

Nonoptimal levels of suburbanization

Philip E Graves

Department of Economics UCB 256, University of Colorado, Boulder, CO 80309-0256, USA;
e-mail: Philip.Graves@colorado.edu

Received 21 March 2002; in revised form 30 July 2002

Abstract. Suburbanization has many causes, among which is the attempt to relocate to acquire a more desirable vector of local public goods. The traditional economists' procedure for valuing public goods involves vertical aggregation of marginal willingness to pay, at a given income level. This approach is flawed by failing to recognize that individuals will not work for goods that cannot be acquired individually with higher incomes. There will be a parallel input market failure any time there is a public good output market failure, thus the 'given income' of the traditional valuation method is too low. Hence, traditional valuation methods result in underprovision of local public goods (for example, parks, safety, education, and environmental quality) at the urban centers. As a consequence, there will be nonoptimally large levels of suburban sprawl with substantial resulting welfare loss.

1 Introduction

Noneconomists almost universally despise suburban sprawl, while economists generally are more ambivalent. The issue, as seen by the economist, is quite complex and hinges critically on the nature of externalities. If all externalities (crime, air and water pollution, congestion, noise, and so on) were properly internalized in each separate market, the economists' presumption would be that there would be no role for urban planners; they could only make things worse. Land would be allocated to the highest valued use, and the city's edge would just be where it turns out to be ... but wherever that turns out to be would be optimal.

Clearly, externalities are not fully internalized in cities, and it might even seem unclear whether, on net, total externalities rise or fall when households relocate from the urban center to the suburban fringe. There will generally be longer commutes, with increasing levels of pollution and congestion in the suburb, and perhaps throughout the urban area. But it is also the case that significant population movement from the central city reduces population density there, with perhaps corresponding declines in external damages received by those who remain. I speculate that most economists would guess that, on the margin, observed suburbanization levels have costs greater than benefits. Yet, an apparently strong 'revealed preference' argument (for many decades, massive population outflows from urban to suburban areas have occurred) must certainly carry some weight. Might not the private benefits of living in the suburbs exceed the full social costs of doing so, despite concomitant externalities?

The present paper represents an attempt to demonstrate that observed high levels of suburbanization represent a socially nonoptimal urban pattern. That is, neoclassical urban economic models (for example, Fujita, 1989; Wheaton, 1976) typically predict concentric rings of diminishing intensity of residential housing as one approaches the city edge from the center. It is argued here that private market equilibria result in cities that cover too great an area, having population densities, that are too low in the urban centers. Moreover, even if we had a Samuelsonian local public goods optimum for every public good in the urban core, the extent of any *remaining* suburbanization would still be nonoptimally large.

In section 2, I present a discussion of a flaw in the valuation of public goods, drawing heavily from more in-depth treatments [see Graves (2001) or Flores and Graves (2001), the latter representing the more formal treatment]. In section 3 the results of section 2 are applied to the case of suburbanization. It is seen here that suburbanization largely represents the purchase of a substitute for underprovided local public goods in the urban core. Critically, this underprovision is not merely that we have failed to provide ‘Samuelson-optimal levels’ of public goods, but rather that the Samuelson condition is, itself, theoretically flawed. In the closing section 4 I summarize and discuss potential extensions.

2 An input market failure and its implications for optimal public goods provision

Samuelson (1954) argued that the proper amount of public goods would be forthcoming when the aggregated marginal willingness to pay of all households benefiting from the good was equated to the marginal cost of its production. Indeed, if one could tax each household at its marginal willingness to pay, the Pareto-superior amount of both public and private goods would seem to be forthcoming. He was not optimistic about arriving at that optimal public good level, because of important problems of demand revelation leading to what is now called the ‘free-rider’ problem [for potential solutions to this problem, see Clarke (1971) Groves and Ledyard (1977)].

