What is urban economics? I will refrain from torturing myself with this issue. Any definition will be too wide for some and too narrow for others. My definition will be simple, but appropriate for my interests and background: urban economics is scholarly writing that has specifically urban content and contains specifically economic analysis. Everyone will notice that my definition leaves room for judgment, which I will exercise freely.

My concern, as the title indicates, is with themes and only incidentally with authors. I propose to ask: What subjects has urban economics been concerned with, and how have they changed through time? What has been the division of labor between theoretical and applied research? And most important, with what subjects has progress been made, and what significant problems are still outstanding?

Where does one look for urban economics publications? One journal, the Journal of Urban Economics, whose twenty-fifth birthday 1999 was, has exactly that title. Two or three other journals have those or similar words in their titles. Several scholarly journals are devoted each to housing, transportation, and analysis of local government, and each such subject has a continuum of journals and magazines that blend into trade publications. Per-
haps the most difficult distinction is between real estate analysis and urban economic analysis. Three or four scholarly journals are devoted to real estate and another three or four to housing and housing finance analysis. In addition, many journals are devoted to related subjects: urban planning, urban geography, regional science and economics, urban government, and so forth. Journals devoted to human resources, labor relations, poverty, and racial problems often contain urban economics articles. And of course urban economics research appears in general economics journals and in books, monographs, and reports of various kinds. Finally, as in the case of international capital flows, there is a home bias in a survey of urban economics. My home bias tells me that most urban economics is published in, or at least promptly translated into, English. But home bias is also geographical. Much urban economics is published in India, Japan, Korea, and other countries. One could certainly write a long survey of urban economics publications from or about India.

By definition, urban economics did not begin before economics, which I date with Adam Smith. In the nineteenth and early twentieth centuries, Johan Heinrich von Thünen, Walter Christaller, Henry George, August Lösch, and Alfred Marshall published more or less systematic urban economic analyses.2 In the first half of the twentieth century, the list is long enough to require a separate survey of the period. This particular survey arbitrarily covers only the second half of the twentieth century. Another home bias is that, professionally, I have lived through that period.

Many fine economists have written on urban economics. Some have been part-time urban economists, many have been full-time. Among the part-time workers have been several Nobel laureates. Beyond doubt, urban economics has benefited greatly from the writings of its part-time participants. Simon Kuznets had many insightful observations about urban growth in developing countries. Gary Becker provided the first definitive analysis of the economics of discrimination. A paper by Robert Solow and William Vickrey, and one by James A. Mirrlees, are among the finest and most influential in urban economics. Solow has contributed three or four excellent papers to urban economics, but I believe that the Mirrlees paper is the only one he has written on the subject. Vickrey wrote many papers on urban transportation pricing and investment. Stretching the boundaries of the subject would permit inclusion of Paul Samuelson, whose analysis, certainly spatial if not urban, has

2. Smith (1776), von Thünen (1826); Christaller (1933); George (1874); Lösch (1954); Marshall (1920).
influenced urban research. Most recent is an important paper, I believe his first in urban economics, by Robert Lucas.  

Other urban analyses by Nobel laureates appear somewhat churlish to this reviewer. James Buchanan and Gordon Tullock advocated that entrance fees to urban areas be charged to migrants to compensate for the externalities they generate in urban areas. Migrants do not generate greater externalities than those born in urban areas. More important, externalities imply that governments have not gotten the prices right for urban transportation, polluting discharges, and other local issues; these problems should be attacked directly. Most important, large urban areas generate positive externalities, which are discussed later. And W. Arthur Lewis railed against infrastructure investments in large cities in developing countries, because high land values there made infrastructure excessively expensive, contributing to external debt. He failed to note that land rents are transfers, not resource uses; that land rents are high in big cities, at every stage of development, because land is productive there; and that infrastructure is justified to the extent that it increases land values by more than its cost. Infrastructure investments can, to an approximation, be financed by taxing the resulting increments in land values.  

What Do Urban Economists Do?  

Urban economists do what other economists do: theoretical and applied research. Theoretical analysis of urban resource allocation started in the 1960s with what is usually referred to as the Alonso-Mills-Muth model. Early papers struggled to make sense of simple yet pervasive characteristics of urban areas: greater centrality of employment than of population, declining land values  

3. Kuznets (1955); Becker (1957); Solow and Vickrey (1971); Mirrlees (1972); on urban transportation pricing and investment, notably, Vickrey (1963); Samuelson (1952); Lucas (1999).  


5. See William Alonso (1964); Edwin Mills (1967); Richard Muth (1967). A personal note: My first paper in urban economics, Mills (1967), which has received its share of plaudits, was written in the summer of 1966 while I visited the Rand Corporation. At that time, I had read several early applied studies, Alonso (1964), and a draft that became Muth (1967). I had never heard of von Thünen, the legitimate nineteenth-century precursor to the theoretical models of the 1960s. My immediate inspiration was Clark (1951) and, most important, Solow (1956). It occurred to me to make Solow’s analysis reflect distance from an urban center instead of time. Add a reason for there to be a center and a representation of transportation costs, and the analytical results follow easily.
and population densities with distance from urban centers, and apparently uni-
versal flattening of density-distance functions as time passes and as one moves
from small to large urban areas. Early theoretical models provided important
and testable insights, some of which have withstood the test of time, but most
of which were clumsy special cases. Compare any early theoretical paper with
the abstract, general, and unified models in the survey by Masahisa Fujita or
the surveys by Mahlon R. Straszheim, Konrad Stahl, and Jan Brueckner in
volume 2 of the *Handbook of Regional and Urban Economics.*

Until the late 1960s, there were few high-quality applied-research publi-
cations in urban economics. A fine book on housing by Leo Grebler, David
Blank, and Louis Winnick, and another by Margaret Reid, complete the list
along with a few journal publications. Much of the research until the late
1960s was financed by or undertaken at Resources for the Future (RFF). During
the early 1960s, RFF placed excessive emphasis on regional accounts, undoubt-
edly reflecting endemic faith in input-output analysis at the time.

In the late 1960s the dam broke, and the quantity and quality of applied
urban economic research increased quickly. Richard Muth’s *Cities and Hous-
ing* was certainly the finest combination of theoretical and applied urban
economic research prior to 1970. Jerome Rothenberg first demonstrated
appropriate welfare-economics evaluations of urban renewal, the most promi-
nent government urban program of the time. An idea of the increase in the
quality and quantity of applied urban economic research from the late 1960s
to the late 1970s can be obtained by comparing the survey edited by Peter
Mieszkowski and Mahlon Straszheim in 1979 with the 1968 survey edited by
Harvey Perloff and Lowdon Wingo.

The biblical generation since the late 1970s has seen a flood of publica-
tions on applied urban economic research. The growth rate can be ascertained
by comparing the publications reviewed and referenced in volumes 2 and 3
of the *Handbook of Regional and Urban Economics.* A guess is that length-
ening bibliographies reflect not only growth in publications but also decreasing
costs of computer-generated bibliographies. Specific topics are commented

7. See Goldstein and Moses (1973).
8. Grebler, Blank, and Winnick (1956); Reid (1962).
on below. The growth in applied urban economics publications, which must have been close to a compound 10 percent annual rate, has resulted from obvious changes. First and foremost has been the decreasing cost of computing, especially since personal computers became available to almost every scholar during the 1980s. Closely related has been the dramatic increase in the quantity, quality, and availability of urban data. Better and more easily available econometric techniques and software have also spurred the revolution. A by-product that has frequently been referenced has been the rapid growth of publications coauthored by scholars in different institutions and countries, undoubtedly entailing economies of scale and scope. A corollary is that technological unemployment must have increased among research assistants; perhaps some of these individuals should be employed to verify published empirical results.

**Topical Survey**

Publications can be assigned to subject matter categories in many ways. I have compiled a classification of papers that appeared in the *Journal of Urban Economics* in 1978, 1988, and 1998. The first volume of the *Journal of Urban Economics* was published in 1974. The classification is in table 1. I believe that the nine categories are intuitively appealing and correspond to the way that urban economists tend to think about subfields in their specialty. The difficulty of classifying papers increases with the fineness of the categories. For that reason and because the total sample is only 129 papers, I have used the nine-subject classification. The only difficult task in my classification was assigning papers between categories 1 and 2. Many spatial analyses are primarily about housing and many housing analyses are also spatial. My classification has been subjective and some readers might classify some papers differently.

Several other comments can be offered. The number of papers per year in the *Journal* has increased only modestly, but the number of papers on similar subjects but in other journals has increased greatly. That is particularly true of housing, for which the number of outlets for scholarly papers has increased rapidly, but the same is true for “real estate,” for which the number of high-quality papers and journals has also increased rapidly.

Category 8 includes all papers on relationships between an urban area and other places—for example, papers on rural-urban or interurban migration, or
papers on the size and spatial distributions of urban areas. Category 9 includes all papers on countries that are not members of the Organization for Economic Co-operation and Development (non-OECD). I put these in a separate category regardless of subject category because good papers on those countries have increased rapidly, although the Journal is not a primary outlet. The same is true of papers on transportation, crime, and education. Finally, I believe that the Journal, under both of its editors, has always welcomed submissions of papers, theoretical or applied, on all subjects that could reasonably be called urban economics.

The strongest suggestion in table 1 is that there is a trend away from spatial analysis. I believe this results from two causes. First is that the easy problems in spatial analysis, specifically the monocentric model, have been solved. Multicentric models are much more realistic but much more difficult. Second is the trend toward more and better applied papers referred to on page 4. Housing papers have been consistently numerous, despite the growth of competing outlets—again a tribute both to the importance of and interest in the subject and to the growth of applied work. Analyses of government behavior in urban areas, most but not all of these concerning local governments, have constituted the big growth sector in the table. The small growth of papers in category 4 surprised the author. The shrinkage of attention to crime probably reflects economists’ overall attitudes. Good papers on economic aspects of education have certainly increased since the late 1970s, and many are published in other journals. Likewise, growth in the number of


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Spatial analysis</td>
<td>30.8</td>
<td>22.7</td>
<td>8.7</td>
</tr>
<tr>
<td>2. Housing</td>
<td>20.5</td>
<td>27.3</td>
<td>19.6</td>
</tr>
<tr>
<td>3. Government sector</td>
<td>20.5</td>
<td>22.7</td>
<td>30.4</td>
</tr>
<tr>
<td>4. Labor, poverty, race</td>
<td>10.3</td>
<td>4.5</td>
<td>13.0</td>
</tr>
<tr>
<td>5. Crime</td>
<td>5.1</td>
<td>0.0</td>
<td>2.2</td>
</tr>
<tr>
<td>6. Transportation</td>
<td>5.1</td>
<td>2.3</td>
<td>6.5</td>
</tr>
<tr>
<td>7. Education</td>
<td>0.0</td>
<td>4.5</td>
<td>4.3</td>
</tr>
<tr>
<td>8. Interurban, rural-urban</td>
<td>7.7</td>
<td>9.1</td>
<td>8.7</td>
</tr>
<tr>
<td>9. Developing countries</td>
<td>0.0</td>
<td>6.8</td>
<td>6.5</td>
</tr>
<tr>
<td>Total number of papers</td>
<td>39</td>
<td>44</td>
<td>46</td>
</tr>
</tbody>
</table>

Source: Author’s calculations.
papers on developing countries probably reflects the rising quality of papers on the subject. Needless to say, my sample is too small for the results to be more than suggestive; better subject-matter analysis would require a survey of more journals, which would require a more subtle classification between urban and other papers and, for some journals, between economics and other specialties.

