

POLLUTION AND LAND USE: OPTIMUM AND DECENTRALIZATION*

March 2008

Oded Hochman, Department of Economics, Ben Gurion University of the Negev
Gordon Rausser, Robert Gordon Sproul Distinguished Professor, University of California,
Berkeley
Richard Arnott, Distinguished Professor, Department of Economics, University of California,
Riverside

*The authors would like to thank the seminar participants at THEMA (Paris) and Boston College for their constructive comments on this paper.

1. Introduction

For many decades and throughout much of the world, the tension between industrial pollution and households has been crucial in urban economies. Some cities, e.g., Washington, D.C., have introduced green buffer zones near densely populated areas while others have intermixed industry and households (e.g., Lima, Shanghai, Bangkok, Moscow, etc.) What cities have come to recognize is that space can and should be used as a means of controlling pollution. Separating polluter and pollutee typically reduces pollution damages but leads to increased commuting costs. When pollution damages are low relative to transport costs, separation into industrial and residential areas is uneconomic and the uniform distribution of industry and housing over space is economic. However, above a certain threshold, the division of housing and industry into separate residential and industrial areas becomes desirable. As pollution damages relative to transport costs rise, increasing separation into larger areas becomes more efficient. This paper focuses on the role of space in the control of pollution externalities. Accordingly, we concentrate on pollution from stationary sources and avoid dealing with congestion and vehicle emissions for which separation by space does not reduce damages.¹

The existing papers on spatial pollution from stationary sources (Tietenberg [20,21,22], Henderson [5,6,7], Hochman and Ofek [9], and Baumol and Oates [4]) have the common weakness that they all take the pattern of land use between housing and industry as fixed, assuming that housing is in one zone and industry is in another. This paper relaxes this assumption, treating as endogenous the pattern of land use. Specifically, this paper characterizes the optimal resource allocation and joint location of polluting firms and their

¹ Note that this applies only to commuting inside the residential and industrial zones. Travel on interstates and freeways near residential areas can be considered stationary and our results apply. Indeed, the increasingly common bypasses around highly populated areas separate expressways and dwellings, resulting in green buffer zones between the expressways and residential areas.

workers' housing around a circle, as well as policies that decentralize the optimum. To eliminate those factors which are not essential for isolating the role of land use in pollution control we: *i*) specify a city without a predetermined center; *ii*) assume a constant returns to scale production function so that production processes are not the source of any endogenous separation and agglomeration of housing and industry; *iii*) assume a ring-shaped city to avoid dealing with edge-of-city effects; and *iv*) assume all workers to be identical and all firms to be identical to avoid the complications introduced by heterogeneity. Accordingly, if pollution does not exist, a uniform layout of the city emerges with factories and houses intermixed.

In studies where agglomeration is due to positive production externalities such as external scale economies, (e.g. see Lucas [10], Lucas and Rossi-Hansberg [11] and Rossi-Hansberg [18]), an industrial zone is located in the midst of a residential zone and intensities of land use increase with proximity to the joint center of the two zones. Such a layout is the result of a balance between two forces of attraction operating on two land uses: the primary attraction between firms due to scale economies and the attraction between households and industry caused by commuting costs.

Contrary to the above studies, in our model of pollution externalities the spatial layout results from a balance between two opposing forces: one is the repulsion of households from polluting industry and the other is the attraction between households and industry caused by increasing-with-distance commuting costs. The balance between the attraction and repulsion forces leads to the agglomeration of the two land uses into a set of alternating industrial and residential zones. The intensity of land use in each zone increases with proximity to the center of the zone where the density peaks. Furthermore, depending on the specific parameters, empty buffer zones may exist between the industrial and the residential zones.

Additionally, in our model for every specified level of commuting costs, there can be an infinite number of local optima but only one global optimum. When commuting costs are very low, the global optimum entails a single industrial zone and a single residential zone. When commuting costs rise above a certain threshold, the global optimum changes to an allocation with two (or more) industrial zones in each of which industry agglomerates, and two (or more) residential zones in each of which households agglomerate. As commuting costs continue to rise, successive thresholds are reached, each with more industrial zones and an equal number of additional residential zones, until a final threshold is reached above which the global optimum is a uniform allocation of mixed residential and industrial land uses without commuting.

We also investigate decentralization of the global optimum. Spatially differentiated Pigouvian taxes per unit emission levied on industrial polluters will not generally support the optimum in either the short run (fixed household and firm locations) or the long run (endogenously determined locations). Whether or not the model's solution entails separating land uses, only if the dispersion function is linear in emissions or if locations are predetermined and fixed and the dispersion function is convex in emissions will the typical Pigouvian taxes offered in the literature (Baumol and Oates [4], Rausser and Lapan [17] Spulber [19]) be optimal. Henderson [5] showed the insufficiency of Pigouvian taxes, proposing an additional lump-sum tax along with the Pigouvian tax. Hochman and Ofek [9] proved that the optimum can be achieved by levying a tax on each unit of industrial land equal to the spatial aggregate of added damages contributed by that unit of land. In a non-spatial model, Polinsky [16] demonstrated the failure of the Pigouvian tax and also derived a tax equal to the added damages caused by a firm. Our analysis shows, under more general conditions than considered in previous papers, that a spatially differentiated added-damages tax is sufficient to achieve the

global optimum. We also argue that with our specifications a laissez-faire solution will always yield an inefficient allocation without zoning and without commuting.

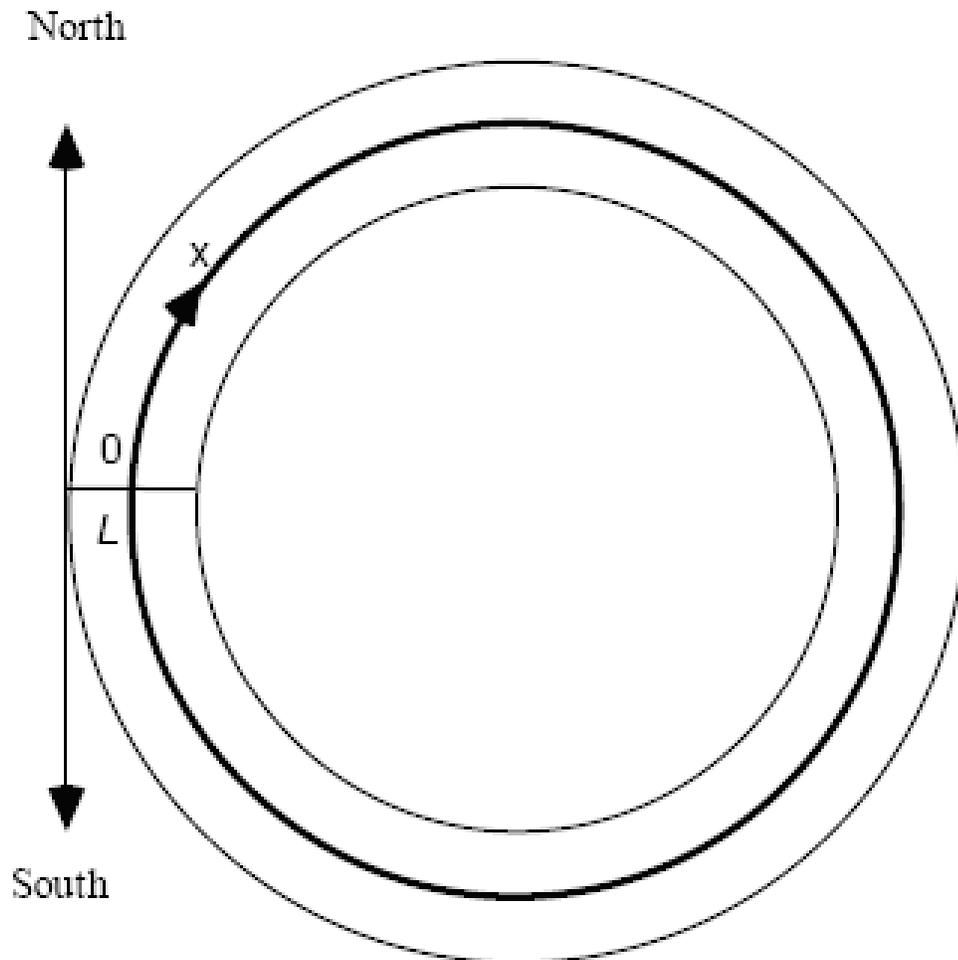
The following section presents the model. Section 3 specifies the social optimum problem. Section 4 derives and investigates conditions for a local optimum, and the price system that supports it. To gain insight and intuitive understanding, section 5 investigates a number of special cases using bid-rent analysis. Section 6 characterizes the local optima where the number of zones is predetermined, based on the interpretation of the special cases, and section 7 describes the global optimum. Section 8 presents several concluding remarks.

2. Model Specification

Assume a ring-shaped featureless strip of land of unit width. Let L be the circumference of the circle equidistant from the two boundary circles of the ring (see Fig. 1); as a result, L is also the total area of the ring. This circle is the location axis in the ring. The point due west on this ring is arbitrarily chosen as the origin. The clockwise distance from the origin is designated by x ; $x=0$ and $x = L$ are the two coordinates of the origin and $0 \leq x \leq L$. Only circumferential travel is costly.

Fig. 1

A Ring-Shaped City



Firms produce a (numéraire) composite good, using a constant-returns-to-scale, neoclassical production technology, with land and labor inputs and pollution emissions as a by-product. In particular, output per unit distance at x is $F(a(x), a(x)n(x), a(x)e(x)) = a(x)f(n(x), e(x))$, where $a(x)$ is the proportion of land occupied by industry at x , $n(x)$ the number of workers per unit of industrial land at x , $e(x)$ the quantity of emissions per unit of industrial land at x , and

$f(n(x), e(x))$ the output per unit of industrial land at x , where the intensive (per unit land) production function $f(n, e)$ fulfills $f(\lambda n, \lambda e) < \lambda f(n, e)$ for $\lambda > 1$.

Each household commutes with transport cost per unit distance of t units of composite good, to a firm to which it supplies one unit of labor. A household derives utility from land and the composite good, and disutility from the pollution concentration at its residence. In particular, the household at x receives utility $U(h(x), z(x), c(x))$, where $h(x)$ is lot size, $z(x)$ composite good, and $c(x)$ the concentration of pollution; the utility function is quasi-concave in h , z , and c .

The economy is open in the sense that households migrate freely between the economy and the rest of the world so that

$$(1) \quad U(h(x), z(x), c(x)) = U_0$$

at all settled locations, where U_0 is the exogenous utility level.

In general, the concentration of pollution at x is a functional depending on the spatial distribution of emissions, characterized by $e(y)$ and $a(y)$, as well as all y and x , i.e., $c(x) = C(\langle e(y) \rangle, \langle a(y) \rangle, \langle y \rangle, x)$, where y is another index of location and $\langle \rangle$ around a function denotes the entire range of the function's values, i.e., for all $0 \leq y \leq L$. To make the model analytically tractable, we make three simplifying assumptions:

$$i) \quad C(\langle e(y) \rangle, \langle a(y) \rangle, \langle y \rangle, x) = C(\langle a(y)D(e(y), x - y) \rangle),$$

where $D(\cdot)$ is the pollution dispersion function. According to this assumption doubling the land area at y devoted to industry has the same effect on pollution concentration at x as doubling $D(\cdot)$.

$$ii) \quad C(\langle a(y)D(e(y), x-y) \rangle) = \int_0^L a(y)D(e(y), x-y)dy .$$

This assumption means that the pollution concentration at x is additive² in the pollution contributions from different y .

$$iii) \quad D(e(y), |x-y|) = \begin{cases} D^+(e(y), x-y) & \text{for } y \in [x - \frac{L}{2}, x] \\ D^-(e(y), y-x) & \text{for } y \in [x, x + \frac{L}{2}] \end{cases}$$

This specification allows pollution emissions at y to affect the concentration of pollution at x differently, depending on whether pollution travels clockwise to x from ($y \in [x - \frac{L}{2}, x]$) or counterclockwise from ($y \in [x, x + \frac{L}{2}]$).

Combining these assumptions gives

$$(2) \quad c(x) = \int_{x-L/2}^x a(y)D^+(e(y), x-y)dy + \int_x^{x+L/2} a(y)D^-(e(y), y-x)dy .$$

$D^+(e, x-y)$ has the following properties:

$$(3) \quad \frac{\partial D^+(e, x-y)}{\partial e} = D_1^+ > 0, \frac{\partial D^+(e, x-y)}{\partial (x-y)} = D_2^+ < 0, D^+(0, \cdot) = 0, \text{ and}$$

$$D^+(e, x-y) = 0 \text{ for } x-y \geq \frac{L}{2} .$$

² As shown in Arrow *et al.* [3] and the references contained therein, economies or diseconomies of scale can exist in the assimilative powers of the environment when the density of concentrations at a given location gets close to a breakdown point of biological systems. This means that, contrary to this assumption, concentration at a given location is not just the addition of contributions from different sources, but is a function of concentration and emissions levels at different locations. Indeed, regulatory agencies have been employing complex nonlinear simulation models to represent the emission/dispersion process (see, for example, Allegrini and De Santis [1] and the NTIS, US Department of Commerce [15]).

Our specification does not allow the contribution of pollution at y to the pollution concentration at x to depend in a non-linear fashion on the emissions at some other location z (as did the specification of Tietenberg [22] and Henderson [5]). We address later whether the policy results we derive are affected by our simplifying assumptions concerning the form of $c(\cdot)$.

Analogous properties are assumed for $D^-(\cdot)$, with $y - x$ replacing $x - y$, and it is furthermore assumed that $D^+(e, 0) = D^-(e, 0)$.

The simplest reasonable concentration function would have $c(x) = \int_0^L a(y)e(y)g(|x - y|)dy$, where $g(\cdot)$ is the distance-decay function. Our more general specification allows pollution to be directionally asymmetric, the pollution contribution from y to depend not only on the total emissions at y but also on the intensity of emissions.

