CAUSES OF SPRAWL: A PORTRAIT FROM SPACE*

MARCY BURCHFIELD
HENRY G. OVERMAN
DIEGO PUGA
MATTHEW A. TURNER

We study the extent to which U. S. urban development is sprawling and what determines differences in sprawl across space. Using remote-sensing data to track the evolution of land use on a grid of 8.7 billion 30 × 30 meter cells, we measure sprawl as the amount of undeveloped land surrounding an average urban dwelling. The extent of sprawl remained roughly unchanged between 1976 and 1992, although it varied dramatically across metropolitan areas. Ground water availability, temperate climate, rugged terrain, decentralized employment, early public transport infrastructure, uncertainty about metropolitan growth, and unincorporated land in the urban fringe all increase sprawl.

I. INTRODUCTION

In a recent survey by the Pew Center for Civic Journalism [2000], 18 percent of Americans said urban sprawl and land development were the most important issue facing their local community—the top response, tied with crime and violence. However, there was a key element of disagreement. Respondents to this survey were almost evenly split (40 versus 52 percent) between those wanting local government to limit further development to the infilling of already built-up areas and those wanting local government also to allow more scattered development in previously undeveloped areas. Despite this widespread interest,

* For helpful comments and suggestions we thank three anonymous referees, William Fischel, Masahisa Fujita, John Hartwick, Vernon Henderson, John Landis, William Strange, and, in particular, Edward Glaeser. We also received helpful comments from seminar participants at the University of California Berkeley, Harvard University, Universitat Pompeu Fabra and Stanford University, and at conferences organized by the Regional Science Association International, the Centre for Economic Policy Research, the Association of Environmental and Resource Economists, and the Lincoln Institute. We are very grateful to Ferko Csillag for his advice on remote-sensing data. Also to the U. S. Geological Survey, and in particular to Stephen Howard and James Vogelmann, for early access to preliminary versions of the 1992 data. Vernon Henderson and Jordan Rappaport kindly provided us with data on metropolitan population 1920–1990, Matthew Kahn with data on employment decentralization, and Jacob Vigdor with data on streetcar usage. Kent Todd helped with ARC Macro Language scripts. While working on this project, Burchfield was an Master's student at the Department of Geography, University of Toronto. Funding from the Social Sciences and Humanities Research Council of Canada (Puga and Turner) and from the Centre de Recerca en Economia Internacional and the Centre de Referència d'Economia Analítica (Puga), as well as the support of the Canadian Institute for Advanced Research (Puga) and the National Fellows program at the Hoover Institution (Turner) are gratefully acknowledged.

The Quarterly Journal of Economics, May 2006

spatial development patterns are the dimension of sprawl that we know the least about. We have some understanding of what determines urban growth (see, e.g., Glaeser, Scheinkman, and Shleifer [1995], Overman and Ioannides [2001], and Black and Henderson [2003]) and the decentralization of economic activity within cities [Glaeser and Kahn 2004]. However, we know almost nothing about the extent to which development is scattered or compact, how this varies across space or what determines that variation. This paper is concerned with this key aspect of sprawl.

Existing data sets are not well-suited for studying the scatteredness of development. To improve our understanding, we construct a new data set by merging high-altitude photographs from around 1976 with satellite images from 1992. From these data, for units that are square cells of 30×30 meters, we know whether land was developed or not around 1976 and in 1992, as well as details about the type of developed or undeveloped land. Our data set consists of 8.7 billion such 30×30 meter cells for a grid covering the entire conterminous United States.

Using these data, we provide basic facts about the extent of urban land development. Our main focus, however, is on the spatial patterns of residential land development—in particular, whether residential development is sprawling or compact. This involves capturing the extent to which residential development in urban areas is scattered across otherwise undeveloped land. In sprawling areas, much of the land immediately surrounding the average house will not itself be developed. Conversely, in areas where development is compact, there will be a high proportion of developed land in the immediate vicinity of the average house. To measure this, for each 30×30 meter cell of residential development, we calculate the percentage of undeveloped land in the immediate square kilometer. Averaging this measure across all *developed* cells in a metropolitan area gives us an index of sprawl for the metropolitan area: the percentage of open space in the square kilometer surrounding an *average* residential development. We calculate this index for all metropolitan areas and then examine the reasons why sprawl differs across space.

Regarding overall development, we find that only 1.9 percent of the United States was built-up or paved by 1992. Two-thirds of this was already in urban use by 1976, while the remaining one-third was developed subsequently. Turning to spatial patterns, only 0.3 percent of 1992 residential development is more than one kilometer away from other residential development. On

the other hand, at a finer spatial scale, our measure of sprawl shows that 43 percent of the square kilometer surrounding an average residential development is undeveloped. Thus, while residential development almost never leapfrogs over large extensions of undeveloped land, it is also not particularly compact on average. Moreover, contrary to widespread claims, the extent to which average residential development is scattered was essentially unchanged between 1976 and 1992. That is, while we have seen an increase in the amount of residential development, that development was not any more biased toward sprawling areas in 1992 than in 1976. The same is not true of commercial development: this appears to have become considerably more biased toward sprawling areas in the time period under study. While spatial patterns of residential development did not vary much between 1976 and 1992, our sprawl index indicates that there are dramatic variations across metropolitan areas. Much of this paper is devoted to describing this variation and to investigating the ability of the various theories of urban economics to explain it.

We start with the monocentric city model and its generalizations. Consistent with these theories, factors that increase the importance of the central business district decrease sprawl. Thus, cities sprawl less if they specialize in sectors, such as business services, that tend to be centralized in the average city. The commute to the city center also plays a role, with cities built around public transportation more compact than cities built around the automobile. Patterns of past growth in the metropolitan area also affect sprawl. Cities with higher historical population growth rates sprawl less. Among other things, in fast growing cities small undeveloped plots do not stay undeveloped for long. Greater historical uncertainty about growth also causes more sprawl as developers withhold land to better adapt it to future needs.

We next consider physical geography. Despite technological progress, the physical environment continues to play an important role in shaping cities. Sprawl increases substantially with the presence of water-yielding aquifers in the urban fringe: such aquifers allow people to sink a well and locate far from other development without bearing the large costs of extending municipal water lines. Regarding physical barriers to development, high mountains close to development constrain urban expansion and tend to make development more compact. Hills and small-scale terrain irregularities, on the other hand, encourage scattered

development. Finally, factors that increase the value of open space, a temperate climate in particular, increase sprawl. In all, physical geography alone explains about 25 percent of the cross-city variation in our sprawl measure.

We turn finally to political determinants of sprawl. There is more sprawl in cities where a large proportion of undeveloped land lay outside of any municipality. In contrast, municipal fragmentation has no effect, suggesting that developers are often leapfrogging out of municipal zoning and building regulations altogether, rather than playing municipalities against each other. Public finance also plays an important role. There is more sprawl in places where larger intergovernmental transfers mean that local residents bear less of the cost of extending infrastructure to service new scattered development.

II. Data and Methodology

We construct our core data from two fine-resolution data sets describing land cover and land use (i.e., the physical features that cover the land and what those features are used for) across the conterminous United States for the mid-1970s and the early 1990s. The most recent data set, the 1992 National Land Cover Data [Vogelmann, Howard, Yang, Larson, Wylie, and Driel 2001] classifies the land area in 1992 into different land cover categories mainly on the basis of Landsat 5 Thematic Mapper satellite imagery. The earlier data set, the Land Use and Land Cover Digital Data [U. S. Geological Survey 1990; U. S. Environmental Protection Agency 1994], derives mainly from high-altitude aerial photographs taken circa 1976.[1]

Despite the different technologies used to construct the two data sets, the processes are fundamentally similar. For the 1992 data, first-pass boundaries of contiguous areas with similar land cover are generated by grouping together contiguous cells with similar vectors of reflectance values recorded by satellite imagery. Aerial photographs and ancillary data are then used to refine these boundaries and to assign land cover codes. For the 1976 data the initial boundaries are drawn directly on the basis of the aerial photographs, and then these photographs and ancillary data are used to assign land cover codes. While the 1970s data

1. These photographs were collected over the period 1971–1982, but the most common date is 1976, which is also the median year.

have been available for over a decade, the 1990s data only became available in 2001 and are the most current land use data available for the nation. The Data Appendix describes in more detail the process followed by the U. S. Geological Survey (USGS) and the U. S. Environmental Protection Agency (EPA) to construct each of the data sets, as well as the way in which we have completed and integrated them. Our resulting data set has units of observation which are square cells of 30 $\times$ 30 meters situated on a regular grid. For each of the approximately 8.7 billion cells that make up the conterminous United States, we know the predominant land cover and land use circa 1976 and in 1992. Land is categorized as residential development; commercial and industrial development and transportation networks; water; bare rock and sand; forest; range and grassland; agricultural land; or wetlands.

Figure I presents a map of the United States derived from our data. This map shows, in yellow, the stock of land that was already built up circa 1976, and in red, new urban land built between circa 1976 and 1992. Land that remained nonurban in 1992 appears gray with shaded relief, and water is marked blue. This map reveals a number of noteworthy aggregate features. Perhaps the most striking is that the United States is overwhelmingly unoccupied. In fact, our data show that only 1.9 percent of the land area was either built up or paved by 1992. Two-thirds of this developed land was already in urban use around 1976, one-third was developed subsequently. Developed area grew at a very high rate (2.5 percent annually, 48 percent over sixteen years), but new development absorbed only a very small proportion of undeveloped land (0.6 percent over sixteen years).[2]

Our estimate that only 1.9 percent of the United States was developed by 1992 is slightly lower than previous estimates. Typically, these estimates use the partition of the territory into "urban" and "rural" made by the U. S. Census Bureau for administrative purposes. In the 1990 census, 2.5 percent of the conterminous United States was classified as urban. Using this figure systematically overstates the extent of built-up land in population centers by counting the entire area as developed when it need not be. At the same time it ignores development housing the

2. Our data also allow us to look at the development rate of different types of undeveloped land. There is no large bias toward the development of any particular type of land.

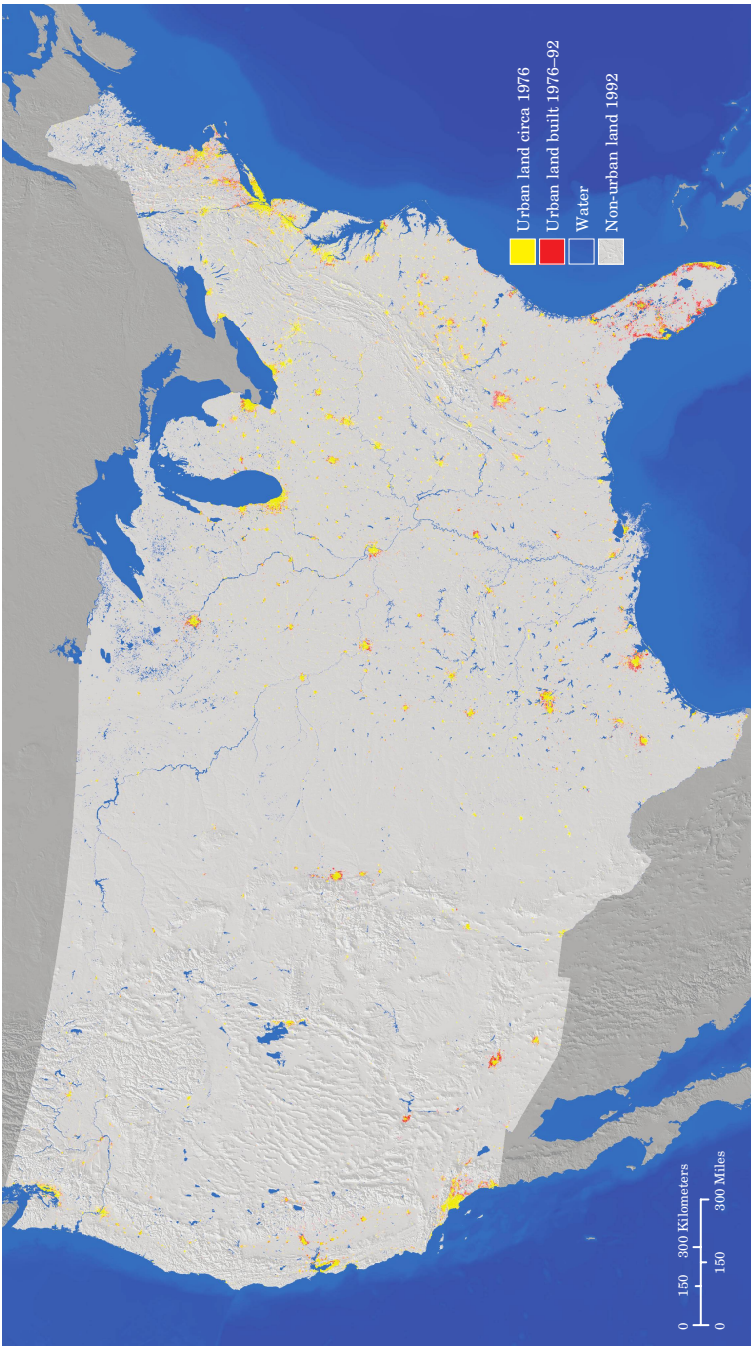

FIGURE I
Urban Land in the Conterminous United States

one-quarter of the population that was classified as rural in 1990. Some recent studies (e.g., Fulton, Pendall, Nguyen, and Harrison [2001]) estimate built-up land using National Resource Inventory (NRI) data, assembled by the U. S. Department of Agriculture on the basis of remote-sensing data for a relatively small sample of U. S. nonfederal land. According to these data, 2.9 percent of the United States was urban or built-up by 1992. The main reason why this estimate is larger than our figure is that, in the NRI data, the boundaries of urban and built-up areas are drawn in such a way that they incorporate substantial amounts of undeveloped land. In particular, all undeveloped land located between buildings or roads that are up to 500 feet (152 meters) apart is classified as built-up [U. S. Department of Agriculture 1997]—contrast this with our 30-meter resolution. However, the main advantage of our data is that it allows us to measure the scatteredness of development, the key concern of this paper. In contrast, neither census urban/rural boundaries nor the NRI allows this; in fact, the NRI is not available at the substate level since "[data at the county level] do not meet NRI reliability standards because of the small sample sizes for geographic units of that size" [U. S. Department of Agriculture 2001, p. 21].[3]

