 11, 22044. URL: https://doi.org/10.1038/s41598-021-01477-y, doi:10.1038/s41598-021-01477-y.

Road network metric	**Unit**	**Description**
Local griddedness	Node	Number of quadrilaterals touching a node divided by its degree (3)
Road density	Edge	km road per km built-up area
Orientation entropy	Edge	Entropy of edge orientation angles, discretized into bins of 5° (3; 1)
Mean degree	Node	Mean number of edges per intersection
Dead end rate	Node	Percentage of nodes of degree 1 (3)
Percentage degree 4+	Node	Percentage of nodes with degree 4 or higher (3)

Table S1: Urban statistics used to compare city growth correlations.