3. The Social Optimum Problem

We are now in a position to set up the social optimum problem. The objective function is *net city surplus*, the amount of the composite good left over after commuting expenses and consumption of the composite goods by the city's workers. This is maximized subject to the open city constraint (1), land utilization constraints, and constraints describing the technologies of pollution concentration (2) and commuting.

We start with the commuting technology. We impose as an assumption an obvious property of the social optimum that cross commuting does not occur. Thus, all households living at a particular location commute to work in the same direction. Define $T(x)$ to be the number of workers who cross x clockwise on the journey to work or minus the number who cross counterclockwise. With clockwise commuting, increasing x by dx increases $T(x)$ by $\frac{b(x)}{h(x)}dx$, the number of residents between x and $x + dx$, where $b(x)$ is the proportion of residential land at x , minus $a(x)n(x)dx$, the number of workers there. With counterclockwise commuting, increasing x by dx increases the number of workers crossing x by $\left(a(x)n(x) - \frac{b(x)}{h(x)} \right) dx$. Since,

however, $T(x)$ is measured negatively with counterclockwise commuting, for travel in either direction:

$$(4a) \quad \dot{T}(x) = \frac{b(x)}{h(x)} - a(x)n(x),$$

where a dot above a function indicates differentiation with respect to x , with

$$(4b) \quad T(0) = T(L) = 0$$

$T(0) = T(L)$ since the total number of households in the city equals the number of workers, and $T(0) = 0$ forces the origin to be a point not crossed by workers (this entails no loss of generality since, as shown later, every solution has at least two points commuters do not cross). Eqs. (4a) and (4b) together imply the commuting constraint:

$$(4c) \quad T(x) = \int_0^x \left[\frac{b(y)}{h(y)} - a(y)n(y) \right] dy.$$

The relevant *land-utilization constraints* are

$$(5a,b,c) \quad a(x) + b(x) - 1 \leq 0 \quad a(x) \geq 0 \quad b(x) \geq 0;$$

when the first constraint is not binding, at least some land at x is vacant.

Net city surplus is given by

$$(6) \quad S = \int_0^L \left[af(n, e) - \frac{b}{h}z - |T|t \right] dx.$$

The first term on the RHS is the aggregate production of the composite good, the second the aggregate consumption of the composite good by the city's residents, and the third the aggregate commuting expenses, which are calculated as the number of commuters who travel

across the interval $[x, x + dx]$, $T(x)dx$, times t , commuting cost per unit distance, and summed over locations.

Maximization of S in (6), subject to (1), (2), (4a), and (5), with (4b) as terminal conditions, provides the necessary and sufficient conditions for *parochial*³ efficiency. In the next section we describe and interpret these conditions, which constitute a subset of the necessary conditions for Pareto optimality for the economy as a whole (see Hochman [8]). Clearly, since we are dealing with a non-convex problem some of the variables might have corner solutions and there might be more than one such local optimum. These issues are addressed in the determination of the global optimum (see Sections 6 and 7).

According to the Second Welfare Theorem, any Pareto optimal allocation with convex production functions and quasi-concave utility functions can be decentralized as a price quasi-equilibrium with transfers (see Mas-Colell *et al.* [12], Proposition 16.D.1). In the presence of externalities this means that competitive markets support the optimum with government intervention limited to corrective taxes and lump sum income redistributions. In the rest of the paper we shall consider allocations that are locally optimal. And when we discuss decentralization of a local optimum, we shall use the term *supporting price system* to refer to a price vector that supports the local optimum allocation under minimal government intervention.⁴

4. The Local Optimum Solution and its Supporting Price System

³ The efficiency concept we employ—which we term *parochial efficiency*—is only for our city, not for the rest of the economy. Parochial efficiency is necessary for global efficiency. Note that in our model residents have no unearned income and all the profits of the city go to non-residents.

⁴ There are three groups of agents in our economy: households, firms, and a government or city developer. The government/city developer owns the land and sets the tax rates, receiving land rents and tax revenue as income. The rest of the economy is competitive with households and firms being price-takers. Thus, we define a competitive equilibrium with taxes to be a gross-of-tax price vector (over emissions, wage, and land rents at different locations) such that: i) each household maximizes its utility subject to its budget constraint; ii) each firm maximizes its profit; and iii) land and labor markets clear.

The constrained maximization problem described above, as well as the necessary conditions for a local optimum, is given in Appendix A. In this section, we present and interpret each condition in turn and indicate its implications for market decentralization.

- *Employment*

If $\Psi(x)$ represents the co-state variable corresponding to $T(x)$ in the commuter equation of motion (4a), then $\Psi(x)$ is the social cost of placing a household at x . From the optimization condition (7), employment is determined by setting the marginal productivity of labor equal to $\Psi(x)$, i.e.

$$(7) \quad a(x) \{f_1(n(x), e(x)) - \Psi(x)\} = 0. \quad \left(f_1 \equiv \frac{\partial f}{\partial n}, \text{etc.} \right)$$

The fact that $a(x)$ multiplies the expression in (7) means that the equality of the expression in the brackets to zero must hold only where industry is located (not necessarily exclusively).

Choosing optimally the number of commuters yields

$$(8) \quad \Psi(x) = \text{sign}(T(x))t,$$

where

$$(9) \quad \text{sign}(\chi) = \begin{cases} +1 & \text{iff } \chi > 0 \\ 0 & \text{iff } \chi = 0 \\ -1 & \text{iff } \chi < 0. \end{cases}$$

The function $\text{sign}(T(x))$ is constant as long as $T(x)$ does not change its sign. Therefore, along a segment where the sign of $T(x)$ remains constant, (8) indicates that $\Psi(x)$ is a linear function of x and increases (decreases) by t per unit distance. Thus $\Psi(x)$ is the shadow wage at locations of employment and in locations where industry does not exist $\Psi(x)$ equals the shadow wage minus commuting costs.

Let $w(x)$ be the local net earnings (LNE) at location x in the supporting market solution. In a location where an industry is sited, $w(x)$ is the wage rate, and in a location where there is no industry, $w(x)$ is the wage rate where the household works minus commuting cost to the workplace. It follows that $w(x)=\psi(x)$.

- *Residential Land*

$$(10) \quad U_h/U_z = \rho(x) - \mu(x); \mu(x) \geq 0, \mu(x)b(x) = 0.$$

Note that $\rho(x)$ is the multiplier on the land constraint (5a), and is interpreted as the shadow (land) rent at location x and $\mu(x)$ is the slack variable for $b(x)$ in (5b). Thus, (10) states that, at location x , the marginal rate of substitution between land in residential use and the composite good, which can be interpreted as the rent on land in residential use, equals the shadow rent when at least some land there is in residential use, and is less than the shadow rent otherwise.

Define $r(x)$ to be the land rent in the supporting price system. It follows from (10) that $r(x) = \rho(x)$ at residential locations.

- *Household Budget Constraint*

$$(11) \quad b(x) \left[\Psi(x) - z(x) - \frac{U_h}{U_z} h(x) \right] = 0.$$

Eq. (11) states that at all residential locations, the (net-of-commuting-cost) social benefit of locating a household there, $\Psi(x)$, equals the social opportunity cost of doing so. In the corresponding market equilibrium, the equation is the household budget constraint, that the household's income net of commuting cost equals its expenditure on residential land and the composite good.

- *Pollution concentration*

$$(12) \quad \eta(x) = -\frac{b(x) U_c(x)}{h(x) U_z(x)}.$$

Eq. (12) indicates that $\eta(x)$, the shadow price of pollution concentration at x or alternatively the marginal damage of pollution concentration there, equals minus the population density times the marginal rate of substitution between pollution concentration and the composite good.

• *Pollution Emissions*

$$(13a) \quad a(x)[f_2(x) - M(x)] = 0,$$

where

$$(13b) \quad M(x) = \int_x^{x+L/2} \eta(y) D_1^+(e(x), y-x) dy + \int_{x-L/2}^x \eta(y) D_1^-(e(x), x-y) dy.$$

$M(x)$ is the marginal damage from pollution emitted at x . A unit increase of pollution emitted at x augments the concentration at y by $D_1^i(e, |y-x|)$, $i=+$ or $-$. $M(x)$ is obtained by multiplying this increase in pollution concentration at y by $\eta(y)$, the marginal damage caused by a unit concentration there, and summing over all possible y . Thus, (13a) states that at all industrial locations the damage from a unit increase in emissions equals the value of the additional output created. In the supporting price system, $M(x)$ equals the Pigouvian tax at x .

• *Industrial Land*

$$(14a) \quad f[n(x), e(x)] - \Psi(x)n(x) - Q(x) = \rho(x) - \gamma(x); \quad \gamma(x) \geq 0, \quad \gamma(x)a(x) = 0,$$

where $\gamma(x)$ is the slack variable for $a(x)$ in (5c), and

$$(14b) \quad Q(x) = \int_x^{x+L/2} \eta(y) D^+(e(x), y-x) dy + \int_{x-L/2}^x \eta(y) D^-(e(x), x-y) dy$$

is the additional damage caused by the total emissions from a unit area of land at x . The first term on the left-hand side of (14a) is industrial output per unit of land and the second the wage

bill. The first term on the right-hand side of (14a) is the shadow price of the land utilization constraint (5a). That $\gamma(x)$ is non-negative requires that the left-hand side of (14a) not exceed $\rho(x)$, which in turn must fulfill

$$(15) \quad \rho(x) \geq 0; \quad \rho(x)[1 - a(x) - b(x)] = 0.$$

Thus, (14) states that at all industrial locations, land rent equals residual income (revenue less the wage bill).

If in the supporting equilibrium $Q(x)$ is levied as a tax per unit of industrial land and there is no Pigouvian tax imposed, at all locations with industry $\rho(x)$ equals the land rent $r(x)$.⁵ In order to satisfy (13a), however, at first glance it appears that $M(x)$ should be levied as a per unit emission tax, and that only when the dispersion functions are linearly homogeneous in e is the tax burden the same in the two cases. The following proposition resolves this apparent difficulty.

Proposition 1: To achieve efficiency in a market economy by taxing pollution emissions, a tax per unit of industrial land must be levied at every industrial location. This tax must equal the added damages caused by the pollution emissions from this unit of land $Q(x)$.

Proof: If an industrial producer pays $Q(x)$ for emitting $e(x)$ per unit land, wages of $w(x)n(x)$, and land rent of $r(x)$, a long-run equilibrium with zero profits will satisfy both (13) and (14)

since $\frac{\partial Q(x)}{\partial e(x)} = M(e(x))$.

Corollary 1: Each local optimal allocation has a supporting equilibrium with its own price system and corrective pollution taxes; so too does the global optimal allocation.

⁵ Note that in the supporting equilibrium wages are the only source of income of residents and land rents and proceeds from taxes go to nonresidents.

In subsequent sections we shall use the supporting price system and the supporting equilibrium relations together with the optimum relations to characterize the global optimum. The key elements of the supporting equilibrium can be insightfully expressed in terms of the industrial and residential bid-rent functions at the optimum. Specifically:

Definition 1: Define $R_I(x)$, the industrial bid-rent function, as

$$(16) \quad R_I(x) \equiv f(n(x), e(x)) - \Psi(x)n(x) - Q(x),$$

where $n(x)$, $e(x)$, $\Psi(x)$ and $Q(x)$ are evaluated at a local optimum.

This bid-rent function follows from (14) and indicates the maximum amount industry can pay for land at x without suffering losses when $Q(x)$ in (14b) is imposed as a tax per unit of industrial land.

Definition 2: Given $c(\cdot)$, $\Psi(\cdot)$, and U_0 , define $R_h(\cdot)$, the household or residential bid-rent function, as

$$(17) \quad R_h(x) \equiv \max_{z(x), h(x)} \frac{\Psi(x) - z(x)}{h(x)} \quad s.t. \quad U(h(x), z(x), c(x)) = U_0.$$

From (10), $R_h(x) = \frac{U_h(x)}{U_z(x)}$. $R_h(x)$ is the maximum amount a household can afford to pay per unit of land, consistent with achieving utility U_0 .

Using equations (14) and (16) we obtain

$$(18) \quad R_I(x) \leq r(x), \quad R_I = r(x) \Leftrightarrow a(x) > 0.$$

Similarly, from equations (10), (11) and (17),

$$(19) \quad R_h(x) \leq r(x), \quad R_h = r(x) \Leftrightarrow b(x) > 0.$$

Equations (18) and (19) imply that an activity (production or consumption) takes place at a given location if and only if its bid rent there equals the market land rent. Finally, from (18) and (19) the land rent $r(x)$ can be determined by

$$(20) \quad r(x) = \max[0, R_h(x), R_I(x)].$$

The above definitions and relations imply the following bid-rent rule⁶ :

Lemma 1 (Bid-rent Rule): Consider the residential and industrial bid-rent functions at a local optimum. Near a point of intersection of the two functions, only the land use with the larger derivative (with respect to x) of its bid-rent function is located immediately clockwise of the intersection point and only the other land use is located counterclockwise of it. If the two derivatives are equal at the point of intersection, the two bid-rent functions coincide in a neighborhood of this point and housing and industry may coexist there. Locations where both bid rents are negative are empty buffer zones.

The proof is provided in Appendix B (available on the internet).

Corollary 2: Pigouvian taxes are distortive when $D_{II} \neq 0$.

The proof is based on two principles: First, that the marginal payment per unit of emission e is the same and optimal under the two tax regimes (see Proposition 1); and second, that total emission payments at a particular location are different in the two tax regimes when $D_{II} \neq 0$, implying that the industrial bid-rent function under Pigouvian taxation is not optimal. Thus, if the industry is located optimally (e.g. by zoning regulations), its emissions are optimal in the two tax regimes. However, the allocation of land between industry and residence can be suboptimal if only Pigouvian corrective taxes are used.

