Available online at www.sciencedirect.com
SCIENCE DIRECT
Review of Economic Dynamics 7 (2004) 69–106
Review of Economic Dynamics
www.elsevier.com/locate/red

Optimal urban land use and zoning

Esteban Rossi-Hansberg

Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, USA

Received 6 November 2002; revised 14 May 2003

Abstract

The paper studies the optimal distribution of business and residential land in a circular city. Once the optimum is characterized, we analyze the effect of changes in commuting costs and externality parameters. We also propose policies like labor subsidies, land taxes and zoning restrictions that can implement the efficient allocation as an equilibrium, or close the gap between the optimal and equilibrium allocations. The results show that business land is more concentrated at the center of the city in the optimum and that higher commuting costs increase the difference between optimal and equilibrium allocations.

JEL classification: D62; H23; R12; R14

Keywords: Land use; Spatial production externalities; Commuting costs; Zoning

1. Introduction

All economic theories of cities have to take a stand on the key force that agglomerates firms and agents in a compact geographical area we call a city. Production externalities are one of the leading agglomeration mechanisms used in the literature. They have formed the basis of many urban models that have undoubtedly contributed to our understanding of the structure of cities. The equilibrium allocation in these models is not efficient, so society can gain from urban policies that distort equilibrium land use structure. City governments have used zoning policies as well as different types of subsidies and taxes to take advantage of these potential gains. The practical intent to materialize these benefits highlights the need for a conceptual framework to analyze and evaluate urban policies. In this paper we

E-mail address: erossi@stanford.edu.

1094-2025/$ – see front matter
doi:10.1016/S1094-2025(03)00056-5

characterize the optimal distribution of urban land, and use this analysis to design policies that improve the efficiency of equilibrium allocations.

Our focus is on allocations that maximize total production net of total consumption in the city. Or, equivalently, total rents. The theory determines the distribution of business and residential land together with employment and residential densities at all locations in the city. The two main forces that determine the optimal allocation of land in our theory are spatial production externalities and commuting costs. The first force agglomerates firms in clusters. The second disperses producers, so that workers can live close to their work places. The trade-off between these two forces leads to optimal allocations that differ depending on the size of commuting costs and the form and degree of external effects. Low transport costs lead to a central business district with high employment density, while higher transport costs result in low employment density or residential areas at the center. If spillovers decline faster with distance, producers concentrate closer to the center of the city, and business areas become smaller but with higher employment densities.

We use the model to analyze potential policies to improve efficiency in the city. We find that location specific labor subsidies are sufficient to implement the optimal allocation as an equilibrium. In contrast, zoning restrictions cannot implement the optimal allocation, but they can improve the efficiency of the equilibrium allocation, especially when commuting costs are high.

Our results show that business land and employment are more concentrated in the optimal than in the equilibrium allocation. In equilibrium, the higher commuting costs the more mixed areas in the city.[1] In the optimum, mixed areas will disappear and the land use structure will resemble a Mills city (a central business center surrounded by residential areas, following Mills, 1967) for reasonably high commuting cost. Even higher commuting cost will yield multiple business or residential sectors. We will prove that it is *never* optimal to have mixed areas in the city.

The higher concentration of the optimal employment density is illustrated in Fig. 1, where we have plotted the Lorenz Curve of employment concentration in New York City.[2] We find parameters so that the equilibrium Lorenz Curve fits the data as well as possible. We then use the same parameter configuration to compute the Lorenz Curve for the optimum. The figure suggests that employment in New York City should be much more concentrated.

Another implication of the model is that the difference between the Pareto Optimal allocation and the equilibrium allocation is smaller as the cost of commuting decreases. In equilibrium, when commuting costs are low, the allocation resembles a Mills city. The same is true in the optimal allocation, but with more concentrated business areas. In contrast, high commuting costs imply different optimal and equilibrium land use structures. Of course, the optimal employment density is greater than the equilibrium density, since the effect on other firms of one firm employing more workers is internalized.

[1] A mixed area is defined as a section of the city where workers live at, or very close to, their work places.

[2] The data comes from the 1992 Economic Census Zip Code Statistics. We calculate total employment in each Zip code by adding employment in the service, retail and manufacturing sectors. Employment data for the manufacturing sector is only given as a range; we use the middle of the range as the employment observation in this sector.

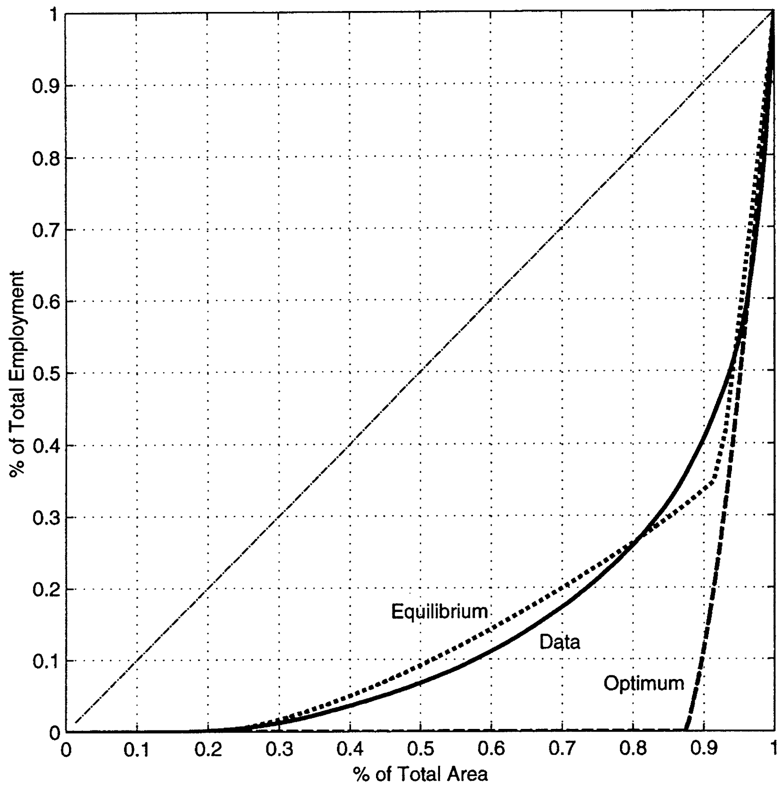

Fig. 1. Lorenz curves of employment in New York city.

There are two sources of gains from reducing commuting costs: a direct effect of agents paying less commuting costs per mile commuted and an indirect effect via the concentration of business areas. Policies like road construction, improving public transportation and parking space provision should consider this additional gain. The indirect effect has been mostly forgotten in the literature since we lacked a framework to study land use structures within cities. The need for this framework is also evident in the design of zoning restrictions. This paper tries to provide that framework.

Borukhov and Hochman (1977), Dixit (1973) and Hochman and Pines (1982) study the problem of optimal allocation of workers within a city. These three papers obtain densities of workers and households, but they impose the land use structure by assumption. Other papers like Stull (1974) and Helpman and Pines (1977) impose a land use structure in the form of a business sector in the middle of the city surrounded by a residential area, but obtain the optimal size of the business sector endogenously. However, these papers use the assumption that agents commute to the center of the city instead of their actual work places. This assumption turns out to be an important simplifying assumption since it separates the problem of location of firms from the problem of location of residences. Fujita and Thisse (2002) analyze one dimensional model of a city with constant employment and residential densities. We relax both of these assumptions. The model is a two dimensional model since densities and externalities take into account that economic activity locates in a circular plane, a disc. However, we consider only symmetric allocations. None of these papers designs policy instruments to improve efficiency.

Other papers like Fujita and Ogawa (1982), Berliant et al. (2002) and Lucas and Rossi-Hansberg (2002) (from now on LRH) determine only the equilibrium internal structure of the city and not the optimum. The difference is crucial, as pointed out above, since in these models agglomeration is generated using production externalities.[3] In particular, we will use the same production externalities as in Lucas (2001). Fujita and Ogawa (1982) and Berliant et al. (2002) assume a one dimensional city and constant densities of firms and residents. Although the latter article does introduce location specific capital. LRH relaxes both of these assumptions. We use the results in LRH extensively. They constitute the equilibrium allocation to which we compare the results of our model. Although the basic frameworks are identical, the way in which the equilibrium is calculated in LRH differs substantially from the way in which the optimum is calculated in this paper. In LRH, wage no arbitrage conditions are derived and they constitute an important part of the solution algorithm. For the optimum these conditions are not satisfied.

The paper is organized as follows. We present the model and study some of the necessary conditions for an optimum in the next section. In Section 3 we provide a solution algorithm and prove the existence of a solution to the optimization problem. Section 4 presents some analytical results and the optimal labor subsidies and zoning restrictions. Some numerical exercises that help us illustrate predictions of the theory are presented in Section 5. Section 7 concludes. In Appendix A we prove the results presented in the text.

2. The model

We model a circular city where a single good is produced using land and labor. People consume goods and residential land. There is a production externality that depends on where other firms locate. Workers allocate one unit of time between working and commuting to their workplaces. Commuting is costly in time.

The objective is to maximize the value of land in the city. That is, we want to maximize total production net of residents' consumption (from now on 'net output'). We will consider only *symmetric allocations*, where everything depends *only* on the distance from the center.[4] Let $n(r)$ be the number of workers per unit of business land at a distance r from the center (from now on 'location r'), $N(r)$ the number of residents per unit of residential land at location r, $\theta(r)$ the proportion of land used for business purposes and $c(r)$ and $\ell(r)$ consumption and units of land per person.

[3] See Anas et al. (1998) and Fujita et al. (1999) for other ways to generate agglomeration in urban economics models.

[4] The assumption of symmetry is a strong assumption. Casual observation of metropolitan areas is enough to realize that people do not commute only to and away from the center to go to work. Nevertheless, the assumption is an important simplifying assumption of our theory. Without it, we can not reduce the external effect to a one dimensional function and hence the stock of unhoused workers cannot be constructed sequentially as we will do below.

Firms produce using land and labor. At location r, production per unit of land of a firm that employs $n(r)$ workers per unit of land is given by

$$g\big(z(r)\big)f\big(n(r)\big),$$

where $z(r)$ denotes the productivity of a firm at location r. Consider a circular city with radius S. Productivity at each location is determined by an external effect of employment that declines exponentially with distance. As in Lucas (2001), we let

$$z(r) = \int_0^S \varphi(r,s)s\theta(s)n(s)\,\mathrm{d}s,$$

where $\varphi(r,s)$ is given by

$$\varphi(r,s) = \delta\int_0^{2\pi} \mathrm{e}^{-\delta x(r,s,\phi)}\,\mathrm{d}\phi \quad \text{and} \quad x(r,s,\phi) = \big[r^2 - 2\cos(\phi)rs + s^2\big]^{1/2}.$$

That is, productivity at a particular location is a weighted average of employment at other locations. The particular form of the production externalities is arbitrary. Fujita and Thisse (2002) derive this type of external effects from knowledge spillovers between firms. It is possible to show that $z(r)$ is also the reduced form of different types of agglomeration effects. For example, if agglomeration effects are the results of differences in local demand and increasing returns.

There are many empirical papers that test for spatial production externalities. Some examples are Ciccone and Peri (2002), Ciccone and Hall (1996), Dekle and Eaton (1999), Ellison and Glaeser (1997), Glaeser et al. (1992), Henderson (2001), Jaffe et al. (1993), Moretti (2002), and Rauch (1993). All these papers find evidence of agglomeration effects. In particular, Dekle and Eaton (1999) find that the external effects decline with distance, particularly in the financial services sector. Jaffe et al. (1993) also provide evidence on patents that suggests that interactions decline with distance. Glaeser et al. (1992) find evidence that supports external effects among different industries and not only inside particular industries. The same is true for Ciccone and Hall (1996) that show that the density of economic activity determines productivity. This is important for us, since we have a one good model and so the external effect affects all firms.[5]

Agents derive utility from residential land and consumption of goods. The utility of an agent that consumes $c(r)$ goods and occupies $\ell(r)$ units of land is given by

$$U\big(c(r),\ell(r)\big).$$

[5] Another way of introducing agglomeration effects in urban economics models is via increasing returns and monopolistic competition. Fujita and Thisse (2002) present a good summary of this literature. We believe that both types of agglomeration effects result in identical optimal land use structures. However, this statement requires a proof which is left for future research. Using external effects or agglomeration effects internal to the firm will have different policy implications.

Commuting is costly in time. Each worker is endowed with one unit of labor. This unit is spent working and commuting. An agent that works at location r, but lives at location s, spends

$$1 - \kappa|r - s| \approx \mathrm{e}^{-\kappa|r-s|}$$

units of time working.

Define the state variable $H(r)$ as the stock of unhoused workers at location r. Given housing and employment at locations $[0, r)$, $H(r)$ keeps track of how many workers need to be housed at locations between r and S (the rest of the city) in order for every worker in the city to have a place to live. We construct this stock as in LRH using a differential equation with boundary condition $H(0) = 0$. At any location r, to calculate the change in $H(r)$ in an interval dr, we need to compute the net number of unhoused workers in the interval,

$$2\pi r\big[\theta(r)n(r) - \big(1 - \theta(r)\big)N(r)\big]\,\mathrm{d}r.$$

We also need to take into account that, if $H(r)$ is positive, we are housing workers dr miles farther away from their work places. This implies that they commute for longer and therefore supply less time for work. Since employment at locations $[0, r)$ is given, we need to house more workers in order for them to supply the needed amount of work. In particular, we need to house $\kappa H(r)\,\mathrm{d}r$ more workers. The above reasoning leads to the following differential equation,

$$\frac{\mathrm{d}H(r)}{\mathrm{d}r} = 2\pi r\big[\theta(r)n(r) - \big(1 - \theta(r)\big)N(r)\big] + \kappa H(r)$$

for $H(r) \geqslant 0$.