But, there is a problem with the Samuelson condition for social optimality that has gone unrecognized until quite recently. The income level was taken as given in the argument leading up to the marginal conditions, and that turns out to be of critical importance. At the heart of economics is the presumption that we work to get the things we want, for (apart from philanthropy or theft) we can only acquire the things we want through the supply of our labor and other inputs into the production process. This works fine for ordinary private goods; we work up to the point where the marginal value of goods is equal to the marginal value of the foregone leisure associated with their acquisition.⁽¹⁾

However, in the case of public goods, we have a very different situation. To the extent that we value public goods, we also realize that getting extra income to buy them will accomplish nothing. There is no market in which we can buy, say, reduced global CO₂ levels or endangered species preservation. We know that we will get whatever level of public goods that we collectively decide to have (ideally, or so it might at first seem, determined according to the conditions laid out by Samuelson). But we also know that any one individual’s desires will have a negligible effect on that. So, because the supply of labor to get the public good would be costly, yet would have no benefits to the supplier, the income that *would have been generated if we could buy public goods as we can private goods is not generated*. In other words, even in a perfect demand-revelation setting, we are in a second-best world, because people are failing to optimally generate the income that would have gone toward public good purchases.

Indeed, and ironically, in a world of heterogeneous preferences, the more important public goods are, relative to private goods, in one’s preferences, the lower will be the observed income.⁽²⁾ Consumers desiring private goods will want to generate the income

⁽¹⁾ Indeed, the point is more general. In any decision involving additional income (for example, schooling or training) the marginal benefits of expected future goods gained are compared with the costs of the training. The focus here, for simplicity, is only on short-run work – leisure decisions.

⁽²⁾ Marginal willingness to pay, in situations involving real money, typically looks small relative to stated willingness to pay in hypothetical situations. Traditionally, this is viewed as ‘hypothetical bias’ in the contingent valuation literature. The argument here suggests that it is the real-world payments that may be ‘wrong’, because the wrong income is being generated. The big ‘hypothetical’ numbers might be *paired with* respondent increases in work effort to get the necessary income

to buy the many things they desire. Rational individuals who care a lot about the environment or any other public good will generate very little income because they cannot get what they want, on the margin, regardless of income generated. Hence they will substitute toward leisure which they *can* affect, until the marginal value of leisure is equated to their (fairly low) marginal valuations of ordinary private goods.⁽³⁾

3 Urban local public goods provision and suburbanization

The discussion of the preceding section ignores the fact that many public goods have an important spatial component and, partly as a consequence, one might wish to explore what difference the presence of private market complements and substitutes for local public goods makes. The use of CO₂ or species preservation in the previous section implicitly assumed that there were no good substitutes (for example, for a warmer planet or for a lost whale). However, in the urban setting, this is no longer the case. Consider, for example, the amenity associated with a local park. If an urban area fails to provide optimal levels of convenient parks for its residents, the residents can leave the city, buying a large suburban backyard as a private good substitute. Hence, the fact that many public goods have a spatial component, suggests that suburbanization can be seen, at least in part, as a spatial reaction to the failure to produce optimal quantities of local public goods in the urban core.⁽⁴⁾

Consider individuals living in the urban core. The urban core typically has many problems (crime, pollution, congestion, noise, lack of open space, etc) that are, at least partially, of a local public good nature. Nonoptimally low levels of such public goods are provided in the urban core, with core residents undergenerating income. Were we to get to the Samuelsonian optima for those goods, those levels would, per section 2, still be too low.

However, were there *perfect* private good substitutes for all public goods, the correct amount of income *would* be generated and the, now larger, flow of private goods would be purchased with the larger incomes generated. Perfect private market substitutes effectively eliminate the public goods problem; income generated would be optimal, the proper amounts of all goods would be produced, and there would be no problems of either demand revelation or input market failure.