If one regards the distinction between categories one and two as arbitrary and adds the entries together, the percentage share declines from 51.3 to 50 to 28.3 during the twenty-year period. That decline is counterbalanced by a 50 percent increase in the share of papers related to government behavior and by modest increases in the shares of several other categories. What accounts for the increase in papers devoted to analysis of government behavior? Certainly not a decrease in alternative outlets available. Nor does the reason appear to be a dramatic increase in data available. Of course, more government data are available in electronic form, but that is as true of data related to the private sector. I believe the reason is a dramatic increase in analyses that relate to the positive theory of government. Many more economists now analyze how governments actually behave instead of ending papers with inane comments on “policy implications.” At the local level, for example, governments have vastly overregulated housing and transportation for parochial reasons; yet until recently economists continued to identify “market imperfections” and to urge additional government intervention without asking whether the history of government involvement in such areas makes it likely that additional intervention will make things better or worse. Analyses of the motivation and behavior of government agents and the relations between government and private agents are the keys to progress in research on governments.

**Progress in Urban Economics**

In this section I trace the development of several specific subjects of urban economic research. My goal has been to discuss important subjects on which there has been a substantial and persistent literature displaying a coherent line of development. In each subsection, I finish with my perceptions of important but inadequately studied topics.

Selection of topics is inevitably subjective, reflecting my reading, my interests, and my judgments, and probably reflecting some home bias.
Formation, Size, and Spatial Distributions of Urban Areas

This subsection reviews several closely related topics: why urban areas are formed, why they grow, rural-urban migration, and the size and spatial distributions of urban areas.

What is an urban area? In most countries, definitions are similar: a contiguous area with a population of at least 2,500 to 25,000 (occasionally 50,000) people in which a majority of resident workers are in nonagricultural sectors, and possibly having a minimum population density. An important peculiarity of U.S. data is that, although the minimum urban population is only 2,500 people and about 2 percent of workers work on farms, about 27 percent of the population of workers is rural. The apparent implication is that about 25 percent of workers are rural residents not working directly on farms. Furthermore, most of the 25 percent do not appear to be engaged in farm-related activities.14 No other country appears to have a rural nonfarm workforce of 20 percent. The peculiar U.S. data are probably mostly illusion. Outside metropolitan areas, census employment data are provided by place of residence. In metropolitan areas, data are by both place of residence and place of work. We also know that a few more workers commute into than out of metropolitan areas on the way to work, but the difference is not large relative to the 20 percent of rural nonfarmworker residents. The reconciliation probably is that most of the rural and non–farm-resident workers commute to jobs in relatively small nearby urban but not metropolitan communities.

Why do urban areas exist? The simplest answer is the correct answer. Urban areas exist because proximity among diverse activities economizes on the cost of moving goods, people, and messages. This characterization covers early commercial urban areas, medieval fortress urban areas, medieval cathedral-based centers, and the largest metropolitan complexes in the late twentieth century. Even if scale economies internal to a single manufacturing plant justified employment of 2,500 workers, the location would not be an urban area unless commuting costs induced the employees to live close to the plant. Likewise, even if people’s sole interest were in social interaction with each other, 2,500 people would not form an urban area unless social interaction were more costly at greater distances between residents. Alternatively, suppose that 2,500 fishermen who catch fish over a large part of the ocean live in a hamlet because it is at the best natural harbor in the region; then the reason for the urban cen-

ter is that the proximity of fishermen’s residences near the harbor makes the
costs of shipping fish from that place less than the costs of shipping them from
scattered sites.

Works by the great urban historians Paul Bairoch, William Cronon, Paul
Hohenberg and Lynn Lees, Adna Weber, Charles N. Glaab and A. Theodore
Brown, and many others, as well as those by most modern urban economists,
have all used this explanation of urban areas, mostly implicitly, but I have not
seen a succinct statement.\textsuperscript{15} Modern urban economists typically attribute the
advantages of urban areas to economies of scale, scope, and agglomeration.\textsuperscript{16}I have just shown that economies of scale, and by extension of scope, are not
sufficient reasons for urban areas. Although much progress has been made
during the last quarter century in understanding and measuring agglomera-
tion economies, I believe it remains the great mystique of urban economics.
The definition of \textit{agglomeration economies} is economies from proximate
locations or urban size that are external to individual firms. Thus agglomera-
tion economies do not include economies of scale and scope, which are internal
to firms.

Agglomeration economies is one of the few subjects in economics in which
measurement preceded theory. Several imaginative measurements were made
with U.S. data in the 1970s:\textsuperscript{17} these have since been repeated with better data
and estimation techniques for the United States and a few other countries.\textsuperscript{18}Although data and conceptual approaches vary greatly, conclusions are remark-
ably uniform: each doubling of the size of an urban area increases total factor
productivity by 5 to 10 percent. Although intuition suggests that there must
be a limit to agglomeration economies, no careful study seems to have pro-
vided convincing evidence of a limit.

What could account for this almost incredible finding? Several sources have
been prominently suggested. The economist Edgar Hoover suggested a prob-
abilistic phenomenon: if employers’ demands for similar kinds of labor are

\begin{thebibliography}{9}
\bibitem{15} Bairoch (1988); Cronon (1991); Hohenberg and Lees (1985); Weber (1899); Glaab and
Brown (1967).
\bibitem{16} For a survey, see the paper by Randall W. Eberts and Daniel P. McMillen in Cheshire
\bibitem{17} For references, see the paper by Eberts and McMillen in Cheshire and Mills, eds. (1999).
\bibitem{18} See Shukla (1996) for Indian estimates and for references to estimates for other coun-
tries. Agglomeration economies are mentioned in many studies of urban areas in developing
countries, often as a catch-all phrase. Remarkably enough, World Bank (1999), devoted to
“knowledge for development,” does not mention the term.
\end{thebibliography}
only imperfectly correlated, proximate locations can reduce per-worker unemployment rates and hence competitive wage rates.\textsuperscript{19}

This appears to be a reasonable source of agglomeration economies in that it does not result from scale or scope economies that are internal to particular plants, but after more than half a century I have not seen any estimation, and my intuition is that it is unimportant beyond an urban size of one or two million in population.

Urban economists’ recent predilection is to attribute agglomeration economies to “nonexcludable public goods” (a redundancy) such as access to a large labor pool or capital market or to “infrastructure” such as roads, power, water supply, and sewage treatment. Access to large input suppliers may have the statistical advantage just discussed, but it is otherwise difficult to imagine how large input markets could be sources of market failure. Infrastructure is subject to scale economies but is certainly excludable. Governments perversely underprice roads (see page 28) and most other products they produce, but metering by electronic means and pricing by fuel taxes or by metered travel are certainly methods of excluding “free riders” from use of such goods. Piped water, power, and sewage treatment are quintessential private goods. Fire protection can be priced by insurance, as health care typically is. The criminal justice system is almost inherently a public good, but evidence of substantial scale economies is scant. Education is excludable and is provided free of charge by governments mostly for equity reasons, and is priced at all levels by private providers. Local urban goods and services provided by governments may be underpriced and subject to scale or scope economies, but I am unable to identify any such services other than criminal justice that satisfy the technical definition of public goods.

The strongest case for a source of agglomeration economies is “knowledge spillovers.” The concept was prominent in the work of Alfred Marshall, is often attributed to Jane Jacobs, and has recently been the subject of measurement.\textsuperscript{20} Careful work requires a distinction between basic and applied research and development, spillovers being more effectively prevented (by patents, trademarks, company secrets, nondisclosure agreements, and other means) the farther toward the practical end of the spectrum one goes. One should also

\textsuperscript{19} Hoover (1948). Mills (1972) is sometimes given credit for this idea. Although I included Hoover (1948) in the chapter reading, I failed to include the acknowledgement, and the reference was omitted from later editions. I apologize.

\textsuperscript{20} Marshall (1920); Jacobs (1969); for recent surveys, see Helsley and Strange (1999) and the paper by Eberts and McMillen in Cheshire and Mills, eds. (1999).
distinguish between the generation and the application of new knowledge. Also important is the relationship, presumably complementary, between universities and industrial research centers. Nevertheless, the rapid growth of high-tech centers (most prominently Route 128, Silicon Valley, the Research Triangle, Austin, and Tsukuba in Japan) during recent decades provides a strong prima facie case that applied research and innovation benefit from proximity. It is worthy of note that all the above centers were on the peripheries of metropolitan areas when they started and that all reportedly have substantive interactions with nearby universities.

Of course, proximate locations of horizontally related firms do not prove that there is market failure. Urban firms are located near each other in many sectors (law, medicine, finance, advertising) in which no one suggests market failure. Firms locate near each other if some transactions benefit from proximity. That firms benefit from face-to-face communication accounts for extremely high densities of central-business-district employment, where land values may be at least an order of magnitude greater than their value twenty-five to forty miles away and no market failure is presumed.

An apparently unnoticed implication of agglomeration economies is that market forces result in metropolitan areas that are too small for social efficiency. If the most important source of agglomeration economies is technology concentrated on peripheries of metropolitan areas, the suggested government policy is subsidization of such firms to locate in distant suburbs!

My guess is that the most important reason for measured increases in total factor productivity with respect to urban size is unpriced and unmeasured inputs, namely time spent in commuting and other business travel. Of course, such travel is a business input and presumably affects input and output prices, but only a brave soul would expect that it would show up accurately in data on which urban productivity analysis is based. Presumably, message transmission is an increasingly important part of the urban transportation-communication syndrome. Message transmissions also cost money, but electronic transmission is increasingly both cheap and independent of distance in our knowledge-intensive economy. John Naisbitt and others have forecast an increasingly exurban economy as a result.21 Jess Gaspar and Edward Glaeser have insightfully hypothesized that electronic and face-to-face communication are complementary as falling communication costs permit more information-intensive production.22 For example, it seems highly likely that

widespread use of telephone communication has promoted urbanization, including complementary face-to-face communication, during the last century. Furthermore, electronic communication has been important for thirty to forty years, and as yet there is no evidence that urban growth in the United States or elsewhere is off its long-term path. Finally, university administrations have shown little enthusiasm to encourage faculty to deliver electronic lectures from their Aspen retreats to students scattered worldwide.