To clarify this result, consider the case of $D_{II} < 0$. Then $Q > eM$, which implies that levying a Pigouvian emissions' tax, $M(x)$, is insufficient to support the optimum. For a solution in which industry and housing are intermixed and the Pigouvian tax eM is levied, the industrial bid-rent function rises above its optimal value, which in turn causes r , the land rent, to be too high as well. Since r , a supporting price, is higher than its optimal value, the allocation supported is not

⁶ See proof in Appendix B.

optimal. In this event, in each location more land is allocated to the industry than in the optimum and less land is allocated to housing.

For the same situation, i.e., $D_{11} < 0$ and the Pigouvian tax eM levied, but with residential and industrial land separated, once again the industrial bid-rent function and hence land rents are higher than their optimal values, but now only in the industrial area. This leads to a larger-than-optimal industrial zone, and a residential zone which is smaller and more heavily polluted than optimal.⁷

When $D_{11} > 0$ (and hence $Q < eM$) under the Pigouvian tax the industrial bid-rent function is lower than optimal. Whether the land uses coexist at the same location or are separated, industry occupies less land, produces less output and pays lower wages than at the optimum.⁸ When $D_{11} = 0$, then $Q = eM$ and the allocation is optimal in the two tax regimes.

How robust is the optimality of the per unit land corrective tax $Q(x)$ introduced in Proposition 1? From (14) it can be seen that its optimality hinges on the assumptions of constant returns to scale in production and the additivity of the dispersion function. When these assumptions are relaxed, we conjecture that the corrective tax still equals the pollution damages

⁷ Henderson [5] has shown that in a spatial setting over the short run, Pigouvian taxes are efficient when the dispersion function is weakly convex in emissions ($D_{11} \leq 0$). In a non-spatial model, Spulber [19] and Baumol and Oates [4] have shown that Pigouvian taxes provide the proper incentive for firms to produce the optimal output in the short run by using the optimal mix of inputs. Spulber has also argued that when the damage function is convex in emissions, Pigouvian taxes provide the proper incentives for entry and exit of firms in the long run. However, Pigouvian taxes fail to achieve efficiency in our spatial framework because the generated externality does not cause the actual damages. The emissions are the direct external effects of the production process, but what causes the damages are concentrations. Concentrations are created by emissions from different sources via non-linear (dispersion) functions. It is clear from equations (14) that if $\eta(y)$ could be levied as a tax per unit of concentration contributed by the firm, efficiency would be attained. This means that Pigouvian taxes are efficient when levied on concentrations rather than on emissions. However, producers create emissions, and only when the relation between emissions and concentrations is linear can taxes on emissions be optimal. Accordingly, a necessary condition for Pigouvian taxes to be effective is that the accumulation process of concentrations from different sources be additive in emissions, the external effect itself. This will occur only when $D_{11}=0$, a result rarely satisfied (see footnote 1).

⁸ This claim seems more plausible in view of Polinsky's results [16].

added by the firm, i.e. total pollution damages with the firm's emissions minus pollution damages without it, but that the unit of taxation is the firm.⁹

5. Laissez Faire and Special Cases¹⁰

Thus far, we have derived necessary conditions for a local optimum. The remainder of the paper explores how the spatial structure of the global optimum changes as t , the unit commuting cost, increases. This section concentrates on special cases. Section 5.1, depicting the laissez-faire allocation, is a detour from the investigation of the social optimum. Section 5.2 provides some preliminary definitions, and section 5.3 considers three special cases with one extreme parameter value, which give insight into the economic determinants of the globally optimal spatial structure. In section 6 we examine how an increase in unit commuting costs t affects the solution with a given number of zones, and in section 7 how commuting costs affect the number of zones in the global optimum.

5.1 Laissez Faire

The laissez-faire allocation corresponds to the equilibrium with no government intervention. This allocation is inefficient and is the only allocation considered in the paper that is not related to a local optimum. Under this allocation each atomistic firm, taking pollution concentrations as given, locates right next to its workers, reasoning that by doing so it eliminates their community costs and hence can pay them a lower wage. Thus, there is no commuting and the

⁹ Polinsky [16] provides a non-spatial example where Pigouvian taxes fail to achieve efficiency. In his analysis of strict liability and negligence, Polinsky utilizes a partial equilibrium model almost identical in its mathematical exposition to that of Spulber's model, with one small difference. In Polinsky's model, 'care' (the equivalent of negative emissions in our and Spulber's models) reduces external damages caused by the individual firm, i.e. the amount of care provided by a firm is an argument with a negative effect in a separate damage function of the individual firm, which transforms emissions of each firm into monetary terms. These individual money damages are then accumulated to obtain the total social damages. In Spulber's model, the emissions of the individual firms are added first and the accumulated amount of emissions is then converted to monetary terms via a single social damage function. Both models are correctly specified and the differences in their specifications follow from differences in the issues examined. These differences lead to what appear to be contradictory results of the two models; while in Spulber's model Pigouvian taxes provide long-run efficiency, in Polinsky's model they do not. In Polinsky's model, the separate damage functions introduce the non-linearity which in our model is introduced via the dispersion function.

¹⁰ Proofs and technical elaboration of some cases appear in Appendices B and C, available on the internet.

marginal productivity of emissions is zero. Firms collectively fail to take into account that the added emissions resulting from their location choices lead to wage increases that are needed to compensate for added pollution damages. Such allocations can be found in practice only in rural villages since agriculture is one of the few industries with technology close to constant returns to scale.

5.2 Preliminaries to the Special Cases

There are two relevant principal solution types¹¹. One is an interior solution in which land use is mixed; i.e., $a(x) > 0$, $b(x) > 0$ and $a(x)+b(x)=1$ for all x . Any such *mixed allocation* satisfying the necessary conditions of the previous section is a local optimum.¹² In this case the two bid-rent functions $R_I(\cdot)$ and $R_H(\cdot)$ coincide everywhere. All other possible locally optimal allocations are corner solutions and involve separation of industrial and residential land.

Definition 3: A separated allocation is an optimal allocation in which industrial and residential land use are strictly separated. Thus there are industrial zones and residential zones. An empty area with no land use is also allowed; such an area is termed a buffer zone

In each separated allocation, industrial and residential zones alternate, perhaps separated by buffer zones.¹³ A buffer zone exists between an industrial and a residential zone if there is a

¹¹ The globally optimal land allocation may be an “empty” city, i.e., no households and no industry. This outcome will occur if in all local optimum solutions the maximized surplus is negative. Namely, the price of the city’s export product is insufficient to maintain the predetermined utility level of the city residents. In the following analysis, only non-empty allocations, i.e., $N > 0, S \geq 0$, which satisfy the necessary and sufficient conditions are considered.

¹² Note that often in problems involving inequalities only one type of extremum can result. Here only local maxima can occur. To see this, note that a solution with a positive S cannot be a local minimum since $a(x)$ and $b(x)$ can be reduced continuously while maintaining their ratio intact and thus reducing S until it disappears. Since we can increase density and commuting distances indefinitely, we can always increase a deficit ($-S$) indefinitely too.

¹³ In the types of land use patterns analyzed in the paper, all land at a particular location is either vacant or occupied. Zones with partially occupied and partially empty locations are also a possibility, however. Housing in a partially empty location cannot arise, since positive marginal utility of housing implies that households in a particular location will use all or none of the available land at that location. In the case of industry, when $D_{22} = 0$ the linear homogeneous production functions together with diminishing marginal productivity imply that there is no empty space where industry is located, since by keeping constant overall emissions, $(a(x)e(x))$, as well as overall labor, $(a(x)n(x))$, and expanding industry across the entire space in that location x (i.e. $a(x)=1$) output can

segment of land between the two zones in which the two bid rents are negative. This may occur if, at these locations, concentration levels are too high and wages too low to support the predetermined economy-wide household utility level *and* if for the specified emission taxes and wages the industry suffers losses. There cannot be allocations with buffer zones between two residential zones or between two industrial zones.

In practice, separation into industrial and residential zones is also the result of scale economies in production and in the consumption of collective goods. However, buffer zones are unique to pollution. Note that in practice buffer zones are often green areas since plants, especially trees, help reduce pollution. Moreover, large highways and freeways in the vicinity of densely populated areas (e.g., city bypasses), are also stationary sources of noise and air pollution. Indeed, along many of these roads we observe green buffer zones near densely populated areas (e.g., Washington, D.C., Portland, Oregon). In still other cities buffer zones often exist between waste collection industries (landfills) and households (e.g., Dallas, Texas).¹⁴

Definition 4: A no-crossing location (NC location) is a location in either an industrial or a residential zone not crossed by commuters.¹⁵

Lemma 2: The value of the function $T(\cdot)$ at a NC location is zero.

Lemma. 3: In each residential and each industrial zone there is one and only one NC location (which may sometimes be extended to a NC area).

be increased without changing inputs. Furthermore, when $D_{22} > 0$, the reduction in emissions density also leads to a reduction in the contribution to concentrations, which in turn strengthens the tendency to fill in empty space or leave it entirely empty. However, when $D_{22} < 0$, reducing the density of emissions while keeping their total at the given location constant increases concentrations. Consequently, we cannot rule out the possibility that increases in the concentrations will outweigh the effect of diminishing returns in production and result in an optimal solution with industrial zones which are only partially occupied. In the subsequent analysis we shall disregard this case.

¹⁴ For permanent versus temporary diminution in bid-rent schedules for households located near landfills, see McCluskey and Rausser [13,14].

¹⁵ The NC locations are theoretical tools intended to help us in the coming analysis. In practice, because of irreversible investments, historical trends, social connections and the fact that there is more than one worker per household, such locations are unobserved.

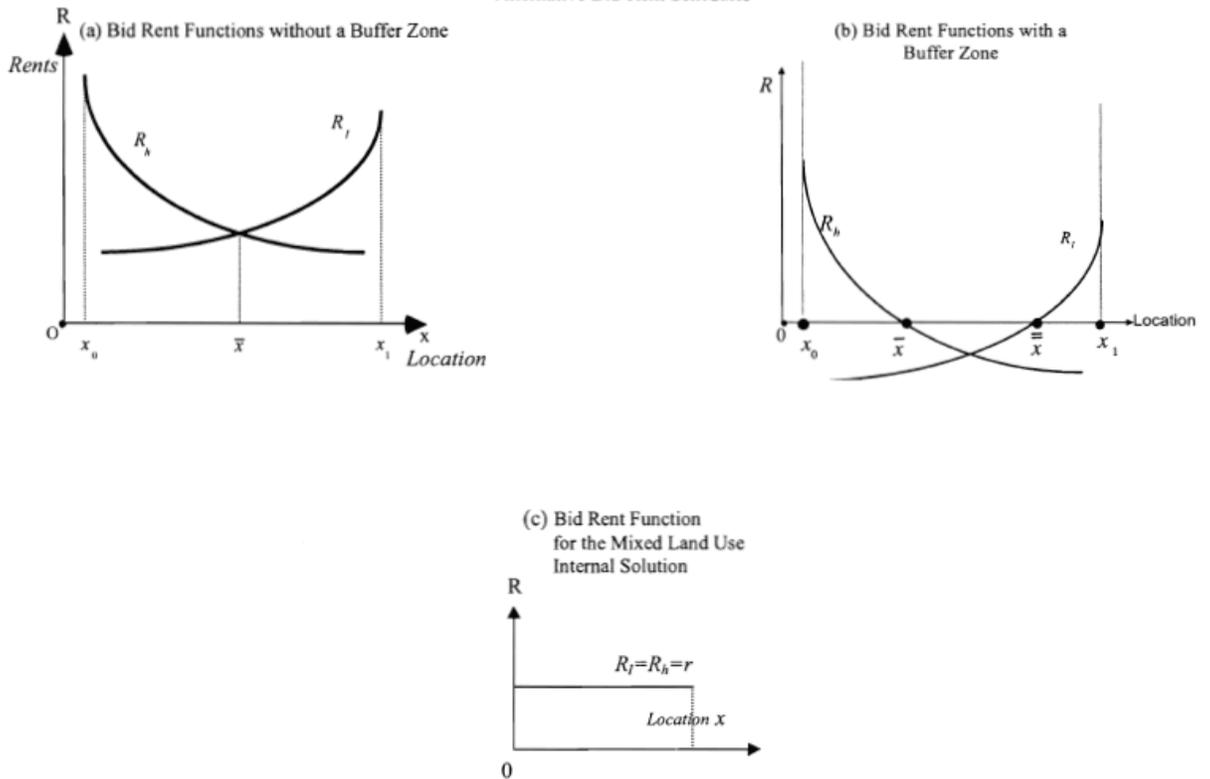
Proof: Since paths of commuters cannot cross, in a residential zone there must be a location where all those living clockwise of this location commute clockwise and all those living counterclockwise of this location commute counterclockwise. In an industrial zone there must be a location at which commuters employed clockwise of that location commute counterclockwise and *vice versa*. Each of these NC points can extend to an empty segment. Two or more NC locations with only residential and no industrial space between them cannot exist, since occupants between such locations would have to cross one of the NC points when commuting. Similarly, two NC points with only industrial and no residential space between them cannot exist.

Definition 5: An autonomous area (AA) is the area between two consecutive NC points.

An autonomous area includes part of a residential zone and part of an industrial zone, and all households who reside in an AA also work there and *vice versa*. If the allocation includes buffer zones, each AA includes an empty buffer zone between its residential and industrial zones. The concept of an autonomous area is essentially based on NC points and as a result exists in theory only.

Without loss of generality, in what follows the origin will be placed at an NC point where residents are located.