If $H(r)$ is negative, as we increase r we are employing workers farther away from their residences, so we need to reduce the amount of employment necessary to provide every resident with a job. So, for $H(r) < 0$,

$$\frac{\mathrm{d}H(r)}{\mathrm{d}r} = 2\pi r\big[\theta(r)n(r) - \big(1 - \theta(r)\big)N(r)\big] - \kappa H(r).$$

The maximization problem can then be stated as:

Problem (O). *Choose functions $n(r)$, $N(r)$, $\theta(r)$, $c(r)$, $\ell(r)$, $z(r)$ and $H(r)$ so as to maximize*

$$\int_0^S 2\pi r\big[\theta(r)g\big(z(r)\big)f\big(n(r)\big) - \big(1 - \theta(r)\big)N(r)c(r)\big]\,\mathrm{d}r \tag{1}$$

subject to, for all $r \in [0, S]$,

$$1 \geqslant \theta(r) \geqslant 0, \tag{2}$$

$$U\big(c(r), \ell(r)\big) \geqslant \bar{u}, \tag{3}$$

$$n(r), N(r) \geqslant 0, \tag{4}$$

$$\ell(r) = \frac{1}{N(r)}, \tag{5}$$

$$H(0) = 0 \quad \textit{and} \quad H(S) \leqslant 0, \tag{6}$$

$$z(r) = \int_0^S \varphi(r, s) s\theta(s) n(s)\,\mathrm{d}s. \tag{7}$$

Constraint (2) states that the proportion of land used for business purposes has to be between 0 and 1. (3) guarantees that every household gets at least a reservation utility $\bar{u}$. One useful way of thinking about this constraint is to let the city be part of a big country. Then $\bar{u}$ is the utility that an agent could get if he migrates to some other city. Constraint (4) guarantees that the density of workers and residents is positive and (5) determines the amount of land per person. (6) guarantees that the stock of unhoused workers at the boundary of the city is zero. That is, that all workers are housed in the city. The last constraint determines the external effect as discussed above.

The Hamiltonian for the problem, without considering (3) and (7), is given by

$$\begin{aligned} G(r, H, n, N, \theta, c, z, \lambda) = {} & 2\pi r\bigl[\theta(r) g\bigl(z(r)\bigr) f\bigl(n(r)\bigr) - \bigl(1 - \theta(r)\bigr) N(r) c(r)\bigr] \\ & - \lambda(r)\bigl[2\pi r\bigl[\theta(r) n(r) - \bigl(1 - \theta(r)\bigr) N(r)\bigr] + \kappa\bigl|H(r)\bigr|\bigr]. \end{aligned}$$

Including constraints (3) and (7) we can built a Lagrangian given by

$$\begin{aligned} L(r, H, n, N, \theta, c, z, \lambda, \xi, \mu) = {} & G\bigl(r, H(r), n(r), N(r), \theta(r), c(r), z(r), \lambda(r)\bigr) \\ & + \xi(r)\left[U\left(c(r), \frac{1}{N(r)}\right) - \bar{u}\right] \\ & + \int_0^S \mu(s)\left[\int_0^S \varphi(s, r) r\theta(r) n(r)\,\mathrm{d}r - z(s)\right]\mathrm{d}s, \end{aligned}$$

where $\xi(r)$ is the Lagrange multipliers associated with (3), $\lambda(r)$ is the co-state variable associated with (6) and $\mu(r)$ is the Lagrange multiplier associated with constraint (7). The co-state variable $\lambda(r)$ represents the cost of employing an extra worker at location r. It is the costs of giving the worker land at some location in $(r, S]$ and enough consumption to obtain utility $\bar{u}$.

The first-order conditions after eliminating the Lagrange multipliers $\xi(r)$ and $\mu(r)$ are, for all $r \in [0, S]$,

$$g\bigl(z(r)\bigr) f'\bigl(n(r)\bigr) \leqslant \lambda(r) - \int_0^S s\theta(s) g'\bigl(z(s)\bigr) f\bigl(n(s)\bigr)\varphi(s, r)\,\mathrm{d}s, \tag{8}$$

$$c(r) + \frac{U_\ell}{N(r) U_c} \geqslant \lambda(r), \tag{9}$$

$$\lambda(r)\big[n(r)+N(r)\big] \gtreqless \big[g\big(z(r)\big)f\big(n(r)\big)+N(r)c(r)\big] + \frac{n(r)}{2\pi}\int_0^S s\theta(s)g'\big(z(s)\big)f\big(n(s)\big)\varphi(s,r)\,\mathrm{d}s, \tag{10}$$

where the first two conditions hold with equality if $n(r) > 0$ and $N(r) > 0$. In the last condition, the left-hand side is greater if $\theta(r) > 0$, smaller if $\theta(r) < 0$, and equal to the right-hand side if $\theta(r) \in (0, 1)$.

These equations plus the constraints

$$U\left(c(r), \frac{1}{N(r)}\right) = \bar{u}, \qquad \int_0^S \varphi(r,s)s\theta(s)n(s)\,\mathrm{d}s = z(r),$$

and (6), form the system of first-order conditions.

The last terms (8) represents the effect on other locations of hiring an extra worker at location r. Since this term is positive, if f is concave, $n(r)$ will be larger than if the externality was not internalized. It is also useful to understand the intuition behind condition (10). Rewrite it as

$$g\big(z(r)\big)f\big(n(r)\big) + \chi(r) - \lambda(r)n(r) \gtreqless N(r)\big[\lambda(r) - c(t)\big],$$

where

$$\chi(r) = \frac{n(r)}{2\pi}\int_0^S 2\pi s\theta(s)g'\big(z(s)\big)f\big(n(s)\big)\varphi(s,r)\,\mathrm{d}s.$$

Then, the first term on the left hand side is total output per unit of land if location r is a business sector. The second is the total gain in output at other locations from producing at location r, caused by the external effect, and the third represents the cost of employing $n(r)$ workers. On the right hand side, $N(r)[\lambda(r) - c(t)]$ are the benefits in terms of net output of housing $N(r)$ worker at location r. As we mentioned before, $\lambda(r)$ has the interpretation of the unit cost, land plus consumption, of employing an extra worker. So $[\lambda(r) - c(t)]$ is the gain per resident of assigning the land at location r for housing. Hence, the first-order condition compares the benefits, in terms of net output, of land being used for business or residential purposes.

The Maximum Principle provides an extra condition on the behavior of the co-state variable. Namely,

$$\frac{\partial\lambda(r)}{\partial r} = -\frac{\partial G(r)}{\partial r} = \begin{cases} -\kappa\lambda(r) & \text{if } H(r) > 0 \\ \kappa\lambda(r) & \text{if } H(r) < 0 \end{cases} \quad \text{for all } r \in [0, S], \tag{11}$$

and Pontryagin Maximum Principle also requires λ to be continuous. The system of first order conditions can be solved for n, N, c, z given $\lambda(0)$ and θ. Using (6) we can then solve for the value of $\lambda(0)$. Given this, we can use the condition above to calculate λ at all locations r where $H(r) > 0$ or $H(r) < 0$. If $H(r) = 0$, then $\partial\lambda(r)/\partial r = -\kappa\lambda(r)$

if $\theta(s) = 1$ for $s > r$ arbitrarily close to r and $\partial\lambda(r)/\partial r = \kappa\lambda(r)$ if $\theta(s) = 0$ for $s > r$ arbitrarily close to r. If $\theta(s) \in (0, 1)$ for $s > r$ arbitrarily close to r, then if

$$\theta(r)n(r) - \bigl(1 - \theta(r)\bigr)N(r) > 0, \quad \text{we know that} \quad \frac{\partial\lambda(r)}{\partial r} = -\kappa\lambda(r),$$

since we are accumulating unhoused workers so $H(s) > 0$ and λ is a continuous function. If

$$\theta(r)n(r) - \bigl(1 - \theta(r)\bigr)N(s) < 0, \quad \text{then} \quad \frac{\partial\lambda(r)}{\partial r} = \kappa\lambda(r),$$

since we are decreasing the amount of unhoused workers. Finally, if

$$\theta(r)n(r) - \bigl(1 - \theta(r)\bigr)N(r) = 0,$$

we can use this equation to solve for $\partial\lambda(r)/\partial r$.

The above description implies that we can solve for all variables in terms of θ. The next section provides an algorithm to find the optimal land use structure, summarized by θ.

3. Optimal land use structure

All the exposition until now shows how to obtain the different controls, the state, and the co-state given a function of land use proportions, θ. But it does not provide an algorithm to find the optimal land use structure. Notice that the first-order condition with respect to $\theta(r)$ can be used only when $\theta(r) \in (0, 1)$ since in both other cases constraint (2) is binding. Optimization theory does not provide a good way of dealing with corner solutions, one possible way of dealing with this problem is to compute the value of the objective function for all possible combinations (given a minimum lot size). This solution is very costly and would not add to the understanding of the problem at hand. We proceed in a different way. The first-order condition in (9) implies that if, for example, location r is assumed to be a pure business area, and it is actually optimal for this area to be a pure business area, the constraint $\theta(r) \leqslant 1$ must be binding. Hence

$$\frac{[g(z(r))f(n(r)) + N(r)c(r)] + \chi(r)}{[n(r) + N(r)]} > \lambda(r).$$

Define $t(r)$ by

$$t(r) \equiv \frac{[g(z(r))f(n(r)) + N(r)c(r)] + \chi(r)}{[n(r) + N(r)]}.$$

Using the same type of reasoning for all possible assumptions on $\theta(r)$, we know that at the optimum,

$$\begin{aligned} &\theta(r) = 1 && \text{iff} \quad t(r) > \lambda(r), \\ &\theta(r) \in (0, 1) && \text{iff} \quad t(r) = \lambda(r), \quad \text{and} \\ &\theta(r) = 0 && \text{iff} \quad t(r) < \lambda(r). \end{aligned}$$

Every time we propose a solution we need to check if the above inequalities hold, if they do not, the allocation is not a solution to Problem (O).

In order to find a maximum, a solution to Problem (O), it seems that a strategy that assumes some land use structure and then tests if the solution is actually a maximum is not going to take us too far. In this section we will provide a general solution algorithm to find the solution of this problem. The basic idea is to start with some initial guess for the function θ, compute the corresponding densities, productivity function, and state variables, and then use the previous equations to obtain a new function θ.

Define an operator $T : M \to M$ (where M denotes the space of measurable function $f(r) : [0, S] \to [0, 1]$) that maps land use proportion functions θ into land use proportion functions. In particular we let

$$(T\theta)(r) = \begin{cases} 1 & \text{if } t(r) > \lambda(r), \\ \frac{N(r)}{N(r)+n(r)} & \text{if } t(r) = \lambda(r), \\ 0 & \text{if } t(r) < \lambda(r). \end{cases}$$

Notice that in the case where $t(r) = \lambda(r)$, we let $(T\theta)(r) = \frac{N(r)}{N(r)+n(r)}$. This has to be the case for

$$\theta(r)n(r) - \big(1 - \theta(r)\big)N(r) = 0$$

to hold. If this equation is not satisfies we know that $\lambda(r)$ will be decreasing at rate κ or growing at rate κ, and so the equality can only hold for a set of Lebesgue measure zero. If the constraint is satisfied, then we may have $H(r) = 0$ for an interval of locations with positive Lebesgue measure, a mixed sector.

A solution that satisfies the first-order condition for θ, will have to satisfy

$$T\theta = \theta.$$

If we can find a function θ that is a fixed point of the operator T, then we have found an allocation that satisfies all the necessary conditions for a maximum (See Proposition 3).

The existence of a solution that satisfies the necessary conditions, and the existence of a fixed point of operator T, is proven in the following propositions and theorem. All the proofs of theorems and propositions are presented in Appendix A. To prove the results we impose the following assumptions.

Assumption A. (i) g and f are continuously differentiable and strictly increasing.

(ii) For any function θ,

$$g\left[\int_0^S \varphi(r, s)s\theta(s)n(s)\,\mathrm{d}s\right] f\big(n(r)\big)$$

is strictly concave in $n(\cdot)$.

(iii) U is continuously differentiable, strictly increasing in both arguments, concave, and $U(\cdot, 0) < \bar{u}$.

(iv) κ, δ, and γ are all positive and finite.

(v) There exists an arbitrarily small ε such that for any $r_1 > r_2 > r_3$ such that $|r_1 - r_3| < \varepsilon$, $|\theta(r_1) - \theta(r_2)| + |\theta(r_2) - \theta(r_3)| < 2$.

Two of these assumptions require an explanation. Part (ii) guarantees that the objective function is concave in $n(\cdot)$. Part (iii) guarantees that lot sizes are uniformly bounded above zero as we increase κ arbitrarily.

Proposition 1. *Under Assumption A, given θ there exists a unique set of functions $\{n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ that satisfies the first-order conditions and the Maximum Principle conditions.*

Proposition 2. *Under Assumption A, there exists a function θ^* such that*

$$T\theta^* = \theta^*.$$

Proposition 3. *Under Assumption A, there exists a set of functions $\{\theta^*, n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ that satisfies the first-order conditions, the Maximum Principle conditions and $T\theta^* = \theta^*$.*

We now prove that the solution consists only of pure business and pure residential areas.

Theorem 1. *Under Assumption A, the optimal land use structure has no mixed areas. That is $\theta^*(r) \in \{0, 1\}$, except for sets with zero Lebesgue measure in $[0, S]$.*

Mixed area require two conditions, $H(r) = 0$ and that the value of adding an extra worker equals the value of housing an extra resident, $t = \lambda$. The idea behind the proof of Theorem 1 is that this can happen, but only at some specific r and not for an interval with positive Lebesgue measure. This is called a transversal crossing, the two values may cross but their slope with respect to r is different.