My strong guess is that sufficiently fine measurement of input and output prices and quantities would show approximately constant total factor productivity with respect to sizes of urban areas with more than a few hundred thousand in population.

I would like to make three suggestions regarding future research on formation of urban areas and agglomeration economies. First, analysts must be more careful in their use of the term public goods. The term has a meaning. It refers to services that are nonexcludable and nonsatiable. Second, they must distinguish carefully between agglomeration economies and economies that are internal to plants and firms. Third and most important, they must think and estimate carefully regarding the kinds of technologies that are likely to generate significant underpriced external benefits from proximity.

Urban areas grow or shrink by three processes. The first is natural growth—the excess of births over deaths—of the urban population. The second is migration from rural areas, other urban areas, and other countries. The third—a significant part of urban growth in many rapidly urbanizing developing countries—is reclassification from rural to urban of places on the peripheries of urban areas as the density and occupational structure of the peripheral areas change.

Rural-urban migration has received far more study than the other two phenomena. Rural-urban is by far the largest kind of migration in most developing countries, whereas in many industrialized countries in which 70 to 80 percent of the population is urban, urban-urban migration is larger.23 Much migration literature is country-specific. A voluminous literature has been produced on migration within India, for example. Generally speaking, fine studies have been produced on migration in many developing countries.24 U.S. migration studies are too numerous to survey, but Michael Greenwood has been among

23. Substantial migrations, in many cases urban-rural, are the results of extreme violence and severe depressions. These events are ignored in this paper.

the most productive economists who have written on domestic migration, and
his Migration and Economic Growth in the United States reviews literature
up to 1981.\textsuperscript{25} A large literature on transatlantic migration exists. Transatlantic
migration contributed greatly to both U.S. urban and rural growth in the late
nineteenth and early twentieth centuries. The least studied transatlantic migra-
tion has been the slave trade.\textsuperscript{26}

Although theoretical analyses of rural-urban migration have been pub-
lished, most empirical studies are based more on common sense than on
theory. Findings depend on data availability and on the fineness of studies.
But most of the important causes of rural-urban migration appear to be com-
mon among times and places. Rural-urban migration is the property of youth.
Many youth go to urban areas to pursue more or better education than is avail-
able in rural areas. Most do not return. Not surprisingly, the richest rural
residents rarely migrate, but otherwise migrants span the income distribution
from desperately poor to relatively well off. In many countries, young men
migrate to urban areas and then, after they have completed their education and
improved their living standards, marry women from their home regions and
bring them to the city. An interesting aspect of migration, especially in low-
income areas, is an informal contractual aspect. Male migrants may send
remittances back to their rural families—sometimes because of loyalty, some-
times to retain informal rights to property or employment in case they want
to return, sometimes as repayment for financial assistance provided by the
family while they establish themselves in the urban area.

Comparison of rural and urban wage rates is complicated by the fact that
urban living costs exceed rural living costs, but the range of goods and ser-
VICES available is much greater in urban than in rural areas. Benefit from the
broader range of goods is the factual basis of the “bright lights” attraction of
urban migrants. Rural wages are difficult to measure in poor countries because
much farm production is consumed where it is produced, making valuation
difficult. Unemployment data are predominantly urban in most countries, and
are nearly meaningless in poor urban places because literal lack of earnings
is life-threatening for the poorest people. Rural-urban migration has slowed

\textsuperscript{25} Greenwood (1981).

\textsuperscript{26} Between 1500 and 1825, approximately 11 million slaves moved from Africa, some-
times via Europe, to the Western Hemisphere, compared with about 2 million whites. The slave
trade was by far the biggest trade flow of many European and American countries for three
centuries. The slave migration was mostly, but not exclusively, rural-rural. The definitive study
is Thomas (1993), see pp. 804, 805 for estimated statistics.
in many relatively high-income developing countries.27 The countries that the World Bank classified as upper-middle income had an average 73 percent urban population in 1995.28

I have three suggestions regarding future research on U.S. migration: First, researchers should perform all analyses in a simultaneous-equation framework, because people attract jobs and jobs attract people. Second, they should distinguish carefully between amenities that are public goods (for example, good weather) and those that are private goods (for example, golf courses). Benefits of public-good amenities have limited spatial effects and presumably are capitalized in land values and thus have little effect on migration. This may not always be the case, but researchers should make this distinction carefully. Third, they would do well to read the paper by Joseph Gyourko, Matthew Kahn, and Joseph Tracy in volume 3 of the Handbook of Regional and Urban Economics before proceeding with further work.29

I conclude this subsection with brief comments on the size and spatial distributions of urban areas. When people migrate, they almost invariably go to an urban area; they may also come from one.

Size and spatial distributions are of course related, as Paul Krugman has emphasized.30 A geographically small country heavily dependent on trade and with only one fine natural harbor is likely to have a very different urban size and spatial distribution than a large country with many harbors and a large internal market, such as the United States. Likewise, a mountainous country such as Japan is likely to have small urban areas in interior valleys and a few large urban areas at natural harbors.

A cottage industry has grown in estimating the size distribution of urban areas. Care must be taken in interpreting such studies, because some employ city-size data, which for most of the twentieth century have had little to do with sizes of generic urban areas. City data are better approximations to a metropolitan concept in some countries than in the United States, and no official metropolitan data existed in the United States prior to 1940.

Although the first estimates appeared long before his work, George Zipf clarified the rank-size rule for U.S. city-size distributions.31 (A Pareto distribution with a unit exponent, it has the irresistible property that the nth largest
city has one-$n$th the population of the largest city, for all integer values of $n$ in the data.) Subsequent studies have shown that estimated parameters of the Pareto (or log-normal) distribution vary by country and by choice of urban unit (countries with fewer than about 15 million people are typically excluded from most estimates). However, it is astounding that U.S. and other estimates, using the best available approximations to metropolitan areas, find the Pareto distribution with an exponent close to one to be a good fit with data between about 1800 and 1980. During that time, the U.S. urban population increased from about 6 percent to about 75 percent, industrialized massively, and bought or stole real estate that multiplied its size manyfold. Ingenious explanations have been put forward for a Pareto or log-normal distribution, some probabilistic and some economic, but the persistence of a near unit distribution parameter cries out for explanation. Sukkoo Kim’s recent computer analysis merits study and follow-on research.

Uniform geography might contribute to a stable distribution of urban sizes over time, but U.S. geography has evolved dramatically since the early 1800s as boundaries have expanded westward. There has been some speculation about the spatial distribution of urban areas. The three metropolitan areas that have been the largest in the United States for many decades are spaced so that about a third of the population is located closest to each; Moscow and Vladivostok are at opposite ends of Russia; Bombay, Delhi, Calcutta, and Madras are geographically distributed about the same as the Indian population; and Canada’s large metropolitan areas are distributed about uniformly along the inhabitable southern sliver of that country. (Why is there no large Maritime Provinces urban area?) Large urban areas do not like mountains or deserts and do like coasts and navigable rivers.

Analysis beyond these banal observations would require a careful model of the interactions between a metropolitan area and its surrounding countryside (central-place theory), the interactions with other nearby metropolitan areas, and many separate geographical characteristics. Such modeling might be great fun for computer enthusiasts, but I am not sure it has great social value. In principle, careful analyses of size and spatial distributions might shed

32. See the paper by Paul Cheshire in Cheshire and Mills, eds. (1999, pp. 1339–69).
light on the determinants of economic growth. Governments influence urban sizes and locations. Undoubtedly, many national capitals are larger than social optima. U.S. state and national capitals have been located on political considerations, many state capitals having been located simply to be near the state’s geographic center. It has already been indicated that if agglomeration economies are real and substantial, some or all metropolitan areas are excessively small. Finally, some governments attempt to curtail growth of large metropolitan areas, especially by restricting migration of poor people. It appears there have been few attempts to model influences of governments on the size and spatial distributions of cities.

**Spatial Structure of Urban Areas**

The spatial structure of urban areas refers to locations of various sectors within urban areas. The thirty-five-year history of this analysis is monotonic: going from simple and relatively unrealistic to complex and more realistic models.

The Alonso-Mills-Muth model is very simple and easily tested. It starts by assuming something that imposes centrality on the urban area, located exogenously at distance = 0. “Something” was sometimes thought to be a natural harbor, or a rail center, or a market, a place to which all the urban area’s production had to be shipped for sale. Harbors are more or less exogenously located, but rail centers and product markets are endogenous (and not centralized)—and none of the three has served as a powerful centralizing force, at least in most urban areas, during most of the twentieth century. The assumption that all production had to be shipped to distance = 0 is really a deus ex machina to provide centrality. Production in the urban areas was assumed to be competitive, with a constant returns-to-scale, often Cobb-Douglas, production function. Producers located themselves endogenously and typically shipped their output to the deus ex machina at a constant transportation cost per unit mile. Labor, capital, and land were typically the inputs in production. Capital was purchased competitively from outside the urban area.

The urban area’s labor market was competitive, and workers’ housing was produced competitively, also with constant returns, and was located endoge-
nously. Workers paid a constant per-mile commuting cost between home and work. Simple models did not have a production system for local transportation. The urban area’s output was sold at a fixed price at the deus ex machina.

This canonical model generates several theorems about long-run equilibrium: production occupies contiguous space centered on the deus ex machina; land rent equilibrates to make producers’ profits just zero, and rents decline monotonically with distance from the deus ex machina; a well-defined boundary exists between the production and residential areas; all workers occupy a contiguous residential area, are equally well off, with house rents and workers’ density declining monotonically within the residential area, house rents decreasing with distance so as to offset increasing commuting costs; land rents are equal at the touching edges of the production and residential areas; the outer edge of the urban area is where residential land rents equal the rent of land in nonurban (presumably agricultural) use; and worker density and land rent functions flatten if the wage rate increases exogenously or if per-mile commuting costs fall. If the urban area is open, workers’ common utility level equals that which can be obtained elsewhere; if the urban area is closed (with a locally exogenous worker population), then the common utility level is set by the model. Cobb-Douglas production functions and simple housing demand equations provide land-rent and density-distance functions that can be estimated.39

Complexity remains modest if several worker types with different competitive wage rates are introduced; the commuting system is assumed to require land, capital, and travel-time inputs; and a simple retail sector is introduced that can locate in the residential sector.

Estimates reveal some predictive accuracy of the simple models, at least at moderately high levels of spatial aggregation. Population density invariably falls with distance, as do house prices (and undoubtedly actual and imputed rents for equivalent houses), and the population density–distance function flattens if wages increase or if per-mile commuting costs fall, and, as simple models predict, high-wage workers live farther from the central business district than do low-wage workers. This is a good record of richness and accuracy for a simple model.

