Density of Large Urban Areas in the U.S. and Barriers to Urban Expansion, 1950-2020

John R. Ottensmann
Indiana University-Purdue University Indianapolis
john.ottensmann@gmail.com
urbanpatternsblog.wordpress.com
May 2023

Abstract

Housing unit densities are examined in 56 large urban areas in the United States defined in a consistent manner from 1950 to 2020. The mean density declined slightly over this period but this masks tremendous variation across the urban areas. Some of the most dense urban areas at the start experienced large drops, but substantial numbers of areas had increases in density, some large. Densities across regions changed dramatically, with mean densities for urban areas in the West rising from only slightly above the South to the highest by 2020, well above the Northeast and the Midwest which were highest in 1950. Density and density change are related to the size of the urban area (number of housing units), and change is also related to change in size and (negatively) to density at the start. The effect of potential barriers to expansion on density is investigated, with strong, significant effects of water and mountains on urban area densities.

Introduction

Density of population and housing is one of the characteristics that makes a place urban. This sociologist Louis Wirth (1938) listed this along with size and heterogeneity as the three defining features of urbanism. The density of an urban area has been shown to be related to a wide range of things from transportation and energy use (Newman and Kenworthy 1989) to the cost of public services (Ladd 1992) and possible aspects of human health (Zhao and Kaestner 2010). Urban sprawl has been identified as a negative aspect of the growth of urban areas, with low densities and scattered, leapfrog development being among the defining characteristics of sprawl (Harvey and Clark 1965; Ewing 1997). Indeed, density has frequently been used as a measure of the degree of sprawl (Pendall 1999; Fulton, *et al.* 2001; Anthony 2004; Angel, *et al.* 2010).

The densities of large urban areas would be expected to be greater where expansion has been limited by the presence of barriers to urban expansion such as water or mountains. Such barriers limit the supply of land available. The increased price would result in less land being used in the development of housing and therefore higher densities.

Examining the densities of urban areas over time has been difficult because of the limitations of the available data. Density can be examined for cities (Bryan, Minton, and Sarge 2007) but these do not include the entire urban area, and their boundaries can change over time. Metropolitan Statistical Areas are delineated using entire counties and include varying and sometimes very large amounts of rural area, making densities for these areas unusable. The census Urbanized Areas are better candidates but how they have been defined has changed greatly over time.

This research examines the housing unit densities of large urban areas over a seventy-year period and the effects of barriers to urban expansion on those densities. It begins with the description of the dataset for the 56 largest urban areas in 2020 delineated using census tract data in a consistent fashion from 1950 to 2020. Measures of potential barriers limiting urban expansion—water, mountains, wetlands, protected lands, and arid climate—are described next. This is followed by the examination of density and change over the period, variation by region, and urban area characteristics associated with density and change in density. The final major section addresses the barriers to urban expansion with models demonstrating the effects of water and mountains on density and change and discussion of the other barriers to urban expansion.

The *Urban patterns 2* data

The *Urban patterns 2* dataset includes housing unit counts for census tracts from 1950 to 2020 that have been used to delineate 56 large urban areas in the United States for each census year. Data for 2010 and 2020 are from the Census and from the National Historical Geographic Information System (Manson, *et al.* 2022). Data from the censuses from 1970 to 2000 are from a unique dataset from the Urban Institute and Geolytics (2003) with the data normalized to 2000 census tract boundaries. Housing units for 1950 and 1960 are estimated from the data on housing units by year built from later years, taking the numbers built before 1950 and 1960 as the estimates of the numbers present in those years. These estimates include error resulting from changes to the housing stock over time, especially the loss of some units, but analyses suggest that the estimates are reasonable for two decades back in time. Census tract boundaries for 2020 are used for the dataset. The census tract relationship files are used to estimate values for the 2020 tracts from data for earlier years. Detailed documentation of the dataset and listings of all data sources are included in Ottensmann (2023a).

Urban areas consist of contiguous census tracts that meet urban criteria. Some large areas of continuous urban tracts include what should reasonably be considered two or more urban areas. Areas in the northeastern United States are a major example.

To distinguish separate urban areas, the Combined Statistical Areas (CSAs) are used (and MSAs that are not included in a CSA). CSAs are used rather than the more commonly used MSAs as they better represent the full extent of urban areas. The CSAs are only used to identify the urban areas, such as Philadelphia, New York, and Hartford. The boundaries are established at the locations where the urban areas have become contiguous as they have expanded. The urban areas included in the dataset are the 56 areas containing more than 300,000 housing units in 2020.

The criteria defining the urban areas are as close as possible to those being used for delineating the 2020 census Urban Areas, which include what were formerly called Urbanized Areas (U.S. Census Bureau 2022). A census tract is considered to be urban and is included in an urban area if it has a housing unit density greater than 200 housing units per square mile. To include urban territory that is nonresidential, a tract is also included if over one-third of its area has impervious surface of 20 percent or more. An additional condition is that a tract is only considered to be urban if it had been designated as urban for the following census year. This is to provide a pattern of cumulative expansion of the urban areas. This direction has been chosen rather than the reverse (if urban, then urban later) because the more recent data are considered to be more accurate.

Some of the urban areas include two or more areas that were originally separate but that have since growth together. Areas that are sufficiently large are considered to be urban centers and are included in an urban area with tracts assigned to one of the urban centers. The Dallas-Fort Worth area is an example. As the areas become contiguous, tracts are assigned to the center growing more rapidly toward the other and to provide more continuous, less irregular boundaries. Areas are considered separate urban centers and are included in an urban area if the number of housing units in 2020 exceeds 16 percent of the total units in the urban area. This cutoff was established by identifying as candidates any initially separate area deemed large enough to potentially be considered an urban center and then setting the threshold. The smallest urban centers in relation to the total size of the urban area are Providence, with Boston; Tacoma, with Seattle; and High Point, with Greensboro and Winston-Salem. Next highest, at 11 percent are Port Charlotte in the Sarasota-Bradenton area and Winter Haven in the Orlando area. The names given to the urban areas include the names of additional urban centers that have been included.

Barriers to urban expansion

The argument that barriers to urban expansion can result in increased density of an urban area is a straightforward case of supply and demand. In reducing the amount of land on which development can profitably take place, demand for the remaining land becomes greater, increasing the price. The higher price reduces the quantity of land used in the development of housing, increasing density.

Note that the more general term "barriers to urban expansion" is being used in this paper rather than calling such areas "undevelopable land" as some others have done. Whether land is developed is a function of both the demand for development as well as the difficulty of its development. The houses hanging off hillsides in Los Angeles show that development is possible (at substantial cost) on land that in other settings might be considered undevelopable. The Back Bay area in Boston was literally that—a bay—before it was filled in during the middle of the nineteenth century. Barriers make development more difficult and less likely but are not presumed to be inviolable.

This section begins by considering some other discussions of barriers to urban expansion. It then proceeds to present the barriers to be examined in this paper and describe the measures used. These include water, mountains, wetlands, protected lands, and an arid climate.