Mixed areas can be part of the equilibrium city structure, as shown in LRH. In equilibrium, wages at the center can adjust so that the two necessary conditions for a mixed area are satisfied. The difference is that in the efficient allocation changes in $\lambda(0)$ will also change $\chi(r)$ (the effect of employing an extra worker at location r in the production of other firms in the city), which in turn changes the gains from using the location for business purposes. Thus, in general there is no $\lambda(0)$ that satisfies the conditions necessary for the existence of a mixed area.

The next proposition uses Theorem 1 to prove existence of an efficient allocation.

Proposition 4. *Under Assumption A, there exists a set of functions $\{\theta^*, n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ that solves Problem (O).*

Figure 2 illustrates the fixed point of the iterations of t and λ, that in turn determine the fixed point of θ. We use the parameter values presented in Section 5 as the base case, in particular $\kappa = 0.005$. The initial land use proportions function is $\theta_0(r) = 1$ for $r \in [0, 2]$ and $\theta_0(r) = 0$ for $r \in (2, 10]$. The fixed point of the operator is given by $\theta^*(r) = 1$ for $r \in [0, 3.5]$ and $\theta^*(r) = 0$ for $r \in (3.5, 10]$. We have guaranteed the existence of a fixed point, but not the convergence of the iterates of θ to that fixed point. Hence, the fixed point could depend on the initial function θ_0. We computed the same experiment with other

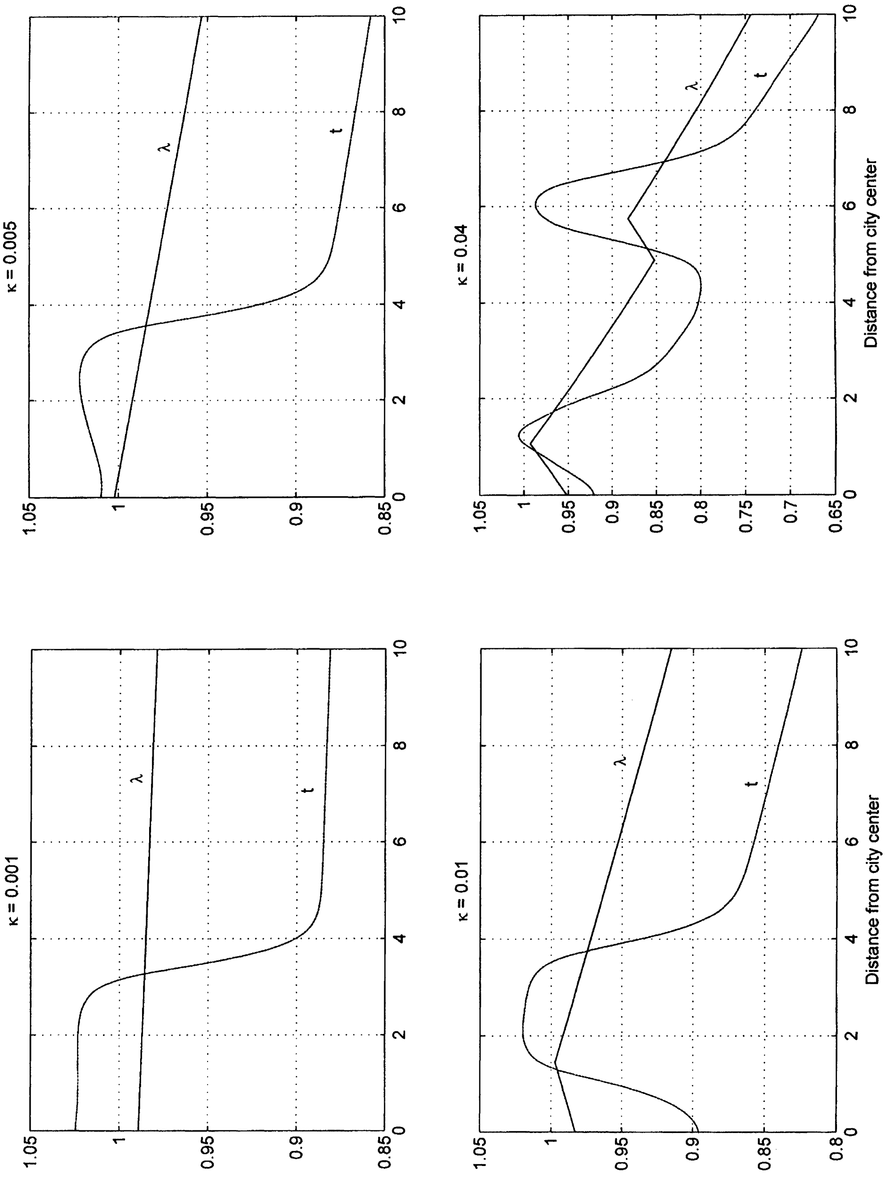

Fig. 2. Fixed point of t and λ.

initial functions and obtained the same result. We also computed the value of the objective function for different locations of the borders between the business and residential sectors, and in all case obtained smaller values. All of this leads us to believe that this is the actual solution to Problem (O) and not a local maximum or saddle point.

The method proposed is more efficient than looking at all possible combinations of land use structures. Even if we only look at a restricted set (e.g. only one business sector). The reason is that each iteration of the operator is equivalent to testing a set of boundaries, but it allows us to move in the right direction and without testing all boundaries in between two iterations. It also guarantees that the solution found is an extreme point of Problem (O) over all possible combinations of land use structures. Finally, if different initial land use proportion functions θ_0 are used, and all of them converge to the same fixed point, it gives some assurance that the fixed point is unique and therefore a global maximum of the problem.

Figure 2 also presents exercises for three other κ values: 0.001, 0.01, and 0.04. The κ values where chosen to illustrate different possible allocations of land in the model: a central business district, with high ($\kappa = 0.001$) and low ($\kappa = 0.005$) densities, a ring of businesses ($\kappa = 0.01$) and two rings of businesses ($\kappa = 0.04$). If $\kappa = 0.001$, the center of the city is used for business purposes and densities are high at the center of the city. For $\kappa = 0.01$ agents will commute to and away from the center in order to get to work since there is a residential core. Hence, λ first increases at rate κ, has a kink, and then starts decreasing at rate κ. Finally, when $\kappa = 0.04$, the center is residential and we obtain two rings of business areas, together with rings of residences between business areas and at the boundary. All examples use the same initial land use function as the base case. How does the land use structure change as we increase commuting costs and decrease externalities? The results suggest that the solution will have several rings of businesses with rings of residences in between them. Larger commuting costs and lower externalities will be associated with *more* rings of businesses.

4. Optimal subsidies and zoning policies

The analysis of optimal urban structure can be used to design optimal policies to improve the efficiency of equilibrium allocations in the same setup. The first step is to define the marginal product of labor and the value of land in the optimum. With these two concepts in hand, we then proceed to design and characterize the effect of different policies. The main problem is to find policies that implement the optimum as an equilibrium.

One argument against designing policies using our framework, is that is abstracts from congestion costs. Commuting costs are the only force balancing agglomeration. Many papers in the urban economics literature focus on congestion and the policies recommended to reduce its costs (see for example Dixit, 1973; Anas and Xu, 1999; or Wheaton, 1998). We do not model congestion directly, although it can be modeled as part of the externality parameters. δ can be thought of as consisting of two parts, a positive part that makes firms want to be close to each other, and a negative part that makes firms want to be far apart (the congestion effect). In this paper we focus on positive δ, so the positive spillover is more important than the negative one resulting from congestion. In this framework, if δ is not positive, firms do not agglomerate in cities. So the fact that in reality there cities do exists implies that $\delta > 0$. Some evidence on the need of high positive δ values is presented in Lucas (2001). This reasoning leads us to believe that the omission of congestion in the model does not invalidate the results in the paper. Nevertheless, further attempts to include congestion in this framework could provide interesting insights.

4.1. Marginal product of labor and the value of land

Given n^*, N^*, z^*, θ^* and λ^* we can calculate the land rents and wages that, in an equilibrium context, would make agents and firms at a particular location choose the same amount of workers, land, and consumption, as the planner. These are not equilibrium wages and land rents since for these prices agents and firms have incentives to move to other locations. In business areas, the value of land and the marginal product of labor can be computed using

$$q(r) = g\big(z^*(r)\big) f\big(n^*(r)\big) - w(r)n^*(r)$$

and the implied first-order condition of the firm's problem

$$w(r) = g\big(z^*(r)\big) f'\big(n^*(r)\big).$$

Solving the system, the value of land and the marginal product of labor can for $\theta(r) = 1$ be given by

$$\begin{aligned} q(r) &= g\big(z^*(r)\big) f\big(n^*(r)\big) - g\big(z^*(r)\big) f'\big(n^*(r)\big)n^*(r), \\ w(r) &= g\big(z^*(r)\big) f'\big(n^*(r)\big) = \lambda^*(r) - \int_0^S s\theta^*(s) g'\big(z^*(s)\big) f\big(n^*(s)\big)\varphi(s,r)\,\mathrm{d}r, \end{aligned}$$

where the second equality in the second equation comes from the first-order condition of Problem (O). Notice that the marginal product of labor in pure business sectors is lower than the equilibrium wage where wages are given by

$$w(r) = \lambda^*(r).$$

The fact that firms internalize the externality results in a lower marginal product of labor. To take full advantage of spillovers, optimal employment is greater than in equilibrium which implies a lower marginal product, given decreasing returns to labor. The slope of λ^* has to be the same here than in equilibrium, although not the level.

The consumer problem is given by

$$\begin{aligned} & w(r) = \max\left[c^*(r) + q(r)\left(\frac{1}{N^*(r)}\right)\right], \\ & \text{s.t.} \quad \bar{u} \geqslant U\left(c^*(r), \frac{1}{N^*(r)}\right) \end{aligned}$$

and the corresponding first-order condition is

$$q(r) = \frac{U_\ell}{U_c}.$$

Solving the system, the value of land and the marginal product of labor for $\theta(r) = 0$ are

$$q(r) = \frac{U_\ell\left(c^*(r), \frac{1}{N^*(r)}\right)}{U_c\left(c^*(r), \frac{1}{N^*(r)}\right)} = N^*(r)\lambda^*(r) - N^*(r)c^*(r),$$

$$w(r) = c^*(r) + \frac{U_\ell\left(c^*(r), \frac{1}{N^*(r)}\right)}{U_c\left(c^*(r), \frac{1}{N^*(r)}\right)}\left(\frac{1}{N^*(r)}\right) = \lambda^*(r),$$

where the second inequality in the first equation follows from the first-order condition.

4.2. Optimal subsidies

The definition of equilibrium in LRH determines that in equilibrium firms maximize profits, residents minimize expenditures given a reservation utility, wage no arbitrage conditions are satisfied, land is assigned to its highest value (highest bid rent), everyone is housed, and the externality function, z, is constructed as described in constraint (7). Hence if we want to implement the optimal allocation as an equilibrium, we need to check that the wage no arbitrage conditions are satisfied by the optimal marginal product of labor and that land is assigned to its highest value at the optimal value of land. All other conditions are constraints in Problem (O).

In order to implement the optimal allocation as an equilibrium, we need to make sure that the policy that we design satisfies the wage no arbitrage conditions. these conditions are given by

$$w(r) = K_1 e^{-\kappa r} \quad \text{if } H(r) > 0 \quad \text{and} \quad w(r) = K_2 e^{\kappa r} \quad \text{if } H(r) < 0,$$

for some constants K_i, $i = 1, 2$.

In our case, λ^* does satisfy these conditions since, as we discussed before,

$$\frac{\partial \lambda^*(r)}{\partial r} = \begin{cases} -\kappa\lambda^*(r) & \text{if } H(r) > 0, \\ \kappa\lambda^*(r) & \text{if } H(r) < 0. \end{cases}$$

Nevertheless, the optimal marginal product of labor function $w(r)$ does not. We have shown that in business areas,

$$w(r) = \lambda^*(r) - \int_0^S s\theta^*(s) g'\left(z^*(s)\right) f\left(n^*(s)\right)\varphi(s, r)\,\mathrm{d}r.$$

So a policy that satisfies the wage no arbitrage conditions is a labor subsidy τ to firms of the form

$$\tau(r) = \begin{cases} \int_0^S s\theta^*(s) g'\left(z^*(s)\right) f\left(n^*(s)\right)\varphi(s, r)\,\mathrm{d}r & \text{if } \theta^*(r) > 0, \\ 0 & \text{if } \theta^*(r) = 0. \end{cases}$$

The second condition that we need to check, for any proposed policy, is that land is assigned to its highest value. That is, we need

$$q(r)|_{\theta(r)=1} > q(r)|_{\theta(r)=0} \quad \text{for a business sector,}$$
$$q(r)|_{\theta(r)=1} = q(r)|_{\theta(r)=0} \quad \text{for a mixed sector, and}$$
$$q(r)|_{\theta(r)=1} < q(r)|_{\theta(r)=0} \quad \text{for a residential sector.}$$

One can show that this is always the case in the optimal allocation. For this, suppose that

$$\frac{[g(z^*(r)) f(n^*(r)) + N^*(r)c^*(r)] + \chi(r)}{[n^*(r) + N^*(r)]} > \lambda^*(r),$$

or

$$g\bigl(z^*(r)\bigr) f\bigl(n^*(r)\bigr) - \lambda^*(r)n^*(r) + \chi(r) > N^*(r)\lambda^*(r) - N^*(r)c^*(r).$$

From the definition of $q(r)|_{\theta(r)=1}$, we know that

$$\begin{aligned} q(r)|_{\theta(r)=1} &= g\bigl(z^*(r)\bigr) f\bigl(n^*(r)\bigr) - w(r)|_{\theta(r)=1} n^*(r) \\ &= g\bigl(z^*(r)\bigr) f\bigl(n^*(r)\bigr) - \lambda^*(r)n^*(r) + \chi(r). \end{aligned}$$

Notice also that

$$w(r)|_{\theta(r)=0} = c^*(r) + q(r)|_{\theta(r)=0}\left(\frac{1}{N^*(r)}\right),$$

which implies that

$$q(r)|_{\theta(r)=0} = N^*(r)w(r)|_{\theta(r)=0} - N^*(r)c^*(r) = N^*(r)\lambda^*(r) - N^*(r)c^*(r).$$

Hence in a business sector of the optimal allocation $q(r)|_{\theta(r)=1} > q(r)|_{\theta(r)=0}$ is always satisfied. The proof is analogous for the cases of mixed and residential sectors.

