

Report Part Title: Changing Urban Landscapes in the United States:

Report Title: Urban Sprawl

Report Author(s): Thomas J. Nechyba and Randall P. Walsh

Lincoln Institute of Land Policy (2004)

Stable URL: <http://www.jstor.com/stable/resrep18527.3>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



*Lincoln Institute of Land Policy* is collaborating with JSTOR to digitize, preserve and extend access to this content.

JSTOR

## **Changing Urban Landscapes in the United States: 1900-2000**

The central theme of urban development over the past century is surely the increasing trend toward suburbanization, as central cities have struggled to hold onto households and jobs. In explaining this trend, the urban economics literature casts a primary focus on the role of declining transportation costs and rising incomes, with supporting forces emerging from various government tax, expenditure and zoning policies. The local public finance literature, on the other hand, emphasizes the desire of mobile households to segregate based on preferences for local taxes and amenities as well as the desire by such households to take advantage of peer externalities. After reviewing some basic facts, we discuss the potential causes for and consequences of sprawl that emerge from each of these two overlapping literatures. Both these literatures then inform a brief discussion of the phenomenon of edge cities and the very different nature of sprawl in the United States and Europe.

### **Some Facts on Urban Sprawl and Suburbanization**

Urban sprawl can take on different forms. It may involve low-density residential developments or so-called “edge cities” (clusters of population and economic activity at the urban fringe) that give rise to business activity like office buildings, retail, and even manufacturing. It can take the form of planned communities that have their own “downtown” or are aligned to a lake or park. Or it can occur as individual houses pop up across formerly rural landscapes. A common way to document the presence of urban sprawl over time, however, is to look first at the evolution of rural and urban population levels, and then to look within urban areas at the evolving relationship between suburbs and central cities.

Figure 1 documents the dramatic transformation of a primarily rural population in 1790 to one that became increasingly centered in cities over the course of the nineteenth and early twentieth centuries. Only slightly more than 5 percent of the U.S. population lived in urban areas in 1790, a figure that had tripled by 1850 and surpassed 50 percent by 1920. By the 2000 Census, 79 percent of all Americans lived in areas designated as “urban” by the Census Bureau. The nineteenth and early part of the twentieth century can thus broadly be characterized as a time during which the industrial revolution transformed an agrarian economy into one that became increasingly dominated by cities. Much of the remainder of the twentieth century, on the other hand, witnessed the accelerated growth of suburbs within urban areas – with seeds for that growth sown prior to World War II.

Data documenting the rise of suburbs become readily available only with the 1950 Census when the Census Bureau first defined “urbanized areas” to include only those areas that truly represent built-up urban or suburban census blocks (as opposed to “metropolitan statistical areas” which often include large areas of unused land) and divided populations within these areas into suburban and central city populations. Within the largest urbanized areas, central cities contained close to 65 percent of urbanized

populations in 1950 – suggesting that suburbs had already reached substantial sizes (35 percent of urbanized populations) by mid-century.

Combining data on urbanized areas with populations larger than 1 million, Figures 2 and 3 then paint a stark picture of how cities have evolved since 1950. Both in terms of population (Figure 2) and land area (Figure 3), central cities within the urbanized areas have remained relatively stagnant while suburbs have experienced and are continuing to experience enormous growth. Approximately 65 percent of the urbanized population lived in central cities in 1950, with the remaining 35 percent residing in suburbs. By 1990, these percentages had flipped, with central city populations down to 35 percent of populations within these urbanized areas. The total land occupied by central cities has fallen from roughly 40 percent to 20 percent of urbanized areas during the same time. Finally, Figure 4 shows how population densities in both central cities and suburbs have declined over this period, suggesting that both increases in overall population as well as declines in city densities have contributed to expanding urban footprints.

### The Monocentric City Model: Falling Transportation Costs, Rising Incomes and Expanding City Footprints

Much of our understanding of this urban growth can be derived from the “monocentric city model” (Alonzo, 1964; Muth, 1969; Mills, 1967), which explains urban spatial structure as arising from the trade-off between commuting costs and land rents. In equilibrium, this trade-off requires lower land rents at the urban edge to offset increased commute costs—with the declining rent gradients leading to declining density gradients as one moves out from a metropolitan area’s central business district to the urban boundary. While the model captures the basic fact that downtown real estate is typically more expensive than equivalent land in the suburbs, it does not offer large insights into the development of the microstructure of the urban landscape. The urban economics literature that uses this model highlights the role of declining transportation costs – primarily cars on the consumer side and trucks on the producer side, combined with public infrastructure investment in roads – to explain the general decline in city density and expansion of city footprints, or urban sprawl, experienced over the last century.<sup>1</sup>

In this view, the advent of the automobile and accompanying lower transportation costs became the primary catalyst of sprawling cities through much of the twentieth century (Glaeser and Kahn, 2003). By 1910 the number of registered automobiles in the U.S. had passed the 500,000 mark, and in 1920 car registration reached eight million. In 1922, the first suburban, auto-oriented shopping center was constructed in Kansas City, Missouri (Williams, 2000), and by 1952 a majority of households in America owned at least one car (Glaeser and Kahn, 2003). The percentage of workers that drove to work stood at 64 percent in 1960 and rose to 78 percent by 1970 and 84 percent in 1980. It is difficult to imagine large increases in suburbanization without this rise of the automobile, even if other causes have contributed to the sprawling of cities in the presence of the automobile.

