



Curbing excess sprawl with congestion tolls and urban boundaries[☆]

Alex Anas^{a,*}, Hyok-Joo Rhee^b

^a State University of New York at Buffalo, Amherst, New York 14260, USA

^b Seoul National University of Technology, Seoul, Republic of Korea

Accepted 31 March 2006

Available online 21 June 2006

Abstract

In modern urban areas jobs and residences are connected by commuting and by discretionary travel and consumers sort in land and labor markets according to idiosyncratic attachment, rents, wages and access to products. Our numerical simulations of a spatial general equilibrium model show that unpriced traffic congestion creates excess sprawl causing daily personal travel to be 8 min or 13% longer than optimal. We juxtapose congestion tolls, favored by economists, and urban boundaries, favored by planners, as two alternative policies for eliminating this excess. As explained in Anas and Rhee [Anas, A. and Rhee, H.-J. 2005, When are urban growth boundaries not second-best policies to congestion tolls?, in press in the *Journal of Urban Economics*], in our dispersed city, a boundary of any stringency is absolutely harmful and is not a second-best policy to congestion tolls. To curb the excess, the planner's boundary must be very stringent causing large distortions in land, labor and product markets while leaving the congestion unpriced. Such a boundary's deadweight loss is 15% of money income and 70 times higher than the benefits of first-best congestion tolls that curb the sprawl by pricing travel efficiently. This result is robust in sensitivity

[☆] Alex Anas was supported by award SES9816816 ("Infrastructure and Metropolitan Development") from the National Science Foundation and a fellowship from the Lincoln Institute of Land Policy. Hyok-Joo Rhee was supported by the Seoul National University of Technology and the State University of New York at Buffalo. The paper was presented at a conference organized by the Regional Science Association International, in a seminar at the Federal Reserve Bank of Philadelphia and at a session of the American Environmental and Resource Economics in the 2005 annual meeting of the Allied Social Science Association. We thank Richard Arnott, Antonio Bento, Jan Brueckner, Peter Gordon, David Pines, Kenneth Small, Tony Smith, the seminar audiences and two anonymous referees for their comments.

* Corresponding author. Tel.: +1 716 645 2121x415; fax: +1 801 749 7805.

E-mail addresses: alexanas@buffalo.edu (A. Anas), rheehj@snut.ac.kr (H.-J. Rhee).

analyses. Boundaries are efficient when compactness is valued, but tolls are still needed to reduce congestion.

© 2006 Elsevier B.V. All rights reserved.

JEL classification: C35; C68; D58; D61; H23; H41; R14; R41; R51; R52

Keywords: Urban sprawl; General equilibrium; Traffic congestion; Urban boundaries

1. Introduction

Burchfield, Overman, Puga and Turner (2006) use satellite photography data to document that the percent of land urbanized in the lower 48 United States has grown from a mere 1.30% of the total land area in 1976 to 1.92% in 1992. The annualized growth rate is 2.48%, 2.5 times the 0.98% annualized growth rate of population in the same period.¹ Virtually all of this growth in urbanized land has occurred in the suburbs, as recently noted by Nechyba and Walsh (2004). This has caused suburban traffic congestion levels to increase and urban open spaces to become scarcer. These trends, in turn, have fueled increasingly popular sentiments against urban expansion, commonly referred to as *urban sprawl*, and the advocacy of more compact cities. The issue is now considered so important that many states such as Oregon, Maryland and New Jersey are implementing aggressive land use policies to curb future urban expansion.

The adoption of automobiles and the availability of roads and highways are among the major drivers of urban sprawl (Downs, 1992; Dunphy, 1997; Glaeser and Kahn, 2004). Nechyba and Walsh (2004) put it as follows:

“It is difficult to imagine large increases in suburbanization without this rise of the automobile, even if other causes have contributed to the sprawling of cities in the presence of the automobile.” (p. 182)²

The key issue is the extent to which sprawl is the outcome of efficient market processes or of market distortions. The part of sprawl that is caused by distortions may be called *excess sprawl* (Brueckner, 2000) and its reduction or elimination is a reasonable policy goal if this can be done without creating other more harmful distortions. The automobile-related distortion causing the excessive sprawling of land use is unpriced traffic congestion. Economists have known for decades that the lack of marginal cost pricing in urban transportation makes road travel cheap (Walters, 1961) and that the misuse of cost–benefit rules by transportation planners has resulted in too much highway-building (Krauss et al., 1976; Wheaton, 1978), causing too much urban expansion into suburban areas where land was initially cheap. Empirical studies of congestion tolls, on the other hand, assumed that land use remained unchanged (Keeler and Small, 1977). On the theoretical side, numerically solvable urban models have been used since the 1970s to examine the effect of hypothetical congestion tolling on urban land use (see, for example, Mills, 1972; Solow, 1972; and, most recently, Wheaton, 1998). However, all of these studies assume

¹ In the most urbanized state, New Jersey, the same percentage changed from 18.09% to only 20.79% and in one of the most rapidly urbanizing, Florida, from 4.44% to 8.90%.

² The growth in households and in incomes has also contributed to sprawl (Margo, 1992), as has the flight from the poor public services and the higher crime rates of the central cities (Cullen and Levitt, 1999). American land values are much lower than those in Western Europe, Japan and South Korea (Mayo, 1998) and since land is a normal good, the income and substitution effects of the cheap land boost the demand for private land lots especially by the wealthier Americans.

artificially that all employment is located at a point or small circle in the center of a city and equally artificially treat only commuting, ignoring discretionary personal travel. On the basis of these stylized models it is generally understood that congestion tolls would make cities more compact eliminating the excess travel time caused by unpriced congestion and the excess sprawl that comes with it.

While economists advocate the corrective pricing of traffic to induce modifications in travel behavior (Small, 1992), planners and policy makers favor land use controls that directly limit the expansion of cities, or high density development, or subsidizing development in central cities or building mass transit systems to compete with highways.³ Virtually all of these anti-sprawl policies advocated by planners are aimed at creating urban areas that are spatially more compact in the long run (Ewing, 1997). It is widely presumed that doing so would likely reduce sprawl by almost any measure. For the same population, a more compact (less sprawled) city would have higher land use densities and lower daily average travel times because average distances between urban sites would be reduced. While planners view the reduction in travel times and distances as a benefit, they may not consider that congestion and the opportunity cost of land within urban areas would also increase. This has been the case in Portland, Oregon (Knaap, 1985) and in South Korea cities (Son and Kim, 1998).

In this paper we will compare the long run costs and benefits of two alternative anti-sprawl policies aimed at reducing the daily average travel time (DATT) per worker which is the measure of sprawl used in this article. One of these is the policy of levying optimal congestion tolls on traffic, favored by economists. The second is a planned urban boundary (also known as an “urban growth boundary” or UGB) that reduces the urban radius and beyond which there is a greenbelt in which urban development is not allowed or is severely limited. A good example is the UGB around Portland, Oregon.⁴

In order to reach credible policy conclusions about the relative effectiveness and efficiency of the two policies, congestion tolls versus urban boundary, we must depart from the tradition of assuming all jobs to be located in the center and of ignoring discretionary travel. To these ends, we will use an appropriate theoretical general equilibrium model of urban land use presented in Section 2 that treats the interactions of land, labor and product markets (Anas and Xu, 1999). In this modern model, jobs and residences are dispersed and mixed throughout the urban area, and their spatial distributions and land use densities are determined endogenously and are variable throughout the city. Workers, in the model, decide where in the urban area to work, how much labor to supply, where to reside and how much land to rent there. Because of idiosyncratic variation in tastes, these decisions generate a wide variety of residence–job location patterns in equilibrium as is commonly observed. While the joint choice of a job and a residence location determines the length of the commute, in our model workers also decide how many trips to make to the various parts of the city in order to purchase the goods and services offered there. Thus, they make trade-offs between commuting and discretionary travel. This gamut of jointly determined travel decisions generates traffic congestion and the level of congestion, in turn, affects travel, location and consumption choices.

³ Kain (1991) has documented how federal subsidies for transit systems have resulted in excessive, even deceptive claims by planners about the benefits and potential patronage of such systems.

⁴ Boulder, Colorado, is another example in the US. London had a greenbelt in the late 1930s, while Moscow, New Delhi, Ottawa and Tianjin have also followed the practice. Zoning and greenbelt policies with similar effects are also common in the United Kingdom’s Town and Country Planning System (Cheshire and Sheppard, 2002) and in other European countries, and have also been tried in Seoul (Lee and Linneman, 1998) and other Korean cities (Son and Kim, 1998).

Using such a model suitable for modern cities, we establish several results. First, we show that as employment and residences decentralize out of the center, commutes are shortened because average distances between jobs and residences are reduced as documented empirically by Gordon, Kumar and Richardson (1989) and as demonstrated in the general equilibrium model of Anas and Xu (1999). But we also show that as employment decentralizes, the number of discretionary trips and the time spent on them increases sharply because more shopping opportunities become available closer to residences.⁵

Second, road planners in our model can finance roads by levying a head tax or by implementing first-best congestion tolls. When tolls are used, roads become self-financing and the initial aggregate over-allocation of land to roads is self correcting in the long run with less roads being allocated to the suburbs and more to the heavily congested centers. In Section 3 we show that the excess sprawl induced by unpriced congestion and initially misallocated roads result in a daily average travel time of commuting plus discretionary round trips per worker that is about 8 min or 13% too long. The congestion tolls eliminate this excess travel or excess sprawl by making the job and residence distribution of the city more compact. In doing so, efficiency gains of about 0.22% of average incomes are achieved.⁶

We compare the effects of tolls to those of an urban boundary as an alternative anti-sprawl policy *designed to realize the same reduction in travel time that is obtained from tolls*. We show that such an equivalent boundary entails a huge greenbelt since the boundary does nothing to directly change the marginal cost of travel. The equivalent boundary is so stringent that it would have to expropriate about 47% of the urban land causing the inward relocation of about 27% of the population. This would entail big upward adjustments in densities, land rents and traffic congestion, especially in the suburban communities where land supply is most curbed by the boundary and where the boundary also reduces the road capacity used in local trips. In our simulations, the deadweight loss of the equivalent urban boundary amounts to 15% of average incomes and is 69.6 times larger than the efficiency gains of tolls. These numbers are found to be extremely robust to sensitivity analyses in which we vary the initial allocation of land to roads and the variance of the idiosyncratic taste dispersion among consumer–workers.

It is important to understand why the effects of the equivalent urban boundary are so dramatic. In contrast to congestion tolls, the boundary is a blunt instrument and does nothing to directly increase the per-mile marginal cost of travel to its socially desirable level. Thus, the boundary fails to modify travel behavior significantly. Instead, the boundary seeks to induce travel time reductions by shortening trip lengths. This is done by curtailing low density residential and commercial land use (and the associated road capacity) in the suburbs. Residential and commercial activities are forced inward into a more compact pattern. This is especially harmful to residents who show an attachment to their suburban communities. Such attachment is captured in our model because residents are horizontally differentiated and matched to their communities of choice according to an idiosyncratic component of preference. Making land use compact increases suburban congestion because about the same number of dispersed commuting and

⁵ The rise of discretionary travel and travel time relative to that of commuting is documented by Nelson and Niles (2000).

⁶ These gains may be compared to those of Arnott and MacKinnon (1978), 0.068% of income, and of Sullivan (1983), 1.91% of income or Segal and Steinmeier (1980) who also find small effects. These are monocentric models which assume that all jobs are located at the city's center or central area, all land is owned by absentee landlords, consider only commuting travel to the center and ignore labor and product markets and idiosyncratic tastes.

discretionary trips must now crisscross a more confined area and with less road capacity in the suburbs.

The features described above are not captured by the traditional monocentric models with congestion mentioned earlier.⁷ In these models there are no discretionary trips, everyone commutes to the center and residential communities are perfect substitutes, residents feeling no attachment to their chosen communities. In such settings, an urban boundary that is not too tight can result in a welfare improvement bounded from above by the first-best efficiency gains of congestion tolls. In an informal commentary on congestion tolls and urban boundaries, Brueckner (2000) stated this as follows:

“...a UGB may be much too stringent, needlessly restricting the size of the city and leading to an inappropriate escalation in housing costs and unwarranted increases in density.” (p. 167)

While this statement recognizes that an urban boundary can easily be set inefficiently (too stringently), it also suggests that a *not-too-stringent* boundary can be a second-best when congestion tolls are not available. This conjecture is derived from a monocentric intuition about cities. The intuition suggests that in a city in which people make no discretionary trips going in all directions but commute to the center only, a not-too-stringent urban boundary would reduce total travel miles to the center resulting in less congestion, keeping the road capacity profile unchanged.⁸ But, as we shall see, in a non-monocentric model such as ours in which discretionary trips disperse in all directions from a place of residence, restricting the size of the city and reducing suburban road capacity by imposing a boundary can increase suburban congestion. Then, we will see that a boundary of any stringency could be inefficient. If residents have idiosyncratic attachment to their suburban communities, as they do in our model, the boundary compounds the inefficiency by raising rents and local congestion even more, forcing some residents to communities with which they are not as well matched and causing the remaining ones to pay higher rents. In Anas and Rhee (2005), a related paper, we carefully discuss some of the reasons why an urban growth boundary is not always a second-best to congestion tolls.

In Section 4, we extend the analysis by assuming that urban workers value the peripheral greenbelt created by a boundary as a pure public good. Since American urban areas are already surrounded by vast open spaces, one should be skeptical of a claim that extra open space at the peripheries of cities has any amenity value. But anti-sprawl planners have offered anecdotal evidence that residents, especially suburbanites, are increasingly concerned about urban sprawl. If so, then the creation of greenbelts by imposing boundaries around cities would be valued because they increase compactness and urban densities. We show that the excess-sprawl-curbing greenbelt of Section 3 would be efficient if *at the margin*, each worker were willing to pay about \$8 per year per acre of greenbelt. Yet, because the greenbelts increase suburban congestion, the need for tolls on traffic is not eliminated and tolls are still an important part of a Pareto efficient strategy when such efficient boundaries are used.