The greatest failure of the model is the predicted location of all businesses in contiguous space in the central business district (CBD). That prediction becomes less accurate every decade, and in the 1990s only about 10 percent of metropolitan-area employment was located in the CBD in many U.S. met-

39. See the paper by Jan Brueckner in Mills, ed. (1987).
In fact, I believe that the remarkable fact is not that the chimp types so badly, but that it types at all; the broad predictions from simple models remain more accurate than I would have expected, given the massive dispersion of employment in U.S. metropolitan areas and the pervasiveness in the United States of fragmented local government jurisdictions. The second greatest deficiency of the simple model is the absence of a government sector, discussed below.

Starting in the 1960s, dispersed employment was gradually introduced into metropolitan models. Rarely cited in recent years, computer simulation models provided the first important analysis. John Kain spearheaded that subsector.40 Starting in the mid-1970s, dispersed employment was introduced in conceptual models that modified the above-described monocentric model. Michelle White dominated that subsector.41 Early computer and analytical models simply introduced exogenously located employment subcenters scattered around the metropolitan area and away from the CBD. Investigation then centered on the effects of employment subcenters on residential locations and commuting. Still hardly studied is why subcenters form. Presumably the reasons have to do with exhaustion of scale, scope, and agglomeration economies in the CBD; residential suburbanization attracting employment to suburbs; decreasing costs of communication at considerable distances; and congestion, crime, taxes, and so forth in central cities. Even less studied is why subcenters locate where they do. Some were small urban areas before the metropolitan area spread out to engulf them, and subcenters grew there for reasons that put them there to begin with. Some affected and were affected by locations of transportation routes. Presumably there are other influences, such as locations of residences of desirable workers and business-friendly local governments.

A final comment is that identification of employment subcenters depends very much on the fineness of measurement. At a sufficiently gross density, only the CBD shows up. At a very fine density, dozens of subcenters show up in a large metropolitan area such as Los Angeles or Chicago.42 The subcenter surrounding Chicago’s O’Hare Airport has about the same employment as the CBD, but at lower density. Using a gross-density cutoff, only the CBD shows up; at a lower density, O’Hare appears; at a still lower density, two O’Hare subcenters appear.43

40. See the paper by John Kain in Mills, ed. (1987).
41. See the paper by Michelle J. White in Cheshire and Mills, eds. (1999, pp. 1375–410).
42. See Anas, Arnott, and Small (1998), and references therein.
43. See McMillen and McDonald (1998).
There has been an enormous literature on residential suburbanization, starting with the above-referenced deduction from the simple monocentric model that high-income residents are likely to reside farther from the CBD than lower-income residents and that high-income metropolitan areas are likely to have flatter residential-density functions than low-income metropolitan areas. Most of the residential suburbanization literature has been written without regard for interactions between residential and employment suburbanization.

The central controversy in the population suburbanization literature has been the relative importance of basic economic causes (rising income, falling commuting costs, employment suburbanization, lower suburban taxes, and suburban land-use controls suitable to the wishes of high-income residents) compared with central-city social problems (racial discrimination, poverty, crime, poor schools). There can be little doubt that both sets of causes are important; the difficult issue is how much. As in many applied economics specialties, an element of ideology affects this analysis. Those with great faith in markets tend to place greatest faith in the former category, whereas those with less faith in markets place greatest faith in the latter category. My home bias is certainly toward faith in markets.

That basic economic causes are at work is indicated not only by statistical studies of postwar U.S. suburbanization, but also by the fact that suburbanization has pervaded present OECD countries for a century or more and rapidly developing countries for most of the postwar years. Most such countries have had few or no race-related central-city social problems that might cause suburbanization.

That racial and other central-city social problems are also important is also virtually certain. Although racial discrimination has been illegal since the 1960s and is almost certainly less virulent in the 1990s than it was in earlier decades, only a braver person than I would claim that it no longer prevents suburbanization of blacks and Hispanics. My judgment is that race is now a more important determinant of who suburbanizes than of how many people suburbanize. If discrimination prevents minorities from suburbanizing, more whites probably suburbanize than would have in the absence of discrimination.

44. For precise conditions, see the paper by Jan Brueckner in Mills, ed. (1987).
45. See Mieszkowski and Mills (1993) for a survey.
46. See Mills and Tan (1980) and references therein.
47. See Yinger (1995).
Although claims that U.S. suburbanization is excessive have long existed, sprawl and smart growth have recently become political buzzwords. Anthony Downs has long been the intellectual leader of antisprawl advocates. Remarkably enough, welfare economics and notions of the social efficiency of competitive markets are almost never mentioned in antisprawl literature. There is no reason to doubt that U.S. metropolitan land, housing development, housing, and housing finance markets are competitive. Instead, antisprawl writers simply point to the greater suburbanization of U.S. metropolitan areas than of those in Europe (and they occasionally point to even more pronounced centralization in Asia) as evidence of excessive U.S. suburbanization, lamenting congestion on arterial highways, auto pollution, and excessive disappearance of farmland into suburban developments. Benefits of suburbanization are almost never mentioned.

The farmland shortage claim cannot be taken seriously; agricultural surpluses have prevailed for several decades and are the likely future scenario. In any case, there is a market for land at urban peripheries and there is no reason to doubt that relative urban and rural prices can preserve farmland. Even at low suburban densities, urban densities exceed rural densities, so each rural-urban migrant results in a net increase in the rural land available for agriculture. Antisprawl advocates never point out that there is no intrinsic positive correlation between commuting distances and suburbanization. Although it has not happened in the United States, it would be easy to imagine a scenario in which suburbanization of employment and population reduced commuting distances (see page 28). Another false issue raised by antisprawl advocates is that suburbanization eats up excessive amounts of open space. The appropriate technique to preserve open space is for governments to buy land or development rights according to the wishes of voters to be taxed for the purpose. If governments simply prevent development by regulation, they impose the costs of open space on owners who are denied the right to develop their land. That procedure is likely to induce governments to preserve excessive amounts of open space. The Endangered Species Act is a contemporary example of the effects of governments imposing the costs of open-space preservation on landowners. U.S. metropolitan areas are indeed more suburbanized than those in most other countries. One reason is that the United States has large supplies of inexpensive rural land near most metropolitan areas. More important, most countries (all of northern Europe, all of East Asia, most of South Asia,
and Canada) place severe government restrictions on conversion of rural land to urban development. The result is certainly more compact urban areas; typically greater, not less, traffic congestion; and, never mentioned by antisprawl advocates, much more expensive housing. In Canada, which strictly controls metropolitan expansion, house prices (at least in Toronto and Vancouver) are 50 percent greater relative to residents’ income than in the United States; in Europe and Asia, house prices are several times greater relative to incomes than in the United States.50 It is difficult to imagine that U.S. antisprawl proposals could fail to raise house prices relative to income. My newspaper reading indicates that advocacy of suburban growth controls by political officials has already begun to strengthen the exclusionary proclivities of suburban governments.

Future research might begin to follow up on the fragmentary data now indicating that many minorities began to suburbanize in the 1980s. There are some key questions to be addressed: How many minorities have suburbanized? Which metropolitan areas have seen these changes? Which suburbs have seen the most increases, and for what reasons? Has minority suburbanization been a cause of the apparent increase in white settlements in central cities since the mid-1990s? If so, is it because Americans have become racially less tense, because central cities have become more attractive, or because there is no longer any place for whites to “hide”? It appears that job growth in central cities is not an important factor. The proportion of metropolitan jobs located in suburbs appears to have continued to increase, even during the rapid employment growth of the late 1990s.

**Housing**

As was seen in table 1, housing has been a consistently strong interest of urban economists. The survey reported there far understates the volume of publications on the subject since many new scholarly journals, professional publications, and trade magazines have begun and expanded during the past thirty years. Between them, the Department of Housing and Urban Development and the Federal National Mortgage Association publish at least a half-dozen housing periodicals. Housing research is extensive not only in the United States but also throughout OECD countries.51 There is also a vast hous-

---

50. See Malpezzi and Mayo (1997).
51. See the paper by Christine M. E. Whitehead in Cheshire and Mills, eds. (1999, pp. 1559–85).
ing literature related to developing countries. The World Bank has made progress in understanding even the chaotic housing markets in China and Russia.  

In the United States, housing is not only a crucial part of living standards and social status, but also is about half of the market value of privately owned produced fixed capital; it is a smaller part of fixed capital production since housing lasts much longer than most industrial plants and equipment.

Early housing research, such as the work of Leo Grebler, David Blank, and Louis Winnick, focused on time series and cross-sectional measurement of housing prices and quantities. Margaret Reid provided the first sophisticated estimates of the income elasticity of housing demand. Increasingly sophisticated estimates of price and income elasticities were published during the 1960s and 1970s. Although estimates varied disturbingly, by the 1980s most specialists apparently believed that income elasticities were typically somewhat below one and price elasticities were somewhat above minus one-half. These estimates can explain indications that for many decades a nearly constant fraction of income was spent on housing and that, as incomes have risen, housing prices rose somewhat faster than the overall price level. The persistent increase in the relative price of housing probably results from migration from cheap rural housing to expensive urban housing and from increases in urban land values as urban areas grow. It seems that construction costs have not risen faster than the price level. The massive worldwide data compilation under the direction of Stephen Mayo suggests strongly that housing-demand equation parameters vary little among countries. Mayo and Stephen Malpezzi assume that income elasticities are close to one and that demand is price-inelastic, and analyze the dramatic differences in hous-

52. See the paper by Stephen Malpezzi in Cheshire and Mills, eds. (1999, pp. 1791–845).
54. Reid (1962).
55. For a survey through the late 1970s, see the paper by John Quigley in Mieszkowski and Straszheim (1979).
56. See the paper by Christine M. E. Whitehead in Cheshire and Mills, eds. (1999, pp. 1559–85). The paper by Richard Voith in this volume cites evidence that the price elasticity of urban land is about –1.6. Muth (1969) and others have shown that reasonable estimates of the elasticity of substitution between land and structural capital in housing production imply that the price elasticity of demand for land might be an order of magnitude times the price elasticity of demand for housing. The Gyourko-Voith estimate appears to be consistent with this analysis. Indeed, the same analysis applied to business offices explains why the demand for proximity generates building heights and land values in CBDs that are one or two orders of magnitude greater than they are a few miles away. Yet office rents vary much less than proportionately.
57. See Malpezzi and Mayo (1997).
ing price-income ratios in terms of supply-side differences, especially government policies.

Housing is a complex and differentiated commodity. Although government and private housing price indexes have been compiled for decades, specialists have long known that they have been badly flawed. Hedonic analysis, applied in other areas much earlier, began to be applied to housing in the early 1970s. The goal is to ascertain how house prices depend on house characteristics and how characteristics coefficients vary over time and space. An important issue is the extent to which characteristics coefficient differences result from differences in demand or supply. Dozens of applied studies have appeared in this and other countries. I detect little tendency for estimates of characteristics prices to converge.