We are ready to formally state the result that the subsidy proposed above implements the optimal allocation as an equilibrium.

Theorem 2. *A labor subsidy τ of the form*

$$\tau(r) = \begin{cases} \int_0^S s\theta^*(s) g'\bigl(z^*(s)\bigr) f\bigl(n^*(s)\bigr)\varphi(s,r)\,\mathrm{d}r & \text{if } \theta^*(r) > 0, \\ 0 & \text{if } \theta^*(r) = 0, \end{cases}$$

implements the optimal allocation as an equilibrium.

Proof. We have shown above that with this subsidy the optimal land values and marginal products of labor satisfy the wage no arbitrage condition and the condition that land is assigned to its highest value. Since all other parts of the definition of equilibrium in LRH are either imposed as constraints of Problem (O), or implied by the definition of optimal land value and marginal product of labor, the equilibrium allocation with labor subsidies is an optimum. □

The subsidy proposed in Theorem 2 is a subsidy that the government should pay firms so that they face lower labor costs, and hence hire more workers (in equilibrium average wages will increase, and since reservation utility is fixed, residential rents). One may ask how, in this model, a government would pay for the subsidy. One possibility is to levy a flat tax on land owners. The implied tax is given by

$$\tau^{+} = \frac{1}{2\pi S}\int_{0}^{S} \tau(s)\,\mathrm{d}s$$

per unit of land. This tax will not distort rents and would balance the city government budget. Notice that in this model we have absentee landlords so we are not actually modeling the decision of buying land in the city or in different cities.

In reality we do not observe these types of subsidies, and although this does not undermine their optimality, it is puzzling that we do not see any efforts in this direction. A closer look at City Government policies may give us a possible answer to the problem. For example, parking lots construction by government agencies may be a way of reducing the costs of working at business centers, thereby actually subsidizing workers in these areas. Highway improvement and construction and investments in public transportation are other interesting examples.

4.3. *Zoning policy*

Suppose the city government can determine land use in different sections of the city. That is, the city government can determine if some section of the city is a business area, a residential area or a mixed area (the government has the power to determine $\theta(r)$). In this case, we can ask what the best set of restrictions is, and if zoning is an effective policy to increase net output in a city.

The optimal land use does not solve the optimal zoning policy problem. This problem is given by Problem (O) plus the constraints that the optimal marginal product of labor declines or increases at rate κ throughout the city. That is, the optimal zoning problem implies solving for the optimal land use structure restricting the allocation to be an equilibrium given θ. We do not solve that problem in this paper, but we do analyze the case where we impose the optimal land use as a restriction. This is what we call the equilibrium with zoning restrictions. The exercise yields a lower bound for the benefits of zoning restrictions since the optimal land use structure is a feasible θ.

The equilibrium with zoning restrictions will not internalize the externality, and hence will not be optimal, but it will take advantage of higher external effects caused by firms being closer to each other (in Theorem 1, we proved that there are no mixed areas in the optimal land use structure). Hence, net output will increase. How much of the way between the optimum and the equilibrium will this type of restrictions take us is a numerical issue and will be studied in the next section.

One important feature of the equilibrium with zoning restrictions is that land rents will not be continuous at the boundary between business and residential sectors. The zoning restriction will be binding, since the equilibrium land use structure is different than the optimal land use structure. So there are incentives for some land owners to lobby for changes in the zoning restrictions. Given the restrictions, some firms may, for example, be ready to pay higher rents in residential sectors than residents. Of course, if the restrictions are lifted altogether, we return to the equilibrium outcome, and since the external effect is lower in equilibrium, land rents will go down. In the next section we will presents examples of the equilibrium with zoning restrictions where the discontinuities in land rents can be observed.

5. Numerical exercises

In this section we use numerical exercises to illustrate some of the equilibrium possibilities of the model. We will set the different parameters values following the previous literature and then vary key parameters to illustrate their effect on optimal city structure, in particular κ and δ. These two parameters are crucial in determining the land use structure within a city and the results are very sensitive to their values. The objective of these exercises is not to obtain numerical implications that match particular characteristics of a specific city, but to illustrate how optimal allocations and policies react to changes in key fundamental parameters. The model is admittedly very stylized and should be used as a conceptual framework to understand the interaction of forces that are important in determining urban structure, not to obtain quantitative implications.

In all the numerical exercises we use a Cobb–Douglas specification for the utility function,

$$U(c,\ell)=c^{\beta}\ell^{1-\beta},$$

and the production function,

$$f(n)=An^{\alpha}.$$

We also let

$$g(z)=z^{\gamma}.$$

In the numerical exercises shown below we use the following parameter values:

α	β	γ	A	$\bar{u}$	κ	δ	S
0.95	0.9	0.04	1	1	0.005	5	10

Caselli and Coleman (2001) show that the share of land rentals in nonfarm business income in the US is around 5%, hence we let $\alpha=0.95$. Roback (1982) estimates the share of income going to mortgage payments to be 0.178, since this includes structures (payments for the actual house not only land) we reduce the number to 0.1 and let $\beta=0.9$. Lucas (2001) uses Japanese Land Rent data to argue that a reasonable value of γ is 0.04. A and $\bar{u}$ are set arbitrarily at 1.[6] The city size is also set arbitrarily at $S=10$ miles.[7] These

[6] Another possibility is to calibrate $\bar{u}$ such that the population size is about 100,000. In a sample of the largest 1083 cities in the USA, the average population size was 105412 according to the 1990 Census. We could also calibrate A such that the average per capita monetary income is about 15000 US dollars. In 1990, the average monetary income per capita in a sample of 1083 cities was 14867 US dollars.

[7] The average land area in 1990 of a sample of 1083 cities was 38.78 sqm.

parameter values are fixed throughout the paper. We will perform comparative statics with all the remaining parameters. For $\delta = 5$ see Lucas (2001) and LRH, for $\kappa = 0.005$, notice that in a sample of 1083 cities, people commute 20.57 minutes on average to and from work. If they work around 8 hours a day (including commuting), this implies that they spent 4.29 percent of their time commuting. So if they commute on average about 10 miles it implies a $\kappa = 0.00429$. Since some of these calculations are very loose, we will perform numerical exercises for a wide range of values of κ.

In order for Assumption A to hold for this particular specification, we need $\alpha, \beta, \gamma > 0$, $\beta, \alpha < 1$, and $1 - \alpha > \gamma$. Clearly these restrictions are satisfied for the base case parameter values presented above. For an interpretation of $1 - \alpha > \gamma$ in the urban economics literature, see Lucas (2001) or Fujita et al. (1999). We formalize this statement in the following lemma, proven in Appendix A.

Lemma 1. *If U, f and g are Cobb–Douglas as specified above, $\alpha, \beta, \gamma > 0$, $\beta, \alpha < 1$ and*

$$1 - \alpha > \gamma,$$

then Assumption A is satisfied.

We solve the nonlinear system of equations in Section 2 using MATLAB. The algorithm solves the system given a function θ. We let $\Delta r = 0.01$ and so we have a system of $S5/\Delta r + 1$ equations in the same number of unknowns. As a result we obtain the optimal densities n^*, N^*, the optimal consumption per person, c^*, the optimal productivity function z, and the co-state λ, at all locations and as a function of θ. To find the optimal θ we iterate using the operator T until the sequence of θ functions converges. As noted before this sequence does not have to converge, although it does in all the simulations presented.

For the base case parameters, we obtain a business area with a radius of 3.5 miles. The rest of the city is residential. In LRH, the equilibrium city for the same parameter values has a mixed sector in the middle with a radius of about 4 miles, followed by a residential area of about two miles, and a residential sector at the boundary. Hence, business areas are more concentrated at the optimum. The densities of workers and households, and marginal product of labor and land values are presented in Fig. 3.

Figure 3 shows the main forces that are at work to determine the optimal allocation. The density of workers is relatively high at the center, since business sectors are relatively close in all directions which increases the external effect. As we move away from the center the external effect decreases. Even though the external effect decreases, firms are closer to the residential areas. Being closer to residential areas implies that less production has to be spent in commuting costs. The result is an increase in the amount of workers at these locations. The higher employment density increases the productivity of firms nearby. The productivity function and the optimal values of land follow this same pattern. The density of households declines since it is a function of λ. As we move away from the center, the cost of housing workers increases since commuting costs increase with distance. Consumption decreases too, since these households are living in larger land lots ($\ell = 1/N$ grows) and their utility is constant at $\bar{u}$. The marginal product of labor follows first the

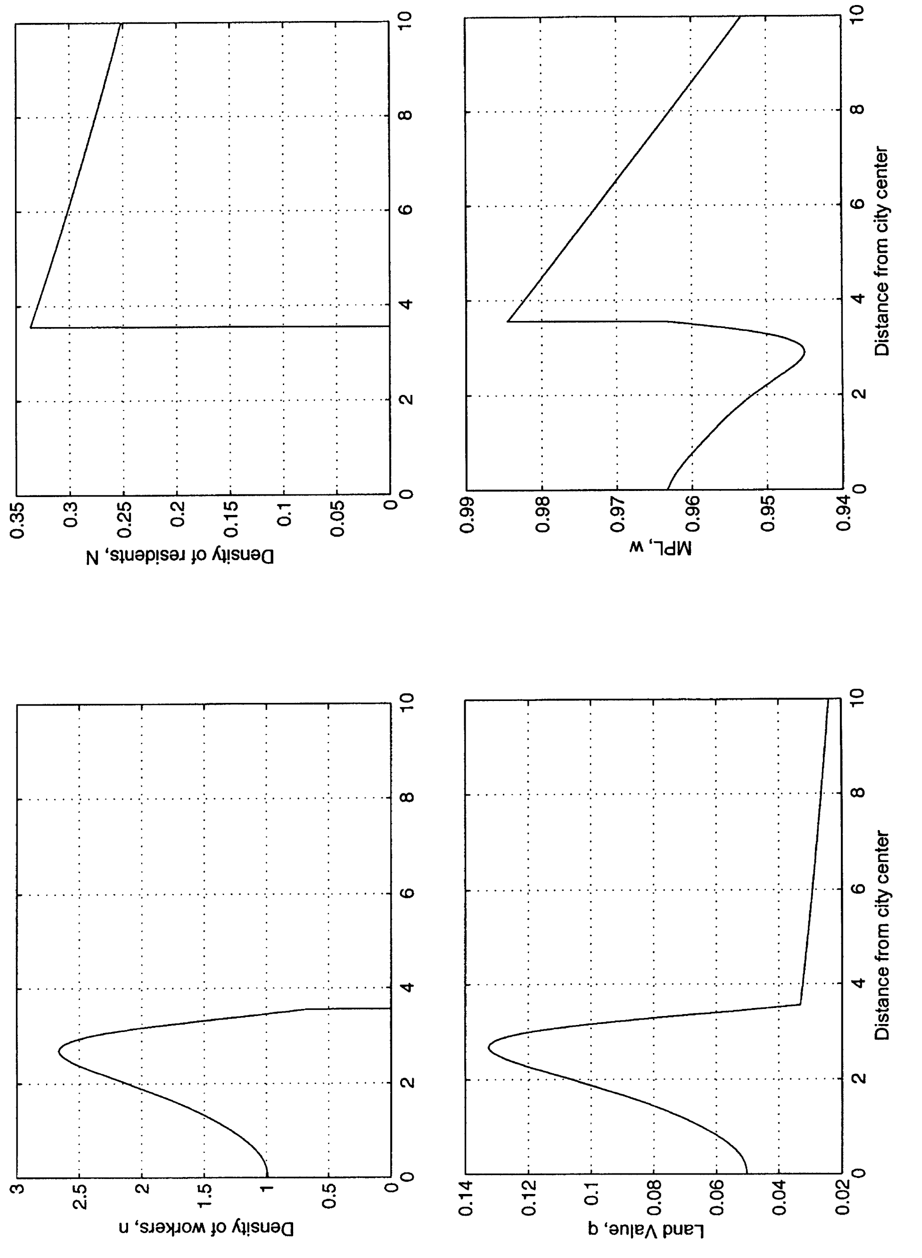

Fig. 3. Optimal allocation for $\kappa = 0.005$.

same pattern as the density of workers and then the pattern of the density of residents. Since $H(r)$ is positive for all r, λ decreases at rate κ throughout the entire city. At each location in business sectors the optimal marginal product of labor is lower than λ, since firms hire more workers than the profit maximizing amount. At the optimum we are internalizing the external effect. In an equilibrium allocation $w(r) = \lambda(r)$. The trade-off between higher external effects and proximity to residential sectors is also illustrated in Fig. 4 for commuting costs $\kappa = 0.001$.

If we increase κ to 0.01, the high commuting cost imply that the monocentric city is not optimal. In this case the business sector is a ring around the center of the city. The rest of the city, that is the center and the areas near the boundary are residential. Since commuting costs are higher, proximity to residents becomes more important than the spillovers from nearby firms. Agents leaving near the center will commute *away* from the center to go to work. Agents living close to the boundary will commute *to* the center to go to work, as in all other cases. These results are presented in Fig. 5.