Fig. 2
Alternative Bid-Rent Schedules



With these definitions, we can now relate the bid-rent functions to the qualitative spatial structure. Let x_0 and x_1 be two consecutive NC points, the former in a residential zone and the latter in an industrial zone. Fig. 2(a) depicts a separated allocation without buffer zones, with \bar{x} being the boundary between the residential and the industrial zones at which the two bid-rent functions intersect. Industry occupies locations where $R_i > R_h$, residences occupy locations where $R_h > R_i$, and rents must be non-negative where $R_i = R_h$. In Fig. 2(b) the bid-rent functions result in a buffer zone. Both bid rents are zero at the boundaries of the buffer zone $\bar{\bar{x}}$ and $\bar{\bar{x}}$ and remain non-positive everywhere over the zone. In Fig. 2(c) the two bid rents are constant and coincide everywhere over the autonomous area; industry and housing coexist everywhere, each at its own constant density. In this case all points are NC-points.

Assumption 1 (Symmetric Dispersion Assumption): $D^+(e, y) = D^-(e, y) \equiv D(e, y)$ for all $e > 0$ and $y > 0$.

Henceforth we shall restrict our analysis to a more specific case in which dispersion is symmetric. The assumption is that pollution spreads clockwise and counterclockwise in the same way. This may happen in practice with respect to air pollution if throughout the year the wind blows in each direction with equal probability. The model can be solved under other assumptions (e.g., $D^+(\cdot) = 0$ and $D^-(\cdot) > 0$), but the solution under each assumption is different and space limitations dictate that we present only the symmetric case.

5.3 Special Cases

From the above definitions, the global optimum for three special cases can be characterized, each with one parameter having an extreme value (zero or infinity). These special cases capture the essence of the solution in general and indicate the range of possible outcomes.

5.3.1. Case Zero: Zero commuting cost

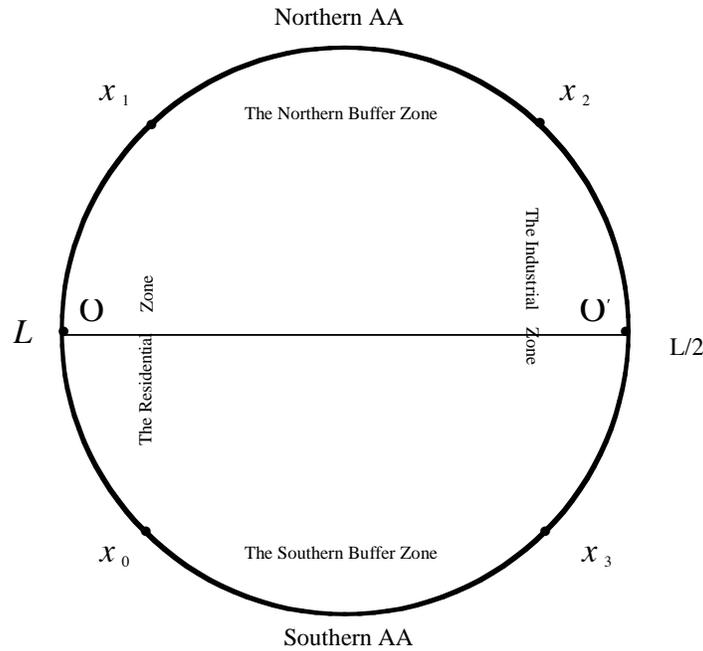
In this case all parameters are presumed to be finite and strictly positive except t , commuting cost per unit distance, which is assumed to be zero. Zero commuting costs imply the same constant shadow wage, $\Psi(x)$, everywhere. As in the general case, pollution causes positive damages that increase with concentrations at any level of consumption, i.e., $-\infty < U_c(h, z, c) < 0$. A superscript zero designates variables for this case, e.g., $r^0(x), R_I^0(x)$ and $R_H^0(x)$ specify respectively the rent function, the industrial bid-rent function and the residential bid-rent function.

Since pollution decays with distance, the greater the distance between polluter and pollutee, the lower the concentrations experienced by the pollutee, which leads to a higher utility level for a given level of the composite good and housing. Since commuting costs are zero, this separation does not involve any loss of resources. Under these conditions, separate industrial

and residential zones arise. Hence, $a^0(x)b^0(x) = 0$ for all x . Moreover, there is only one industrial zone and one residential zone, and there may or may not be buffer zones between these zones. Increasing the distance between polluter and pollutee reduces concentrations and therefore generates benefits without increasing costs. Any solution with many industrial zones can be restructured as a single industrial zone without decreasing the distance between any residential and any industrial location and with some distances increasing. An analogous argument holds for many residential zones.

If there is empty space in the midst of one of the occupied zones, land uses can be moved from the boundaries of the zone to fill it. Such reallocations increase the distances between the two land uses and thus reduce effective concentrations without entailing any cost. As a result, an allocation with empty space in the midst of an occupied zone cannot be optimal. Therefore for all x in an occupied area (a residential or industrial zone), $a^0(x) + b^0(x) = 1$. This condition together with $a^0(x)b^0(x) = 0$ implies that if one of these two variables is positive, its value must be one.

Fig. 3
The Spatial Layout of the Two Autonomous Areas Case



Two NC points emerge, one, the origin O , in the residential zone, and the other O' , in the industrial zone. $T(x)$ is zero at both NC points. Fig. 3 depicts the layout of the city ring with the boundaries of the different zones designated by x_i^0 , $0 \leq x_i^0 \leq L$, $i = 0,1,2,3$. x_0^0 is the southern boundary of the residential zone and the northern boundary of the southern buffer zone and x_2^0 is the northern boundary of the industrial zone, etc. Note if there are no buffer zones, $x_1^0 = x_2^0$ and $x_3^0 = x_0^0$. The symmetric dispersion assumption ($D^+(\cdot) = D^-(\cdot)$) implies that the allocation has OO' as an axis of symmetry.

The bid-rent functions are equivalent to the residual income per unit land in each location (eqs. (16) and (17)). The density of land use is an increasing function of the rent. When $t=0$,

the rent together with the density of land use peak at the center of each zone. The centers of each zone are also the NC points and the boundaries between the AAs. In the industrial zone, since the center is the pollution-generating location furthest from all residential locations, the optimal tax Q is at its lowest level, and R_l at its maximum there; analogously, since the center of the residential zone is least affected by pollution, R_h is at its maximum there. The lowest rents and densities are at the boundaries. Also, rents rise monotonically within an occupied zone from the boundary to the center. In a buffer zone and its boundaries, rents vanish, as does all economic activity. Small mining towns are an example of such a complete separation. Commuting costs are not zero but are negligible compared to pollution damages.

5.3.2. Case One: Pollution has no ill effects

This case presumes $U_c = 0$ and $0 < t < \infty$; namely pollution has no ill effects and commuting costs are positive. We designate the solution of this case by superscript 1. Since pollution causes no damage, the optimum entails zero commuting costs — each household lives and works at the same location. Since conditions are the same everywhere, symmetry implies that $a^1(x) = \bar{a} > 0$, $b^1(x) = \bar{b} > 0$ and $\bar{a} + \bar{b} = 1$ for all x . As a result, the land rent and wage rate are spatially constant. Rural villages are an example of this case. Villages are also an example of a laissez faire allocation, however, here the allocation is optimal.

5.3.3. Case Two: Mixed allocation

In this case, pollution causes ill effects, i.e. for all positive arguments $-\infty < U_c(h, z, c) < 0$, and t is positive and finite. As in the previous case, laborers reside next to their workplace and all variables are spatially constant. The allocation is not a separated allocation and thus concentrations are affected only through the production process. The solution for this case, distinguished by the superscript 2, is a local optimum for all t and a global optimum when t is sufficiently large.

6. Local Optima¹⁶

In general, local optima can be either mixed allocations (e.g., Case Two) or separated allocations. Case Two allocations are interior solutions since the variables $a(x)$ and $b(x)$, along with all other variables, obtain values in the interior of their domain of definition, while in the separated allocations $a(x)$ and $b(x)$ obtain boundary values and are therefore corner solutions. For a given set of parameter values, there might be several local optima that are corner solutions, each with a different number of autonomous areas. Even for a given number of AAs, there may also be more than one local optimum.

In this section we characterize a general separated allocation and its supporting price system, by investigating the changes due to an increase in commuting costs in a local optimal solution with two symmetric AAs, as depicted in Fig. 3. Similar relationships exist when the number of AAs is larger. As a reminder, we repeat the definition of the LNE, this time formally.

Definition 6: Local net earnings (LNE), $w(x)$, equal, in the supporting equilibrium, wages net of commuting costs for a household living at x . At an industrial location $w(x)$ is the wage rate. At a location where there is no industry, $w(x)$ is the wage rate net of commuting costs to the workplace.

Lemma 4: In an AA, $w(x)$, is a linear function of x and the absolute value of its slope equals t , viz.

$$(21) \quad w(x;t) = \Psi(x;t) = \begin{cases} \psi(x_2^0;t) + (x - x_2^0)t, & \text{for } 0 \leq x \leq L/2 \\ \psi(x_3^0;t) + (x_3^0 - x)t, & \text{for } L/2 \leq x \leq L \end{cases}$$

¹⁶ Proofs of the Lemmas and Propositions of this section not presented in the text appear in Appendix C in the internet version of the paper.

where x_2^0 and x_3^0 are the boundary points in special Case Zero of the industrial area in the northern and southern parts of the AA, respectively, as depicted in Figure 3, and $\psi(x_2^0; t)$ (also equal to $\psi(x_3^0; t)$), which we refer to as the intercept of $w(x; t)$, is a function of t but not of x .

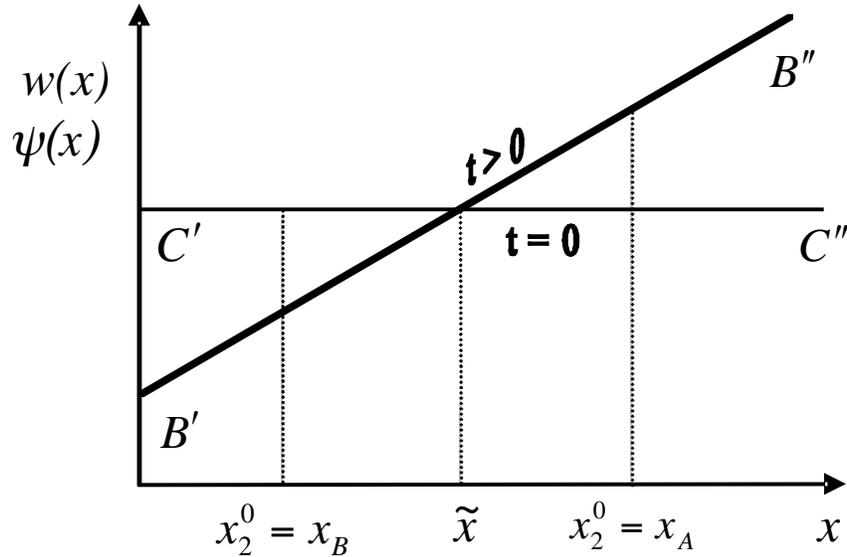
In what follows, we deal only with the northern AA while keeping in mind that the southern AA is symmetric, with OO' as the axis of symmetry (see Fig. 3). Lemma 4 implies that if t is positive, the LNE in the industrial zone increases at the rate of t per unit distance when moving from the boundary of the industrial zone towards the NC point. The opposite occurs when moving away from the boundary into the residential or buffer zone. Lemma 4 is a standard result in models with separation into distinct industrial and residential zones.

Corollary 3: *An increase in t augments the multiplier of x in $\Psi(x; t)$, and moves the intercept*

$\psi(x_2^0; t)$ by the shift factor, $\frac{\partial \psi(x_2^0; t)}{\partial t}$. The shift factor can be positive, negative, or zero,

depending on the model's details.

Fig. 4
Spatial Patterns of Two Local
Net Earning Functions



For ease of exposition, henceforth we restrict the analysis only to cases without buffer zones. Notable differences arising from the presence of buffer zones will be commented on in footnotes. Fig. 4 demonstrates the possible effects of an increase of t on $\psi(x_2^0; t)$ in the northern AA. The line $B'B''$ depicts $\Psi(x; t)$ with $t > 0$, $C'C''$ depicts $\Psi(x; 0)$, and \tilde{x} denotes the location where $B'B''$ and $C'C''$ intersect. Since the local net earnings (LNE) function pivots around \tilde{x} , we refer to \tilde{x} as the *pivot point*. There are two cases which differ according to whether \tilde{x} lies to the left or right of x_2^0 — the boundary between the residential and the industrial zone with $t=0$. When \tilde{x} lies to the left of x_2^0 , then the shift factor is positive and the wage rate increases in the AA throughout the industrial zone, while the local net earnings (LNE) in the residential zone increases near the boundary and decreases near the no crossing (NC) point 0. When \tilde{x} lies to the right of x_2^0 — the shift

factor is negative — the LNE decreases throughout the residential zone while the wage rate in the industrial zone falls near the boundary and increases near the NC point $L/2$.