There is a substantially greater likelihood that private market substitutes will exist for local public goods than for ordinary public goods (for example, parks versus species preservation). This is because one cannot ‘free ride’ if consumption of the public good requires the purchase of the land offering the public good. The Tiebout (1956) model provides a mechanism by which local public goods can be efficiently

⁽²⁾ continued

(for people generally realize that they have to pay for the things they want). The number recorded by the researcher as current income, in stated preference studies, may not be appropriate. Hence, the stated preference numbers could be more valid than generally presumed—in the proper conceptual experiment.

⁽³⁾ They may also pursue less remunerative occupations (for example, working with environmental groups and the like) that offer an alternative means of impacting provision of public goods.

⁽⁴⁾ A local public good is like a pure public good (for example, national defense, species preservation) in that it is nonexcludable and nonrivalrous. But in addition local public goods have the property that their benefits are received over a ‘relatively small’ geographical area, as, for example, a park or walking access to the city center. Local public goods may be provided by any level of government (for example, school district, county, or city) or, indeed, may be provided by a private developer in a planned urban community, an example being the joint provision of a golf course and housing lining the fairways.

provided if there are sufficiently of many jurisdictions, without important jurisdictional spillovers or scale economies.⁽⁵⁾

It should, however, be emphasized that the prospect of a perfect Tiebout world is extremely dubious in the present urban–suburban spatial context. This follows from the bundling of both nongovernmental (for example, access to the center) and governmental location-specific public goods of a large number and variety. Such bundling inevitably results in an internal contradiction. The contradiction stems from three mutually inconsistent requirements in the model. First, there must be many jurisdictions to allow each household to find its perfect bundle of local public goods. But, with many jurisdictions, cross-jurisdictional spillovers would defeat that efficiency. That is, desires for a specific, say low, level of crime in one jurisdiction would be expected to drive criminals into nearby jurisdictions, or a polluter at the downwind edge of one jurisdiction might prevent those in an adjacent jurisdiction from acquiring the air quality they desire. This problem is compounded by the presence of scale economies in the provision of some local public goods. Hence, jurisdictions will always be either too large, too small (and, typically, both simultaneously, depending on the public good involved), or in the wrong place (for example, too far from the center) from the standpoint of household demands for particular local public goods bundles. By way of illustration, a location that is close to the center, in a school district with high-quality schools, with clean air and spacious parks might well not exist (there will inevitably be nonconvexities, multiple equilibria, and so on).

Although not perfect, however, a Tiebout-like spatial substitute for directly improving levels of urban crime, air quality, noise, and congestion, and so on is provided by the option of moving to the suburbs. As a consequence, many people *will* generate the income to acquire the higher levels of *some* local public goods in the suburban areas. But suburbanization, with associated lengthy commutes and reduced goods consumption choices (restaurants, cultural amenities, etc) is a poor substitute for direct increases in the provision of urban public goods. And, in the cases of some public goods (such as pollution in areas like Los Angeles), suburbanization might fail to deliver much in improved consumption levels, relative to the desires of the residents. The inability to disbundle will imply that the perfect substitute private good does not exist for local central city public goods, hence the problem of section 2; failure to generate the right amount of income remains.⁽⁶⁾

We observe, as a consequence, nonoptimal movements to the suburbs because of the failure to have the right amounts of public goods provided in the urban core—even if we had both perfect demand revelation and perfect information in applying the Samuelson aggregation rule.

⁽⁵⁾ One of the desirable features of the Tiebout model is that it was perceived as a possible solution to the inefficiencies associated with majority voting as a means of deciding the levels of public goods to provide. It is, interestingly, an even better solution to median voter efficiency problems than has been realized. This is because the free-riding median voter will not have generated the optimal level of income, as discussed in section 2. However, voting ‘with one’s feet’ will require that individual households generate the income, in competition with others, to acquire local public goods; hence they will generate that income. But only in a perfect Tiebout world will the proper income be generated, as discussed in the text.