The monocentric city model also suggests that rising incomes have led to decreasing city densities to the extent that the income elasticity of demand for housing and land is

sufficiently large relative to the income elasticity of commuting costs.<sup>2</sup> Empirical evidence on the role of rising incomes in urban sprawl is provided by Margo (1992) and Brueckner (2000, 2001), with Margo's analysis suggesting that as much as half of the increase in suburbanization between 1950 and 1980 can be explained by rising incomes. Thus, the monocentric city model ultimately relies on the combined effect of increasing income and lower transportation costs to explain the phenomenon of suburbanization and sprawl.

While a variety of public policies have been suggested as potentially important contributors to urban sprawl within the monocentric city model, the empirical evidence suggests that these played at best a minor role—at least to the extent to which they did not contribute directly to enabling the rise of the automobile (such as through the construction of roads). Urban development may have been affected by the New Deal's creation of the Home Owners Loan Corporation and the Federal Housing Administration (FHA) (Jackson, 1981) and by increased post-World War II support for mortgage insurance programs through the FHA and the Veterans Administration (Williams, 2000). The federal deductibility of mortgage interest lowers the price of housing disproportionately for higher income families in higher tax brackets (Voith, 1999), potentially reinforcing the tendency of higher income households to suburbanize and commute. However, the U.S. General Accounting Office (1999) provides a skeptical review of the evidence of the extent to which such federal policies have created urban sprawl, at least as compared to the impact of lower transportation costs and higher incomes. At a more local level, the use of property taxes, as opposed to pure land taxes that do not tax improvements of land, provides incentives for low-density development in a monocentric city model (Brueckner and Kim, 2003). Again, it seems unlikely that local property taxation has played anything other than a supporting role in generating sprawl, especially in light of the fact that when interacted with minimum zoning rules, such taxes may in practice exhibit many of the features of land taxes (Fischel, 2001a).<sup>3</sup>

### The Tiebout Local Public Finance Model: Local Public Goods, Peer Externalities and Segregation

The monocentric model is a useful starting point for studying urban patterns and almost certainly leads to the appropriate identification of the primary historical cause of urban sprawl, but the empirical evidence suggests strongly that residential location choices *within* metropolitan areas are made on the basis of many factors other than transportation and commuting costs, such as local schools, crime rates and other local public amenities. These local public finance considerations play an important role in the debate over the costs and benefits of urban sprawl even if they themselves are not primary causes for the growing city footprints and declining population densities attributed to sprawl. Awareness of such considerations in residential location choices arose from Tiebout's (1956) classic article that suggests how people may sort themselves into different local jurisdictions (potentially within the same metropolitan area) based on their tastes for local amenities.

These sorting effects can be separated into two broad categories for purposes of our discussion of sprawl: those which *pull* people out of central cities because of attractive features of suburbs, and those which *push* people out of central cities because of inner city

problems. The pull side of the Tiebout coin emphasizes how relatively mobile families form new cities in the suburbs in part to create communities comprised of households with similar willingness to pay for the provision of public goods or with other characteristics considered “desirable.” When combined with the ability to zone (Fischel, 2001a,b), such suburban jurisdictions provide the additional “advantage” of permitting residents to exclude those who are thought to bring with them either negative fiscal externalities (in the sense that they will free ride on tax payments by others) or negative peer externalities in forms like higher crime rates or lower school quality. The push (or “flight from blight”) side of the Tiebout coin, on the other hand, refers to the hypothesized propensity of relatively high-income residents to leave the central city in response to higher inner city crime rates, lower quality schools and general fiscal distress within the central business district. Thus, inner city residents may wish to leave central cities not because they seek to form or join a particular (more homogeneous) suburb but rather to escape inner city problems. Cullen and Levitt (1999), for instance, find that a 10 percent increase in crime corresponds to a 1 percent decline in central city population. Other related recent work includes Leicknenko (2001) who looks at regional location and climate, Sigleman and Henig (2001) who investigate different preferences between races, and Adams et al. (1996) who focus on more general “central city hardship conditions.” The research literature has paid less attention to the role of potential “pull” Tiebout forces within sprawling cities.

