FACULTY OF LAW, ECONOMICS AND FINANCE

CREA Discussion Paper 2019-03

Economics

Center for Research in Economics and Management
University of Luxembourg

Green Urban Areas

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers

Pierre M. Picard, CREA, Université du Luxembourg
Thi Thu Huyen Tran, CREA, Université du Luxembourg

March, 2019

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
162A, avenue de la Faïencerie
L-1511 Luxembourg

UNIVERSITÉ DU LUXEMBOURG

Green Urban Areas

Picard P.M. and Tran T.T.H. [*†]

March 8, 2019

Abstract

This paper studies the size and location of urban green areas across city spaces. Urban green areas offer amenities that affect residential choices, land consumption and land rent. This paper discusses the socially optimal sizes and locations of urban green areas within a city and their decentralized allocation through land markets. The main result is that the share of land dedicated to urban green areas is a concave function of the distance to the city center. This result is confirmed by the empirical study of urban structures in the 305 largest EU cities. The importance of urban green areas is finally assessed by a counterfactual analysis, where 50% of urban green areas are removed in each city.

Keywords: Urban green areas, urban spatial structure, land use policy, amenities, optimal locations, public facilities, structural estimation

JEL Classification: C61, D61, D62, R14, R53

[*]Department of Economics, Center for Research in Economic Analysis, University of Luxembourg. Rue de la Faiencerie, 162A, 1510 Luxembourg. Correspondence to Huyen Tran <thithuhuyen.tran@uni.lu>.

[†]We would like to thank Jan Brueckner, Caruso Geoffrey, Michael Pfluger, Jos van Ommeren and all participants at UEA 7^{th} European Meetings, UEA 13^{th} North American Meetings, 64^{th} Annual North America Meetings of the Regional Science Association, 57^{th} ERSA congress, ITEA Annual Conference 2017, Cities and the Environment Workshop at the University of Potsdam 2018, Eureka Seminar at VU Amsterdam, and DSEF Jamboree at the University of Luxembourg for useful comments and suggestions. All our apologies for omissions and errors.

1 Introduction

Urban green areas play a crucial role in the debate on sustainable cities. They are an important part of any urban area whose quality and quantity are prime concerns for environmental sustainability. Recent research has confirmed the relationship between urban parks and the well-being of city's residents. Brack (2002) and Strohback et al. (2012) find a strong influence of urban natural ecosystems in reducing air and noise pollution and CO_2 absorption in Australia and Germany. Heidt and Neef (2008) suggest economic benefits by showing a significant increase in the nearby property values. Access to nearby urban parks helps reduce stress and improve psychological well-being[1] and increasing physical activities.[2]

In this paper, we study the geographical distribution of green urban areas in cities and compare our theoretical results with empirical observations of green urban areas. More precisely, we study the optimal level of green urban areas in urban spaces. Because green urban areas are land-intensive and offer very localized amenities, we find that the relationship between the share of surface devoted to green urban areas is a concave function of the distance to the city center. On the one hand, the opportunity cost of land is too expensive in the city center for the planner to implement many green areas there. On the other hand, residents are too sparse at the city edges to give planners the incentives to invest in such very localized amenities. We confirm our results using the European Environment Agency's Urban Atlas data on the land use of 305 EU cities with more than 100,000 inhabitants. These data describe land use and cover across Europe using harmonized Earth Observations (EOs), which are combined with Eurostat Urban Audit statistical data. The data represent a unique source of reliable and comparable European urban planning data. As far as we know, this is the first paper that uses both a theoretical model and empirical estimation of European urban land use. This result is robust to many variations of the land use specifications, city structure specifications, and city and country characteristics.

We further estimate residential land use and use the estimated parameters to study the value of green urban areas. To assess the value of green urban areas, we study counterfactual exercises, where fifty percent of urban green areas are removed. We show that open cities lose more than 6% of their population if those areas are left unused. The total loss for landlords is approximately €150 million per city if green areas are not converted into residential land. Converting those urban green areas into residential land, however, would increase residential

[1]See Ernstson (2012), Woo et al. (2009), Chiesura, A. (2004), IFPRA (2013)
[2]See Cohen at al. (2006, 2007) and Evenson et al. (2013)

surfaces and increase the total housing market value by approximately €50 million. In closed cities, residents lose utility, which can be restored by a subsidy of nearly one-tenth of their net income. This exercise, based on a dramatic decline in green urban areas, suggests that those areas provide highly valuable amenities to residents. Furthermore, our approach allows us to quantify the impact of green areas across cities with various incomes and population sizes and across locations within cities. To our knowledge, this paper is the first urban economics contribution that quantifies the welfare value of green urban areas.

Our contribution relates to several strands of the economics literature. First, green urban areas share the nature of a local public good. Since Tiebout (1956), economists have discussed the issue of resident mobility and democratic decision over local public goods (voting with feet). Fujita (1986), Cremer et al. (1986) and Sakashita (1987) discuss the problem of the optimal location of local public goods and find that local public goods should spread to equidistant locations. Berliant et al. (2006) endogenize the public good provision and location in cities where households have inelastic land use. Optimal public good providers are found again to be equidistant and to serve basins of residents of the same size. Yet, in contrast to this paper, those studies are conducted under the assumption of no land use in the production of local public goods and/or no endogenous choice of residential land plots. Because green urban areas are rather land-intensive, it is important to study how land use affects public goods. Furthermore, the distribution of residents is not uniform across urban landscapes and not exogenous to the local amenities given by green urban areas. This paper focuses on the relationships between the endogenous distribution of residents and green areas.

Finally, this paper links to the urban economics literature regarding the effect of open spaces on urban form. Wu and Plantinga (2003) investigate the effect of an open space on the surrounding urban structure. They, however, treat the location and size of this open space as exogenous. Warziniack (2010) considers voting on the location of a single open space when the geographical distribution of households is exogenous. Lee and Fujita (1997) and Yang and Fujita (1999) examine the effect of a greenbelt, which has an exogenous location. Yang and Fujita (1999) consider the effect of open spaces at the neighborhood level and conclude that the equilibrium open space provision is uniform across the distance to the city center. Such results contrast with our empirical analysis that shows that the share of green urban areas is not constant across the city space. Our model with endogenous locations and choices of residential space enables an explanation of this pattern.

Parallel to our question is the issue of unoccupied urban spaces, which are often seen as green areas. In contrast to green urban areas, open spaces are not maintained for human activities. Unoccupied land has been primarily justified by the leapfrogging effect. Capoza and Helsley (1990) and followers root this effect in the commitment of building decisions and the resulting option value of urban land. Turner (2005) explains unoccupied land by the negative externalities of dwellings in their direct neighborhoods. Walsh (2007) discusses and estimates the protection and regulation of open spaces in Wake County (California, USA), which expands the discussion beyond monocentric city frameworks. Caruso et al. (2007) simulate market equilibria with discrete house slots and a fixed housing consumption, which lead to open spaces. In contrast, our paper discusses a continuous model where households decide their locations and slots and where open green areas are costly and planned as in many EU cities. Urban green areas, such as parks or trees planted in rows, have maintenance and land opportunity costs that are incurred by society.

The organization of this paper is as follows: Section 2 presents the theoretical model and discusses the social optimal allocation of green urban areas and the decentralization through the land market. Section 3 is devoted to our empirical approach and results. We first provide evidence on the concave shape of the share of green urban areas, then estimate residents' land choice, and finally quantify the economic benefits of green urban areas. The last section concludes the paper.

2 Theoretical model

We consider a circular monocentric city hosting a central business district (CBD) and a mass N of individuals. We denote by $b \in \mathbb{R}_0^+$ the distance between the CBD and the city border. The population density is defined as the number of individuals in a unit of area at distance r from the CBD and is denoted by the function $n : [0, b] \to \mathbb{R}^+$, which varies across the city. In this paper we focus on green urban areas that are closely accessible to the local community around its location. Green urban areas provide quick and frequent access to greenery, quiet, children's parks, socialization areas, etc. We consider the few blocks in the vicinity of a green urban area as our unit of area or patch and model the urban area in a continuous fashion. In a unit of area at distance r from the CBD, green urban areas offers a service $x : [0, b] \to [0, \overline{x}]$, $\overline{x} \in \mathbb{R}_0^+$, to the local community living in the vicinity. This service brings a level of amenity $a = \alpha x(r)$, although it necessitates the use of a fraction of land $\beta x(r)$

and maintenance costs $\gamma x(r)$. The parameters $\alpha, \beta, \gamma \in \mathbb{R}^+$ distinguish the amenity, land use and maintenance factors that affect green urban areas. Hence, the fraction of land used for residential purposes is given by $1 - \beta x$, and the maximum service level $\overline{x}$ is bounded by $1/\beta$. We assume absentee landlords, and the outside opportunity value of land is given by the agricultural land rent $R_A \in \mathbb{R}^+$. For simplicity, we consider that rural areas beyond the city border consist of private properties that do not provide green urban area service for city dwellers (e.g., private crop fields, fenced areas, etc.). We denote the land supply at distance r from the CBD by $\ell : [0, b] \to \mathbb{R}^+$ (e.g., $\ell = 2\pi r$ if the city lies in a plain disk). In summary, land at distance r from the CBD includes a surface $\beta x(r)\ell(r)$ of maintained green urban area and a residential area $[1 - \beta x(r)]\,\ell(r)$, and it hosts $n(r)\ell(r)$ residents who all benefit from the green urban area amenity $\alpha x(r)$.

Individuals consume a quantity z of nonhousing composite goods and a quantity s of residential space, while they benefit from the amenity a of a green urban area. They are endowed with the utility function $U(z, s, a)$, which is assumed to be concave and increasing for each variable. We assume that demands for nonhousing composite goods, residential space and amenity are gross substitutes such that U has negative second derivatives and positive cross derivatives. As individuals are homogeneous, they work and earn the same income $w \in \mathbb{R}^+$ in the CBD. Workers incur a total commuting cost $t : [0, b] \to \mathbb{R}^+$ with $t(0) = 0$ and $\mathrm{d}t/\mathrm{d}r > 0$. The price of composite good z is normalized to 1 without loss of generality. From this point on and whenever there is no confusion, we dispense the functions a, ℓ, n, s, t, x, z and R with reference to distance r.

We first study the social optimal allocation and then the land market equilibrium.

2.1 Social Optimum

In an ideal world, green urban areas and residential structures should be combined to balance their social benefits and costs. Analysis of the social optimal structure of residential and green urban areas provides urban planners with viable directions for urban planning. Towards this aim, we assume a benevolent social planner who controls residential and green urban plots across the city.

As in Herbert and Steven (1960), we assume that the planner desires to set the same utility target $u \in \mathbb{R}$ for all urban residents.[3] She (the planner) minimizes the cost in the

[3] This assumption avoids Mirrlees's discussion of the *unequal treatment of equals,* has a close link to competitive land equilibrium and yields first and second welfare theorems (Fujita 1989).

city

$$C=\int_0^b (tn+zn+R_A+\gamma x)\ \ell \mathrm{d}r,$$

subject to the target constraint $U(z,s,\alpha x)=u$ and land use constraint $sn=1-\beta x$. The total city population results from the accumulation of population density across the city: $N=\int_0^b n\ \ell \mathrm{d}r$. The planner chooses the profiles of consumption (z,s) and spatial allocations (n,x) as well as the border b. Since wages w are exogenous, this is equivalent to the maximization of total surplus $\mathcal{S}=wN-C$. After substitution of the population and land use constraints, this provides

$$\mathcal{S}=\int_0^b \left[\frac{w-t-z}{s}(1-\beta x)-R_A-\gamma x\right]\ell \mathrm{d}r. \tag{1}$$

The planner then chooses the variables (z,s,x,b) that maximize $\mathcal{S}$ s.t. $U\geq u$.

The optimal consumptions (z,s) are given by the pointwise maximization of (1), which is equivalent to the set the maximum of the residential land value:

$$V\equiv \max_{z,s}\frac{w-t-z}{s}\ \text{ s.t. } U(z,s,a)\geq u. \tag{2}$$

Since the objective function in this expression decreases with z and U increases with it, the constraint is binding. We denote the consumption $\widetilde{z}(s,a,u)$ as the solution of $U(z,s,a)=u$. Because the utility function increases for all variables, we obtain $\widetilde{z}_s=-U_s/U_z<0$, $\widetilde{z}_a=-U_a/U_z<0$ and $\widetilde{z}_u=1/U_z>0$, while concavity of utility yields $\widetilde{z}_{ss}>0$, where the subscripts denote partial derivatives. Denoting an individual's net income (net of commuting cost) as

$$y\equiv w-t,$$

the problem simplifies to

$$V=\max_s \frac{y-\widetilde{z}(s,a,u)}{s}.$$

The optimal use of residential space is given by the solution of the following first-order condition:

$$\widetilde{z}(s,a,u)-s\widetilde{z}_s(s,a,u)=y.$$

Because $\widetilde{z}-s\widetilde{z}_s>0$ and $(\partial/\partial s)(\widetilde{z}-s\widetilde{z}_s)=-s\widetilde{z}_{ss}<0$, this condition determines the optimal residential space $\widehat{s}(y,a,u)$. We denote the optimal consumption of commodity goods by the function $\widehat{z}(y,a,u)=\widetilde{z}\left[\widehat{s}(y,a,u),a,u\right]$ and the optimal bid land value as $V(y,a,u)$. Ceteris paribus, $\widehat{s}$ increases with decreasing y, and since $\widetilde{z}_s < 0$, $\widehat{z}$ increases with increasing y. By the envelop theorem, the residential land value V rises with increasing y and a and decreasing u.

The planner's problem can then be rewritten as

$$\max_{x,b} \mathcal{S} = \int_0^b \left[V(y, \alpha x, u) (1 - \beta x) - R_A - \gamma x \right] \ell \mathrm{d}r. \tag{3}$$

Pointwise differentiation w.r.t. x provides the necessary condition for green urban area service,

$$\alpha V_a (1 - \beta x) - \beta V - \gamma = 0, \tag{4}$$

where V is evaluated at $(y, \alpha x, u)$. Using the land use constraint, we obtain the following optimality condition:

$$\alpha s n V_a = \beta V + \gamma. \tag{5}$$

This condition expresses the planner's balance between the benefit of green area amenities (LHS) and the costs of green urban land and its maintenance (RHS). Let $x^* (y, u)$ be the optimal profile of the green urban area service. Note that x^* never reaches its upper bound $\overline{x} = 1/\beta$. If it did, the population density n would fall to zero, and green urban areas would lead to maintenance and land costs but no amenities (zero LHS in (5)). It can nevertheless be that $x^* = 0$ if the LHS is smaller than the RHS for all $x \in [0, 1/\beta)$. In summary, $x^* \in [0, 1/\beta)$. For the sake of conciseness, we assume in this section that the second-order condition holds and concentrate our discussion on interior solutions.[4]

Interestingly, expression (5) can be recovered as Samuelson's optimality condition of public goods after some mathematical transformations (see Appendix A):

$$\alpha \frac{U_a}{U_z} n = \beta \frac{U_s}{U_z} + \gamma \tag{6}$$

where U_a/U_z is the marginal rate of substitution between commodities and amenities and U_s/U_z that between commodities and residential spaces. The Samuelson's optimality condition states that the sum of marginal rates of substitution for green area amenities equates the maintenance cost γ plus the marginal rate of substitution for residential land. This last element is novel in the context of public goods theory. In our context, it applies at the level

[4]The second-order condition is given by $\alpha^2 V_{aa} (1 - \beta x) - (1 + \alpha) \beta V_a < 0$. This condition holds provided that $\widehat{V}_{aa} < 0$.

of the patch because externalities are localized at this level. A green urban area is a local public good because its amenities equally benefit the residents localized in its patch, which has a population density n. The same local park indeed serves many residents. Green urban areas differ from usual (spaceless) local public goods in their land intensity. This has an impact on locations with lower usage of residential spaces because low space consumption is usually associated with higher marginal rates of substitution for space U_s/U_z. To our knowledge, this tradeoff has not been highlighted in the literature.

Fixing the variables z, s and n, we observe that a lower amenity parameter α, higher land use parameter β and higher maintenance cost parameter γ entice the planner to reduce the green urban area service x and also its land area αx. We also distinguish between the effects of population density and use of space. On the one hand, areas with low population density should accommodate smaller shares of green urban areas because they benefit fewer people. On the other hand, areas with small residential plots imply high marginal rates of substitution for space and should also be provisioned with smaller shares of green urban areas. In general, population density is low at the city edges, and residential plots are small near CBDs. Hence, the planner is enticed to set smaller shares of green urban areas at the city edges and CBDs and larger shares in intermediate locations, which is the idea that we will explore in the empirical section.

2.2 Comparative statics

Comparative statics on the service of green urban areas can be obtained by totally differentiating (4). Noting that $V_a > 0$, it is easy to see that

$$\frac{\mathrm{d}x^*}{\mathrm{d}\beta}, \frac{\mathrm{d}x^*}{\mathrm{d}\gamma} < 0,$$

such that land and maintenance costs have negative impacts on the service of green urban areas to residents. Other comparative statics are ambiguous as

$$\begin{aligned}
\frac{\mathrm{d}x^*}{\mathrm{d}\alpha} > 0 &\iff \left(\left(1 + \alpha V_{aa}\right)\left(1 - \beta x^*\right) - \beta V_a\right) x^* > 0, \\
\frac{\mathrm{d}x^*}{\mathrm{d}y} < 0 &\iff \alpha V_{ay}\left(1 - \beta x^*\right) - \beta V_y < 0, \\
\frac{\mathrm{d}x^*}{\mathrm{d}u} < 0 &\iff \alpha V_{au}\left(1 - \beta x^*\right) - \beta V_u < 0.
\end{aligned}$$

First, the effect of the amenity parameter is ambiguous. It can be shown from the first expression that when green urban area services provide very few amenities and use small land pieces ($\alpha, \beta \to 0$), the optimal service x^* and surface βx^* rise with the amenity parameter α. In fact, at very low levels, a higher α gives the planner incentives to raise service x^* because the parameter raises the effectiveness of the service. However, at high levels, a higher parameter α substitutes for service level x and entices the planner to reduce it.