Prior consideration of barriers

Location of an urban area on a coastline makes the water an obvious barrier to urban expansion that has long been recognized. For example, Mills (1972) presents a method for estimating the density gradient for an urban area using only two population values, for the central city and the entire urban area, and the radius of a circle approximating the boundaries of the city. The derivation requires the areas involved, and for coastal urban areas Mills subtracts the proportion of the circle that is water.

In an influential article in *USA Today*, El Nasser and Overburg (2001) examine the extent of sprawl in America's urban areas using a measure combining density and density change. They discuss factors associated with sprawl. Availability of water is seen to affect sprawl as the need to obtain water from centralized water systems keeps development close to the edge of the developed areas. Geographic barriers such as oceans and mountains cause areas to grow more compactly. And federal land surrounding Las Vegas is described as an urban growth boundary.

Looking at broader patterns of urban development, Lang, Popper, and Popper (1995, 1997) argue that the presence of mountains and the arid climate has led to settlement patterns in the American West (west of approximately the 100-degree meridian) to differ from those in the East. Their analysis focuses on comparisons of the states in the two parts of the country, with no further consideration of terrain or climate. Lang (2002) returns to this arguing further that aridity, public and Indian lands, and slope affects the density and change in density for urban development in the West. He compares several measures of the urban area settlement patterns, demonstrating differences between East and West but with no consideration the those specific barriers. And Lopez (2014) also refers to the need for public water and the presence of government-owned land as affecting sprawl in urban areas in the West.

The extent of sprawl as measured by the population density of developed land in metropolitan areas is addressed by Fulton and his coauthors (2001). They cite factors associated with lower levels of sprawl (higher densities) including houses with sewer service, areas with slope over 15 percent, "substantial wetlands," and "much land owned by government." The latter three measures come from the National Resources Inventory. The only results presented, however, consists of the listing of these factors in tables. In addition, they curiously report that greater use of public water is associated with lower densities, opposite to the arguments made for arid lands.

Burchfield and coauthors (2006) consider sprawl to consist of scattered development (though the scattering of development produces lower densities when considered in larger areas). They produce measures of the degree of scattering for all metropolitan areas in 1976 and 1992 using land cover data derived from aerial photography and satellite imagery. They explicitly measure the extent of two barriers to urban expansion on land within 20 kilometers of areas that were "mostly developed" in 1976. These are mountainous areas identified using elevation range and a ruggedness index calculated from elevation data and the percent of land overlying aquifers, allowing development without public water service. They find statistically significant relationships between these measures and sprawl in the expected directions.

Saiz (2010) focuses not on density and sprawl but on housing supply elasticities. He hypothesizes that these are also affected by barriers including land lost to sea and internal water and wetlands and land with slopes exceeding 15 percent. He obtains measures of these for the areas within a 50-kilometer radius of the central city centroid. These are combined to give a single measure of undevelopable land which proves to be a significant predictor.

Sprawl is measured by Paulsen (2014) using four measures associated with density and change from 1980 to 2000 that are constructed for urban areas defined using census block group data. The list of barriers considered includes water and wetlands, steeply sloped land, and areas protected from development. These are combined to create a measure of undevelopable land for the entire metropolitan area in which each urban area is located. He also finds statistically significant relationships in the directions hypothesized.

Finally, an earlier paper on the density of urban areas over time using an earlier dataset (Ottensmann 2015) more informally considers the effects of barriers on density. Whether expansion of an urban area might be affected by barriers such as mountains or wetlands is determined subjectively by visual examination of maps, with different levels of restriction considered. Whether areas are impacted significantly by either mountains or wetlands is significantly related to density and density change. The presence of government-owed lands and arid climate is also considered, but these overlap with mountains such that it was not possible to identify independent effects.

Measures of barriers to urban expansion

In this paper, five potential barriers to urban area expansion are considered: water, mountains, wetlands, protected lands, and arid climate. The measures of these for the 56 large urban areas are described in this section.

The first four barriers represent physical characteristics of specific territory that can inhibit the expansion of an urban area. This raises the question of the areas in which the presence of such barriers should be measured. The presence of, for example, mountainous areas many miles from the existing urban area is likely to have little or no effect on urban expansion and density. Rather, it seems more appropriate to focus on the extent to which areas adjacent to the existing urban area might serve as barriers. In the studies cited above, Burchfield, *et al.* (2006) consider areas within 20 kilometers of the developed urban area. Saiz (2010) likewise limits consideration to the area within 50 kilometers of the city, though using a single value for urban areas of widely varying sizes does not seem appropriate. (A 50-kilometer circle would exclude significant portions of the larger urban areas in the dataset.)

This research will consider the extent of the barriers in a ring five miles wide around each urban area in 2020. Too wide and one is increasing the possibility of including barrier territory having less affect on urban expansion. But too narrow and the possibility increases for urban expansion to continue across the barrier. Visual examination of the patterns of the barriers around the urban areas suggests that the choice may not be that critical for those urban areas likely to be affected by the barriers, as many of the barriers are larger.

The width of five miles was chosen by examining the irregular patterns of some urban areas obviously impacted by the presence of mountains. Examples: The Los Angeles urban area has extended out along both sides of the Santa Ana mountains in Orange and Riverside Counties. These mountains form a tongue projecting into the urban area with urban development on both sides. The width of this area, which his not been breeched by urban development, ranges from about five to nine miles. Similar situations exist in other urban areas where the undeveloped area ranges in width from under four to over nine miles. On the other hand, several examples exist in the Los Angeles area of urban area expansion continuing over mountainous areas three to four miles wide. And the eastern portion of the Santa Monica Mountains separating the main part the Los Angeles area from the San Fernando Valley, about 2.5 to 4.98 miles wide, is completely urbanized.

The first barrier is water, which refers to the portions of the rings within the oceans, Gulf of Mexico, and Great Lakes. To be more precise, this includes all areas outside the land area of the United States, so also included area land areas in Canada and Mexico adjacent to the Detroit, San Diego, and El Paso urban areas. These international borders are considered to be a comparable barrier to urban expansion. While urban development occurs on the other side, as in Windsor across from Detroit, the Detroit urban and metropolitan areas are not considered to extend into Canada. (Urban development across international borders presents an interesting question, but it will not be pursued here.) The national boundary and coastline used is from the National Historical Geographical Information System (Manson, *et al.* 2022) and is used for their data products. It is highly detailed, including inlets, bays, barrier islands, and the wider mouths of rivers. The water barrier measure is the proportion of the area of the ring outside this boundary.

Steeply sloped land, used in other studies, is not an unreasonable indicator of the presence of mountains. However, varying areas of less steeply sloped land are present higher up in mountainous areas. Three datasets providing more comprehensive identification of mountains are provided on the Global Mountain Explorer website (U.S. Geological Survey 2023). The third of these, by Karugelle, *et al.*, is used here. This is the most recent, was created with the highest spatial resolution data, is the one most specifically aimed at classifying land forms, and is arguably the most sophisticated of the three. It also falls between the other two in the amount of land classified as mountainous. The 4-kilometer raster dataset is used, with all 4 mountain classes. The measure is the proportion of the pixels in each ring that are designated as mountains.

The wetlands measure use the data from the National Wetlands Inventory (U.S. Fish and Wildlife Service 2023). This includes surface waters, so those areas are also included among the barrier areas. The areas of the wetlands polygons within the rings as a proportion of the total areas of the rings is used here.