In Fig. 2 we presented the fixed point of t and λ. When $\kappa = 0.005$, the fixed point of t and λ is such that $t(r) \geqslant \lambda(r)$ at the center (the business district). For locations farther away from the center $t(r) < \lambda(r)$ and so $\theta(r) = 0$ (the residential area). That is,

$$\theta^*(r) = \begin{cases} 1 & \text{if } r \in [0, 3.5], \\ 0 & \text{if } r \in [3.5, 0], \end{cases}$$

is a fixed point of the operator T for $\kappa = 0.005$. For $\kappa = 0.001$ and $\kappa = 0.01$ the figure shows how the fixed point of t and λ imply the land use structure described above. In Fig. 2 we also presented a numerical example for $\kappa = 0.004$. The result is a land use structure with two rings of businesses. In what follows we drop this case and focus on comparative statics where we increase κ to 0.01 and decrease κ to 0.001 relative to the base case.

Given our focus on symmetric allocations, we can ask how important is the restriction that the whole land area, πS^2, has to be used to construct only one city. We can address this question numerically. We start with the optimum presented in Fig. 5, break down the total land area of the city in two, and compute the optimum for each of the smaller cities assuming no external effects between them. The results imply that the sum of the two small cities net output is smaller than the net output of the initial city by 13%. We repeated this exercise by breaking down the original city in four and eight cities. The sum of net output is decreasing in the number of cities. The smaller the land area, the more concentrated the business sectors. In particular, for the cases of 4 and 8 cities we obtain monocentric land use structures.[8]

In Fig. 6 we present the optimal density of workers (times θ) resulting from the optimal land use structure (same as in Fig. 2) and the optimal density of workers given the equilibrium land use structure. In both curves the external effect is internalized. That is, the dashed line is not the equilibrium density of workers for this parameter configuration, but the optimal density of workers restricting land use to be as in equilibrium. For $\kappa = 0.005$, equilibrium land use is given by a mixed area of 4.2 miles at the center, surrounded by a

[8] It is clear that the result will not hold if we start doubling the size of the city. The extra commuting costs and the implied less concentrated business areas will eventually offset the gains in productivity implied by the larger city.

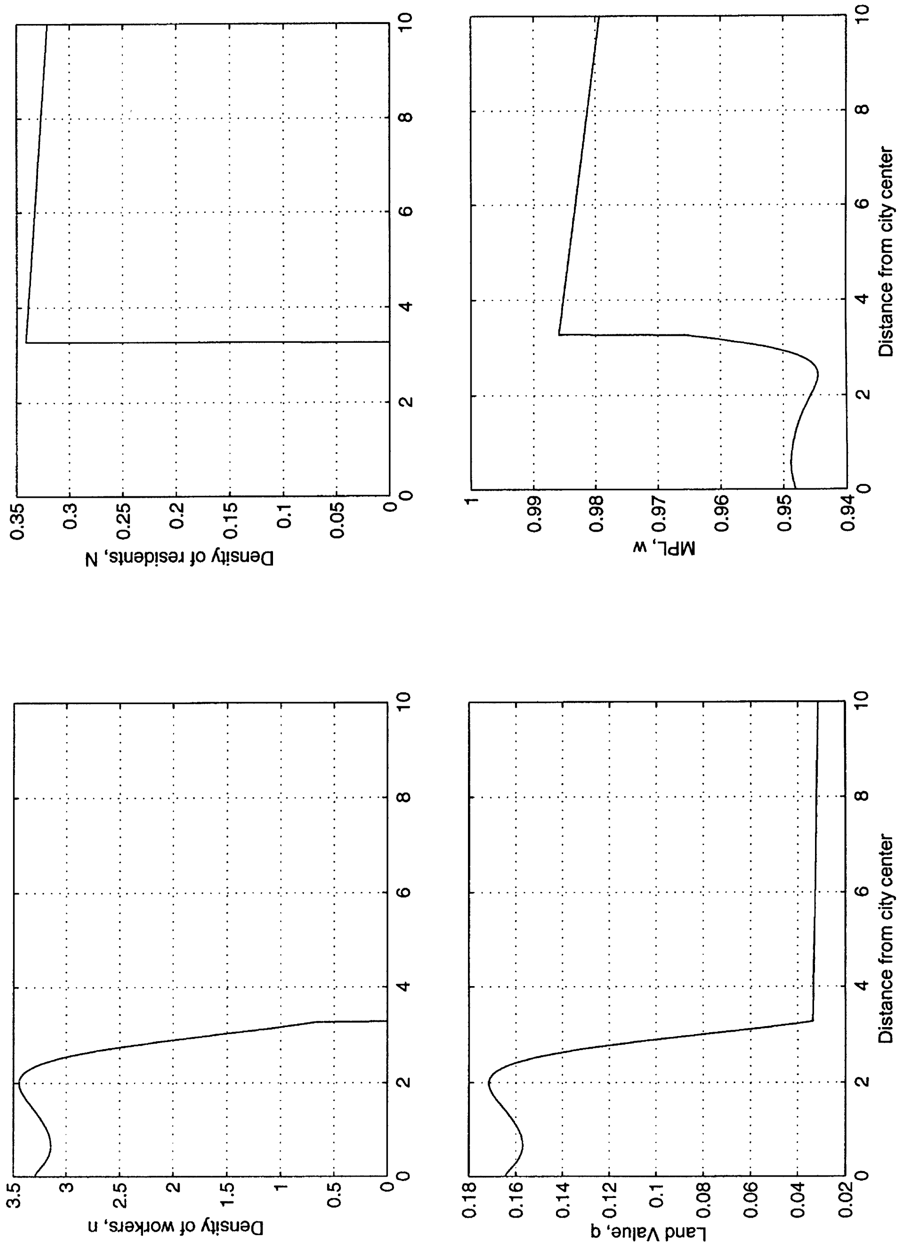

Fig. 4. Optimal allocation for $\kappa = 0.001$.

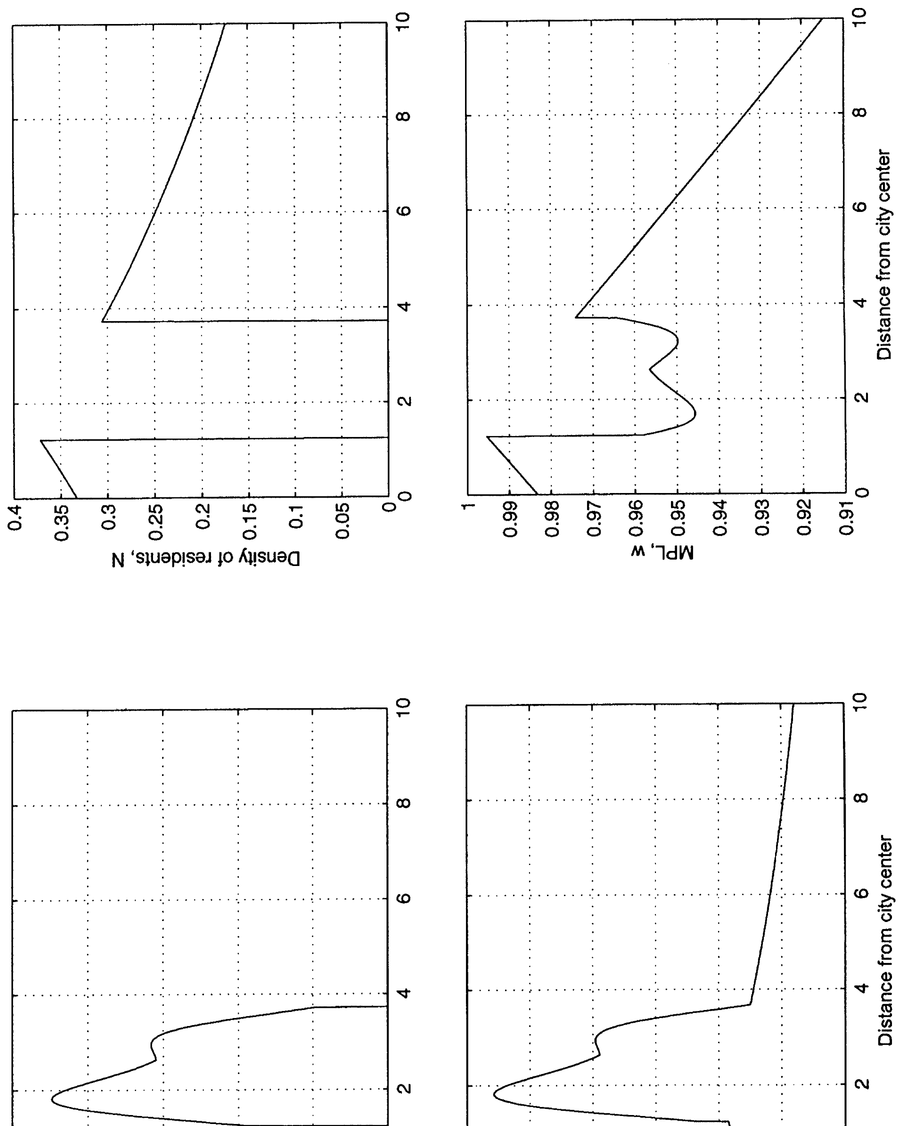

Fig. 5. Optimal allocation for $\kappa = 0.01$.

business area of 2 miles. The rest of the city is residential. The exercise illustrates the pure effect of changes in land use on the density of workers. Endogenizing land use to obtain optimal employment densities is crucial. Net output (total output minus total consumption) increases by 9.5%. One may argue that in order to evaluate this effect correctly it may also be important to look at the total number of residents in the city. This number increases by 10%.

Figure 6 shows the results of the same exercise for $\kappa = 0.001$ and $\kappa = 0.01$. In these cases net output increases by 2.1% and 28% respectively. The total number of residents increases by 1.2% and 13%. The changes in net output and total number of residents suggest that the higher commuting costs, the larger the effect of changes in the land use structure. Hence, the exercises suggest that zoning restrictions or labor subsidies that result in equilibrium allocations that are similar to the optimal allocation imply a higher gain, the higher commuting cost.

Another key parameter in the model is the externality parameter δ. A higher δ implies that the external effect decays faster with distance. Notice that a larger δ, as can be seen in the equations presented in Section 2, does not imply that the external effect is lower everywhere. This, since the function φ is also multiplied by δ. The idea is to make the external effect decay faster without changing the mean external effect.

In Fig. 7 we present the density of workers (multiplied by θ) for $\delta = 5$, 10, and 15. A higher δ increases the amount of workers at the center when $\kappa = 0.001$ and 0.005 (the cases where there is a business district at the center). The intuition is that a higher δ increases the relative advantage of locating at the center, where firms are nearby in all directions. Lucas (2001) provides some evidence of the need of high δs to fit Japanese land rent data.

Notice that for all κ values used in the figure, a higher δ results in a business sector that is more concentrated. The intuition for this result is clear. A higher δ results in more localized external effects and so, in order to take advantage of the externality, the optimal solution yields a more concentrated business sector. For the case of $\kappa = 0.01$, the concentration of business areas may eventually lead to a business sector in the middle, nevertheless we believe that values of δ much higher than 15 are not reasonable for a one sector model like this.[9] If we were to include different sectors, we may use a very large δ for an industry like financial services (see Dekle and Eaton, 1999, for some evidence on this), which may result in this industry locating at the center.

5.1. Numerical results on optimal subsidies

In Fig. 8, we present the optimal marginal product of labor and the function λ for the different κ values. The subsidy $\tau(r)$ is given by the difference between the two curves. The subsidy is never negative, it is actually a subsidy. The reason is that the externality faced by firms is always positive, so in equilibrium they always hire less workers than in the optimum. The total subsidy -the area between the two curves- is larger the larger κ. This

[9] If $\delta = 5$, the external effect declines by half every 300 yards.

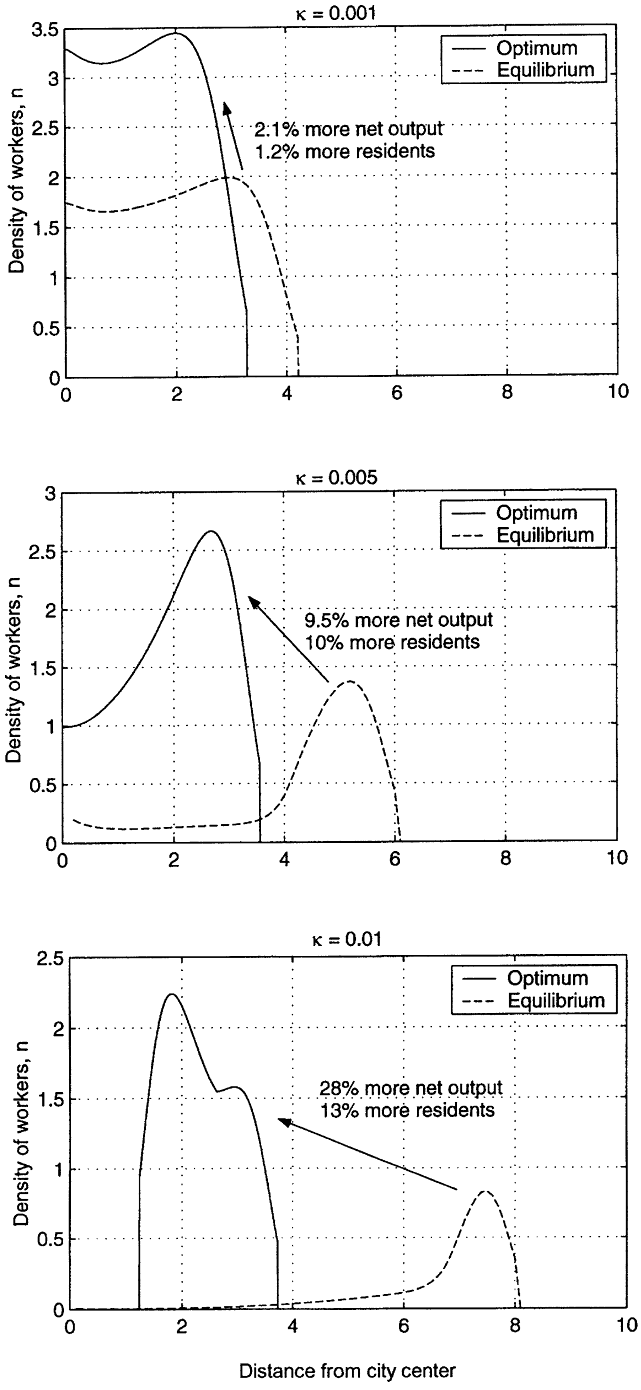

Fig. 6. Optimal density of workers.