Net income y is given by a worker's wage minus his/her commuting cost, which increases with distance to the CBD. Therefore, comparative statics on y highlight the effect of distance to the CBD. The effect of net income can be deduced from the second expression as follows:

$$\frac{\mathrm{d}x^*}{\mathrm{d}y} > 0 \iff \alpha n^* \frac{\mathrm{d}}{\mathrm{d}\ln y}\left(\frac{U_a}{U_z}\right) + \alpha n^* \frac{U_a}{U_z}\frac{\mathrm{d}\ln \widehat{s}}{\mathrm{d}\ln y} - \beta \frac{\mathrm{d}}{\mathrm{d}\ln y}\left(\frac{U_s}{U_z}\right) > 0. \tag{7}$$

The effect of higher net income on green area services depends on three factors: first, on the income elasticity of demand for residential spaces $\mathrm{d}\ln \widehat{s}/\mathrm{d}\ln y$; second, on the marginal rates of substitution between commodities and green urban areas U_a/U_z; and finally on the reaction of those marginal rates to increases in income $(\mathrm{d}/\mathrm{d}\ln y)\,(U_a/U_z)$ and $(\mathrm{d}/\mathrm{d}\ln y)\,(U_s/U_z)$. The latter reactions are related to the Engel curves in spaces (z, s) and (z, a). It can be shown that the marginal rate of substitution rises (falls) with income in those spaces if the Engel curves rise and bend upward (downward). In other words, higher income raises more demand for residential spaces s and amenities a than demand for commodities z. As a result, the effect of net income on green urban area services depends on the balance between the income effects on the demands for amenities and space. If income effects are identical (as will be the case below under the Cobb-Douglas preferences), the marginal rates of substitution are invariant to income, and the above inequality holds for all net incomes. The optimal green urban area service then rises with net income. As a result, this optimal service also rises with wages and falls with distance from the CBD. In the end, the relative importance of each income effect is still an empirical issue for which we have found no information in the literature.

A similar comparative exercise can be performed on the impact of utility target u. One simply substitutes y for u in condition (7). The effect of target utility therefore depends on how it affects the use of residential space and the above marginal rates of substitution.

2.3 City border and population

Finally, the planner sets the city border so that the first-order condition w.r.t. b,

$$V(y, \alpha x^*, u)\,(1-\beta x^*) - \gamma x^* = R_A, \tag{8}$$

holds, where y and x^* are evaluated at $r = b$. Since x^* maximizes the LHS, the latter should be no smaller than $V(y, 0, u)$. A sufficient condition for an optimal border $b^*(u)$ is that the LHS lies above zero at $r = 0$ and decreases for increasing r. That is,

$$V(w, 0, u) \geq R_A, \tag{9}$$

$$\left(-V_y \frac{\mathrm{d}t}{\mathrm{d}r} + \alpha V_a\right)(1-\beta x) < \beta V + \gamma \quad \text{for } r \in [0, b^*], \tag{10}$$

where the second line is evaluated at $(y, \alpha x^*, u)$. We assume that these conditions hold. Comparative statics can be obtained by totally differentiating (8). Since (8) decreases with x, the optimal border b^* increases with the parameters that increase the value of the LHS of (8) and reduce its RHS. Recalling that $V_y, V_a > 0 > V_u$ and $y = w - t$, it follows that

$$\frac{\mathrm{d}b^*}{\mathrm{d}\beta}, \frac{\mathrm{d}b^*}{\mathrm{d}\gamma}, \frac{\mathrm{d}b^*}{\mathrm{d}u}, \frac{\mathrm{d}b^*}{\mathrm{d}R_A} < 0 < \frac{\mathrm{d}b^*}{\mathrm{d}\alpha}, \frac{\mathrm{d}b^*}{\mathrm{d}w}. \tag{11}$$

Hence, As in the literature, cities also expand with higher wages and *cities spread when green urban areas provide higher amenities, use smaller pieces of land and require lower maintenance costs.*shrink with higher agricultural rent and utility costs.

The city population is given by $N = \int_0^{b^*(u)} [1 - \beta x^*(y,u)] / \widehat{s}(y, \alpha x^*, u)\ \ell \mathrm{d}r$, with $y = w - t$. Finally, the total surplus is given by

$$\mathcal{S}^*(u) = \int_0^{b^*(u)} [V(y, \alpha x^*, u)\,(1 - \beta x^*(y,u)) - R_A - \gamma x^*(y,u)]\,\ell \mathrm{d}r. \tag{12}$$

Therefore, by the envelop theorem, the change in surplus is given by $\mathcal{S}_u^*(u) = \int_0^{b^*(u)} V_u\,(1-\beta x^*)\ \ell \mathrm{d}r$, which is negative since $V_u < 0$. Higher utility targets u reduce the city surplus.

2.4 Regulation and Henry George theorem

We are now equipped to discuss the impact of migration restriction and land regulation in cities. In practice, utility targets are determined by city planners (city officials and representatives) through their land regulation and migration policies. On the one hand, a city planner may opt for unrestricted migration so that u is determined by the outside utility level, e.g., $\overline{u} \in \mathbb{R}$. The population densities and levels adapt to migration pressure, and the city generates a surplus $\mathcal{S}^*(\overline{u})$, which we assume to be positive (otherwise, the planner has no incentive to create the city without external funding). On the other hand, a city planner may opt to restrict land use and population as to maximize incumbent residents' utility. Then, he/she targets the highest possible utility subject that is compatible with a positive surplus: $\mathcal{S}^*(u) \geq 0$. As the surplus decreases with increasing u, the highest utility, say u^*, is reached for a zero surplus: $\mathcal{S}^*(u^*) = 0$. In this case, green urban areas can be self-financed by land value. Indeed, $\mathcal{S}^*(u^*) = 0$ can be written as

$$\int_0^{b^*(u^*)} (1 - \beta x^*)\,(V - R_A)\ \ell \mathrm{d}r = \int_0^{b^*(u)} x^*\,(\gamma + \beta R_A)\ \ell \mathrm{d}r,$$

which shows an exact balance between the aggregate differential residential land value $(V - R_A)$ and the land and maintenance costs of green urban areas (the functions V and x^* being evaluated at $(y, \alpha x^*, u^*)$). This balance is a reminiscence of the Henry George theorem, by which a confiscatory tax on land would by itself finance a city's public goods, provided that the city reaches the size that maximizes residents' utility. This paper adds two new elements to the standard version of this theorem: *optimal green urban areas, first, are very localized public goods that provide unequal amenities and, second, require uneven land areas through the city.*

It must be emphasized that self-financing takes place at the city level and not at the patch level. There indeed exist *cross-subsidies across urban dwellers*. Indeed, at any distance r from the CBD, residents create a value V of their residential land, while they generate a green urban area maintenance cost γx and a land opportunity cost R_A on the patch. This approach yields a resident's net value equal to $S \equiv V\,(1 - \beta x) - R_A - \gamma x$, which is the integrand of (3). Using the first-order condition (4), one readily obtains that $\mathrm{d}S/\mathrm{d}r = V_y \mathrm{d}y/\mathrm{d}r = -V_y \mathrm{d}t/\mathrm{d}r < 0$. Hence, a resident's net value falls with increasing distance from the CBD. When the city management maximizes residents' utility and imposes a zero surplus such that $\mathcal{S}^* = \int_0^{b^*} S\ \ell \mathrm{d}r = 0$, it is clear that residents close to the city center bring positive net values, while those away from it bring negative values. Hence, the central population subsidizes the green urban areas at urban edges.

2.5 Competitive land market equilibrium

In most modern cities, residents freely choose their residential locations and spaces. They make their decisions according to the land rent values signaled in the urban land market. We here discuss the equilibrium allocations in a competitive land market for an exogenous profile of green urban services $x : [0, b] \to [0, 1/\beta)$ and amenities $a = \alpha x$. A household's budget constraint is given by $z + sR + t \leq w$, where $R : [0, b] \to \mathbb{R}^+$ is the land rent function of distance to the CBD.

In a competitive land market equilibrium, each land slot is awarded to the highest bidder, and individuals have no incentives to relocate within and out of a city. Therefore, they reach the same utility level u^e, where the superscript e refers to the equilibrium value. Households bid up to $(w - z - t)/s$ for each unit of residential space. Their bid rent $\psi : [0, b] \to \mathbb{R}^+$ is a function of distance r from the CBD such that

$$\psi = \max_{s,z} \frac{y - z}{s} \quad \text{s.t.} \quad U(z, s, a) \geq u^e, \tag{13}$$

where net income $y = w - t$ is a function of distance to the CBD. As individuals compete for land, they raise their bids to make their participation constraint binding and obtain the equilibrium utility level u^e. Note that (13) is equivalent to the social optimal consumption choice (2). Therefore, households' optimal consumptions are given by the functions $\widehat{s}(y, a, u^e)$ and $\widehat{z}(y, a, u^e)$, and the bid rent, by $\widehat{\psi}(y, a, u^e) = V(y, a, u^e)$. The bid rent inherits the properties of V. That is, $\widehat{\psi}_y, \widehat{\psi}_a > 0$, while $\widehat{\psi}_u < 0$.

A competitive land market equilibrium is defined as the set of functions (z, s, R, n) and scalars (b, N, u^e) satisfying the following four conditions. First, individuals choose their optimal consumptions: $z = \widehat{z}(y, a, u^e)$ and $s = \widehat{s}(y, a, u^e)$. Second, land is allocated to the highest bidder: $R = \max\{\widehat{\psi}(y, a, u^e), R_A\}$, with $R = \widehat{\psi}(y, a, u^e)$ if $n > 0$, and $R = R_A$ if $n = 0$. Third, the land market clears everywhere: $n\widehat{s}(y, a, u^e) = (1 - g)$ if $n > 0$. Finally, the total population conforms to its density: $N = \int_0^b n \, 2\pi r \mathrm{d}r$. Here, N is taken as exogenous in a closed city model, while u^e is exogenous in an open city.

Within a city, equilibrium land rents are given by the winning bids such that $R = \widehat{\psi}(y, a, u^e)$. Since bid rents ψ increase with net income y and amenities a, the equilibrium land rent R falls with distance from the CBD but rises with the proportion of green urban area. Importantly, *at the social optimal amenity a and utility level u, consumptions in a competitive land market match the social optimal ones exactly, while land bid rents ψ and land rents R match social land values V.* This finding is reminiscent of the social optimum property in Herbert and Steven's (1960) model, where competitive land market equilibria are socially optimal. The land market is then allowed to decentralize the choices of land and commodity consumption. However, this applies only if green urban areas are optimally set in our framework.

In equilibrium, land rents must exceed R_A for any location $r \in [0, b)$ and be equal to it at the equilibrium city border b^e. To simplify the exposition, we assume that $R^e(r)$ crosses R_A from above at $r = b^e$, which occurs if $\widehat{\psi}(y, a, u^e)$ lies above R_A in the CBD and falls in r. A sufficient condition is given by

$$\widehat{\psi}(w, 0, u^e) > R_A,$$
$$-\widehat{\psi}_y \frac{\mathrm{d}t}{\mathrm{d}r} + \alpha \widehat{\psi}_a \frac{\mathrm{d}x}{\mathrm{d}r} < 0.$$

Given that $\widehat{\psi} = V$, these conditions compare to the social optimal ones, (9) and (10), except that they do not include the land and maintenance costs of green urban areas. Residents do not consider those costs in a competitive equilibrium. After some reshuffling, this gives

$$w > \widehat{z}(w, 0, u^e) + R_A \widehat{s}(w, 0, u^e), \tag{14}$$
$$\frac{\mathrm{d}t}{\mathrm{d}r} > -\alpha \widetilde{z}_a \frac{\mathrm{d}x}{\mathrm{d}r}. \tag{15}$$

These sufficient conditions imply that urban productivity is sufficiently high for a city to exist in the absence of green urban areas and that green urban areas do not have too steep density profiles or do not yield too much spatial variation in amenities. Sufficiently high wages w and a low amenity parameter α guarantee these conditions. Under conditions (14) and (15), a spatial equilibrium exists. The equilibrium city border b^e is given by the unique solution of the land arbitrage condition: $R(b^e) = R_A$. This border coincides with the social optimal one since $R = \widehat{\psi} = V$.

The equilibrium population density is equal to $n^e = (1 - g)/\widehat{s}(y, a, u^e) \geq 0$, while the equilibrium population aggregates the population density across the urban area as

$$N^e = \int_0^{b^e} \frac{1-\beta x}{\widehat{s}(y,a,u^e)}\, 2\pi r \mathrm{d}r.$$

The following proposition summarizes the above discussion.

Proposition 1 *Suppose that conditions (14) and (15) hold. Then, a competitive land market equilibrium exists and is unique. Furthermore, the land and commodity consumption in the competitive land market equilibrium and social optimum coincide if green urban areas are provided at the socially optimal levels and if utility levels u and u^e match.*

A competitive land market is a powerful mechanism to decentralize consumption decisions. It can be checked that land tax does not affect goods and land consumption choices; thus, land taxes may be used by city planners to finance green urban areas.

To obtain more analytical results, we focus on a narrower class of Cobb-Douglas preferences.

2.6 Cobb-Douglas preferences

We define our workhorse model with the Cobb-Douglas utility

$$U = z^{1-\phi-\varphi} s^{\phi} e^{a\varphi}$$

with $\phi, \varphi, (1-\phi-\varphi) \in (0,1)$. We compute $\tilde{z} = (us^{-\phi}e^{-a\varphi})^{\frac{1}{1-\phi-\varphi}}$, which gives

$$\widehat{z} = \frac{1-\varphi-\phi}{1-\varphi}\, y, \tag{16}$$

$$\widehat{s} = \left(\frac{1-\varphi}{1-\varphi-\phi}\right)^{\frac{1-\varphi-\phi}{\phi}} \left(uy^{-(1-\varphi-\phi)}e^{-a\varphi}\right)^{\frac{1}{\phi}}, \tag{17}$$

$$V = \kappa^{-1}\left(uy^{-(1-\varphi)}e^{-a\varphi}\right)^{-\frac{1}{\phi}} \tag{18}$$

where $\kappa = (1-\varphi)^{\frac{1-\varphi}{\phi}}(1-\varphi-\phi)^{-\frac{1-\varphi-\phi}{\phi}}$. Condition (6) becomes

$$e^{\frac{\alpha\varphi}{\phi}x}\left(\alpha\varphi - \beta\phi - \beta\alpha\varphi x\right) = \kappa\gamma y^{-\frac{1-\varphi}{\phi}} u^{\frac{1}{\phi}}. \tag{19}$$

As shown in Appendix A, there exists a unique interior optimal service level $x^* > 0$ if the green area amenities per surface unit are sufficiently large $\alpha/\beta > \phi/\varphi$, the maintenance cost γ is sufficiently low and the net income y is sufficiently high. Otherwise, there is a corner solution $x^* = 0$.

It is also shown that for any interior solution x^*,

$$\frac{\mathrm{d}x^*}{\mathrm{d}y} > 0 \quad \text{and} \quad \frac{\mathrm{d}^2x^*}{\mathrm{d}y^2} < 0.$$

Hence, since $y = w - t(r)$, the optimal share of green urban areas $g^* = \beta x^*$ increases with wage w and decreases with distance from the CBD r. The optimal share is also a concave function of wage and distance from the CBD. Finally, since $\widehat{s}$ falls in y and a and because $a = \alpha x^*$ increases with y, $\widehat{s}(y, \alpha x^*, u)$ also falls with the latter. Then, the population density increases with higher wages and falls with longer distances to the CBD.

Proposition 2 *Under the Cobb-Douglas preferences, population density is a decreasing function of distance from the CBD, while the share of green urban areas is a decreasing and concave function of this distance.*

The monotonicity of the share of green urban areas is specific to the Cobb-Douglas preference. It can be shown that the share of green urban areas increases and then decreases with distance from the city center if one assumes Polak preferences, e.g., $U = z^{(1-\varphi-\phi)}\,(s - s_0)^{\phi}\,a^{\varphi}$, where $s_0 > 0$ is an individual's minimum use of residential space; this is also a concave function under hyperbolic preferences of the form $U = a + z - 1/s$.[5] Thus, as our main theoretical prediction, we retain the fact that the share of green urban areas is a concave function of the distance from the CBD.

3 Empirical Analysis

In this section, we compare the model prediction to the actual green urban patterns in European cities.

[5]See Supplementary Appendix, available on request.

3.1 Data

In this paper, we use the dataset on urban land use from the Urban Atlas 2006, implemented by the Global Monitoring for Environment and Security (GMES) service and provided by the European Environment Agency (EEA), for the time period 2005-2007. The dataset offers a high-resolution map of land use in European urban areas, containing information derived from Earth observations and backed by other reference data, such as navigation data and topographic maps. The Urban Atlas uses Earth observation satellite images with 2.5 m spatial resolution.[6] According to the GMES, the dataset covers the functional urban areas (FUAs) of the EU cities with at least 100,000 inhabitants.[7] FUAs include land with both commuting distance and time below the critical levels defined by Eurostat.[8] The dataset includes all capital cities and covers nearly 300 of the most populous towns and cities in Europe (EU 27).[9] Figure 1 displays the urban areas covered by these cities.

The Urban Atlas provides a classification of city zones that allows for a comparison of the density of residential areas, commercial and industrial zones and extent of green areas. In this paper, we use the data on "green urban areas" (class 14100), which are defined as artificial nonagricultural vegetated areas. They consist of areas with planted vegetation that is regularly worked and/or strongly influenced by humans. More precisely, first, green urban areas include public green areas used predominantly for recreational use (gardens, zoos, parks, castle parks, cemeteries, etc.). Second, suburban natural areas that have become and are managed as urban parks are included as green urban areas. Finally, green urban areas also include forest and green areas that extend from the surroundings into urban areas with at least two sides being bordered by urban areas and structures and containing visible traces of recreational use. Importantly, for our study, green urban areas do not include private gardens within housing areas, buildings within parks, such as castles or museums, patches of natural vegetation or agricultural areas enclosed by built-up areas without being managed as green urban areas. It must be noted that green urban areas belong to the Urban Atlas' class of "artificial surfaces", which include all nonagricultural land devoted to human activities.[10]

[6] GMES maps have a 100 times higher resolution than traditional maps in the Corine Land Cover inventory produced since 1990.