The Protected Areas Database (U.S. Geological Survey (USGS) Gap Analysis Project (GAP) 2022) identifies a range of land areas that could to varying degrees be barriers to urban expansion. First is government-owned land that presumably is not available, such as National Parks and Forests and military bases, state parks, and local parks. Conservation easements where development is restricted are included. Finally, and more ambiguous, are what is termed designated areas, such as areas within the designated boundaries of a National Forest that are privately owned. The sizes of many of these areas may make then unsuitable for extensive urban urban expansion or attempts to develop may face obstacles. For these reasons, it seems reasonable to include such designated areas as barriers to urban expansion as well. One exception occurs in Oklahoma with designated tribal lands. With the state's unique history, major areas in the state are so designated. These include areas not only within the ring surrounding Oklahoma City but within the developed Oklahoma City urban area as well. Given this, these tribal lands are excluded and not considered to be protected lands. The various areas designated as protected lands are merged, as territory can fall into multiple protected areas, for example, a wilderness area within a National Forest. The final measure is the proportion of the land area of the ring within these protected areas.

Arid climate as a barrier to urban expansion is distinctive in not being a physical characteristic of the landscape. Rather, the lack of water can limit or prevent development away from the existing urban area that would rely on wells. Instead, new development is forced to locate close to the existing urban area where existing water lines are incrementally extended to provide service by a larger utility that in many cases obtains the water from a considerable distance. The presence of an arid climate is indicated by the average annual precipitation. Since precipitation can vary somewhat throughout an urban area, especially those with considerable variation in elevation, continuous, interpolated climate data (USA PRISM Climate Group 2023) is used to estimate the mean average annual precipitation across the 2020 urban areas. The distribution of these values shows a large gap between the 11 urban areas with average annual precipitation of less than 22 inches and the remaining 45 urban areas receiving an average of 32 inches or more. The former group is classified as having an arid climate for the analysis.

Housing unit densities, 1950-2020

This section begins with basic descriptions of housing unit densities and change from 1950 to 2020 and and how these vary across regions. This is followed by some simple models of the relationships of density and change in density to the sizes of urban areas, changes in the sizes of urban areas, and the densities of urban areas. While of interest in themselves, these models also serve as the starting point for the consideration of the effects of barriers to urban expansion on density and change.

Description

Basic descriptive statistics on housing unit densities for the 56 large urban areas in each census year from 1950 to 2020 are presented in Table 1. These results provide a very mixed picture. Mean density shows small declines through 2000, followed by an uptick in the final two decades. This would seem to be only very weak confirmation of the observation by others of decreasing densities of urban areas over time (Angel, *et al.* 2010).

Looking at the extremes in each year shows very different patterns of change over time. Change in the maximum housing unit density is most dramatic, plummeting from about 3,000 housing units per square mile in 1950 to just under 2,000 by 2020. The minimum, on the other hand, rises from 400 to 645 in the first two decades before dropping back down to 514 by the end of the seventy-year period. Densities became

Table 1. Summary statistics for housing unit densities of large urban areas, 1950-2020.

Year	Mean	Standard deviation	Minimum	Maximum
1950	1,234	510	400	3,016
1960	1,167	387	488	2,541
1970	1,154	347	645	2,227
1980	1,141	307	589	2,062
1990	1,158	316	546	1,944
2000	1,093	312	513	1,889
2010	1,098	326	521	1,955
2020	1,121	341	514	1,988

more uniform across the urban areas, with the standard deviation dropping from over 500 at the beginning to values something above 300 after 1960.

It helps to look at the urban areas with the highest and lowest densities. Table 2 lists the six urban areas having the highest housing unit densities and the six areas with the lowest densities in both 1950 and 2020. (Six areas are just over ten percent of all of the large urban areas.) The list of the most dense urban areas in 1950 would not be that much of a surprise to those familiar with urban America. New York was by far the densest urban area at just over 3,000 housing units per square mile. It was followed by three other very large urban areas—Philadelphia, San Francisco Oakland-San Jose, and Chicago—along with two more of the older areas in the country, Buffalo and New Orleans. All had densities somewhat above 2,000 units per square mile.

The list of the urban areas with the highest densities in 2020 is likely to be more of a surprise to many. New York remains on the list but in second place, edged out by Las Vegas, both with densities below 2,000 housing units per square mile, below all of the top 6 in 1950. San Francisco-Oakland-San Jose retains its position in third place but is followed by Los Angeles and Miami-Fort Lauderdale-West Palm Beach with nearly identical densities, all ranging from 1,834 to 1,859 units per square mile. San Diego is the sixth densest urban area with a density about 200 lower. New York and San Francisco-Oakland-San Jose were the only areas among the most dense in both years. New York accomplished this in spite of a huge decline by starting with by far the highest density in 1950. San Francisco-Oakland-San Jose experienced a modest drop in density over the period but maintained its position given the overall lower densities in 2020.

Only one area, Greenville-Spartanburg, was among the areas having the lowest densities in both years. But nearly all of the areas in both years shared one thing in

Table 2. Urban areas with the highest and lowest housing unit densities in 1950 and 2020.

1950		2000	
Area	**Density**	**Area**	**Density**
New York	3,016	Las Vegas	1,988
Philadelphia	2,182	New York	1,961
San Francisco-Oakland-San Jose	2,136	San Francisco-Oakland-San Jose	1,859
Chicago	2,133	Los Angeles	1,855
Buffalo	2,115	Miami-Fort Lauderdale-West Palm Beach	1,834
New Orleans	2,065	San Diego	1,621
* * *		* * *	
Nashville	684	Atlanta	746
Greenville-Spartanburg	666	Raleigh-Durham	684
Albuquerque	659	Birmingham	664
Orlando	625	Charlotte	664
Sarasota-Bradenton	498	Greensboro--Winston-Salem--High Point	603
Cape Coral-Fort Myers-Naples	400	Greenville-Spartanburg	514

common. They were in the South, in the area east of the Mississippi River. The one exception was the Albuquerque urban areas with a very low housing unit density in 1950 but not in 2020. One thing that characterized the five least dense urban areas in 1950 was their small size. Each had about 30,000 or fewer housing units, with the Cape Coral-Fort Myers-Naples having a truly tiny 2,720 units. This compares to a mean of about 250,000 housing units for all of the 56 urban areas in that year.

Further insight is provided by the lists of the areas experiencing the greatest increases in density from 1950 to 2020 and the greatest density declines, shown in Table 3. Based on the Las Vegas urban area emerging as most dense in 2020, it is not surprising that it experienced the greatest increase in density of about 1,300 units per square mile, jumping from a modest density of 684 units per square miles in 1950. Miami-Fort Lauderdale-West Palm Beach likewise made the list of the most dense areas in 2020 with the second greatest increase of over 700 and San Diego went up by over 600. After starting with the lowest density in 1950, Cape Coral-Fort Myers-Naples had

Table 3. Urban areas with the greatest density increases and the greatest density declines from 1950 to 2020.