The total derivative of the functions $R_k, k = h, I$ with respect to t is given by

$$\frac{dR_k(x)}{dt} = \frac{\partial R_k(x)}{\partial t} + \sum_i \frac{\partial R_k(x)}{\partial \chi_i} \frac{\partial \chi_i}{\partial t}, \text{ where } \chi_i \text{ are the controls and shadow prices of the system.}$$

Lemma 5. *In the optimum allocation, both bid rents are functions of x and of t , and*

i. *In the residential zone*
$$\frac{dR_h(x)}{dt} = \frac{\partial R_h(x)}{\partial t} = \frac{\partial \Psi(x;t)/\partial t}{h(x)}$$

ii. *In the industrial zone*
$$\frac{dR_I(x)}{dt} = \frac{\partial R_I(x)}{\partial t} = -n(x) \frac{\partial \Psi(x;t)}{\partial t}.$$

Lemma 5 demonstrates that commuting costs affect the bid-rent functions only through

$\Psi(x;t)$; the terms for the other controls disappear. The change in the industrial bid rent at a

given location is negatively related to the change in the wage rate, with the factor of

proportionality $n(x)$, the local labor density. The change in the residential bid rent at a given

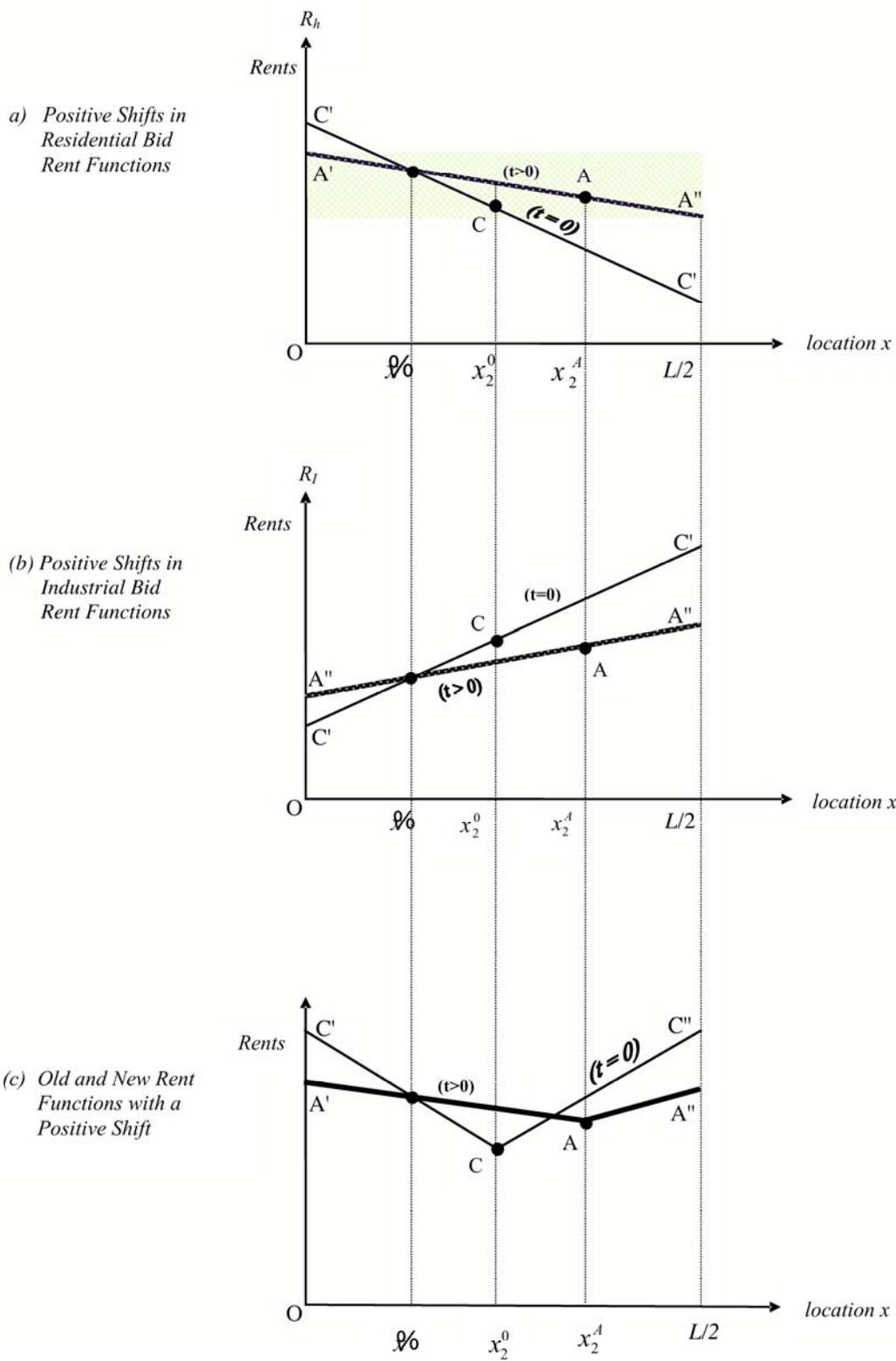
location is proportional to the change of the LNE, with the factor of proportionality $\frac{1}{h(x)}$, the

local residential density. Lemma 5 also implies that at the point where $\frac{\partial \Psi(x;t)}{\partial t} = 0$, i.e. at the

pivot point, the bid rent functions remain unchanged as well. Commuting costs cause both bid-

rent functions to pivot around \tilde{x} .

Fig. 5



In this pattern the increase in t causes the residential bid rent to increase near the boundary (x_2^0) and decrease near the origin; in Fig. 5(a), the line $A'AA'$ represents $R_h(x,t)$, and $C'CC'$, $R_h^0(x,0)$. In this pattern the increase in t causes the industrial bid rent to fall everywhere in the industrial zone, less so near the boundary (x_2^0) and more so near $L/2$; the line $A''AA''$ in Fig. 5(b) represents $R_l(x,t)$, and $C''CC''$, $R_l^0(x,0)$. The rent functions before and after the shift, depicted in Fig. 5(c), are the upper envelope curves of the before and after bid-rent functions. The boundary moves right from x_2^0 to x_2^A .

The increase in t causes the residential zone to expand and the industrial zone to shrink by the same amount. In the industrial zone, the rent declines everywhere, less near the boundary (x_2^0) and more near $L/2$. In the residential zone, the rent function increases near the boundary (x_2^0) and decreases near the origin.

Pattern II (a negative shift factor) emerges when t increases from zero to $t > 0$ and the pivot point \tilde{x} is right of the boundary point x_2^0 .

With this pattern, the increase in t causes the residential bid-rent function to decline throughout the residential zone, more near the origin than close to the boundary; in Fig. 5(d), the line $B'B$ represents $R_h(x,t)$ where $R_h(x,t) \geq R_l(x,t)$, and $C'C$ represents $R_h(x,t)$ where $R_h^0(x,0) \geq R_l^0(x,0)$. The industrial bid-rent function increases near the boundary of the industrial zone and decreases near the origin. In Fig. 5(d), BB'' represents $R_l(x,t)$, where $R_l(x,t) \geq R_h(x,t)$, and CC'' represents $R_l^0(x,0)$, where $R_l^0(x,0) \geq R_h^0(x,0)$. The boundary point between the two zones moves left from x_2^0 to x_2^A so that the industrial zone expands and the residential zone shrinks. The rent falls everywhere except near the boundary of the industrial zone.

In both patterns, the density of population in the residential zone and the density of employment in the industrial zone move in the same direction as the corresponding rents, while the total AA's population declines. Which of these patterns occurs depends on the production and pollution dispersion technologies, as well as tastes.¹⁷

Since the commuting and pollution effects influence the industrial bid rent function in the same direction, it always decreases with distance from the NC point at a decreasing rate. However, the commuting and pollution effects influencing the residential bid rent, unlike the case of the industrial bid rent, run in opposite directions and thus the residential bid rent may either decrease or increase with distance. Nevertheless, at the boundary the slope of the residential bid rent in the direction away from the boundary is always higher than the slope of the industrial bid rent. Note that when commuting costs are low compared to pollution damages (e.g., case zero) the slope of the residential bid rent when moving away from the boundary is increasing. Conversely, in the direction of the industrial zone the industrial bid rent's slope is increasing and higher than that of the residential zone. Otherwise, housing will outbid the industry in the industrial zone and industry outbid housing in the residential zone, which is impossible. In general, close to the residential NC point the residential bid rent may become convex, decline, or even become negative. However, this outcome cannot occur in the global optimum, since a solution with a larger number of smaller zones is then more efficient. With zero commuting costs (case zero), the rent function is at a local maximum at the NC points, gradually declining towards the boundary where it reaches its lowest level.

¹⁷ The above analysis can be modified straightforwardly to cover the situation when there is a buffer zone in the AA. The results are broadly similar. In a buffer zone case, however, an increase in t may cause both occupied zones to shrink and the buffer zone to expand. This will occur whenever $\psi(x, t + \Delta t)$ intersects $\psi(x, t)$ in the buffer zone. The intuition is that with an expanding buffer zone, both commuting distances and emissions increase with offsetting effects on pollution concentrations, while employment density decreases and wages increase. Indeed, within the residential zone the LNE and the rent may be lower but pollution concentrations can fall due to the increased distance.

Lemma 6: When t increases, the absolute value of the slope of the rent function falls except perhaps near the boundary, where it may increase.

Lemma 6 states that an increase in t causes the rent function (see Fig. 5) to become flatter everywhere except perhaps for pattern I over $[x_2^0, x_2^A]$ and for pattern II over $[x_2^A, x_2^0]$.

The above analysis pertains only to the case of two AAs. Symmetry dictates through Lemma 7 that the analysis can apply with no significant changes to any even number of AAs.

Lemma 7: For each $t > 0$, there may exist a local optimum solution with $2m$ zones, where m can be a subset of (or all) the integers fulfilling $1 \leq m \leq \infty$. If $2m$ is the fixed number of AAs in an allocation, so is the number of NC points which are located at the boundaries of the AAs (and in the middle of the occupied zones). All AAs are of the same size with an area of $L/2m$ and each AA is the mirror image of its neighbor AAs. The qualitative results discussed previously in this section of the effects of an increase of t on the internal structure of an AA hold for the general case as well.

We now summarize the role of the pivot point in the following corollary.

Corollary 4: If the pivot point lies inside the industrial zone, the zone expands with t and the residential zone contracts. Both population and employment density decline, but since more land is allocated to industry than before and less to housing, employment density decreases proportionally more.

If the pivot point lies inside the residential zone, the zone expands with t and the industrial zone contracts. Population density then declines proportionally more than employment density.

Corollary 4 reveals that the location of the pivot point reflects where it is efficient to allocate more land when the density of population/employment changes. Obviously, the allocation of land and with it the location of the pivot point depends on whether land is more useful in production or in consumption.

In addition to the symmetric solutions, there may be asymmetric solutions with AAs of varying sizes. Disregarding problems of indivisibility, each of the AAs in an asymmetric solution will also appear in a symmetric solution of a different number of zones. Of any two such symmetric solutions one is superior to the other and therefore superior to the asymmetric solution. We will ignore the case where we are indifferent between the two solutions and their mixture. Thus, there is always a symmetric global optimum.

The parameter t does not always have a finite upper bound above which a separated local optimum does not exist (note that $m = \infty$ is equivalent to case two, the solution of mixed land uses). It might occur that as t grows larger, buffer zones disappear and the density of land use in the occupied areas becomes lower and more concentrated around the boundaries of the occupied zones. A further increase in t may cause the centers of the occupied zones to become empty, thus changing the no-crossing points to no-crossing segments. And when t approaches infinity, the actual occupied areas in the zones shrink towards the boundaries, approaching zero but never completely disappearing while t is finite. An allocation with a no-crossing segment in the middle of the two occupied areas cannot be a global optimum, however, since a solution with a larger number of fully occupied smaller zones is clearly more efficient. As a result, as t approaches infinity, the global optimum always entails a mixed allocation.

The following Proposition concerns the decentralization of a local optimum.

Proposition 2: To implement the allocation of a given local optimum solution, a developer (or local government) has only to choose an origin and impose on each unit of land the optimal corrective tax of the supporting market allocation corresponding to the local optimum. Market competition will allow any local optimum to be supported as a competitive equilibrium.

Proof: In this section we have shown that in each decentralized local optimum both industry and housing have different bid rent functions defined over the entire city. We also showed that

industry outbids residents in industrial zones, residents outbid industry in residential zones, and in buffer zones both the industry and the residents do not bid. Since the bid rents reflect the maximum amount each sector is willing to pay under these conditions, the desired local optimum is the only outcome which can result from competition between industry and residents.

7. Global Optimum.¹⁸

We have argued that a particular economy may have multiple local maxima, each corresponding to a qualitatively different spatial configuration. We now develop results concerning the global maximum. We define net surplus functions, each of which is indexed by the integer number of zone pairs and t . Each of these surplus functions is declining in t , and the rate of decline is lower the larger the number of zone pairs. Plotting these net surplus functions against t , their upper envelope indicates how the globally optimal spatial configuration varies with t .

Initially, when $t=0$, the global optimum consists of two AAs (case zero). In this case, $R_l(x_2,0) > 0 > R_h(x_1,0)$. It is possible that any increase in t , even an infinitesimal one, will cause the internal allocation to become the global optimum. In this section, only cases in which separation is the global optimum for at least some positive t are investigated. For simplicity, we assume that for a given t each positive integer m , where $2m$ is a given number of AAs in the solution, has no more than one local optimum¹⁹.

Definition 9: Let $S^*(m,t)$ designate the maximized surplus of a local separated optimum solution with commuting cost $t \geq 0$ and $2m$ AAs, m being a positive integer.

¹⁸ Proofs of the Lemmas and Propositions of this section not presented in the text appear in Appendix C of the internet version of the paper.

¹⁹ In general, there may exist more than one local optimum for a given number of zones. The assumption made here of a single local optimum for each m simplifies the exposition, but it is not difficult to extend the analysis to the more general case.

In what follows we investigate the number of zones in the global optimum by finding the m that maximizes S^* for a given t .