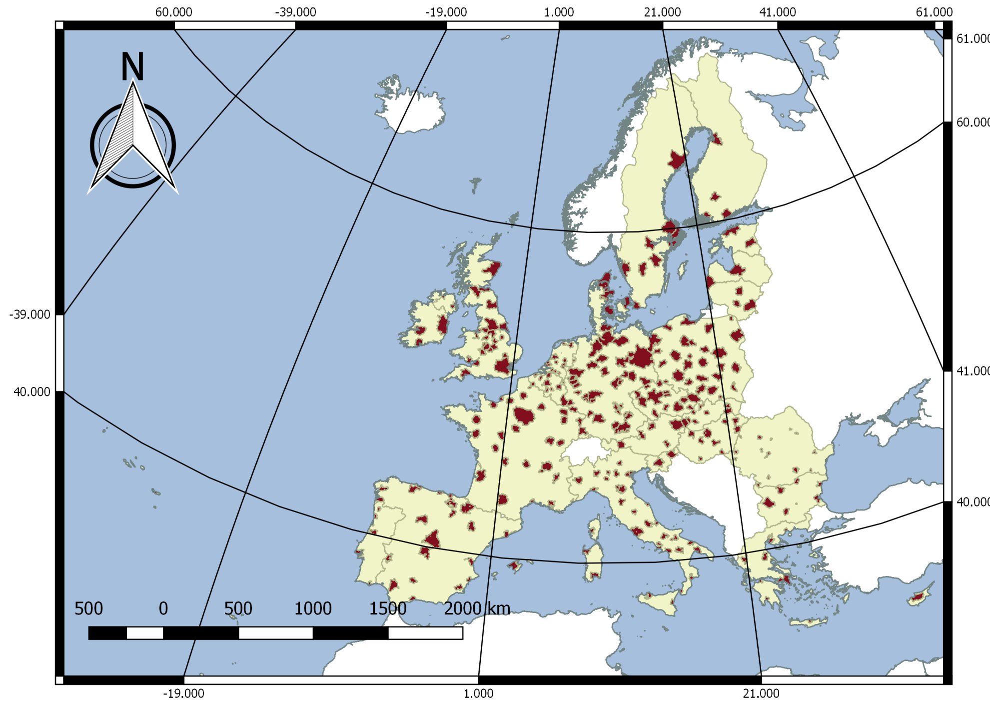
[7] See the definition in the Urban Audit database and European Environmental Agency, GMES Urban Atlas Metadata. Link: https://land.copernicus.eu/local/urban-atlas (accessed on Jan 25, 2018).

[8] See the definition in the Urban Audit in EEA, 2015, and the details in Appendix B.

[9] Austria, Belgium, France, Germany, Bulgaria, Cyprus, the Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom.

[10] In addition to green urban areas, artificial surfaces include urban areas with dominant residential use,

Figure 1: GMES 2006 maps



Maps of GMES 2006

Created using GMES 2006map boundary, EU map boundary data in Quantum GIS

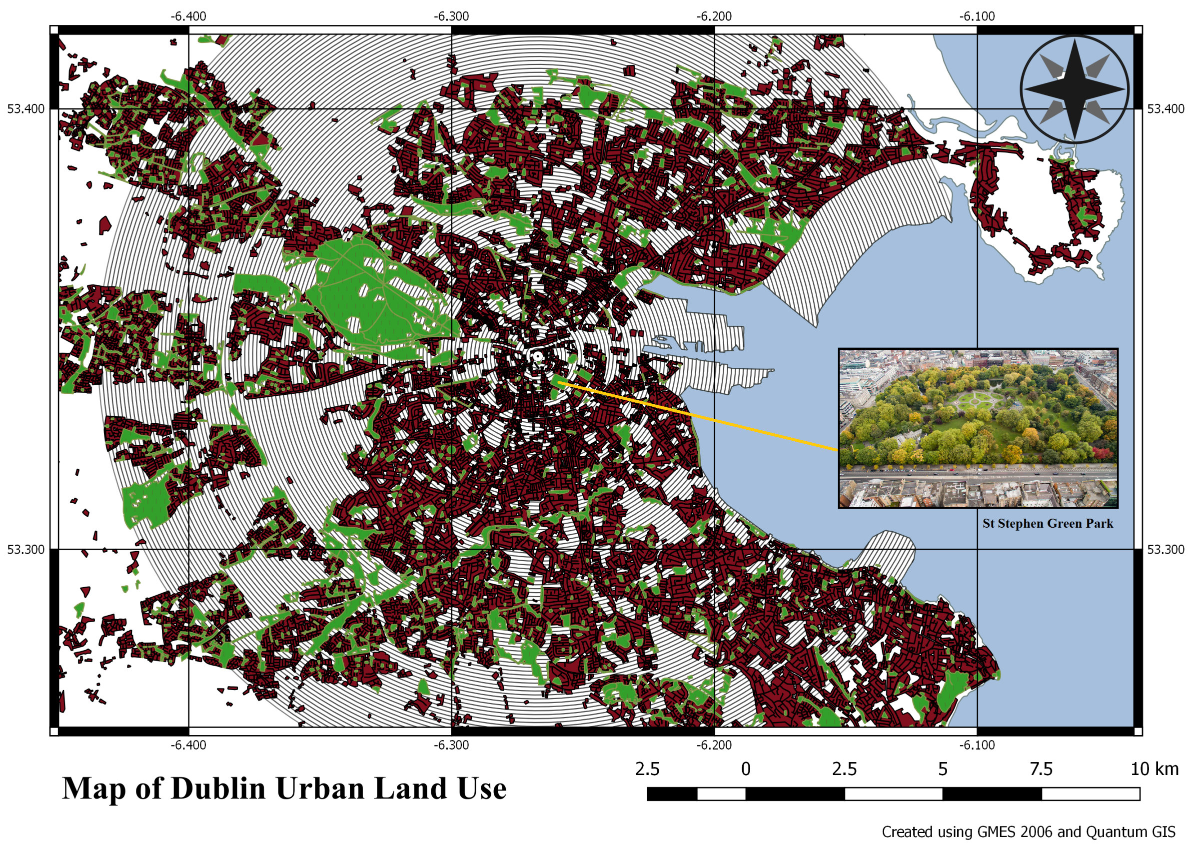
This class is distinguished from the agricultural, seminatural areas and wetlands, forest areas and water areas devoted to nonurban activities.

We select the (oldest) town hall locations as the CBDs. Then, we create a set of annuli (rings) around each CBD at 100 m intervals. We define the "annulus land area" as the intersection of the annulus and the land within the urban zone area reported by the GMES. This area includes artificial surfaces, agriculture, seminatural areas, wetlands and forest but does not include water areas because those seas and oceans are not appropriate for potential human dwellings. We then compute the share of green urban area as the ratio of the surface of green urban area to the total land in the annulus land area for each annulus . Figure 2 displays the annuli and the land use of green urban areas (green color) for Dublin.

Whereas urban theoretical models usually assume a neat frontier between residential

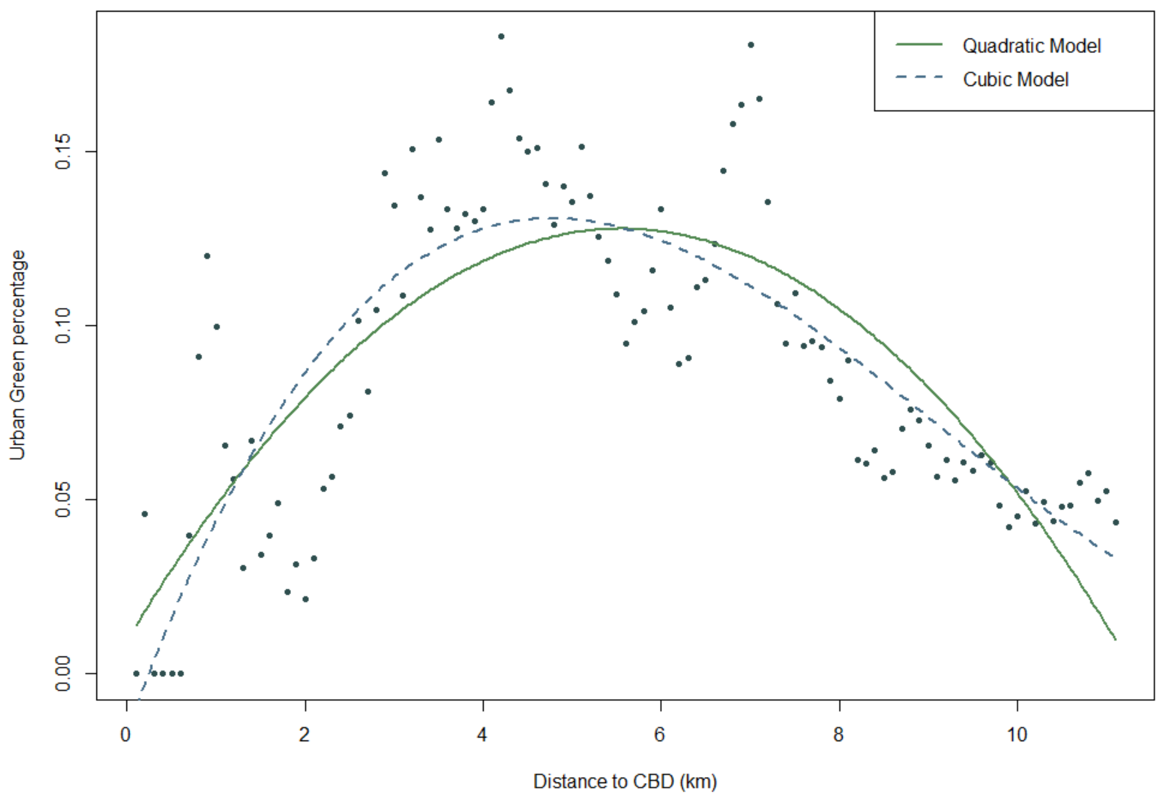
innercity areas with central business district and residential use, industrial, commercial, public, military and private units, transport units, mines, dump and construction sites, and sports and leisure facilities.

Figure 2: Dublin Land Use maps



and nonresidential spaces, urban data do not provide a clear separation between residential locations and agricultural areas and forests. In this paper, we choose to fix the city borders to the annulus for which the ratio of residential surfaces over the annulus land area falls below 20%. Residential surfaces include urban areas with dominant residential use and innercity areas with central business district and residential use. They are shown in red in Figure 2. As shown in the sequel and Appendix B, the use of other thresholds does not lead to qualitative differences in our empirical results. We define the distance from the CBD, $dist$, as the distance from the CBD of the annulus and the relative distance, $rdist$, as this distance from the CBD divided by the distance between the CBD and the city border. We also include several controls that do not depend on the relative distance to the CBD. A country dummy vector accounts for a country's specific urban regulations and wealth. We also divide the sample into three city groups: small cities, with a population below a half million; medium-sized cities, with a population between a half and one million; and large

Figure 3: Dublin Green Urban Space



cities, with a population exceeding one million. The dummy vector *city_size* includes the fixed effect on each city group.

In addition to the GMES, we use the population density from the European population grid.[11] We calculate the population mass at each distance to the city center and redistribute the population to the residential area in each annulus.[12] Because the Eurostat population grid does not cover Cyprus, we exclude Lefkosia, Cyrus. For the income level of a city, we use the household net income at the NUTS2 level, as reported in the Eurostat's Regional Economic Accounts, which provides the finest detail on household net income. Our results are robust to the use of city's per capita GDP at the NUTS3 level.[13] Other measures of cities'

[11]For more information on the European population grid, please check the technical report of the GEOSTAT 1A project from Eurostat. Link: http://ec.europa.eu/eurostat/documents/4311134/4350174/ESSnet-project-GEOSTAT1A-final-report_0.pdf/fc048569-bc1c-4d99-9597-0ea0716efac3 (Accessed on May 30, 2018).

[12]Residential areas are called 'urban fabrics' in the GMES.

[13]See Supplementary Material.

exogenous geographical characteristics are taken from the E-OBS database.[14] We control for these exogenous geographical characteristics because they may affect residential choices. We finally measure the city populations as the number (millions) of inhabitants living in the city and greater city (CGC) areas, as defined and reported in the Eurostat databases.[15] Table 1 presents the summary statistics.

[14]The E-OBS database provides detailed data on the daily temperature, daily precipitation, sea level pressure and elevation across Europe. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu).

[15]For more details, see metadata files for *urb_esms* in the Urban Audit database of the Eurostat website.

Table 1: Summary statistics

	average	sd	min	max	observations
City border (km)	4.3	3.2	1.0	24.0	305
City area (km^2)	84	174	1	1809	305
Number of annuli	43	32	10	240	305
Population in FUA (millions)	0.79	1.29	0.06	12.10	301
Population in CGC (millions): all cities	0.44	0.79	0.03	8.17	305
small cities	0.20	0.11	0.03	0.50	240
medium sized cities	0.64	0.13	0.51	0.98	36
large cities	2.25	1.66	1.01	8.17	29
Total share of green urban area in city (UGS) (%)	6.5	4.3	0.6	42.6	305
Highest share of UGS (%)	19.2	8.8	2.4	70.0	305
Highest share of UGS (%) (kernel smoothing)	9.5	5.0	0.4	42.2	305
Distance of highest share of UGS (km)	1.3	1.5	0.1	15.8	305
Distance of highest share of UGS (km) (kernel smoothing)	1.7	1.8	0.1	15.5	305
GDP per capita (€1000/hab.)	26.88	13.06	6.00	83.70	305
Household income (€1000/hab.)	15.46	5.63	3.70	30.90	304
Density (hab./100m^2)	0.44	0.28	0.10	1.96	303
Residential Space (100m^2)	0.98	0.44	0.21	2.47	304
City Geographical Controls					
Elevation (m)	212	210	-2	1,614	305
Average temperature at Jan 01 (^{o}C)	2.30	4.62	-8.48	15.57	305
Average temperature at July 01 (^{o}C)	19.18	12.27	12.27	28.70	305
Average daily precipitation (mm/day)	1.91	0.60	0.48	4.45	305
Share of Urban Green Land (%)	304	6.57	4.38	0.62	42.69
Share of Residential Land (%)	304	34.03	4.99	12.85	47.42
Share of Industrial and Public Land (%)	304	16.61	6.02	2.26	47.57
Share of Sport and Leisure Land (%)	304	3.76	3.21	0.00	12.79
Share of Forest Land (%)	304	5.14	6.23	0.00	32.57
Share of Agricultural Land (%)	304	16.29	10.75	0.00	52.13
Share of Forest Land within 100m buffer (%)	304	1.43	2.04	0.00	13.42
Share of Agricultural Land within 100m buffer (%)	304	6.15	5.66	0.00	33.19

Note: The GMES database released on May 2010 reports only 301 FUAs for the time period 2005-2007. Cities without FUAs reporting are Wrexham and Derry (UK), and Gozo (Malta). Aix-en-Provence shares the same FUA as Marseille. We use the Nadaraya-Watson Gaussian Kernel to smooth variations of annuli values.[16] GDP per inhabitants and Household income are taken from Regional Economic Accounts from Eurostat at NUTS3 and NUTS2 level respectively. Note that in Eurostat database, household income level exists only at NUTS2 level. In eurostat database for household income at NUTS2, there is no data for Luxembourg (NUTS2 code LU00); therefore, there is only 304 cities instead of 305 cities. The number of inhabitants in each annuli is calculated based on Eurostat Population Grid. As Eurostat Population Grid 2006 does not cover Cyrus; hence, we also drop the city cy001l_lefkosia in our database. The total number of annuli are calculated for 303 cities excluding lu001l_luxembourg and cy001l_lefkosia. For city geographical controls, we take into account the average for period 1995-2010 for each city.

[16]The Nadaraya-Watson estimate is given by $[\Sigma_{i=1}^{n} K(x-x_i)g_i] / [\Sigma_{i=1}^{n} K(x-x_i)]$ where $K(x) = \left(\sqrt{2\pi}h\right)^{-1} \exp(-x^2/2h^2)$. Here, we choose bandwidth $h = 2$ (km). More details, see Hansen (2009).

On average, green urban areas account for nearly 6% of the total surface of city areas. Cities have a rather heterogeneous share of green urban areas. Figure 4 displays a histogram of the shares of green urban areas across the studied cities. There is a large dispersion in the average green urban areas across EU cities. In Figure 4, the city with lowest share of green urban area (0.62%) is Limerick, Ireland, and the one with the highest share (42.6%) is Karlovy Vary, Czech Republic. The latter is a spa resort city, which offers many green areas to its visitors. The former city includes few land surfaces classified as green urban areas because it also has many agricultural and seminatural lands that can be used for urban green amenities. These outliers do not affect our results. Spots with the highest densities of urban green areas are located, on average, at $1.3 - 1.7$ kilometers from the CBD.