Greatest density increase	
Area	**Density increase**
Las Vegas	1,304
Miami-Fort Lauderdale-West Palm Beach	759
Albuquerque	656
San Diego	626
Phoenix	544
Cape Coral-Fort Myers-Naples	544

Greatest density decline	
Area	**Density decline**
New York	1,055
Philadelphia	1,007
Buffalo	988
Milwaukee	857
New Orleans	773
Chicago	770

an increase of over 500, still leaving it with below average density in 2020. The six areas on the list of those with the greatest density increases are all either in the West or in Florida.

Turning to the areas with the great declines in density from 1950 to 2020, New York was the leader with a drop of over 1,000. But Philadelphia experienced nearly as large a drop. Five of the six areas having the largest decreases in density were among the areas with the highest densities in 1950. Milwaukee was the other area having a large drop in density. Among the areas with the greatest declines in density, all but New Orleans were located in the Northeast and Midwest.

This has focused on the urban areas with the highest and lowest densities and density declines. To provide a broader picture of how densities changed across the range of urban areas, plots of urban area densities by year for five urban areas are shown in Figure 1. The urban areas included were selected to reflect the full range of density change across the urban areas. New York, as noted above, had the minimum change in density (the greatest density decline) from 1950 to 2020. Omaha was at the first quartile of density change, dropping by just over 400 persons per square mile. Atlanta was the median area (one of the two) with a small decrease of 139. Dallas-Fort Worth had the third quartile increase of nearly 200. And as also noted above, Las Vegas had the maximum increase in density over 1,300 housing units per square mile.

Figure 1 illustrates how the five urban areas experienced very different trajectories in their densities from 1950 to 2020. New York, in red, shows the extreme drop in density (though with a slight uptick near the end of the period. Omaha, in blue, also trended downward but in a much less dramatic fashion. With Atlanta, in green, one

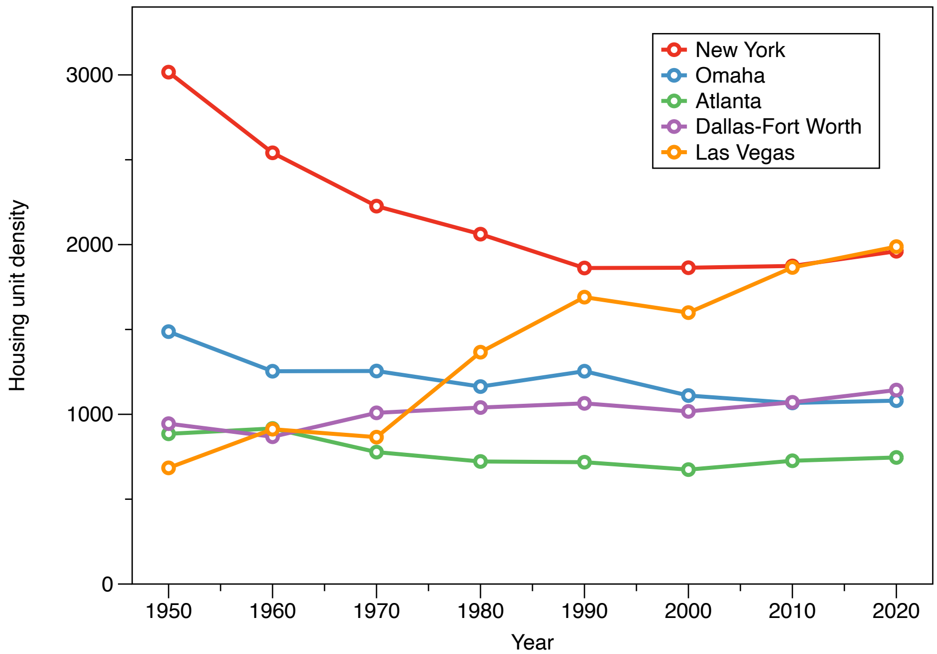

Figure 1. Housing unit densities from 1950 to 2020 for the urban areas with the minimum, first quartile, median, third quartile, and maximum change in density from 1950 to 2020.

can see some decline but it is not very great. Dallas-Fort-Worth, in purple, saw density increase gradually. And then Las Vegas, in orange, shows the greatest change in either direction, slightly surpassing New York in 2020.

Very apparent from the lists of the areas with the highest and lowest densities and density changes is the variation that occurs by region. For this, we are using the locations of the urban areas in the four census regions with one exception. The Washington-Baltimore urban areas lies in the South census region, is south of the Mason-Dixon line in slave-holding states. But for consideration of the largest urban areas in the United States, Washington-Baltimore is clearly a part of the nearly continuous area of urban development extending up to Boston, an area often referred to as the Northeast Corridor or megalopolis. For this reason, the Washington-Baltimore urban area is classified here as being located within the Northeast region.

Figure 2 plots mean densities by region in each of the census years from 1950 to 2020. The changes are dramatic. The urban areas in the Northeast, in red, and the Midwest, in blue begin the period with average densities far higher than in the other two regions, about 1,800 and 1,600 versus about 950 and 1,100 for the South and West. Their average densities steadily decline, pretty much in lockstep with each other, until

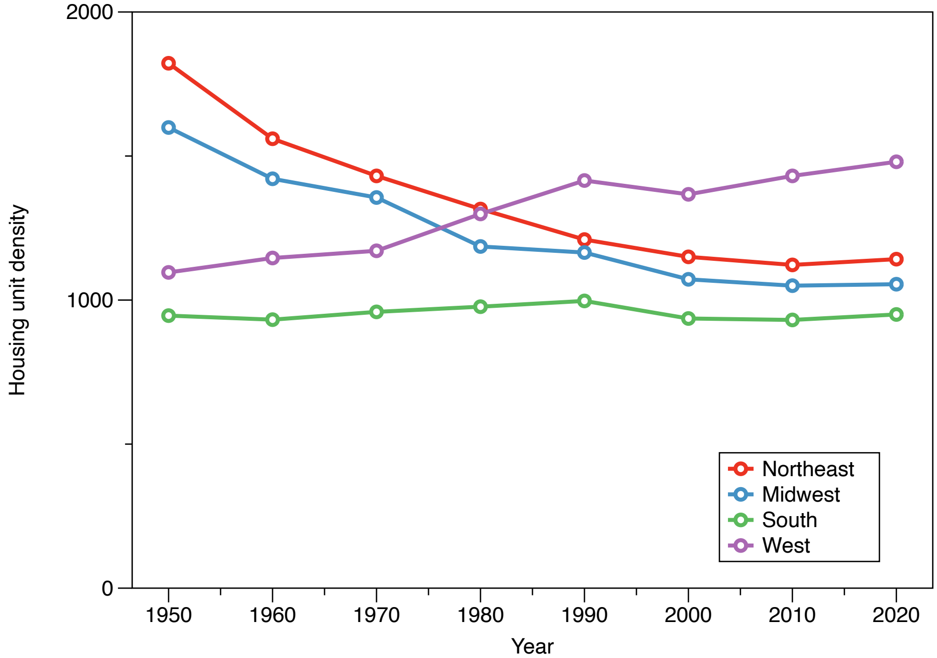

Figure 2. Mean housing unit density for urban areas in each region from 1950 to 2020.

by 2020 they are both somewhat over 1,000, not that much higher than the mean density in the South.