⁷ Bento et al. (2006) model the effects of an urban boundary and of various tax policies on land rents in such a traditional monocentric city. In their model, all production and jobs are assumed concentrated in the city's center and rents are not redistributed. Therefore, anti-sprawl policies benefit absentee landlords. Their model also does not consider congestion or the allocation of land to roads, treats commuting travel but ignores discretionary travel.

⁸ Kanemoto (1977), Arnott (1979) and Pines and Sadka (1985) all showed that in a monocentric city, with an arbitrary exogenous allocation of land to roads and unpriced congestion, the *shadow* rent on land at the edge of the city is lower than the rent on farming. Since the border of the city occurs where urban and farming market rents are equal, this finding suggests that the second-best optimal border should be closer to the center than the market border.

2. The model⁹

We model a circular city in which travel happens only in the radial direction. Since all travel is assumed to be radial, it is sufficient to model an angular *unit section* along a diameter. We divide such a section into 11 concentric zone wedges as shown in Fig. 1. The innermost zone 6 is 2 miles in diameter and straddles the geometric center. The eight zones surrounding zone 6 (four in each direction) are each 2 miles wide and the last zones (one at each end) represent the suburban–exurban fringe and are 5 miles wide. We may think of zones 4–8 as the central city and zones 1, 2, 3 and 9, 10, 11 as the suburbs. At the end of the edge zones (1 and 11) there is a natural boundary (e.g. seashore, mountains) that is an ultimate limit to urban expansion.

As we shall see later, consumer–workers view the land in the 11 zones as horizontally differentiated and imperfectly substitutable areas of distinct character, so that they may be considered towns or “communities” for the purposes of residing, working or shopping there. For example, land may be flat in one zone, hilly in another, treed in a third etc. We ignore intra-zonal differences in land: land parcels within the same zone are identical and viewed as perfect substitutes. Although we assume that travel occurs only in the radial direction, it can cross the center to terminate at the other side of the city. For example, a worker could reside in zone 3 and work (or shop) in zone 7, on the other side. Reverse commuting is also possible so that a worker who resides in zone 5 could hold a job in zone 1.¹⁰ To calculate trip distances and travel times, all trips are assumed to originate from the midline of a zone. In each zone the markets will allocate land to residences, to production and — with the help of planners — to roads. Residential, production, employment and traffic densities and the rent on land are all determined endogenously and, in our symmetric equilibrium, will peak in the geometric center, zone 6, which is the most accessible place to labor and where land is most limited. We may force all production to zone 6 only for comparison with the monocentric city literature.

We assume that there is a total of G_0 acres of open undeveloped space somewhere beyond the edge zones 1 and 11. The land in zones 2–10 can be allocated only to roads, production and residences, while the land in the edge zones has a fourth possible use: some of it, say G acres ($\frac{G}{2}$ on each side) can be kept out of development as a greenbelt by instituting a planned urban boundary as shown in Fig. 1. In that case, the total amount of open space becomes $G_0 + G$. In the absence of a boundary policy all of the land in the edge zones 1 and 11 will be in urban use and the city will extend all the way to the natural limit. Under a boundary policy the planner can choose the area G and the boundary stringency this implies so as to maximize welfare when workers care either about having extra open space at the urban fringe or, equivalently, when they value city compactness or dislike *urban sprawl*. However, planners can and do institute growth boundaries even when workers do not care. Workers in our model do not visit the open space whether they care about G or not. In the case that they do care, G is a *pure public good* that can be enjoyed by all simply because it is there. Since no one visits it, putting it somewhere other than the fringe would needlessly lengthen travel through the open space without creating any offsetting benefits.

⁹ The Appendix contains a full notational glossary and Table 1 lists the key parameter values we have used in the simulations reported here.

¹⁰ It has been known, for decades, that both reverse commuting and crossing the “center” are very common. For example, one can reside in Manhattan and work in New Jersey (reverse commuting) or reside in New Jersey and work in Queens, crossing Manhattan daily (crossing the center). Both phenomena are ignored in all formal urban models except in Anas and Xu (1999) and the one we use here.

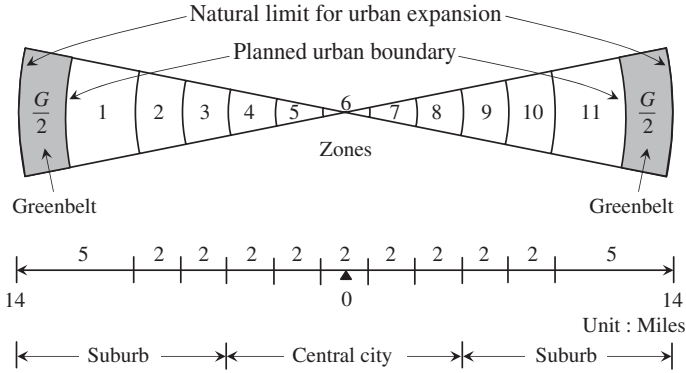


Fig. 1. Unit section of circular city.

2.1. Producer equilibrium

We assume that a distinct composite good is produced and sold in each of the 11 zones. Pretend that each composite good is sold at a community shopping center located in the midline of the zones in which it is produced. Workers can buy it only by traveling there. Production is constant returns to scale and Cobb–Douglas, labor and land being the only two substitutable inputs. Hence, the number of firms in a zone is indeterminate. Let $\mathbf{X}=(X_1,\dots,X_{11})$ be the vector of aggregate outputs of the composite goods produced in the zones and M_v and Q_v be the aggregate labor and land inputs in zone v . The production function in zone v is $X_v=EM_v^\delta Q_v^\mu$ where E is the scale parameter and $\delta+\mu=1$. Let r_v be the land rent, w_v the wage rate and p_v the mill price of the good produced in zone v . Then the conditional input demands for labor and land are:

$$M_v(p_v, w_v, X_v) = \frac{\delta p_v X_v}{w_v}, \tag{1}$$

$$Q_v(p_v, r_v, X_v) = \frac{\mu p_v X_v}{r_v}. \tag{2}$$

Free entry in each zone ensures zero profit and price equals unit average (and marginal) cost. Hence,

$$p_v = \frac{w_v^\delta r_v^\mu}{E \delta^\delta \mu^\mu}. \tag{3}$$

2.2. Worker (consumer) equilibrium

Consider a worker residing in zone i and working in any zone j . (i, j) will be such a worker’s commuting arrangement and his choices — as we shall see — are conditional on it. He derives utility from lot size q_{ij} rented at the zone of residence i , from leisure l_{ij} , from the quantities of the 11 zone-specific composite goods Z_{ijv} which he shops by traveling to their zones of production v , and — as already explained — from the open space $G_0 + G$, G_0 outside the urban area and G in the edge zones. Each worker has H hours per year to allocate to work, travel and leisure. The number

of work days per year (two-way commutes) is d . The costs of travel per trip from zone i to j organized in a matrix, are time spent $\mathbf{g}=[g_{ij}]$ and congestion tolls paid $\mathbf{t}=[t_{ij}]$ per trip, respectively. Commuting and shopping require two-way trips and each originates from and ends at home. Total travel time per consumer is then given by $T_{ij} = 2dg_{ij} + \sum_{\forall v} 2g_{iv}f/z_{ijv}$ for a worker with home–work pair (or commuting arrangement)(i, j), where f is a constant: the number of trips per year needed to purchase a unit z_{ijv} . $p_{ijv} \equiv p_v + 2f(t_{iv} + w_j g_{iv})$ is the effective price (full opportunity cost) per unit of the composite good purchased from v for a worker residing at i and working at j , inclusive of the full opportunity cost of the required travel (time plus tolls). As explained in our discussion of the firm, r_i is the land rent in i , and w_j is the wage rate earned by working in zone j .

For a given home–work pair or commuting arrangement (i, j), the worker solves the utility maximization problem:

$$\max_{\{z_{ijv}, q_{ij}, l_{ij}\}} U_{ij} = \eta \ln(G_0 + G) + \alpha \ln \left(\sum_{\forall v} z_{ijv}^\rho \right)^{1/\rho} + \beta \ln q_{ij} + \gamma \ln l_{ij} + u_{ij}, \tag{4}$$

$\eta \geq 0, 0 < \alpha, \beta, \gamma < 1, \alpha + \beta + \gamma = 1$, subject to the budget constraint $\sum_{\forall v} p_{ijv} z_{ijv} + r_i q_{ij} + w_i l_{ij} = \Omega_{ij}$, where Ω_{ij} , the *full endowment income* of the worker net of the *full cost of commuting*, is the sum of potential wage income minus the opportunity cost of commuting, plus rent dividends from land, D :

$$\Omega_{ij}(w_j, t_{ij}, g_{ij}, \mathbf{r}, G) = w_j H - 2d(t_{ij} + w_j g_{ij}) + D(\mathbf{r}, G). \tag{5}$$

Each worker earns income and owns an equal share of the entire land of the city, collecting a net-of-tax dividend $D(\mathbf{r}, G)$.¹¹ Let A_k be zone k 's total area. Then $D(\mathbf{r}, G) = \frac{\sum_{k=1,11} r_k (A_k - \frac{G}{2}) + \sum_{k=2, \dots, 10} r_k A_k}{N} - h$ is the net-of-tax rent dividend for each worker,

where h is a head tax per worker if such a tax is used to pay for roads. The consumer–worker's *money income*, on the other hand, is his labor supply (total time endowment net of all travel time and leisure) multiplied by the wage rate, plus the rent dividend. We denote money income as $I^{ij} = w^j (H - T^{ij} - l^{ij}) + D(\mathbf{r}, G)$ for a worker employed at j and residing in zone i . So the budget constraint can be written as $\sum_{\forall v} (p_v + 2ft_{ij})z_{ijv} + r_i q_{ij} + 2dt_{ij} = I_{ij}$, equivalent to (5).

By assuming $0 < \rho < 1$ in the utility function, the marginal utility of each composite good at zero consumption is infinite, so that at any prices, the worker is always better off by consuming a strictly positive amount of the composite good from each zone. This reflects the worker's taste for variety as in Dixit and Stiglitz (1977). This means that in this model, as in Anas and Xu (1999), the consumer travels and shops in all the zones to satisfy his taste for variety over the zone-specific goods. This is a force which, without assuming increasing returns, causes an agglomeration in which employment and residential densities will peak at the geometric center of the space where accessibility to all the zones is maximized.

Where appropriate, below, we use bold letters to refer to the vector of rents, wages and the mill and effective prices of products: $\mathbf{r} \equiv (r_1, \dots, r_{11})$, $\mathbf{w} \equiv (w_1, \dots, w_{11})$, $\mathbf{p} \equiv (p_1, \dots, p_{11})$, $\mathbf{p}_{ij} \equiv (p_{ij1}, \dots,$

¹¹ Note that unlike the partial equilibrium models of standard urban economics, ours is a fully closed general equilibrium model. There is no exogenous non-urban use of the land. What is distributed to the workers is the entire land rent not the "differential land rent" of the partial equilibrium models. In these partial models, it is assumed that land is purchased from the farmers at its agricultural price, but the rental income obtained by the farmers is not included in the welfare criterion.

p_{ij11}). For a given home–work pair (i, j) and open space $G_0 + G$, the Marshallian demands for goods, land and leisure are obtained by solving (4). They are:

$$z_{ijv} \equiv z_{ijv}(\mathbf{p}_{ij}, \Omega_{ij}) = \alpha \frac{p_{ijv}^{1/(\rho-1)}}{\sum_n p_{ijn}^{\rho/(\rho-1)}} \Omega_{ij}, \tag{6}$$

$$q_{ij} \equiv q_{ij}(r_i, \Omega_{ij}) = \beta \frac{\Omega_{ij}}{r_i}, \tag{7}$$

$$l_{ij} \equiv l_{ij}(w_j, \Omega_{ij}) = \gamma \frac{\Omega_{ij}}{w_j}. \tag{8}$$

The indirect (optimized) utility function is then $U_{ij} = V_{ij}^* + u_{ij}$ where

$$V_{ij}^* \equiv V_{ij}(\mathbf{p}_{ij}, r_i, w_j, \Omega_{ij}, G) = \eta \ln(G_0 + G) + \ln \Omega_{ij} - \frac{\alpha(\rho-1)}{\rho} \ln \left(\sum_{\forall v} p_{ijv}^{\frac{\rho}{\rho-1}} \right) - \beta \ln r_i - \gamma \ln w_j. \tag{9}$$

The u_{ij} are idiosyncratic constants measuring attachment to commuting arrangement (i, j) and will now be explained. Each worker compares all 121 home–work pairs (i, j) choosing the most preferred. Therefore, commuting arrangement choices are described by a probabilistic model of sorting: $\Psi_{ij}(\mathbf{p}, \mathbf{r}, \mathbf{w}, \Omega) = \Pr[U_{ij} > U_{nm}; \forall (n, m) \neq (i, j)] = \Pr[V_{ij}^* + u_{ij} > V_{nm}^* + u_{nm}; \forall (n, m) \neq (i, j)]$. Ψ_{ij} is the probability that a randomly selected worker most-prefers the home–work pair (i, j) . Assuming the u_{ij} are i.i.d. and Gumbel with mean zero, variance σ^2 and dispersion parameter $\lambda = \frac{\pi}{\sigma\sqrt{6}}$, the choice probabilities of sorting among the commuting arrangements will be multinomial logit (McFadden, 1973):

$$\Psi_{ij}(\mathbf{p}, \mathbf{r}, \mathbf{w}, \Omega) = \frac{\exp \lambda V_{ij}^*}{\sum_{\forall (n,m)} \exp \lambda V_{nm}^*}, \sum_{\forall (i,j)} \Psi_{ij}(\mathbf{p}, \mathbf{r}, \mathbf{w}, \Omega) = 1. \tag{10}$$

The dispersion parameter, λ , is important. At one extreme, as $\lambda \rightarrow \infty$ ($\sigma \rightarrow 0$), taste idiosyncrasies vanish and all workers choose identically. In this case, the Ψ_{ij} corresponding to the highest U_{ij} goes to one and all others go to zero, unless there are ties for the highest U_{ij} in which case the tied (i, j) are chosen with the same probability. At the other extreme, as $\lambda \rightarrow 0$ ($\sigma \rightarrow \infty$), idiosyncrasies swamp the systematic part of utility, V_{ij}^* , that depends on prices. Then workers choose randomly (the Ψ_{ij} are all equal to $1/121$). The case of finite λ (σ) has empirical validity. We noted earlier that in observing actual commuting patterns, it is easy to see that many possible commuting arrangements are used. Such a pattern known as *the wasteful commuting hypothesis*, first noted by Hamilton (1982), is readily explained by assuming idiosyncratic tastes as we do, but cannot be explained using the uniform-tastes assumption of the traditional models which are equivalent to assuming that $\lambda = +\infty$ ($\sigma = 0$), ignoring idiosyncratic tastes. The dispersal in the population of consumer–workers, of idiosyncratic preferences over the commuting arrangements is a force which in addition to the taste to shop in every zone and without increasing returns, imparts an agglomeration advantage to the geometric center if more workers prefer to work at the center and reside in the periphery than preferring to reside and work in the periphery.