Tiebout sorting within metropolitan areas, whether of the push or pull variety, is prevalent and empirically well documented. The extent to which sprawl is *caused* by such inter-urban and suburban sorting (made easier by lower transportation costs), however, seems very much an open question. Put differently, the empirical importance of local amenities (such as schools and low crime rates) in determining household residential location decisions is beyond question. A long capitalization literature dating back to Oates (1969) has demonstrated the importance of such amenities in housing prices, as have more recent neighborhood discontinuity studies (Black, 1999) and discrete choice models (Nechyba and Strauss, 1998; Bayer et al., 2002). But little research to date has investigated to what extent this sorting has contributed to expanding city footprints rather than representing a consequence of such expansion as cars enable households to consider wider geographic areas in which to reside. Either way, Tiebout sorting remains an important component to the discussion of urban sprawl as it informs our understanding of how expanding cities develop within the shifting urban boundaries. We will have more to say on this in the next section as we discuss equity and efficiency concerns raised by the interaction of sprawl and Tiebout sorting.

The most studied empirical link between sorting and sprawl relates to the relationship between sprawl and racial segregation. In principle, sprawl may both aggravate and lessen racial segregation within urban areas. On the one hand, minorities – because of lower incomes as well as possible exclusionary suburban policies -- may become increasingly segregated in central cities as suburbs grow, or they may live in segregated suburbs when they do move out of central cities. On the other hand, fast-growing metropolitan areas may give rise to an emerging minority middle class that can afford to move into suburbs – potentially decreasing racial segregation, especially in light of the fact that segregation within central cities is often quite stark. The empirical evidence is mixed, although we

interpret recent evidence to be more supportive of the hypothesis that greater Tiebout mobility leads, all else being equal, to greater segregation. Several authors including Lewis (1973), Yinger (1995), and Powell (2000) have found a direct relationship between urban development and racial segregation, while others have found no link between density and segregation (Cutler, Glaeser, and Vigdor, 1999), and yet others have found that fast-growing metro areas have experienced a sharper decline in racial segregation than slow-growth areas (Glaeser and Kahn, 2003). Bayer, McMillan and Rueben (2002) document a tendency of households to seek to reside near households of similar race/ethnicity – suggesting that, to the extent to which sprawl contributes to an increase in Tiebout sorting, it may contribute to an increase in racial segregation. Alesina, Baqir and Hoxby (forthcoming) find evidence suggesting that greater racial heterogeneity leads to greater numbers of local governments within metropolitan areas – and greater segregation between these sub-metropolitan boundaries.

### Edge Cities: Multiple Centers and Further Sorting

As residential sprawling and suburbanization solidified over the course of the twentieth century, the last few decades also witnessed a growing trend towards “edge cities,” with multiple employment centers located throughout many metropolitan areas. Edge cities pose difficulties for models of urban patterns based either on transportation or on sorting, and their increasing empirical importance has led to developments of alternative polycentric city models that endogenize the formation of employment centers outside the central business district (Anas, Arnott and Small, 1998; Brueckner, 1979; McDonald and McMillan, 2000; Henderson and Mitra, 1999). While some researchers have focused on patterns of dense employment subsectors at the outskirts of cities (Brueckner, 1979, Henderson and Mitra, 1999), Glaeser and Kahn (2003) suggest that edge cities typically represent relatively low density employment areas that accompany low density suburbanization. The formation of edge cities or decentralized employment centers raises efficiency and equity concerns that link to similar issues raised by the Tiebout literature below and must be balanced against the potential for lost agglomeration opportunities at the urban core.

In addition, edge cities may contribute to the “spatial mismatch hypothesis” first analyzed by Kain (1968), which suggests that job suburbanization has led to a disconnect in locations between jobs and low-income residential developments that are inhabited by less mobile households. In cities with little public transportation (Raphael and Stoll, 2001), this spatial mismatch may suppress employment opportunities for the poor who do not have access to the transportation technologies (cars) that drive the sprawling of cities and jobs. It remains difficult to determine whether jobs follow people or people follow jobs, although the evidence to date suggests that the former may be the case more than the latter (Stainnes, 1977; Glaeser and Kahn, 2003).

### Comparative Study of Urban Sprawl: The United States and Europe

The United States and the nations of Western Europe have experienced markedly different patterns of urban development, which suggests the complexity of attempting to explain



urban sprawl. Both regions experienced strong growth in income levels over the twentieth century. But the U.S. urban landscape resulted from a combination of car purchases, large public investments in road infrastructure, limited public investment in central cities, the existence of much population heterogeneity within cities, and low cultural barriers to household mobility. Western Europe, on the other hand, may be viewed as investing relatively more in public transportation within cities, expending greater resources on maintaining central city amenities and developing within a culture that is less willing to consider residential mobility as an avenue to improve household welfare -- all within cities that are more homogeneous in terms of population characteristics. Both as a cause and as a result of its European-style decisions, Germany, for example, took until 1970 to reach the same level of car ownership that the United States had reached in 1920 (Glaeser and Kahn, 2003).