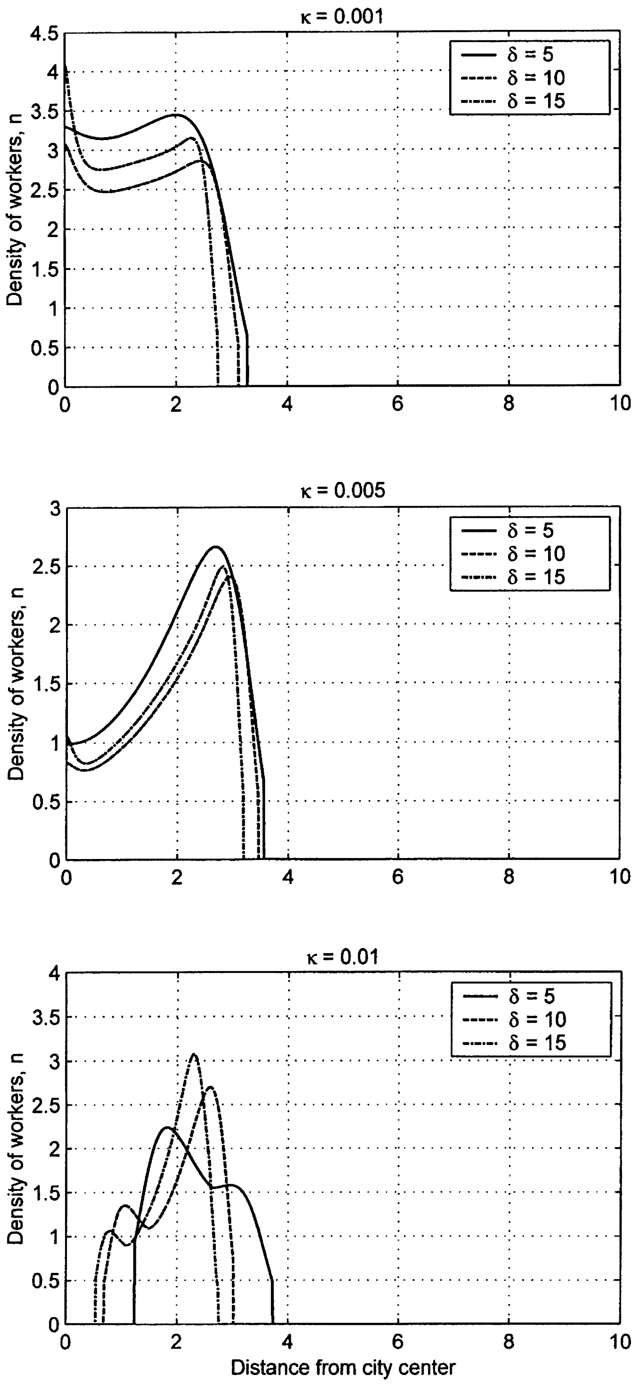

Fig. 7. The effect of δ.

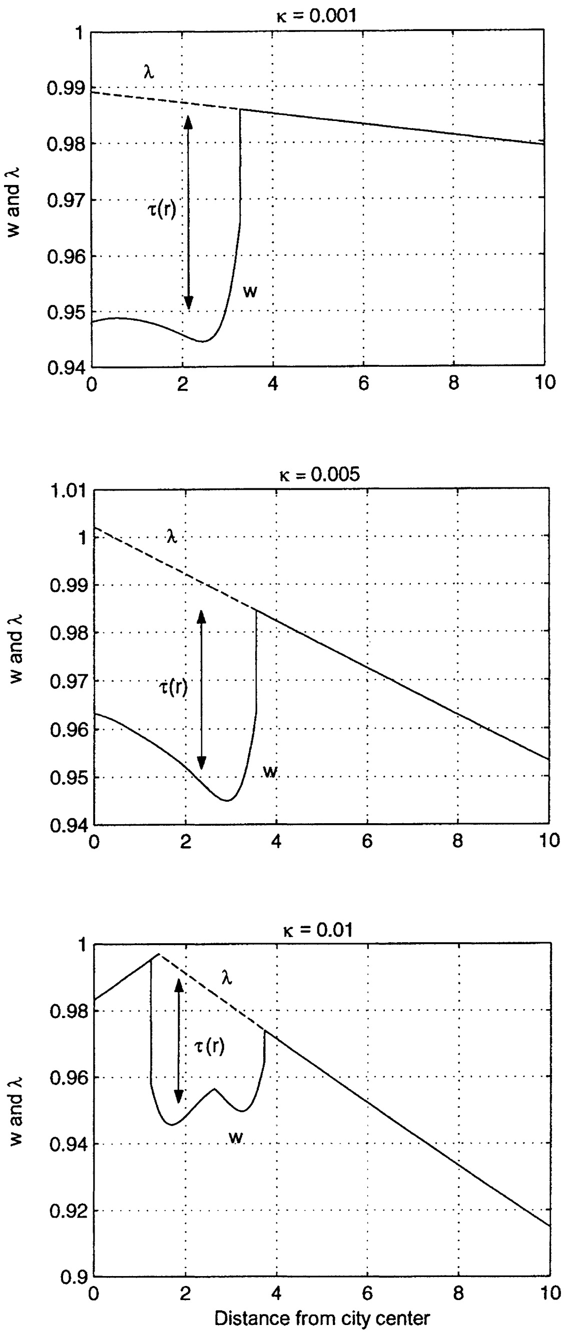

Fig. 8. Labor subsidy.

is true in absolute value but not as a percentage of total net output. We saw before that the gains from this type of subsidy are larger, the larger κ.

There are two effects that determine the total bill of subsidies. On one hand, the higher κ, the more important the difference between the equilibrium and optimum land use structure, which increases the total subsidy. On the other, the external effect is lower the higher κ, and so the subsidy needed to internalize the externality is smaller. The result of this trade-off is the area depicted in these figures.

5.2. Numerical results on zoning restrictions

In Section 4, we discussed some of the theoretical implications of imposing zoning restrictions. There are two important issues that we want to address using numerical exercises. The first is the extend in which zoning laws can reduce the difference between equilibrium and optimal net output. The second is the discontinuity of land rents at the boundaries between business and residential areas.

We calculate the equilibrium with zoning restrictions for the three sets of parameter values we have used thus far. With regard to the first question, for $\kappa = 0.001$, with zoning laws increase net output by 2.48% with respect to the equilibrium and imply 5.49% less net output than in the optimum. For $\kappa = 0.005$, the equilibrium allocation with zoning results in 6.83% more net output than in equilibrium and 5.51% less than in the optimum. The corresponding numbers for $\kappa = 0.01$ are 21.58% and 6.70%. These numbers show that the higher κ, the larger the gains from using zoning laws. The effectiveness of zoning increases with commuting costs. The relative difference in net output between the efficient allocation and the zoning equilibrium also increases slightly with κ.

Figure 9 shows employment densities (multiplied by θ) at the optimum and zoning equilibrium, as well as optimal values of land and land rents in the zoning equilibrium. For $\kappa = 0.001$ and $\kappa = 0.005$, the optimal land use structure is monocentric, so land rents are discontinuous at one location. For $\kappa = 0.01$ land rents are discontinuous at two locations (we have two residential sectors, one at the center and the other at the boundary of the city). One interesting feature of these figures is that land rents and densities are always higher near the sector boundaries in the zoning equilibrium. The reason is that, on one hand, these locations are low productivity (external effects are low), so the effect of internalizing the externality is low. On the other hand, housing extra workers is more costly in goods at the optimum, since there are more residents so residential land plots are smaller than in equilibrium and the reservation utility is fixed. At the boundary, the second effect dominates so densities in the zoning equilibrium are higher. However, productivity is larger in the efficient allocation than in the zoning equilibrium everywhere in the city.

6. Conclusions

Let us try to summarize the characterization of the optimal allocation. When commuting costs are low, the center of the city is used for business purposes and the rest of the city is residential. The density of employment at the center is relatively high. As we increase commuting cost, we obtain lower densities of employment at the center, but the city

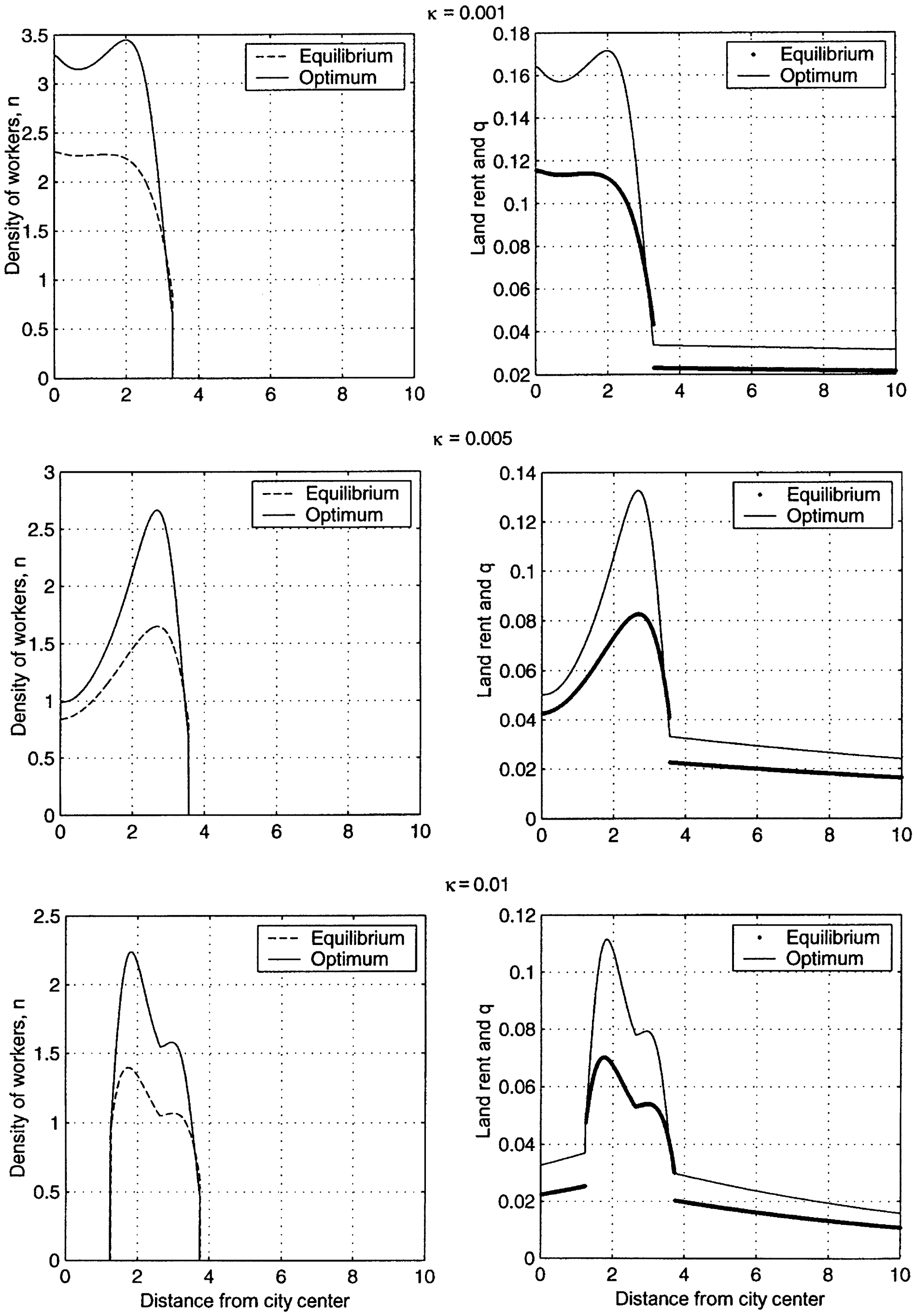

Fig. 9. Optimum and equilibrium with zoning restrictions.

remains a Mills city. Further increases in commuting costs result in a residential area at the center surrounded by a ring of business areas and then a ring of residential areas. In contrast, in the equilibrium allocations presented in LRH, low commuting cost imply a Mills city with high densities of employment at the center. As commuting costs increase, the center becomes mixed, with pure business and residential areas surrounding it.

Efficient and equilibrium allocations differ since we use production externalities as the agglomeration force. Other theories in urban economics have proposed agglomeration effects that are not external but internal to the firm. The allocations characterized in this paper can be viewed as *equilibrium* allocations of urban theories without externalities but with agglomeration effects that take the form in (7). However, the validity of the policies proposed does depend crucially on the presence of external effects.

Theorem 1 shows that there are no mixed areas in the efficient allocation. In the equilibrium allocation, however, there are mixed areas at the center of the city. Hence, we should observe mixed areas in actual cities, but those areas are inefficient according to our model. Chicago is a good example of mixed areas at the center. Chicago residents commuting by foot to work live mostly in downtown Chicago.

As we argued above, the results in this paper can be viewed as the equilibrium allocation of a theory with internal agglomeration effects. Theorem 1 then implies that, if agglomeration effects are internal to the firm, the equilibrium allocation has no mixed areas. Hence, the presence of mixed areas in actual cities is evidence of production externalities.