Mean densities in the South, in green, show little change over the period, staying just under 1,000. The South consistently has the lowest average densities for its urban areas among the four regions.

The large declines in mean densities in the Northeast and Midwest are mirrored by the large increases among the urban areas in the West, in purple. The average urban area density in the West jumps from about 1,100 in 1950 to nearly 1,500 by 2020, by far the highest mean density among the four regions. This raises the question of how it came to be that this region has the densest urban areas, which will be addressed at least in part later in examining the effect of barriers to urban expansion.

The differences in mean densities by region in each year from 1950 to 2020 and the differences in mean change in density by region over this period are all highly statistically significant.

Basic models predicting density and density change

Demand for locations closer to the center will be greater in larger urban areas. Thus it is reasonable to expect that the density of an urban area and change in density

Table 4. Regression models predicting density in 1950 and 2020 using the size of the urban area (standard errors in parentheses).

	Density 1950		Density 2020	
Log housing units in year	319.1	***	217.9	***
	(29.0)		(49.2)	
Constant	-2,448.3	***	-1,859.8	**
	(336.4)		(674.0)	
R^2	0.692	***	0.267	***

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

will be related to the size of the area and its change. This is taken as the starting point for developing simple models predicting density and density change.

First comes the prediction of the density of urban areas. The number of housing units in the urban areas is the measure of size, with the natural logarithm of housing units used. The change in the number of housing units from some earlier time is also evaluated as a predictor. The log of housing units proves to be statistically significant in the prediction of density. Change in the number of housing units is not significant. The final, very simple models use only the log of housing units as the predictor of density.

Table 4 presents the results of the regressions of density on the log of housing units for 1950 and 2020. Results are significant at the 0.001 level in both years. The relationship is especially strong in 1950, with the an R^2 of 0.69, so over two-thirds of the variation in density across the 56 urban areas is accounted for by size. This drops to 0.27 in 2020, still highly significant, and the regression coefficient for log of housing units drops from 319 in 1950 to 218 in 2020. So the relationship has weakened. Results for the intermediate years show generally steady declines in the value of R^2 and the regression coefficient over time.

Considering density over some period of time, size at the start of the period and the change in size might be expected to be related to change in density. In addition, given the dramatic declines in density observed for some of the most dense urban areas in 1950, density at the start of the period could be negatively related to the change in density. Again, the log of housing units is taken as the measure of the size of the urban area and the change in the log of housing units over the period is the measure of change.

Table 5 gives the results for models using these variables to predict the change in density from 1950 to 2020, the entire period, and from 1950 to 1960 and from 2010 to 2020, the first and last decades. The models for density change over the entire period

Table 5. Regression models predicting change in density from 1950 to 2020 and in the first and last decades of the period (standard errors in parentheses).

	Change in density 1950-2020		Change in density 1950-1960		Change in density 2010-2020
Log housing units at start of period	114.9	*	67.5	**	-3.4
	(56.7)		(24.3)		(10.3)
Change in log housing units over period	283.0	***	222.1	*	28.3
	(63.0)		(83.3)		(89.9)
Housing unit density at start of period	-0.573	***	-0.359	***	0.043
	(0.136)		(0.061)		(0.027)
Constant	-1,336.8	*	-531.1	*	18.0
	(664.5)		(258.1)		(129.3)
R^2	0.707	***	0.625	***	0.053

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

and for the first decade are both highly significant, with R^2 values of 0.71 and 0.63. The regression coefficients for all three of the predictors are statistically significant at the 0.05 level or better. Directions of the relationships are as expected. Density change is positively related to log of housing units in 1950 and the change in the log of housing units over the respective periods and is negatively related to housing unit density at the start of the period in 1950. Regression coefficients on log housing units and density in 1950 are lower for the second model predicting change over just the first decade, expected as the amount of change over the single decade will generally be less than the change over the entire 70-year period.

The third model predicting the change in density in the final decade from 2010 to 2020 is not even close to being significant, having an R^2 of only 0.05. Looking at the models for change in each of the intervening decades shows all but the model for the final decade to be statistically significant, but with R^2 dropping in an irregular fashion from a high of 0.63 for the first decade to lows of about 0.15 for the decades 1990 to 2000 and 2000-2010. At least one regression coefficient is significant at the 0.05 level or better in all but the final decade, but only for the model predicting change in density from 1950 to 1960 are the coefficients for all three predictors significant. The set of variables does well at predicting change in density over the entire period from 1950 to 2020 and for the earlier decades, but less well in the later decades.

These models add to understanding the relationships of density and density change to some basic characteristics of the urban areas. They also play a significant role as the starting points for the assessment the effect of the barriers to urban expansion on density, controlling for the effects of the basic urban area characteristics.

Effects of barriers to urban expansion on density

The effects of five potential barriers to the expansion of urban areas are considered—water, mountains, wetlands, protected lands, and arid climate. Some basic description of the barriers precedes the results. These are models predicting density and density change using water and mountains as the barriers to expansion. This is followed by discussion of the effects of the remaining barriers.

The barriers to urban expansion

As described above, the measures for the first four barriers are the proportions of the land in the five-mile wide ring around the 2020 urban areas that are covered by the barrier. (For arid climate, the measures are based on mean annual precipitation in inches.) Not surprisingly, the distributions are highly skewed. The majority of the urban areas have little or no land impacted by the barriers. The median proportions are zero for water and mountains, 0.06 for wetlands, and 0.12 for protected lands. For three-quarters of the areas, the barrier proportions are 0.27 or less. This makes the normal descriptive statistics not particularly helpful in describing the barriers.

A better sense of the extent of the barriers is to consider a few of the urban areas that have the greatest barriers. For water, Honolulu, Norfolk-Virginia Beach, and Sarasota-Bradenton have the highest proportions in the barriers, with values ranging from 52 to 62 percent. New Orleans, Orlando, and Jacksonville have the largest proportions in wetlands, 32 to 45 percent. Mountains, protected lands, and arid climate are considered together. The five urban areas with the highest proportion mountains range from 60 to 97 percent. For protected lands the percentages go from 55 to 85. And for arid climate the lowest five have mean precipitation less than 11.9 inches. The reason for looking at the three barriers together is their overlap: Las Vegas, Tucson, and Albuquerque are on all three lists of the top five and Phoenix and Salt Lake City are on two of the lists.

To further examine the relationships among the barriers, Table 6 gives the correlations among the barriers measures. A modest negative correlation of -0.27 exists between mountains and wetlands that is barely significant and not surprising. Nor is the similar positive correlation of water with wetlands. Wetlands are much more highly correlated with precipitation at 0.45, which makes perfect sense. And now for mountains, protected lands, and precipitation: The correlations among these measures

Table 6. Correlations among barrier measures.

	Water	Mountains	Wetlands	Protected lands
Mountains	-0.082			
Wetlands	0.258	-0.273 *		
Protected lands	-0.034	0.854 ***	-0.034	
Precipitation	0.235	-0.740 ***	0.448	-0.669 ***

** p < 0.05 ** p < 0.01 *** p < 0.001*

range in magnitude from 0.67 to 0.85, with the correlations with precipitation being negative, of course. These high correlations will become an issue in developing models to estimate the effects of the these barriers on density and change. The discussion of the urban areas having the highest values for the different barriers suggests significant regional differences (though this would have already been obvious to anyone with the most basic knowledge of North American geography). Table 7 gives the mean values for the barrier measures for the urban areas in each of the regions.