While we seek in this paper to address only the issue of urban sprawl in the United States, we suspect that greater insight into the causes of urban sprawl within the United States could be obtained from a better understanding of why cities in other developed societies look very different. A notable beginning for such an exploration is offered by Brueckner, Thisse and Zenou (1999), who focus on amenity explanations for the difference between European and U.S. cities. A more general framework may yet explain the stereotypical European and U.S. city outcomes as two possible urban equilibria resulting from different underlying primitives faced by different cities.

### **Urban Sprawl: Negative Externalities and Equity Issues**

Those concerned about growing city footprints and lower city densities, or urban sprawl, associate these urban changes with a panoply of perceived ills in the urban and exurban landscape: the loss of open space, urban decay, unsightly strip mall developments, urban air and water pollution, traffic congestion, low-density housing developments, the loss of a sense of community, patchwork housing developments in the midst of agricultural land, increasing reliance on the automobile, the separation of residential and work locations, and a general spreading of urbanized development across the landscape. For economists, this range of objections is difficult to evaluate. Certain topics in the list, like traffic congestion, are well-recognized as possibly involving market and policy failures. Other arguments seem more rhetorical than real. It seems unlikely, for instance, that advocates of anti-sprawl measures worry primarily about the encroachment of city footprints on farm land, especially in light of the fact that forests encroach significantly more on farmland than do cities (Glaeser and Kahn, 2003). Other arguments like "reliance on the automobile" seem more like aesthetic judgements, favorable for some people and unfavorable for others, that don't enter an economist's social welfare calculation in an obvious way. Still other arguments, like a greater quantity of low-density housing, may point to potential environmental and other externalities but also suggest social benefits of urban sprawl since they involve a greater consumption of land and housing.

Our strategy here is not to attempt to compile an overall list of benefits and costs of urban sprawl but rather to focus on those that emerge most directly from the monocentric city and Tiebout models discussed above. Within the context of the monocentric city model, Brueckner (2000, 2001) identifies three major potential externalities that developers may

fail to internalize: traffic congestion (including its implications for pollution), the loss of open space at the urban fringe, and unrecovered infrastructure costs associated with new low-density development. We forego a detailed discussion of unrecovered infrastructure costs given the ease with which policy can address this through appropriately set “impact fees” charged to land developers, and we begin with discussions of traffic congestion and pollution. Within this discussion, we point out that relationships which may appear clear-cut – between city densities, congestion and pollution as well as between policy prescriptions and consequences – become less clear as the analysis escapes the strict confines of the monocentric city model. This lesson remains important in our discussion of landscape amenities and open space that follows. We conclude this section with a discussion of externalities and equity concerns that do not arise in the monocentric city model but appear when sprawl is viewed through the lens of the Tiebout model.

### Traffic Congestion and Sprawl

Empirical estimates suggest that traffic congestion in the United States causes substantial costs to urban residents. For instance a recent study of 75 U.S. urban areas finds that the total increase in commuting costs from congestion in these areas is \$69.5 billion, an average of \$520 per person (Shrank and Lomax, 2003). Of these costs, \$60 billion are attributed to the time costs associated with longer commute times and \$9.5 billion are attributed to increased fuel consumption resulting from congestion-induced drops in fuel efficiency.

Urban economists speak with a fairly unified voice on the issue of congestion – usually prescribing peak-load taxes and/or toll roads as the appropriate method for internalizing congestion externalities. We have two concerns regarding this conventional wisdom. First, we are concerned, as is suggested by Downs (1999), that Americans may not be willing to incur the levels of congestion taxes that would be required to make meaningful reductions in peak-hour traffic.<sup>4</sup> Our second concern is related to the distinction between mobility and accessibility.

When congestion is treated as “the problem,” its inverse -- ease of mobility within the metropolitan area – emerges as the societal goal. However, it is far from clear that congestion (or its inverse – ease of mobility) is the key argument in household utility functions -- aside from the extent to which congestion relates directly to air pollution which we turn to in the next section. Easier mobility (or declining congestion) is, after all, valuable only to the extent to which it is synonymous with easier *access* to the goods and services valued by households within metropolitan areas. Ease of mobility and ease of access are thus far from synonymous. Levine and Garb (2002) argue, for instance, that increased metropolitan mobility could actually be associated with more time and money spent in travel in the long run because greater mobility facilitates even more decentralized metropolitan development. The key insight here is that -- ignoring pollution externalities -- we should only value reduced congestion (increased ease of mobility) to the extent that it leads to an increase in accessibility over the long run. Spending the same amount of time driving longer distances on uncongested roads is a dubious social gain.



Replacing the congestion metric with an accessibility metric, however, highlights the importance of understanding the impact of congestion policies on land use patterns. Policies that facilitate compactness of urban development, mixed land uses, or development clustered around high quality public transportation could all increase access without reducing congestion per se. The impact of congestion taxes on urban structure, for instance, is not clear. Yinger (1993) argues that imposition of congestion pricing would lead to a denser urban area. Levine and Garb (2002), on the other hand, suggest that over time congestion pricing that reduces the number of trips into the urban core can lead to decreased accessibility as businesses respond to decreased consumer traffic by relocating to more remote locations only accessible by car. Similarly, Rufolo and Bianco (1998) argue that the costs of implementation of congestion pricing will likely result in taxation of only the most congested roads – typically those connecting suburbs to the inner city. Because locations of businesses and commuters are endogenous in the long run, raising the cost of commuting to the inner city without raising the cost of commuting between suburban locations may induce a decentralization of activity – and an increase in sprawl.