An alternative to hedonic indexes that has been analyzed during the 1990s is “repeat sales” indexes. In very large data sets compiled for secondary mortgage market securities, some dwellings sell two or more times. Assuming that the physical characteristics of a given dwelling do not change between sales, a precise price index can be estimated. (Many papers have appeared in the Journal of Real Estate Finance and Economics and in Real Estate Economics.) Repeat sales indexes pertain to time-series changes, but not to cross-sectional differences. An additional limitation is that second “sales” are frequently refinancings and resale “prices” are in fact appraisals.

It has long been recognized that dwelling prices are asset prices, whereas the cost of occupying a dwelling for a year is its rent or imputed rent or user cost. Dwelling prices are set by supply and demand and converge to land values plus depreciated current construction cost in the long run. Rents of rental dwellings are set by agreement between tenants and landlords, usually annual contracts but not infrequently unwritten agreements with no explicit duration. Imputed rent for an owner-occupied dwelling equals the sum of mortgage interest plus forgone returns on the owner’s equity plus insurance, maintenance, and repair costs plus real estate taxes less the excess of the dwelling’s value at the end of the year over its value at the beginning of the year. This equation appears in many articles in Real Estate Economics and other housing-related journals. Mortgage interest and real estate taxes are deductible for owner-occupiers who itemize, so their after-tax values are relevant. Long-run ownership costs also include the now negligible capital gains taxes on owner-occupied dwellings, and transactions costs (financial and time costs) of buying and sell-

ing. Long-run costs of ownership include the present value of all such costs over the ownership period. If dwellings and people lasted forever, owner-occupants would be indifferent to exogenous marketwide house price increases, since capital gains would be offset by increases in imputed rents, other things equal.

Such data permit returns to owner-occupancy to be calculated ex post or to be forecast ex ante. Most calculations of returns to owner-occupancy overstate the after-tax returns because they fail to take account of the 80 percent phaseout on the federal tax form of deductions (and exemptions) for relatively high-income owners, to whom the deductions are ostensibly worth most. (Phaseout increases the effective marginal federal tax rate of many owner-occupiers with incomes between about $150,000 and $250,000 by about one percentage point.) Deductions associated with owner-occupancy are worth almost nothing to those with family incomes much over $200,000.

Nevertheless, owner-occupancy is still a good investment for many high-income families because they have access to dwellings on which capital gains are substantial, to low-interest mortgages, and to optimal debt-equity ratios. Low-income families can obtain down payment assistance from government programs and government mortgage insurance, permitting owner-occupancy with virtually no equity. But deductibility is not worth much to families in the lowest tax bracket, and they typically lack access to dwellings with substantial capital-gains potential. Downside risk is negligible for low-income owner-occupiers because they can walk away from a dwelling whose value has fallen by, say, 5 or 6 percent, with little risk of pursuit by lenders or insurers, but with possible risk to credit ratings. Owner-occupancy is not very risky for families with annual incomes between about $50,000 and $150,000, if they can plan for long-term tenure and are sophisticated enough to refinance when interest rates fall and to optimize debt-equity status.

Deregulation, technology, and the secondary mortgage market (which has increased competition in mortgage origination by permitting entry of nondepository institutions) have made housing finance highly competitive, and it has been much studied, especially in the *Journal of Real Estate Finance and Economics*. The United States has as fine a system of housing finance as any country in the world. In most countries, governments restrict competition in mortgage finance. Some of the worst cases are in Asia, where governments restrict entry of private lending institutions and do not issue or permit modern mortgages, apparently to discourage diversion of savings from industrial investment to “unproductive” housing. India and Korea are examples.
I conclude with brief comments on housing supply. Housing supply comes from the standing stock as owners put existing dwellings on the market, and from net additions to the stock. Net additions are mostly production less demolitions, but also include net conversion from nonhousing uses (for example, shops converted to dwellings) and net creation of dwellings from conversion of single-family to multifamily dwellings. Some demolition is intentional, as when land values rise and a single-family dwelling is demolished and replaced by a high-rise apartment. But most demolition results from fires, floods, earthquakes, arson, and so forth. Moves from one dwelling in the standing stock to another provide no net supply.

In the United States, no one has any market power over the standing stock of dwellings. Likewise, the construction industry is dominated by small firms. The largest national housing construction firms supply only a small percent of annual construction, and most housing is constructed by firms that build fewer than 100 dwellings per year. Entry is easy and many builders switch to repair, maintenance, and rehabilitation when the market for new dwellings is poor. Over a period of a few years, housing supply is virtually perfectly elastic. Only two qualifications are needed. First, local governments restrict construction (see the discussion on page 31). Second, as urban areas grow, land prices and therefore housing prices rise. Thus the long-run industry supply of the stock of an urban area’s housing is less than perfectly elastic, but prices of the flow of construction do not depend on the rate of construction in the long run.

Jerome Rothenberg and his colleagues provide detailed analysis of housing supply depending on assumptions made about the time of adjustment and how input prices vary. The durability of housing makes supply elasticities asymmetric between increases and decreases. Absent delays imposed by the need to obtain local government approvals, substantial increases in housing supply can occur in a few months. Decreases, in contrast, depend on the age of the housing stock in question and may require many years. Decreases in demand result in intentional neglect and abandonment if costs of demolition and clearing exceed the value of the cleared land, as was common in poor neighborhoods in earlier decades.

Given that the United States now has quite good legal definitions of condominium property rights, rental versus owner-occupancy is almost entirely

60. See Rothenberg and others (1991) for a fine survey of the entire sector.
demand-driven. The only important constraint is that governments sometimes prohibit conversion of single-family dwellings into apartments (or, equivalently, occupancy by people unrelated by blood or marriage) in neighborhoods of single-family dwellings. Such laws are frequently honored in the breach when economic conditions dictate multiple occupancy.

Beyond doubt, forty years of intensive research have improved understanding of markets for housing and housing finance. I can raise three issues for future research. First, at U.S. commuting costs, dwelling location hardly depends on work location within a broad range of distances and times spent commuting (see the discussion on page 28). My impression from the data is that commuting between similar subcenter suburbs is about equal in volume in both directions. What, then, do dwelling locations depend on? Two-worker families make a difference, but it is doubtful that this is the primary explanation, at least in large metropolitan areas. Second: Why is housing development increasing in some central cities, certainly including Chicago? Is it because of the increasing numbers of suburban empty-nesters? Is it because of falling crime rates and improving amenities in central cities? A third area that cries out for study is manufactured housing. The term refers to any dwelling entirely or mostly produced in a factory and then moved to a site. It encompasses between 20 and 25 percent of annual housing production. Most of these are similar to condominiums in that they are separately owned but located on sites that are owned by a landlord. As is the case in condominiums, rules are established that control housing-unit conditions and residents’ behavior. This may result in low crime rates and low costs. This subsector is virtually unstudied by scholars.

Transportation

As was shown in table 1, urban transportation economics has occupied a small but consistent fraction of papers in the *Journal of Urban Economics*. My judgment is that urban transportation economics publications have also been at a consistent volume in other outlets. *The Urban Transportation Problem* is still the most important book ever published on the subject. Its most important defect is a curious neglect of roadway congestion, which has been analyzed more than any other urban trans-

portation issue during the last quarter century or so. But it was the first public-
lication to provide clear reasons for the suspicion, which still endures, that
fixed-rail commuter systems should have only a very limited role in U.S. met-
ropolitan areas.

I believe that the important issues related to urban transportation are by
now well understood. Functions are now known for estimating costs and
benefits of alternative urban transportation investments, for estimating and ana-
lyzing congestion costs, in the long run and the short run, and even for esti-
mat ing the relative wear and tear on roads by various kinds of vehicles.
Alternative techniques of road pricing (electronic metering and charges for
road use by time and place, fuel taxes, conventional and electronic tolls) have
all been studied. A curiosity is that the United States retains high-transaction-
cost tolls on roads built by taxpayers’ money.

Only about 9 percent of metropolitan work trips are by public transit; about
83 percent are by private vehicle, and the remainder are by foot, bicycle, and
so forth. No careful benefit-cost study has indicated that benefits exceed costs
of any postwar U.S. urban fixed-rail commuting investment (witness Balti-
more, Atlanta, Dallas, Washington, San Francisco, and Los Angeles). Both
U.S. and World Bank studies for developing countries indicate that road invest-
ments have much greater benefits relative to costs.

Commuter bus services can be socially viable if they can provide frequent
rush-hour service in suburban collection areas, high-speed express service on
corridors, and good destination-area distribution. Those are stringent condi-
tions, but they must be met if buses are to compete with cars in time and money
costs. Privatization of bus services is resisted, even though private companies
typically operate more flexibly and at lower cost.

There is strong opposition to the dominance of car-based commuting,
and some of it is virulent. Given U.S. suburban residential densities and
the partially resulting origin-destination diversity, there can be little doubt
that a dominant car-based system is privately and socially optimal. Oppo-
nents of our car-based system claim that low-density suburbs are in good
part the result of auto commuting and that we should have made major invest-
ments in radial fixed-rail systems starting in the early postwar period. Many

66. See Small (1992) and the paper by Kenneth A. Small and José A. Gomez-Ibanez in
68. See Winston and Shirley (1998).
69. See Downs (1992) for a reasoned statement.
such advocates fail to realize that radial fixed-rail commuter systems encourage suburbanization, as London demonstrates. Some recognize the connection and advocate direct controls that would mandate high-density suburbs. Even those advocates fail to acknowledge the higher housing prices that would result.

Nevertheless, there can be no doubt that road use is far underpriced in U.S. metropolitan areas. Optimal congestion pricing would require much greater road-use fees.70 Even if metropolitan road systems were optimized by modest highway construction in suburbs and much better traffic control systems in built-up areas (sequenced traffic signals, reverse-direction streets and lanes, much higher charges and more restrictions for on-street parking, much better traffic law enforcement), metropolitan road-use prices are much lower than roads’ opportunity costs (current land values, replacement costs of improvements, operating costs, all converted to a vehicle-mile basis). In an optimized system with little congestion, I have estimated that fuel taxes should be perhaps 10 times their current level.71 The result would be U.S. fuel prices comparable to those in much of Western Europe (about $3.50 to $4.00 per U.S. gallon). My estimate could easily be off by 25 percent in either direction. If fuel taxes were substantially increased, taxpayers would need to be assured of a decrease in or abolition of some other tax, such as the real estate tax, to avoid powerful opposition. (Fuel taxes could be partially refunded on the income tax to residents who neither live nor work in metropolitan areas.)

What would be the effect of much higher fuel prices? Certainly there would be less driving and less use of large vehicles. Undoubtedly there would be a reduction in frivolous driving and in the large amount of suburb-to-suburb commuting indicated by similar congestion levels in both directions on circumferential highways. Workers who work in suburb B but live in A would tend to move to B, and vice versa if A and B are similar, as many suburban communities are. I doubt that the wishes of public transit advocates for much denser suburbs and much greater reliance on public transit would be realized, and I believe that most urban economists oppose increases in regulatory controls on residential locations or densities.