To measure overall welfare, we use the expected value of the maximized utilities. Under the Gumbel distributional assumptions made above, welfare is:

$$\begin{aligned}
 W(G, \mathbf{p}, \mathbf{r}, \mathbf{w}, \mathbf{\Omega}) &= E \left[\max_{\forall(i,j)} (V_{ij}^* + u_{ij}) \right] = \frac{1}{\lambda} \ln \sum_{\forall(i,j)} \exp(\lambda V_{ij}^*) \\
 &= \sum_{\forall(i,j)} \Psi_{ij} V_{ij}^* + \left(\frac{\sigma\sqrt{6}}{\pi} \right) \sum_{\forall(i,j)} \Psi_{ij} (-\ln \Psi_{ij}).
 \end{aligned}
 \tag{11}$$

The first term on the right side is the mean of the non-idiosyncratic utilities. The second term (positive) is the mean utility surplus from taste dispersion which increases with the variance of the idiosyncratic tastes, because more dispersion means higher upsides for the u_{ij} , inducing higher levels of maximized utilities on average. Also note from (11) that, other things being equal, adding distinct commuting arrangements improves expected welfare as a higher variety of home–work pairs means that idiosyncratic tastes are better matched on average. When we force all production to zone 6 and then relax that constraint, the number of (i, j) pairs utilized in equilibrium will jump from 11 to 121.

2.3. Transport equilibrium

The model’s transport sector requires three processes to be mutually consistent. *First*, zone-to-zone traffic flows must be derived from the choices of workers and firms discussed above. *Second*, congestion tolls or a head tax must be computed appropriately in each case so that the travel times and monetary costs are, in fact, the ones perceived by workers. *Third*, the allocation of land to roads must be determined. We discuss each process.

2.3.1. Traffic flows

Two types of daily traffic flows from any residence in i to any destination in j : commuting trips $F_{ij}^w = N\Psi_{ij}$ where j is the place of work and shopping trips $F_{ij}^s = \frac{1}{d} \sum_{\forall s} N\Psi_{is}f_{z_{isj}}$, where j is a shopping zone (s being the workplaces of workers residing in i). N is the number of workers in the city. The total daily flow from i to j is $F_{ij} \equiv F_{ij}^w + F_{ij}^s$. We ignore all trip-scheduling issues, assuming that the traffic is uniform throughout the day. Then, $F_i = F_{ii} + \sum_{\forall j \neq i} (F_{ij} + F_{ji}) + 2 \sum_{n=1}^{i-1} \sum_{m=i+1}^{11} (F_{nm} + F_{mn})$ is the traffic in an interior zone ($1 < i < 11$). The first term is intra-zonal traffic and the second is traffic originating and ending in the zone. These are assumed to travel half a zone length. The third term is the sum of traffic flows passing through that zone. It is factored by two because it traverses the entire zone. When i is an edge zone ($i=1$ or $i=11$) there is no through traffic so the last term is dropped.

2.3.2. Travel times and congestion tolls

Let g_i be the travel time per mile of a trip in zone i and let t_i be the congestion toll per mile of a trip in zone i . As noted, intra-zonal trips are assumed to traverse half a zone length. Denoting zone length as Δ_i (for edge zones it is net of the width of area $G/2$), $g_{ii} = (\Delta_i g_i / 2)$, $t_{ii} = (\Delta_i t_i / 2)$ for intra-zonal, and $g_{ij} = (\Delta_i g_i + \Delta_j g_j) / 2 + \sum_{s=i+1}^{j-1} \Delta_s g_s$, $t_{ij} = (\Delta_i t_i + \Delta_j t_j) / 2 + \sum_{s=i+1}^{j-1} \Delta_s t_s$ for inter-zonal trips ($i \neq j$). Suppose that road capacity K_i in zone i is proportional to R_i , the land allocated to roads in zone i , so that $K_i \equiv \zeta R_i$, ($\zeta > 0$). The time needed for a trip to travel a mile in zone i is given by the congestion function $g_i = a \left[1 + b \left(\frac{t_i}{K_i} \right)^c \right]$, where b and c are positive coefficients and a is the

inverse of the free (of congestion) traffic speed. For each trip, the planner charges a traveler crossing zone i the congestion toll t_i per mile (unless, as we shall see, the roads are financed by a head tax). The efficient toll is given by the difference between the marginal social time–cost that the traveler is imposing on other travelers and his private time–cost. Since $g_i F_i$ hours per mile are spent by the traffic in i , the marginal social time–cost is $\partial(g_i F_i) / \partial F_i = abc(F_i / K_i)^c + g_i$. Therefore, the first-best efficient congestion toll per mile, *in units of hours* in zone i , is $\hat{t}_i \equiv \partial(g_i F_i) / \partial F_i - g_i = abc(F_i / K_i)^c$. For the correct monetary toll, we value \hat{t}_i at the average wage rate of those travelers delayed by the traffic in i . $\bar{w}_{ij} = [w_j F_{ij}^w + \frac{1}{d} \sum_{n \neq i} w_n N \Psi_{in} z_{inj}] \div F_{ij}$ is the average wage rate of those who flow from i to j , the weights being the shares of work and shopping trips. Then, the average wage rate of the travelers crossing zone i is $\bar{w}_i = [\bar{w}_{ii} F_{ii} + \sum_{j \neq i} (\bar{w}_{ij} F_{ij} + \bar{w}_{ji} F_{ji}) + 2 \sum_{n=1}^{I-1} \sum_{m=i+1}^{11} (\bar{w}_{nm} F_{nm} + \bar{w}_{mn} F_{mn})] \div F_i$ (for an edge zone, the third term in the last equation is dropped out because there is no through traffic), and finally the monetary toll such a traveler should be charged is $t_i \equiv \hat{t}_i \bar{w}_i = abc(F_i / K_i)^c \bar{w}_i$.

2.3.3. Land allocation to roads

Since the aggregate time–cost of travel is homogeneous of degree one in traffic volume and in the land allocated to roads, it is well known that tolls will break even on each piece of road. Therefore, R_k^* the land allocated to roads in zone k is calculated so that toll revenues in zone k are equal to the cost of renting the land at its opportunity cost, r_k . Thus $dF_k t_k \Delta_k = r_k R_k$ from which $R_k(r_k, \mathbf{w}, F_k) = (\frac{\Delta_k abc \bar{w}_k}{r_k})^{\frac{1}{1-c}} F_k$, and $t_k = (\frac{abc \bar{w}_k}{(\Delta_k \Delta_k)})^{\frac{1}{1-c}} r_k^{\frac{1}{1-c}}$. When congestion tolls are not levied, the amount of land in roads is determined arbitrarily by the planner (as described earlier) as some exogenous \hat{R}_k . In that case we assume that a head tax, h , is levied to finance roads so that the citywide road budget is balanced by $Nh = \sum_k r_k \hat{R}_k$.

2.4. General equilibrium

General equilibrium will be computed for a completely closed city. This means that there is no trade with other cities, that the city’s population, N , is exogenous and that aggregate rent is redistributed to workers who are the initial owners of all the land. At general equilibrium, some urban boundary and, hence, greenbelt G , may be set by the planner (see Fig. 1). Subject to this, the market for land, for labor and for the zone’s output must clear in each zone, firms must make zero economic profits and, *at the same time*, these markets and the transport sectors must be in mutual equilibrium, as discussed in Section 2.3. In the case where the head tax is used to finance roads, transport sector equilibrium determines the head tax so that the city-wide road budget is balanced, and the traffic flow and corresponding congested travel time in each zone are such that the traffic flows are consistent with the firm and worker choices from the labor, land and product markets. In the case where congestion tolls are used to finance roads, the optimal toll in each zone is determined as was discussed in Section 2.3 with the road budget being balanced in each zone.

Suppose that we solved the optimization problems of workers and firms under either tolls or head taxes as was described in Sections 2.1 and 2.2. We will then have obtained the demand functions (1), (2) of the firms and (6)–(8) of the workers and the full income functions (5) $\Omega = [\Omega_{ij}]$, for each work–residence pair. Using these in the zero excess demand conditions for *each zone k*:

$$X_k - N \sum_{\forall(i,j)} \Psi_{ij}(\mathbf{p}, \mathbf{r}, \mathbf{w}, \Omega) z_{ijk}(\mathbf{p}_{ij}, \Omega_{ij}) = 0 \quad (\text{product market}), \tag{12}$$

$$N \sum_{\forall i} \Psi_{ik}(\mathbf{p}, \mathbf{r}, \mathbf{w}, \boldsymbol{\Omega})(H - T_{ik} - l_{ik}(w_k, \Omega_{ik})) - M_k(p_k, w_k, X_k) = 0 \text{ (labor market)}, \tag{13}$$

$$\hat{A}_k - N \sum_{\forall j} \Psi_{kj}(\mathbf{p}, \mathbf{r}, \mathbf{w}, \boldsymbol{\Omega}) q_{kj}(r_k, \Omega_{kj}) - Q_k(r_k, p_k, X_k) - R_k = 0 \text{ (land market)}. \tag{14}$$

Zero profit conditions given by (3) also hold in *each zone* and we restate them here for completeness:

$$p_k - E^{-1} \delta^{-\delta} \mu^{-\mu} w_k^{\delta} r_k^{\mu} = 0 \text{ (zero profits)}. \tag{15}$$

In the case where roads are optimally allocated and financed by congestion tolls, the road allocation, R_k , used in (14), is given by the optimal road function $R_k(r_k, F_k)$ derived in Section 2.3. In the case where the head tax is used, $R_k = \hat{R}_k$, an arbitrary road allocation. In (14), $\hat{A}_k = A_k$ for $k=2, \dots, 10$ and $\hat{A}_k = A_k - \frac{G}{2}$ for $k=1, 11$.

Given the open space G , the equilibrium conditions (12)–(15) together with the transport sector equilibrium depend on just four vectors of endogenous variables: \mathbf{X} , \mathbf{r} , \mathbf{w} , and \mathbf{p} which are the outputs, rents, wages and product prices in each zone. We have 11 goods markets and $2 \times 11 = 22$ factor (land and labor) market equilibrium conditions as well as the 11 zero profit conditions. In total, therefore, we have 44 equations with the same number of unknowns.¹² Observe from Section 2.3 that traffic flows are homogeneous of degree zero in (\mathbf{r}, \mathbf{w}) and that congestion tolls are homogeneous of degree one in (\mathbf{r}, \mathbf{w}) . It can be further verified by inspection that the excess demand system (12)–(15) is homogeneous of degree zero in $(\mathbf{r}, \mathbf{w}, \mathbf{p})$ and that the Walras Law holds. Relative prices can be computed for an equilibrium to be supported, but the price level is arbitrary. Instead of determining the price level via a numeraire price, we adjust the price level so that the average money income $I \equiv N \sum_{\forall (i,j)} \Psi_{ij} [w_j (H - T_{ij} - l_{ij}) + D]$ matches the same value in each solution of the model.¹³

2.5. Welfare analysis

The land use planner’s role will be to set the urban boundary, the stringency of which determines the greenbelt allocation, G . We will examine two kinds of planning behavior. In Section 3 we assume that $\eta=0$ and the planner finds a $G>0$ so that DATT, defined later as the daily average travel time in the city and the measure of sprawl, meets a target level which is the value of DATT when $G=0$ but congestion tolls are levied. In Section 4, $\eta>0$ and the land use planner finds $G^* \geq 0$ to maximize the value of the expected welfare given by (11) subject to the markets and the transport planner satisfying the general equilibrium conditions. Let $W^o \equiv W(G^o, \mathbf{p}^o, \mathbf{r}^o, \mathbf{w}^o, \boldsymbol{\Omega}^o)$ be the expected welfare evaluated at some base (o) equilibrium and let $W^p \equiv W(G^p, \mathbf{p}^p, \mathbf{r}^p, \mathbf{w}^p, \boldsymbol{\Omega}^p)$ be the welfare at a policy (p) equilibrium (with tolls or boundary or both). Since a unique path integral from $(G^o, \mathbf{p}^o, \mathbf{r}^o, \mathbf{w}^o, \boldsymbol{\Omega}^o)$ to $(G^p, \mathbf{p}^p, \mathbf{r}^p, \mathbf{w}^p, \boldsymbol{\Omega}^p)$ cannot be calculated, welfare changes can be measured either by calculating an equivalent variation, EV, or a

¹² The algorithm that solves for the general equilibrium is written in GAMS and is described in detail in an unpublished Appendix which is available upon request from Alex Anas. The algorithm finds an equilibrium by means of an extremely accurate iterative procedure that continues until the values of all endogenous variables obtained from successive iterations are within $1 \times 10^{-12}\%$ and each excess demand is also within $1 \times 10^{-12}\%$ of the level of demand. Convergence is almost always achieved in 150–250 iterations of \mathbf{X} , \mathbf{r} , \mathbf{w} and \mathbf{p} .