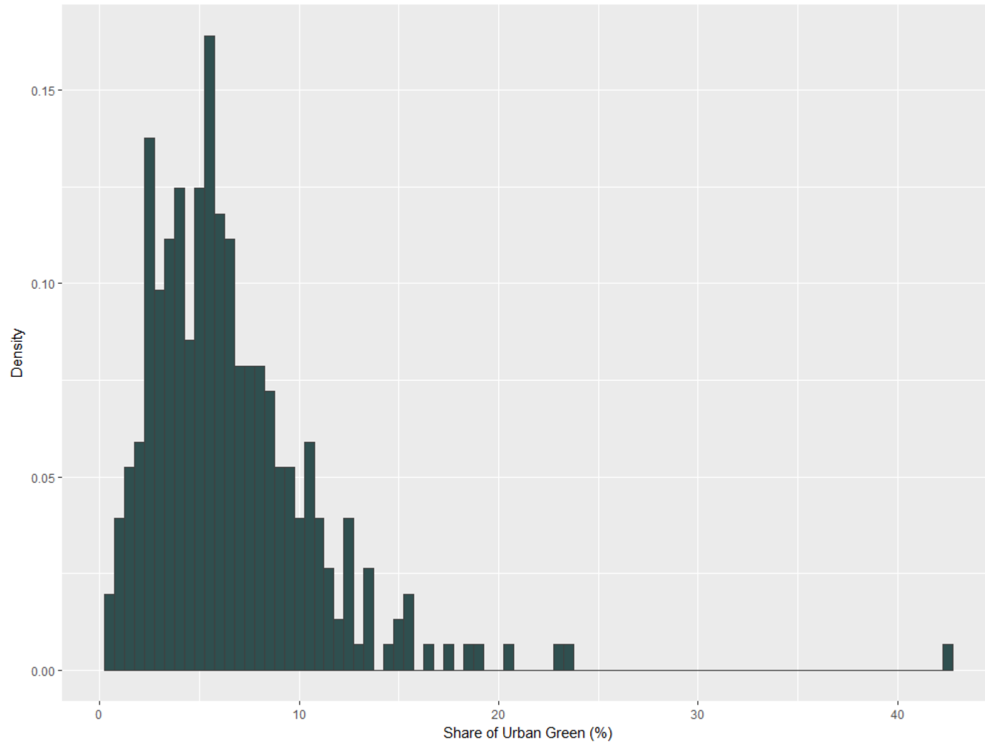
In this paper, we mostly use the household income that measures the *per capita* income net of all income taxes and at the NUTS2 level. Household incomes vary greatly across EU cities, from €3,700 per inhabitant to €30,900 per inhabitant. The average income is €15,460. Household income represents a bit more than one-half of the per capita production value (NUTS3), which reflects the high tax wedge between production cost and net income in the EU. City elevation also varies greatly, from two meters below sea level in Amsterdam, Netherlands, to $1,614$ meters above sea level in the mountainous city of Innsbruck, Austria. European cities belong to a mild climate zone, with temperatures varying between -8 and $+28$ degrees Celsius at the lowest and highest day temperature in winter and summer (measured on January 1 and July 1, respectively, for the period 1995-2000).[17] The average population density is approximately 4,400 inhabitants per square kilometer and ranges from 1,000 to 9,800 inhabitants per square kilometer. These numbers are reasonable because the database concentrates on the core of urban areas with no agricultural or seminatural land use.

3.2 Urban green area profiles

In this subsection, we compare our theoretical predictions with the empirical properties of green urban areas in EU cities. According to the theoretical model, the social optimal land share devoted to green urban areas is a concave and possibly nonmonotonic function of distance to the CBD; it first increases and then decreases as one moves away from the

[17] We use observations from the E-OBS databases from the EU-FP6 project (for details, see the references). Our samples do not contain some northern European cities in Iceland and Norway.

Figure 4: Average Share of Green Urban Space Distribution



CBD. This pattern reflects the tradeoff between the high land values in the center, which make green urban areas too costly, and the too sparse population in the periphery, which associates green urban areas with too low social benefits. The aim of this subsection is to test the concavity of the land share of green areas in the studied European cities. In addition, we test whether this share is nil at the CBD. We propose the following reduced form model:

$$g_{ijc} = \eta_1 \; rdist_{ijc} + \eta_2 \; rdist_{ijc}^2 + \eta_3 \; rdist_{ijc}^3 + X_j + \epsilon_{ijc}, \tag{20}$$

where g_{ijc} is the land share of green urban areas in annulus i (ranked according to distance to the CBD) of city j in country c. We study both quadratic and cubic models, where the coefficient η_3 is constrained to zero in the first case. To allow for comparison across cities, we define the covariate of the relative distance of an annulus $rdist_{ijc}$ and add the city fixed effects X_j as controls. Urban green areas are likely to lie on adjacent annuli such that green urban area densities may not be independent observations, which biases the estimation. Additionally, more distant annuli aggregate more surface; thus, the estimation may suffer from heteroskedasticity. Hence, we report the Huber-White heteroskedastic-consistent estimation of the standard residual errors. Table 2 presents the results from the regression of equation (20).

Table 2 reports a negative and significant correlation with the square of the distance to the CBD, which suggests that the hypothesis of a concave profile for the share of green urban areas g should not be rejected. Columns (1) to (3) report the results with and without country and city controls. As can be seen, these controls do not have a large effect on the amplitude and significance of results. Columns (4) to (6) present the results for the subsamples of small cities (population below a half million), medium-sized cities (population between a half and one million) and large cities (population exceeding one million), respectively. The signs are not altered, which corroborates the idea of concave profiles. A formal testing of concavity requires an examination of the p-value of a one-sided t-test of the respective coefficients. In Appendix B, we show that the p-values are very low; thus, the joint hypothesis of $\eta_1 > 0$ and $\eta_2 < 0$ cannot be rejected at the 99% confidence. We also show that the results are robust to various observation weightings, which suggests that misspecification issues can be excluded.

Columns (4) to (6) also show that the (absolute values of) amplitudes of the coefficients increase with city size. The shares of green urban areas reach higher levels in larger cities. Indeed, we can compute the average shares of green urban areas in the CBD ($rdist = 0$), at its peak location ($rdist = -\eta_1/(2\eta_2)$) and at the city border ($rdist = 1$). The average shares of green urban areas in the CBD are given by the intercepts of the regression models, which are computed as the following averages of the city fixed effects in Columns (4) to (6): 0.067***, 0.059***, and 0.058***. With this information, we deduce that, on average, the shares of green urban areas in small cities (Column (4)) rise from 6.7%, peak at 7.7% at $rdist = 33\%$ of the border and fall back to 3.6% at the city border. The shares in medium-sized cities (Column (5)) rise from 5.9%, peak at 9.0% at 46% of the border and fall back to 4.6%. Finally, large cities (Column (6)) have shares that rise from 5.8%, reach their peaks at 9.9% at 45% of the border and fall back to 3.7%. Finally, since the share of green urban areas over an entire city is given by $\int_0^1 (\eta_0 + \eta_1\xi + \eta_2\xi^2)\,\mathrm{d}\xi = \eta_0 + \frac{1}{2}\eta_1 + \frac{1}{3}\eta_2$, we can compute the average shares to be 6.7%, 7.7% and 8.1% for small, medium-sized and large cities, respectively.

Table 2: Profile of green urban spaces (EU27): baseline results

	Dependent variable: Share of Urban Green Space											
	Quadratic						Cubic					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Distance	0.102***	0.102***	0.102***	0.062***	0.134***	0.181***	0.297***	0.292***	0.294***	0.232***	0.299***	0.424***
	(0.016)	(0.016)	(0.016)	(0.018)	(0.040)	(0.037)	(0.038)	(0.038)	(0.039)	(0.047)	(0.082)	(0.089)
Distance_square	−0.127***	−0.127***	−0.127***	−0.093***	−0.147***	−0.202***	−0.598***	−0.588***	−0.591***	−0.499***	−0.550***	−0.802***
	(0.014)	(0.014)	(0.014)	(0.015)	(0.035)	(0.034)	(0.080)	(0.079)	(0.080)	(0.097)	(0.174)	(0.196)
Distance_cubic							0.307***	0.300***	0.302***	0.263***	0.264**	0.396***
							(0.048)	(0.047)	(0.048)	(0.058)	(0.108)	(0.120)
Country FE		✓						✓				
City FE			✓	✓	✓	✓			✓	✓	✓	✓
Sample of cities	All	All	All	Small	Medium	Large	All	All	All	Small	Medium	Large
Observations	13,091	13,091	13,091	7,982	2,106	3,003	13,091	13,091	13,091	7,982	2,106	3,003
Adj. R^2	0.046	0.148	0.369	0.367	0.338	0.383	0.056	0.158	0.379	0.374	0.346	0.402
df	13,088	13,063	12,785	7,741	2,068	2,972	13,087	13,062	12,784	7,740	2,067	2,971
F Stat.	79.63***	20.50***	26.13***	20.32***	30.05***	63.00***	54.35***	19.60***	27.13***	20.82***	30.37***	66.07***

Note: Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. Standard errors are clustered at city level and reported in parentheses. The row "df" reports the degree of freedom. The table shows results from the regression of share of green urban space within each annulus on the relative distance and its square to the city center using control variables and a constant. For comparison purpose, the relative distance of an annulus is given by its distance from city center divided by the distance of the farthest annulus in the city. Columns (2) to (3) sequentially includes controls on country and city fixed effects. Columns (4) to (6) restrict the sub-samples to small cities (population lower than half million), medium size cities (population lower between half and one million) and large cities (population larger than one million). Same strategies are applied to cubic regressions (7) to (12).

Our model predicts a concave profile for the share of green urban areas rather than the quadratic profile in Columns (1) to (6). Columns (7) to (9) present the same regression analysis as in Columns (1) to (3) for a cubic regression model. One can observe that the coefficients are significantly different from zero. The results suggest a hump-shaped profile for the share of green urban areas, with the coefficients for the square of distance being significantly negative. In this cubic model, local concavity is given by the sign of the second derivative $g'' = 2\eta_2 + 6\eta_3 \times rdist$, with $rdist \in [0, 1]$. Since η_2 is negative in Table 2, the share of green urban areas is certainly a concave function in areas sufficiently close to the CBD. However, because η_3 is positive in Table 2, concavity fails at greater distances from the CBD. Using the results in Column (9), the convexity coefficient is equal to $g''(0) = 2\eta_2 = -1.182$ in the CBD and $g''(1) = 2\eta_2 + 6\eta_3 = 0.630$ at the city edge. Hence, the profiles are increasing and concave near the CBD but convex at city edges. However, such convexity is not inconsistent with our theory because the share of green urban areas must have convex kinks when reaching zero. Finally, using Column (9), the share of green urban areas reaches a maximal value at 33% of the distance between the CBD and city border (solution of $g' = \eta_1 + 2\eta_2 \times rdist + 3\eta_3 \times rdist^2 = 0$).

We run a series of robustness checks, and the results are presented in Table 3. We study variations around the setup of Column (3) in Table 2, with both city and country fixed effects. First, as Figure 3 shows, there exists substantial serial correlation in the share of green urban areas, which questions the assumption of homoskedatisticity. Column (1) presents similar results under OLS without correction for heteroskedasticity and therefore suggests that heteroskedasiticy is not an important issue. Second, we check issues of truncation and size observation units. The surface areas of the annuli rise linearly with distance to their center. Therefore, the annuli about a CBD measure green urban areas within smaller surfaces and may have much more variability, as would be the case if an identically sized park were randomly dropped on the annuli. Furthermore, small-surface annuli are supposed to include the true city center but may miss this objective if true city centers are slightly away from the city hall locations used as the city centers. The observed share of green urban areas close to CBDs can then be more volatile and biased. To check for this issue, we aggregate the three most central 100-meter annuli into one larger central ring and the next two annuli in another ring, while we leave the other annuli intact. The results are presented in Column (2) and do not qualitatively differ from the baseline model. In Column (3), we also include the robustness check for urban green profiles, where all large parks are excluded. Since our model focuses on the effect of local urban green areas within a neighborhood, it might not apply to very large parks that have global effects on city inhabitants. Therefore, we further exclude all the parks exceeding one square kilometer,[18] and the regression results do not significantly different from our baseline regression.

[18] We also checked other criteria for this size threshold for large parks in our model, such as the 99^{th} percentile level; however, the results were rather similar. For more details, please check the Supplementary Materials.

Table 3: Profile of green urban spaces: robustness

	Dependent variable: Share of Urban Green Space											
	OLS	Wider central	Exclude	Border 15%		Monocentric cities		Contiguity		Other measures		
		annulli	big parks	Raw	Smooth'd	<1Mo	OECD	50km	100km			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10	(11)	(12)
Distance	0.102***	0.068***	0.092***	0.065***	0.005	0.078***	0.102***	0.080***	0.055	0.108***	0.087***	0.126***
	(0.006)	(0.019)	(0.015)	(0.015)	(0.014)	(0.017)	(0.029)	(0.021)	(0.043)	(0.020)	(0.027)	(0.030)
Distance_square	−0.127***	−0.100***	−0.120***	−0.116***	−0.078***	−0.105***	−0.128***	−0.112***	−0.089**	−0.105***	−0.094***	−0.111***
	(0.005)	(0.016)	(0.013)	(0.014)	(0.013)	(0.014)	(0.025)	(0.018)	(0.036)	(0.018)	(0.024)	(0.027)
City FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sample of cities	All	All	All	All	All	Small+med.	EU21	50km	100km	All	All	All
Observations	13,091	12,179	13,091	16,819	22,507	10,088	5,094	8,634	2,826	13,091	13,091	13,091
Adj. R^2	0.369	0.435	0.349	0.399	0.442	0.363	0.300	0.387	0.301	0.340	0.306	0.326
df	12,785	11,873	12,785	16,513	22,201	9,811	5,011	8,425	2,758	12,785	12,785	12,785
F Stat.	26.13***	31.73***	22.42***	37.65***	59.54***	21.79***	27.58***	27.23***	19.18***	23.10***	19.89***	21.75***

Note: Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. Standard errors are clustered at city level and reported in parentheses. The row "df" reports the degree of freedom. The table shows results from the regression of share of green urban space within each annulus on the relative distance and its square to the city center using control variables and a constant. The relative distance is normalized to one for farthest annulus for comparison purpose. Columns (1) includes simple OLS without corrections; other columns are controlled for heteroskedasticity. Column (2) merges the central annuli in wider areas. Column (3) exludes all big parks with the size above 1 kilometre square. Columns (4) and (5) define city borders with a lower threshold (15%) on the un-smoothed and smoothed profile of residential density. Columns (6) and (7) isolate monocentric cities respectively by using cities with less than one million inhabitants and by excluding the polycentric cities reported by OECD. Columns (8) and (9) exclude contiguous cities with inter-city distances lower than respectively 50 and 100 km. Column (10) measures the share of green urban space with respect to only artificial surfaces. Column (11) further excludes transport infrastructure, etc. Column (12) measures it with respect to only urban fabrics.

Third, in Table 2, city borders have been defined by the locations where the share of residential space reaches a threshold of 20%. This definition resulted from our tradeoff between theory and data. In theory, a city border is well defined and has a zero residential density, although it is not well defined in the data where the residential density never reaches zero. The use of a threshold that is too high certainly undershoots the actual distances between the CBD and city borders. Therefore, we extend the definition of a city border with a lower threshold of 15%. This extension is shown in Column (4), where the number of observations rises to 16,851 annuli. The results remain qualitatively the same. However, for this threshold value, the random variations in the share of residential spaces lead to fluctuations and downward biases in city border values. Column (5) displays the results obtained when the share of residential spaces is smoothed (with the same kernel smoothing as in Table 1). The number of observed annuli rises to 22,549. The results remain qualitatively the same, except for the coefficient of the linear distance term, which becomes not significant. This result reflects a decrease in the slope of the measured share of green urban areas and is explained by the fact that Column (5) includes new observations with no green urban areas. These new observations with zeros at far distance from the CBD reduce the slope of the share of green urban areas.

Fourth, the dataset may not match the monocentric city hypothesis of the theoretical model because it includes polycentric cities and contiguous cities. The next columns of Table 3 reduce this mismatch. Column (6) reports results for the set of 27 small and medium-sized EU cities with populations less than 1 million individuals (keeping fixed country effects). This approach eliminates the largest cities that are prone to host multiple subcenters. Column (7) focuses on monocentric cities using the OECD study on metropolitan urban polycentricity, keeps the 21 countries that are common to our GMES database and excludes reported polycentric cities (see Appendix B for details). Similarly, in our theory, cities are spatially separated, which is not always the case in the data. Columns (8) and (9) report results with the subsamples of cities that are at least 50 km and 100 km apart, respectively, keeping fixed country effects. The first distance usually corresponds to the extent of urban labor market areas. The second distance makes sure that daily commuting between cities is unattractive. As can be seen in Table 3, Columns (6) to (9) do not qualitatively deviate from the baseline results.

One may question to which land functionality green urban areas should be compared. Our theoretical model discusses the split between green urban areas and residential land. In the above empirical model, we have extended the areas for residential functionality to all human dwellers' functionalities. Accordingly, our above baseline empirical analysis used a measure of the share of green urban areas consisting of the ratio of the area of green urban areas as the numerator and the area of artificial surfaces, agriculture, seminatural areas, wetlands and forest as the denominator. Therefore, the denominator includes many potential land functions. The last columns of Table 3 present the results on alternative measures for this denominator, which increasingly narrow the comparison down to residential areas. Column (10) displays the share of green urban areas when we keep only the artificial surfaces in this denominator. This approach eliminates agricultural areas, wetlands and forests. Next, we compare green urban areas to the land used exclusively for human activities. Column (11) reports the results with the denominator measuring the land for residences, offices and green urban areas (i.e., urban fabrics, industrial, commercial, public, military and private units and green urban areas). This approach eliminates roads, railways, ports, airports, mines, construction sites, land with no use and sports and leisure facilities. Finally, Column (12) is even more restrictive by concentrating on only urban fabrics and green urban areas. The regression coefficients remain stable despite important variations in the definition of the share of green urban areas.

Finally, we run the regression (20) for each city in our sample and count the number of cities for which the concavity property holds. We observe only 4.59% (10.49%) of cities where we cannot reject convexity ($\eta_2 > 0$) for p-value < 0.01 (p-value < 0.1).