The theory emphasizes that there is an indirect effect, via changes in land use, of reducing commuting costs. The numerical exercises show that the effect is potentially important. City governments should try to include estimations of this gain in their evaluation of projects and policies that reduce commuting costs.

Theorem 2 shows that using labor subsidies the optimal allocation can be implemented as an equilibrium. Other policies like zoning restrictions do not implement the optimum but do improve the production possibilities of a given city. The numerical exercises show that zoning restriction are much more useful when commuting costs are high.

As Fig. 9 shows, land rents are discontinuous in the presence of zoning restrictions. Firms would like to move to residential areas but they are not allowed. The reason is that optimal zoning restrictions always *concentrate* business areas in order to increase productivity. Hence, with optimal zoning restrictions we know that firms, and *not* residents, will have incentives to lobby for the elimination of zoning restrictions. Therefore, our provides a test of the optimality of current regulation.

Given our focus on optimal allocations, our theory does not provide empirical implications on land use structure. However, the parallel equilibrium theory does. It is important to test the implications of our theory since we have made several policy recommendations. Their validity depends directly on the extent to which this model captures important elements of reality. A test of these implications is, nevertheless, beyond the scope of this paper and left for future research.

Acknowledgments

I am particularly grateful to Robert E. Lucas Jr. for his support, guidance and advice. I also thank Fernando Alvarez, Marcus Berliant, Jan Brueckner, Jeff Campbell, Guillaume Carlier, Ivar Ekeland, Lars P. Hansen, Boyan Jovanovic, Nancy Stokey, Mehmet Yorukoglu and seminar participants at Boston University, Brown University, UCSB, the University of Chicago, the University of Michigan, the Minneapolis FED, the SED Meetings in Stockholm, Stanford University, the University of Toronto, Vanderbilt University, the editors of this journal, and an anonymous referee for helpful comments. The paper benefitted from discussions with Carlos da Costa, Claudio Irigoyen, Juan Santalo, Rodrigo Soares, and Ivan Werning.

Appendix A. Proofs

In this appendix we prove all the theorems, propositions and lemmas not proved in the different sections of the paper. The order is the same as the order in which the results are presented in the text.

Proposition 1. *Under Assumption A, given θ there exists a unique set of functions $\{n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ that satisfies the first-order conditions and the Maximum Principle conditions.*

Proof. Under Assumption A, given θ, the problem is a strictly concave problem with constraints that define a closed, convex and bounded set in $\{n^*, N^*, c^*, \ell^*, z^*, H^*\}$. To see this, notice that, since the total amount of land is given by $2\pi S^2$, part (iii) of Assumption A guarantees that $0 < \ell(r) < 2\pi S^2$ and so that $c(r)$ is strictly positive and bounded. $N(r)$ is then also strictly positive and bounded by (5). Assumption A part (ii) together with (6) result in $n(r)$ strictly positive and bounded. Since all these properties hold for every $r \in [0, S]$, they imply that $z(r)$ and $H(r)$ are positive and bounded and that the functions $n^*, N^*, c^*, \ell^*, z^*, H^*$ also inherit these properties. Hence by Mangasarian Sufficiency Theorem there exists a unique solution to Problem (O), given θ. The first-order conditions and Maximum principle conditions are then necessary and sufficient, so there exists a solution that satisfies them given θ. □

Proposition 2. *Under Assumption A, there exists a function θ^* such that $T\theta^* = \theta^*$.*

Proof. Since clearly T does not have to be monotone for all sets of parameters values, we cannot base the proof on convergence of monotone operators. We will use Schauder's fixed point theorem in the version proven in Zeidler (1986, Vol. 1). We need to prove that the operator T maps a nonempty, compact, convex subset of a Banach space into itself and that T is a continuous operator in that same space. For this, first define the space $L_1(\mathbb{R})$, a Lebesgue space, by the set of all measurable functions $f : \mathbb{R} \to \mathbb{R}$ with $\|f\|_1 < \infty$, where

$$\|f\|_1 = \int_{\mathbb{R}} \left| f(x) \right| \mathrm{d}x.$$

Zeidler (1986, Vol. 1) provides a proof that $L_1(\mathbb{R})$ is a Banach space, since clearly there are two functions, elements of $L_1(\mathbb{R})$, that differ only by a set of measure 0. Next, define M as the set of all measurable functions $f:[0,S]\to[0,1]$ with bounded variation. Since $\|f\|_1 < S < \infty$, $M \subset L_1(\mathbb{R})$, M is bounded and M is clearly nonempty. To show that M is convex, notice that for $f_1, f_2 \in M$ and $\alpha, \beta \in [0,1]$ with $\alpha+\beta = 1$, $0 < \alpha f_1(x) + \beta f_2(x) \equiv f_3(x) < 1$, all $x \in [0,S]$. Thus, since the set of bounded variation functions on a given interval is a linear space (see, for example, Kolmogorov and Fomin, 1975), $f_3 \in M$.

We now prove that M is a compact subset of $L_1(\mathbb{R})$. For this we use the compactness theorem of Riesz–Kolmogorov (see Zeidler, 1986, Vol. 2B). The theorem states that a bounded subset of $L_1(\mathbb{R})$ (say M) is relatively compact if and only if it is 1-mean continuous, i.e., for each $\varepsilon > 0$, there exists a $\vartheta(\varepsilon) > 0$ such that

$$\sup_{u\in M} \int_{\mathbb{R}} \left|u(x+h) - u(x)\right| \mathrm{d}x < \varepsilon,$$

provided $|h| < \vartheta(\varepsilon)$. We set $u(x) = 0$ outside S.

We have shown that M is a bounded subset of $L_1(\mathbb{R})$. Notice that if the function u that achieves the supremum is continuous, the condition is trivially satisfied. Without loss of generality assume that the function u that achieves the supremum is an indicator function. Then, if the number of discontinuity points in u is finite (say $D < \infty$),

$$\sup_{u\in M} \int_{\mathbb{R}} \left|u(x+h) - u(x)\right| \mathrm{d}x \leqslant hD < \varepsilon$$

for $h < \varepsilon/D$. If the number of discontinuity points is infinite, then the set S^D at which $|u(x+h) - u(x)| = 1$ is of measure 0 since u has finite variation by Assumption A, part (v). Hence M is relatively compact. Since both $[0,S]$ and $[0,1]$ are closed intervals, and lot sizes are bounded below by Assumption A, part (v), the set M is also closed and hence M is compact.

$T: M \to M$, follows from the definition of T and the fact that if θ has bounded variation, λ and T are continuous and have bounded variation, which implies that $T\theta$ has bounded variation. Thus, the first set of conditions has been proven.

We now turn to the proof that T is a continuous operator in $\|\cdot\|_1$. That is, we need to prove that if a sequence $\{\theta_i\}_{i=1}^{\infty}$ is such that

$$\lim_{i\to\infty} \int_0^S \left|\theta_i(r) - \theta(r)\right| \mathrm{d}r = 0,$$

it implies that

$$\lim_{i\to\infty} \int_0^S \left|T\theta_i(r) - T\theta(r)\right| \mathrm{d}r = 0.$$

In particular, because of the definition of T, we need to show that if $\theta_i(r) \to \theta(r)$ in $\|\cdot\|_1$, $\lambda(\cdot;\theta_i) \to \lambda(\cdot;\theta)$ and $t(\cdot;\theta_i) \to t(\cdot;\theta)$ where this convergence can be in the sup-norm ($\|\cdot\|$), (here we use a notation that emphasizes the dependence on the function θ).

That is, we need to show that λ and t are continuous in θ in $\|\cdot\|$. To show that this is sufficient, notice that if this is the case, $t-\lambda$ will be continuous and hence an indicator function that is equal to 1 for r's such that $t(r)-\lambda(r)\geqslant 0$ and equal to 0 for r's such that $t(r)-\lambda(r)>0$ will be continuous in $\|\cdot\|_1$. This, since the r's such that $t(r)-\lambda(r)=0$ will change continuously with θ.

Showing that λ and t are continuous in θ is equivalent to showing that the system of first-order and Maximum Principle conditions for θ_i converge to the system for θ in the sup-norm. We will do that by showing that all the terms that involve θ in the system converge. θ appears in three terms in the system, we will analyze each term in turn. For the first term,

$$
\begin{aligned}
&\Biggl\| \int_0^S s\theta_i(s) g'\bigl(z(s;\theta_i)\bigr) f\bigl(n(s;\theta_i)\bigr)\varphi(s,r)\,\mathrm{d}s \\
&\quad - \int_0^S s\theta(s) g'\bigl(z(s;\theta)\bigr) f\bigl(n(s;\theta)\bigr)\varphi(s,r)\,\mathrm{d}s \Biggr\| \\
&\quad \leqslant S2\pi\delta \max_{s\in S}\bigl[g'\bigl(\max\bigl[z(s;\theta_i),z(s;\theta)\bigr]\bigr) f\bigl(\max\bigl[n(s;\theta_i),n(s;\theta)\bigr]\bigr)\bigr] \\
&\qquad \times \Biggl\| \int_0^S \bigl[\theta_i(s)-\theta(s)\bigr]\,\mathrm{d}s \Biggr\| \\
&\quad \leqslant S2\pi\delta \max_{s\in S}\bigl[g'\bigl(\max\bigl[z(s;\theta_i),z(s;\theta)\bigr]\bigr) f\bigl(\max\bigl[n(s;\theta_i),n(s;\theta)\bigr]\bigr)\bigr] \\
&\qquad \times \int_0^S \bigl|\theta_i(s)-\theta(s)\bigr|\,\mathrm{d}s,
\end{aligned}
$$

where we are using the fact that for every $\theta\in M$, $z(s;\theta)$ and $n(s;\theta)$ are finite (see the proof of Proposition 1), that $\varphi(s,r)\in[0,2\pi\delta]$ and $s\leqslant S$.

For the second term,

$$
\begin{aligned}
&\Biggl\| \int_0^S \varphi(r,s)s\theta_i(s)n(s;\theta_i)\,\mathrm{d}s - \int_0^S \varphi(r,s)s\theta(s)n(s;\theta)\,\mathrm{d}s \Biggr\| \\
&\quad \leqslant S2\pi\delta \max\bigl[n(s;\theta_i),n(s;\theta)\bigr] \Biggl\| \int_0^S \bigl[\theta_i(s)-\theta(s)\bigr]\,\mathrm{d}s \Biggr\| \\
&\quad \leqslant S2\pi\delta \max\bigl[n(s;\theta_i),n(s;\theta)\bigr] \int_0^S \bigl|\theta_i(s)-\theta(s)\bigr|\,\mathrm{d}s.
\end{aligned}
$$

For the third term, notice that

$$\begin{aligned}
&\sup_{s\in[0,S]}\big|2\pi s\big[\theta_i(s)n(s;\theta_i)+\big(1-\theta_i(s)\big)N(s;\theta_i)\big]\\
&\quad-2\pi s\big[\theta_i(s)n(s;\theta)+\big(1-\theta_i(s)\big)N(s;\theta_i)\big]\big|\\
&\quad\leqslant 2\pi s\sup_{s\in[0,S]}\max\big[n(s;\theta_i)-n(s;\theta),N(s;\theta_i)-N(s;\theta)\big],
\end{aligned}$$

which does not involve θ_i or θ explicitly. So since the other two terms converge, n and N are continuous functions of θ, and so the term above can be made arbitrarily small by choice of i. Hence t and λ are continuous functions of θ and so by the argument above, T is a continuous operator. Schauder's fixed point theorem yields the result. □

Proposition 3. *Under Assumption A, there exists a set of functions* $\{\theta^*, n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ *that satisfies the first-order conditions, the Maximum Principle conditions and* $T\theta^* = \theta^*$.

Proof. By Proposition 1, there exists a unique solution that satisfies the first-order and Maximum Principle conditions given θ. Proposition 2 shows the existence of a θ^* such that $T\theta^* = \theta^*$. Hence there exists a set of functions $\{\theta^*, n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ that satisfies all the necessary conditions for a maximum. □

Theorem 1. *Under Assumption A, the optimal land use structure has no mixed areas. That is* $\theta^*(r) \in \{0, 1\}$, *except for sets with zero Lebesgue measure in* $[0, S]$.

Proof. For a mixed sector, we need

$$H(r)=0,\qquad \theta(r)n(r)-\big(1-\theta(r)\big)N(r)=0,\quad\text{and}\quad t(r)=\lambda(r).$$

For any $r > 0$ and $\varepsilon > 0$ such that $H(r-\varepsilon) \neq 0$, given the value of $\lambda(r)$ and the value of $\chi(r)$, plus the first-order condition with respect to $n(r)$ and $N(r)$, the system is over-identified ($\lambda(r)$ follows one of the differential equations given by the Maximum Principle). The reason is that we are imposing two extra conditions, $H(r) = 0$ and $t(r) = \lambda(r)$. The third, $\theta(r)n(r) - (1-\theta(r))N(r) = 0$ determines $\theta(r)$. One condition could be satisfied by determining the value $\lambda(0)$ but we are still left with one condition. This equation can be satisfied *only* for a given value of $\lambda(r)$ and $\chi(r)$, which implies that except for sets of measure zero in the parameter space, $t(r) = \lambda(r)$ or $H(r) = 0$ cannot be satisfied and so mixed areas cannot be part of the solution. Notice that the argument above holds *only* for mixed areas with positive Lebesgue measure, since for sets with zero Lebesgue measure $\theta(r)n(r) - (1-\theta(r))N(r) = 0$ does not have to hold and so we loose one of the equations.

There is a special case we still need to analyze. Since $H(0) = 0$, by (6), it could be the case that we start with a mixed area. Suppose that r is a mixed area, in order for the three conditions above to hold, $\lambda(r)$ will be determined by $\lambda(r) = t(r)$. $\theta(r)$ is then given by

$$\theta(r)=\frac{N(r)}{N(r)+n(r)}.$$

Notice that if $\lambda(r) = t(r)$, the value of $\bar{L}(r, \theta(r) = 1) = \bar{L}(r, \theta(r) = 0)$ where $\bar{L}$ is the Lagrangian evaluated at the maximizing values of n, N, z, c, λ and H given θ.