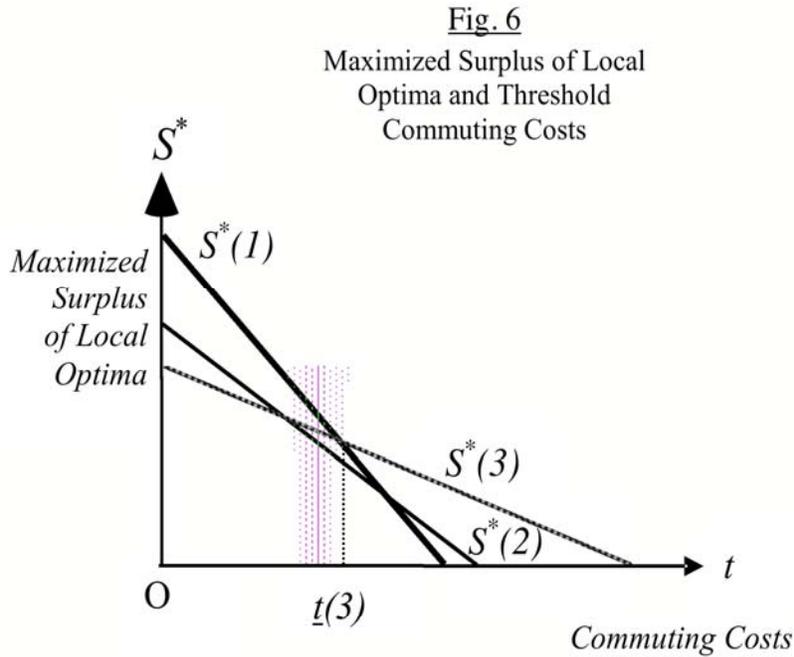
Lemma 9: The following are properties of the function $S^*(m, t)$:

$$(i) \quad \partial S^*(m, t) / \partial t = -\int_0^L |T(x)| dx \leq 0$$

$$(ii) \quad \left| \frac{\partial S^*(m_i; t)}{\partial t} \right| > \left| \frac{\partial S^*(m_j; t)}{\partial t} \right| \text{ if } m_j > m_i$$

and for all t for which both functions are positive and well-defined

$$(iii) \quad S^*(m_i, 0) > S^*(m_j, 0) \text{ for all pairs fulfilling } m_j > m_i.$$



Lemma 9 reveals (see Fig. 6) that: (i) the slope of $S^*(m, t)$ in the (S, t) plane is non-positive and strictly negative as long as m is finite (since $|T(x)|$ in a separated solution is positive almost

everywhere); (ii) in the (S,t) plane at a given t , $S^*(m,t)$ is steeper for smaller m (because $|T(x)|$ attains higher values in larger zones); and (iii) the intercept on the S axis of $S^*(m,t)$ in the (S,t) plane is decreasing with m (since pollution damages increase and therefore the value of $S^*(m,0)$ decreases with the number of zones — see case zero for the proof).

The corollary below now follows directly from Lemma 9:

Corollary 5 In the (S,t) plane, two $S^*(m,t)$ curves with different m 's may intersect in the positive orthant at most once (see Fig. 6).

The above corollary follows directly from (ii) in Lemma 9, which implies that the smaller is m , the steeper is S^* for a given t .

Lemma 10: For a given t , the global optimum allocation is the local optimum allocation for which $m^*(t) = \operatorname{argmax}_m S^*(m,t)$ (note that $m^*(t)$ may be infinite for all t).

The lemma above follows from the definition of S^* and the nature of the global optimum for a given t .

Definition 10: Let $\hat{S}^*(t) = S^*(m^*(t), t)$ be the global optimum value of the surplus as a function of t .

Lemma 10 implies that $\hat{S}^*(t)$ is the upper envelope curve of all the $S^*(m,t)$ in the (S,t) plane. Lemma 9, Corollary 5 and Lemma 10 provide the basis for Proposition 3.

Proposition 3: Let $2m^*(t)$ be the number of AAs in the global optimum solution of the problem with commuting costs t . The function $m^*(t)$, defined in Lemma 10, is a non-decreasing step function of t , $0 \leq t \leq \infty$.

The proof of Proposition 3 is straightforward. Lemma 9 and Corollary 5 imply that two $S^*(m,t)$ curves intersect only once in the (S,t) plane, and the one with lower m intersects the

other from above. Thus, if $m_i < m_j$ and both are in the global solution, m_i will be associated with lower t than m_j .

Corollary 6: If $t_i \begin{matrix} > \\ = \\ < \end{matrix} t_j$, then $m^*(t_i) \begin{matrix} \geq \\ = \\ \leq \end{matrix} m^*(t_j)$.

Let i index the order of m in the global optimum, $i = 1, \dots, I$.²⁰

Corollary 7: For $i \in (1, \dots, I)$, the set of all t 's for which $m_i = m^*(t)$ is a connected segment of the non-negative t axis. The intersection of each consecutive pair of segments is a single point and the union of all I segments of t 's exhausts the half line $t \geq 0$. Finally, with $i > j$, if $m^*(t') = m_i$ and $m^*(t'') = m_j$, then $t' > t''$.

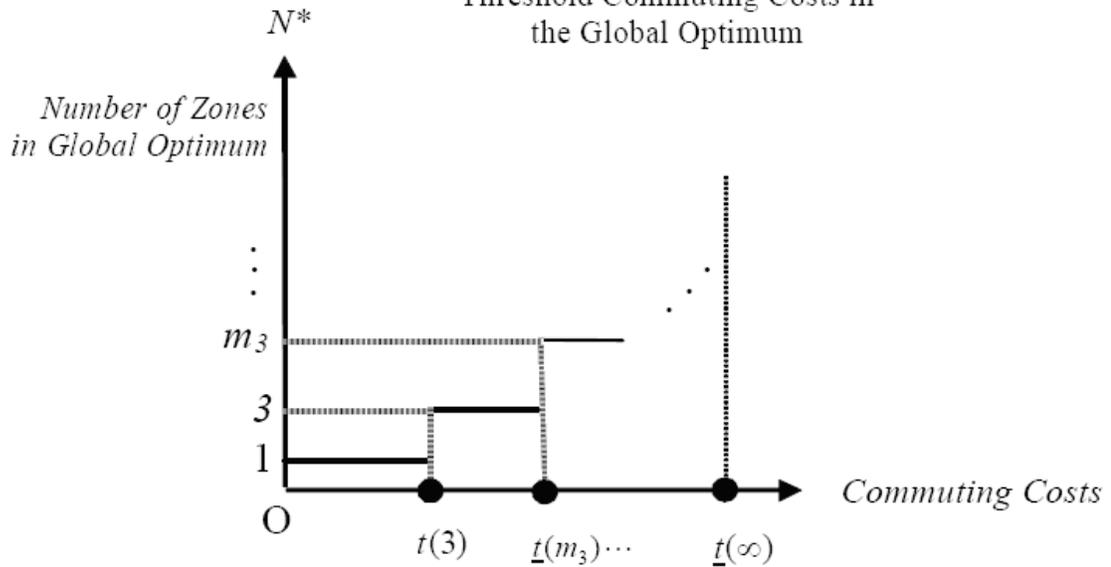
Corollary 7 indicates that the number of zone pairs (m) at the global optimum is a non-decreasing function of commuting costs. I , the number of m 's in the global optimum for some subset of t , can be any positive integer or infinity.

To complete the characterization of the global solution, the concept of a commuting cost threshold is introduced.

Definition 11: Define $\underline{t}(m_i)$, $i = 1, 2, \dots, I$, to be the commuting cost threshold of an allocation with $2m_i$ AAs. $\underline{t}(m_i)$ is the lowest commuting cost in which $2m_i$ zones are the number of AAs in the global optimum, i.e., $\underline{t}(m_i) = \min\{t \mid m^*(t) = m_i\}$.

²⁰ For example, with $I=10$, and $(m_1, \dots, m_{10}) = (1, 4, 6, 9, 11, 14, 15, 19, 23, 25)$, in the global optimum for $t=0$ there are two AAs and the global optimum for infinite t has 50 AAs.

Fig. 7
Optimal Number of Zones and
Threshold Commuting Costs in
the Global Optimum



In Fig. 7, $\underline{t}(m_i)$, $i = 1, 2, \dots, I$ are the jump points of the step function $m^*(t)$, and in Fig. 6 they are the values of t at the intersection points of the S^* curves in the global optimum. Note that (m_1, \dots, m_I) is a set of increasing, not necessarily consecutive, positive integers whose number I may or may not be infinite.²¹

Proposition 4: The Threshold Theorem.

- (i) When t increases and reaches $\underline{t}(m_i)$, the number of zones in the global optimal allocation increases from m_{i-1} to m_i and remains at this level until t reaches $\underline{t}(m_{i+1})$.
- (ii) $m_I = \infty$ always, even when I is finite and $\underline{t}(\infty) \leq \infty$.
- (iii) $m_1 = 1$ and $\underline{t}(1) = 0$.

²¹ In the paper we assume full divisibility of all variables, as is often done in urban economics and other branches of economics. The number of zones is an exception because it must be an integer variable. However, the zone's size still satisfies this assumption. In practice, when residential zones become too small to contain even a single household, the optimum is either the solution with a smaller number of zones, each with a single household, or the mixed solution which is also a local optimum, whichever is more efficient.

Proof: Part (i) follows from the definition of $\underline{t}(\cdot)$ as the lower bound of all t 's having the same number of zones in the global optimum, and $m_i > m_{i-1}$ by construction; (ii) follows from the fact that the mixed solution, which is equivalent to a solution with an infinite number of zones, is always the solution when t becomes sufficiently large to deter commuting, so the value of $S^*(\infty, t)$ is independent of t ; and (iii) follows from the fact that case zero is always the solution when $t = 0$.

The following proposition concerns implementation of optimal corrective taxes, this time in the global optimum.

*Proposition 5: To achieve global efficiency, including the optimal zoning allocation, a developer (or a local government) has only to levy at every location x the corrective tax per unit of land of the global optimum solution. The global optimal corrective tax is the corrective tax $Q(x)$ for the particular local optimum solution that is the global optimum for the given t .*²²

The proof follows directly from Proposition 2.

8. Concluding Remarks.

This paper characterized the social optimum in a spatial economy with pollution, and its decentralization. Its main innovation over previous literature is that land use is completely endogenous. The model is of a ring-shaped economy with residential and industrial land use. Employing a constant-returns-to-scale technology, firms use land and labor to produce a composite good, with emissions as an undesirable by-product. Households supply fixed labor to firms, to which they commute, derive utility from housing and the composite good, and

²² Note that the optimal corrective taxes are very complicated to compute, especially since taxes vary from one location to another. Regulations on emissions, as suggested in Hochman and Ofek [9] may prove more practical to implement.

suffer disutility from the concentration of pollution at their place of residence. The optimal land use pattern is then determined by the tradeoff between pollution and commuting costs.

At one extreme, when transport costs are high, firms and households are completely intermixed; commuting costs are eliminated but pollution concentration at residential locations is high. At the other extreme, when pollution is highly noxious, households crowd together at one end of the ring and firms at the other end, with buffer zones in between; commuting costs are high but pollution concentrations at residential locations low. Between these two extremes is a wide range of possibly optimal land use patterns — different numbers of pairs of residential and industrial zones, perhaps with buffer zones.

Our specification of the mapping from the spatial distribution of pollution emissions to the spatial distribution of pollution concentrations is more general than previous specifications in the literature, though still not completely general. Under our specification, the global optimum can be decentralized with a spatially differentiated tax per unit of industrial land set equal to the additional damages caused by the total emissions from the unit of land, evaluated at the social optimum. A spatially differentiated Pigouvian tax — a tax on emissions — will not decentralize the optimum.

To focus on essentials, our model contained only two forces whose interaction determines the pattern of land use. Commuting costs are an attractive force between residences and firms, pollution a repulsive force. As well, we considered an *ex ante* homogenous space to abstract from edge effects/spatial inhomogeneities. However, our model could be enriched to incorporate other forces affecting optimal land use, such as the scale-economies-agglomeration (SEA) treated in Lucas and Rossi-Hansberg [11]²³.

²³ First, consider the laissez-faire allocation when the scale-economies-agglomeration à la Rossi-Hansberg are introduced into a ring-shaped city without pollution. When transport costs are zero, there will be one industrial and one residential area, with two NC points, one located at the midpoint of each area. The two boundaries between

Future research should investigate how pollution from stationary sources interacts with the other forces which have been identified in the literature as affecting the pattern of land use: returns to scale in production, spatial inhomogeneity, linkages, product variety, spatial interaction, traffic congestion, and automobile pollution. A natural extension for further research is to allow population groups to differ in skills, wages and thus utility levels with distances from polluting firms related to household income levels.

these two areas are straight lines, and the industrial and residential areas are completely separated. The rent function in the industrial zone has a maximum at the NC point in the middle of the zone and monotonically declines as a decreasing rate when moving towards the boundary. At the boundary, the rent function is kinked and becomes constant throughout the residential area. The density of employees follows the industrial bid-rent function, having its maximum in the center and declining towards the boundary. The density of households is constant in the residential area.

When commuting costs (in Lucas and Rossi-Hansberg, they are of the iceberg type) increase, the whole rent function shifts down and in the residential area slopes downward from the boundary to the NC point. With further increases in commuting costs, the boundary line becomes a boundary zone within which industry and households are mixed. As commuting costs are further increased, at some point, the two areas will split into four, and so on.

Now augment the model with two areas so that factories pollute. The rent function and the density will decline in the two areas, especially near the boundary zone, which shrinks and becomes a boundary line once again. The rent function in the residential area becomes flatter. As pollution becomes increasingly severe, rent and density decline even more near the boundary line and may disappear, with a buffer zone appearing. The rent function in the residential zone near the boundary line may increase before it starts declining again. When optimal pollution taxes are levied, rents and density increase and the buffer zones shrink or disappear.