3.3 Residents' land uses

In this subsection, we estimate residents' land use. Toward this aim, we use the model with Cobb-Douglas preferences because of their popularity and convenient properties in urban economics.[19]

We suppose the presence of observable heterogeneity χ in the preference for land plots or specific characteristics of locations as well as unobservable heterogeneity or measurement errors ϵ. The utility function becomes

$$U(s, a, z, \chi, \epsilon) = \epsilon \chi z^{1-\phi-\varphi} s^{\phi} (e^{a})^{\varphi}.$$

We assume that the transport cost $t = w(1 - e^{-\tau(r)})$; thus, the net income is given by

[19]The results are similar when using hyperbolic preferences (Mossay and Picard 2012). See Supplementary Material.

$y = w - t = we^{-\tau(r)}$, where $\tau(r)$ is a function of distance to the CBD. For simplicity, we assume the quadratic form $\tau(r) = \tau_1 \times r + \tau_2 \times r^2$. Green area amenities are given by $a = \alpha x$, which can be written as a function of green urban areas g as $a = \alpha g/\beta$, where α and β are green amenity and land use intensity parameters, respectively. We can further standardize the amenity value $\alpha = 1$ and consider β as land use intensity to provide one unit for green amenities. Taking the natural logarithm of (17) and adding the heterogeneity and error terms, we obtain the following residential land use:

$$\ln s = \ln \left(\frac{1-\varphi-\phi}{1-\varphi} \right)^{-\frac{1-\varphi-\phi}{\phi}} + \frac{\varphi}{\phi} \ln \beta - \frac{1-\varphi-\phi}{\phi} \ln w$$
$$+ \frac{1-\varphi-\phi}{\phi} \tau_1 r + \frac{1-\varphi-\phi}{\phi} \tau_2 r^2 - \frac{\varphi}{\phi\beta} g + \frac{1}{\phi} \ln \bar{u} - \frac{1}{\phi} \ln \chi - \frac{1}{\phi} \ln \epsilon.$$

Accordingly, residents have larger land plots for cities with smaller incomes w, larger distances between residences and the CBD r, smaller green urban areas g, higher outside utility $\overline{u}$, and smaller observable characteristics χ.

From these results, we build a regression model of residential land use

$$\ln(\mathrm{s}_{ijc}) = \vartheta_0 + \vartheta_1 \ln w_{jc} + \vartheta_2 \mathrm{dist}_{ijc} + \vartheta_3 \mathrm{dist}^2_{ijc} + \vartheta_4 g_{ijc} + \vartheta_{5c} I_c + \vartheta_6 X_{jc} + \vartheta_7 A_{ijc} + \varepsilon_{ijc}$$

for the observations of annulus i in city j of country c. We measure the city wage w_{jc} by the per capita household net wage in the NUTS2 areas[20] and the green urban areas g_{ijc} by the land share of green urban areas (as in Table 2, Column (1)). Given language, cultural and administrative barriers, we consider that individuals freely move across cities only within the same country. Thus, the country utility level is captured by the vector of country dummies I_c. Finally, vector X_{jc} controls for observed city characteristics, such as elevation, rainfall and temperature. Vector A_{ijc} controls for observed amenities in each annulus, such as the shares of sport leisure facilities and industrial lands and the shares of forest and agricultural lands within a 100 m distance from the residential areas.

[20] In this text, city wages are measured by the incomes net of taxes at the NUTS2 level. Net incomes closely reflect the budget constraints faced by residents in their land use choices. However, NUTS2 encompasses larger areas than the cover of many cities, which may downward bias city income values. In Appendix C, we perform the same analysis with the production value at the NUTS3 level, which includes taxes. The results are similar except the values should be interpreted differently.

A potential endogeneity issue arises because the choices for residents' land use and planers' green urban areas are intertwined. Indeed, urban planers are expected to organize green urban areas as a function of surrounding population densities and therefore residents' land use. To control for such a reverse causality, we use the historical level of urban green areas as indicative of the current ones. The main idea behind using historical urban green area information is that once an urban green area is developed, it is rarely changed. In fact, many urban green areas in Europe were provided decades ago and have remained intact. Examples are Hype Park in London, created around the 16^{th} century by Henry VIII and originally intended for hunting, and the 'Jardin du Luxembourg' was first built as a private garden of Queen Marie de Medici in the early 17^{th} century. Both private parks were later converted to public green areas by public authorities. Thus, we can mitigate the reverse causality using the data for old parks to predict the locations of current public green areas. Toward this aim, we use the Corine Land Cover 1990 database, which unfortunately does not cover all GMES Urban Atlas countries. As a result, the regression results exclude cities in the UK, Sweden and Finland.[21] The details for the first-stage regression are reported in Appendix C, which confirm that the historical levels of urban green areas are a good proxy or predictor of the current levels.

The results are reported in Table 5. Columns (1) to (4) display OLS estimates without instrument variables. In all columns, the coefficient estimates are consistent with our model predictions: residents use larger land plots for smaller city income, larger distance between residences and the CBD and smaller green urban areas. The results are robust after controlling for country fixed effects, city geographical conditions, such as elevation, rainfall and temperature (see Column (2) and (3)) and different types of amenities within annuli (see Column (4)). We applied the same level of controls for IV regression, which is shown in Column (5) to Column (8). The IV regression reports slightly stronger effects of urban green areas on the residential slot size than those of the OLS regressions, which is intuitive because the historical level of urban green areas was lower than the current level. We also apply the Wu-Hausman test for endogeneity (reported in Appendix C). The Wu-Hausman

[21] Corine Land Cover (CLC) 1990 does not cover the UK, Sweden and Finland. The database covers Austria, Belgium, Bulgaria, Croatia, the Czech Republic, Denmark, Estonia, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Montenegro, the Netherlands, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, and Turkey, out of which 23 countries are included in our data. For details, see Corine Land Cover 1990 Metadata: https://land.copernicus.eu/pan-european/corine-land-cover/clc-1990?tab=metadata (Accessed May 02, 2018). To our knowledge, CLC 1990 is the oldest land use database that systematically covers all of Europe.

test coefficient is not significant at the 90% confidence level, meaning that we can confirm the alternative hypothesis of no endogeneity at the 90% confidence level, which further implies that endogeneity may not be a critical issue in our analysis. Both the OLS and IV results show significant coefficients for the share of green urban area amenities g for approximately $2.24 - 2.25$ before including the control and $1.69 - 1.86$ after including all other controls. This finding implies that, ceteris paribus, residents in annuli with no green urban areas use 14% more land than those residing in annuli with a 7% share of green urban area.[22] Population densities are reduced in the same proportions. According to this empirical estimation, green urban areas are an important factor explaining the use of residential land and population density.

[22]We compute $s_{ijc}/s_{i'jc} = e^{\vartheta_3(g_{ijc} - g'_{ijc})}$, with $g_{ijc} = 0$ and 0.07.

Table 5: Residents' land use

	Dependent variable: Ln Residential Space							
	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	−0.507***	−1.448***	−1.517**	−0.896	−0.505***	−1.420***	−1.478**	−0.868
	(0.143)	(0.356)	(0.631)	(0.556)	(0.145)	(0.360)	(0.635)	(0.560)
Ln Household Income	−0.074	−0.758***	−0.784***	−0.735***	−0.074	−0.752***	−0.778***	−0.730***
	(0.071)	(0.173)	(0.167)	(0.151)	(0.072)	(0.174)	(0.168)	(0.151)
Distance to CBD	1.502***	1.412***	1.455***	0.885***	1.497***	1.415***	1.448***	0.879***
	(0.237)	(0.220)	(0.212)	(0.213)	(0.237)	(0.221)	(0.212)	(0.213)
Distance to CBD square	−0.607***	−0.549***	−0.553***	−0.314**	−0.604***	−0.556***	−0.551***	−0.312**
	(0.130)	(0.130)	(0.133)	(0.129)	(0.130)	(0.130)	(0.133)	(0.127)
Share of Urban Green	−2.241***	−2.052***	−1.941***	−1.695***	−2.253***	−2.274***	−2.153***	−1.866***
	(0.400)	(0.254)	(0.251)	(0.214)	(0.431)	(0.285)	(0.273)	(0.237)
Country FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
City Geographical Controls	No	No	Yes	Yes	No	No	Yes	Yes
Annuli Amenity Controls	No	No	No	Yes	No	No	No	Yes
Sample	All	All	All	All	All	All	All	All
Observations	10,853	10,853	10,853	10,853	10,853	10,853	10,853	10,853
Adjusted R^2	0.207	0.468	0.483	0.579	0.197	0.464	0.480	0.576
Residual Std. Error	0.574	0.470	0.464	0.418	0.578	0.472	0.465	0.420
df	10,848	10,827	10,823	10,819	10,848	10,827	10,823	10,819
F Statistic	19.35***	26.02***	24.35***	34.07***	18.23***	24.60***	24.45***	34.67***

Note: Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. Standard errors are clustered at city level. The row "df" reports the degree of freedom. Here, we use the Household income taken from Regional Economic Accounts from Eurostat at NUTS2 level with adjustment to purchasing power standard (PPS) as the proxy for city income level, and it is measured in €100,000. The distance to CBD is measured in 10 kilometres. The inverse of residential space is calculated by dividing the number of inhabitants in each annuli with annulus areas net of its urban green space. We exclude Cyrus and Luxembourg as the Eurostat population grid database does not cover Cyrus and the household income data at NUTS2 of Eurostat does not cover Luxembourg. United Kingdom and Finland are also excluded as they are not covered by Corine Land Cover 1990. City boundary is chosen at 20% cut-off point. For city control, we take into account the elevation, average rain fall, average temperature in Jan 01 and average temperature in July 01 for period 1995-2010. City amenity controls include the share of industrial, sport and leisure land use as well as the share of forest and agriculture land within 100 meters buffer from residential area.

3.4 Counterfactual Analysis

In this subsection, we use the previous regression model to quantify the value of green urban areas. We recover all parameters of our theoretical model and run several counterfactual analyses. In particular, we build counterfactuals where half of the green urban areas are deleted in every annulus and are either left unused or converted to new residential land. We can then evaluate the changes in the residential land use and consumption of goods, population density, land rents and utility levels for each city. To express utility changes more intuitively, we measure the cost to residents by their incentives to leave the city and wage compensation (subsidy or tax) that they must receive to keep their utility levels. By the same token, we discuss the distribution of the effect of green urban areas between cities and within them. We consider this analysis a useful exercise because it informs policy makers about the impact of urban green areas on city structures and sizes.

We recover the model parameters from the estimated coefficient of residents' land use using the values of ϑ_0, ϑ_1, ϑ_2, ϑ_3 and ϑ_4 from Column (8) in Table 5. Country utility levels are recovered from the parameters ϑ_{5c} and the constant term ϑ_0. Our baseline model and counterfactuals use the observed distance to the city center, city and country caracteristics and local (non-green) amenities. The baseline model simulates the variables under study using those estimated parameters and the observed characteristics (distance to CBD, wage, green urban areas, ...). The counterfactual exercises investigate the impact of canceling 50% of the urban green areas in each annulus of each city, keeping the same observed characteristics. Counterfactual exercise 1 considers open cities where utility levels and unobserved heterogeneity are maintained. This helps us discuss a long-term and unregulated perspective, where urban planners do not impose restrictions on workers' mobility within and between cities. Counterfactual exercise 2 considers closed cities with exogenous city populations. The study of closed cities can be appropriate in evaluating policy changes that occur simultaneously in all cities, such as changes in EU policies.[23] Here, our aim is to discuss a midterm or regulated perspective, where urban planners are able to restrict workers' mobility between cities but allow residents' land use to change. To give a relevant measure of utility change, we also compute the *compensating variation wage* as the city wage that maintains the baseline utility level when we remove green urban areas. Counterfactual exercises 1 and 2 hinge on the assumption that empirical model residuals reflect land heterogeneity that is

[23] Cheshire and Shepard (2002) also use the closed city model to analyze the welfare effects of policy changes in the UK.

unobserved to the econometricians but observed and used by residents in their land plot size choices. Such heterogeneity is reported in the counterfactual results. This assumption may be strong, as it imposes strong information on behalf of residents. Therefore, we also take the opposite view in counterfactuals 3 and 4 where the residuals consist of measurement errors that can be observed neither by the econometricians nor the residents. The details of the counterfactual analyze are relegated to Appendix D.

The results of the baseline model are displayed in Table 6(a). The first column reports the number of cities in the baseline exercise. Every other column reports the average and standard deviation over the city averages imputed from the baseline model. The second and third columns show the consumptions of composite goods and housing by households, while the fourth column displays the net income. The difference between this column and the sum of the two previous columns accounts for commuting costs. On average, individuals have €15,220 as disposable income and spend €6,920 for housing expenses, which account for approximately 45% of their net income. Such a figure is slightly above the average housing costs in European cities, which are approximately a quarter of the household income for both European rural and urban areas. The literature reports a range between 18% and 32%, with higher levels for urban areas and renters (Fahey et al, 2004, Davis and Ortalo-Magne, 2011). Our model differs from this literature because we do not take into account housing furniture and maintenance (5% of housing costs in Eurostats, 2015), consider city cores, which have more expensive housing locations, and finally do not model the construction process. The last two columns of Table 6(a) report the average residential area and green urban area across cities, the latter being about one-fifth of the former. The rows address the cases when we consider the regression model error as a spatial amenity for the residents (spatial heterogeneity) and when we do not (no heterogeneity). The difference between the two cases is not large.

Table 6 (a): Counterfactual analysis: open and closed cities

	Cities number	Composite Goods (Z) (€1000)	Housing Rent ($R \times s$) (€1000)	Income (W) (€1000)	Residential Area (km^2)	Green Area (GA) (km^2)
Spatial heterogeneity	264	5.05	6.92	15.22	25.29	4.52
		(1.87)	(2.56)	(5.87)	(48.91)	(9.87)
No spatial heterogeneity	264	5.03	6.89	15.22	25.29	4.52
		(1.86)	(2.55)	(5.87)	(48.91)	(9.87)

Note: The standard deviation is reported in the parenthesis. Household income is taken from Eurostat at NUTS2 level and is measured on purchasing power standard (PPS) at 1000€. More details on PPS measure, please check Eurostat technical documents.

Our main results are displayed in Table 6(b), which shows the baseline model (first row) and the counterfactual exercises (other rows) for cases with and without spatial heterogeneity. The table structure is the same as that of Table 6(a). Every column reports the average and standard deviation of the city averages over annuli imputed in the baseline and counterfactual exercises. Consider the first row, which presents our baseline model and permits comparison with the literature. The first five columns display the imputed residential surfaces, land rents on units of residential plots and green urban areas, population and relative utility. The average city size of 0.31 million inhabitants is consistent with the statistics that most European cities are medium-sized (European Commission and UN-Habitat (2017), Urban Audit, Eurostat). Residents' average use of space is approximately 95 m^2; the measure is reasonable given that we consider the core of the most populous cities in the EU, which are the densest areas of the most urbanized parts of the EU.[24] The land rent per square meter is 93.41 €/m^2/year on average for all cities. On average, the land values of green urban areas (102.32 €/m^2/year) are higher than residential land prices. These values are imputed from the residential land price associated with each annulus. Because urban green areas are concentrated at close and intermediate distances to CBDs, they are surrounded by more expensive residential land plots.

Consider, now, counterfactual exercise 1, where one removes 50% of the green areas in every annulus of open cities. In open city systems, the utility of city inhabitants is exogenous, but the change in green amenities affects the urban structures. Suppose, furthermore, the case where the removed land is not (re)used, as indicated in the second row of Table 6(b). To keep the same utility level, residents must compensate for the decrease in urban green amenities by larger residential plots, which implies that a share of the population must migrate out of the city. On average, city residents raise their land use from 95 to 100 m^2 (a rise of 5.3%), and the city population falls from 0.31 to 0.29 million (a loss of 6.5%). Land rents fall from 93.41 to 87.11 € per m^2 and year (a fall of 6.74%). We compute the total loss in the land market to be approximately €147 million for an average city.[25]

Suppose now that 50% of green urban areas are converted into residential land, as shown in the third row of Table 6(b). In an average city, there is a new land supply of 2.26 km^2 (half of 4.52km^2) on top of the baseline residential land supply of 25.29 km^2, a rise of 8.9%

[24] They are more densely populated than US cities.

[25] To estimate the total land rent loss, we multiply the city residential area of each annulus with the per-square-meter land rent loss between the baseline and counterfactual models, aggregate over the city, and compute the average over all cities.

(see Table 6(a)). This rise is slightly more than the 5.3% space compensation that residents required without land conversion. As a result, the additional land supply attracts new city dwellers, and the city population rises to 0.32 million on average. Residential land rents rise slightly to 87.64 € per m^2 and year because the new land is supplied at more central locations with higher values. We compute that, compared to the baseline model, the housing market increases its total value by nearly €55 millions per year and city when we take into account converted areas.