⁽⁶⁾ There are possibly many reasons, unrelated to local public goods provision, why particular households might wish to locate in the suburbs (bigger lots for large families, nearness of the countryside, and so on). This is not critical to the argument, in the sense that the text claims about inappropriate urban local public goods provision are overlaid upon whatever other reasons for suburbanization exist. The qualitative implication that more suburbanization occurs than is socially desirable continues to hold.

More formally, define X_0 , L_0 , S_0 , and Q_0 as the initial quantities of ordinary goods, leisure, lot-size, and the collectively determined public good.⁽⁷⁾ Assume initially that all economic agents are located in a homogeneous central city, effectively ignoring the spatial nature of cities. The representative agent will have failed to give up leisure to generate income to buy public goods, even if that agent has very high marginal benefits of the public good, because doing so will not enable any more of the good to be acquired. City residents will have adjusted leisure (and private goods), whose benefits they can capture by foregoing income, purchasing more of each than is optimal.⁽⁸⁾ That is,

$$U_0 = U(X_0, L_0, S_0, Q_0) \ll U(X^*, L^*, S^*, Q^*) = U^*, \quad (1)$$

where, under independence, $X_0 > X^*$, $L_0 > L^*$, $S_0 = S^*$, $Q_0 < Q^*$, and the double inequality indicates that a potentially large disparity exists between the constrained utility values (when the public good cannot be purchased with incremental income) and the unconstrained values.⁽⁹⁾ Lot size, S , is assumed the same in the initial case because there is a fixed number of people and suburbs do not yet exist.⁽¹⁰⁾

With suburbanization offering a (poor) substitute for in situ provision of public goods, we will observe a utility-enhancing (relative to U_0) increase in optimal levels of income generated to buy in private land markets the higher levels of local public goods available in the suburbs. In the traditional monocentric urban modeling context, the additional income will be buying not only higher (but not necessarily optimal) levels of local public goods, but will also be buying longer commute times, a joint bad. Additionally, suburban residents will be foregoing other desirable aspects of locating at the urban center (restaurant variety, cultural activities, and other density-dependent goods). Because rent compensation would result in equilibration of utility across the city–suburb location choice, we have

$$\begin{aligned} U(X_0, L_0, S_0, Q_0) &< U(X_S, L_S, S_S, Q_S) \\ &= U(X_C, L_C, S_C, Q_C) \ll U(X^*, L^*, S^*, Q^*). \end{aligned} \quad (2)$$

That is, the utility associated with moving to the suburbs will be greater than if that option were unavailable, but there is still a potentially large divergence between utility at either suburb or center vis-à-vis optimal in situ public goods provision. Rents in the suburbs will be bid up from initial agricultural levels, while those in the city will fall relatively, as people move to the suburbs. The lot-sizes will be larger in the suburbs than in the center, that is, $S_S > S_C$. Whether income generated and goods consumption

⁽⁷⁾ Vector notation for both ordinary goods and for public goods is suppressed for simplicity, as is potential capitalization of local public goods in labor markets in addition to land markets.

⁽⁸⁾ The text discussion implicitly assumes that private and public goods are independent (as would be the case with Cobb–Douglas utility functions). More general cases are discussed in Flores and Graves (2001).

⁽⁹⁾ Note that the disparity might have been relatively small when we first began providing public goods collectively. For example, when environmental programs were first promulgated, as the environment quality freely available from nature came to be viewed as inappropriately low, the spread between optimal and actual utility might have been small. But, as income and population have grown over time, the marginal values of public goods *relative* to those of the ever-increasing quantities of private goods has doubtless risen sharply.