### City Density, Urban Sprawl and Pollution

The link between urban sprawl and air pollution has two components: increases in emissions per mile traveled related to traffic congestion, and increases in vehicle miles travelled, which are linked to lower density development. The link between vehicle traffic and air pollution is clear. In the United States in 2001, on-road vehicles accounted for 37 percent of total nitrogen oxides which play a major role in the formation of ground-level ozone, particulate matter, haze and acid rain; 27 percent of volatile organic compounds, which react with nitrogen oxides to form ground-level ozone; and 62 percent of total emissions of carbon monoxide, which is a particular threat for individuals who suffer from cardiovascular disease (USEPA National Emissions Inventory, <<http://www.epa.gov/ttn/chief/trends>>). Also, between 1970 and 2001, total vehicle miles traveled increased 151 percent from 1.1 trillion miles to 2.8 trillion miles. Over the same period, miles traveled by passenger cars and motorcycles increased by over 75 percent (from 920 billion miles to 1.63 trillion) (U.S. Department of Transportation National Highway Statistics 2001).

While these numbers are striking, the links between sprawl and pollution are not as straightforward as it might seem. First, the impact of reduced highway congestion on air pollution is a function of the type of congestion. A U-shaped relationship exists between emissions such as nitrogen oxides per vehicle mile travelled and car speed, with average speeds of 30 MPH representing the trough of the U (Final Facility Specific Speed Correction Factors, USEPA EPA420-R-01-06. November, 2001). Second, the link between development density and car pollution is similarly unclear. As discussed above, density itself is not necessarily related to spatial accessibility – implying that vehicle miles traveled per individual within a metropolitan area depend as much on micro-features of the area as on overall density. For instance, the appearance of edge cities, while leading to a less dense metropolitan area, may also result in a decline in commuting (and thus vehicle miles travelled per individual) as jobs are more decentralized within the urban area (Glaeser and Kahn, 2003).

It is therefore not surprising that the empirical literature on the link of air pollution and city density yields no strong results. The majority of the empirical studies claiming to document a clear link between density and pollution fail to account for other important variables such as income and household demographics. Those studies that use micro-level data and attempt to control for these other factors generally conclude that the relationship between density and travel behavior is weaker and less certain than is often claimed in the popular press (Pickrell, 1999, Crane, 2001). As pointed out by Glaeser and Kahn (2003), U.S. urban air pollution on the whole has been on the decline since the 1970s, with increasingly cleaner cars supplanting older, more polluting. Thus, even if there were a clearer link between sprawl and pollution, policies unrelated to sprawl seem to be more effective in addressing urban pollution problems.

### Landscape Amenities and Open Space within Metropolitan Areas

The rapid urbanization of open space is often identified as one of the key problems associated with sprawl. Using satellite imagery data and aerial photography, Burchfield et al. (2003) estimate that the urbanized percentage of the U.S. landscape increased by 47.7 percent between 1976 and 1992 – an annualized rate of 2.48 percent. However, as the authors of the study point out, from a policy perspective these numbers can be misleading. While urbanization is increasing rapidly, the actual percentage of the U.S. that is urbanized remains quite small – only accounting for 1.92 percent of the entire U.S. land area as of 1992. Given the very low percentage of the U.S. landscape that is urbanized, why is the loss of open space receiving such prominence in the debate over sprawl?

When used to analyze questions of urban sprawl, monocentric city models typically do not include an explicit specification of household tastes for open space and are thus not well-equipped to analyze welfare losses from sprawl related loss of open space. In addition, the macro-nature of the monocentric city model limits the definition of open space to rural land at the urban fringe. However, evidence from the empirical literature suggests that open space *within* suburbs is significantly more important to households than open space at the urban fringe. A wide array of empirical work, aided greatly by the recent advent of Geographic Information System (GIS) technology, has investigated how housing prices are related to proximity to various uses – and thus how households value different types of land uses. The work on landscape amenities has identified positive values for proximity to public parks, privately owned open space, the natural land cover immediately surrounding household locations and access to natural views.<sup>5</sup> But the relationship between open space and housing prices seems to be non-linear. Geoghegan et al. (1997), for instance, find that housing prices are increasing in the percent of open space within a 0.1 km ring and decreasing in the percent of open space within 1 km ring surrounding the house. Acharya and Bennett (2001) find that house prices are increasing in the percentage of open space at a decreasing rate within both a 1/4 mile and a 1 mile radius, and Walsh (2003) finds open space to provide positive amenity flows at high levels of development and negative flows at low levels of development.