While our data show that only 1.9 percent of all land was developed by 1992, this aggregate number masks large differences across states. Data for individual states are reported in Table I. The first two columns show the percentage of all land in each state that was urban by 1992 and by 1976. The third column reports the percentage of 1976 nonurban land converted to urban between 1976 and 1992. The last three columns report the percentages accounted for by each state of U. S. urban land in 1992, of U. S. land area, and of U. S. 1976 nonurban land converted to urban between 1976 and 1992.[4] One particularly interesting aspect of this heterogeneity is that coastal states both had high initial percentages of urban land and also experienced relatively fast growth. More detailed analysis shows that land within 80

3. A number of other papers use detailed geographical data similar to our own (e.g., Mieszkowski and Smith [1991], Rosenthal and Helsley [1994], Geoghegan, Wainger, and Bockstael [1997], Geoghegan [2002], and Irwin and Bockstael [2002]), but each focuses on a particular city or small area.

4. We correct for photographs not taken in 1976 by first determining the portions of each county photographed in any given year, then estimating the percentage of urban land in each of these county portions by assuming a constant local annual growth rate over the period, and finally aggregating up to the state and national levels.

TABLE I
THE EXTENT OF LAND DEVELOPMENT

State	% of state's land area urbanized by 1992	% of state's land area urbanized by 1976	% of state's 1976 nonurban land urbanized 1976–1992	State's % of U. S. urban land 1992	State's % of U. S. land area 1992	State's % of U. S. new urban land built 1976–1992
Alabama	1.39	1.02	0.37	1.24	1.71	1.01
Arizona	0.79	0.44	0.35	1.58	3.82	2.14
Arkansas	1.25	0.72	0.54	1.15	1.76	1.51
California	2.85	2.14	0.73	7.81	5.25	6.01
Colorado	0.89	0.49	0.40	1.62	3.50	2.25
Connecticut	16.30	9.89	7.12	1.38	0.16	1.67
Delaware	7.18	5.94	1.32	0.25	0.07	0.13
DC	68.13	67.21	2.80	0.07	0.00	0.00
Florida	8.93	4.45	4.68	8.52	1.83	13.10
Georgia	2.52	1.59	0.94	2.58	1.96	2.92
Idaho	0.39	0.19	0.20	0.57	2.80	0.89
Illinois	3.70	2.87	0.86	3.61	1.87	2.50
Indiana	3.38	2.60	0.80	2.14	1.21	1.52
Iowa	2.49	0.95	1.56	2.40	1.85	4.57
Kansas	0.98	0.67	0.31	1.41	2.76	1.37
Kentucky	1.84	1.33	0.52	1.29	1.35	1.11
Louisiana	2.62	1.80	0.84	1.99	1.46	1.92
Maine	1.41	0.96	0.45	0.76	1.04	0.74
Maryland	7.82	6.72	1.18	1.36	0.33	0.58
Massachusetts	17.34	12.35	5.70	2.42	0.27	2.13
Michigan	3.20	2.39	0.84	3.22	1.93	2.52
Minnesota	1.62	0.82	0.80	2.27	2.69	3.44
Mississippi	1.24	0.79	0.45	1.03	1.59	1.13
Missouri	1.95	1.24	0.72	2.38	2.33	2.65
Montana	0.21	0.10	0.11	0.55	4.92	0.90
Nebraska	0.60	0.41	0.20	0.82	2.61	0.82
Nevada	0.34	0.13	0.21	0.66	3.71	1.23
New Hampshire	4.56	2.49	2.12	0.64	0.27	0.89
New Jersey	20.57	17.78	3.39	2.54	0.24	1.06
New Mexico	0.34	0.22	0.12	0.72	4.06	0.75
New York	5.48	4.59	0.93	4.50	1.57	2.24
North Carolina	4.19	2.97	1.26	3.60	1.65	3.21
North Dakota	0.46	0.16	0.30	0.56	2.34	1.12
Ohio	5.27	4.41	0.90	3.80	1.38	1.90
Oklahoma	1.51	1.10	0.41	1.83	2.33	1.53
Oregon	0.76	0.53	0.23	1.27	3.21	1.17
Pennsylvania	4.13	3.37	0.79	3.23	1.50	1.83
Rhode Island	17.99	14.11	4.52	0.34	0.04	0.23
South Carolina	3.43	2.30	1.15	1.83	1.02	1.84
South Dakota	0.37	0.16	0.22	0.50	2.56	0.90
Tennessee	2.76	1.99	0.78	2.09	1.45	1.79
Texas	1.83	1.10	0.73	8.44	8.85	10.28
Utah	0.54	0.34	0.20	0.73	2.58	0.82
Vermont	2.91	1.47	1.46	0.30	0.20	0.45
Virginia	3.48	2.77	0.73	2.43	1.34	1.52
Washington	2.23	1.44	0.79	2.62	2.25	2.82
West Virginia	1.32	0.97	0.35	0.56	0.81	0.45
Wisconsin	1.84	1.28	0.57	1.77	1.84	1.65
Wyoming	0.21	0.11	0.09	0.36	3.28	0.49
United States	1.92	1.29	0.63	100.00	100.00	100.00

kilometers of the ocean or Great Lakes accounts for only 13.4 percent of the total land area, but contained 45.6 percent of developed land in 1976. This share declined slightly to 44.2 percent in 1992, but coastal areas still accounted for 41.3 percent of 1976–1992 urban development. Interestingly, the evolution of the coastal concentration was quite different for residential and commercial development. While the share of residential land within 80 kilometers of the coasts increased from 46.6 percent to 48.5 percent, the share of commercial land fell from 43.2 percent to 34.3 percent. The shift of commercial development away from coast is consistent with Holmes and Stevens' [2004] findings on changes in the location of large U. S. manufacturing plants. This decline in the coastal concentration of commercial land together with the rise in that of residential land can also be seen as supporting the argument made by Rappaport and Sachs [2003] that amenity considerations are increasingly important relative to production considerations in driving coastal concentration.

Zooming in, Figures IIa and IIb depict development for four areas: Atlanta (top of Figure IIa), Boston (bottom of Figure IIa), around San Francisco (top of Figure IIb), and around Miami (bottom of Figure IIb). As before, urban land circa 1976 is marked in yellow and 1976–1992 urban development in red, but nonurban land is now split according to its 1992 cover. These maps reveal some of the complex spatial details of the land development process. Atlanta, the epitome of sprawl, experienced an extraordinary amount of development from the mid-1970s and both recent and older development are very scattered. Boston had less recent development and contains a much more compact old urban core. However, the suburban development that took place since the mid-1970s is, by some measures, even more scattered than in Atlanta. Development in San Francisco and neighboring metropolitan areas is much more compact than in either Atlanta or Boston, although looking closely at the map one can see green speckles marking the presence of parks within the yellow-colored old development. New development respected these urban parks but remained contiguous to earlier development, as evidenced by the red on the fringe of pre-1970s development. Miami, like most of Florida, experienced spectacular growth in the amount of developed land, but unlike Atlanta, this recent development either infilled portions of undeveloped land within earlier development (notice there are fewer urban parks than in San Francisco) or took place contiguously with previously built-up areas. Figure III

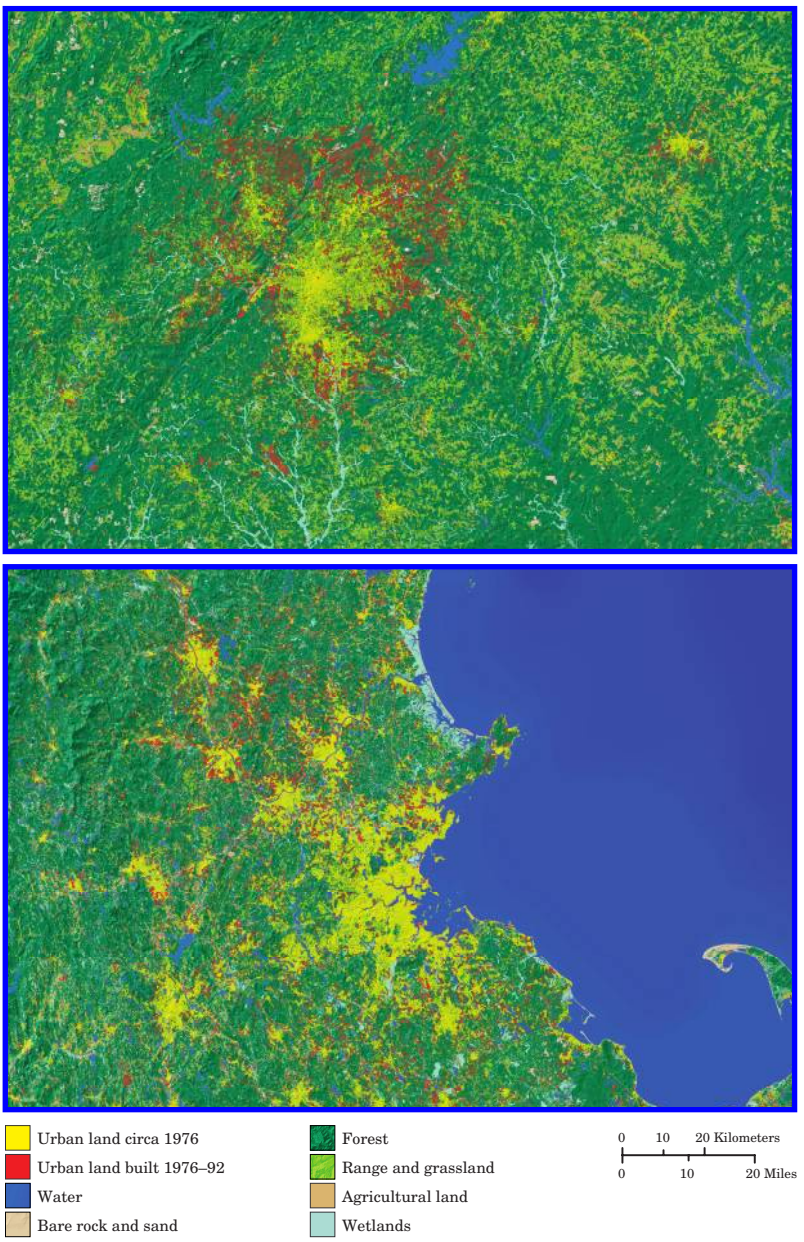

FIGURE IIa
Urban Land in Atlanta, GA (Top Panel) and Boston, MA (Bottom Panel)