REFERENCES

- [1] I. Allegrini, F. De Santis (Eds), *Urban air pollution: Monitoring and control strategies*, [Proceedings of October 1994 NATO Advanced Research Workshop], Springer, New York, 1996.
- [2] R. Arnott, Optimal city size in a spatial economy, *Journal of Urban Economics* 6 (1979) 65-89.
- [3] K. Arrow, B. Bolin, R. Costanza, P. Dasgupta, C. Folke, C.S. Holling, B-O. Jansson, S. Levin, K-G. Maler, C. Penings and D. Pimentel, Economic growth, carrying capacity, and the environment, *Science* 268 (1995) 520-521.
- [4] W.G. Baumol, W. E. Oates, *The theory of environmental policy*, 2nd ed., Cambridge University Press, Cambridge, 1988.
- [5] J.V. Henderson, Externalities in a spatial context, *Journal of Public Economics* 7 (1977) 89-110.
- [6] J.V. Henderson, *Economic theory and the cities*, 2nd ed., Academic Press, New York, 1985.
- [7] J.V. Henderson, Effects of air quality regulation, *American Economic Review* 86 (1996) 789-813.
- [8] O. Hochman, Land rents, optimal taxation, and local fiscal independence in an economy with local public goods, *Journal of Public Economics* 15 (1981) 59-85.
- [9] O. Hochman, H. Ofek, A theory on the behavior of municipal governments: The case of internalizing pollution externalities, *Journal of Urban Economics* 6 (1979) 416-431.
- [10] R.E. Lucas, Jr., Externalities and cities, *Review of Economic Dynamics* 4 (2001) 245-267.
- [11] R.E. Lucas, Jr., E. Rossi-Hansberg, On the internal structure of cities, *Econometrica* 70 (4) (2002) 1445-1476.
- [12] A. Mas-Colell, M. D. Winston, J. R. Green . *Microeconomic Theory*, Oxford University Press, Oxford, 1995.
- [13] J.J. McCluskey and G.C. Rausser, Stigmatized asset value: Is it temporary or long-term? *Review of Economics and Statistics* 85 (2003) 276-285.
- [14] J.J. McCluskey and G.C. Rausser, Hazardous waste sites and housing appreciation rates, *Journal of Environmental Economics and Management* 45 (2003) 166-176.
- [15] National Technical Information Service (NTIS), US Department of Commerce, "Air Quality Models", 1997.
- [16] M.A. Polinsky, Strict liability vs. negligence in a market setting, *American Economic Review* 70 (1980) 363-367.
- [17] G.C. Rausser, H. Lapan, Natural resources, goods, bads, and alternative institutional frameworks, *Resources and Energy* 2 (1979) 293-324.
- [18] E. Rossi-Hansberg, Optimal urban land use and zoning, *Review of Economic Dynamics* 7 (2004) 69-106.
- [19] D.F. Spulber, "Effluent regulation and long run optimality, *Journal of Environmental Economics and Management* 12 (1985) 103-116.
- [20] T.H. Tietenberg, Derived demand rules for pollution control in a general equilibrium space economy, *Journal of Environmental Economics and Management* 1 (1974) 3-16.
- [21] T.H. Tietenberg, Comment, *American Economic Review* 64 (1974) 461-462.

- [22] T.H. Tietenberg, Spatially differentiated air pollutant emission charges: An economic and legal analysis, *Land Economics* 54 (1978) 265-277.

Appendix A (for Publication)

Derivation of the First-Order (Kuhn-Tucker) Conditions for a Local Optimum

Let L be the Lagrangean of the model, where the variables, constraints and shadow prices are as defined in sections 2 – 4 of the paper.

$$\begin{aligned}
 (A\ 1) \quad L = & \int_0^L [a(x)f(n(x), e(x)) - (\frac{b(x)}{h(x)}z(x) + |T(x)|t)]dx + \int_0^L \lambda(x) [U(h(x), z(x), c(x)) - U_0] dx \\
 & + \int_0^L \eta(x) \left\{ c(x) - \int_{x-L/2}^x a(y)D^+[e(y), x-y]dy - \int_x^{x+L/2} a(y)D^-[e(y), y-x]dy \right\} dx + \\
 & + \int_0^L \zeta(x) \left\{ T(x) - \int_0^x \left[\frac{b(y)}{h(y)} - a(y)n(y) \right] dy \right\} dx - \int_0^L [\rho(x)(a(x) + b(x) - 1) - \gamma(x)a(x) - \mu(x)b(x)] dx.
 \end{aligned}$$

It should be noted that $\zeta(x)$, the shadow price of the commuting constraint, is different from $\Psi(x)$, the co-state of $T(x)$ as defined in the text. We elaborate below on the relation between the two. The necessary conditions are as follows.^{24, 25}

$$(A\ 2) \quad n(x) \quad a(x)[f_1(x) + \int_x^L \zeta(y)dy] = 0$$

$$(A\ 3) \quad e(x) \quad a(x)[f_2(x) - \int_x^{x+L/2} \eta(y)D_1^+(e(x), y-x)dy - \int_{x-L/2}^x \eta(y)D_1^-(e(x), x-y)dy] = 0$$

²⁴ The variable of differentiation is noted on the left-hand side of each equation. Note that a function with a number as a subscript indicates derivations of the function with respect to the variable of the order of the subscript.

²⁵ With a slight abuse of notation, $f(x) \equiv f(n(x), e(x))$, etc.

$$(A 4) \quad a(x) \quad f(x) - \left[\int_x^{x+L/2} \eta(y) D^+(e(x), y-x) dy + \right. \\ \left. + \int_{x-L/2}^x \eta(y) D^-(e(x), x-y) dy \right] + n(x) \int_x^L \zeta(y) dy - \rho(x) + \gamma(x) = 0$$

$$(A 5) \quad h(x) \quad \frac{b(x)}{h(x)^2} z(x) + \lambda(x) U_h(x) + \frac{b(x)}{h(x)^2} \int_x^L \zeta(y) dy = 0$$

$$(A 6) \quad b(x) \quad -\frac{z(x)}{h(x)} - \frac{1}{h(x)} \int_x^L \zeta(t) dt - \rho(x) + \mu(x) = 0$$

$$(A 7) \quad z(x) \quad -\frac{b(x)}{h(x)} + \lambda(x) U_z(x) = 0$$

Since $|T(x)| = [\text{sign}(T(x))]T(x)$, differentiation of L with respect to $T(x)$ yields:²⁶

$$(A 8) \quad T(x) \quad -[\text{sign}T(x)]t + \zeta(x) = 0$$

$$(A 9) \quad c(x) \quad \lambda(x) U_c(x) + \eta(x) = 0$$

Define the co-state of $T(x)$ to be $\Psi(x)$:

$$(A 10) \quad \Psi(x) \equiv -\int_x^L \zeta(y) dy$$

Then

$$(A 10') \quad \Psi'(x) = \zeta(x) = [\text{sign}T(x)]t.$$

²⁶ $\text{Sign}(c)$ is differentiable and its derivative equals zero everywhere except at $x = 0$ where the derivative is not defined. The function sign enables differentiation of $|T(x)|$ everywhere except at $x = 0$.

Substituting out $\zeta(x)$ from the above equations using (A 10), and then eliminating $\lambda(x)$ from the equations by substituting from (A 9), we obtain the necessary conditions as specified in the text.

Appendix B (not for publication)

Proofs and Derivations

Proof of Lemma 1:

Consider the boundary points of the zones in an optimal allocation, e.g. x_1 and x_2 , $x_1 < x_2$, as depicted in Fig.3 of the paper. The bid-rent rule implies that at such boundary points the slopes with respect to distance (designated by a dot over the function) of the bid-rent functions must fulfill $\dot{R}_I(x_2) \geq \dot{R}_H(x_1)$, otherwise the allocation is not optimal. Suppose, only for the sake of proving a contradiction, that $\dot{R}_I(x_2) < \dot{R}_H(x_1)$, then industry outbids housing in the residential zone and housing outbids industry in the industrial zone. Switch the location of a single household at x_1 and a firm occupying the same amount of land at x_2 . Since industry outbids housing in the residential zone and *vice versa* in the industrial zone, this transfer increases total rents. It also increases total pollution damages since it shortens distances between polluters and pollutees. Total pollution taxes therefore increase as well. According to the Henry George rule ($\int_0^L [r(x) + Q(x)] dx = S$), total rents and optimal taxes together constitute the goal function. Since the switch increases the goal function, the initial allocation is not optimal; a contradiction. Hence $\dot{R}_I(x_2) \geq \dot{R}_H(x_1)$.

Calculating the Spatial Derivatives of the Bid Rent Functions

We differentiate the bid-rent functions with respect to distance x . First we differentiate Eq. (17) at locations where $a(x) > 0$ and substitute (7), (8), (13), and (14b) into the result to yield:

$$\begin{aligned}
 \dot{R}_I &= f_2 \dot{a} - n \dot{a} - \dot{Q} \\
 \text{(B1)} \quad &= -\text{sign}[T(x)] \{ n + \{ \eta(x + L/2) [D^+(e(x), L/2) - D^-(e(x), L/2)] \\
 &+ \int_x^{x+L/2} \eta(y) D_2^-(e(x), y - x) dy \} - \int_{x-L/2}^x \eta(y) D_2^-(e(x), x - y) dy
 \end{aligned}$$

Use has also been made of the continuity assumption $D^+(e, 0) = D^-(e, 0)$ and

$\eta(x + L/2) = \eta(x - L/2)$ (because $(x + L/2)$ and $(x - L/2)$ are the same point.)

By differentiating (1) with respect to x and substituting $R_h(x) = \frac{U_h(x)}{U_z(x)}$ into the result, we get the

expression $R_h \frac{d}{dx} \left(\frac{U_c}{U_z} \right) = R_h \frac{d}{dx} \left(\frac{U_c}{U_z} \right) - h\eta = 0$, where the second equality is obtained by substitution of (12) with $b(x)=1$ into the first equality. We then differentiate (17) and substitute

$R_h(x) = \frac{U_h(x)}{U_z(x)}$ into the result to obtain the first equality:

$$(B2a) \quad R_h = \frac{U_c}{h} - \eta = \frac{1}{h(x)} [\text{sign}[T(x)] + \frac{U_c(x)}{U_z(x)}]$$

The second equality is obtained after substitution of (12) and (8) into the previous term.

Differentiating (2) with respect to x yields

$$(B2b) \quad \frac{d}{dx} \left(\frac{U_c}{U_z} \right) = \int_{x-L/2}^x a(y) D_2^+[e(y), x-y] dy - \int_x^{x+L/2} a(y) D_2^-[e(y), y-x] dy \\ - a(x+L/2) [D^+(e(x-L/2), L/2) - D^-(e(x+L/2), L/2)]$$

where once more we have made use of the facts that $x + L/2 = x - L/2$ and that

$D^+(e, 0) = D^-(e, 0)$. Substituting (B2b) into (B2a) yields the desired expression for R_h .

Buffer Zones and Boundary Conditions In a Two AA's Case

In Fig.3 of the paper the following chain of inequalities holds:

$$(B3) \quad 0 < x_1 \leq x_2 < x_3 \leq x_0 < L$$

where x_i , $i = 0,1,2,3$ are the boundaries of the different zones. A necessary condition for optimal

boundaries is $R_f(x_2) = R_h(x_1)$ and $R_f(x_3) = R_h(x_0)$. If $x_1 = x_2$ and $x_3 = x_0$, buffer zones do not

exist and there is only a residential zone and an industrial zone. However, if in the optimum only strong inequalities hold between the boundaries specified in (B3), the solution also includes buffer zones.

A segment of the ring is a buffer zone if for all x of the segment, $R_I(x) \leq 0$, $R_h(x) < 0$ (by saying that $R_h(x) < 0$ we mean that if $\underline{z}(x)$ fulfills $U(\infty, \underline{z}(x), c(x)) = u_0$, then

$\Psi(x) - \underline{z}(x) (=h(x)R_h(x)) < 0$). At the boundary of a buffer zone and an industrial zone $R_I(x) = 0$, and at the boundary of a residential zone and a buffer zone $R_h(x) = 0$. Additional necessary conditions for the general zoning case are:

$$(B4) \quad \begin{aligned} &R_I(x) < R_h(x) > 0 \text{ for } x_0 \leq x \leq L; \text{ and } 0 \leq x \leq x_1 \\ &0 < R_I(x) > R_h(x) \text{ for } x_2 \leq x \leq x_3 \end{aligned}$$

and the following conditions are specific to buffer zones.

$$(B5) \quad \begin{aligned} &R_h(x) \leq 0, R_I(x) \leq 0 \text{ for } x_1 \leq x \leq x_2 \text{ and } x_3 \leq x \leq x_0 \\ &R_h(x_0) = R_h(x_1) = R_I(x_2) = R_I(x_3) = 0 \text{ and} \\ &R_h(x_2) \leq 0, R_h(x_3) \leq 0 R_I(x_0) \leq 0 R_I(x_1) \leq 0 \end{aligned}$$

Since we use the assumption $D^+(e, y) = D^-(e, y)$ (Assumption 1) and disregard problems of indivisibility and multiple optima, there is complete symmetry between north and south. That is OO' , the line through the origin and the second NC point, divides the circle into two halves and serves as an axis of symmetry between two mirror images. Thus $x_0 + x_1 = L = x_2 + x_3$; see Fig. 3, which depicts a case where this assumption holds.