There is no intrinsic connection between suburbanization and commuting. Employment suburbanization matched by housing suburbanization (in which workers lived near subcenters in which they worked) could result in shorter commuting trips than would be possible in a metropolitan area where employ-

ment was mostly in the CBD. Higher fuel taxes would encourage such matching of worker residences and subcenters.

I can identify some connections needing further research. There are almost no studies of the response of car travel to fuel prices. A spurt of publications based on data from the gasoline crises in 1973 and 1980 was revealing, but recent studies are scarce. There is no doubt that the demand for urban car travel is inelastic with respect to fuel prices. Nevertheless, an increase in fuel prices to about two to three times their current levels might reduce urban car travel by 15 to 20 percent. Most of the reduced commuting travel would probably result from suburban workers’ moving closer to work—to a considerable extent, they would be effectively swapping houses. Since commuting places the greatest stress on metropolitan road systems, research should focus there. To what extent and how quickly would owners substitute smaller cars and export their depreciated sport-utility tanks to Mexico? How many workers would move closer to work? How many would carpool or take transit modes to work? Would residential densities increase close to suburban subcenters? Would some CBD workers move to central cities?

The effects of increased road-use charges would inevitably interact with local-government land-use controls. Increased residential densities near work centers would depend on increased flexibility in land-use controls.

*Local Governments*

As was shown in table 1, papers on local governments have become increasingly common in the *Journal of Urban Economics*. Undoubtedly, the volume of publications in other outlets has increased proportionately. Local public finance has been a specialty in economics for many decades, and many great economists have made perceptive comments on the subject. But postwar U.S. economic analysis of local government behavior can, with only moderate exaggeration, be said to start with Charles Tiebout’s “Pure Theory of Local Expenditures,” one of the most referenced scholarly papers written in the post-war period.72 The fact that most of this section is about work based on Tiebout reflects a home bias, since Tiebout analysis is almost exclusively a provincial U.S. topic. It pertains to a system in which local governments, especially in metropolitan areas, are fragmented and largely independent of higher levels of government. Although U.S. local governments have no constitutional existence, in practice they have more independence to raise and spend money and

---

72. Tiebout (1956).
to regulate activities within their jurisdictions than have local governments in almost any other country. Big-city mayors who are believed to be able to influence the votes of their constituents in national elections also have strong influence in Washington.

The question that has dominated the analysis of local public choice in the United States for more than forty years can be stated simply: Are local governments socially efficient in a metropolitan area in which there are many local governments, people move freely among local jurisdictions, and local governmental taxes and expenditures are determined by majority vote of each jurisdiction’s residents? A preliminary comment is that no one believes that a system of autonomous local governments is likely to engage in income redistribution that is optimal from the point of view of the entire society. That is why income redistribution should be, and largely is, financed by higher-level, mostly federal, government programs.

A fine survey, with a skeptical conclusion, appears in a paper by Stephen Ross and John Yinger in the *Handbook of Regional and Urban Economics.*

The champion of the view that fragmented local governments are likely to be socially efficient is Dennis Epple, ten of whose papers with various coauthors are referenced by Ross and Yinger; two other papers have become available only recently.

Epple and his colleagues have produced variants of the following remarkable model: There is a fixed number of local governments in a metropolitan area, each with fixed but not necessarily equal land area. There is a fixed continuum of metropolitan-area residents’ incomes, with fixed but not equal incomes before local-government taxes and expenditures, not dependent on community of residence. Residents have well-behaved utility functions in a composite good, housing, and local-government services. Residents take prices of housing, the composite good, and each community’s tax-expenditure function parameters as fixed. Households are mobile among communities and locate where their utility is maximal. Tax-expenditure parameters are set by the median voter in each community. Communities are in equilibrium in that housing supply and demand are equated and no resident can obtain greater utility by a move.

Epple’s theorem is that households sort themselves among communities in such a way that communities can be numbered so that a continuous segment of incomes is found within each community, the highest-income resident in

74. Epple and Sieg (1999a) and (1999b).
one community and the lowest-income resident in the next are indifferent between residences in the two communities, and income levels ascend as one moves to higher-numbered communities. It does not seem to matter whether tax-expenditure functions are dependent on house values or are linear in income within each community (equals are treated equally within communities). Epple and his coauthors have ingeniously estimated and tested the model and find considerable consistency between facts and theory.  

One remarkable fact about the model of Epple and his coauthors is that, as in Tiebout’s theory, there are no land-use controls in the model and dwellings are built to be optimal for residents who occupy them, yet no low-income free riders sneak across borders and pay low taxes levied on modest dwellings, receiving benefits from the richer community’s generous spending. How many dozen papers on the Tiebout model assume the contrary?

The first thing to say is that this is an equilibrium model and no one believes that an entire metropolitan area is in residential equilibrium. But postwar suburbanization may be a move toward equilibrium. Second, what about central cities, which contain about one-third of the metropolitan population? Their income levels span the entire metropolitan distribution, although their mean incomes are typically lower than those of suburbs. High-income people have been moving from central cities for half a century, so perhaps equilibrium is on the horizon. But central cities do not fit easily into the Tiebout model. Third, if the Epple model tells the entire story, it is unclear why suburbs (and central cities) take land-use controls so seriously. Of course, jurisdictional locations do matter, and perhaps that is enough to invalidate the conclusions of the Epple model. In addition, even though equilibrium might not be much different if all jurisdictions dropped land-use controls, things might be much different if any single jurisdiction abandoned them. If that is so, land-use controls must have badly distorted resource allocation in metropolitan areas. One cannot attend many zoning hearings in high-income suburbs without concluding that keeping out the low-income hordes is really at issue. Again, things might be very different if all high-income suburbs abandoned controls simultaneously, but I do not think this is likely. Next to schools, land-use controls are the biggest political issue in some suburbs; they are fought over bitterly, with considerable expenditure of time and money.  

75. See especially Epple and Sieg (1999b).
76. See Fischel (1985).
Land-use controls stratify communities directly, according to housing demand, but also indirectly, according to income, family size, and race. Stratification has many faces. Children from low-income families probably perform worse if they live in neighborhoods that are predominantly low-income. In addition, low-income children are probably more expensive to educate than high-income children. Another aspect of the same syndrome (believed by some education specialists) is that children’s success in school and in the workplace depends on their preschool environment. Residence in a neighborhood not dominated by other low-income residents may significantly improve the preschool environment of children from low-income families. Alternatively, it is possible that tension might result from neighborhood mixtures of incomes, family backgrounds, and races and that the outcome may be harmful to low-income children. Land-use controls also stratify residents by the propensity to commit crimes and by the frequency of teenage pregnancies. More mixing because of less zoning could affect such behaviors of any or all groups involved.

Dennis Epple and Richard Romano have recently applied the basic Epple model to the issue of school choice. The main purpose of proposed educational voucher systems is to increase competition among schools for students in order to improve educational quality. Vouchers are likely to have that effect. But vouchers also permit more stratification by student quality than is likely if schools have substantial monopoly power, as they currently have, at least in central cities. Epple and Romano show that stratification would be the likely result of almost any voucher system.

It is clear that zoning is the result not just of the median voter’s preferences, but also of residents who vote and lobby, interacting with developers who lobby and make campaign contributions. Zoning almost certainly follows the market, at least when market forces are sufficiently strong. That may make resource allocation better or worse than it would be under the Epple median-voter model.

My guess is that land-use controls represent an important breakdown of the median-voter model. In many jurisdictions, owners of residences and businesses near a disputed property have a disproportionate effect on land-use controls on that property. Just as owners of expensive single-family detached dwellings go to great lengths to prevent rezoning of nearby property for high-rise apartment dwellings, even if it would be indisputably good for the entire

78. See Freeman (1999).
community, similar results occur if a business that would be good for the community tries to locate in a high-income neighborhood. In some jurisdictions, a city council member has a disproportionate influence on land-use controls within his or her district, and his or her reelection is not affected by those who do not live there because they have been zoned out.

Finally, this observer believes that Leviathan is at work to an unknown extent within local governments. Public school bureaucracies have undue influence on public school operations. Gangs that peddle illegal substances corrupt the criminal justice system. Contracting between local governments and private companies to haul waste and to build and repair buildings and streets is subject to corruption. Local government offices are overstaffed with political supporters of elected officials. Elected officials distort and hide facts from the electorate. (How else to account for massive cost overruns on virtually all government infrastructure projects?) The local criminal justice system is blatantly discriminatory against racial minorities.\textsuperscript{79}

I have omitted private racial discrimination from the above examples because, although it is a violation of many state and local laws, it is primarily a federal matter.\textsuperscript{80}

I propose a number of questions for future research. First: What exactly are the consequences of land-use controls, and how much waste do they cause? The above discussion has focused on housing, but similar analysis is needed for businesses. Small high-income suburbs sometimes permit almost no businesses. Everybody has cars, no residents need to work in the community, and it is little trouble to shop in a mall in a nearby community. I have nothing against malls in a country in which suburbanites own vehicles that are ideal for one-stop shopping and freight haulage. But if most suburbs zone out most businesses, the result is wasteful travel for shopping. Perhaps the median-voter model applies to business zoning, but I suspect that high-income suburbs are not only excessively homogeneous regarding race and income but also excessively exclude businesses. Second: Is there anything to my Leviathan story? No one can live long in Chicago, Philadelphia, Detroit, New Orleans, Washington, or many other U.S. central cities and believe that the local governments are producing the demands of the median voter at minimum cost. Why not? Are these cities too big? Are their residents so poorly educated that they cannot control their governments? Many characteristics of central-city governments appear to be impossible to explain, either by the Tiebout model

\textsuperscript{79} See Thernstrom and Thernstrom (1997).
\textsuperscript{80} See Yinger (1995).
or by any theory of rational social choice. Surely more research on how and why central-city governments allocate resources is close to the highest priority for economic research.

How Far Have We Come? Where Are We Going?

I believe it is beyond dispute that urban economics has made progress on every major topic that has been addressed. We understand much better why and where urban areas form and grow. I have indicated important deficiencies in the analysis of what we refer to as agglomeration economies, but estimation and analysis are far ahead of where they were in 1960. The size and spatial distributions of urban areas have been studied intermittently, but further progress, especially related to the spatial distribution, will be difficult.

Analysis of the spatial organization of urban areas hardly existed in 1960. We now understand how and why suburbanization of residences and employment occurs. We still have much to learn about subcenter formation and growth, but important progress has been made.

Despite the careful work of urban economists on suburbanization, uninformed antisprawl proposals have swept the political scene since the late 1990s. Perhaps the proposals will fade, as the witch-hunt in child care centers did by the mid-1990s, but the potential for harm is enormous. Urban economists could undertake careful estimation of market failure related to suburbanization, of effects of government failures to “get the prices right,” and of optimal government programs to ameliorate resource misallocation.

Understanding of urban housing markets is far ahead of its status in 1960. We now basically understand housing supply and demand and most basics of housing finance. We have much to teach governments in developing countries about the benefits of a modern housing finance system.