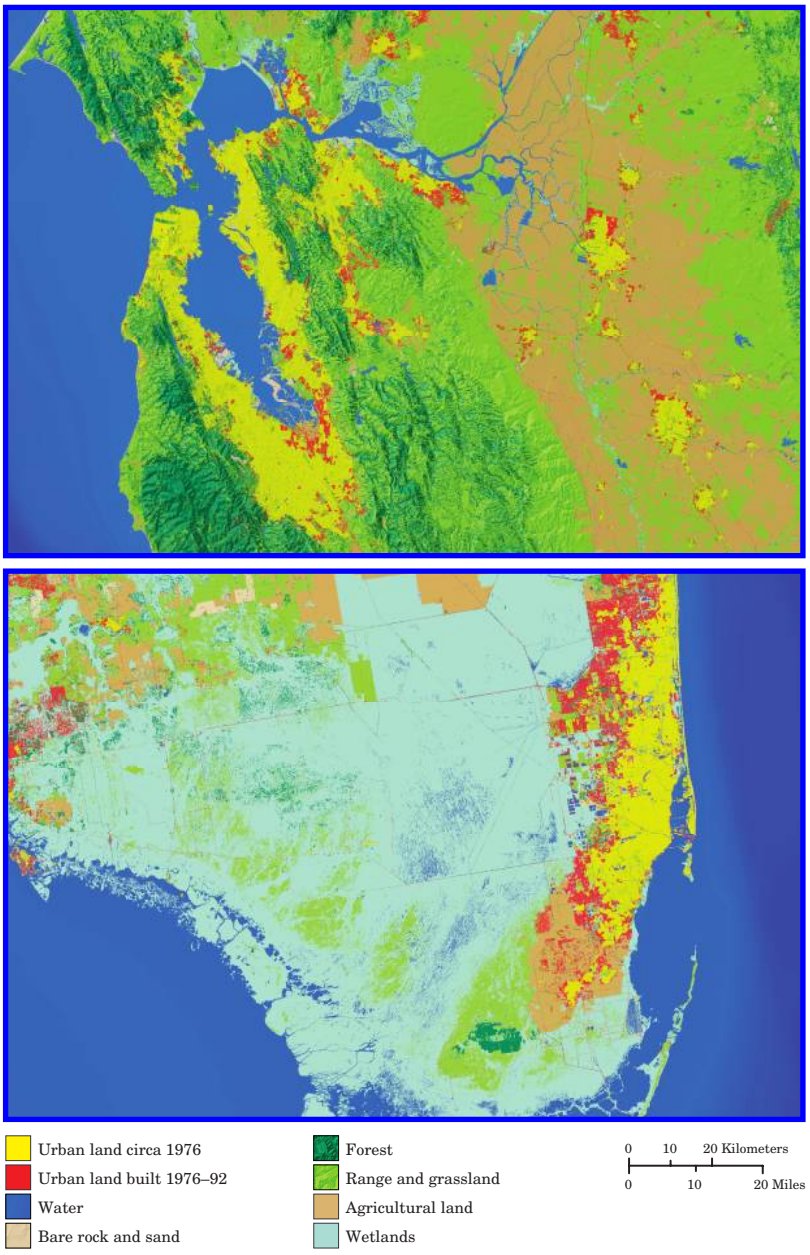

FIGURE IIb
Urban Land in San Francisco, CA (Top Panel) and Miami, FL (Bottom Panel)

FIGURE III
Urban Land and Aquifers in San Antonio and Austin, TX (Top Panel), and Urban Land and Incorporated Places in Saint Louis, MO (Bottom Panel)

presents two additional maps, depicting development in the area encompassing the Austin-San Marcos and San Antonio metropolitan areas (top panel, ignore for now the location of aquifers discussed later), and in the Saint Louis metropolitan area (bottom panel, drawn at a different scale to show details of the location of development relative to municipal boundaries, also discussed below). In terms of urban sprawl, these two areas are somewhere in between the scatteredness of Atlanta and Boston and the compactness of San Francisco and Miami.

To summarize such differences in the extent to which development is scattered or compact, we develop a measure of sprawl. The first step is to find a relevant spatial scale at which to conduct our analysis. To this end, we start by checking how often residential development leapfrogs over more than one kilometer of undeveloped land. It almost never does: only 0.3 percent of all residential development was more than one kilometer away from other residential development in 1992. Even for recent (1976–1992) development, the figure was only 0.5 percent.[5] Thus, if large amounts of development are scattered rather than compact, this is happening at spatial scales less than one kilometer. This means we need to exploit the full spatial resolution of our data and look within the immediate kilometer of development.

We proceed as follows. To measure the extent of sprawl, for each 30-meter cell of residential development, we calculate the percentage of open space in the immediate square kilometer.[6] We then average across all residential development in each metropolitan area to compute an index of sprawl. For instance, to calculate a sprawl index for the new development that took place between 1976 and 1992 in each metropolitan area, we identify 30-meter cells that were not developed in 1976 but were subject to residential development between 1976 and 1992, calculate the percentage of land not developed by 1992 in the square kilometer containing each of these 30-meter cells, and average across all such newly developed cells in the metropolitan area. We also

5. This tiny amount of long-distance leapfrogging has, however, significantly reduced peoples' ability to "get away from it all." The percentage of U. S. land more than five kilometers way from any residential development dropped from 58.1 percent in 1976 to 47 percent in 1992.

6. For computational reasons, rather than looking at the square kilometer centered on each 30-meter cell, we construct a grid made up of square blocks of 30-meter cells each measuring approximately one square kilometer (990 $\times$ 990 meter squares so that each one contains an integer number, 1089, of our underlying 30-meter cells). The percentage of open space is then calculated for the one-kilometer cell block in which each 30-meter cell is located.

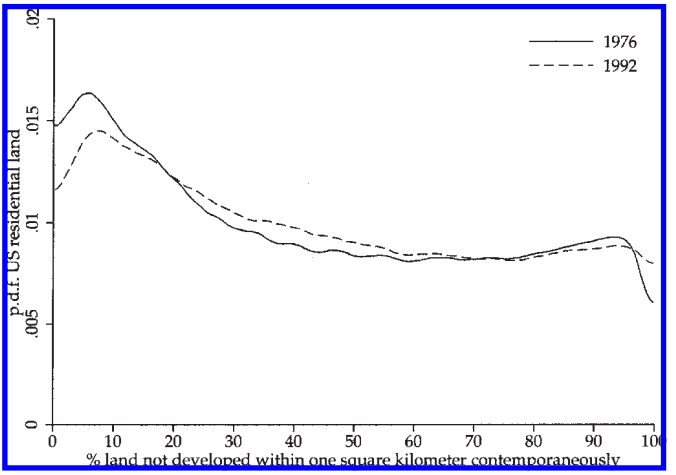

FIGURE IV
Probability Function of 1976 and 1992 U. S. Residential Land across Areas with Different Degrees of Sprawl

perform similar calculations to calculate a sprawl index for the stock of development in 1976 and in 1992. This provides a very intuitive index of sprawl: *the percentage of undeveloped land in the square kilometer surrounding an average residential development.*[7]

III. SPRAWL ACROSS THE UNITED STATES

We start by examining the spatial structure of urban development for the United States as a whole. Figure IV plots the probability density function showing the distribution of residen-

7. One could imagine configurations of development for a particular square kilometer with a large percentage of open space that we might not want to characterize as sprawl. However, in metropolitan areas these are rare enough that they do not drive our sprawl index. In particular, since we average across all residential development rather than across all land in the metropolitan area, a square kilometer that is average for the nation at 43 percent undeveloped is counted 620 times (57 percent of 1089 cells) when averaging across the metropolitan area. On the other hand, a square kilometer with just one isolated developed cell is only counted once. Thus, the index is not driven by rare instances of isolated houses but by the groups of houses with an intermediate mixture of developed and undeveloped land surrounding them. For computational reasons, it is too difficult to work with buffers of less than one kilometer around houses. We have, however, tried other summary statistics, such as the percentage of undeveloped land in the square kilometer surrounding the median (instead of the average) residential development, and found almost identical results.

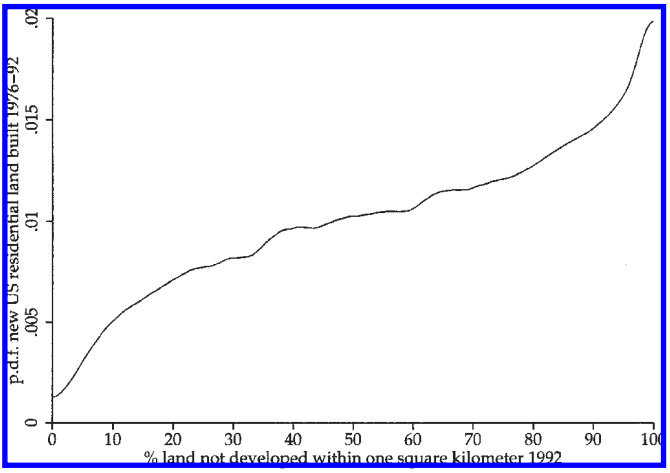

FIGURE V
Probability Function of 1976–1992 U. S. Residential Development across Areas with Different Degrees of Sprawl

tial development across areas with different degrees of sprawl. Consider the distribution for 1976, given by the solid line. The area under the line between any two values in the horizontal axis is the probability that a randomly picked 30-meter cell classified as residential development in 1976 lies in a square kilometer where the percentage of land not developed in 1976 lies between those two values. The figure shows that residential development was almost uniformly distributed across areas where between one-third and all of the surrounding square kilometer was not developed, but overall residential development was skewed toward more compact areas.

Figure V shows that this is not the case for new development that occurred between 1976 and 1992. The figure plots the probability density function for this new development across areas with different degrees of sprawl at the end of the period.[8] The figure shows that, in contrast to the stock of residential development in 1976, the flow of new residential development between

8. Note that is the amount of *final* development near new development that distinguishes sprawling from compact areas. The easiest way to see this is to consider a city that grows in a completely contiguous way. All new development occurs in areas that are initially almost entirely undeveloped but end up being completely developed.

1976 and 1992 was biased toward sprawling areas. Thus, new development does tend to be scattered at small spatial scales.

We suspect that it is some perception that the flow of new development is more scattered than the initial stock that often leads people to conclude that development is more sprawled than in the past. However, looking back at Figure IV, we see that, for the dimension of sprawl that is our focus, this is not the case. The dashed line showing the distribution of the stock of 1992 residential development across areas with different degrees of sprawl is almost identical to the solid line for the 1976 stock. In fact, on average, 42 percent of the land in the square kilometer surrounding residential development was open space circa 1976. Remarkably, this figure remained almost unchanged at 43 percent in 1992. Thus, while a substantial amount of scattered residential development was built between 1976 and 1992, overall residential development did not become any more biased toward such sprawling areas.

To reconcile these apparently conflicting tendencies, note that the distribution of the final stock of development across different degrees of sprawl is *not* the result of adding the distribution of the flow of new development to the distribution of the initial stock. The reason is that, by adding the flow of new development to the initial stock, the distribution of the initial stock becomes shifted to the left as infilling makes formerly sprawling areas more compact. Figure VI further illustrates the importance of this infilling of areas that were partially developed to start with. It plots the average intensity of 1976–1992 residential development (i.e., the percentage of nonurban land turned residential) in areas with different percentages of open space in the immediate square kilometer in 1976. The figure shows that it is areas that were about half undeveloped in 1976 that were subject to the most intense subsequent residential development.

Pulling all this together, what do we learn about recent residential development and common perceptions of sprawl? It helps to consider how the environment might have changed near a hypothetical house located in a medium-density suburb. The open space in the immediate neighborhood of this house will most likely have been partly infilled. Areas initially more compact, presumably closer to downtown, will have experienced less change. Undeveloped areas farther out may now be scattered with low density development. To the family living in this house,

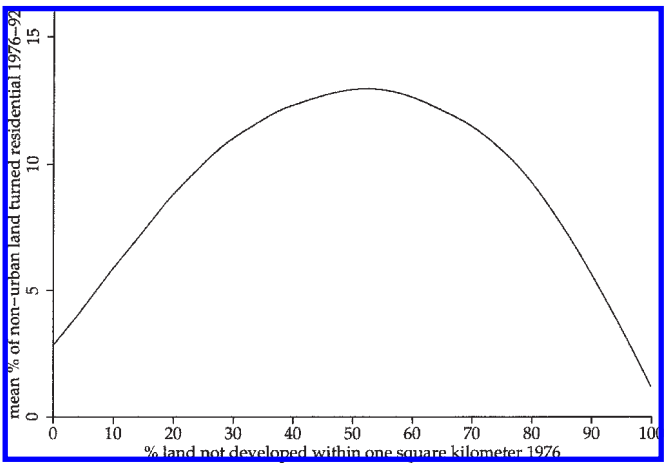

FIGURE VI
Mean Percentage of Nonurban Land Turned Residential 1976–1992 by Initial Percentage of Nonurban Land within One Square Kilometer

the pattern of residential development around them is very different from the one they experienced in the 1970s. However, if we zoom out and look at the city from a distance, we see little change, at least in the proportions of sprawling and compact development: the new city is just like an enlarged version of the old city.

While our focus is on residential sprawl, it is of interest to compare the distribution of residential land with that of commercial land. As it turns out, while the sort of places where Americans live has not changed substantially, the places where they shop and work has. Figure VII is a counterpart to Figure IV giving the distribution of commercial land (including industrial land and transportation networks) across areas with different percentages of developed land nearby. Looking first at the solid line, we see that the distribution for the stock of commercial land in 1976 is clearly bimodal. Commercial development in the 1970s was biased toward areas that were either very compact or very sprawling. Presumably the very compact commercial development is office buildings located downtown, while the scattered development is factories and malls located on the outskirts of town.