¹³ We do not prove uniqueness of the equilibrium analytically, but we have explored it numerically. Starting from different starting points, the algorithm converges to the same equilibrium under the same parameters. This holds for a broad range of parameters.

compensating variation, CV, or by an approximated worker (consumer) surplus change. To find the EV or the CV we would solve for them from $W(G^o, \mathbf{p}^o, \mathbf{r}^o, \mathbf{w}^o, \Omega^o + \text{EV}) = W(G^p, \mathbf{p}^p, \mathbf{r}^p, \mathbf{w}^p, \Omega^p)$ and from $W(G^o, \mathbf{p}^o, \mathbf{r}^o, \mathbf{w}^o, \Omega^o) = W(G^p, \mathbf{p}^p, \mathbf{r}^p, \mathbf{w}^p, \Omega^p - \text{CV})$, respectively. For a welfare improving policy, we expect EV and $\text{CV} > 0$. Our approximation of the worker–consumer surplus change is $\Delta \text{CS} = \frac{W^p - W^o}{\text{MUI}}$ where MUI is a calculated marginal utility of income. Since utility is not linear in income, the marginal utility of income, $\frac{1}{\Omega_{ij}}$ varies by the policy as well as by (i, j) . We calculate the MUI as the weighted average:

$$\text{MUI} \equiv \frac{1}{2} \sum_{\forall(i,j)} \left[\Psi_{ij}^p \left(\frac{1}{\Omega_{ij}^p} \right) + \Psi_{ij}^o \left(\frac{1}{\Omega_{ij}^o} \right) \right].$$

2.6. Calibration

The chosen parameters are shown in Table 1. The predetermined open space $G_0 = 200,000 \text{ m}^2$ (or 49.42 acres) which is quite arbitrary and inconsequential. $G_0 > 0$ is needed to avoid the outcome in the utility function that $\ln(G_0 + G) = -\infty$ when $G = 0$. The 1.5° wedge (unit section) of Fig. 1 contains 5000 workers. If each worker supports two dependents, then a fully circular urban area would have 1,800,000 residents. According to the Bureau of Economic Analysis statistics, wage income in 2002 accounted for 86% and rent revenues for 14% of national income excluding

Table 1
Base calibrated values of parameters

Geography (see Fig. 1)

Unit section radians: 1.5° in each direction
Zone length: 2 (5) miles for each inner (edge) zone
Maximum possible radius of the city: 14 miles
 $G_0 = 200,000 \text{ m}^2$ (49.42 acres)

Production

$\delta = 0.86$ (labor cost share), $\mu = 0.14$ (land cost share), $E = 1$ (scale factor)

Resident–workers

$N = 5000$ resident–workers in a unit section (see Fig. 1)
Assumed population. $3N = 15,000$ residents per unit section
 $d = 250$ work days per year
 $H = 250 \text{ days/year} \times 24 \text{ h/day} = 6000 \text{ h/year}$ time endowment
 $\alpha = 0.36$, $\beta = 0.15$, $\gamma = 0.49$, $\rho = 0.6$, $\lambda = 4.0$
See Table 5 for sensitivity analysis of λ (10.0, 15.0 and 20.0)
 $\eta \geq 0.0$ (0 in Section 3, > 0 in Section 4, see Table 6)
Average annual money income = \$40,000

Transport (traffic congestion function)

$a = 1/45 \text{ h per mile}$ (45 miles per hour),
 $b = 50$, $c = 2$, $f = 1/13$, $\zeta = 1.1$

Road allocation (see Fig. 2)

10% below efficient level in zones 4–8 (central city),
10% above efficient level in zones 1–3 and 9–11 (suburbs).
See Table 4 for sensitivity analysis of the road allocation

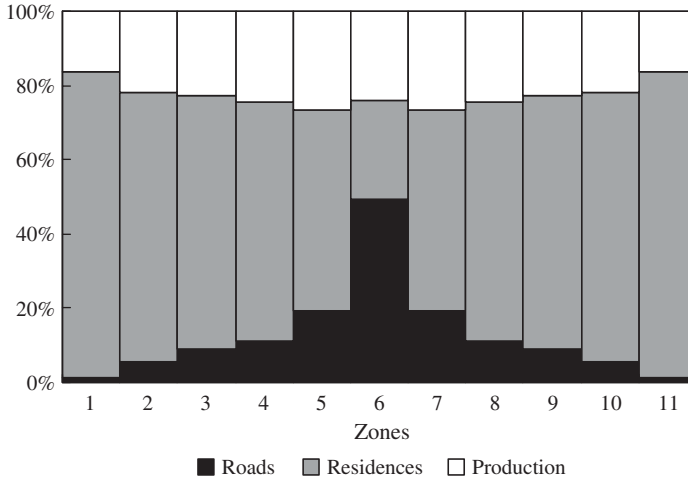
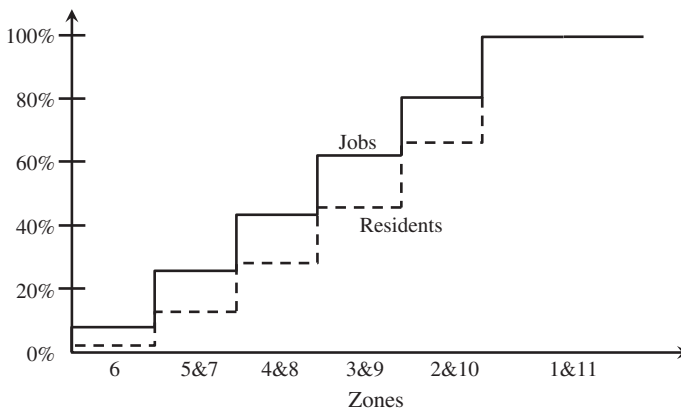


Fig. 2. Land use.

firm profits, interest income and taxes. Preferences are set so that 36% of the full endowment (but not monetary) income goes to consumption goods, 15% to land (this may be reasonable given that durable housing, missing from our model, maybe considered as complementary to land) and 49% to leisure. The elasticity of substitution among the zone-specific commodities is 2.5 ($\rho=0.6$). Travel speed free of congestion is 45 miles per hour. The congestion function, stated earlier, resembles the Bureau of Public Roads form with minor differences in coefficient values. We assume that in the base situation, roads are allocated and funded inefficiently as explained in the Introduction. More precisely, we conjecture that as cities have expanded outward from their centers pre-World War II, planners initially did not anticipate the extent of future urban growth and the growth in auto use, allocating too little land to roads. Later as city expansion reached areas that are now suburbs, cheap land, the rapid adoption of the automobile and unpriced congestion encouraged excessive road building creating an over-allocation of land to roads in the suburbs. To



NOTE: Residents = 3 × resident-workers

Fig. 3. Cumulative distribution of residents and jobs.



NOTE: Residents = 3 × resident-workers

Fig. 4. Gross job and residential densities.

reflect this reality, we assumed that land-in-roads in the central city (in each of zones 4–8) is ten percent below first-best efficient, while in the suburbs (zones 1–3 and 9–11) roads are allocated ten percent more land than is first-best efficient. We assume that, in the base situation, a head tax is used to fund this suboptimal road pattern. This inefficient road allocation pattern will be separately set for a city with all employment centralized as well as a city with dispersed employment.

Figs. 2–5 present aspects of the *Base Equilibrium Land Use Pattern*, based on the parameters and assumptions of Table 1. In the *Base*, roads are allocated inefficiently and funded by the head tax as explained, $\eta=0$, and there is no boundary. The base equilibrium land use profile of Fig. 2 illustrates that the percentage of land in each zone allocated to roads falls sharply as a function of the zone’s distance from the center, while the land allocated to residences rises with employment falling only somewhat. The overall pattern is consistent with real cities in which about half of the jobs are in the central city and the other half in the suburbs. Fig. 3 shows the



NOTE: Residents = 3 × resident-workers

Fig. 5. Net job and residential densities.

cumulative distribution of jobs and residences by distance, from which we can see that the job distribution is more centralized. Figs. 4 and 5 are the profiles of the gross and net residential and job densities. Net density is measured as residents (or jobs) in a zone divided by the aggregate land in that zone occupied by residents (or jobs). Gross density is the number of worker–residents (or jobs) divided by the zone’s total area.

3. Congestion tolls versus urban boundaries

As explained, we want to compare the effects of congestion tolls and urban boundaries as alternative anti-sprawl policies. In this section workers do not directly care about urban compactness. Hence, we set $\eta=0$ in the utility function (4). The case of $\eta>0$ will be examined thoroughly in the next section. Since $\eta=0$, all resource misallocation arises solely from the unpriced congestion. If first-best congestion tolls are levied as discussed in Section 2.3 (the economist’s solution to the excess sprawl problem), the road misallocation is corrected, travel is reduced and excess sprawl is eliminated. The key questions are: *if an urban boundary (excess travel time) is used instead of congestion tolls (the planner’s solution to excess sprawl), how tight a growth boundary is required to reduce the excess sprawl to zero? How much of the urban land does such a boundary expropriate into a greenbelt? How big are the changes in rents and densities inside the boundary? And, how big is the deadweight loss of using the equivalent boundary, relative to the efficiency gains of congestion tolls?*

The efficient allocation is found by solving for the equilibrium in which the congestion tolls are levied to replace the head tax and roads are allowed to adjust so that in the long run the tolls collected in each zone cover the rent on the land used for roads in that zone. This reflects the well-known fact that, with a constant returns traffic flow technology, tolls are the correct mechanism for decentralizing the first-best allocation of land to roads. When an urban boundary is imposed, the inefficient road allocation is not changed and roads continue to be funded by the head tax. This is to reflect that, in the absence of tolls, planners cannot optimize the road profile because they do not know the utility functions of the consumers.

As we have already noted, we will measure sprawl as the daily average travel time (DATT) per worker. This is computed as the total annual time spent for commuting and making discretionary (shopping) trips divided by the number of workers in the city and by the number of days in a year. More precisely, it is:

$$\text{DATT} \equiv \sum_{\forall(i,j)} \Psi_{ij} \left[2g_{ij} + d^{-1} \sum_{\forall v} 2g_{iv} f_{z_{ijv}} \right]. \quad (16)$$

Measured this way, sprawl is the opposite of compactness that urban planners seem to be advocating as a social goal. Since, in the base situation, travel is not priced at its full social marginal cost, the resulting general equilibrium urban form will be too sprawled and this will result in excessive DATT.

Table 2 is the road map for the simulations that are reported in Sections 3 and 4. The base equilibrium results, based on the parameters in Table 1, are summarized in Tables 3a and 3b. These tables present urban area as well as zonal averages or aggregates and show how the congestion tolls and the boundary modify the base situation. The tables are for the case in which all production and jobs are forced to zone 6 (Table 3a) and the case in which they are allowed to disperse and mix with residences in all eleven zones (Table 3b). The former case is intended to mimic a monocentric urban form with zone 6 as the CBD in which all commuting and shopping

Table 2
Simulations of Sections 3 and 4

| Simulation | | Preference for greenbelt, η | Greenbelt, G | Road finance | Road allocation |
|--------------------------------------|----------------|-------------------------------------|---|--|------------------------------------|
| Base | Sections 3 & 4 | =0 | =0 | Head tax | Inefficient (<i>see Table 1</i>) |
| Urban boundary | Sections 3 | =0 | >0 <i>Set to eliminate excess travel in Base</i> | Head tax | Remains inefficient |
| Congestion tolls | Sections 3 | =0 | =0 | First-best tolls <i>Eliminate excess travel in Base</i> | Becomes efficient |
| Congestion tolls & urban boundary | Section 4 | ≥ 0 | ≥ 0 <i>Set efficiently as a pure public good</i> | First-best tolls <i>Eliminate excess travel within boundary</i> | Becomes efficient |

trips terminate. The latter case is a dispersed city in which travel through the center and reverse commuting are allowed and, hence, all 121 (11×11) commuting arrangements (i, j) and all 1331 ($11 \times 11 \times 11$) commuting–shopping arrangements (i, j, v) are utilized in the equilibrium. The comparison of the dispersed and centralized-employment cities illuminates the effects of job decentralization ongoing since World War II in U.S. metropolitan areas.

3.1. The case of centralized production (jobs)

In the base case with the suboptimal land allocation to roads, when all jobs are forced to locate in zone 6, our model is still more general than the monocentric city models commonly used in the literature. Zone 6 (the CBD) accommodates not just firms but also residences and roads. Land uses by type are not fixed in zone 6 and, hence, net densities in a zone vary between firms and workers. Also, the output of the CBD will adjust in response to rent and wage changes and the wage, product price and output of the CBD are endogenous not exogenous as in most monocentric models.

The rent distance and the residential density distance correspondences are both very steep (Table 3a) with CBD rents 162 times those of the edge zones. Congested car speed near the CBD is only 6.25 miles per hour whereas at the fringe zones it is almost 42 miles per hour, near the free flow speed. Since roads are financed by the head tax, there is no marginal pricing in the Base. The marginal cost of travel per mile (MCT in the tables) consists of just the time lost (valued at the wage rate) and falls more than six fold from \$1.67 per mile in the CBD to \$0.25 per mile in the fringe.