Table 6(b): City structure under closed and open scenarios $g = 0.5 \times g_0$

		Cities	s	Land Rent	Green value	Pop.	$U/\bar{U}^{\mathrm{b}}$	Comp. W	Total loss in housing market	Total loss in wage comp.	Total loss in pop.
		(number)	($100m2$/hab)	(€/m^2/y)	(€/m^2/y)	(mil.)		(€1000)	(€ $mil.$)	(€ $mil.$)	$1,000hab.$
Baseline 0 (spatial heterogeneity)		264	0.95	93.41	102.32	0.31	1.00				
			(0.44)	(68.95)	(75.71)	(0.69)	(0.36)				
$g = 0.5 \times g_0$											
Counterfactual 1: Open City											
Urban green conversion	No	264	1.00	87.11	92.73	0.29	1.00		146.61	0	20.38
			(0.46)	(63.11)	(66.70)	(0.64)	(0.36)		(407.11)		(52.53)
	Yes	264	1.00	87.64	92.73	0.32	1.00		-54.74	0	-8.24
			(0.46)	(63.46)	(66.70)	(0.71)	(0.36)		(126.62)		(18.39)
Counterfactual 2: Closed City											
Urban green conversion	No	264	0.95	93.36	99.38	0.29	0.94	16.65	1.82	580.56	0
			(0.44)	(69.00)	(72.98)	(0.64)	(0.34)	(6.46)	(16.97)	(1926.01)	
	Yes	264	1.03	84.47	89.26	0.32	1.03	14.57	3.45	-198.56	0
			(0.47)	(60.75)	(63.28)	(0.71)	(0.38)	(5.71)	(15.30)	(544.27)	
Baseline 0 (no spatial heterogeneity)		264	0.88	93.70	100.81	0.28	1.00				
			(0.34)	(59.34)	(61.55)	(0.50)	(0.36)				
$g = 0.5 \times g_0$											
Counterfactual 3: Open City											
Urban Green Conversion	No	264	0.93	87.37	91.53	0.29	1.00		130.20	0	18.13
			(0.35)	(54.04)	(54.44)	(0.64)	(0.36)		(311.97)		(41.07)
	Yes	264	0.93	87.60	91.53	0.32	1.00		-50.81	0	-7.58
			(0.35)	(54.03)	(54.44)	(0.71)	0.36)		(115.20)		(16.41)
Counterfactual 4: Closed City											
Urban Green Conversion	No	264	0.88	93.64	98.01	0.29	0.94	16.65	2.26	511.77	0
			(0.34)	(59.40)	(59.64)	(0.64)	(0.34)	(6.47)	(13.72)	(1483.03)	
	Yes	264	0.96	84.72	88.58	0.32	1.03	14.61	2.88	-183.74	0
			(0.36)	(51.84)	(52.31)	(0.71)	(0.37)	(5.70)	(11.47)	(484.95)	

Note: The standard deviation is reported in the parenthesis. Household income is taken from Eurostat at NUTS2 level and is measured on purchasing power standard (PPS) at 1000€. More details on PPS measure, please check Eurostat technical documents. Population is calculated in million inhabitants. Utility is measured relatively to baseline average. Basically, we measure

$$\frac{U_{jc}}{\bar{U}^{baseline}} = \frac{U_{jc}}{\frac{1}{N}\sum_{jc} U_{jc}^{baseline}} = \frac{u_{jc}^{1/\hat{\phi}}}{\frac{1}{N}\sum_{jc}\left(u_c^{baseline}\right)^{1/\hat{\phi}}}$$ where u_{jc} is calculated as in the above section.

Counterfactual exercise 2 allows us to discuss the impact of reducing urban green areas by half in closed cities where city planners prohibit migration. As predicted by our theoretical model, when there is no conversion of land, the utility of all residents decreases once we reduce the level of urban green amenities. Specifically, the average utility decreases from 1 to 0.94. This decrease requires an increase in the baseline annual net income of €15,200 to the compensating-variation income of €16,650, an increase of €1,430 (9.4%). Multiplying this figure by the city population, we obtain a subsidy of €580 million for an average city. Residential land rents decrease only by a small amount from 93.41 €/m^2/year to 93.36€/m^2/year, providing a total loss of €1.82 million per year.

Suppose now that green urban areas are converted into residential land. Then, the residential land supply increases, land rents decreases, and residents can use more land to compensate for the loss of green area amenities. Specifically, the land rents drop by 9.6% to 84.47 €/m^2/year, and the total loss in housing market increases to €3.45 millions per year. However, city residents enjoy both lower land rent and larger residential land plots, which increase their average utility level. They obtain a slightly higher average utility level (increase by 3%) and require a smaller compensating-variation wage of €14,570 per year to maintain their level of utility, which is equivalent to an income reduction of €650 per year.

The bottom panel of Table 6(b) replicates the above analysis when we replace the assumption of unobserved spatial heterogeneity by that of measurement errors. It can be seen that most effects are similar. The main differences lie in the level of averages and standard deviations of our variables of interest. On the one hand, the standard deviations are naturally smaller because residents are no longer assumed to consider spatial amenity variations in their choices. However, the reduction in the standard deviation is not drastic, which shows that the model is already well explained by the independent variables of the regression model. On the other hand, the assumption of measurement errors also alters the average values of imputed variables. For instance, in the baseline model of the bottom panel, the residential populations, land uses and land rents are smaller than in the baseline model of the top panel. This is because those variables are nonlinear functions of the error term ε under the Cobb-Douglas preferences.[26]

What type of cities are more sensitive to removing green urban areas? Where in the city are the changes more important? To answer these questions, we compare the impact of

[26]The expression of land use $\hat{s}(w, r, g, u, \epsilon)$ includes a term e^{ε} that is a convex function of ε. Similarly, the expression for the population density $1/\hat{s}(w, r, g, u, \varepsilon)$ includes a term $e^{-\varepsilon}$, which is also a convex function of ε.

removing or converting green urban areas between cities of different incomes and population sizes as well as between within-city locations at different distances to the CBD. Towards this aim, Table 7 reports the baseline wage (first row), the changes in the compensating-variation wages to sustain constant utility (next four rows), the baseline land rent to landlords (sixth row), and the landlords' losses (last four rows) when we group cities by income quartiles (first four columns), by population size quartiles (next four columns) and by quartiles of relative distances to the CBD (last four columns). The positive changes in compensating-variation wages can be interpreted as subsidies required to maintain the residents at their baseline equilibrium levels. Table 7 presents the results for open and closed cities with and without the conversion of green urban areas to residential land. All figures are aggregated from the same counterfactual exercise with unobserved spatial heterogeneity and with the 50% reduction in the green urban areas presented in Table 6(a) and (b).

Table 7: Counterfactual analysis: between- and within-city statistics

		Between cities								Within cities			
		City Income Quartiles				City Population Size Quartiles				Distance to CBD Quartiles			
		0-0.25	0.25-0.50	0.50-0.75	0.75-1	0-0.25	0.25-0.50	0.50-0.75	0.75-1	0-0.25	0.25-0.50	0.50-0.75	0.75-1
Baseline wage (€1000/hab/y)		8.19	12.27	17.71	22.93	12.36	14.97	15.85	17.72	16.71	16.62	16.64	16.54
		(2.03)	(1.23)	(1.16)	(2.23)	(4.47)	(5.91)	(6.19)	(5.57)	(5.92)	(5.91)	(5.92)	(5.92)
Increase in comp.var.wage(€1000/hab/y)													
Counterfactual 1: Open City													
Urban green conversion	No	0	0	0	0	0	0	0	0	0	0	0	0
	Yes	0	0	0	0	0	0	0	0	0	0	0	0
Counterfactual 2: Closed City													
Urban green conversion	No	0.77	1.18	1.67	2.09	1.28	1.37	1.33	1.71	1.60	1.85	1.47	1.12
		(0.54)	(1.12)	(0.97)	(0.92)	(1.29)	(0.89)	(0.78)	(1.09)	(1.75)	(1.53)	(1.27)	(1.31)
	Yes	−0.36	−0.66	−0.71	−0.90	−0.83	−0.64	−0.51	−0.64	−0.29	−0.43	−0.61	−0.92
		(0.41)	(0.84)	(0.77)	(0.54)	(1.04)	(0.60)	(0.42)	(0.49)	(0.83)	(0.67)	(0.63)	(1.03)
Land rent baseline (€/m^2/y)		49.73	94.59	104.09	126.23	73.45	85.55	100.08	114.56	176.38	114.34	80.06	63.85
		(25.29)	(56.73)	(93.81)	(57.47)	(43.00)	(52.39)	(73.99)	(90.54)	(131.70)	(85.43)	(66.05)	(63.93)
Landlord loss (€/m^2/y)													
Counterfactual 1: Open City													
Urban green conversion	No	3.27	6.30	7.49	8.22	5.21	5.53	6.54	7.90	11.24	8.69	5.23	3.89
		(2.95)	(6.21)	(9.40)	(7.31)	(5.86)	(4.79)	(8.09)	(8.80)	(16.55)	(10.29)	(7.55)	(9.75)
	Yes	2.97	5.35	6.79	8.09	4.01	5.02	6.22	7.83	11.24	8.69	5.23	3.89
		(2.69)	(5.20)	(8.57)	(8.41)	(4.43)	(4.35)	(7.41)	(9.58)	(16.55)	(10.29)	(7.55)	(9.75)
Counterfactual 2: Closed City													
Urban green conversion	No	0.05	0.08	0.04	0.01	0.01	0.03	0.06	0.09	0.08	1.27	−0.06	−0.45
		(0.08)	(0.17)	(0.21)	(0.43)	(0.12)	(0.18)	(0.16)	(0.43)	(14.95)	(7.05)	(3.34)	(5.14)
	Yes	4.69	9.04	10.51	11.67	8.43	7.87	8.83	10.73	15.47	11.71	7.49	5.92
		(5.15)	(10.02)	(13.61)	(10.99)	(10.58)	(7.24)	(8.83)	(10.73)	(19.39)	(13.30)	(10.42)	(13.85)

Note: The standard deviations are reported in the parenthesis. Here, we report only the case with heterogeneity; for non-heterogeneity case (like in counterfactual exercises 3 and 4), the results are similar. The loss for landlord is accounted as the difference between new land rent equilibrium and land rent in baseline model 0.

Let us consider first the case of open cites (counterfactual exercise 1). The reduction in green urban area amenities harms residents who partly leave the city. Because of free migration, city residents keep the same utility as in the country side and ask for no compensation to stay in cities, which is why the second and third rows in Table 7 display a set of zeros for the change in compensating-variation wages. By contrast, landlords lose money. If urban green areas are not converted to residential land, they lose €3.27 and €8.22 per m^2 and year in cities belonging to the bottom and upper income quartiles, respectively (see seventh row). Similarly, they lose €5.21 and €7.90 per m^2 and year in cities belonging to the bottom and upper population quartiles, respectively. This result is explained by the fact that land value, city size and income are positively correlated. Landlords also lose €11.24 per m^2 and year in the central city quartile but only €3.89 per m^2 and year in the city periphery quartile, indicating that land rents decrease with distance from the CBD. This pattern remains approximately the same if urban green areas are converted to residential land (eighth row). In this case, the above figures decrease by approximately €1 per m^2 and year in the lowest income and population size cities but only slightly for the highest ones. The conversion of green urban areas mitigates the conclusions only to a small extent.

Let us now consider closed cities in which migration is restricted and half of the green urban areas is removed (conterfactual exercise 2). Suppose, initially, no land conversion (fourth row in Table 7). To stay in the city, residents require an increase in compensating-variation wages, or subsidy, of €770 per year for the bottom city income quartile and €2,090 per year for the top one. This increase represents up to 9.1% and 9.6% of the baseline net incomes. These subsidies also increase with city population. One can check that larger cities require proportionally higher subsidies, which result from the higher losses incurred by the residents in larger cities. The subsidy is not monotonic with distance to the city center: it first increases from €1,600 to €1,850 per year when one moves from the first to the second distance quartiles and then drops to €1,120 for the last distance quartile. This pattern reflects the geographical distribution of the share of green urban areas (see Section 3.2). It can finally be seen that landlords are not substantially harmed by the reduction in green urban areas when cities are closed and land is not converted (ninth row).

Finally, suppose that the green urban areas are converted to residential land (fifth row), which increases the residential land supply and compensates residents for the lack of green area amenities. The negative changes in the compensating-variation wages indicate that residents are better off in this situation. Low-income cities would accept lower compensating-variation wages and would therefore pay a tax of €360 per year in the lowest city income quartiles and €900 per year in the highest. As shown in the table, this benefit is larger for peripheral residents. Finally, landlords are negatively affected by the additional supply of residential land (see tenth row). They are more impacted in the richest and the largest cities and at the most central locations.

4 Conclusion

In this paper, we discuss the patterns of urban green areas in cities from theoretical and empirical perspectives. Urban green areas mainly include green areas maintained for recreational purposes by non-private human institutions (typically, municipalities). Green urban areas provide residents with amenities that have the property of local public goods and high land intensity. We find that the optimal provision of urban green areas is a nonmonotonic concave function of the distance to CBDs. It results from the balance between the higher opportunity cost of land near CBDs and the lower population density at city edges.

This property is confirmed by our study of the urban land use in the 305 most populous urban EU areas. We use detailed maps of urban land use from the European Environment Agency (GMES) to study the spatial configurations of urban green areas. Our study shows a concave and hump-shaped profile of urban green areas with respect to distance to the CBD. The result is robust to many variations in the land use specifications, city structure specifications, and city and country characteristics.

We finally quantify the value of green urban areas by presenting a set of counterfactual exercises, where half of the green urban areas are removed. We estimate that, on average, open cities lose more than 6.5% of their population and that landlords lose €147 millions in each city and year if the green urban areas are not converted into residential land. If they are converted, the total residential land supply increases by 8.9%, which is sufficient to compensate locals with additional residential space and to attract new city dwellers. Compared to our baseline model, the housing market increases its total value by nearly €55 millions per year and city. In closed cities where the green urban areas are not converted, city governments need to offer an average compensation of €$1,430$ per person and year to the residents for them to maintain their utility levels. To our knowledge, this paper is the first urban economics contribution that quantifies the benefit of green urban areas in such a way.

References

Alonso W. (1964). Location and Land Use. Harvard Univ. Press, Cambridge, MA, 1964.

Brander L.M. and M.J Koetse (2011). The value of urban open soace: Meta-analyses of contingent valuation and hedonic pricing results. Journal of Environmental Management 92(10), 2763-2773.

Brueckner J., J-F. Thisse, and Y. Zenou (1999). Why is central Paris rich and downtown Detroit poor? An amenity-based theory. European Economic Review 43, 91-107.

Capozza, D. R., & Helsley, R. W. (1990). The stochastic city. Journal of urban Economics, 28(2), 187-203.

Caruso G., D. Peeters, J. Cavailhes, and M. Rounsevell (2007). Spatial configuration in a periurban city. A cellular automata-based microeconomic model. Regional Science and Urban Economics 37(5), 542-567.

Cremer H., A. De Kerchove, and J-F. Thisse (1985). An economic theory of public facilities in space. Mathematical Social Sciences 9, 249-262.

Cribari-Neto F. and Da Silva W.B. (2011). A New Heteroskedasticity-Consistent Covariance Matrix Estimator for the Linear Regression Model. Advances in Statistical Analysis, 95(2), 129-146.

Davis, M.A, Ortalo-Magne, F. (2011). Household Expenditures, wages, rents. Review of Economic Dynamics 14 (2), 248-261.

E-OBS database (2017 release version). We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu).

EU DG-ENVI (2012). The multi-functionality of green infrastructure. Science for Environmental Policy, In-depth Reports.

Fahey, T., Nolan, B., Maitre, B. (2004) Housing expenditures and income poverty in EU countries. Journal of Social Policy 33 (3), 437-454.

Fujita M. (1986). Optimal Location of public facilities - Area Dominance Approach. Regional Science and Urban Economics 16, 241-268.

Fujita M. and J-F. Thisse (2002). Economics of Agglomeration: Cities, Industrial Location, and Regional Growth. Cambridge Univ. Press.

Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, M. New. 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation.

J. Geophys. Res (Atmospheres), 113, D20119, doi:10.1029/2008JD10201

IFPRA (International Federation of Parks and Recreation Administration) (2013). Benefits of Urban Parks - A systematic review. Copenhagen, IFPRA. 68p.

Kabisch N. and D. Haase (2014). Green justice or just green? Provision of green urban spaces in Berlin, Germany. Landscape and Urban Planning 122 (0), 129-139.

Lee C. and M. Fujita (1997). Efficient configuration of a greenbelt: Theoretical modelling of greenbelt amenity. Environment and Planning A 29, 1999-2017.

Mossay P. and P.M. Picard (2011). On spatial equilibria in a social interaction model. Journal of Economic Theory, 146(6), 2455-2477.

OECD (2013). Definition of Functional Urban Areas (FUA) for the OECD metropolitan database. Link: https://www.oecd.org/cfe/regional-policy/Definition-of-Functional-Urban-Areas-for-the-OECD-metropolitan-database.pdf (acessed on Mar 01, 2018).

Picard P.M. and T. Tabuchi (2013). On microfoundations of the city. Journal of Economic Theory 148(6), 2561-2582.

Purdy J. (2015). After Nature - A politics for the Anthropocene. Harvard Univ. Press.

Sakashita N. (1987). Optimum Location of Public Facilities under the influence of the land market. Journal of Regional Science 27.

Sheshire P., and Sheppard S. (2002). The welfare economics of Land Use Planning. Journal of Urban Economics, 2002, vol. 52, issue 2, 242-269.

Tiebout C. M. (1956). A Pure Theory of Local Expenditures. The Journal of Political Economy 64(5), 416-424.

Thisse J-F. and D. Wildason (1992). Public facility location and urban spatial structure - Equilibrium and welfare analysis. Journal of Public Economics, 48, 83-118.

Turner M. (2005). Landscape preferences and patterns of residential development. Journal of Urban Economics 57, 19-54.

Walsh R. (2007). Endogenous open space amenities in a locational equilibrium. Journal of Urban Economics 61, 319-344.

Warziniack T. (2010). Efficiency of public goods provision in space. Ecological Economics 69, 1723-1730.

White H. (1980). A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity. Econometrica 48, 817-838.

Wu J. and A. Plantinga (2003). The influence of public open space on urban spatial structure. Journal of Environmental Economics and Management 46, 288-309.

UN DESA (2014). World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352).

Appendix A: Proofs

Samuelson Rule We can transform expression (5) in the Samuelson's optimality condition of public goods. Plugging the optimal residential space condition $\widetilde{z} - s\widetilde{z}_s = y$ in the land value $V = (y - \widetilde{z})/s$ gives $V = -\widetilde{z}_s$. At the same time, applying the envelop theorem on V yields $V_a \equiv -\widetilde{z}_a/\widehat{s} > 0$. Finally, using $\widetilde{z}_s = -U_s/U_z$ and $\widetilde{z}_a = -U_a/U_z$, expression (5) can be written as (6).

Cobb Douglas preferences Under extreme value theorem, there always exists a global maximum for a continuous function on a compact set; therefore, there always exists solution for city planner as x is always within the domain $[0, \frac{1}{\beta}]$.