Turning to the dashed line in Figure VII, we see that, unlike residential land, commercial land has become more biased over

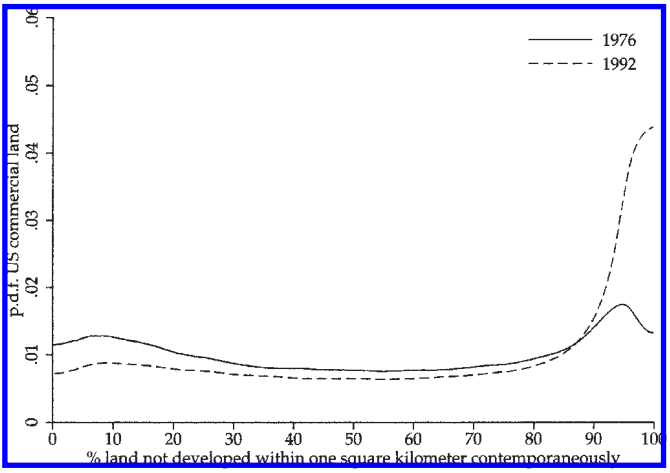

FIGURE VII
Probability Function of 1976 and 1992 U. S. Commercial Land across Areas with Different Degrees of Sprawl

time toward areas with little nearby development.[9] This finding is consistent with the view that the decentralization of housing in the United States had already reached its peak by the mid-1970s, whereas it is only more recently that employment, and especially manufacturing employment, has shifted from city centers to suburbia [Glaeser and Kahn 2001; Holmes and Stevens 2004].

III.A. Sprawling and Compact Cities

Earlier in this section we showed that the distribution of U. S. residential development across areas with different degrees of sprawl remained almost unchanged between 1976 and 1992. Analogous distributions for individual metropolitan areas also show small differences across time for most areas but very large differences across areas at either point in time. We can summarize these differences using our sprawl index, the percentage of

9. It is worth noting that the proportions of residential and commercial land in overall U. S. urban land remained unchanged between 1976 and 1992 at about 70 percent and 30 percent, respectively. Hence, our findings do not reflect changes in the relative magnitudes of residential and commercial development but rather changes in their locations. We note that commercial land includes roads and industrial land, so this result may partly reflect new (or misdated) roads in rural areas as well as newly constructed factories and shopping malls. More detail on this issue is available in the Data Appendix.

TABLE II
SPRAWL INDICES FOR METROPOLITAN AREAS WITH POPULATION OVER ONE MILLION

Metropolitan area	Sprawl index for 1992 residential land	Sprawl index for 1976 residential land	Metropolitan area	Sprawl index for 1992 residential land	Sprawl index for 1976 residential land
Atlanta	55.57	57.77	Minneapolis-St. Paul	32.07	31.34
Boston	47.64	44.72	New Haven	39.11	38.68
Buffalo	39.92	37.87	New Orleans	32.29	33.92
Charlotte	52.73	51.12	New York	28.75	28.47
Chicago	31.76	31.21	Norfolk	40.82	44.07
Cincinnati	47.79	47.45	Orlando	40.02	39.39
Cleveland	36.84	36.24	Philadelphia	42.51	43.03
Columbus	41.20	41.59	Phoenix	27.54	34.94
Dallas	28.08	26.65	Pittsburgh	57.70	56.71
Denver	28.63	28.63	Portland	44.90	43.38
Detroit	33.28	30.47	Rochester	48.80	48.11
Greensboro	52.94	51.45	Sacramento	34.93	30.72
Hartford	41.34	42.23	Salt Lake City	31.90	32.88
Houston	38.15	38.93	San Antonio	32.77	29.58
Indianapolis	39.66	37.68	San Diego	45.63	45.40
Kansas City	35.32	34.33	San Francisco	30.48	29.81
Los Angeles	35.41	32.95	Seattle	46.97	45.03
Memphis	27.40	28.72	St. Louis	43.44	40.62
Miami	20.73	20.03	Tampa	36.01	34.84
Milwaukee	35.33	33.85	Washington-Baltimore	49.81	50.68

Each sprawl index measures the percentage of undeveloped land in the square kilometer surrounding an average residential development in each metropolitan area in the corresponding year (1992 or 1976). For instance, the sprawl index for 1992 residential land is computed by calculating the percentage of land not developed by 1992 in the square kilometer containing each 30-meter cell classified as residential land in 1992 and averaging this percentage across all cells classified as residential land in 1992 in the metropolitan area.

undeveloped land in the square kilometer surrounding an average residential development, which corresponds to the mean of these distributions. For each MSA with 1990 population greater than one million, Table II lists the sprawl index calculated for final (1992) and initial (1976) residential development.

A comparison of the figures provided in Table II with the maps presented in Figures IIa and IIb shows that the sprawl indices provide a good summary of development patterns. Overall development in Boston is substantially less scattered than in Atlanta, reflecting its much more compact old urban core (47.64 percent open space in the square kilometer around the average residential development in Boston in 1992, compared with 55–57 percent in Atlanta). However, the scatteredness of recent suburban development in Boston has made this metropolitan area somewhat less compact on average than it used to be: the per-

centage of undeveloped land around the average residential development increased from 44.72 to 47.64 between 1976 and 1992. Development in San Francisco is much more compact: only 30.48 percent of the square kilometer around the average residential development in San Francisco was not itself developed in 1992. Miami is even more compact than San Francisco, reflecting the greater presence of concrete and asphalt as opposed to small parks in between residential buildings: there was a mere 20.73 percent undeveloped land in the square kilometer around the average residential development in Miami in 1992.

Table II shows that, even at the metropolitan area level, the extent of sprawl is very stable over time. However, Table II also reveals the spatial heterogeneity of development patterns that is suggested by Figure I. The square kilometer around the average residential building in Atlanta or Pittsburgh is nearly 60 percent open space. In Miami this number is just over 20 percent. Before we turn to explaining this extraordinary variation across metropolitan areas, it is worth having a brief look at how our index compares with measures looking at alternative dimensions of sprawl.

III.B. Correlation with Other Measures of Sprawl

Given that, until now, data to directly measure the scatteredness of development have been unavailable, median lot size has often been used as a proxy for metropolitan areas where this is known. We would expect places with a large median lot size also to have relatively scattered development as measured by our sprawl index, since residential developments built on larger lots tend to have a higher ratio of open space to built-up area. However, our index also captures the presence and size of undeveloped land in between built-up lots. Table III shows the correlation between our index for 1992 residential development and several other measures of sprawl for metropolitan areas with more than one million inhabitants in 1990.[10] The correlation between our index for 1992 residential development and median lot size in 1994–1998 is 0.52.

The scatteredness or compactness of residential development, while an important dimension of sprawl, is not the only

10. For most smaller metropolitan areas data are not available for these alternative indices. Correlations reported in Table III are for the largest 40 metropolitan areas listed in Table II (with the exception of median lot size which is based on 38 out of 40). See the Data Appendix for further details.

TABLE III
CORRELATION MATRIX FOR VARIOUS METROPOLITAN AREA SPRAWL MEASURES IN THE 1990S

	Sprawl (scatteredness) index	Median lot size	Miles driven per person	% Employment over 3 miles from CBD
Sprawl (scatteredness) index	1.000			
Median lot size	0.521	1.000		
Miles driven per person	0.271	0.187	1.000	
% Employment over 3 miles from CBD	−0.070	0.011	−0.073	1.000

The sprawl (scatteredness) index is the measure of sprawl used throughout this paper: the percentage of undeveloped land in the square kilometer surrounding an average residential development in each metropolitan area in 1992. Median lot size compiled from the metropolitan data contained in the American Housing Survey 1994–1998. Average number of miles driven per person calculated from the 1995 Nationwide Personal Transportation Survey [U. S. Federal Highway Administration 1995]. Share of employment located more than three miles away from the central business district in 1996 from Glaeser and Kahn [2001].

one. "Sprawl" is also used to describe cities where people need to drive large distances to conduct their daily lives, or cities where employment is very decentralized [Glaeser and Kahn 2004]. People in more scattered metropolitan areas do tend to drive longer distances: the correlation between our index and the average miles driven per person is 0.27. However, there is almost no correlation between the extent to which residential development is scattered and that to which employment is decentralized (measured by the share of employment located more than three miles away from the central business district, as calculated by Glaeser and Kahn [2001]). This low correlation between measures of employment decentralization and other measures of sprawl is also noted by Glaeser and Kahn [2004].

These correlations are of interest for three reasons. First, because they highlight the complexity of spatial patterns of development. Second, they indicate the importance of interpreting our results appropriately. We determine the factors that lead to sprawl in the sense of scattered development. There is no reason to think that these factors will also explain the other features of the spatial patterns of development, such as how much people drive or the extent to which employment is decentralized. Finally, the table points out the difficulty of interpreting composite sprawl indices (e.g., Ewing, Pendall, and Chen [2002]). Given the low correlations between measures of different aspects of sprawl,

when these are combined into a single number it is hard to know what is measured, let alone explain its determinants.

IV. Urban Economic Theory and the Causes of Sprawl

To investigate the determinants of sprawl, we turn to urban economic theory for guidance. Unfortunately, there is no unified model that tells us what determines the extent to which development is scattered or compact. Instead, the difficulties involved in working with general equilibrium models where space is explicitly modeled have led urban economists to develop many special models to tackle particular issues. A few of these models have been written specifically to study some aspect of sprawl. Most of them have, however, been written with a different purpose in mind, yet also have implications relevant for sprawl. In this section we survey this literature in order to formulate hypotheses about the causes of sprawl.

IV.A. The Monocentric City Model and Its Generalizations

The most widely used theoretical construct in urban economics is the monocentric city model, which deals with the determinants of variations in the intensity of residential urban development. This model derives from the pioneering contributions of Alonso [1964], Mills [1967], Muth [1969], and Wheaton [1974] (see Brueckner [1987] for an elegant synthesis). The monocentric city model assumes that all employment in the city takes place at a single center, the central business district. Residential development around that center is then shaped by the trade-off between convenient commuting close to the center and affordable housing farther away. Equal utility across residential locations implies that housing prices decline with distance to the city center to offset higher commuting costs. Equal profits for developers, who combine land and capital to produce housing, imply a similar gradient for land prices. Substitution in response to declining land and housing prices leads to larger dwellings with lower capital to land ratios (i.e., less tall, more spacious units and larger yards) as one moves away from the center.

The extent to which U. S. metropolitan areas can be characterized as monocentric has declined over time. The proportion of jobs located in central cities fell from about 75 percent in 1950 to about 45 percent in 1990 [Mieszkowski and Mills 1993], and metropolitan areas have become increasingly polycentric [Anas,

Arnott, and Small 1998]. Beginning with the contributions of Fujita and Ogawa [1982] and Imai [1982] and continuing more recently with Lucas and Rossi-Hansberg [2002], a number of papers have extended the monocentric city model to endogenously derive monocentric as well as polycentric urban structures. In these models, cities specializing in sectors with stronger agglomeration economies, due to externalities in the transmission of information, tend to be monocentric while those with weaker agglomeration economies are more likely to be polycentric (see Chapter 6 in Fujita and Thisse [2002]). In their study using ZIP-code employment data, Glaeser and Kahn [2001] show that the extent of employment decentralization does indeed vary widely both across cities and across sectors. In addition, sectors such as business services where communication is particularly important do tend to be more centralized. General equilibrium models of systems of cities building on Henderson [1974, 1987] also show that cities specializing in sectors with stronger agglomeration economies have more expensive land, which offsets the higher wages resulting from agglomeration economies. Substitution away from land then implies higher buildings with smaller units and yards, i.e., more compact development. Thus, a crucial implication of the monocentric city model is that *cities specializing in sectors where employment tends to be more centralized will be more compact.*

A second prediction arising from the monocentric city model is that *lower transport costs within a city will result in more dispersed development.* A greater ability to use the car for commuting not only reduces transport costs, but also eliminates the fixed costs associated with public transport [Glaeser and Kohlhase 2004]. Both these effects contribute to sprawl.