On average the regions have between 9 and 21 percent of their ring areas with water, so the presence of this barrier is broadly distributed and differences across the regions are not significant. Mountains, on the other hand, are concentrated mainly in the West, with an average presence of 57 percent versus 0 to 3 percent for the urban areas in the other regions. Obviously differences are highly significant. The presence of wetlands is greatest with the South averaging 13 percent, are in the middle at about 8

Table 7. Mean proportions of land in rings around urban areas water, mountains, wetlands, and protected lands and mean annual precipitation by region.

	Northeast	Midwest	South	West
Water	0.214	0.093	0.134	0.137
Mountains***	0.033	0.000	0.019	0.569
Wetlands	0.082	0.084	0.132	0.053
Protected lands***	0.127	0.064	0.116	0.422
Precipitation***	45.0	39.1	48.2	21.9

** p < 0.05 ** p < 0.01 *** p < 0.001*

percent in the Northeast and Midwest, and are lowest in the West. The West has the highest average for protected lands at 42 percent, but the other regions have moderate amounts, with the Midwest lowest at 6 percent. This may result from the extent to which the lands surrounding the large urban areas are primarily agricultural. Finally, mean precipitation for the areas in the West is 22 inches, with the areas in the other regions averaging close to double that or more. As discussed earlier, an alternative measure for arid climate designates as arid the urban areas with average annual precipitation of less than 22 inches, as a large gap exists between those and the areas with higher levels of precipitation. Using this criterion, 11 urban areas have an arid climate. Ten of these areas are in the West. with that region having only three urban areas that do not have an arid climate. The one urban area outside the region with an arid climate was El Paso. And while El Paso, in Texas, is considered to be in the South, it is about as far west as Albuquerque and is farther west than Denver.

Water and mountains as barriers to urban expansion

Models are presented that use water and mountains as barriers to urban expansion to predict housing unit density and density change in the 56 large urban areas. The measures are the proportion of the areas of the five-mile rings around the 2020 urban areas that are outside the coastline (or land boundary) of the United States and the proportion identified as mountains. The starting points are the models predicting density and density change presented earlier using the log of the number of housing units in the urban area, change in the log of housing units, and density at the start of the period of change.

The base model predicting the housing unit density of the urban areas in 2020 had a single predictor, log of housing units, and an R^2 value of 0.27 (Table 4). First, the water and mountains barrier variables are each added separately to that model, with the regression model results shown in Table 8. Each is highly significant at the 0.001 level. Water as a barrier increases R^2 to 0.35. The mountains barrier performs even better, with the value of R^2 at 0.55.

The final model includes both water and mountains in the model together along with the log of housing units. Both remain highly significant with the regression coefficients are not significantly changed from the values when in the model separately. R^2 jumps to 0.67. This model is accounting for two-thirds of the variation in density across the urban areas.

Consider the effect of an increase of 0.20 in either of the barrier variables. Given that the barrier measures go up to 0.62 for water and 0.97 for mountains and over a quarter of the urban areas have values greater than 0.20, this is not an unreasonably large difference to consider. A change of that amount in the proportion of the ring area water would be associated with an increase of about 130 housing units per square mile.

Table 8. Regression models predicting density in 2020 using the water and mountains barriers (standard errors in parentheses).

	Density 2020	Density 2020	Density 2020
Proportion ring area water	557.0 *	— —	660.4 ***
	(210.3)		(152.3)
Proportion ring area mountains	— —	708.4 ***	752.7 ***
		(122.5)	(106.5)
Log housing units in 2020	185.9 ***	208.2 ***	169.7 ***
	(48.2)	(38.9)	(34.8)
Constant	-1,499.5 *	-1,829.6 **	-1,400.4 **
	(653.7)	(532.8)	(471.4)
R^2	0.352 ***	0.550 ***	0.670 ***

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

The difference between no water barrier and the maximum value of 0.62 predicts density to be over 400 units per square mile higher. With the larger regression coefficient, the predicted effects of a change in the mountains barrier is even larger. An increase of 0.20 would be associated with an increase in density of 150. The difference between no mountains and the maximum of 0.97 for Las Vegas is 730 units per square mile. Given this, it should not be that surprising that Las Vegas has the highest density of all of the urban areas in 2020.

These results consider the effects of the barriers on density in 2020. The barriers reflect the proportions of the areas of the rings around the 2020 urban areas, so they are measures of the extent to which an urban area was up against the barriers in that year. In earlier years the urban areas were smaller and additional land would have been available, areas into which the urban area could and did expand. Given this, the expectation is that the effects of the barriers on density would have been less, though they still should have had some effect as their presence indicated limitations on the availability of land in the long run.

Regressions comparable to the final model for 2020 including both water and mountains as barriers were carried out for each of the earlier years back to 1950. Since the interest is primarily in the estimated regression coefficients for water and mountains, only these values along with R^2 are shown for each year in Table 9. The results confirm the expectations. The regression coefficients tend to decline from their maximum values in 2020 to the smallest values in 1950. For water, this is a drop from

Table 9. Regression coefficients for water and mountains predicting density in each year from 1950 to 2020.

Year	Regression coefficients		R^2
	Proportion water	Proportion mountains	
1950	174.0	-96.6	0.699 ***
1960	256.9	53.2	0.688 ***
1970	298.5	98.9	0.641 ***
1980	386.8 *	343.0 **	0.604 ***
1990	454.2 **	546.7 ***	0.593 ***
2000	567.8 ***	574.6 ***	0.657 ***
2010	545.6 ***	714.9 ***	0.680 ***
2020	660.4 ***	752.7 ***	0.670 ***

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

660 to 174 and for mountain, from 753 to -97 (though this negative value is not significantly different from zero. In the final three years, 2000 to 2020, the coefficients for both water and mountains were highly significant with p-values less than 0.001. Conversely, for the first three years from 1950 to 1970, the barriers did not have a statistically significant effect on density. The two intervening years saw some decline in the levels of significance compared to the later years.

Attention now turns to the change in density. The initial models predict the change in housing unit density from 1950 to 2020 for the urban areas. The base model had three significant predictors, the log of housing units in 1950, the change in the log of housing units over the period, and the housing unit density in 1950. That model had a very high R^2 value of 0.71 (Table 5), so it already accounted for a very large proportion of the change in density.

As with the models predicting density, the first two regression models show the effect of adding the water and mountain barrier variables individually to the base model, with these results in Table 10. Including the proportion of the ring area water only slightly improves the fit of the model, raising R^2 only to 0.73. The water barrier is significant at the 0.05 level, but including this in the model causes the log of housing units in 1950 to become not significant. The mountains barrier has a larger effect, increasing R^2 to 0.85. That barrier is highly significant and the log of housing units remains significant as it was in the base model.

Table 10. Regression models predicting change in density from 1950 to 2020 using the water and mountains barriers (standard errors in parentheses).