We denote the marginal welfare by $F(x) - G(y)$ where $F(x) = e^{Ax}(B - Cx)$ and $G(y) = \kappa\gamma y^{-\frac{1-\varphi}{\phi}} u^{\frac{1}{\phi}} > 0$ are the LHS and RHS of the FOC condition (19) while $x \in [0, 1/\beta]$, $A = \alpha\varphi/\phi > 0$, $B = \alpha\varphi - \beta\phi$, and $C = \alpha\beta\varphi > 0$. If $B \leq 0$, then $F(x) - G(y) < 0$ for $x \in [0, 1/\beta]$, so that the optimal service is the corner solution: $x^* = 0$. If $B > 0$ then, $F(-\infty) = 0$, $F(0) > 0$ and $F(\infty) = -\infty$ while $F(x)$ has a unique maximum at $\bar{x}$ such that $F(\bar{x}) \geq F(x)$ and $F(\bar{x}) > 0$. It can be checked that $F(1/\beta) < 0$. As a result $F(x)$ has a single root for $x < 1/\beta$. If $AB < C \iff \bar{x} < 0$, $F(x)$ is a decreasing function and there exists a single root x' for $F(x) - G(y) = 0$ iff $F(0) > G(y)$. The optimal service is then $x^* = x'$. If $AB > C \iff \bar{x} > 0$, $F(x)$ is a bell-shaped function and there exist two roots for $F(x) - G(y) = 0$ iff $F(\bar{x}) > G(y)$. The highest root x'' has $F'(x'') < 0$ and gives the optimal service: $x^* = x''$. To sum up, there exists an interior optimal service $x^* > 0$ if $F(\max(0, \bar{x})) > G(y)$. At the interior optimal service x^*, $F'(x^*) < 0$.

To get the comparative statics w.r.t. y on the interior optimum x^*, we denote $K(x, y) \equiv F(x) - G(y)$. So, $K_x(x^*, y) = F(x^*) < 0$, $K_{xy}(x^*, y) = 0$, $K_y(x^*, y) = -G'(y) > 0$, and $K_{yy}(x^*, y) = -G''(y) < 0$. Then, we get $\mathrm{d}x^*/\mathrm{d}y = -F_y/F_x > 0$ and $\mathrm{d}^2x^*/\mathrm{d}y^2 = -\left(K_{yy}K_x - K_yK_{xy}\right)/\left(K_x\right)^2 = -K_{yy}/K_x < 0$.

Appendix B: Data and robustness check for urban green space profiles

Definitions

We first summarize the definitions of urban area entities. Eurostat uses three level of spatial units that are based on clusters of high density grid cells and urban cores. Following Eurostat a *high density grid cells* are defined as population grid cells of one kilometer square with at least $1,500$ inhabitants. A *cluster of high density cells* is a set of high density grid cells that are each surrounded by at least five other high density cells (over the eight surrounding cells). Clusters exclude the high density grid cells that are not connected or isolated. An *urban core* is a cluster of high density cells that totals at least $50,000$ inhabitants. The first level of spatial unit is the *City*. It is related to an urban core and defined by the local administrative boundary so that more than 50% of inhabitants live inside the associated urban core.[27] The second level of spatial unit is the *Greater City*, which is created when the urban population resides far beyond the local administrative boundaries. Greater cities like Greater Manchester, Greater Nottingham and Greater Paris have been defined with alternative but close definitions. In most cases, a Greater City contains a single City. The *City, Greater City* (CGC) includes those two levels. A *Functional Urban Area* (FUA) combines the city area and its commuting zone, as defined in the EU-OECD functional urban area definition (OECD, 2013). A FUA includes the "working catchmen area" of a city and is defined as the collection of all surrounding municipalities with at least 15% of their employed residents working in the associated urban core. Figure B1 presents the three different levels of spatial units for Dublin.

Hypothesis testing of Table 2

Table B1 reports the p-value associated to Table 2 columns (1) to (6). Those values are extremely small, corroborating the strong non-monotone relationship.

[27]There are some exception to this rule when the geography is disrupted by a river, a lake, fjords, or steep slopes etcetera, making it hard to recover the urban core. In this case, the City can be added to cover this urban centre.

Figure 5: Figure B1: Eurostat levels of spatial units (Source: Eurostats)

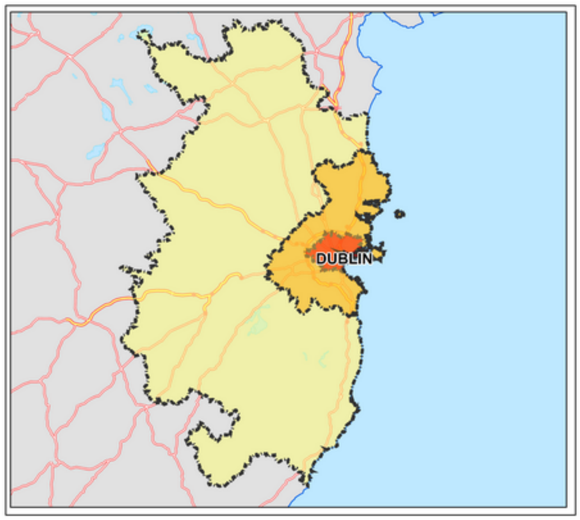


Table B1: P-values for one-sided t-test for baseline models

	(1)	(2)	(3)	(4)	(5)	(6)
$H_0 : \beta_0 > 0$	1.23×10^{-58}	1.39×10^{-6}	2.01×10^{-160}	2.00×10^{-24}	3.55×10^{-10}	1.28×10^{-24}
$H_0 : \beta_1 > 0$	4.16×10^{-11}	4.99×10^{-11}	7.88×10^{-11}	2.29×10^{-4}	4.31×10^{-4}	4.90×10^{-7}
$H_0 : \beta_2 < 0$	5.14×10^{-20}	5.60×10^{-20}	1.31×10^{-19}	5.11×10^{-10}	1.55×10^{-5}	2.68×10^{-9}

Note: The covariance matrix for residuals used for hypothesis testing is clustered at city level and heteroskedastic-robust.

Polycentric cities

In the GMES EU27 database, some cities develop in a polycentric way and host several urban cores that are physically separated but economically connected. Column (6) in Table 3 excludes EU polycentric cities using an OECD study on polycentric cities. In the latter, OECD assigns a polycentric status to a FUA using the percentage of residential population commuting from one urban core to another. A polycentric city includes two or more urban cores, which are connected and attract at least 15% of each other's population as workforce (OECD, 2013).

Furthermore, the column (6) of Table 3 focuses on the 21 countries that are common to the 27 countries in GMES EU27 database and the 23 countries in the OECD study. This excludes six EU27 countries (Bungaria, Lithuania, Latvia, Malta, Romania and Cyrus) and 2 OECD countries (Switzerland and Norway). Table B2 reports the population statistics of monocentric and polycentric cities for those 21 countries. There are 18 FUAs with two urban centres, and 6 FUAs with more than two urban cores (Barcelona, Paris, Lyon, Amsterdam, Stockholm and London). We will exclude all those FUAs in our sample. Finally, since the OECD study concentrates on cities with more 500,000 inhabitants, we keep only the cities with same sizes in the robustness of column (6) of Table 3. This excludes the GMES 27 'small' sized cities that we used in other robustness analyses.

Table B2: Descriptive statistics for EU metropolitan forms in OECD database

	average	sd	min	max	observations
Monocentric Cities					
Population of metro area (thousands)	1235	890	445	4399	87
Population of city area (thousands)	772	633	90	3467	87
Duocentric Cities					
Population of metro area (thousands)	1713	1609	561	7079	18
Population of city area (thousands)	1187	1274	314	5264	18
Polycentric Cities (with more than three centers)					
Population of metro area (thousands)	5786	5030	1096	12401	6
Population of city area (thousands)	4487	4105	1331	9942	6

Note: Table B2 includes 21 countries: All EU27 except Bungaria, Lithuania, Latvia, Malta, Romania and Cyrus. Metro population is computed from Census 2010 according to OECD metropolitan boundary maps in 2001. The metropolitan population is very similar to the FUA population provided in Urban Audit database of Eurostat. For more details, see OECD Metropolitan Explorer database, version June 2016 (OECD, 2016).

Table B3 provides further robustness analysis on the city border determination. It reports regressions coefficient when the city border is defined by various shares of residential density: 15%, 20%, 25% and 30%. Coefficient signs are unaffected by those definitions in both the quadratic and cubic models.

Table B3: Profile of green urban areas (EU27): Robustness with different levels of cut-off

	Dependent variable: Share of Green Urban Area							
	Quadratic				Cubic			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border cut-off	15%	20%	25%	30%	15%	20%	25%	30%
Distance	0.065***	0.102***	0.119***	0.111***	0.303***	0.294***	0.264***	0.260***
	(0.015)	(0.016)	(0.018)	(0.020)	(0.037)	(0.039)	(0.041)	(0.045)
Distance_square	−0.116***	−0.127***	−0.123***	−0.098***	− 0.697***	−0.591***	−0.470***	−0.449***
	(0.014)	(0.014)	(0.016)	(0.018)	(0.079)	(0.080)	(0.085)	(0.093)
Distance_cubic					0.380***	0.302***	0.225***	0.226***
					(0.048)	(0.048)	(0.052)	(0.056)
Constant	0.099***	0.100***	0.098***	0.093***	0.078***	0.082***	0.085***	0.080***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Sample of cities	All	All	All	All	All	All	All	All
Observations	16,819	13,091	10,188	8,156	16,819	13,091	10,188	8,156
Adj. R^2	0.399	0.369	0.344	0.327	0.416	0.379	0.349	0.331
df	16,513	12,785	9,882	7,850	16,512	12,784	9,881	7,849
F Statistic	37.65***	26.13***	18.50***	13.97***	40.09***	27.13***	18.84***	14.21***

Note: Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. Standard errors are clustered at city level and reported in parentheses. The row "df" reports the degree of freedom. The table shows results from the regression of share of green urban area within each annulus on the relative distance and its square to the city center using control variables and a constant. The relative distance is normalized to one for farthest annulus for comparison. All columns include controls for country and city size. The percentage of cut-off points is defined as the annuli with percentage of urban fabric areas over the total area of the city within the annuli smaller than the percentage of the cut-off points indicated.

Table B4 presents regressions with alternative weighting schemes for observations. Larger cities are less numerous but include larger populations and larger numbers of annuli, which could influence the results. Table B4 reports the regression coefficients corresponding to Table 2 with weights proportional to the city population (columns (1) and (4)) and the two different measures of city population density (columns (2)-(3) and (5)-(6)). Coefficients are invariant to weighting specifications, which suggests low risk of misspecification.

Table B4: Profile of green urban areas (EU27): Robustness with weighted OLS

	Dependent variable: Share of Urban Green Space					
	Quadratic			Cubic		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	0.166***	0.074***	0.070***	0.344***	0.267***	0.278***
	(0.011)	(0.016)	(0.019)	(0.029)	(0.040)	(0.047)
Distance_square	−0.187***	−0.065***	−0.057***	−0.625***	−0.523***	−0.551***
	(0.009)	(0.016)	(0.019)	(0.061)	(0.094)	(0.111)
Distance_cubic				0.288***	0.295***	0.318***
				(0.038)	(0.063)	(0.074)
Constant	0.084***	0.079***	0.079***	0.068***	0.062***	0.060***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)
City Size FE	✓	✓	✓	✓	✓	✓
Sample of cities	All	All	All	All	All	All
Weighting	City Pop.	Pop. Density 1	Pop. Density 2	City Pop.	Pop. Density 1	Pop. Density 2
Observations	13,117	13,117	13,117	13,117	13,117	13,117
Adj. R^2	NA	0.111	0.115	NA	0.118	0.123
df	13086	13086	13086	13085	13085	13085
F Stat.	69.30***	27.99***	27.47***	68.17***	28.77***	28.20***

Note: Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. Standard errors are reported in parentheses. We include different weighting strategy. Weights are proportional to the city population (columns (1) and (4)). Columns (2) and (5) include population density measured by dividing total city population with area of whole city. Columns (3) and (6) measure city population by dividing city population to areas of artificial urban fabric, which mostly used for residential purpose.

Appendix C: First stage regression and Wu-Hausmann Test for IV regressions

In this section, we report the first stage regression and the Wu-Hausman test for IV regression. We use the historical level of urban green spaces (land use code 141) in Corine Land Cover in 1990 as our instrument variable. To our knowledge, Corin Land Cover (CLC) was the first systematized the land use over whole Europe, and its earliest version was in 1990. However, there are two issues with CLC 1990. First, CLC 1990 did not cover UK, Sweden and Finland as those three countries only appeared in later version of Corine Land Cover in 2000 onward. Therefore, we need to drop the city samples which belong to those three above countries. Second, as CLC cover not only urban area but also the rural and all lands in Europe. Hence, its resolution is much less precise than GMES Urban Atlas that covers only urban areas. There are also discrepancies in these two databases. To decrease the discrepancies, we use the land cover in CLC 2006 and GMES Urban Atlas 2006 and correct for the discrepancies between these two sets. We assume that the changes between Corine 2006 and Corine 1990 is the evolution of urban green, while the difference between Corine 2006 and GMES 2006 are just discrepancies in measurement. We adjust the Corine 1990 with these measurement errors before using it in the first stage regression.

As showed in Table C1, the coefficient between adjusted urban green in Corine 1990 is a very good predictors for the current level of urban green. The slope is highly significant and is around 0.8. R^2 is at $0.76-0.78$. The Wu-Hausman test coefficient for endogeneity ($\hat{v}_{error}$) is not significant at 90% of confidence level, meaning that we can confirm the alternative hypothesis of no endogeneity at 90% confidence level, which further implies that endogeneity may not be a critical issue in this analysis.

Table C1: First stage regression and Wu-Hausman test for endogeneity

First stage regression

	Dependent variable: Share of green area in GMES 2006			
	(1)	(2)	(3)	(4)
Adjusted Share of green area in Corine 1990	0.802***	0.794***	0.792***	0.791***
	(0.027)	(0.027)	(0.027)	(0.027)
Adjusted R^2	0.760	0.782	0.784	0.785

Wu-Hausman test for endogeneity

	Dependent variable: Ln Residential Space			
	(1)	(2)	(3)	(4)
$\hat{v}_{error}$	-0.421	0.603	0.581	0.429
	(0.885)	(0.472)	(0.450)	(0.429)
Country FE	No	Yes	Yes	Yes
City Geographical Controls	No	No	Yes	Yes
Annuli Amenity Controls	No	No	No	Yes
Sample	All	All	All	All
Observations	10,853	10,853	10,853	10,853

Note: We use the following procedure to test for endogeneity. First Stage: $g_{ijc} = \alpha g_{Corine90} + \vartheta' Z + v_{ijc}$. Second Stage: $\ln s_{ijc} = \vartheta Z + \vartheta_4 g_{ijc} + \vartheta_{err} \hat{v}_{ijc} + \varepsilon_{ijc}$ where Z is the vector $\begin{pmatrix} 1 & w & dist & I_c & X_{jc} & A_{ijc} \end{pmatrix}$ and $\hat{v}_{ijc}$ is the residuals from first stage regression. Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. The row "df" reports the degree of freedom. For city control, we take into account the elevation, average rain fall, average temperature in Jan 01 and average temperature in July 01 for period 1995-2010. The observations are all annuli of all cities covered by both GMES and Corine Land Cover 1990. Other variables are those from original regression (GDP per capita at purchasing power standard and distance to CBD). Regression (1) to (4) are corresponding to IV regression (5) to (8) in Table 5 in the main text respectively.

Appendix D: Counterfactural analysis

We recover the model parameters from the estimated coefficient of residents' land use using the values of ϑ_0, ϑ_1, ϑ_2, ϑ_3 and ϑ_4 from Column (8) in Table 5. Country utility levels are recovered from the parameters ϑ_{5c} and the constant term ϑ_0.[28]. From the theoretical model, we recover the residential space, composite goods and the residential land rent as

$$
\begin{aligned}
\hat{s}(w,r,g,u,\varepsilon) &= (we^{-\hat{\tau}_1 r-\hat{\tau}_2 r^2})^{\hat{\vartheta}_1}\ u\ e^{g\hat{\vartheta}_4+\hat{\vartheta}_6 X_{jc}+\hat{\vartheta}_7 A_{ijc}+\hat{\varepsilon}_{ijc}} \\
\hat{z}(w,r) &= we^{-\hat{\tau}_1 r-\hat{\tau}_2 r^2}\left(\frac{-\hat{\vartheta}_1}{1-\hat{\vartheta}_1}\right) \\
\hat{R}(w,r,g,u,\varepsilon) &= \frac{we^{-\hat{\tau}_1 r-\hat{\tau}_2 r^2}-\hat{z}(w,r)}{\hat{s}(w,r,g,u,X,\epsilon)}
\end{aligned}
$$

where we use $\hat{\tau}_1 = \hat{\vartheta}_2/\hat{\vartheta}_1$, $\hat{\tau}_2 = \hat{\vartheta}_3/\hat{\vartheta}_1$, and $(1-\hat{\varphi}-\hat{\phi})/(1-\hat{\varphi}) = -\hat{\vartheta}_1/(1-\hat{\vartheta}_1)$. We first define the baseline model and the counterfactual exercises. Both baseline and counterfactuals use the observed distance to the city center r_{ijc} and amenities X_{jc} and A_{jc}. The baseline model includes the observed city wage w_{jc}, the share of green urban areas g_{ijc}, the estimated values of country utility $\widehat{u}_c$ and the unobserved heterogeneity or measurement error $\widehat{\varepsilon}_{ijc}$. Formally, we set the baseline model values to $\hat{s}^0_{ijc} = \hat{s}(w_{jc}, r_{ijc}, g_{ijc}, \widehat{u}_c, \widehat{\varepsilon}_{ijc})$, $\hat{z}^0_{ijc} = \hat{z}(w_{jc}, r_{ijc})$ and $\hat{R}^0_{ijc} = \hat{R}(w_{jc}, r_{ijc}, g_{ijc}, \widehat{u}_c, \widehat{\varepsilon}_{ijc})$.