We still lack critical evaluation of government housing programs. I suspect that typical local-government land-use controls do more harm than good. Perhaps I am wrong, but estimates of benefits and costs of alternative control programs would be possible. Public housing is about the only government housing program that has been studied carefully, and it has been found wanting. Housing vouchers were studied carefully twenty years ago and were found to be preferable to all government supply-side programs;81 but almost nothing has been written about them or alternatives for two decades.

81. See Bradbury and Downs, eds. (1981).
Urban transportation issues have been carefully studied. We know how to compute appropriate pricing and investment benefits and costs. Congestion and its pricing have been carefully studied. I do not believe that agreement is widespread with my belief that moderate road and traffic control investments and much higher fuel taxes would be better and vastly cheaper than sophisticated congestion pricing based on electronic metering. The issues could be studied.

The vast outpouring of literature has contributed much to our understanding of local governments. However close we may be to a Tiebout equilibrium, it is by now clear that powerful market forces propel at least suburban communities in that direction. The implication was formerly drawn forcefully that the ultimate result would be destitute and alienated central cities composed mostly of minorities. In the late 1990s at least some central cities—apparently those with relatively good governments—have begun to revive. Why? What should government do, or stop doing, to reinforce the process? Why do some central-city governments zone down or zone out high-rise residential and business developments? Congestion is the usual answer, but it is unlikely. Zoning out high-income high-rises induces more high-income residents to reside in suburbs, from which they commute to CBDs on congested arterials. Their employers also are likely to move, depriving the central city of both high-income residents and well-paid jobs. This pattern exists in Chicago. Is it common? Do central-city governments simply not understand the natural consequences of their actions, or is another motive at work?

Social Problems and Urban Economic Research: Three Proposals

Some readers may be offended at my slight treatment in this paper of what I will call social problems in U.S. metropolitan areas and especially in central cities: poverty, crime, racial discrimination, and education. However they are designated, the most intractable colonial and U.S. social problems have revolved around the great American trauma of the last three or four centuries: racial relations. Whether justified or not, neglect has been intentional. First, adequate treatment would have made the paper much too long. Second, sociologists and political scientists have contributed more than economists to analysis of many social problems, and I am not competent to sort out economists’ contributions or to survey the entire subject. Of course, economists have made important contributions to analysis of these social problems. However, at the risk of irritating colleagues, I venture the opinion that reading superb surveys such as that of Q. James Wilson and Richard J. Herrnstein or that of Stephan Thernstrom and Abigail Thernstrom leaves one with the impression
that economists’ contributions have been more slight than they should have been. In this section I briefly propose three topics that I believe economists have a comparative advantage in analyzing.

Economists are the masters of quantitative benefit-cost analysis of government programs. One way to reduce crime is to decriminalize certain activities. In terms of the numbers of people involved, drugs are the prime candidate. Of the one percent of the U.S. adult population that is in prison, a larger number has been convicted of drug offenses than of any other crime. Some of those criminals could lead productive lives if drugs were decriminalized. Furthermore, decriminalization—presumably making drugs available on prescription—would reduce the profitability of sales efforts in marketing illegal drugs and might reduce consumption. Drugs probably provide a large part of the revenue of street gangs and help to corrupt the criminal justice system, as Prohibition did in the 1920s. International comparisons are never conclusive, but insight could be obtained from them. I do not have a firm opinion about decriminalization, but I believe that careful benefit-cost analysis could contribute to the debate.

A completely different social issue is projecting who would benefit from suburban growth controls. I argued on page 21 that the important effect would be to increase the relative price of urban housing by regulatory limits on supply. Existing owners would make capital gains when controls were instituted, but gains would be offset by proportionate increases in imputed rents, for them or their heirs. Older owners who would soon move to less-expensive housing even in the absence of controls would benefit, at least if they had no great concern for their heirs. That would seem to be a small consideration. Likewise, farmers or other owners of land on which housing development became illegal would suffer capital losses, of course offset by proportionate reductions in imputed rent unless they anticipated moves to urban areas. It is difficult to imagine a significantly important political group that could be mobilized to support proposed suburban growth controls. Why such controls exist in so many countries then becomes a mystery. Of course, governments confuse the issue by claiming land “shortages,” apparently relying on citizens’ lack of understanding of elementary supply and demand. Determining what groups might benefit from suburban growth controls is an issue in political economy on which economic analysis could shed light.

A third example of a social issue to which economists should be able to contribute comes from James Heckman. Education specialists have compiled increasingly strong evidence that preschool intervention is crucial if children from deprived backgrounds are to do well in school and in subsequent work experience. The claim is that unless children begin to develop cognitive, psychological, and social skills by about the age of six, later development of such attributes will be much more difficult. The subject cries out for benefit-cost analysis. The first issue is program design. Preschool programs are quite different interventions than reducing class sizes or introducing labs or computers in middle school. It is also important to design interventions so as not to violate rights of parents or children. The second issue is benefit-cost analysis of alternative programs or of no preschool intervention or of interventions in later secondary or postsecondary education.

Forecasts

I believe that in the near future urban economic research will move toward a general synthesis of how, when, and where subcenters form and grow, and how that process interacts with residential suburbanization and transportation improvements. The process can be accompanied by considerable reductions in commuting travel if governments take appropriate actions. Presumably such research will entail revival of interest in the type of large computer simulations that John Kain, Alex Anas, and others undertook in the 1970s and 1980s, but with much better data and theories and much more powerful computers. Such research may constitute a renaissance of urban spatial analysis in coming years.84

It is not difficult to predict that interest in urban governments will continue apace. It is dangerous to believe that a problem has been solved, but I believe that we now have the deepest possible conceptual understanding of the Tiebout model, thanks mostly to the work of Epple and his coauthors. Details of empirical estimation and testing can and will certainly continue to appear, but a key issue is simply how far toward a Tiebout equilibrium suburbs have progressed. More interesting, I believe, is how governments perform, and could perform, in distinctly non-Tiebout-like central cities. Am I correct in believing that central-city governments represent interests of constituents much less accurately than do suburban governments? If so, why? Most important, how much unde-

sirable stratification has been strengthened by suburban land-use controls? Perhaps, to close on the most optimistic note possible, central city–suburban problems may be solved by continual out-migration of central-city minorities and in-migration of suburban whites, until by 2050 central cities and suburbs will be almost uniformly and fully integrated.
Comments

Dennis Epple: When I was a Ph.D. student at Princeton University, the two faculty members whose work I found most exciting were Edwin Mills and Wallace Oates. It is an honor to share this forum with them.

Writing a history of urban economic analysis would be a daunting task for most of us, but Mills handles it with evident ease. He organizes research in urban economics into nine subject areas. It is an appealing classification that begins with the topic that unifies the field: spatial analysis. It includes eight other topics that divide urban research into relatively distinct subject areas: housing; government sector; labor, poverty, race; crime; transportation; education; interurban, rural-urban; and developing countries. The paper goes on to provide a concise summary of research efforts to date and highlights key remaining issues for research.

I like the paper very much. It not only provides a summary of the state of research, but also offers opinions about the issues within the research, pointing out where Mills thinks work is on target and, in some cases, where he thinks it is wrongheaded. The author has also been extremely modest in this history of research in urban economics, understating the central role that he has played in the field through his tremendous research contributions, his text, his role in founding and editing the Journal of Urban Economics, and his influence on his students and on virtually everyone else who has worked in the field. The editors of these new Brookings-Wharton Papers on Urban Affairs chose well in giving Mills the leadoff role to initiate the journal.

It might be helpful if I highlight some of the open research issues raised in the paper. The review begins with fundamental questions for the field of urban economics.

Why do urban areas exist? What is the reason for the measured increase in total factor productivity as urban areas increase in size? Mills’s answer to the first question is that proximity economizes on the costs of moving goods, people, and messages. His conjecture regarding the second is that unpriced and
unmeasured inputs are the source of measured increases in total factor productivity as urban areas grow. The paper emphasizes in particular time spent in commuting and other business travel. This conjecture strikes me as providing a valuable focus for researchers looking to pin down the elusive sources of economies in urban areas. Mills also takes some stands about things that are probably not the source of increasing factor productivity, and he highlights key issues for this area, such as the importance of distinguishing between external economies and economies internal to firms.

The paper considers another question central to urban economics: What determines the size distribution of cities? Mills cites evidence of the pervasiveness of the rank-size rule within various countries and across countries for long periods of time. I find this particularly compelling because I know that Mills reads this literature with a skeptical eye. Indeed, I recall that some time ago he made the point that if you have two countries where the rank-size rule holds, then it cannot hold if the data for two countries are combined. The fact that this relationship is so robust over time and across countries makes it a central regularity that, as Mills says, cries out for explanation.

Regarding issues related to location within metropolitan areas and to the zoning activities of local governments, Mills takes to task those of us who neglect land-use controls in modeling household-location choices and local government tax and expenditure policies. He conjectures that land-use controls represent a massive breakdown of the median-voter process because residences and businesses near a disputed property have disproportionate influence on land-use controls for the property. He also expresses the view that the outcomes that result are often not the outcomes that are best when viewed from the perspective of the larger interests of the community. I wish I could say that I think his criticism is wrong, but, to borrow a phrase Mills uses elsewhere, it would take a braver person than I to assert that he is wrong about this. Several different researchers have taken a crack at modeling zoning from one angle or another. These efforts have produced models of zoning that generate many interesting insights. However, I do not believe that we have a good political-economy model of zoning. I have in mind the kind of model that can be made part of an analytic structure with which we can characterize the sorting of people among communities in metropolitan areas, or characterize the distribution of housing consumption across households and across communities.

The paper’s discussion of the problem highlights the challenges for modeling. It is daunting to build a collective-choice model of how zoning is set.
One difficulty is in modeling the way that money and votes interact in determining political outcomes. This seems particularly important in conflicts between residents of a community and prospective developers. Residents have the votes, but developers and builders may be willing to spend a good deal of cash to influence political outcomes. It is hard enough to work with a model in which only one or the other force is present, but it is really challenging to devise sensible models that balance these very different forms of political activity. The other problem, as Mills notes, is that the traditional majority-rule model may not be very helpful if decisions about zoning for a parcel are dominated by owners of nearby parcels. In any case, I agree that models that characterize metropolitan equilibrium while neglecting zoning may be missing an important piece of the problem. Developing a model that responds to this challenge strikes me as an important item on the research agenda for urban economics.

A second topic that I want to comment on is stratification of the metropolitan population by income, race, and other characteristics. The paper discusses research focused on attempting to disentangle the relative importance of economic factors and central-city social problems in influencing suburbanization. Peter Mieszkowski and Edwin Mills have also developed this issue more fully in their paper reviewing this topic. If I recall correctly, Miezkowski and Mills made the point that a key issue is whether central-city social problems are simply a consequence of the concentration of disadvantaged households or whether the concentration tends to foster increased social problems via peer effects.