One obvious interpretation of these results is that there exists a trade-off between open space amenities and access to commercial services. In particular, if open space proximity is valued highly but decays very quickly with distance, while the value of accessibility to commercial areas decays more slowly, the pattern observed in the data would arise. The evidence then seems to suggest that households value access to commercial activity (that is, not too much open space everywhere around them) but prefer open space in the immediate vicinity of their residence. There seems to be no particular reason to believe that households in general place great value on open space on the urban fringe.

While it is important to recognize that the studies cited above cannot identify amenity values that are shared equally across the land market (since such value would not appear in differential sales prices of land), the empirical analysis suggests that open space is, at least to some extent, a quite local public good. As a result, the within-metropolitan area creation of open space can be studied within the context of Tiebout-type models in which developers have an incentive to create local public goods (such as parks, small lakes and recreation areas) to drive up prices for land they are developing near such spaces. The more local the nature of the open space amenities and the larger the scale of development, the stronger will be the incentives for private developers to provide efficient levels of open space.<sup>6</sup> When the scale of individual developments is not large enough to capture spillovers associated with open space amenities (such as dense development in inner cities), local political processes may combine with city-wide policies aimed at internalizing local externalities in the creation and maintenance of open spaces through land purchases (such as large central city parks). In other cases, institutions such as private land trusts and homeowners' associations may emerge and internalize the same local open space externalities.

Inclusion of different types of open space within models of the urban landscape – and linkage of such open space to empirically well-grounded motivations for private land development -- thus becomes important to a fuller treatment of the relationship between sprawl and welfare. An explicit introduction of preferences for open space into the urban literature is furthermore likely to alter the policy prescriptions beyond recommendations of impact fees and the use of land taxes. While such policy instruments may indeed be key to preserving open space at the urban fringe, it is unclear how they would address what appears to be the much larger concern for the preservation and creation of smaller open spaces *within* sprawling city boundaries.

Walsh (2003) provides a glimpse at the complexity of urban land use policy in a spatially explicit general equilibrium model of land markets in North Carolina's Wake County. In this model, an urban growth boundary is shown to freeze development in a relatively low land price region and thus result in increased urban density. An alternative policy that freezes development in Wake County's fastest-growing suburb, on the other hand, decreases development density. Yet both types of policies are advocated by proponents of anti-sprawl measures as they seek to use policy to create denser cities. Walsh also points out that different land purchase policies – say, the choice between protecting forested wetlands, setting aside large undeveloped tracts, and establishing urban parks – might have

quite different effects on the density and open space decisions of private land developers. These issues invite further research.

### Tiebout Sorting, Peer Externalities and “Categorical Equity”

The prevalence of Tiebout sorting, documented in the extensive empirical local public finance literature alluded to earlier, and the greater scope for such sorting in sprawling cities, has both efficiency and equity implications for sprawling metropolitan areas. In our discussion of open space, for instance, we suggested that open space might be viewed as a local public good – and that land developers have incentives to provide such goods in various quantities across the urban landscape to meet different consumer demands. While such a market process may generate more efficient distributions of open space within the urban landscape, it would no doubt result in differential levels of open space consumption, correlated in part with household income, as consumers sort according to their demand for open space. Yet some might object that such differential consumption, while efficient, might carry with it the burden of an inequitable distribution of access to open space, with poorer households having less access than those with greater means.

Such distributional considerations take on particular importance when high moral or legal claim is given to the proposition that at least a high minimum level of certain local public goods has to be delivered to all. While access to open space may or may not qualify for such a claim, other goods clearly do, like schooling, protection from crime and environmental hazards, and access to neighborhoods that are broadly functional. Equity considerations that focus on such particular categories of local public goods will henceforth be referred to as “categorical equity” considerations.

Categorical equity concerns in sprawling cities arise in part because the ability to move out of central cities and into suburbs, and the ability to choose among suburbs, is not uniform across population and income groups. Within U.S. metropolitan areas, the poor generally live in central cities while middle-income individuals live in suburbs (Margo, 1992; Mieszkowski and Mills, 1993; Mills and Lubuele, 1997; Glaeser, Kahn and Rappaport, 2000). While over 17 percent of central city residents are poor, the same is true for fewer than 7 percent of suburban residents. Public transportation has been identified as an important explanatory variable (Glaeser, Kahn and Rappaport, 2000; LeRoy and Sonstelie, 1983) – suggesting that for those using public transportation, the car-centered suburbs may simply not be an option. Furthermore, political boundaries within metropolitan areas correlate strongly with differences in poverty rates across those areas – suggesting politically created distortions through zoning, housing markets and local public finance considerations consistent with Tiebout’s model (Glaeser, Kahn and Rappaport, 2000). This documented geographic separation and lower mobility of poorer households is likely to give rise to a variety of social problems in poor areas and is of independent concern to the extent to which jobs are sprawling out of cities along with residences (Mills and Lubuele, 1997; Glaeser and Kahn, 2003).