The standard monocentric city model thus predicts scattered development, due to large yards, in cities specialized in sectors where employment is less centralized and where it is easier to use a car. However, a key feature that the standard monocentric city model cannot explain is leapfrog development where parcels of land are left undeveloped while others farther away are built up. Urban economists have followed two strategies to extend the monocentric city model to account for equilibrium leapfrogging. The first is to assign an amenity value to public open space so that individuals may be willing to incur the additional commuting costs associated with locating farther away from the city center in order to have open space near their home. Scattered development

then takes the form of equilibrium leapfrogging, where remote areas are developed before central areas and residential development is mixed with undeveloped parcels [Turner 2005]. An immediate implication is that characteristics that make public open space more attractive will increase sprawl. While the same is true about private open space, there is one important regard in which public open space differs from private: the control that the residential owner has over subsequent development. If moving is costly, the willingness to trade off commuting costs against access to public open space will depend on expectations of how long that space will stay undeveloped. In areas where population is growing fast, a rational agent anticipates that nearby vacant land will be developed sooner, and thus is not willing to incur large additional commuting costs to gain access to it. Thus, *cities that have been growing faster will tend to experience less sprawl.*

The second strategy that urban economists have followed to account for equilibrium leapfrogging is to consider dynamic urban models where housing is durable and redevelopment costly. The core argument is that it may be optimal to postpone development of certain parcels so that in the future they can be developed in a way that better suits contemporaneous needs [Ohls and Pines 1975; Fujita 1976; Mills 1981] (see Fujita [1983] and Brueckner [2000] for reviews]. Uncertainty is particularly interesting in this context. In a model with uncertainty, Capozza and Helsley [1990] argue that developers should delay development until the value of the built-up land compensates for, not just the value of land in the best alternative use plus conversion costs (as in Arnott and Lewis [1979]), but also the option value of not developing in the face of uncertainty. Bar-Ilan and Strange [1996] extend Capozza and Helsley [1990] to allow for the fact that there are often long lags between the decision to build and the completion of construction. In this framework, uncertainty about urban growth translates into greater rent uncertainty the farther away a parcel is from the city center. In the presence of construction lags, an increase in uncertainty can encourage some landowners to choose earlier development. Thus, when leapfrogging occurs, *leapfrogging is greater the greater the uncertainty about future urban growth.*

IV.B. When Space Is Not a Featureless Plain

Urban economists typically explain the clustering of people on the basis of agglomeration economies. While there are many

microeconomic foundations for such economies (see Duranton and Puga [2004] and Rosenthal and Strange [2004] for reviews), perhaps the simplest is the existence of large indivisible public facilities [Buchanan 1965]. One example of particular interest is municipal water systems. What makes these shared water systems different from other public facilities is that in certain locations there is an alternative individual provision not subject to the same indivisibilities.

Most households in the United States get their water through the nearest municipal or county water supply. Extending water systems to service new scattered development in the urban fringe requires substantial infrastructure investments, the cost of which is typically borne by developers through connection fees and ultimately reflected in housing prices. For instance, to finance part of the $127 million cost of a twenty-mile pipeline to suburban development in Denver's South Metro area, the East Cherry Creek Valley Water and Sanitation District decided to charge $24,000 to connect new homes[11]—about one-seventh of the contemporaneous median house value in the Denver metropolitan area.

In places where water-yielding aquifers are pervasive, developers can sink a well instead of connecting to the municipal or county water supply. Fifteen percent of households in the United States get their water from private household wells [U. S. Environmental Protection Agency 1997]. According to the National Ground Water Association [wellowner.org], the average construction cost of a private well is approximately $4500. Private wells are rarely used in areas subject to compact development partly because in these areas they are often unsafe, or disallowed, by municipal regulations and partly because infill development in compact areas is typically subject to low water connection fees given the large number of connections per mile of pipe. However, low-cost private wells can facilitate scattered development in the urban fringe provided there is an aquifer from which to pump out water.

The top panel of Figure III illustrates the relationship between aquifers and sprawl with a map of San Antonio (located in the southwest of the map) and Austin (northeast), in Texas. Only part of these cities overlies an aquifer—the Edwards-Trinity

11. "Suburbs plan $127 million water system," *Rocky Mountain News,* December 19, 2003.

aquifer system—outlined and crosshatched in white. Households southeast of the "bad water line" plotted as a white dotted line cannot safely draw water from a well. The San Antonio Water System charges developers one-time connection fees per dwelling unit that range from $500 in some central areas to $24,000 in an eastern suburb. However, developers building in areas overlying the aquifer can sink a well and avoid the water connection fee or even build in areas where a connection to the municipal supply is not available. The map shows that most new development in San Antonio since the mid-1970s (marked in red) has taken place above the Edwards aquifer and that this development is much more scattered than that which does not overlay the aquifer. Austin shows a similar pattern.

Urban models typically treat space as a featureless plain to better focus on economic mechanisms, particularly agglomeration economies. The presence of aquifers is a particularly interesting dimension of underlying heterogeneity in the physical landscape precisely because of the way it interacts with agglomeration economies: *wherever aquifers underly the urban fringe, household water can be obtained without the large increasing returns associated with public water systems and this facilitates scattered development.* We now turn to other features of the physical landscape that are likely to matter for sprawl.

Nature can also contain sprawl through physical barriers hindering urban expansion. For instance, the mountains bordering Los Angeles are often mentioned as the main barrier to further expansion of its sprawling suburbs, and this has led to the coining of the phrase "sprawl hits the wall" [Southern California Studies Center and Brookings Institution 2001].

In studying the effect on sprawl of mountains located in the urban fringe, we need to be careful to separate large-scale from small-scale terrain irregularities. This is because one would expect mountains and hills to have opposite effects. When an expanding city hits a mountain range, further scattered development in the urban fringe becomes very costly. This encourages infilling and leads to increasingly compact residential patterns. On the other hand, small-scale irregularities in the urban fringe presumably have the opposite effect. When terrain in the urban fringe is rugged, steep hillsides where development is more costly alternate with flat portions where development is less costly. Thus, *we would expect rugged terrain to naturally encourage*

scattered development. In contrast, *high mountains in the urban fringe are likely to make development more compact.*

In our discussion of the monocentric city model and its extensions, we saw that characteristics that make open space more attractive are likely to encourage both larger yards and more frequent undeveloped parcels providing public open space. Thus, a third hypothesis related to the physical landscape is that *cities with a pleasant temperate climate experience more sprawl.*

IV.C. Political Geography

In his excellent book on the economics of zoning, Fischel [1985] devotes substantial attention to the political geography of zoning. There, he discusses a possible relationship between jurisdictional fragmentation and the restrictiveness of zoning: if a small number of municipalities dominate a metro area, they may exploit their monopoly power on behalf of incumbent residents to restrict the supply of land and increase property values. He concludes that such a relationship is unlikely to be of practical importance for three reasons. First, large jurisdictions also tend to internalize the pros as well as the cons of development (e.g., a large jurisdiction is more likely to house the construction workers building new residences as well as the neighbors trying to stop these). Second, there are few instances of areas with highly concentrated municipal structures in the United States. Third, legal and practical restrictions limit the ability of even very dominant jurisdictions to act as monopolists.

Rubinfeld [1978] and Katz and Rosen [1987], among others, stress differences between zoned and unzoned areas instead of competition between zoned areas of different sizes. These differences are illustrated in Figure III. This map of municipal boundaries in Saint Louis (outlined in white) is very similar to one included in Fischel's [1985] analysis, except that we add land use data. This reveals that the most striking feature is actually the different character of new development on incorporated versus unincorporated land. A disproportionate share of 1976–1992 development happens in unincorporated areas that were close to existing development but just beyond the municipal boundaries as they were circa 1980. This development is also more dispersed than that on incorporated land. Many other metropolitan areas show a similar pattern. There is a good reason for this: almost every zoning law includes the provision that whenever regulations differ, the most restrictive rules apply. In unincorporated

areas, only county and state planning regulations generally apply, while incorporated places add their own zoning restrictions and growth controls. *To the extent that there are unincorporated areas on the urban fringe, developers can escape municipal regulation by building outside municipal boundaries, and this facilitates sprawl.*

Finally, in Tiebout's [1956] model, zoning is the means by which communities can limit and shape immigration and development to suit the cost structure of local public goods. If local public services are more costly when development is scattered, then aversion to scattered development should be less strong, and *sprawl should be more prevalent where local taxpayers pay a smaller share of local government expenses.*

V. The Causes of Sprawl

Our review of the urban economics literature in the preceding section suggests that cities will sprawl more if

- they specialize in sectors where employment is not typically located close to the city center;
- they were built around the car rather than around public transport;
- they have experienced slow population growth;
- there is greater uncertainty regarding their future population growth;
- aquifers underlie a greater fraction of their urban fringe;
- they are not surrounded by high mountains;
- terrain in their urban fringe is rugged;
- their climate is temperate;
- they begin with substantial unincorporated areas on the urban fringe;
- local taxpayers pay a smaller share of local government expenses.

In this section we test these predictions by regressing our sprawl index for new development in individual metropolitan areas on initial metropolitan area characteristics. The dependent variable in our regressions is therefore the percentage of undeveloped land in the square kilometer around an average 1976–1992 residential development in each metropolitan area (i.e., we identify 30-meter cells that were not developed in 1976 but were subject to residential development between 1976 and 1992, calculate the percentage of land not developed by 1992 in the square

kilometer containing each of these 30-meter cells, and average across all such newly developed cells in the metropolitan area). Figures for the 40 largest metropolitan areas are reported in the final column of Table II, although we will consider all U. S. metropolitan areas in our regressions.

We focus on this dependent variable because concerns about new development drive most of the public debate about sprawl. In addition, we wish to avoid the obvious endogeneity issues that would arise if we instead used as dependent variable our sprawl index for the 1976 or 1992 stocks of development. However, as we show later, results using the sprawl index for the stock of 1992 development are very similar.

The spatial units of observation are individual metropolitan areas (although, obviously our calculations of the sprawl index and various explanatory variables still need to use the full spatial resolution of our data). We use the Metropolitan Statistical Area and Consolidated Metropolitan Statistical Area definitions (New England County Metropolitan Area definitions for New England). Since these are county-based definitions, care is needed when measuring the initial characteristics of areas where new development might take place. This is particularly important in the western part of the country, where counties are sometimes very large and consequently metropolitan area boundaries are often drawn much less tightly around the developed portion of metropolitan areas than in the East. We therefore restrict calculations for geographical variables to the "urban fringe," defined as those parts of the metropolitan area that were mostly undeveloped in 1976 but are located within 20 kilometers of areas that were mostly developed in 1976.[12] Given that we isolate the urban fringe in this manner, it makes sense to start with fairly wide metropolitan area boundaries before we cut out areas far away from initial development. We therefore use 1999 definitions [U. S. Bureau of the Census 2000]. We include all 275 metropolitan areas in the conterminous United States in our regressions.

12. Mostly developed areas are those where over 50 percent of the immediate square kilometer was developed in 1976. The choice of twenty kilometers as a threshold was guided by visual inspection of maps showing the evolution of land use in all metropolitan areas. A buffer of twenty kilometers around areas that were already mostly developed in 1976 includes 98 percent of 1976 residential development and 99 percent of subsequent residential development in metropolitan areas.

TABLE IV
THE DETERMINANTS OF SPRAWL

	Regression results				Summary statistics	
	(1)	(2)	(3)	(4)	Mean	St. dev.
Centralized-sector employment 1977	−1.270 (0.517)**	−1.194 (0.526)**	−0.922 (0.599)	−0.462 (0.489)	22.65	1.14
Streetcar passengers per capita 1902	−1.723 (0.507)***	−1.918 (0.553)***	−1.762 (0.520)***	−1.822 (0.535)***	21.53	62.54
Mean decennial % population growth 1920–1970	−6.072 (1.854)***	−5.528 (1.839)***	−6.241 (2.187)***	−4.686 (1.367)***	24.54	22.42
Std. dev. decennial % population growth 1920–1970	3.169 (1.315)**	3.208 (1.210)***	3.419 (1.424)**	2.482 (1.005)**	15.72	23.42
% of urban fringe overlying aquifers	1.222 (0.473)***	1.090 (0.507)**	0.945 (0.539)*	1.720 (0.484)***	30.43	37.96
Elevation range in urban fringe (m.)	−1.609 (0.946)*	−1.166 (1.023)	0.914 (1.117)	−1.731 (0.815)**	542.43	737.02
Terrain ruggedness index in urban fringe (m.)	1.252 (0.746)*	1.267 (0.746)*	1.108 (0.767)	2.195 (0.741)***	8.84	10.10
Mean cooling degree-days	−6.512 (1.562)***	−5.415 (1.657)***	−6.440 (2.359)***	−6.157 (1.564)***	1348.43	923.13
Mean heating degree-days	−4.986 (1.341)***	−4.768 (1.381)***	−3.051 (2.632)	−6.966 (1.360)***	4580.79	2235.66
% of urban fringe incorporated 1980	−1.363 (0.455)***	−1.558 (0.451)***	−1.708 (0.464)***	−1.629 (0.422)***	5.21	5.05
Intergov. transfers as % of local revenues 1967	1.075 (0.633)*	1.070 (0.682)	1.136 (0.679)*	2.206 (0.596)***	37.17	10.65
Bars and restaurants per thousand people		0.176 (0.783)		1.51	0.41	
Major road density in urban fringe (m./ha.)		−0.179 (0.698)			0.87	0.36
% population growth 1970–1990		−1.916 (0.910)**			35.29	45.46
Herfindahl index of incorporated place sizes		−0.274 (0.652)			0.32	0.26
Latitude			−2.083 (2.731)		37.57	5.22
Longitude			−5.221 (2.700)*		−91.18	13.52
Census division fixed effects			Included			
Constant	111.375 (11.503)***	108.895 (11.870)***	90.467 (21.441)***	75.050 (10.907)***		
Observations	275	275	275	275		
R^2	0.405	0.418	0.469	0.404		

The dependent variable in columns (1), (2), and (3) is our sprawl index for 1976–1992 development, which has mean 64.51 and standard deviation 10.90. The dependent variable in column (4) is our sprawl index for 1992 development, which has mean 46.54 and standard deviation 10.82. The regressions are run for all 275 metropolitan areas in the conterminous United States. Coefficients give the impact on the index of a one-standard-deviation increase in the corresponding variable. Numbers in brackets report heteroskedastic-consistent standard errors. ***, **, and * indicate significance at the 1 percent, 5 percent, and 10 percent level, respectively.