3.2. The effects of dispersed and mixed jobs and residences

When we allow jobs to disperse freely among all eleven zones, their equilibrium distribution becomes determined by the zero profit condition and the markets for product and labor. At equilibrium, goods are produced with labor intensive choice of inputs near the central zones where land is expensive, while goods are produced with land intensive choice of inputs in low density “factories” in the suburbs. In each zone, the labor to land ratio used in production is endogenously determined. Figs. 2–5 show salient aspects of the resulting dispersed equilibrium. The effects of job decentralization can be observed by comparing the corresponding rows in Tables 3a and 3b. Rents in the CBD fall from \$36 to \$7.5 per square

Table 3a
Simulations of city with central production

| Units | <u>G Width</u> miles | <u>G Area</u> acres | <u>Radius</u> miles | <u>Output</u> units/wkr | <u>Welfare</u> utils/wkr | <u>CV</u> \$/wkr/yr | <u>Roads</u> acres |
|--|------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|------------------------------|
| <i>Base</i> | 0 | 0 | 14 | 1443.2 | 6.2255 | 0 | 115.1 |
| <i>Tolls</i> | 0 | 0 | 14 | 1421.2 | 6.2262 | 54.7 | 115.7 |
| <i>Boundary</i> | 4.282 | 1701.7 | 9.718 | 1449.6 | 6.1472 | -6128. | 103.1 |
| Zonal results | 6 | 5 or 7 | 4 or 8 | 3 or 9 | 2 or 10 | 1 or 11 | Areawide |
| % Roads (base) | 27.496 | 18.580 | 9.443 | 6.673 | 3.697 | 0.704 | 3.504 |
| Speed miles/h | 6.251 20.50% - 0.23% | 25.925 9.23% 1.34% | 29.701 7.38% 3.87% | 35.107 - 3.93% 5.72% | 36.161 - 3.39% 10.79% | 41.837 - 1.22% - 59.9% | 27.619 3.24% - 1.87% |
| Marginal cost of travel (MCT) \$/mile | 1.669 120% - 0.72% | 0.402 58.82% - 2.27% | 0.351 46.81% - 4.65% | 0.297 55.62% - 6.31% | 0.288 49.25% - 10.6% | 0.249 17.29% 147% | 0.534 83.41% 4.13% |
| Discretionary trips trips/worker/mon. | 9.546 - 0.08% - 0.14% | 9.489 - 0.41% - 0.13% | 9.384 - 0.97% - 0.10% | 9.295 - 1.42% - 0.06% | 9.215 - 1.84% 0.01% | 9.092 - 2.17% 0.64% | 9.251 - 1.53% 0.44% |
| Discretionary travel h/day/worker | 0.147 - 17.1% 0.09% | 0.181 - 15.7% - 0.20% | 0.244 - 14.2% - 0.81% | 0.297 - 12.4% - 1.47% | 0.344 - 10.7% - 2.29% | 0.416 - 9.0% - 9.92% | 0.322 - 11.2% - 9.77% |
| Commuting h/day/worker | 0.320 - 17.0% 0.23% | 0.397 - 15.4% - 0.07% | 0.542 - 13.3% - 0.70% | 0.666 - 11.2% - 1.41% | 0.778 - 9.01% - 2.30% | 0.953 - 7.00% - 10.5% | 0.728 - 9.63% - 10.3% |
| Aver. travel time h/day/worker | 0.467 - 17.0% 0.2% | 0.578 - 15.5% - 0.1% | 0.786 - 13.6% - 0.7% | 0.963 - 11.6% - 1.4% | 1.122 - 9.5% - 2.3% | 1.369 - 7.6% - 10.3% | 1.050 - 10.1% - 10.1% |
| Net res. density people/acre | 407.35 1.37% 1.51% | 16.339 3.44% 21.59% | 9.755 1.71% 21.64% | 7.328 - 0.11% 21.72% | 5.929 - 0.55% 21.83% | 2.564 - 0.87% 370% | 4.751 0.00% 115% |
| Net job density workers/acre | 422.81 4.57% 0.92% | | | | | | 422.813 4.57% 0.92% |
| Rent \$/sq.ft./year | 35.918 3.07% 0.62% | 1.437 5.02% 20.53% | 0.854 3.00% 20.59% | 0.639 0.95% 20.69% | 0.515 0.31% 20.85% | 0.221 - 0.18% 368% | 0.604 1.89% 108% |
| Wage \$/h | 10.431 - 0.39% - 0.96% | | | | | | 10.431 - 0.39% - 0.96% |

(NOTE: Under *Zonal Results*, for each reported variable, the first rows show the *Base* equilibrium values. Rows two and three show the percent changes from *Base* to the *Tolls Policy* and to *Equivalent Boundary Policy*, respectively.)

foot as jobs and residences both decentralize. The ratio of central to fringe land rents decreases from 162 in the centralized city to about 27 in the dispersed. Net residential densities near the center are reduced by 79.5% while net residential densities near the fringes increase by 23.4%. Although both jobs and residences decentralize, jobs remain considerably more centralized than residences. When jobs are forced to locate in the CBD, rents are very high. Once firms are free to move, they seek places with lower rents where land to labor substitution increases the

Table 3b
 Simulations of city with dispersed production and mixed land use

| Units | <u>G Width</u> miles | <u>G Area</u> acres | <u>Radius</u> miles | <u>Output</u> units/wkr | <u>Welfare</u> utils/wkr | <u>CV</u> \$/wkr/yr | <u>Roads</u> acres |
|---|------------------------------|----------------------------|----------------------------|----------------------------|-----------------------------|----------------------------|----------------------------|
| <i>Base</i> | 0 | 0 | 14 | 2385.2 | 7.5433 | 0 | 153.0 |
| <i>Tolls</i> | 0 | 0 | 14 | 2355.6 | 7.5444 | 87.6 | 152.0 |
| <i>Boundary</i> | 3.792 | 1538.1 | 10.208 | 2285.7 | 7.4659 | -6160. | 134.3 |
| Zonal results | 6 | 5 or 7 | 4 or 8 | 3 or 9 | 2 or 10 | 1 or 11 | Areawide |
| % Roads (base) | 49.189 | 19.316 | 11.131 | 8.927 | 5.533 | 1.217 | 4.658 |
| Speed miles/h | 13.058 23.12% 1.57% | 22.795 14.70% 1.36% | 27.059 11.07% 2.17% | 33.241 -3.43% 3.25% | 34.650 -3.34% 6.85% | 41.399 -1.34% -49.4% | 27.664 5.25% -2.55% |
| Marginal cost of travel (MCT) \$/mile | 0.815 0. 85.03% -2.85% | 467 59.71% -2.64% | 0.394 49.36% -3.43% | 0.322 62.41% -4.49% | 0.310 55.89% -7.91% | 0.260 19.71% 93.77% | 0.426 61.25% 0.86% |
| Discretionary trips trips/worker/mon. | 15.467 -1.13% -4.44% | 15.465 -1.02% -4.45% | 15.439 -0.97% -4.48% | 15.386 -1.05% -4.51% | 15.302 -1.26% -4.54% | 15.092 -1.55% -4.03% | 15.290 -1.24% -4.17% |
| Discretionary travel h/day/worker | 0.395 -11.9% -11.0% | 0.397 -12.6% -10.8% | 0.413 -13.5% -10.2% | 0.440 -13.8% -9.5% | 0.480 -13.2% -8.8% | 0.572 -11.5% -13.9% | 0.482 -12.6% -13.9% |
| Commuting h/day/worker | 0.504 -11.3% -4.93% | 0.513 -12.0% -4.82% | 0.545 -12.6% -4.55% | 0.594 -12.3% -4.27% | 0.660 -11.2% -4.21% | 0.805 -9.1% -10.6% | 0.659 -10.8% -9.92% |
| Aver. travel time h/day/worker | 0.899 -11.6% -7.61% | 0.910 -12.2% -7.43% | 0.958 -13.0% -6.99% | 1.034 -12.9% -6.48% | 1.140 -12.0% -6.15% | 1.377 -10.1% -11.9% | 1.142 -11.6% -11.6% |
| Net res. density people/acre | 83.716 8.56% 12.26% | 22.613 2.03% 13.89% | 13.110 1.05% 14.43% | 9.739 -0.75% 14.68% | 7.764 -0.91% 14.88% | 3.164 -1.05% 235.2% | 6.001 -0.57% 101.1% |
| Net job density workers/acre | 100.74 12.10% 9.60% | 24.786 4.89% 11.30% | 13.830 3.74% 11.88% | 10.064 1.63% 12.16% | 7.901 1.27% 12.41% | 3.036 0.97% 252.2% | 7.921 2.14% 71.15% |
| Rent \$/sq.ft./year | 7.472 10.00% 11.03% | 2.018 3.40% 12.63% | 1.169 2.39% 13.16% | 0.867 0.52% 13.40% | 0.690 0.26% 13.59% | 0.280 0.01% 232.1% | 0.599 1.63% 88.89% |
| Wage \$/h | 9.526 -0.41% -0.10% | 10.220 -0.23% -0.16% | 10.525 -0.23% -0.17% | 10.698 -0.17% -0.17% | 10.832 -0.22% -0.18% | 11.363 -0.40% -5.64% | 10.641 -0.29% -1.32% |

(NOTE: Under *Zonal Results*, for each reported variable, the first row shows the *Base* equilibrium values. Rows two and three show the percent changes from *Base* to the *Tolls Policy* and to *Equivalent Boundary Policy*, respectively.)

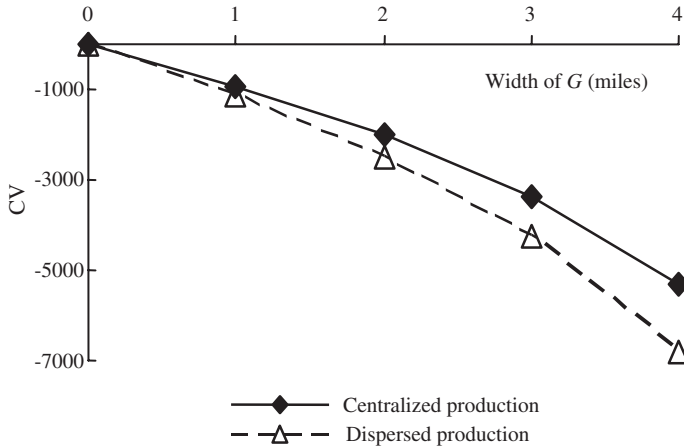
marginal product of labor. This is seen in Table 3b by observing that the wage rate which equals the marginal revenue product of labor at equilibrium, increases by 19.2% as we move from center to fringe. As jobs decentralize, workers do the same because they can get jobs away from the CBD while also renting larger residential lots where land is cheaper. The mutually interdependent decentralization of jobs and residences is reflected both in the average length of commutes as well as the daily average travel time (DATT) spent on discretionary

trips. The daily average commute time per person decreases by 9.5%, while the daily average discretionary time spent on shopping trips increases by 49.7%. The sum of the two is the daily average travel time (DATT), our measure of sprawl. This increases by 8.8%. Thus, in the city with centralized employment, workers spend on average a total of 63 min on the road per day (about 44 min commuting and about 19 min shopping) whereas in the dispersed city they spend an average of almost 69 min per day (about 40 commuting and almost 29 shopping). The commuting time decreases on average because job dispersion brings jobs closer to residences.¹⁴ The discretionary time increases because job dispersion increases the accessibility of shopping opportunities to residences as well as the variety of shopping destinations, causing the number of discretionary trips per month per person to increase by 65% from 9.25 to 15.29. In the monocentric case, no traffic crosses the central zone, since there are no shopping opportunities or jobs on the other side of the city. In the dispersed city, on the other hand, some traffic must cross the center because workers shop everywhere with some probability and can hold jobs at the other side of the CBD. Because there is more traffic in the dispersed case, more land is allocated to roads in every zone and a lot more in the central zone (49.2% rather than 27.5%).

3.3. *The effects of imposing congestion tolls*

Imposing congestion tolls on the base situations described above raises the marginal cost of per mile travel (MCT in the tables) and has several effects. First, as we have already discussed, the replacement of the head tax with congestion tolls alters the road allocation pattern from the misallocated one assumed for the bases case to the efficient. All that is needed for this to happen is to add to road capacity in the zones where toll revenue exceeds the cost of renting the land in roads and to reduce road capacity where the opposite occurs. This will cause the initially inefficient road allocation to increase by 11.11% in each of the five central city zones and to fall by 9.09% in each of the six suburban zones. Beyond this, the effects of the tolls in the centralized versus the dispersed city will be quite different. To see this, recall that in the centralized city workers can react to congestion tolls by moving closer to the CBD and thus reducing travel distance to economize on travel cost, or by making fewer discretionary trips to the CBD. Producers cannot relocate out of the CBD, so their response to tolls is confined to a market adjustment in wage. More workers choose residence in or near the CBD to avoid tolls. This reduces the CBD land available to producers and therefore the marginal product of labor. Hence, the competitive CBD wage offered falls, although the effect is slight. In the dispersed city, there are more margins of adjustment. In response to tolls, workers can still move closer to jobs, and make fewer and shorter discretionary shopping trips. Meanwhile, though, producers can move closer to workers, which improves access to jobs and increases labor supply allowing producers to pay lower wages; and which increases worker access to shopping and the demand for product by reducing the effective price paid for product. But this is not all, because as producers decentralize they substitute land for labor and the marginal product of labor increases making fringe wages about 19% higher than central wages. All of these effects are simultaneously treated in our model. Because of these added margins of adjustment, the imposition of congestion tolls generates higher benefits in a dispersed rather than a centralized city. Imposing congestion tolls in the centralized city amounts to an efficiency gain of \$54.7 per

¹⁴ These daily commute times generated by our model and their changes under decentralization are consistent with those empirically observed (e.g. Gordon et al., 1989). Discretionary travel times and how they changed as urban areas decentralized are documented by Nelson and Niles (2000).