⁽¹⁰⁾ It is possible that the portion of private goods that are housing related might imply larger lot sizes before optimal goods provision, though this is not critical to the argument. When suburbs begin to exist, equilibrium lot sizes will become larger at the edge for the usual monocentric reasons.

in the suburbs will be greater than under the case of optimal in situ central public goods provision is ambiguous.⁽¹¹⁾

The problem, at heart, is one of inability to disbundle locally provided public goods, including access in particular. To get higher levels of environmental or school quality by buying a suburban substitute location, one must accept lower levels of access to work and entertainment. Suburbanites are unable to buy exactly what they want, but only some of the things they want, by the very nature of space.⁽¹²⁾

How great is the empirical relevance of the point made here likely to be in the ‘real world?’ The importance of the argument here is apparent from the hedonic compensation that yields the middle equality in equation (2). With fully informed people and no exhaustion of people types on the margin, the rents or property values at the suburban fringe will leave identically situated households with the same utility in the suburbs as those at the urban center. But, those in the center are continuing to consume the suboptimal amounts of the many urban public goods that the suburbanites are receiving; the only reason they are better off is that the movement to the suburbs will have lowered their rents somewhat.⁽¹³⁾

Hence, commuting costs and losses of other urban public goods largely offset the gains in utility from higher levels of public goods associated with moving to the suburbs. People will have generated more income to enable them to move to the suburbs to get the local public goods they so desperately want, but they do not really make themselves much better off. Rather, while greatly changing their behavior in terms of what they buy, people move along a utility surface that is only slightly shifted outward by movement to the suburbs. Hedonic compensation, if full, guarantees that; there will, of course, be sorting with perhaps entire groups exhausted at the center, hence there could be a nonnegligible utility gain to some suburban movers. Provision of the proper amounts of public goods at the center is, however, clearly preferred to observed patterns of suburbanization.

But history is what it is. How do we decide, now that suburbanization has proceeded at nonoptimally high rates for many decades, *where* to supply marginal increments in local public goods? We can improve streetlighting or crime rates in the city or in the suburbs—where should the public goods be supplied? The spatially optimal distribution (of a much larger than the Samuelson optimal total amount) of public goods will favor the central urban area, because for any given marginal cost, any provision there has more marginal benefits because of greater population density.⁽¹⁴⁾

Pondering the longer run implications of providing the urban core public goods in situ rather than requiring the purchase of poor substitutes at the suburban fringe offers

(11) It would seem clear, however, that issues of what is a ‘good’ versus a ‘necessary bad’ that appear in discussions of national income accounting would become quite relevant in this setting. Moreover, the same type of consideration would apply to leisure activities. Consider, for example, increased commuting that draws equimarginally from work and other leisure activities. The increased commuting would both raise GDP and increase official measurements of ‘leisure’, but would hardly be seen as a good thing by typical agents.

(12) Additional trips to the center and lunchtime or after-work activities alter the quantitative, but not qualitative, importance of being unable to generally purchase the desired local public goods bundle.

(13) Indeed, suburban flight may cause added losses in some urban amenities, offsetting rent reductions.

(14) Current population densities, of course, bear no relation to what they would be if urban public goods had been optimally supplied over time. One might speculate that Tokyo would be considered not very densely settled if urban public goods had always been provided optimally, but policy approaches for some public goods (noise pollution, for example) might also involve less-dense living schemes, as could income effects on lot size. The net impact on average urban density would seem to be clear, however.

further insight into the advantages of properly valuing public goods. Were clean air, noise abatement, quality schools, reduced congestion, and the like provided directly in the right amounts in the central core, the net associated cost increases (and income generation increases) might not be as large as one might initially surmise. As mentioned earlier, people are already generating much income to buy poor suburban substitutes for urban public goods. For example, providing quality urban public schools would eliminate both the expenses associated with sending children to private schools currently (for many occupying the urban centers), and would allow those who moved to the suburbs for better schools to return to the center, saving commuting and other costs.