Differentiating $\lambda(r) = t(r)$ with respect to $\theta(r)$ we obtain (omitting the dependence on r in the notation)

$$\frac{\partial^2 \bar{L}(r, \theta(r))}{\partial \theta(r)^2} = g'(z) f(n) \frac{\partial z}{\partial \theta} + n \frac{\partial \bar{\chi}}{\partial \theta},$$

where $\bar{\chi} = \chi / n$ and we used the first-order conditions to simplify the term. Differentiating the first-order condition with respect to n we obtain

$$g'(z) f'(n) \frac{\partial z}{\partial \theta} + g(z) f''(n) \frac{\partial n}{\partial \theta} + \frac{\partial \bar{\chi}}{\partial \theta} = 0, \tag{A.1}$$

and substituting for $\partial \bar{\chi} / \partial \theta$ in the equation above we obtain that

$$\frac{\partial^2 \bar{L}(r, \theta(r))}{\partial \theta(r)^2} = g'(z) f(n) \left[\bar{k} \left(n + \theta \frac{\partial n}{\partial \theta} \right) \right] \left[f(n) - n f'(n) \right] - g(z) f''(n) n \frac{\partial n}{\partial \theta}.$$

Where we used the fact that

$$\frac{\partial z}{\partial \theta} = \left[\bar{k} \left(n + \theta \frac{\partial n}{\partial \theta} \right) \right], \tag{A.2}$$

where $\bar{k}$ is a non-negative and continuous function of r, with $\bar{k}(0) = 0$. Notice that by Assumption A, f is strictly concave function and so the Lagrangian is a strictly convex function of θ if

$$\frac{\partial n}{\partial \theta} > 0.$$

Notice that since $\bar{L}(r, \theta(r) = 1) = \bar{L}(r, \theta(r) = 0)$, this implies that

$$\bar{L}(r, \theta = 0, 1) > \bar{L}(r, \bar{\theta}) \quad \text{for all } \bar{\theta} \in (0, 1).$$

So the maximum of the Lagrangian is attained when $\theta = 0$ or 1, a contradiction with the assumption that r was a mixed area. This, together with the first part of the proof, yields the result.

We still need to prove that $\partial n / \partial \theta > 0$. For this notice that

$$\frac{\partial \bar{\chi}}{\partial \theta} = \bar{k} \left[\theta g''(z) f(n) \frac{\partial z}{\partial \theta} + g'(z) f(n) + \theta g'(z) f'(z) \frac{\partial n}{\partial \theta} \right].$$

Substituting this in (A.1) and using (A.2), we get

$$\frac{\partial n}{\partial \theta} = \frac{\bar{k}[g'(z) f(n) + \bar{k} \theta g''(z) f(n) n]}{-[\theta^2 \bar{k}^2 g''(z) f(n) + 2\theta \bar{k} g'(z) f'(n) + g(z) f''(z)]}.$$

Notice that the term in brackets in the denominator is the quadratic form of the Hessian and by Assumption A is negative. Hence

$$\partial n / \partial \theta > 0 \quad \text{if } g'(z) > -\bar{k} \theta g''(z) n.$$

Since $\bar{k}(0) = 0$ and $\bar{k}$ is continuous, for r sufficiently close to 0 the condition above is satisfied. So it cannot be the case that we have a mixed area of positive Lebesgue measure at the center.

Notice that if $g(\cdot)$ is constant, $\partial n/\partial\theta = 0$ for all $\theta \in [0, 1]$ and if $g(\cdot)$ is linear then the condition above is clearly satisfied since $g''(z) = 0$. A remark is that in the equilibrium case in LRH since the agents do not control z, $\partial^2 \bar{L}(r, \theta(r))/\partial\theta(r)^2 = 0$ and so mixed areas are possible. □

Proposition 4. *Under Assumption A, there exists a set of functions* $\{\theta^*, n^*, N^*, c^*, \ell^*, z^*, H^*, \lambda^*\}$ *that solves Problem* (*O*)*.*

Proof. We use relaxation theory to prove the Theorem (see Ekeland and Turnbull, 1983). Let $\hat{L}(r, \theta(r), \lambda(r))$ be the Lagrangian defined in Section 2.2 once we substitute the functions $n, N, c, \ell, z,$ and H all as functions of θ and λ. The existence and uniqueness of these functions is proven in Proposition 1. Let

$$\tilde{L}\big(r, \theta(r), \lambda(r)\big) \equiv \min_{f:[0,S]\times[0,1]\times\mathbb{R}\to\mathbb{R}} \sup\big|f\big(r, \theta(r), \lambda(r)\big) - \hat{L}\big(r, \theta(r), \lambda(r)\big)\big|$$

$$\text{s.t.} \quad f\big(r, \theta(r), \lambda(r)\big) \geqslant \tilde{L}\big(r, \theta(r), \lambda(r)\big) \quad \text{and} \quad f\big(r, \cdot, \lambda(r)\big) \text{ concave.}$$

Then

$$\big\{x\colon\ x \in \big[0, \tilde{L}\big(r, \theta(r), \lambda(r)\big)\big] \text{ for some } r \in [0, S]\big\}$$

is the convex hull of

$$\big\{x\colon\ x \in \big[0, \hat{L}\big(r, \theta(r), \lambda(r)\big)\big] \text{ for some } r \in [0, S]\big\}.$$

Hence, for every r, there exists a $\tilde{\theta}(r)$ (maybe many) such that

$$\tilde{\theta}(r) = \operatorname*{argmax}_{\theta} \tilde{L}\big(r, \theta, \lambda(r)\big).$$

Define

$$\hat{\theta}(r) = \operatorname*{argmax}_{\theta} \hat{L}\big(r, \theta, \lambda(r)\big).$$

We need to show that for almost every r, $\tilde{\theta}(r) = \hat{\theta}(r)$, since then

$$\int_0^S \tilde{L}\big(r, \tilde{\theta}(r), \lambda(r)\big)\,\mathrm{d}r = \int_0^S \hat{L}\big(r, \hat{\theta}(r), \lambda(r)\big)\,\mathrm{d}r, \tag{A.3}$$

and so $\tilde{\theta}(r)$ is a solution to Problem (O).

If $\hat{L}(r, 1, \lambda(r)) \neq \hat{L}(r, 0, \lambda(r))$, then $\tilde{L}(r, \theta, \lambda(r)) = \hat{L}(r, \theta, \lambda(r))$ and so $\tilde{\theta}(r) = \hat{\theta}(r)$. When $\hat{L}(r, 1, \lambda(r)) = \hat{L}(r, 0, \lambda(r))$ and so $\tilde{L}(r, 1, \lambda(r)) - \tilde{L}(r, 0, \lambda(r))$, the problem may arise. It then implies that $\tilde{\theta}(r) \subseteq [0, 1]$ and may be a set with positive length. Hence, it may be the case that $\tilde{\theta}(r) \neq \hat{\theta}(r)$. This is a problem if $\hat{L}(r, 1, \lambda(r)) = \hat{L}(r, 0, \lambda(r))$ for some set of positive Lebesgue measure since then $\hat{\lambda}(s) \neq \tilde{\lambda}(s)$, $s > r$, which results in $\tilde{\theta}(s) \neq \hat{\theta}(s)$ for all $s > r$ (where $\hat{\lambda}(\cdot)$ and $\tilde{\lambda}(\cdot)$ are the functions resulting from $\hat{\theta}(\cdot)$ and $\tilde{\theta}(\cdot)$ respectively). But since we showed in Theorem 1 that $\hat{L}(r, 1, \lambda(r)) = \hat{L}(r, 0, \lambda(r))$ only for sets of zero Lebesgue measure, $\tilde{\theta}(r) \neq \hat{\theta}(r)$ only for sets of zero Lebesgue measure and so $\tilde{L}(r, \tilde{\theta}(r), \lambda(r)) \neq \hat{L}(r, \hat{\theta}(r), \lambda(r))$ only for these sets. This implies that the equality in (A.3) is satisfied. □

Lemma 1. *If U, f and g are Cobb–Douglas as specified above,*

$$\alpha, \beta, \gamma > 0, \qquad \beta, \alpha < 1, \quad \textit{and} \quad 1 - \alpha > \gamma,$$

then Assumption A is satisfied.

Proof. The result is obvious except for part (ii). We need to show that

$$\left[\int_0^S \varphi(r,s) s\theta(s) n(s)\, \mathrm{d}s \right]^\gamma n(r)^\alpha \quad \text{is concave in } n$$

for any θ. First notice that

$$z(r) = \int_0^S \varphi(r,s) s\theta(s) n(s)\, \mathrm{d}s \quad \text{is linear in } n.$$

So we need to show that

$$Y(z,n) \equiv z^\gamma n^\alpha \quad \text{is concave in } z \text{ and } n.$$

The second and cross derivatives are given by:

$$\frac{\partial^2 Y(z,n)}{\partial z^2} = \gamma(\gamma - 1) z^{\gamma-2} n^\alpha < 0, \qquad \frac{\partial^2 Y(z,n)}{\partial n^2} = \alpha(\alpha - 1) z^\gamma n^{\alpha-2} < 0,$$
$$\frac{\partial^2 Y(z,n)}{\partial n \partial z} = \alpha\gamma z^{\gamma-1} n^{\alpha-1} > 0.$$

The determinant of the Hessian is then given by

$$\gamma(\gamma-1)\alpha(\alpha-1) z^{2\gamma-2} n^{2\alpha-2} - \alpha^2\gamma^2 z^{2\gamma-2} n^{2\alpha-21}$$
$$= z^{2\gamma-2} n^{2\alpha-2}\big(\gamma(\gamma-1)\alpha(\alpha-1) - \alpha^2\gamma^2\big) = z^{2\gamma-2} n^{2\alpha-2} \alpha\gamma(-\gamma - \alpha + 1) > 0.$$

Hence the determinant of the first minor is negative and the determinant of the second is positive under the assumptions on α and γ in the statement. This implies that $Y(z,n)$ is concave in n and z and so part (ii) of Assumption A is satisfied. □

References

Anas, A., Arnott, R., Small, K., 1998. Urban spatial structure. Journal of Economic Literature 36, 1426–1464.

Anas, A., Xu, R., 1999. Congestion, land use, and job dispersion: a general equilibrium model. Journal of Urban Economics 45, 451–473.

Berliant, M., Peng, S., Wang, P., 2002. Production externalities and urban configuration. Journal of Economic Theory 104, 275–303.

Borukhov, E., Hochman, O., 1977. Optimum and market equilibrium in a model of a city without a predetermined center. Environment and Planning A 9, 849–856.

Caselli, F., Coleman II, W., 2001. The US structural transformation and regional convergence: a reinterpretation. Journal of Political Economy 109, 584–616.

Ciccone, A., Hall, R., 1996. Productivity and the density of economic activity. American Economic Review 86, 54–70.

Ciccone, A., Peri, G., 2002. Identifying human capital externalities: theory with an application to US cities. UPF working paper 661.

Dekle, R., Eaton, J., 1999. Agglomeration and land rents: evidence from the prefectures. Journal of Urban Economics 46, 200–214.

Dixit, A., 1973. The optimum factory town. Journal of Economics and Management 4, 637–651.

Ekeland, I., Turnbull, T., 1983. Infinite-Dimensional Optimization and Convexity. Univ. of Chicago Press, Chicago.

Ellison, G., Glaeser, E.L., 1997. Geographic concentration in US manufacturing industries: a dartboard approach. Journal of Political Economy 105, 889–927.

Fujita, M., Krugman, P., Venables, A., 1999. The Spatial Economy. MIT Press, Cambridge.

Fujita, M., Ogawa, H., 1982. Multiple equilibria and structural transition of non-monocentric urban configurations. Regional Science and Urban Economics 12, 161–196.

Fujita, M., Thisse, J., 2002. Economics of Agglomeration. Cambridge Univ. Press.

Glaeser, E., Kallal, H., Scheinkman, J., Shleifer, A., 1992. Growth in cities. Journal of Political Economy 100, 1126–1152.

Henderson, V., 2001. Marschall's scale economies. NBER working paper 7358.

Helpman, E., Pines, D., 1977. Land use and zoning in an urban economy: further results. American Economic Review 67, 982–986.

Hochman, O., Pines, D., 1982. Adjustment costs and the spatial pattern of a growing open city. Econometrica 6, 1371–1391.

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 63, 577–598.

Kolmogorov, A., Fomin, S., 1975. Introductory Real Analysis. Dover, New York.

Lucas, R.E. Jr., 2001. Externalities and cities. Review of Economic Dynamics 4, 245–267.

Lucas, R.E. Jr., Rossi-Hansberg, E., 2002. On the internal structure of cities. Econometrica 70 (4), 1445–1476.

Mills, E., 1967. An aggregate model of resource allocation in a metropolitan area. American Economic Review 57, 197–210.

Moretti, E., 2002. Human capital spillovers in manufacturing: evidence from plant-level production functions. NBER working paper 9316.

Rauch, J., 1993. Productivity gains from geographic concentration of human capital: evidence from the cities. Journal of Urban Economics 24, 380–400.

Roback, J., 1982. Wages, rents, and quality of life. Journal of Political Economy 90, 1257–1278.

Stull, W., 1974. Land use and zoning in an urban economy. American Economic Review 64, 337–347.

Wheaton, W., 1998. Land use and density in cities with congestion. Journal of Urban Economics 43, 258–272.

Zeidler, E., 1986. Nonlinear Functional Analysis and its Applications. Vols. 1, 2B. Springer-Verlag, New York.