Results are reported in Table IV.[13] Column (1) reports our main specification. Columns (2) and (3) report results including additional variables and controls. Finally, column (4) repeats our main specification using the sprawl index for 1992 development rather than for 1976–1992 development as the dependent variable. To aid comparison across variables, we report standardized coefficients that measure the absolute change in the sprawl index for a one-standard deviation change in each independent variable.

V.A. The Monocentric City Model and Its Generalizations

We begin by examining the link between employment centralization and sprawl. Examining the link directly using a measure of the extent to which employment is centralized in each metropolitan area is clearly problematic: ceteris paribus, more compact cities will have more centralized employment. To avoid this endogeneity problem, we instead measure the extent to which the city is specialized in sectors, such as business services, that in the *average* city tend to be very centralized. To be precise, our measure is the share of employment that would be located within three miles of the central business district if employment in each sector in that metropolitan area was distributed relative to the center as it is in the average metropolitan area. See the Data Appendix for further details on how this variable is calculated. Results are reported in column (1) of Table IV. A one-standard deviation increase in centralized-sector employment decreases the sprawl index by 1.270 points. We see that, consistent with the monocentric city model, cities are more compact if they specialize in sectors that tend to be more centralized in the average metropolitan area.

Peoples' choice of residence might be driven by their leisure activities as well as by their employment. If employment centralization tends to limit the amount of sprawl, perhaps centralized amenities could play a similar role? To examine this possibility,

13. Our sprawl index is bounded between 0 and 100, so one may worry about the validity of OLS estimation. However, the minimum (33.35) and maximum (88.47) values occurring in the data are sufficiently far from the boundary to suggest that this is unlikely to cause problems in practice. This is also reflected in the fact that the minimum (36.01) and maximum (79.05) predicted values lie comfortably within the boundaries. Finally, note that converting to a (0,1) index and running the regression using a logistical transformation only results in some marginal changes to the significance of the results but makes coefficients much harder to interpret.

we tried including various consumer amenity variables used in Glaeser, Kolko, and Saiz [2001], such as live performance venues per capita or restaurants and bars per capita. Results reported in column (2), which includes restaurants and bars per capita, are typical: amenities have no impact and yet including them does not change our results. This is reasonably intuitive. Restaurants and bars are not actually any more centralized than household appliance stores [Glaeser and Kahn 2001]. Performance venues are, but the frequency with which most people go to these is such that their availability might affect the choice of metropolitan area (as suggested by the results on population growth in Glaeser, Kolko, and Saiz [2001]) but not so much the choice of whether to live in a compact or a scattered neighborhood.

We now turn to our prediction that car-friendly cities sprawl more. Naturally, cities developed mostly after the advent of the automobile tend to be much more car-friendly than cities built before 1900 around public transit. We use the number of streetcar passengers per capita in 1902 (from Cutler, Glaeser, and Vigdor [1999]) as a proxy for a historical city center less friendly to car usage. Table IV shows that a one-standard deviation increase in 1902 streetcar usage decreases the sprawl index by 1.723 points.

In addition to the role played by the historical city center, the car-friendliness of a city may also depend on the road density in the urban fringe. Column (2) shows that such a measure (major road density in the urban fringe, calculated from USGS 1980 digital line graphs) has no impact on sprawl, and including it does not change the coefficient on streetcar passengers or other variables. Note that, while more roads may facilitate scattered development, scattered development leads to a less dense road network. Our results suggest that neither of these counteracting effects dominates in the cross section. Using roads early in the study period does not solve this problem since, as we saw earlier, cities with more compact new development tended to also have more compact development in the past.

The third of our predictions concerns the impact of expected population growth on sprawl. In areas where population is growing fast, a rational agent anticipates that nearby vacant land will be developed sooner and, consequently, is not willing to incur large additional commuting costs to gain access to this open space. Developers may expect that cities that have been growing relatively fast in the past will continue to do so in the near future. We therefore proxy expected future population growth using the

metropolitan area's historical mean decennial percentage population growth for the five decades 1920–1970.[14] Historical population growth rates are indeed a good predictor of population growth between the 1970s and 1990s: the correlation between percentage population growth 1970–1990 and mean decennial percentage population growth 1920–1970 is 0.60. Results in Table IV show that areas that have historically seen high population growth rates do, indeed, see less sprawl. A one-standard deviation increase in the historical mean growth rate reduces the sprawl index by 6.072 points.

We interpret this result as telling us something about the value of open space. However, given that historical population growth rates are a good predictor of current population growth rates, this result would also be consistent with fast growing cities using all available land to accommodate their growing population. However, when we add actual 1970–1992 population growth (clearly endogenous, and only introduced as a robustness check) in column (2), we see that this does not explain our results. Faster contemporaneous population growth does make cities more compact, but historical population growth rates continue to have much the same impact on sprawl.

To test our fourth prediction that greater uncertainty regarding future city growth fosters sprawl, we similarly assume that developers consider future local population growth more uncertain in cities that have had more ups and downs in population growth rates over previous decades. Specifically, our measure of uncertainty is the standard deviation of decennial percentage population growth rates 1920–1970 (using the same population time series as above). The results in Table IV show that, as expected, higher uncertainty leads to more sprawl. A one-standard-deviation increase in the standard deviation of decennial population growth rates increases the sprawl index by 3.169 points.

V.B. When Space is Not a Featureless Plain

We now turn to consider the impact of a range of geographical variables. We begin with the prediction that aquifers facili-

14. Constructing a historical series of population data for U. S. metropolitan areas on the basis of county population counts in each decennial census requires tracking changes in county boundaries over time. We did this using a revised version of the County Longitudinal Template of Horan and Hargis [1995] kindly provided to us by Vernon Henderson and Jordan Rappaport.

tate sprawl, by allowing developers to sink a well and avoid the high water connection fees often incurred by scattered development. Results presented in Table IV show that this is indeed the case. A one-standard-deviation increase in the percentage of the urban fringe overlying aquifers (see the Data Appendix for details on this variable) increases the sprawl index by 1.222 points.

We think this result is particularly interesting. Urban economists have long highlighted the importance of indivisible public facilities for agglomeration. However, it is difficult to cleanly identify a role for indivisible public facilities in determining the extent to which development is clustered. Two particular features of water systems help us make a clean identification. First, we can detect their impact in the cross section because certain places (those with aquifers) have an alternative private provision that *is not* subject to the same indivisibilities. Second, the availability of this alternative provision through aquifers is certainly exogenous. This has some interesting policy implications that we consider in the conclusions.

What about terrain? We predict two effects from natural barriers and terrain ruggedness that should work in opposite directions. Coming up with a measure of the presence of mountains in the urban fringe is straightforward. For instance, we can calculate the range in elevation (i.e., the difference between the minimum and the maximum elevation) in the urban fringe. Measuring small-scale terrain irregularities, however, is more difficult because it requires much more geographically detailed elevation data. Given that readily available elevation grids covering the conterminous United States do not have the required spatial resolution, we have assembled a national elevation grid providing the elevation in meters of points 90 meters apart (see the Data Appendix for more detail). Using these data, we calculate the *terrain ruggedness index* originally devised by Riley, DeGloria, and Elliot [1999] to quantify topographic heterogeneity that can act either as concealment for prey or stalking cover for predators in wildlife habitats. This terrain ruggedness index, calculated on the 90-meter elevation grid, gives us a summary statistic of differences in meters of elevation between points 90-meters apart. This captures small-scale topographic heterogeneity using a local counterpart to the global elevation range that we use to capture the presence of mountains.

Turning again to our regression results in Table IV, we see that both mountains and hills have the expected effects. A one-

standard-deviation increase in the elevation range in the urban fringe decreases the sprawl index by 1.609 points. In contrast, more rugged terrain is associated with more sprawl. A one-standard-deviation increase in the terrain ruggedness index increases the sprawl index by 1.252 points.

There are other barriers to urban expansion that could in principle have a similar effect to that of mountains: in particular, proximity to wetlands, public land, or oceans. We have tried numerous measures of all of these in our regressions, and none of them matter empirically. In the case of wetlands and public lands, this is not too surprising. Wetland mitigation banking programs allow developers to build on wetland areas in exchange for financing the preservation or restoration of wetlands elsewhere. For public land, the Homestead Act of 1862 allowed settlers to easily acquire private ownership of public land. As a result, public lands are concentrated in those parts of the nation that have historically been least attractive for setting up a residence. The lack of impact of proximity to the Atlantic and Pacific Oceans, the Gulf of Mexico, and the Great Lakes is more surprising. We have tried hard to find evidence that proximity to these large water bodies reduces sprawl and found none. We conjecture that this is partly because oceans act both as a barrier and as an outdoor amenity: one cannot build on the ocean but proximity to the ocean makes open space more enjoyable. Furthermore, as illustrated by the map of Boston in the bottom panel of Figure IIa, a city can be bounded by the ocean on one side and still sprawl profusely on the other.

Our final prediction regarding the role of geographical variables is that characteristics that make open space less attractive should reduce sprawl. The two most obvious characteristics are whether the city has an extremely hot or cold climate. A standard measure of extreme heat is cooling degree days, a concept used by engineers to calculate the demand for air conditioning. Extreme cold can be similarly measured through heating degree days, used to calculate fuel demand for heating. We use mean annual cooling and heating degree days calculated from climatic normals for the period 1961–1990 (again, see the Data Appendix for more details). The results in Table IV show that both variables have the predicted effect. A one-standard-deviation increase in mean cooling degree days reduces the sprawl index by 6.512 points. While a one-standard deviation in mean heating days reduces the sprawl index by 4.986 points.

We have also checked whether other climatic variables, such as average precipitation, have an impact on sprawl and found no evidence that they do. Finally, we have examined whether sprawl is affected by other characteristics that may change the attractiveness of open space. Variables capturing the percentage of forest or various types of vegetation in the urban fringe have no significant effects. This is in accordance with the literature on the amenity value of vegetation, which finds very mixed results (see Irwin [2002]).

V.C. Political Geography

We turn, finally, to the role that political geography plays in driving sprawl. Estimation results confirm Fischel's assertion that the relationship between jurisdictional fragmentation and the restrictiveness of zoning is unlikely to be of empirical importance. Using a digital representation of the municipal boundaries in effect at the time of the 1980 census [GeoLytics 2000], we have computed various measures of municipal dominance (the ratio of the size of the largest municipality in each metropolitan area to the combined area of other municipalities, a Herfindahl index of municipality sizes, and the inverse of the number of municipalities). Results reported in column (2) for the Herfindahl index are typical: none of these measures have a statistically significant relationship with sprawl when added to our specification.

While competition between zoned areas of different sizes does not appear to matter for sprawl, the differences between zoned and unzoned areas stressed by Rubinfeld [1978] and Katz and Rosen [1987] do. To study the extent to which sprawl is encouraged by unincorporated areas on the urban fringe, that allow developers to escape municipal regulation, we calculate the percentage of the urban fringe incorporated in 1980. Results in column (1) of Table IV show that a one-standard-deviation increase in the percentage of the urban fringe incorporated reduces the sprawl index by 1.363 points. In all, these results suggest to us that the failure of municipal and county governments to harmonize land use regulation is an important contributor to sprawl. Developers, it seems, are often leapfrogging out of municipal regulations altogether rather than playing municipalities against each other.

To examine our final prediction that sprawl increases when local taxpayers bear less of the cost of providing public services to

scattered development on the urban fringe, we include the percentage of local government revenue that were transfer payments from other levels of government in 1967 [U. S. Bureau of the Census 1974]. Table IV shows that this variable has the expected positive effect on sprawl: a one-standard-deviation increase in the percentage of intergovernmental transfers in local revenues in 1967 increases the sprawl index by 1.075 points.

V.D. Physical Geography and Urban Sprawl

Our paper is unusual in its emphasis on the role that physical geography plays in explaining sprawl. In fact, a regression including only our five geographical variables (capturing the role of aquifers, terrain, and climate) explains 23.5 percent of the variation in our sprawl index. As one might expect, several of these variables vary in a quite predictable manner as one moves across the country. To check the extent to which these variables may just capture spatial gradients in the degree of sprawl, column (3) of Table IV reports results when we include the latitude and longitude of the centroid of each metropolitan area as well as fixed effects for nine census regions. Three geographical variables (the two terrain variables and mean heating degree days) are no longer significant at the 10 percent level. Remarkably, our aquifers variable and cooling degree days remain significant. In addition, the impact of all variables not measuring physical geography are essentially unchanged, with the exception of specialization in centralized sectors.