The enhanced values of central locations would raise rents and property values in the center relative to the suburban fringe. This would have the long-run effect of encouraging greater density as capital is substituted for the relatively more expensive land at the center. This, in turn, would facilitate the spread of viable mass transit, including walking, as opposed to the current situation that tends increasingly to discourage alternative transport modes over time. The greater population density would allow many (or perhaps all⁽¹⁵⁾) cities to enjoy the kinds of cultural and restaurant amenities now taken for granted in only a few urban areas. Indeed, if the correct amounts of public goods were provided in central locations, the principle determinant of suburbanization would be the income elasticity of demand for lot size relative to the income elasticity of commuting costs, independent of the local suburban public goods that are at present positively correlated with lot size.

4 Summary and generalizations

An input market free-riding flaw will inevitably accompany the output-market failure resulting from the nature of public goods. The output-market flaw has long been recognized; it is unprofitable to produce goods whose consumption cannot be excluded. And, zero marginal social costs of usage imply that any positive marginal charges would, in any event, nonoptimally discourage individual consumption. But, even governmental provision in what would seem to be the right relative amounts (comparing the vertically summed demands with marginal cost, at a given income) yields the wrong amount of the public good. This is because households will generate the wrong amount of income when generating that income does not allow increments in the public goods they care about. We work to buy the things we want—when we cannot get those things by working, rational people will not work.

In the case of suburbanization, however, people will generate the income that will allow them to buy higher levels of local public goods in the suburbs, but this is a poor substitute for in situ provision of those public goods on efficiency grounds.

The failure to get the right amount of the public good has more general nonspatial implications for the optimal quantities of private goods provided. The suburbanization result generalizes to private goods lacking a spatial nature that are substitutes and complements for public goods. Any private good which is a substitute for a public good will be overproduced relative to its optimal production, were public goods provided in the right (larger) relative amounts. For example, private personal protection goods and services (locks, burglar alarms, protective weapons, costly altered behavior, etc) are overconsumed when the provision of police public good protection is underprovided.

Similarly, though not stressed here, private goods that are complements with public goods will be underproduced when the public goods are not produced in their larger optimal quantities. As the classic example, consider the lighthouse that fails to get built

⁽¹⁵⁾ Some smaller cities might reasonably be expected to disappear when larger urban areas are as desirable as they would be under a longer term scenario.

because the benefits, even with a perfect demand-revelation Samuelson rule, fall a bit short of the costs. Because those benefits are calculated without the income that would have been generated to buy such goods were they available for purchase, it is quite possible that the lighthouse should have been built. But, if the lighthouse were built, the demand for boats (which is, after all, a demand for a stream of *safe* boating experiences) would have been higher. Thus, there will be an underprovision of private goods that are complementary with public goods in the current milieu of free riding in input markets.

Suburbanization has been occurring for many decades. This suburbanization has been widely vilified, but with arguments that were unconvincing to many economists. The present paper provides an argument that suburbanization has proceeded, and continues to proceed, at too rapid a rate. The high growth in income and population (combined with the input market failures emphasized here) after World War 2 would suggest that far higher levels of local goods should have been, and should be, provided in our large urban areas. Producing those local public goods at optimal levels would result in far less suburbanization than has been observed over the past half century.

Acknowledgements. I would like to acknowledge anonymous reviewers for this journal as well as Paul Romer for useful comments on an earlier draft.

References

- Clarke E H, 1971, "Multipart pricing of public goods" *Public Choice* **11** 17–37
- Flores N, Graves P E, 2001, "The valuation of public goods: why we work", manuscript; copy available from the author
- Fujita M, 1989 *Urban Economic Theory* (Cambridge University Press, Cambridge)
- Graves P E, 2001, "Valuing public goods", manuscript; copy available from the author
- Groves T, Ledyard J O, 1977, "Optimal allocations of public goods: a solution to the 'free rider' problem" *Econometrica* **45** 783–809
- Samuelson P A, 1954, "The pure theory of public expenditures" *Review of Economics and Statistics* **36** 387–389
- Tiebout C, 1956, "A pure theory of local expenditures" *Journal of Political Economy* **64** 416–424
- Wheaton W C, 1976, "On the optimal distribution of income among cities" *Journal of Urban Economics* **3** 31–44