NOTE: CVs (\$/worker/year) relative to the base cases corresponding to Table 3a (centralized production) and Table 3b (dispersed production).

Fig. 6. Welfare losses of growth boundaries when the greenbelt is not valued ($\eta=0$).

worker per year (Table 3a). In the dispersed city, under the same parameters of Table 1, the benefit of tolls as measured by the compensating variation rises by about 60% from \$54.7 to \$87.6 (Table 3b). From the tables we can see that — because tolls induce more roads to be built near the center and roads to be retired near the edge — traffic speeds increase in the central zones and decrease in the peripheral zones. Toll causes the marginal cost of travel (MCT) near the center, inclusive of tolls, to increase by about 120% in the centralized case and by about 85% in the decentralized case. This increase tapers off with distance falling to just under 20% at the fringes of the city. In the centralized city, tolls reduce the daily average travel time (DATT) by about 10.1% and in the dispersed city by about 11.6%. Both commuting and discretionary travel time are reduced as expected. The travel reductions due to tolls eliminate a wasteful 7.86 min (6.36 min) of travel per worker per day in the dispersed (centralized) city. The time saved under the toll regime, valued at the average wage rate, is only \$1.39 per day in the dispersed case or \$348.5 per year per worker which is nearly four times the workers' consumer surplus gain of about \$88.

3.4. The effects of the urban boundary

Unlike tolls, boundaries do not provide a mechanism for adjusting the allocation of land to roads.¹⁵ Therefore, the Base land allocation to roads will not be adjusted (except that roads that were beyond the boundary in the Base are no longer used and are converted to “greenspace”) and roads inside the boundary will continue to be funded by the head tax.

¹⁵ Planners could optimize the allocation of roads within the boundary, but this is complex requiring knowledge of utility and production functions. In contrast, the replacement of the head tax with congestion tolls as a road funding mechanism would cause the allocation of land to roads to rectify in the future. In Section 4 we will show the effects of congestion tolls when an urban boundary is already present.

In the simulations reported in this paper, a boundary of any stringency is never second best.¹⁶ This is shown in Fig. 6. Using our accurate algorithm we can calculate the welfare loss resulting from an even 1 meter wide urban growth boundary. As Fig. 6 shows, a boundary of a given stringency is more harmful in an urban area with dispersed production than in one with centralized production. The reason for this is that when production occurs in the edge zones, then the much higher suburban rents caused by the boundary adversely impact suburban product prices by raising the cost of production.

To understand how the urban boundary policy differs from the optimal tolls policy, recall that the boundary does nothing to correct the under-pricing of travel in the base case. Instead, the boundary curbs the excess sprawl caused by the distortion in the road allocation and the under-pricing of traffic, by creating yet another distortion in the land market. Since the growth boundary restricts the supply of land available for urban development, it forces the same population to squeeze into a smaller circular city. This has several effects.

The most evident of these effects is that the boundary causes rents within the city to increase as land becomes scarce. This causes producers to substitute labor for land. The marginal product of labor decreases somewhat and so do wages. The effects of the boundary on trip making, trip lengths, speeds and the daily average travel time (DAT) are more complex. First, because the boundary increases rents and reduces wages, there are income and substitution effects on trip-making. The negative income effect of higher rents occurs because more income must be spent on land and thus less is available for discretionary trips to purchase goods. The negative substitution effect arises because the rent increase is partially capitalized into the mill (retail) prices of goods. Meanwhile, because wages are reduced somewhat and the value of time decreases, the opportunity cost of a discretionary trip decreases and this results in another positive substitution effect on trip making. Second, by forcing firms and businesses closer together while removing road capacity from the edge zones, the boundary induces effects that have opposing consequences on traffic flow and travel speed. On the one hand, average distances between firms and customers are shortened and each potential discretionary trip involving an edge zone becomes shorter. This would reduce the opportunity cost of such trips and the resulting substitution effect would tend to induce more discretionary trip-making. On the other hand, keeping the number of trips constant, the boundary increases congestion by removing road capacity in the edge zones. This would affect discretionary trips that occur locally in the edge zones and all other trips originating or terminating in the edge zones, reducing average speeds. In our simulations, the combined impact of all these effects is such that when the boundary is imposed, the number of discretionary trips falls by 4.17%, while the time spent on discretionary travel falls by 13.9% (Table 3b). Note that the travel speed in the edge zones falls by 49.4% indicating the substantial increase in congestion due to the reduction of road capacity. Since consumers are attached to their communities because of idiosyncratic tastes, many stay in the edge zones despite the higher rents and this increases congestion since the boundary has reduced road capacity. But in the non-peripheral zones speeds increase slightly due to the reduction in discretionary trips. In the central employment case (Table 3a), imposition of the boundary slightly increases discretionary trips on average by 0.44% because residents are packed closer to the center, while reducing discretionary travel time by 9.8%.

Table 4 juxtaposes the efficiency gains induced by the congestion tolls and the welfare losses caused by the urban boundary for the case of dispersed production/employment. This is done for

¹⁶ As pointed out in the Introduction, the reasons why urban boundaries are not always second-best is the subject of our other paper (Anas and Rhee, in press).

Table 4
Welfare gains and losses with various initial road misallocations

| Policy change and welfare gain/loss | | Initial road allocation in the Base city | | | | |
|-------------------------------------|--|--|-------|-------|--------|--------|
| | | 0% | ±8% | ±10% | ±12% | ±14% |
| From Base to Congestion Tolls | <i>DATT per worker reduction (min)</i> | 4.9 | 7.2 | 7.9 | 8.7 | 9.6 |
| | <i>Compensating variation (\$/worker/year)</i> | +38.3 | +71.5 | +88.5 | +107.0 | +130.1 |
| | <i>Change in consumer surplus (\$/worker/year)</i> | +38.7 | +72.1 | +88.2 | +107.7 | +130.8 |
| | <i>Greenbelt width (miles)</i> | 2.431 | 3.493 | 3.792 | 4.094 | 4.389 |
| From Base to Boundary | <i>Compensating variation (\$/worker/year)</i> | -3207 | -5359 | -6160 | -7129 | -8338 |
| | <i>Change in consumer surplus (\$/worker/year)</i> | -3154 | -5208 | -5960 | -6861 | -7972 |

(NOTE: $\pm x$ % denotes a road system of the Base case with x % above the first-best efficient road allocation in the suburbs and x % below in the central city. The case of $\pm 10\%$ road system is the one used in the simulations reported in Tables 3a and 3b; The 0% column is the case of the first-best efficient road allocation.

the optimal road allocation that would occur under tolls in the long run and for four different initial distortions in road allocation ranging from 8% to 14% to check the sensitivity to the initial misallocation. The case of 10% misallocation corresponds to the simulations reported in Table 3b. The larger the deviation from the efficient road allocation, obviously the more beneficial are congestion tolls. Nevertheless, the efficiency gains induced by the tolls vary from 0.18% to only 0.33% of income. Although these efficiency gains are small as a percentage of income, the tolls nevertheless correct an excess sprawl that ranges from 7.2 to 9.6 min per worker per day, and so they are highly effective as an anti-sprawl policy. In contrast, the welfare losses induced by the equivalent urban boundary that eliminates the excess sprawl are quite huge, ranging from 7.5% (when roads are optimal) to about 21% of money income. The deadweight loss increases with the initial road misallocation as does the width of the greenbelt and the amount of land expropriated by the boundary. Most of the deadweight loss caused by the boundary is due to the rent increases within the boundary. The part of the compensating variation that corresponds to the distortion in rents, product prices and wages can be quantified by changing these vectors one at a time and, at each step, calculating a compensating variation relative to the Base situation. Changes in these vectors have a direct effect on full income, and on the Marshallian demands of a worker for lot size, product quantities and leisure, keeping traffic flows and the rent dividend at their Base values. Indirectly, rents and wages also affect the rent dividend and the traffic flows. In the city with dispersed jobs, 87.6% of the compensating variation loss is due to the direct effects and the remaining 12.4% is due to the indirect effects of the price distortions. Rent distortions account for 77% of the direct effects, product price distortions for 18% and wage distortions for the remaining 5%.¹⁷

Finally, we should comment on the physical effects of the equivalent growth boundary policy. From Table 3b, the equivalent greenbelt in the case of the city with job dispersion is 3.792 miles

¹⁷ The path integral is not unique because the indirect utility is not linear in prices and income. Therefore, the reported percentages depend on the order in which rents, prices and wages are changed from their base to their policy values. To account for this potential sensitivity, we computed each of the six permutations and confirmed that the results are within 0.1% of each other. The reported percentages are arithmetic averages of these permutations.

wide (a 27.9% reduction in the radius of the city). It confiscates 1538 acres of fringe land in each unit section of the circular city, amounting to 46.9% of the total land initially available within the 14 mile radius. Note that this percentage is large because in a circular city, the land available rises with the square of the distance from the center and — in any circular city — a lot of the total urban land is near the edge. In the base case, the confiscated land accommodates about 480,589 residents or 26.7% of the population of 1.8 million. Thus the physical adjustment needed to eliminate the excess sprawl via a growth boundary is rather enormous and unpalatable. As already explained, this dramatic adjustment occurs because the boundary does not change pricing but instead reduces travel by rationing the trip lengths of the edge-zone residents. A large number of these suburbanites must be moved inward to get a reduction in daily average travel time (DATT) sufficient to curb all excess sprawl. This contrasts with the subtle effect of the congestion tolls that replace the head tax and work through marginal pricing to achieve an identical reduction in sprawl without interfering with the markets and without limiting where people can settle but by significantly changing the travel cost of each and every resident. Since tolls would replace other taxes (head tax in our model) they should be much more palatable politically once their advantages over urban boundaries are understood.

3.5. The effect of idiosyncratic attachment

As Tables 3a and 3b showed, the urban boundary causes steep rent increases in the edge communities (zones 1 and 11). These steep increases reflect the importance of idiosyncratic attachment in residential location choice. Recall that the consumers who sort into the edge communities are those who have relatively stronger idiosyncratic preferences for living there. To drive them to other communities, rents must rise steeply enough to overcome this attachment. It stands to reason then, that if we weaken the attachment, then the boundary will cause milder rent increases in the edge communities. In our model, the variance of idiosyncratic attachment is controlled by λ , the taste heterogeneity parameter. As λ is increased, the idiosyncratic attachment to community diminishes and prices loom larger in location decisions. Since this is such an important ingredient of our model, we have performed some sensitivity analyses by changing the value of λ . Our objective is to observe two things: (a) how much does the steep rent increase in the edge communities caused by the boundary moderate?, and (b) how much does the deadweight loss of the boundary as a percentage of money income diminish? Table 5 shows the results. Note that in the case of the urban area

Table 5
Effects of idiosyncratic attachment on rent changes and deadweight loss

| λ | <u>Zone 6</u> | <u>5 or 7</u> | <u>4 or 8</u> | <u>3 or 9</u> | <u>2 or 10</u> | <u>1 or 11</u> | <u>Areawide</u> | UGB's CV (% of income) | Equivalent UGB G width (in miles) |
|--|---------------|---------------|---------------|---------------|----------------|----------------|-----------------|---------------------------|--------------------------------------|
| <i>Percent changes in rents per sq. ft. caused by the Equivalent UGB</i> | | | | | | | | | |
| <i>Centralized production</i> | | | | | | | | | |
| 4 | 0.62 | 20.53 | 20.59 | 20.69 | 20.85 | 368.01 | 108.37 | -15.32 | 4.282 |
| 10 | -0.26 | 23.42 | 23.48 | 23.59 | 23.76 | 92.77 | 63.50 | -10.55 | 3.029 |
| 20 | -0.22 | 10.30 | 7.88 | 11.10 | 6.19 | 18.02 | 16.06 | -9.31 | 2.480 |
| <i>Dispersed production</i> | | | | | | | | | |
| 4 | 11.03 | 12.63 | 13.16 | 13.40 | 13.59 | 232.06 | 88.89 | -15.40 | 3.792 |
| 10 | 11.32 | 18.92 | 20.79 | 21.53 | 22.03 | 103.98 | 67.00 | -14.09 | 3.137 |
| 15 | 9.67 | 23.48 | 26.48 | 27.52 | 28.20 | 90.27 | 65.30 | -14.96 | 3.078 |
| 20 | 7.84 | 27.19 | 31.19 | 32.44 | 33.24 | 85.25 | 65.56 | -15.81 | 3.085 |

with centralized production, the rent increase in the edge communities moderates rapidly from 368% to only 18% as λ is increased from 4.0 to 20.0. In the case of the urban area with dispersed production, the rent increase moderates from 232% when $\lambda=4.0$ to 85% when $\lambda=20$. The moderation is not as rapid when production is dispersed as it is when production is centralized, because producers only spend 14% of their costs on land and can easily substitute labor for land without leaving the zone when rents increase. In both the centralized and dispersed production cases, when λ is increased, rent increases in the inner communities are higher as some residents of the edge communities are driven out and relocate inward. With regard to the second question, as Table 5 shows in the next to last column, the deadweight loss as a percentage of money income is smaller when there is less idiosyncratic attachment but it does not fall dramatically, because the large rent increases in the edge zones in the case of high attachment are simply redistributed to other zones in the case of low attachment. This underscores the robustness of our finding that urban growth boundaries are extremely ineffective tools for sprawl reduction as compared to congestion tolls.