	Change in density 1950-2020		Change in density 1950-2020		Change in density 1950-2020	
Proportion ring area water	459.5	*	— —		558.3	***
	(202.4)				(129.8)	
Proportion ring area mountains	— —		765.5	***	800.5	***
			(106.7)		(92.4)	
Log housing units in 1950	101.483		93.887	*	76.644	*
	(54.910)		(40.509)		(35.181)	
Change in log housing units 1950-2020	251.994	***	223.184	***	182.738	***
	(62.130)		(45.623)		(40.471)	
Housing unit density in 1950	-0.626	*	-0.545	***	-0.608	***
	(0.133)		(0.097)		(0.085)	
Constant	-1,114.1		-1,112.3	*	-831.5	
	(646.9)		(474.3)		(414.4)	
R^2	0.734	***	0.854	***	0.893	***

** p < 0.05   ** p < 0.01   *** p < 0.001*

The final model includes both water and mountains. The regression coefficients for each are higher than when in the model alone, with the larger increase of nearly 100 to 558 for the water barrier. Both are highly significant at the 0.001 level. And the value of R^2 goes up to 0.89. This model is accounting for all but about ten percent of the variation in the change in density.

The sizes of the regression coefficients for water and mountains are not that different from those in the model predicting density, so the predicted effects of changes in the values for water and mountains are similar. An increase in water of 0.20 would be associated with an increase over 110 in the change in density, and the difference between no water and the maximum implies a change in density of nearly 350. For a change on 0.20 in mountains, this would be an increase in the change in density of 160 while going from no mountains (the case for most of the urban areas) to the maximum of 0.97 for Las Vegas implies an increase in the change in density of around 775. This would account for a significant share of the change in density of about 1,300 for Las Vegas, the largest change, and the Las Vegas area also had a very large increase in the number of housing units over the period and started with a very low density, so those can account for much of the remaining change in density.

The models in Table 10 predict the change in housing unit density over the entire seventy-year period from 1950 to 2020. For the base model predicting density change, tests were also conducted predicting the change in each individual decade, with results for the first decade and for the final decade reported in Table 5. The model performs well, with high R^2 and all regression coefficients significant when predicting the change in density from 1950 to 1960. But for 2010 to 2020, nothing is significant. Very much the same pattern occurs in predicting density change in each decade including water and mountains. Both water and mountains are significant in predicting density change in the first decade. Neither are significant and R^2 is low predicting change in the final decade. And again performance in the predictions of density change in the decades from 1960 to 2010 falls between these extremes. The water and mountains barrier variables are significant in some of those periods and not in others. As with the base models, an unclear picture emerges in the prediction of density change for the individual decades. The amounts of the change are apparently too small compared to the random variation over the ten-year periods. Looking at density change over the entire seventy-year period allows a more consistent pattern to emerge.

Wetlands, protected lands, and arid climate

Water and mountains as barriers to urban expansion are important predictors of both housing unit density and the change in density from 1950 to 2020. The regression coefficients are significant, both statistically and substantively. For differing reasons wetlands, protected lands, and arid climate did not enhance the prediction and improve the models. Their performance is discussed in this section, focusing on the predictions of density in 2020. Models predicting density in earlier years and density change raise the same issues.

When the proportion of the ring area classified as wetlands is added to the model including water and mountains as barriers to urban expansion, it is not statistically significant and produces a negligible improvement in R^2. Wetlands also is not significant when added to the base model alone, without the other barriers. Various transformations of the wetlands barrier have been tried tried but all are not significant.

This is not to conclude that wetlands are not a barrier to urban expansion. The number of urban areas with substantial proportions of their ring areas in wetlands is small. The area with the sixth-highest proportion wetlands is Minneapolis-St. Paul with only 20 percent wetlands (which also includes lakes as a major element in that area). This illustrates another issue: With the rather low proportions wetlands compared to other barriers, wetlands are unlikely to significantly affect density unless they are combined with other barriers. The Miami-Fort Lauderdale-West Palm Beach area is a clear example, sandwiched between the Everglades and the Atlantic Ocean. In an attempt to capture a combined effect, interaction terms between forms of the wetlands variable and water have been included in models but these are also uniformly not significant. The limited number of urban areas significantly impacted by wetlands is small, so wetlands are not a statistically significant predictor of density and density change.

The problem with estimating the effect of protected lands and arid climate on density and density change arises from the overlap of these barriers with mountains. Urban areas that have high proportions of the ring areas mountains tend to also have high levels of protected lands and arid climates. As shown in Table 6, the correlations of mountains with the proportion protected lands and mean annual precipitation are 0.85 and -0.74. The binary measure of arid climate designates arid climate for areas having mean annual precipitation less than 22 inches. For those eleven urban areas, the mean proportion mountains is 0.59 while for the remaining 45 areas, the mean is 0.04.

The results when including protected lands and arid climate are similar. Starting with protected lands, adding that barrier to a model including water as a barrier but not mountains, then protected lands is highly significant. R^2 is increased to 0.57 compared to 0.35 for the model including only water but not protected lands. Including protected lands in a model that also includes mountains and protected lands becomes completely not statistically significant, and the regression coefficient is actually a small negative value. R^2 with mountains and water but not protected lands is 0.670, substantially greater than the model with water and protected land. Adding protected lands to that model and R^2 remains the same to three significant digits. So mountains appear to be a more effective barrier than protected lands. Given their relationships, it is impossible to discern any independent effect of protected lands in addition to mountains.

Arid climate as a barrier performs similarly to protected lands. The binary measure of arid climate as mean annual precipitation less than 22 inches is used as that performs better than mean annual precipitation in inches. Add that to the model with water as a barrier but not mountains and arid climate is highly significant. R^2 increases to 0.58 from the value of 0.35 with only water as a barrier. But once again, including arid climate in a model with mountains as well as water and arid climate becomes completely not significant with a small negative coefficient. And R^2 again stays at 0.670, no difference to three significant digits. The conclusion for arid climate is the same as for protected lands. Mountains is a more effective barrier and one cannot see any independent effect of arid climate.

These results for protected lands and arid climate do not indicate that they are not barriers to urban expansion. They only indicate that given the close relationship of these two barriers to mountains as a barrier, it is impossible to estimate any separate effect that might exist without those relationships.

Conclusions

The patterns of density and density change across the 56 large urban areas do not allow a simple story to be told, but the emphasis must be on change. The set of urban areas that are most dense in 1950 differ little from the set at the beginning of the twentieth century. These are the large urban areas in the Northeast and Midwest along with several older large urban areas outside those regions. By 2020, the densest urban areas include newer areas, led by Las Vegas. Densities vary far less across the urban areas in 2020.

How urban area densities changed varies widely. The decline in density among some of the largest and most dense areas has received much attention, with the densities in the New York and Philadelphia urban areas dropping by over one thousand housing units per square miles from 1950 to 2020. At the same time, other urban areas experience density increases, with the density of the Las Vegas area increasing by more than New York and Philadelphia declined. And the densities of other areas change little.

Region is the one thing showing clear differences in density and density change across the urban areas. In 1950, the urban areas in the Northeast have the highest average densities, followed closely by those in the Midwest. By 2020 their mean densities drop to only slightly above the average for areas in the South, which remains low over the seventy-year period. Urban areas in the West experience dramatic increases in average densities, rising from slightly above that for the South in 1950 to well above all of the other regions by 2020.