It is worth noting that, while our paper focuses on the causes of sprawl, there is also some public interest in the consequences of sprawl. Studying such consequences empirically will require good instruments, and our physical geography variables seem natural candidates. We also note that our results are robust to a variety of other changes to the specification in addition to those discussed throughout this section. Our regressions include all U. S. metropolitan areas regardless of their size. If we include the initial population of each metropolitan area in our specification, this variable is not significant, and the rest of our results are not affected. Similarly, the inclusion of other insignificant variables, such as various measures of demographic structure, segregation, or historical voting patterns, do not change the robustness of any of the results we report here.

V.E. Stocks Versus Flows of Development

We have seen in Section III that there is very high persistence in the extent to which individual metropolitan areas are either sprawling or compact. In fact, the correlation between the sprawl indices for the 1976 and 1992 stocks of development is 0.96. We might therefore expect the variables that explain how sprawling are the flows of new development to also explain the cross sectional variation in how sprawling is the stock of final (1992) or initial (1976) residential development. Column (4) shows that, with the exception of centralized sector employment, initial characteristics have exactly the same impact on the extent to which final development in the metropolitan area is sprawled. Results (not reported) are very similar for initial development and also do not change when explanatory variables calculated for the urban fringe are instead calculated for the entire metropolitan area.

VI. Conclusions

As with many economic and social processes, a true understanding of the implications of urban sprawl can only come about through the study of both the positive and normative aspects of the urban development process. Much of the current debate has seen people rushing to address normative issues without first having a good understanding of the positive aspects. In contrast, in providing the first detailed description of the process of urban development and its determinants, our paper is quite clearly focused on improving our understanding of these positive aspects.

To summarize, 1.9 percent of the land area of the United States was developed by 1992. Two-thirds of this developed land was already in urban use around 1976, while the remaining one-third was developed subsequently. Our main findings are concerned with whether development is sprawling or compact. We measure sprawl as the amount of undeveloped land surrounding an average urban dwelling. By this measure, commercial development has become somewhat more sprawling during the study period, but the extent of residential sprawl has remained roughly unchanged between 1976 and 1992. In contrast to this stability over time, the extent of sprawl does vary dramatically across metropolitan areas.

We study the factors that determine these large differences

across metropolitan areas. We find that sprawl is positively associated with the degree to which employment is dispersed; the reliance of a city on the automobile over public transport; fast population growth; the value of holding on to undeveloped plots of land; the ease of drilling a well; rugged terrains and no high mountains; temperate climate; the percentage of land in the urban fringe not subject to municipal planning regulations; and low impact of public service financing on local taxpayers.

We are some way away from being able to make firm policy recommendations, but our results do raise some interesting questions for policy in this area. Perhaps the most intriguing issue arises from the connection between aquifers and sprawl. Often the same aquifer will supply water both to municipal water systems and to individual private wells. Private incentives may push for scattered development over the aquifer, where one can sink a well and avoid connection fees to the municipal supply. However, such development may be costly for others, since concrete, asphalt, and other nonpermeable materials hinder the replenishment of the aquifer with rainwater. In such a context, raising impact fees may only worsen the problem. This raises the intriguing possibility that groundwater regulation may provide an important avenue through which policy makers can influence the form of urban development. Another interesting policy implication arises from the fact that disparities between municipal and county regulation are important causes of sprawl. Focus, so far, has been on the fragmented nature of local government, but our results suggest that harmonization of county and municipal land use regulation may actually play a much more important role in influencing the form of urban development. Interestingly, while we find that sprawl is affected by two factors which have received little attention, another (the density of roads) that has received much more attention seems to have little impact. While more car-friendly cities do experience more sprawl, we find that what really matters is not the density of the road network on the urban fringe but instead whether the city center was shaped before the advent of the car. Finally, our results on the transfer share in local revenues suggest that internalizing the fiscal externalities of new development appears to limit urban sprawl.

Of course, these comments are fairly speculative given the current state of our knowledge. Further analysis of economic models of development, and of models which incorporate a taste

for landscape features is warranted, and such analysis should form the basis for future policy recommendations.

Data Appendix

A. Land Use/Land Cover Data

We construct our core data from two remote-sensing data sets. The most recent, the 1992 National Land Cover Data [Vogelmann et al. 2001] are derived mainly from leaves-off (spring/fall) and leaves-on (summer) 1992 Landsat 5 Thematic Mapper satellite imagery. The Earth Resources Observation Systems (EROS) data center of the United States Geological Survey (USGS) converted the raw satellite images to land cover categories. Here we give a brief overview of the process, described in detail in Vogelmann, Sohl, Campbell, and Shaw [1998], Vogelmann, Sohl, and Howard [1998], and Vogelmann et al. [2001].

The Thematic Mapper sensor on the Landsat 5 satellite records data for units that are square pixels of 30×30 meters on a regular grid. We refer to these units as 30-meter cells. The sensor detects electromagnetic radiation reflecting from the earth's surface in seven wavelength bands (four of which are used to construct the data). Combining reflectance information from different bands for each 30-meter cell allows a very precise distinction between land cover features because different types of land cover reflect different amounts of radiation at different wavelengths. For instance, healthy vegetation reflects infrared light to remain cool and wet but absorbs visible light for photosynthesis.

Land cover was classified as follows. First, a computer algorithm was used to find clusters of contiguous 30-meter cells with a similar set of reflectance values over the electromagnetic spectrum. Next, analysts used high-altitude aerial photographs and other census and remote sensing data to match these clusters to land cover classes, to refine the boundaries of these clusters, and to make finer distinctions between land cover classes. Since a single cell may contain multiple land cover types, categorization is based on thresholds. For instance, for a cell to be assigned an urban code at least 30 percent of it must be covered with constructed materials. Using this approach, each 30-meter cell was categorized into one of 21 land cover classes.

Like the 1990s data, the 1970s [U. S. Geological Survey 1990;

U. S. Environmental Protection Agency 1994] classify the conterminous U. S. land area into land use/land cover categories. However, rather than satellite imagery, the 1970s data derive mainly from high-altitude aerial photographs collected between 1971–1982. The most common date is 1976, which is also the median year. The conversion to land use/land cover data was done by the USGS. The U. S. Environmental Protection Agency (EPA) further processed the data to facilitate use in geographic information systems, and we use their version [U. S. Environmental Protection Agency 1994]. We filled gaps in these data to construct the first complete coverage for the conterminous United States.[15]

To construct the 1976 data, analysts studied the photographs and, with the help of ancillary data, traced the boundaries of contiguous areas with similar land cover and assigned one of 37 land cover codes. The rules for drawing these boundaries mean that areas may differ in size and that a single area may contain multiple land cover types. Thus, as before, categorization is based on thresholds. For instance, to be assigned an urban code an area must have at least 20 percent urban cover within 4 hectares (10 acres). The resulting data contain the digitized boundaries of these hand-drawn areas (irregular polygons) and a code describing the preponderant land cover for each of them. U. S. Geological Survey [1990] gives a more detailed description of this process.

While the 1976 and 1992 data are roughly comparable, there are a few differences with implications for our analysis. First, the 1992 data are stored in raster format (assigning a code to each cell on a regular grid) while the 1976 data are stored in vector format (assigning a code and providing coordinates for irregular polygons). They also have different geographical projections. Thus, we converted the 1976 data to the same projection and data

15. The digital version of the land use and land cover data from 1:250,000 scale maps produced by the USGS lacks data for a thirty-by-sixty minute rectangle in the map for Albuquerque and in the map for Cedar City and for a one degree by one degree square in the map for Tampa. For Albuquerque and Cedar City, the USGS had digitized data from the 1:100,000 scale maps corresponding to the rectangles with missing data (Chaco Mesa in the case of Albuquerque, and Kanab in the case of Cedar City). We processed these data with the same computer code used by the EPA for the rest of the nation to completely fill the gaps. For Tampa, the missing data were not available digitally but could be found in the corresponding 1:250,000 scale paper map distributed by the USGS. We digitized this to the same format specifications as the rest of the EPA data. Using the USGS paper and digital distributions of the data and two alternative sources for the EPA distribution, we were also able to correct various instances in which land use codes had become corrupted during processing stages that occurred before we received the data. The data used to fill the three holes in the USGS data are available from http://diegopuga.org/data/sprawl/.

model as the 1992 data, by breaking up each polygon into the 30-meter cells it contains. This yields a data set giving the preponderant land cover/land use of each 30-meter cell in a regular grid covering the entire conterminous United States circa 1976 and in 1992. The second difference is that the data are categorized using classifications with different degrees of detail. For this reason, we work with two urban codes that can be defined in both years: residential; and commercial, industrial, and transportation networks.

The third and most important difference arises from the fact that the 1976 data are slightly less precise than the 1992 data when identifying small features different from their surroundings. Given this, rather than compare the data directly, we use the 1976 data to separate urban land in 1992 into new and old development. Thus, we define old development as land that was classified as urban in both 1992 and 1976. We define new development as land that was classified as urban in 1992, but was not urban in 1976. This procedure largely corrects for the difference, but has the drawback that we cannot capture developed land that is converted to farmland, etc. However, such undevelopment is rare: calculations by the Department of Agriculture suggest that less than 0.8 percent of developed land was undeveloped over the fifteen-year period 1982–1997 [U. S. Department of Agriculture 2000].

One possible source of mismeasurement remains: we may date some development incorrectly, if it is small enough relative to the resolution of our data and different from its surroundings in at least one of the two periods. We cannot provide a precise upper bound on the magnitude of this misdating. However, careful inspection suggests that only one result might be sensitive to this: when we find that commercial development became more biased toward scattered areas, this result is amplified by the fact that land classified as commercial/industrial/transportation in 1992 occasionally includes small rural roads that were too small to register with the 1976 data.

B. Data for Alternative Sprawl Measures

Median lot size was compiled from the metropolitan data contained in the American Housing Survey [U. S. Bureau of the Census 1994–1998]. The metropolitan data in the American Housing Survey cover 47 metropolitan areas, where a sample of householders are interviewed about every six years. Each year,

data for a few metropolitan areas are gathered on a rotating basis until all 47 areas included are surveyed. The cycle then begins again. The American Housing Survey does not survey three metropolitan areas with populations over one million (Greensboro, New Haven, and Orlando), although in the case of Greensboro median lot size in 1995 is available from the City of Greensboro Planning Department 2003. Thus, we have median lot size data for 38 cities with populations over one million.

The average number of miles driven per person in individual metropolitan areas was calculated from the 1995 Nationwide Personal Transportation Survey [U. S. Federal Highway Administration 1995], using the tools to calculate local area statistics described in Reuscher, Schmoyer, and Hu [2001]. The share of employment located more than three miles away from the central business district in 1996 was kindly provided by Matt Kahn from Glaeser and Kahn [2001].

C. Additional Data for the Determinants of Sprawl

The following paragraphs provide details on data sources and construction for several variables used in our regressions. All data required to run these regressions are available from http://diegopuga.org/data/sprawl/.

Centralized sector employment 1977: For each metropolitan area, we use county business pattern data for 1977 to calculate the share of employment in each three-digit SIC sector i, $s_{MSA,i}$. For each sector we know from Glaeser and Kahn [2001] the mean percentage of metropolitan area employment in that sector that is found within three miles of the central business district, $\bar{s}_{3,i}$ (see their paper for details of the calculations). Our measure of centralization of employment is then calculated as $\sum_i s_{msa,i} \times \bar{s}_{3,i}$.

Percentage of the urban fringe overlaying aquifers: We use data from U. S. Geological Survey [2003], originally developed by the USGS to produce the maps printed in the *Ground Water Atlas of the United States* [U. S. Geological Survey 2000]. This contains the shallowest principal aquifer at each point of the United States in a continuous geographical coverage. We exclude shallow sand and gravel aquifers since their high permeability and shallow depth to the water table make them particularly susceptible to contamination from nitrates and other pollutants whose presence in sufficient quantity renders water unsuitable for human consumption [Burkart and Stoner 2002].

Elevation range and Terrain Ruggedness Index in the urban

fringe: We assemble the national elevation grid by merging 922 separate elevation grids from the 1:250,000-scale Digital Elevation Models of the USGS, each of which provides 3-arc-second elevation data for an area of one by one degrees. Let $e_{r,c}$ denote elevation at the point located in row r and column c of a grid of elevation points. Then the Terrain Ruggedness Index of Riley, DeGloria, and Eliot [1999] at that point is calculated as $[\sum_{i=r-1}^{r+1} \sum_{i=c-1}^{c+1}(e_{i,j} - e_{r,c})^2]^{1/2}$. The variable used in the regression is the average terrain ruggedness index of the urban fringe in each metropolitan area.