4. Efficient urban boundaries

So far we assumed $\eta=0$, which means that greenbelts created by imposing urban boundaries have no direct value for urban workers. Indeed, it is highly plausible that workers value open space within the city either in the form of extra space between adjacent homes or as urban parks large and small. It is much more difficult to argue that open space placed at the edges of cities is as desirable as open space that is within the urbanized area since cities are normally surrounded by open spaces anyway. It is more plausible that urban workers expect that certain benefits would come from stemming urban expansion, thus making the city more compact, rather than directly from the additional open space itself at the fringe. To the extent that this may be true, it would be reasonable to assume $\eta>0$ in the context of our model. In this case the imposition of a planned urban boundary creates a greenbelt that has value as a pure public good or, equivalently, it decreases the urbanized area and this is perceived as valuable. It is not at all clear how big η should be since there is no quantitative evidence concerning this sort of preference. It should be clear, though, that one more square foot of greenbelt land placed at the edge of the city is a lot less valuable for a worker than one square foot of additional land placed in that worker’s private lot. Since the marginal rate of substitution between private land and expenditures on goods is measured by the equilibrium rent on land, a worker’s marginal rate of substitution between G and other expenditures should be much lower than the rent on land. This provides a guideline for the calibration of η , albeit only a very rough one.

The land use planner would determine G^* by optimizing the welfare function (11) with respect to G taking as given the prices, wages and rents determined by the markets, and the road allocations determined by the road planner. If tolls are not levied, the road allocation would remain inefficient. If tolls are levied, the road allocation within the boundary would revert to efficient. The first order condition is:

$$\frac{dW}{dG} = \sum_{\forall(i,j)} \frac{\partial W}{\partial V_{ij}} \frac{\partial V_{ij}}{\partial G} + \sum_{\forall(i,j)} \frac{\partial W}{\partial V_{ij}} \frac{\partial V_{ij}}{\partial \Omega_{ij}} \frac{\partial \Omega_{ij}}{\partial G} + \sum_{\forall(i,j)} \frac{\partial W}{\partial V_{ij}} \left(\sum_{\forall v} \frac{\partial V_{ij}}{\partial p_{ijv}} \frac{\partial p_{ijv}}{\partial G} \right) = 0. \tag{17}$$

In this equation, $\frac{\partial W}{\partial V_{ij}} = \Psi_{ij}$ is the fact that the demand function, Ψ_{ij} , is the marginal benefit function, $\frac{\partial V_{ij}}{\partial G} = \frac{\eta}{G_0 + G}$ is the marginal utility of the open space and $\frac{\partial V_{ij}}{\partial \Omega_{ij}} = \frac{1}{\Omega_{ij}}$ is the marginal utility

of income of a worker residing at i and working at j . Also, $\frac{\partial \Omega_{ij}}{\partial G} = -2d \left[\frac{\partial t_{ij}}{\partial G} + w_j \frac{\partial g_{ij}}{\partial G} \right] - \frac{r_1}{N}$, $\frac{\partial V_{ij}}{\partial p_{ijv}} = -\frac{z_{ijv}}{\Omega_{ij}}$ and $\frac{\partial p_{ijv}}{\partial G} = 2f \left(\frac{\partial t_{iv}}{\partial G} + w_j \frac{\partial g_{iv}}{\partial G} \right)$. Using these to evaluate $\frac{dW}{dG}$ and rearranging the terms we obtain a version of the Samuelson condition for the optimal finance of a pure public good:

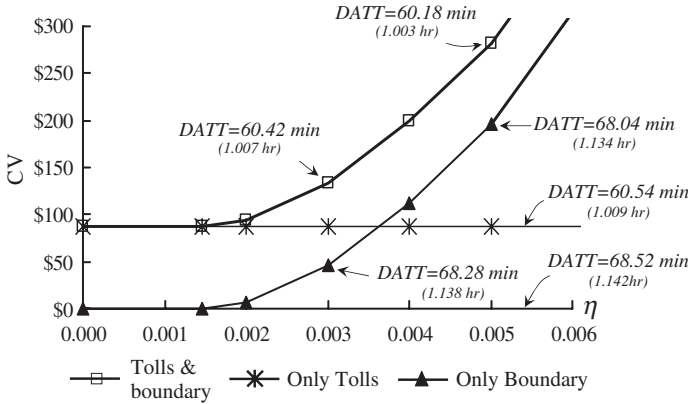
$$N \frac{\frac{\eta}{G_0+G} - 2d \sum_{\forall (i,j) \in S_1} \frac{\Psi_{ij}}{\Omega_{ij}} \left(\frac{\partial t_{ij}}{\partial G} + w_j \frac{\partial g_{ij}}{\partial G} \right) - 2f \sum_{\forall (i,j,v) \in S_2} \frac{\Psi_{ij}}{\Omega_{ij}} \left[z_{ijv} \left(\frac{\partial t_{iv}}{\partial G} + w_j \frac{\partial g_{iv}}{\partial G} \right) \right]}{\sum_{\forall (i,j)} \frac{\Psi_{ij}}{\Omega_{ij}}} = r_1. \tag{18}$$

The right side is the marginal opportunity cost of one more acre removed from urban development by the planned urban boundary, given by the endogenous rent in the edge zone.¹⁸ The left side is the sum over the workers of the ex ante expected marginal benefit of that preserved acre, the SMRS (social marginal rate of substitution). This includes two effects. The first term in the numerator, $\frac{\eta}{G_0+G}$, is the marginal utility of the acre as extra open space, while the two brackets are the marginal utility from the changes in tolls and travel times that occur due to shortened trips in the two edge zones because the boundary compresses edge zone development inward into a smaller area. There are, in fact, two conflicting sources of such net travel savings. On the one hand, the developed part of the zone’s length is reduced making trip lengths in the edge zones shorter but, on the other hand, land available for edge roads is reduced which tends to increase congestion. These edge zone travel cost savings are of two types. Those of *commuters who either reside or work in the edge zones* described by set S_1 of commuting arrangements (i, j) in which i or j (or both) is an edge zone and those of *workers who reside in any zone but shop in an edge zone or reside in an edge zone but shop in any other zone*, described by the set S_2 of shopping arrangements (i, j, v) where i or v (or both) are edge zones. For such situations, $\frac{\partial t_{ij}}{\partial G} = \frac{\partial t_{ij}}{\partial \Delta_1} \frac{\partial \Delta_1}{\partial G}$ and $\frac{\partial g_{ij}}{\partial G} = \frac{\partial g_{ij}}{\partial \Delta_1} \frac{\partial \Delta_1}{\partial G}$ are nonzero but for all other situations they are zero since the urban boundary has no effect on zone length in internal zones and, hence, no direct effect on travel cost. When the effect is nonzero (i.e. $(i, j) \in S_1$ or $(i, j) \in S_2$), we have:

$$\frac{\partial t_{kl}}{\partial \Delta_1} = \begin{cases} \frac{1}{2} \frac{t_1}{c+1}; & \text{if either } k \text{ or } l \text{ is an edge zone} \\ \frac{t_1}{c+1}; & \text{if both } k \text{ and } l \text{ are edge zones} \end{cases}$$

$$\frac{\partial g_{kl}}{\partial \Delta_1} = \begin{cases} \frac{1}{2} \left(g_1 - \frac{c}{c+1} \varphi_1 \Delta_1^{-\frac{c}{c+1}} \right); & \text{if either } k \text{ or } l \text{ is an edge zone} \\ \left(g_1 - \frac{c}{c+1} \varphi_1 \Delta_1^{-\frac{c}{c+1}} \right); & \text{if both } k \text{ and } l \text{ are edge zones} \end{cases}$$

¹⁸ Despite Muth (1961), who showed that the agricultural rent at the urban edge is endogenous, it is common in urban models to assume that r_1 is an exogenous rent on farmland (see also footnote 11). But this ignores the fact that the actual rent at a city’s edge includes a potentially much higher option value because of anticipated future urban development. Although our model is static, it would have been misleading to calibrate it so that r_1 is the rent on farmland. In our base simulation, $r_1 = \$12,197/\text{acre}$ and should be interpreted as reflecting an option rent as well as the rent on farmland. At such a rent, a worker residing at the edge zones occupies a lot of about 3/4 of an acre.



NOTE: DATT: Daily average (per person) travel time; CV: \$/worker/year; The x-axis corresponds to the Base with G=0 and an inefficient road allocation financed with a head tax.

Fig. 7. Benefits of anti-sprawl policies when the greenbelt is valued.

where $\varphi_1 = ab\zeta^{-c}(r_1/\bar{w}_1)^{\frac{c}{c-1}}(abcd\zeta^{-c})^{-\frac{c}{c-1}}$. The zone length is related to the greenbelt's area by $\frac{\partial \Delta_1}{\partial G} = -\frac{1}{\sqrt{\varphi^2 r^2 - 2\varphi G}} < 0$. Finally, the denominator, $\sum_{v(i,j)} \frac{V_{ij}}{\Omega_{ij}}$, in (18) is the expected marginal utility of income and division by it converts the left side to a dollar valued marginal benefit of G .

Fig. 7 and Table 6 report the results of solutions assuming various values of η in the case of the dispersed city. Three policies are juxtaposed in Fig. 7. Each policy is applied to the Base case of Table 2 in which roads are 10% under allocated in the city and 10% over allocated in the suburbs and funded by the head tax. First is the policy of levying congestion tolls and optimizing the road allocation, but setting the greenbelt to zero and thus ignoring that urban workers value a more compact city or extra open space in the city's fringe (labeled "Only Tolls" in Fig. 7). In this case, the benefit as measured by the CV is the constant \$87.6 per worker per year that was encountered in Section 3. Since the part of utility that depends on η , η in G_0 , is unchanged by the congestion toll policy, η does not affect the CV. The second policy determines the optimal urban boundary by optimizing over the possible values of the boundary to find G^* for the given η (labeled "Only Boundary"). We do this by trial and error searching for the value of G that maximizes \mathcal{W} . For each

Table 6
Disaggregation of the marginal benefits of efficient boundaries

| η | G^* (acres) | MRS ¹ (A) | Commute cost savings ² | | | Shopping travel cost savings ² | | | SMRS ³ | Marginal cost ⁴ |
|--------|------------------|-------------------------|-----------------------------------|--------|---------|---|--------|---------|-------------------|----------------------------|
| | | | Toll | Time | Sum (B) | Toll | Time | Sum (C) | | |
| 0 | 0.0 | 0.0000 | 0.0042 | 0.0639 | 0.0680 | 0.0032 | 0.0482 | 0.0514 | 597 | 12,192 |
| 0.002 | 17.5 | 2.3363 | 0.0042 | 0.0640 | 0.0682 | 0.0032 | 0.0482 | 0.0515 | 12,280 | 12,278 |
| 0.003 | 49.6 | 2.3683 | 0.0043 | 0.0641 | 0.0684 | 0.0033 | 0.0483 | 0.0516 | 12,441 | 12,439 |
| 0.004 | 80.8 | 2.4010 | 0.0044 | 0.0643 | 0.0687 | 0.0034 | 0.0484 | 0.0518 | 12,607 | 12,600 |
| 0.005 | 111.3 | 2.4323 | 0.0045 | 0.0644 | 0.0689 | 0.0034 | 0.0485 | 0.0519 | 12,765 | 12,762 |
| 0.010 | 251.9 | 2.5938 | 0.0049 | 0.0652 | 0.0701 | 0.0038 | 0.0489 | 0.0526 | 13,583 | 13,577 |
| 0.161 | 1538 | 7.9426 | 0.0249 | 0.0759 | 0.1007 | 0.0175 | 0.0530 | 0.0705 | 40,589 | 40,599 |

(NOTES: ¹MRS between G and income in \$/worker/year/acre of greenbelt G . ²\$/worker/year/acre of G . ³SMRS: Social marginal rate of substitution (Total marginal benefit= $N[A+B+C]$). ⁴Edge zone rent in \$/acre/year.)

trial value of G , we solve the general equilibrium problem in which the head tax is used to finance roads and the road allocation is the inefficient one. The third policy is the first-best optimal policy of levying congestion tolls while also optimizing the planned boundary (labeled “Tolls and boundary”). In this case, for each trial value of G , we again solve for the general equilibrium while using congestion tolls to finance roads. Once a particular $G^*(\eta)$ is found, we verify that the Samuelson condition (18) is satisfied by summing the three sources of marginal benefit to see that they almost match the rent in the edge zone. These results are shown in Table 6.

Several results are gleaned from Fig. 7. First, when the value of η is very close to zero, it is not optimal to have an urban boundary because the marginal benefits generated by any boundary policy are lower than the marginal costs of such a policy (the left side of the Samuelson condition (18) is always smaller than the edge rent).¹⁹ Second, as η is increased, the benefits of the boundary policy, as measured by the CV, increase and — at some value of η — match the benefits of the congestion tolls policy. For higher η , the urban boundary policy dominates. Third, when tolls and the boundary are used together, the benefits are only slightly subadditive. A fourth and important result relates to Section 3. The amount of sprawl, as measured by the daily average travel time (DAT), is rather insensitive to the value of η and so it is insensitive to the value workers attach to city-compactness or fringe open space. Congestion tolls are far more effective in reducing this measure of sprawl than are urban boundaries. For example, when $\eta=0.005$, we see from Table 6 that each worker values the peripheral greenbelt at \$2.43/acre. In this case, the imposition of a boundary that creates a greenbelt of 111.3 acres per each unit section of the city shown in Fig. 2, reduces DAT only very slightly from 68.5 to only 68.04 min per worker per day, an improvement of only -0.68% . The imposition of tolls achieves a further reduction to 60.18 min per day, or an additional improvement of -11.8% . That is, tolls are responsible for 94% of the improvement achieved by both policies together.