When the quality of locally provided goods is primarily a function of public investment, categorical equity concerns from Tiebout sorting can be addressed straightforwardly

through provision of public resources by state or federal governments to insure the desirable level of local public provision. However, the quality of local public goods may rely less on public financing than on non-financial inputs that derive directly from the composition of local populations. In the case of education, for instance, family and peer externalities may play a powerful role in producing school quality (Gaviria and Raphael, 2001; Hanushek et al., 2003; Vigdor and Nechyba, 2003). It is well-known that public school quality differs across neighborhoods and districts even when observable school inputs such as per pupil spending are equalized. Per pupil spending in California has, for instance, been largely equalized with no evidence for an appreciable narrowing of school quality differences across districts. Analogously, mere spending on public safety does not lead to equal levels of protection from crime, nor does equal public investment in basic infrastructure result in uniformly functional neighborhoods. Similar insights on the potential importance of peer and family effects are emerging from the literature on local crime rates and other neighborhood characteristics (Katz et al., 2001; Solon, Page and Duncan, 2000, Chase-Landsdale et al., 1997). In each of these cases, the level of the public good depends critically on the characteristics of the local population that is being served by public expenditures on the good – giving rise to local peer and neighborhood externalities that shape the true local public good levels.

Models in which mobile household with such different school peer externalities choose residences within existing housing markets that are divided into multiple school districts have been developed by Nechyba (1999, 2000, 2002, 2003a,b,c,e) and Ferreyra (2002). These models are then calibrated or estimated with data from various cities. For example, Nechyba (2002, 2003a) shows that, when school attendance is linked to place of residence, the value of high-quality schools is capitalized into housing prices and considerable residential segregation occurs by income. High quality schools within the model are in large part determined by family and peer effects, and as a result, segregation occurs even if all schools receive the same funding on a per student basis. Such frameworks could equally well model other local public goods that depend on local population externalities – with much the same result.

One general policy conclusion from this work is that sorting effects are highly dependent on institutional frameworks and the micro-structure of urban housing markets. Sorting can lead to increased segregation under certain institutional and policy environments, but these same forces can equally well support desegregation – and potential movement back into central cities – under other policies. For instance, Nechyba (2000, 2003a,b,c,e) and Ferreyra (2002) demonstrate the potential for policies aimed at weakening the link between residential and school choices to impact the incentives of different demographic groups to segregate within a metropolitan environment. Such policies may involve either a fostering of private school markets through vouchers or an increase in public school choice through charter and magnet schools. In each case, middle-income parents in residence-based public school systems have an incentive to locate in lower income districts, to take advantage of low housing prices while avoiding poor public schools by choosing alternatives. To the extent to which Tiebout forces have *pushed* middle income families out of central cities and thus added to sprawl, these same forces can thus, under the right policies, lead to a reduction in sprawl.



Similarly, early efforts to create public housing tended to reinforce segregation patterns, while later efforts to channel resources into more dispersed affordable housing opportunities reversed increasing poverty concentrations within cities in the last decade (Jargowsky, 2003). Experiments such as the “Moving to Opportunity” policies that give residents of low income housing projects in inner cities the option of moving to middle income neighborhoods can thus create more equitable distributions of all types of public services, as such policies attempt to directly shape peer and neighborhood characteristics away from concentrations that reinforce existing inequities within metropolitan areas.<sup>7</sup> Housing policy, like school policy, can thus provide institutional frameworks for within-city development that lead to greater or less segregation – and with it greater or less categorical inequity.

Research that models the links between local public institutions and urban characteristics can clarify how current demographic distributions of populations in cities have emerged and persist. For instance, recent empirical structural work by Bayer, McMillan and Reuben (2002) can differentiate between the different channels through which racial sorting into schools and neighborhoods arises. In their work, housing markets, job locations, and preferences for residential homogeneity – not differences in tastes for quality of education – represent the crucial explanatory forces for racial school segregation patterns observed in the data. This result emphasizes that urban economies arise from a blend of decisions about housing, jobs, schooling, and neighborhoods—and of public and private institutions that shape each of these decisions. Much remains to be done as economists create better connections in their investigations of the impact of urban sprawl on neighborhoods within cities and as their models become more grounded in data and thus more policy relevant.

### **Some Policy Trade-Offs and a New Research Agenda**

Urban sprawl is a multidimensional issue, with multiple causes and effects. Policy trade-offs emerge for three different reasons: First, a tension often emerges between the goal of limiting the expansion of cities (analyzed within the monocentric city model) and the goal of not further disadvantaging the poor who are currently struggling within those cities (often viewed through the lens of the Tiebout model). Little progress can be made in better understanding these tradeoffs without urban models that incorporate the possibility of tension between these goals. Second, urban policy makers face the difficult challenge of operating in what economists refer to as a “second-best” world – a world in which multiple distortions, some arising from market failures and others from existing government policy – already exist. For instance, household location decisions in urban areas are distorted by the link of public schools to residential housing markets – and this existing distortion needs to be taken into account as policies to address sprawl are contemplated. Finally, various trade-offs emerge specifically from the micro aspects of particular social problems – such as, for instance, the tradeoffs that emerge once one learns from environmental economists about the U-shaped relationship between car speed and emissions of pollution.