I want to suggest that, for the same reason, stratification across municipalities in metropolitan areas and across neighborhoods within municipalities is also an important topic that needs more research attention. There are large differences in quality of local public services across municipalities in a typical U.S. metropolitan area. This is particularly so for education. In part, these quality differences can be traced to differences in spending levels. In the Chicago metropolitan area, suburban high schools spend a third more per pupil, on average, than do high schools in the city of Chicago. However, from a research perspective, I think a more important issue than spending differences is the effect of differences in peers. In the Chicago city schools, the average percentage of students from low-income households is 69 percent; the suburban average is 11 percent. And these averages understate the concentration of disadvantaged students within many schools in the city. How

much we worry about such stratification depends on the extent to which peer effects are important.

While the issue of peer effects is not new, there is a good deal of recent theoretical research that proceeds from the presumption that peer effects are important, including work by Roland Benabou, Steven Durlauf, Thomas Nechyba, and Richard Romano and me. While research on local governments has emphasized the stratification that is induced by differences in expenditure across municipalities, peer effects can give rise to stratification even where there are no expenditure differences. Interesting empirical work by Sandra Black shows that housing prices differ across boundaries in urban areas that define which neighborhood school students may attend. A paper by Daniel Aranson investigates whether stratification is lower where expenditure has been made more equal. He finds a measurable reduction in stratification, but it is not large. Both of these findings are consistent with the idea that peer effects are a central force for stratification. But there are other potential explanations. The state-of-the-art empirical contribution on peer effects is the fine paper by William Evans, Wallace Oates, and Robert Schwab. They show that peer effects in teen pregnancy and high-school dropout rates are significant when estimated in the usual way but disappear when one instruments using exogenous factors that influence peer-group composition. This certainly means that the standard method of estimating peer effects that neglects self-selection into groups will not do, and it highlights the importance of developing alternative ways of getting at the magnitude of peer effects.

The final issue I want to comment on is related to Mills’s remarks about the likelihood that Leviathan is present to at least some degree. Mills observes, for example, that public school bureaucracies have undue influence on public school operations. I share this view. I think this is particularly true for central-city school systems, simply because a single bureaucracy typically has control over a substantial fraction of the schools in a metropolitan area. For example, a single district operates the sixty-three high schools in the city of Chicago. The ninety-five high schools in the surrounding suburbs are almost all single-school districts. The competition faced by the suburban schools must surely provide less scope for bureaucratic license than the city district can exercise via control of sixty-three high schools. Caroline Hoxby provides evidence

that increased public school competition does indeed lead to lower per-pupil expenditures and improved educational outcomes.\footnote{Hoxby (forthcoming).}

Mills also provides other examples, and I have one more to add to his list. I am increasingly struck by the amount of resources local and state governments in Pennsylvania spend in the name of economic development. I think much of the expenditure on economic development genuinely derives form a well-intentioned desire to foster economic growth. Expenditure on enterprise zones is a case of particular relevance for urban areas. It strikes me as an area that needs continuing research attention. One of the best empirical papers in this area is a recent paper written at Carnegie-Mellon by John Engberg and Robert Greenbaum.\footnote{Engberg and Robert Greenbaum (forthcoming).} They do a much more micro-level analysis than previous research, using restricted-access data files available through the U.S. Census Data Center. They find that spending on enterprise zones does indeed have a statistically significant positive effect on creation of new jobs. However, they find that there is a nearly identical and statistically significant reduction in old jobs, so net job creation is zero. This is a striking finding that points to the need for theoretical work and additional empirical work to try to understand better the circumstances in which enterprise zones may actually promote development and those in which they simply displace old jobs as they create new ones. There is also the related question of conditions under which enterprise zones would be a sensible policy even when they were not successful in creating net job growth in an area.

In closing, I want simply to reiterate that this is a fine paper. I highly recommend it.

Wallace E. Oates: Edwin Mills’s paper provides a splendid history of research in urban economics. There is surely no more appropriate author for this paper: Mills is widely regarded as the father of urban economics, and the field bears his very strong imprint. Since I have not been very close to much of the recent research in urban economics, I have decided to take a somewhat different tack for my comments by offering some admittedly impressionistic observations on the evolution of the field over the past three decades. It is, in fact, something that I have puzzled over—and I will try to explain the nature of this puzzle and seek some help in understanding it. The discussion at the conference has been very helpful in clarifying several elements of the issue.

To put the matter in perspective, I want to return to the period of the 1960s and early 1970s. This was a period in which two new fields emerged in economics, both of them as a result of the growing social consciousness in the 1960s. These two fields were environmental economics and urban economics. The former had its roots in the so-called environmental revolution with its widespread concern with pollution and general environmental degradation. It gave rise to a spate of new legislation—the Clean Air Act of 1970, the Clean Water Act of 1972, and a host of measures in subsequent years. Parallel to this, the new awareness and rising concern with urban problems likewise ushered in a variety of major new programs to address the decay of the cities: urban renewal programs, model cities, the war on poverty, and several others.

Associated with these concerns, new courses sprang up in the colleges and universities across the country in these two fields, at both the undergraduate and graduate levels. And economists working on these issues during this period often found themselves identified as either urban economists or environmental economists. Moreover, new journals emerged in both of these fields to offer outlets for ongoing research.

The thing that has struck me and that I have puzzled over is what has happened to these two fields in the 1980s and 1990s and why. It is my sense that one of the fields, namely environmental and natural-resource economics, has continued to expand with great energy and public attention, while the other, urban economics, has been less successful in maintaining its earlier momentum.

Let me begin by simply offering some casual evidence for this. First, during the past twenty years, urban problems have faded somewhat from the public view. This is most certainly not to say that these problems were successfully resolved by newly introduced policy measures. Far from it—but for whatever reasons these issues simply did not continue to engender the same level of public attention and concern. In contrast, environmental issues have continued to capture the public attention, both in this country and around the globe. The result has been an ongoing process of new legislation and policy measures aimed at a “sustainable world.”

This divergence has also manifested itself in university curricula and academic research. One finds over this period an ongoing expansion of programs and courses in environmental studies. Economics departments in many universities now offer as a Ph.D. concentration the field of “environmental and natural-resource economics.” And at the undergraduate level, programs of environmental studies (with environmental economics as one component) are ubiquitous throughout universities in this country.
In contrast, the interest in urban economics seems to have waned. Courses have disappeared in many universities. When, for example, I arrived at the University of Maryland at College Park in the late 1970s, we had a very active program of graduate study and research in the field of “regional and urban economics.” There were many students taking degrees and writing dissertations in the field. But that is gone now. We still have an undergraduate course in urban economics, but there is little demand for courses at the graduate level. And my impression is that our experience at College Park is not atypical.

It is also my impression that there has been some difference in the expansion of research efforts. It seems to me that practically every time I turn around, I find a newly introduced journal in environmental and resource economics. This reflects to some extent the thriving and growing body of economists working in the field. Created in 1979, the Association of Environmental and Resource Economists (AERE) expanded rapidly to its present size of nearly 1,000 members with a very active sister organization (EAERE) in Europe. But—although there is still the splendid *Journal of Urban Economics*, initiated by Edwin Mills and now edited by Jan Brueckner—I do not have the sense that there has been the same proliferation of journals and research in urban economics that has occurred in the environmental field.

I must be careful here. As the Mills essay makes clear, there has been a lot of valuable research going on in urban economics. Real advances have been made. And many urban economists have simply moved their base of operations to programs of management or finance, where their new home is often in “real estate economics.”

Nevertheless, I think there is still some substance to my puzzle. It raises the intriguing question of why one field born of social concerns in roughly the same period has grown so much more rapidly than the other. I suspect that the answer lies largely in the extent of public interest and its manifestation in the form of monies to support research. On first glance, this may seem a bit strange. Urban problems certainly have not gone away. But somehow during the 1980s we turned our back on them; we simply chose not to pay a lot of attention to the problems of the center cities. There is some resurgence of concern at present, but the public attention that focused on urban issues in the 1960s and 1970s has apparently not sustained itself.

The visibility of and concern with environmental issues, however, has remained at a very high level, with lots of new legislation and public programs. Just why environmental issues have remained so visible and compelling while urban ones have not is an intriguing question. And I do not have a good answer.
to it. Perhaps it is, in part, the extent of exposure: nearly everyone has contact with the environment in one form or another—it affects high- as well as low-income families. One observer has suggested to me that the environmental movement sapped much of the energy that had been directed to the social concern with poverty and the cities. In her view, the advent of Earth Day in 1970 signaled the shifting of the focus of activist groups from problems of poverty to those of the environment. But we cannot simply wall off the problems in the center cities.

Some part of the answer to my puzzle may also lie in the intellectual evolution of the two fields. The theoretical core of urban economics was the elegant Alonso-Mills-Muth model of urban structure. This monocentric model provided a conceptual framework that generated a number of interesting and testable hypotheses. It provided a kind of intellectual focal point for the field. But as Mills points out, as the field has evolved the model has manifested some serious limitations, both in theory and application. With the ongoing process of decentralization of employment in urban areas, the model has become a less apt description of urban structure (although, as Mills observes, it still retains considerable explanatory power). And efforts to extend the framework to provide a general treatment of metropolitan areas with multiple centers have not been very successful. In short, it seems as though the basic monocentric model, once its basic properties were thoroughly explored, simply played out as a source for further interesting work. It may be that the future of basic theory in urban economics will be of a nonspatial form.

My case here would be a stronger one if I could cite all the major theoretical advances in environmental economics that have maintained its intellectual vitality. Here I must hedge somewhat. The intellectual core of environmental economics was basically the theory of externalities, and that was worked out reasonably well early on. There has been important new work on methods for the valuation of environmental amenities—new ways to understand and measure people’s willingness to pay for a cleaner environment. And during the 1990s there has emerged a challenging new literature on environmental policies in a second-best setting. The work of Lans A. Bovenberg, Lawrence H. Goulder, and others has led environmental economists to take a new, fresh look at some of the basic propositions.

But frankly I find it hard to ascribe the buoyancy of the field of environmental economists mainly to exciting new theory. I come back to the sustained social concern with environmental issues, which translates into funding for research (among other things). In fact, if one looks solely at the issue of global
climate change, one finds virtually a whole industry in its own right. The vast commitment of public funds in the United States (and in other countries as well) to research on this potential threat has drawn many economists (and other scientists as well) into its study. Public attention and the stimulus for scientific work it provides, through funding and by engendering interest, constitute a powerful force with a profound impact on the portfolio of society’s research efforts. It is my perception that this has worked in the favor of environmental and resource economics much more than of urban economics. But this tendency may be in the process of shifting back somewhat. We shall see.
References


Edwin S. Mills


Lösch, August. 1954. The Economics of Location. Yale University Press.


Edwin S. Mills


von Thünen, Johann Heinrich. 1826. *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*. Hamburg, Germany.