Finally, in the last row of Table 6 we show the value of η that would render efficient the enormous greenbelt of Section 3 that eliminated the excess sprawl caused by the congestion externality. Each worker would have to be willing to pay almost exactly \$8 per year per marginal acre of greenbelt. As we saw before, DAT is reduced by 8 min to 60.5 min per worker per day. But because the boundary increases congestion, tolls are still needed and would reduce DAT by another 5 min per worker per day. Thus even in this case of an implausibly extreme anti-sprawl preference, tolls still account for 38.5% of the total improvement possible.

5. Conclusions and extensions

5.1. Conclusions about tolls and UGBs

Although congestion tolls may generate only slight efficiency benefits (0.22% of income in our case), they are a very effective tool for reducing urban sprawl under the long run conditions that we simulated. Equivalent urban boundaries, on the other hand, are not second bests to congestion tolls in general (Anas and Rhee, *in press*) and not second-bests in the simulations reported here (Fig. 6). Boundaries require massive adjustments in the land market and generate

¹⁹ As explained in the Introduction (see also footnote 8), in conventional monocentric models of the urban economy with exogenous agricultural rents, boundaries are a second-best to congestion tolls. Also, in those models all travel goes to the center and thus removing road capacity at the edges does not increase congestion. Also, in the conventional models, alternative locations within an urban area are viewed as perfect substitutes ignoring the attachment of suburbanites to their communities.

a huge deadweight loss compared to the congestion tolls. That happens because the boundary policy ignores the pricing margin and, instead, attempts to reduce daily average travel time (DAT) by significantly rationing land at the suburbs. Even if workers valued peripheral open space, it may be impossible to know how much they are willing to pay for it in the margin. This means that planners would likely make errors in deciding the urban boundary causing quite large deadweight losses. Congestion tolls are also known to be difficult to implement although there are hopeful signs that they might be broadly feasible in the future due to advances in electronic tolling devices. But even if tolls may not be possible, a number of close second bests such as gasoline taxes and targeted parking taxes do exist. Furthermore, road capacities under the discipline of congestion tolling are self-correcting in the long run, whereas urban boundaries do nothing to correct the road misallocation and road planners cannot optimally allocate roads in the absence of tolls. In the short run, when durable buildings make land use density adjustments slow and costly, it would seem that the boundary policy would be even less effective in reducing travel times, whereas tolls can still go a long way by inducing a significant resorting of location decisions if not a substantial modification of building densities. In contrast, with durable buildings, a boundary would still cause a swift rise in rents.

5.2. Extensions of the model

The model we have used in this article has multiple features suitable to the analysis of modern cities. It treats the location of production and of residences as mixed with each other and interdependent on each other not only via commuting but also via discretionary travel (retail trips) and generates a pecuniary agglomeration without assuming increasing returns. Although we have presented results in which the urban economy is completely closed, interactions with the outside world are important for cities. In unreported simulations, we have dealt with imports and exports and also with urban areas open to population migration. We have emphasized the mixing of production and residences. An extension of the model is to endogenize agriculture as an industry and to model the mixing of urban and non-urban use in the exurban fringes of metropolitan areas. Such an extension would result in a more realistic treatment of sprawl as the diffusion of urban land uses into the exurban periphery.

An important weakness of our approach is that we relied entirely on the multinomial logit model as the probabilistic choice structure for the joint choice of residence and employment location. While the logit model has considerable computational benefits, it is important to relax this strict assumption in the future and to allow separate idiosyncratic preferences over job locations and residence locations. The present model assumes that consumers perceive residence–job pairs as distinct uncorrelated alternatives over which they have idiosyncratic preferences. Variants of the nested logit or probit models or the hybrid logit model can be used to test spatial correlations among these idiosyncratic tastes reflecting the underlying zonal characteristics. It would be interesting to see how particular choice models and correlation structures result in different spatial equilibrium rent and density profiles. We also assumed that idiosyncratic tastes reflect pure preferences. In this article, these idiosyncracies stemmed from raw characteristics of the land in the communities (zones), best interpreted as topological or landscape differences. It may be argued that idiosyncratic preferences for individual communities are also related to characteristics of those communities that are endogenous in the spatial equilibrium being modeled. Then, a growth boundary or congestion tolls would change not only travel times, rents, wages, densities and other endogenous variables, but could also change the distribution of idiosyncratic tastes. Development of such choice structures would improve the realism of our model.

Appendix. Notational glossary

The key scalar parameters used in the simulations are listed in Table 1. This glossary defines the vectors and matrices used in the model as well as some other key scalars.

A. Key aggregates

| | |
|-------|---|
| N | total number of workers, |
| I | city-wide average money income, |
| G | amount of land area outside the urban growth boundary, reserved as a greenbelt and split equally between the edge zones 1 and 11, |
| G_0 | amount of greenspace available elsewhere, |
| D | the rent dividend (equal share of aggregate rents) received by each worker, |
| W | expected value of the maximized utilities of workers, |
| CV | compensating variation, |
| EV | equivalent variation, |
| CS | consumer surplus, |
| h | head tax per worker when congestion tolls are not levied to finance roads, |
| DATT | daily average travel time per worker, |
| MCT | per mile marginal cost of travel. |

B. Zones

| | |
|--------------|---|
| \mathbf{A} | (A_1, \dots, A_{11}) : total land available in zone i , |
| Δ | $(\Delta_1, \dots, \Delta_{11})$: length of a zone. |

C. Prices

| | |
|-------------------|---|
| \mathbf{p} | (p_1, \dots, p_{11}) : product (mill) prices, |
| \mathbf{p}_{ij} | $(p_{ij1}, \dots, p_{ij11})$: effective (or delivered) price of product to a worker residing in zone i and employed in zone j , |
| \mathbf{w} | (w_1, \dots, w_{11}) : hourly wage rates, (\bar{w}_{ij}) : average wage rate of all travelers traveling from zone i to zone j and w_{ij}^- : average wage rate of all travelers traversing zone i , |
| \mathbf{r} | (r_1, \dots, r_{11}) : land rents per square foot. |

D. Production

| | |
|--------------|--|
| \mathbf{X} | (X_1, \dots, X_{11}) : outputs, |
| \mathbf{M} | (M_1, \dots, M_{11}) : aggregate labor hours hired by firms, |
| \mathbf{Q} | (Q_1, \dots, Q_{11}) : aggregate land rented by firms. |

E. Transportation

| | |
|----------------|---|
| \mathbf{R} | (R_1, \dots, R_{11}) : land allocated to roads in each zone, |
| \mathbf{K} | (K_1, \dots, K_{11}) : road capacity in each zone, |
| \mathbf{F} | $[F_{ij}]$: flow of total trips originating in residences in zone i and terminating in zone j for work (commuting) plus non-work (shopping), |
| \mathbf{F}^w | $[F_{ij}^w]$: flow of total work trips originating in residences in zone i and terminating in workplaces in zone j (commuting), |

| | |
|-------|--|
| F^s | $[F_{ij}^s]$: flow of total shopping trips originating in residences in zone i and terminating in production places in zone j . |
| F | $[F_1, \dots, F_{11}]$: flow of total trips traversing zone i , |
| g | $[g_{ij}]$: congested travel time from zone i to zone j , |
| g | $[g_1, \dots, g_{11}]$: congested travel time per unit length of zone i , |
| t | $[t_{ij}]$: congestion tolls paid to travel from zone i to zone j , |
| t | $[t_1, \dots, t_{11}]$: congestion toll per unit length of zone i . |

F. Consumer–workers

| | |
|----------|---|
| l | $[l_{ij}]$: leisure hours of a worker residing in zone i and working in zone j , |
| q | $[q_{ij}]$: lot size rented in residence zone i by worker employed in zone j , |
| z | $[z_{ij}]$: product purchased in zone v by a worker residing in zone i , and employed in zone j , |
| u | $[u_{ij}]$: idiosyncratic utility of a worker residing in zone i and employed in zone j , |
| U | $[U_{ij}]$: direct utility of a worker residing in zone i and employed in zone j , |
| V | $[V_{ij}]$: systematic (non-idiosyncratic) indirect utility of a worker residing in zone i and employed in zone j , |
| Ω | $[\Omega_{ij}]$: full endowment income of a worker residing in zone i and employed in zone j net of the full opportunity cost of commuting, plus rent dividends, |
| I | $[I_{ij}]$: money income of a consumer who resides in zone i and is employed in zone j , |
| T | $[T_{ij}]$: total (commuting plus discretionary) travel time of a worker who resides in zone i and is employed in zone j , |
| Ψ | $[\Psi_{ij}]$: probability that a worker will choose to reside in zone i and be employed in zone j . |

References

- Anas, A., Rhee, H.-J., in press. When are urban growth boundaries not second-best policies to congestion tolls? *Journal of Urban Economics*.
- Anas, A., Xu, R., 1999. Congestion, land use and job dispersion. *Journal of Urban Economics* 45, 451–473.
- Arnott, R., 1979. Unpriced transport congestion. *Journal of Economic Theory* 21, 294–316.
- Arnott, R., MacKinnon, J.G., 1978. Market and shadow land rents with congestion. *American Economic Review* 68 (4), 588–600.
- Bento, A.M., Franco, S., Kaffine, D., 2006. The efficiency and distributional impacts of alternative anti-sprawl policies. *Journal of Urban Economics* 59, 121–141.
- Brueckner, J., 2000. Urban sprawl: diagnosis and remedies. *International Regional Science Review* 23, 160–179.
- Burchfield, M., Overman, H.G., Puga, D., Turner, M.A., 2006. Sprawl: a portrait from space. *Quarterly Journal of Economics* 121 (2), 587–633.
- Cheshire, P., Sheppard, S., 2002. The welfare economics of land use planning. *Journal of Urban Economics* 52, 242–269.
- Cullen, J.B., Levitt, S., 1999. Crime, urban blight and the consequences for cities. *Review of Economics and Statistics* 81 (2), 159–169.
- Dixit, A., Stiglitz, J., 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67 (3), 297–308.
- Downs, A., 1992. *Stuck in Traffic*. Brookings Institution and the Lincoln Institute of Land Policy, Washington, D.C.
- Dunphy, R., 1997. *Moving Beyond Gridlock: Traffic and Development*. The Urban Land Institute, Washington, D.C.
- Ewing, R., 1997. Is Los Angeles-style sprawl desirable? *Journal of the American Planning Association* 63 (1), 107–126.
- Glaeser, E., Kahn, M., 2004. Sprawl and urban growth. *Handbook of Urban and Regional Economics*, vol. 4. North-Holland, Amsterdam.
- Gordon, P., Kumar, A., Richardson, H., 1989. The influence of metropolitan spatial structure on commuting time. *Journal of Urban Economics* 26 (2), 138–151.
- Hamilton, B., 1982. Wasteful commuting. *Journal of Political Economy* 90 (5), 1035–1053.
- Kain, J.F., 1991. Deception in Dallas: strategic misrepresentation in rail transit promotion and evaluation. *Journal of the American Planning Association* 56, 184–196.

- Kanemoto, Y., 1977. Cost–benefit analysis and the second-best land use for transportation. *Journal of Urban Economics* 4, 483–503.
- Keeler, T., Small, K.A., 1977. Optimal peak-load pricing, investment, and service levels on urban expressways. *Journal of Political Economy* 85, 1–25.
- Knaap, G.J., 1985. The price effects of urban growth boundaries in metropolitan Portland, Oregon. *Land Economics* 61 (1), 26–35.
- Krauss, M., Mohring, H., Pinfeld, T., 1976. The welfare costs of non-optimum pricing and investment policies for freeway transportation. *American Economic Review* 66 (4), 532–547.
- Lee, C.-M., Linneman, P., 1998. Dynamics of the greenbelt amenity effect on the land market: the case of Seoul's greenbelt. *Real Estate Economics* 26 (1), 107–129.
- Margo, R., 1992. Explaining the postwar suburbanization of the population in the United States: the role of income. *Journal of Urban Economics* 31, 301–310.
- Mayo, S.K., 1998. Land prices, land markets and the broader economy. *Land Lines: Newsletter of the Lincoln Institute of Land Policy*.
- McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), *Frontiers of Econometrics*. Academic Press, New York.
- Mills, E.S., 1972. *Studies in the Structure of the Urban Economy*. Johns Hopkins University Press, Baltimore and London.
- Muth, R.F., 1961. Economic change and urban–rural land conversions. *Econometrica* 29, 1–23.
- Nechyba, T.J., Walsh, R.P., 2004. Urban sprawl. *Journal of Economic Perspectives* 18 (4), 177–200.
- Nelson, D., Niles, J., 2000. Observations on the causes of non-work travel growth. *Transportation Research Board Paper No. 00-1242*, 79th Annual Meeting, January 9–13.
- Pines, D., Sadka, E., 1985. Zoning, first-best, second-best and third-best criteria for allocating land to roads. *Journal of Urban Economics* 17, 167–183.
- Segal, D., Steinmeier, T.L., 1980. The incidence of congestion and congestion tolls. *Journal of Urban Economics* 7 (1), 42–62.
- Small, K.A., 1992. *Urban Transportation Economics*. Harwood, Academic Publishers, Chur, Switzerland.
- Solow, R., 1972. Congestion, density and the use of land in transportation. *Swedish Journal of Economics* 74, 161–173.
- Son, J.-Y., Kim, K.-H., 1998. Analysis of urban land shortages: the case of Korean cities. *Journal of Urban Economics* 43, 362–384.
- Sullivan, A.M., 1983. The general equilibrium effects of congestion externalities. *Journal of Urban Economics* 14, 80–104.
- Walters, A.A., 1961. The theory and measurement of the private and social cost of highway congestion. *Econometrica* 29, 676–699.
- Wheaton, W.C., 1978. Price-induced distortions in urban highway investment. *Bell Journal of Economics and Management Science* 9, 622–632.
- Wheaton, W.C., 1998. Land use and density in cities with congestion. *Journal of Urban Economics* 43, 258–272.