Application of the Bid-Rent Rule to Case Zero:

Consider first the southern boundary of the residential zone, x_0^0 . It is either an intersection point of the two bid-rent curves and there is no buffer zone south of the residential zone, or

$R_h(x_0^0) = 0$, $R_l(x_0^0) < 0$ and an empty buffer zone exists between the residential and industrial zones (see Fig. 3). Since in case zero $t=0$, the first term in the RHS of (B2a) disappears. The second term there depends on $\mathcal{R}(x_0^0)$ given in (B2b). The assumption $D^+(e, y) = D^-(e, y) = D(e, y)$ implies that the last term of $\mathcal{R}(x_0^0)$ is zero. Upon substitution of $x = x_0^0$ and the above symmetry of the dispersion functions assumption into (B2b), we obtain

$$(B6) \quad \mathcal{R}(x_0^0) = \int_{x_0^0 - L/2}^{x_0^0} a(y) D_2[e(y), x_0^0 - y] dy - \int_{x_0^0}^{x_0^0 + L/2} a(y) D_2[e(y), y - x_0^0] dy .$$

There are four cases:

1. No buffer zone ($x_1^0 = x_2^0$, and $x_3^0 = x_0^0$), and $x_1^0 + L - x_0^0 > x_3^0 - x_2^0$
2. No buffer zone, and $x_1^0 + L - x_0^0 < x_3^0 - x_2^0$
3. Buffer zone, and $x_1^0 + L - x_0^0 > x_3^0 - x_2^0$
4. Buffer zone, and $x_1^0 + L - x_0^0 < x_3^0 - x_2^0$

The condition $x_1^0 + L - x_0^0 > x_3^0 - x_2^0$ is that the residential area exceeds the industrial area and *vice versa* when the inequality is reversed. In all cases, the first term on the RHS of (B6) is negative since $a(y) > 0$ for at least a subset of $\left(x_0^0 - \frac{L}{2}, x_0^0\right)$ and zero elsewhere, and since $D_2[\cdot] < 0$.

In cases I and III, the second term on the RHS of (B6) does not exist since the boundaries of the integral are an empty set. Since $\eta(x_0^0) > 0$ (from (12)), from (B2a) in Case I,

$$\mathcal{R}_h(x_3^0) = \mathcal{R}_h(x_0^0) > 0, \text{ and } R_h(x_3^0) = R_h(x_0^0) > 0 \text{ while in Case III,}$$

$$\mathcal{R}_h(x_0^0) > 0, \text{ , } R_h(x_0^0) = 0 > R_h(x_3^0)$$

Now consider the other two cases. The second term of (B6) is derived from the contribution to concentrations at x_0^0 of the segment $\left(x_2^0, x_0^0 + \frac{L}{2}\right)$ of the industrial area. Mirror symmetry

implies that the segment $(x_1^0 + L/2, x_3^0)$ contributes to x_1^0 the same amount of concentrations.

Thus,

$$\int_{x_2^0}^{x_0^0 + L/2} a(y)D^- [e(y), y - x_0^0]dy = \int_{x_1^0 + L/2}^{x_3^0} a(y)D^+ [e(y), x_1^0 + L - y]dy$$

A shift of x_0^0 in the positive direction decreases the left-hand side of the above equality the same way a shift of x_1^0 in the negative direction effects the right-hand side. Thus, the second term on the right-hand side of (B6) can be written as

$$ST \stackrel{def}{=} \int_{x_0^0}^{x_0^0 + L/2} a(y)D_2^- [e(y), y - x_0^0]dy = \int_{x_2^0}^{x_0^0 + L/2} a(y)D_2^- [e(y), y - x_0^0]dy = \int_{x_1^0 + L/2}^{x_3^0} a(y)D_2^+ [e(y), x_1^0 + L - y]dy$$

The first equality above uses $a(y) = 0$ for $y \in (x_0^0, x_2^0 + L)$, note that x_0^0 and $x_0^0 + L$ designate the same point. The second inequality above follows the discussion above.

The first term on the right-hand side of (B6) can be broken down as

$$FT \stackrel{def}{=} \int_{x_0^0 - L/2}^{x_1^0 + L/2} a(y)D_2^+ [e(y), x_0^0 - y]dy + \int_{x_1^0 + L/2}^{x_3^0} a(y)D_2^+ [e(y), x_0^0 - y]dy$$

Subtracting ST from FT yields

$$\mathfrak{E}(x_0^0) = \int_{x_0^0 - L/2}^{x_1^0 + L/2} a(y)D_2^+ [e(y), x_0^0 - y]dy + \int_{x_1^0 + L/2}^{x_3^0} a(y)\{D_2^+ [e(y), x_0^0 - y] - D_2^+ [e(y), x_1^0 + L - y]\}dy$$

When $D_{22} > 0$ the second integral in \mathfrak{E} above is negative as well as the first and therefore the whole expression. When $D_{22} < 0$, however, the second term is positive and therefore the sign of $\mathfrak{E}(x_0^0)$ may theoretically be positive. Note, however, that when $D_{22} < 0$, both $D_2(e, y)$ and $D(e, y)$ disappear for relatively small y , so that the second term of the second integral in $\mathfrak{E}(x_0^0)$ above may be non

existent. If however, if the term does exist, $\mathcal{R}_h(x_0^0)$ can never be positive since then $\mathcal{R}_h(x_0^0)$ is negative which means that the housing bid rent in the residential zone is negative and outside the residential zone it is positive—a contradiction. Therefore either $\mathcal{R}_h(x_0^0)$ is negative or the internal mixed solution holds.

Consider now $\mathcal{R}_I(x_3^0)$. By substituting $t=0$ and $D^+ = D^-$ the first line in the RHS of (B1) disappears. Equation (11) with $b(x)=0$ outside the residential zone implies $\eta(y) = 0$, for $x_1^0 < y < x_0^0$. Consequently, the integrals in the second line of (B1) reduce to the segment $x_0^0 \leq y \leq x_1^0 + L$ only. There are two cases to consider. In the first, the industrial zone is larger than the residential zone; in the second, the opposite is true. When the industrial zone is larger than the residential zone, $\int_{x_3^0 - L/2}^{x_1^0} \eta(y) D_2^-(e(x_3^0), x_3^0 - y) dy = 0$ (none of the pollution from x_0^3 that travels counter-clockwise reaches the residential zone). As a result

$$\mathcal{R}_I(x_3^0) = \int_{x_0^0}^{x_1^0 + L} \eta(y) D_2^+(e(x_3^0), y - x_3^0) dy < 0.$$

When the industrial zone is smaller than the residential zone, the second term of (B1) for x_3^0 is

$$\int_{x_3^0 - L/2}^{x_1^0} \eta(y) D_2^-[e(x_3^0), x_3^0 - y] dy = \int_{x_0^0}^{x_2^0 + L/2} \eta(y) D_2^+[e(x_2^0), y - x_2^0] dy$$
 where the equality in the expression

follows by mirror symmetry. Substituting the above into the remaining second line of (B1) yields

$$\mathcal{R}_I(x_3^0) = \int_{x_0^0}^{x_2^0 + L/2} \eta(y) \{D_2^+[e(x_3^0), y - x_3^0] - D_2^-[e(x_3^0), y - x_2^0]\} dy + \int_{x_2^0 + L/2}^{x_1^0} \eta(y) D_2^+[e(x_3^0), y - x_3^0] dy.$$

When $D_{22} < 0$ the second term in the first integral, as before may be non existent. If, however, the first integral is positive, $\mathcal{R}_I(x_3^0)$ cannot be non-negative because then the industrial bid rent is negative in the industrial zone and positive outside the zone. In that case a zoning solution is impossible and the internal mixed solution holds.

Since x_0^0 and x_3^0 are boundary points they are also points of intersection of bid-rent curves and as such satisfy the bid-rent rule. Our results imply that north of x_0^0 residents outbid industry and south of x_3^0 industry outbids residents. When there are no buffer zones $x_0^0 = x_3^0$, and $R_h(x_0^0) = R_I(x_3^0) = r(x_0^0 = x_3^0) \geq 0$. The rent function at this boundary is not differentiable and has a positive derivative in the positive direction of the x -axis and a negative derivative from the negative direction. Where are buffer zones $r(x_3^0) = R_I(x_3^0) = 0 > R_h(x_3^0)$.

In complete mirror symmetry to the case of x_0^0 and x_3^0 , we can obtain expressions for the slopes of the bid rent functions in x_1^0 and x_2^0 . The results imply that south of x_1^0 residents outbid industry and north of x_2^0 industry outbids residents. When there are no buffer zones $x_1^0 = x_2^0$, and $R_h(x_1^0) = R_I(x_2^0) = r(x_1^0 = x_2^0) \geq 0$. The rent function at this boundary is not differentiable and has a positive derivative in the negative direction of the x -axis and a negative derivative in the positive direction. Where are buffer zones $r(x_2^0) = R_I(x_2^0) = 0 > R_h(x_2^0)$

Appendix C (not for publication)

Further Proofs

Proof of Lemma 4

The commuting cost parameter t appears in the necessary conditions explicitly only in the expression of $\psi(x)$ (Eq. (8)). We already established that choosing the NC point of the residential zone as the origin makes $T(x)$ positive clockwise of the origin up to the second NC point at $L/2$, from which point on $\text{sign}(T(x))$ is negative up to $x=L$. Substituting $+1$ and -1 for $\text{sign}(T(x))$ in the appropriate places in (10) yields, $\psi(x) = \begin{cases} t & \text{for } 0 < x < L/2 \\ -t & \text{for } L/2 < x < L \end{cases}$ which upon integration yields (21). From (7) we know that $\psi(x)$ in the industrial zone is equal to the wage at x and from (11), the budget constraint, that in the residential zone, $\psi(x)$ is the LNE — household earned income after commuting costs have been deducted. From (21) it is clear that the highest wage is at $x=L/2$ (O in Fig. 3). In the residential zone, $\psi(x)$ is independent of work location and depends only on place of residence.

Proof of Corollary 3

Differentiating (21) with respect to t yields

$$(C1) \quad \frac{\partial \psi(x)}{\partial t} = \frac{\partial \psi(x_2; t)}{\partial t} + (x - x_2), \quad \text{for } 0 \leq x \leq L/2,$$

where x_2 is the boundary of the industrial zone and the term $\partial \psi(x_2; t) / \partial t$ represents the change in the wage rate there when t changes. $\partial \psi(x_2; t) / \partial t$ is essentially a shift parameter since it is independent of location.

Proof of Lemma 5

The generalized Henry George rule (see Arnott [2]) implies that the net city surplus satisfies $S = \int_0^L r(x)dx + \int_0^L Q(x)dx$ (see also Hochman and Ofek [9]) where $r(x)$ equals $R_h(x)$ in the residential zone and $R_I(x)$ at the industrial zone. The Envelope Theorem therefore implies

$$\frac{\partial}{\partial \chi_i} \int_0^L r(x)dx = -\frac{\partial}{\partial \chi_i} \int_0^L Q(x)dx \text{ where } \chi_i \text{ is any control variable or shadow price, except for } a(x)$$

and $b(x)$ whose derivatives are everywhere zero except at the boundary points where they are discontinuous. Of these variables only $\eta(y)$ and $e(x)$ appear in $Q(x)$. Since in the residential zone where $Q(x)$ is zero, $r(x) = R_h(x)$ we have $\partial R_h(x)/\partial \chi_i(x) = -\partial Q(x)/\partial \chi_i(x)=0$, and in the industrial zone where $r(x) = R_I(x)$, the non- zero differentials are

$$\partial R_I(x)/\partial \eta(y) = -\partial Q(x)/\partial \eta(y) \text{ and } \partial R_I(x)/\partial e(x) = -\partial Q(x)/\partial e(x) . \text{ However we observe in (12)}$$

that $\eta(y)$ is independent of t , and from the rest of the production equations so is $e(x)$ (actually $\Psi(x)$ is the only variable which depends on t and it does not appear in Q). From (14b) Q is also directly independent of t , hence $\partial Q/\partial t=0$. Consequently by differentiating (16) with respect to t we get

$$(C 2) \quad dR_I(x)/dt = \partial R_I(x)/\partial t = -n(x) \partial \Psi(x)/\partial t .$$

Similarly, by differentiating (17) we obtain

$$(C 3) \quad dR_h(x)/dt = \partial R_h(x)/\partial t = \frac{1}{h(x)} ((\partial \Psi(x)/\partial t))$$

Proof of Lemma 6

Differentiating (C 2) with respect to x yields

$$\begin{aligned}
\partial R(x)/\partial t &= -n(x)(\partial \Phi(x)/\partial t) - R(x)(\partial \psi(x)/\partial t) \\
\text{(C 4)} \quad &= -n(x)\text{sign}(T(x)) + \frac{R(x)}{n(x)}(\partial R_1(x)/\partial t).
\end{aligned}$$

And differentiating (C 3) gives

$$\begin{aligned}
\partial R(x)/\partial t &= \frac{1}{h(x)} \left(\partial \Phi(x)/\partial t - \frac{R(x)}{h(x)} (\partial \psi(x)/\partial t) \right) \\
\text{(C 5)} \quad &= \frac{\text{sign}(T(x)) - \frac{R(x)}{h(x)} (\partial R_h(x)/\partial t)}{h(x)}
\end{aligned}$$

We are still looking at the northern hemisphere in the two AA's case. Consider the RHS of the second equality in (C4). The first term is always negative. The second term has the sign of $\partial R_1(x)/\partial t$ which is negative except in Pattern B near the boundary where R_1 increases with t and with it the whole term. Thus, the RHS of (C4) may or may not increase at such locations, depending on the relative size of the two terms.

A similar argument holds for (C5) but with an opposite sign.

Proof of Lemma 9:

We obtain (i) in the lemma by differentiating (6) with respect to t and utilizing the Envelope Theorem. When m approaches infinity, AA's become infinitesimal and therefore commuting costs approach zero. The solution then approaches the mixed solution of case two. An informal²⁷ proof of (ii) is as follows: An increase in m implies shorter commuting distances and shorter distances for pollution dispersion before concentrations reach residential land use. This implies that in two allocations with the same t , overall commuting costs are lower and overall concentration levels higher in the allocation with more zones. Since an increase in t is costlier, when commuting distances are longer and therefore causes a larger reduction in the surplus, the

²⁷ A formal proof of these statements can be devised along the following lines. Consider (i) in the Lemma. Increasing the number of zones while keeping t constant, shortens commuting distances thus the highest values of $|T(x)|$ are replaced with lower absolute values. Accordingly the total value of the integral is reduced.

function $S^*(m,t)$ is steeper (has a more negative slope) with respect to t the smaller is m . To prove (iii): The smaller is m , the larger is $S^*(m,0)$, because commuting costs are zero for all m while concentrations are lower when m is smaller.