Mean cooling and heating degree days: Our weather variables are calculated from the climatic normals for individual weather stations 1961–1990 contained in the *Climate Atlas of the United States.* Cooling degrees on a given day are zero if the average temperature is below 65°F (about 18°C) and the degrees by which the average temperature exceeds 65°F otherwise. Mean annual cooling degree days are computed by summing cooling degrees over all days in a year. Mean annual heating degree days are similarly calculated by summing degrees below 65°F over all days in a year. We computed metropolitan area mean cooling and heating degree days by averaging climatic normals over all reporting weather stations in each metropolitan area. For the four metropolitan areas that did not contain a reporting station, we averaged data from weather stations within 30 kilometers of the metropolitan area.

NEPTIS FOUNDATION
LONDON SCHOOL OF ECONOMICS AND CEPR
UNIVERSITY OF TORONTO, UNIVERSITAT POMPEU FABRA, CREI, CEPR, AND NBER
UNIVERSITY OF TORONTO

REFERENCES

Alonso, William, *Location and Land Use: Toward a General Theory of Land Rent* (Cambridge, MA: Harvard University Press, 1964).
Anas, Alex, Richard Arnott, and Kenneth A. Small, "Urban Spatial Structure," *Journal of Economic Literature,* XXXVI (1998), 1426–1464.
Arnott, Richard J., and Frank D. Lewis, "The Transition of Land to Urban Use," *Journal of Political Economy,* LXXXVII (1979), 161–169.
Bar-Ilan, Avner, and William C. Strange, "Urban Development with Lags," *Journal of Urban Economics,* XXXIX (1996), 87–113.
Black, Duncan, and Vernon Henderson, "Urban Evolution in the USA," *Journal of Economic Geography,* III (2003), 343–372.
Brueckner, Jan K., "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model," in Edwin S. Mills, ed., *Handbook of Regional and Urban Economics,* Volume II (Amsterdam: North-Holland, 1987), pp. 821–845.

——, "Urban Growth Models with Durable Housing: An Overview," in Jean-Marie Huriot and Jacques-François Thisse, eds., *Economics of Cities: Theoretical Perspectives* (Cambridge: Cambridge University Press, 2000), pp. 263–289.
Buchanan, James M., "An Economic Theory of Clubs," *Economica,* XXXII (1965), 1–14.
Burkart, Michael R., and Jeffrey D. Stoner, "Nitrate in Aquifers Beneath Agricultural Systems," *Water Science and Technology,* XLV (2002), 19–29.
Capozza, Dennis R., and Robert W. Helsley, "The Stochastic City," *Journal of Urban Economics,* XXVIII (1990), 187–203.
City of Greensboro Planning Department, *Greensboro City Data Book* (Greensboro, NC: 2003).
Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor, "The Rise and Decline of the American Ghetto," *Journal of Political Economy,* CVII (1999), 455–506.
Duranton, Gilles, and Diego Puga, "Micro-Foundations of Urban Agglomeration Economies," in Vernon Henderson and Jacques-François Thisse, eds., *Handbook of Regional and Urban Economics,* Volume IV (Amsterdam: North-Holland, 2004), pp. 2063–2117.
Ewing, Reid, Rolf Pendall, and Don Chen, *Measuring Sprawl and Its Impact* (Washington, DC: Smart Growth America, 2002).
Fischel, William A., *The Economics of Zoning Laws: A Property Rights Approach to American Land Use Controls* (Baltimore, MD: Johns Hopkins University Press, 1985).
Fujita, Masahisa, "Spatial Patterns of Urban-Growth–Optimum and Market," *Journal of Urban Economics,* III (1976), 209–241.
——, "Urban Spatial Dynamics: A Review," *Sistemi Urbani,* III (1983), 411–475.
Fujita, Masahisa, and Hideaki Ogawa, "Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations," *Regional Science and Urban Economics,* XII (1982), 161–196.
Fujita, Masahisa, and Jacques-François Thisse, *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth* (Cambridge: Cambridge University Press, 2002).
Fulton, William, Rolf Pendall, Mai Nguyen, and Alicia Harrison, "Who Sprawls Most? How Growth Patterns Differ across the U. S.," Survey series, Brookings Institution, 2001.
Geoghegan, Jacqueline, "The Value of Open Spaces in Residential Land Use," *Land Use Policy,* XIX (2002), 91–98.
Geoghegan, Jacqueline, Linda A. Wainger, and Nancy E. Bockstael, "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS," *Ecological Economics,* XXIII (1997), 251–264.
GeoLytics, *CensusCD 1980, Version 2* (East Brunswick, NJ: GeoLytics, Inc., 2000).
Glaeser, Edward L., and Matthew Kahn, "Decentralized Employment and the Transformation of the American city," *Brookings-Wharton Papers on Urban Affairs* (2001), 1–47.
Glaeser, Edward L., and Matthew E. Kahn, "Sprawl and Urban Growth," in Vernon Henderson and Jacques-François Thisse, eds., *Handbook of Regional and Urban Economics,* Volume IV (Amsterdam: North-Holland, 2004), pp. 2481–2527.
Glaeser, Edward L., and Janet E. Kohlhase, "Cities, Regions and the Decline of Transport Costs," *Papers in Regional Science,* LXXXIII (2004), 197–228.
Glaeser, Edward L., Jed Kolko, and Albert Saiz, "Consumer City," *Journal of Economic Geography,* I (2001), 27–50.
Glaeser, Edward L., José A. Scheinkman, and Andrei Shleifer, "Economic-Growth in a Cross-Section of Cities," *Journal of Monetary Economics,* XXXVI (1995), 117–143.
Henderson, J. Vernon, "The Sizes and Types of Cities," *American Economic Review,* LXIV (1974), 640–656.
——, "General Equilibrium Modelling of Systems of Cities," in Edwin S. Mills, ed., *Handbook of Regional and Urban Economics,* Volume II (Amsterdam: North-Holland, 1987), pp. 927–956.
Holmes, Thomas J., and John J. Stevens, "Spatial Distribution of Economic Activities in North America," in Vernon Henderson and Jacques-François

Thisse, eds., *Handbook of Regional and Urban Economics,* Volume IV (Amsterdam: North-Holland, 2004), pp. 2797–2843.
Horan, Patrick M., and Peggy G. Hargis, *County Longitudinal Template, 1840–1990* (Ann Arbor, MI: Inter-university Consortium for Political and Social Research (ICPSR 6576) 1995).
Imai, Haruo, "CBD Hypothesis and Economies of Agglomeration," *Journal of Economic Theory,* XXVIII (1982), 275–299.
Irwin, Elena G., "The Effects of Open Space on Residential Property Values," *Land Economics,* LXXVIII (2002), 465–480.
Irwin, Elena G., and Nancy E. Bockstael, "Interacting Agents, Spatial Externalities and the Evolution of Residential Land Use Patterns," *Journal of Economic Geography,* II (2002), 31–54.
Katz, Lawrence, and Kenneth T. Rosen, "The Interjurisdictional Effects of Growth Controls on Housing Prices," *Journal of Law & Economics,* XXX (1987), 149–160.
Lucas, Robert E., Jr., and Esteban Rossi-Hansberg, "On the Internal Structure of Cities," *Econometrica,* LXX (2002), 1445–1476.
Mieszkowski, Peter, and Edwin S. Mills, "The Causes of Metropolitan Suburbanization," *Journal of Economic Perspectives,* VII (1993), 135–147.
Mieszkowski, Peter, and Barton Smith, "Analyzing Urban Decentralization—The Case of Houston," *Regional Science and Urban Economics,* XXI (1991), 183–199.
Mills, David E., "Growth, Speculation and Sprawl in a Monocentric City," *Journal of Urban Economics,* X (1981), 201–226.
Mills, Edwin S., "An Aggregative Model of Resource Allocation in a Metropolitan Area," *American Economic Review Papers and Proceedings,* LVII (1967), 197–210.
Muth, Richard F., *Cities and Housing* (Chicago: University of Chicago Press, 1969).
Ohls, James C., and David Pines, "Discontinuous Urban-Development and Economic Efficiency," *Land Economics,* LI (1975), 224–234.
Overman, Henry G., and Yannis M. Ioannides, "Cross-Sectional Evolution of the U. S. City Size Distribution," *Journal of Urban Economics,* XLIX (2001), 543–566.
Pew Center for Civic Journalism, *Straight Talk From Americans—2000* (Washington, DC: Pew Center for Civic Journalism, 2000).
Rappaport, Jordan, and Jeffrey D. Sachs, "The United States as a Coastal Nation," *Journal of Economic Growth,* VIII (2003), 5–46.
Reuscher, Timothy R., Richard L. Schmoyer, Jr., and Patricia S. Hu, *Transferability of Nationwide Personal Transportation Survey Data to Regional and Local Scales* (Oak Ridge, TN: Oak Ridge National Laboratory Center for Transportation Analysis, 2001).
Riley, Shawn J., Stephen D. DeGloria, and Robert Elliot, "A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity," *Intermountain Journal of Sciences,* V (1999), 23–27.
Rosenthal, Stuart S., and Robert W. Helsley, "Redevelopment and the Urban Land Price Gradient," *Journal of Urban Economics,* XXXV (1994), 182–200.
Rosenthal, Stuart S., and William Strange, "Evidence on the Nature and Sources of Agglomeration Economies," in Vernon Henderson and Jacques-François Thisse, eds., *Handbook of Regional and Urban Economics,* Volume IV (Amsterdam: North-Holland, 2004), pp. 2119–2171.
Rubinfeld, Daniel L., "Suburban Employment and Zoning—General Equilibrium-Analysis," *Journal of Regional Science,* XVIII (1978), 33–44.
Southern California Studies Center and Brookings Institution, *Sprawl Hits the Wall: Confronting the Realities of Metropolitan Los Angeles* (Los Angeles, CA: Southern California Studies Center, 2001).
Tiebout, Charles M., "A Pure Theory of Local Expenditures," *Journal of Political Economy,* LXIV (1956), 416–424.
Turner, Matthew A., "Landscape Preferences and Patterns of Residential Development," *Journal of Urban Economics,* LVII (2005), 19–54.
U. S. Bureau of the Census, *County and City Data Book,* 1972 (Ann Arbor, MI:

Inter-university Consortium for Political and Social Research (ICPSR 0061), 1974).
——, *American Housing Survey* (Washington, DC: United States Department of Commerce, Bureau of the Census, 1994–1998).
——, *Statistical Abstract of the United States,* 2000 (Washington, DC: United States Department of Commerce, Bureau of the Census, 2000).
U. S. Department of Agriculture, *National Resources Inventory Data Collection Instructions: Primary Sampling Unit Module II—Farmsteads and Developed Areas* (Washington, DC: United States Department of Agriculture, Natural Resources Conservation Service, 1997).
——, *Summary Report: 1997 National Resources Inventory (Revised December 2000)* (Washington, DC, and Ames, IA: United States Department of Agriculture, Natural Resources Conservation Service, and Statistical Laboratory Iowa State University, 2000).
——, *1997 National Resources Inventory (Revised December 2000): A Guide for Users of 1997 NRI Data Files* (Washington, DC: United States Department of Agriculture, Natural Resources Conservation Service, 2001).
U. S. Environmental Protection Agency, *1:250,000 Scale Quadrangles of Land-use/Landcover GIRAS Spatial Data in the Conterminous United States* (Washington, DC: United States Environmental Protection Agency, Office of Information Resources Management, 1994).
——, *Drinking Water Infrastructure Needs: Survey First Report to Congress* (Washington, DC: United States Environmental Protection Agency, Office of Water, 1997).
U. S. Federal Highway Administration, *Nationwide Personal Transportation Survey* (Research Triangle Park, NC: United States Department of Transportation, Federal Highway Administration, 1995).
U. S. Geological Survey, *Land Use and Land Cover Digital Data from 1:250,000- and 1:100,000-Scale Maps: Data User Guide 4* (Reston, VA: United States Geological Survey, 1990).
——, *Ground Water Atlas of the United States* (Reston, VA: United States Geological Survey, 2000).
——, *Principal Aquifers of the 48 Conterminous United States, Hawaii, Puerto Rico, and the U. S. Virgin Islands* (Madison, WI: United States Geological Survey, 2003).
Vogelmann, J. E., T. Sohl, P. V. Campbell, and D. M. Shaw, "Regional Land Cover Characterization Using Landsat Thematic Mapper Data and Ancillary Data Sources," *Environmental Monitoring and Assessment,* LI (1998), 415–428.
Vogelmann, J. E., T. Sohl, and S. M. Howard, "Regional Characterization of Land Cover Using Multiple Sources of Data," *Photogrammetric Engineering & Remote Sensing,* LXIV (1998), 45–57.
Vogelmann, James E., Stephen M. Howard, Limin Yang, Charles R. Larson, Bruce K. Wylie, and Nick Van Driel, "Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper Data and Ancillary Data Sources," *Photogrammetric Engineering & Remote Sensing,* LXVII (2001), 650–684.
Wheaton, William C., "A Comparative Static Analysis of Urban Spatial Structure," *Journal of Economic Theory,* IX (1974), 223–237.