## Likely Policy Trade-offs in Considering Anti-Sprawl Measures

Since the research integrating elements of the monocentric city model with Tiebout models and models of edge cities is still in its infancy, since second-best considerations are often absent from urban models, and since important structural relationships are not always carefully included in the analysis, it is difficult to make firm policy recommendations based on present research. Nevertheless, we can begin to identify some examples of likely additional policy trade-offs that might emerge in a more thorough analysis for policies often advocated by those concerned with urban sprawl.

Consider, for instance, the common policy prescriptions arising from monocentric city analysis of sprawl that focuses on sub-optimal city densities causing cities to encroach on open space at the urban fringe and leading to excessive congestion. Developers could simply be charged appropriately set impact fees to internalize the cost of public infrastructure from loss of open space at the urban fringe, and congestion taxes could be imposed to internalize externalities from individual location and commuting decisions. However, policies aimed at raising the price of living at the urban fringe, such as impact fees and congestion taxes, may have unintended segregating consequences within the metropolitan area. Such policies might reduce the footprint of cities at the cost of making it relatively more difficult for lower-income households to move to the suburbs. The trade-off that might emerge then involves smaller urban distances to suburbs on the one hand and higher concentrations of the poor in central cities on the other – with predictable consequences for categorical equity within the city. A real analysis of this potential tradeoff awaits more rigorous modeling of the relevant forces that generate such tradeoffs.

Similarly, mass transit accessible to lower-income residents of inner cities is often advocated as a possible policy prescription to minimize the effects of sprawl, but mass transit also provides incentives for city footprints to expand along the rays of the mass transit system, with commuters driving to outer mass transit stations and then commuting on the train or bus. Furthermore, stretching mass transit systems into suburbs is likely to lead to residential income sorting, with the poor concentrated along mass transit access points. This sorting may lead to lower concentration of poverty in inner cities while increasing income segregation outside the central city. Alternative methods of addressing transportation costs for the poor, ranging from subsidized car purchases for the poor (Glaser and Kahn, 2003) to subsidized van or taxi rides might help to open opportunities throughout cities. However, transportation costs are only one of the reasons for spatial sorting, and mass transit may have only a small effect on current levels of segregation. Again, it is difficult to say more without a better and economically more relevant model of cities and suburbs.

Urban growth boundaries have also become a popular proposal to limit sprawl over the past two decades, but while these boundaries can limit the footprint of cities, they simultaneously drive up housing costs and thus impose the largest burdens on lower income households (Quigley and Raphael, 2004). They could furthermore worsen the problems associated with sprawl if designed in such a way as to push sprawl beyond growth boundaries around cities. Similarly, while local zoning can create the types of

neighborhood open spaces that seem to matter most to households, an extensive local public finance literature suggests that powerful local zoning boards may well be used as exclusionary tools to keep local populations segregated (Hamilton, 1975; Fischel, 2001a,b), and such boards are unlikely to internalize potential externalities from loss of open space at the urban fringe.

### Toward a New Approach

Ultimately, we believe that fully understanding the complex set of trade-offs raised by these interconnected policy concerns will require a new integrated modeling approach to the urban problem. While the urban economics literature with its focus on the trade-off between transportation costs and land rents has succeeded in identifying the primary causes for the sprawling of cities in the twentieth century, its models often glance over the micro aspects of how city landscapes within expanding boundaries evolve. Environmental economists have identified the components of the neighborhood landscape that affect housing prices, but have provided less insight to the way that policy activities and land markets interact to give rise to these localized amenities. At the same time, a rich local public finance literature can potentially yield worthwhile insights into the problems faced within sprawling cities, but few local public finance models have incorporated the insights from urban models that make them relevant for an informed discussion of urban sprawl.<sup>8</sup>

A more integrated approach, perhaps built upon recent advances in structural locational equilibrium modeling, could deal more fully with the details of how sprawl affects households and families. Recent innovations in the structural modeling of residential housing and land markets embodied by Epple and Sieg (1999, 2001), Bayer et al. (2002), Walsh (2003), Nechyba (1999) and Fereyra (2002) provide potentially fertile ground. What is unlikely to emerge from such an approach, however, is a set of simple policy prescriptions to combat urban sprawl. Cities and suburbs are complicated economies, and most policies are likely to give rise to similarly complicated tradeoffs. Nevertheless, it is difficult to see how appropriate policy responses to growing concerns about the sprawl of U.S. cities can be formulated without a clearer understanding of these underlying trade-offs.