Looking at the urban area characteristics associated with density and change, the size of the urban areas measured by the log of the number of housing units was related to density in each of the years. The change in density from 1950 to 2020 was related to the size of the area, the change in the size of the area, and negatively to the density of the area in 1950. In other words, higher density urban areas at the start were the biggest losers.

Barriers to urban expansion are expected to be related to higher densities. Urban areas having more water or mountains within five miles have higher densities. The very high density of Las Vegas, surrounded by mountains, and the high density of Honolulu with the highest proportion of the surrounding area water, are illustrations of the effects of these barriers.

Wetlands, protected lands, and arid climate do not show such clear relationships to density. A significant effect was not found for wetlands, likely because of the limited number of urban areas significantly impacted by this barrier. Protected land and arid climate are so closely related to mountains as barriers that no independent effect of these potential barriers is observed.

This analysis considered the overall density of the urban areas. Many additional interesting questions involve the distribution of density within the urban areas.

Densities in the urban core and suburban periphery and the negative exponential decline of density with distance from the center are among the questions to be considered in subsequent papers.

References

Angel, Shlomo, Jason Parent, Daniel L. Civco, and Alejandro M. Blei. 2010. The persistent decline in urban densities: global and historical evidence of 'sprawl.' Lincoln Institute of Land Policy Working Paper. Downloaded from *http://www.lincolninst.edu/publications/working-papers/persistent-decline-urban-densities* on August 22, 2015.

Anthony, Jerry. 2004. Do state growth management regulations reduce sprawl? *Urban Affairs Review* 39, 3: 376-397.

Bryan, Keven A., Brian D. Minton, and Pierre-Daniel G. Sarte. 2007. The evolution of city population density in the United States. *Federal Reserve Bank of Richmond Economic Quarterly* 93, 4: 341-360.

Burchfield, Marcy, Henry C. Overman, Diego Puga, and Matthew A. Turner. 2006. Causes of sprawl: a portrait from space. *Quarterly Journal of Economics* 121, 2: 587-633.

El Nasser, Haya, and Paul Overberg. 2001. A comprehensive look at sprawl in America. *USA Today (*February 22). Downloaded from *http://usatoday30.usatoday.com/news/sprawl/main.htm* on April 9, 2014.

Ewing, Reid. 1997. Is Los Angeles-style sprawl desirable. *Journal of the American Planning Association* 63, 1: 107-126.

Fulton, William, Rolf Pendall, Mai Nguyen, and Alice Henderson. 2001. *Who Sprawls Most?* Brookings Institution.

Harvey, Robert O., and W. A. V. Clark. 1965. The nature and economics of urban sprawl. *Land Economics* 41, 1: 1-9.

Karagulle, Deniz, Charlie Frye, Roger Sayre, Sean Breyer, Peter Aniello, Randy Vaughan, and Dawn Wright. 2017. Modeling global Hammond landform regions from 250-m elevation data. *Transactions in GIS* 21, 5: 1040-1060.

Ladd, Helen F. 1992. Population growth, density and the cost of providing public services. *Urban Studies* 29, 2: 273-295.

Lang, Robert E., Deborah E. Popper, and Frank J. Popper. 1995. Progress of the nation: the settlement history of the enduring American frontier. *Western Historical Quarterly* 26, 3: 289-307.

Lang, Robert E., Deborah E. Popper, and Frank J. Popper. 1997. Is there still a frontier? The 1890 U.S. Census and the modern American West. *Journal of Rural Studies* 13, 3: 377-386.

Lang, Robert E. 2002. Open spaces, bounded places: does the American West's arid landscape yield dense metropolitan growth? *Housing Policy Debate* 13, 4: 755-778.

Lopez, Russell. 2014. Urban sprawl in the United States: 1970-2010. *Cities and the Environment (CATE)* 7, 1: article 7. Downloaded from *http://digitalcommons.lmu.edu/cate/vol7/iss1/7* on June 11, 2018.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2022. *IPUMS National Historical Geographic Information System: Version 17.0* [dataset].

Minneapolis MN: IPUMS. *http://doi.org/10.18128/D050.V17.0*. Data downloaded from *https://www.nhgis.org* on various dates, August to October 2022.

Mills. Edwin S. 1972. *Studies in the Structure of the Urban Economy*. Johns Hopkins Press for Resources for the Future.

Newman, Peter W. G., and Jeffrey R. Kenworthy. 1989. Gasoline consumption and cities: a comparison of U.S. cities with a global survey. *Journal of the American Planning Association* 55, 1: 24-37.

Ottensmann, John R. 2015. Density of large urban areas in the U.S., 1950-2010. Available at *https://urbanpatternsblog.files.wordpress.com/2016/12/density-of-large-urban-areas.pdf*.

Ottensmann, John R. 2023a. Urban patterns of the largest urban areas in the United States, 1950-2020: a new dataset. Available at *https://urbanpatternsblog.files.wordpress.com/2023/02/new-urban-patterns-dataset.pdf*.

Ottensmann, John R. 2023b. Change in the size and distribution of large urban areas in the U.S., 1950-2020. Available at *https://urbanpatternsblog.files.wordpress.com/2023/03/urban-area-size-and-distribution.pdf*.

Pendall, Rolf. 1999. Do land-use controls cause sprawl? *Environment and Planning B* 26, 5: 555-571.

Paulsen, Kurt. 2014. Geography, policy or market? New evidence on the measurement and causes of sprawl (and infill) in US metropolitan regions. *Urban Studies* 41, 12: 2629-2645.

Saiz, Albert. 2010. The geographic determinants of housing supply. *Quarterly Journal of Economics* 125, 3: 1253–1296.

U.S. Census Bureau. 2022. Urban Area criteria for the 2020 Census—final criteria. *Federal Register* 97, 57 (March 24): 16706-16715.

U.S. Fish and Wildlife Service. 2023. *National Wetlands Inventory* (website). Data downloaded from *https://www.fws.gov/node/264847* on January 13, 2023.

U.S. Geological Survey. 2023. *Global Mountain Explorer* 2.0 (website). Data downloaded from *https://rmgsc.cr.usgs.gov/gme/* on Jan 13, 2023.

U.S. Geological Survey (USGS) Gap Analysis Project (GAP). 2022. *Protected Areas Database of the United States (PAD-US) 3.0*. U.S. Geological Survey data release. Data downloaded from *https://www.sciencebase.gov/catalog/item/622262c8d34ee0c6b38b6bcf* on November 8, 2022.

Urban Institute and Geolytics. 2003. *Census CD Neighborhood Change Database (NCDB): 1970-2000 Tract Data*. Somerville, NJ: Geolytics.

USA PRISM Climate Group. 2023. *United States Average Annual Total Precipitation, 1991-2020 (4km)*. Oregon State University. Data downloaded from *https://prism.oregonstate.edu/normals/* on March 13, 2023.

Wirth, Louis. 1938. Urbanism as a way of life. *American Journal of Sociology* 44, 1: 1-24.

Zhao, Zhenxiang, and Robert Kaestner. 2010. Effects of urban sprawl on obesity. *Journal of Health Economics* 29, 6: 779-87.