We now investigate the impact of canceling 50% of the urban green areas in each annulus. In counterfactual exercise 1, we consider open cities where utility levels and unobserved heterogeneity are maintained at the estimated levels $\widehat{u}_c$ and $\widehat{\varepsilon}_{ijc}$. This consideration helps us discuss a long-term and unregulated perspective, where urban planners do not impose restrictions on workers' mobility within and between cities. We then remove half of the green urban areas by setting $g'_{ijc} = 0.5 \times g_{ijc}$. In both cases, we set $\hat{s}^1_{ijc} = \hat{s}(w_{jc}, r_{ijc}, g'_{ijc}, \widehat{u}_c, \widehat{\varepsilon}_{ijc})$ and $\hat{z}^1 = \hat{z}(w_{jc}, r_{ijc})$, while $\hat{R}^1 = \hat{R}(w_{jc}, r_{ijc}, g'_{ijc}, \widehat{u}_c, \widehat{\varepsilon}_{ijc})$. Residents' land use should increase ($\hat{s}^1_{ijc} > \hat{s}^0_{ijc}$) because residents require compensation for the reduction of green area amenities. If green urban areas are left with no use, the total available space remains constant and is given by $\sum_{ijc}(1-g_{ijc})\,\ell_{ijc}$, where ℓ_{ijc} is the land surface of annulus i in city j and country c. Since resident's land use increases, cities host fewer residents. If green urban areas are converted in residential land, city populations may grow if the new supply of land, $g'_{ijc}\ell_{ijc}$, is

[28] As we drop Austria in the country dummies, we have $u_{at}^{1/\phi} = \vartheta_0$, and all other countries as $u_c^{1/\phi} = \vartheta_{5c}+\vartheta_0$

larger than the increase in residents' land demand from $\hat{s}^0_{ijc}$ to $\hat{s}^1_{ijc}$. More formally, population grows if $\sum_{ijc}(1-g_{ijc})\,\ell_{ijc}/s^0_{ijc} < \sum_{ijc}\left(1-g'_{ijc}\right)\ell_{ijc}/s^1_{ijc}$.

In counterfactual exercise 2, we consider closed cities with exogenous city populations. The study of closed cities can be appropriate in evaluating policy changes that occur simultaneously in all cities, such as changes in EU policies.[29] Here, our aim is to discuss a midterm or regulated perspective, where urban planners are able to restrict workers' mobility between cities but allow residents' land use to change. We again remove half of the green urban areas ($g'_{ijc} = 0.5 \times g_{ijc}$). We set $\hat{s}^2_{ijc} = \hat{s}(w_{jc}, r_{ijc}, g'_{ijc}, u^2_{jc}, \widehat{\epsilon}_{ijc})$, $\hat{z}^2_{ijc} = \hat{z}(w_{jc}, r_{ijc})$ and $\hat{R_{ijc}}^2 = \hat{R}(w_{jc}, r_{ijc}, g'_{ijc}, u^2_{jc}, \widehat{\epsilon}_{ijc})$, where u^2_{jc} is the counterfactual city utility level. In the absence of the conversion of green urban areas to residential plots, we set the city utility level u^2_{jc} such that the city population spreads over the baseline residential area; that is, we impose that each u^2_{jc} solves the population identity $\sum_i(1-g_{ijc})l_{ijc}/\hat{s}^0_{ijc} = \sum_i(1-g_{ijc})l_{ijc}/\hat{s}^2_{ijc}$. In the case of land conversion, we set u^2_{jc} such that the city population spreads over the new residential land supply. Then, u^2_{jc} solves the population identity $\sum_i(1-g_{ijc})\ell_{ijc}/\hat{s}^0_{ijc} = \sum_i(1-g'_{ijc})\ell_{ijc}/\hat{s}^2_{ijc}$.[30] However, although utility levels are important concepts in welfare analysis, they are difficult to interpret quantitatively. Therefore, we also compute the *compensating variation wage* w^2_{jc} as the city wage that maintains the baseline utility level when we remove green urban areas, which is equivalent to setting the wage w^2_{jc} such that the above population identities hold with $\hat{s}(w^2_{jc}, r_{ijc}, g'_{ijc}, \widehat{u}_{jc}, \widehat{\varepsilon}_{ijc})$. Under the Cobb-Douglas preferences, this assumption simplifies to the compensating variation wage $w^2_{jc} = w_{jc}(u^2_{jc}/u^0_{jc})^{\frac{1}{\hat{\vartheta}_1\hat{\varphi}}}$.

The above two counterfactual exercises hinge on the assumption that empirical model residuals ε_{ijc} reflect land heterogeneity that is unobserved to the econometricians but observed and used by residents in their land plot size choices. Such heterogeneity is reported in the counterfactual results. This assumption may be strong, as it imposes strong information on behalf of residents. Therefore, we take the opposite view and assume that the residuals ε_{ijc} consist of measurement errors that can be observed neither by the econometricians nor the residents. In that case, residents do not base their decisions on ε_{ijc}, and we set $\varepsilon_{ijc} = 0$ in counterfactual exercises 3 and 4.

[29] Cheshire and Shepard (2002) also use the closed city model to analyze the welfare effects of policy changes in the UK.

[30] This approach yields the counterfactual utility levels $\left(u^2_{jc}\right)^{1/\phi} = \left(u_c\right)^{1/\phi} \frac{\left(\sum_i(1-g_{ijc})\ell_{ijc}/s^2_{ijc}\right)}{\left(\sum_i(1-g_{ijc})\ell_{ijc}/s_{ijc}\right)}$ and $\left(u^2_{jc}\right)^{1/\phi} = \left(u_c\right)^{1/\phi} \frac{\left(\sum_i(1-g'_{ijc})\ell_{ijc}s^2_{ijc}\right)}{\left(\sum_i(1-g_{ijc})\ell_{ijc}/s_{ijc}\right)}$.

Supplementary Material A: NUTS3 incomes

In this section, we report the results of estimation of the structural model using GDP per capita at NUTS3 level as the proxy for city wage. The results are similar to those in Table 5 in main text.

Table A1: Regression Results using GDP per capita at NUTS3 as proxy for city's wage level

	Dependent variable: Ln Residential Space							
	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	−0.686***	−0.899***	−0.834*	−0.230	−0.682***	−0.880***	−0.808*	−0.213
	(0.119)	(0.256)	(0.467)	(0.425)	(0.119)	(0.262)	(0.473)	(0.430)
Ln Household Income	−0.210***	−0.614***	−0.608***	−0.542***	−0.211***	−0.610***	−0.604***	−0.540***
	(0.077)	(0.103)	(0.098)	(0.090)	(0.077)	(0.103)	(0.098)	(0.090)
Distance to CBD	1.596***	1.482***	1.514***	1.004***	1.591***	1.475***	1.509***	0.998***
	(0.239)	(0.220)	(0.212)	(0.219)	(0.240)	(0.221)	(0.213)	(0.220)
Distance to CBD square	−0.610***	−0.515***	−0.518***	−0.322**	−0.608***	−0.512***	−0.517***	−0.320**
	(0.134)	(0.146)	(0.149)	(0.142)	(0.133)	(0.146)	(0.148)	(0.140)
Share of Urban Green	−2.045***	−1.773***	−1.681***	−1.487***	−2.079***	−1.948***	−1.851***	−1.611***
	(0.386)	(0.245)	(0.241)	(0.203)	(0.413)	(0.274)	(0.262)	(0.228)
Country FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
City Geographical Controls	No	No	Yes	Yes	No	No	Yes	Yes
Annuli Amenity Controls	No	No	No	Yes	No	No	No	Yes
Sample	All	All	All	All	All	All	All	All
Observations	10,853	10,853	10,853	10,853	10,853	10,853	10,853	10,853
Adjusted R^2	0.227	0.513	0.525	0.607	0.218	0.510	0.522	0.604
Residual Std. Error	0.567	0.450	0.445 (df = 10823)	0.404	0.570	0.451	0.446	0.406
df	10,848	10,827	10,823	10,819	10,848	10,827	10,823	10,819
F Statistic	19.83***	32.52***	33.78***	46.95***	18.44***	31.44***	32.74***	47.94***

Note:

Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. Standard errors are clustered at city level. The row "df" reports the degree of freedom. Here, we use the GDP per capita taken from Regional Economic Accounts from Eurostat at NUTS3 level with adjustment to purchasing power standard (PPS) as the proxy for city income level, and it is measured in €100,000. The distance to CBD is measured in 10 kilometres. The inverse of space is calculated by divided number of inhabitants in each annuli with the areas within the annuli minus the areas using as urban green (in 100 meters). We exclude Cyrus and Luxembourg as the Eurostat population grid database does not cover Cyrus and the household income data at NUTS2 of Eurostat does not cover Luxembourg. United Kingdom and Finland are also excluded as they are not covered by Corine Land Cover 1990. City boundary is chosen at 20% cut-off point. For city control, we take into account the elevation, average rain fall, average temperature in Jan 01 and average temperature in July 01 for period 1995-2010. City amenity controls include the share of industrial, sport and leisure land use as well as the share of forest and agriculture land within 100 meters buffer from residential area.

Table A1a: First stage regression and Wu-Hausman test for endogeneity

First stage regression

	Dependent variable:			
	Share of green area in GMES 2006			
	(1)	(2)	(3)	(4)
Adjusted Share of green area in Corine 1990	0.797***	0.793***	0.790***	0.789***
	(0.028)	(0.027)	(0.027)	(0.027)
Adjusted R^2	0.759	0.781	0.784	0.784

Wu-Hausman test for endogeneity

	Dependent variable: Ln Residential Space			
	(1)	(2)	(3)	(4)
$\hat{v}_{error}$	−0.330	0.413	0.406	0.241
	(0.911)	(0.455)	(0.438)	(0.407)
Country FE	No	Yes	Yes	Yes
City Geographical Controls	No	No	Yes	Yes
Annuli Amenity Controls	No	No	No	Yes
Sample	All	All	All	All
Observations	10,853	10,853	10,853	10,853

Note: We using the following procedure to test for endogeneity. First Stage: $g_{ijc} = \alpha g_{Corine90} + \vartheta' Z + v_{ijc}$; Second Stage: $\ln s_{ijc} = \vartheta Z + \vartheta_3 \ln(g_0 + g_{ijc}) + \vartheta_{err}\hat{v}_{ijc} + \varepsilon_{ijc}$where Z is the vector $\begin{pmatrix} 1 & w & dist & I_c & X_{jc} \end{pmatrix}$ and $\hat{v}_{ijc}$ is the residuals from first stage regression. Significance levels are denoted by * for p<0.1, ** for p<0.05 and *** for p<0.01. The row "df" reports the degree of freedom. For city control, we take into account the elevation, average rain fall, average temperature in Jan 01 and average temperature in July 01 for period 1995-2010. The observations are all annulus from all cities covered by both GMES and Corine Land Cover 1990. Other variables are those from original regression (GDP per capita at purchasing power standard and distance to CBD). Regression (1) to (4) are corresponding to IV regression (5) to (8) in Table A1 respectively.

Table A3: Counterfactual analysis: open and closed cities

Household expenditure shares

	Composite Goods (Z) (€1000)	Housing Rent ($R \times s$) (€1000)	Income (W) (€1000)	Residential Area (km^2)	Green Area (GA) (km^2)
With heterogeneity	6.39	11.84	26.55	25.29	4.52
	(3.01)	(5.58)	(13.51)	(48.91)	(9.87)
Without heterogeneity	5.03	6.35	11.76	26.55	4.52
	(2.99)	(5.55)	(13.51)	(48.91)	(9.87)

City

structure under closed and open scenarios $g = 0.5 \times g_0$

		Cities (number)	s ($100m2$/hab)	Land Rent (€/m^2/y)	Green value (€/m^2/y)	Pop. (mil.)	$U/\bar{U}^{\mathrm{b}}$	Comp. W (€1000)
Baseline 0 (with heterogeneity)		264	0.95	161.88	177.25	0.31	1.00	
			(0.44)	(127.49)	(140.65)	(0.69)	(0.33)	
$g = 0.5 \times g_0$								
Counterfactual 1: Open City								
Urban green conversion	No	264	1.00	152.18	162.30	0.29	1.00	
			(0.46)	(118.02)	(124.64)	(0.65)	(0.33)	
	Yes	264	0.99	153.16	162.30	0.32	1.00	
			(0.46)	(119.03)	(124.64)	(0.71)	(0.33)	
Counterfactual 2: Closed City								
Urban green conversion	No	264	0.95	161.78	172.54	0.31	0.95	29.56
			(0.44)	(127.61)	(135.29)	(0.69)	(0.31)	(15.28)
	Yes	264	1.03	145.86	154.38	0.31	1.04	24.58
			(0.47)	(111.58)	(115.64)	(0.69)	(0.34)	(12.62)
Baseline 0 (without heterogeneity)		264	0.89	163.66	175.17	0.28	1.00	
			(0.37)	(122.22)	(127.79)	(0.53)	(0.33)	
$g = 0.5 \times g_0$								
Counterfactual 3: Open City								
Urban Green Conversion	No	264	0.94	153.83	160.68	0.27	1.00	
			(0.38)	(112.88)	(114.27)	(0.50)	(0.33)	
	Yes	264	0.94	154.12	160.68	0.29	1.00	
			(0.38)	(112.85)	(114.27)	(0.55)	(0.33)	
Counterfactual 4: Closed City								
Urban Green Conversion	No	264	0.89	163.54	170.64	0.28	0.95	16.65
			(0.37)	(122.30)	(123.46)	(0.53)	(0.31)	(6.47)
	Yes	264	0.97	147.30	153.74	0.28	1.04	14.61
			(0.39)	(106.00)	(107.66)	(0.53)	(0.34)	(5.70)

Note: The standard deviation is reported in the parenthesis. Household income is taken from Eurostat at NUTS2 level and is measured on purchasing power standard (PPS) at 1000€. More details on PPS measure, please check Eurostat technical documents. Population is calculated in million inhabitants. Utility is measured relatively with baseline average. Basically, we measure $\frac{U_{jc}}{\bar{U}^{baseline}} = \frac{U_{jc}}{\frac{1}{N}\sum_{jc} U_{jc}^{baseline}} = \frac{u_{jc}^{1/\hat{\phi}}}{\frac{1}{N}\sum_{jc}(u_c^{baseline})^{1/\hat{\phi}}}$ where u_{jc} is calculated as in above section.

Table A4: Quantiles analysis: between- and within-city statistics

		Between cities								Within cities			
		City Income Quartiles				City Size Quartiles				Distance to CBD Quartiles			
		0-0.25	0.25-0.50	0.50-0.75	0.75-1	0-0.25	0.25-0.50	0.50-0.75	0.75-1	0-0.25	0.25-0.50	0.50-0.75	0.75-1
Baseline wage (€1000/hab/y)		13.20	20.28	27.89	44.83	19.50	25.79	26.67	34.24	30.65	30.36	30.45	30.22
		(2.71)	(2.16)	(2.02)	(12.69)	(6.98)	(14.34)	(11.62)	(15.44)	(15.16)	(15.00)	(15.06)	(15.01)
Increase in comp.var.wage(€1000/hab/y)													
Counterfactual 1: Open City													
Urban green conversion	No	0	0	0	0	0	0	0	0	0	0	0	0
	Yes	0	0	0	0	0	0	0	0	0	0	0	0
Counterfactual 2: Closed City													
Urban green conversion	No	1.38	2.15	3.08	5.42	2.48	2.86	2.71	3.96	3.54	4.15	3.25	2.50
		(1.79)	(1.32)	(1.69)	(2.82)	(2.66)	(2.34)	(1.74)	(2.87)	(4.24)	(3.97)	(3.08)	(2.50)
	Yes	−0.83	−1.59	−2.06	−3.40	−2.12	−1.96	−1.60	−2.21	−1.20	−1.66	−1.94	−2.59
		(1.09)	(1.43)	(1.63)	(2.61)	(2.74)	(1.99)	(1.25)	(1.73)	(2.47)	(2.03)	(1.99)	(2.96)
Land rent baseline (€/m^2/y)		80.16	146.52	178.45	242.40	124.73	152.69	167.74	202.36	349.81	201.97	133.21	102.17
		(37.40)	(95.91)	(140.07)	(147.18)	(86.63)	(110.20)	(135.50)	(156.72)	(307.07)	(157.57)	(118.50)	(113.42)
Landlord loss (€/m^2/y)													
Counterfactual 1: Open City													
Urban green conversion	No	4.10	8.25	10.66	15.80	8.14	8.79	9.65	12.22	19.43	13.57	7.66	5.69
		(4.68)	(7.16)	(11.76)	(15.32)	(11.90)	(8.66)	(12.08)	(12.11)	(29.44)	(13.57)	(11.01)	(16.61)
	Yes	3.58	6.89	9.99	14.44	5.88	7.79	9.12	12.11	19.43	13.57	7.66	5.69
		(4.40)	(6.96)	(10.29)	(13.97)	(6.95)	(7.49)	(11.23)	(13.52)	(29.44)	(13.57)	(11.01)	(16.61)
Counterfactual 2: Closed City													
Urban green conversion	No	0.12	0.14	0.10	0.05	0.00	0.03	0.15	0.23	−0.16	1.99	−0.12	−0.45
		(0.17)	(0.23)	(0.36)	(1.14)	(0.40)	(0.51)	(0.40)	(0.95)	(26.40)	(10.97)	(4.81)	(9.05)
	Yes	6.28	13.39	18.19	26.24	15.44	14.53	15.09	19.03	30.99	20.90	12.81	10.08
		(7.50)	(11.91)	(21.07)	(30.11)	(26.76)	(15.05)	(19.48)	(20.69)	(37.99)	(23.90)	(18.68)	(27.27)

Note: The standard deviations are reported in the parenthesis. Here, we report only the case with heterogeneity; for non-heterogeneity case (like in counterfactual exercises 3 and 4), the results are similar. The loss for landlord is accounted as the difference between new land rent equilibrium and land rent in baseline model 0.