Welfare analysis of cities: utilitarian second-best and the optimal Rawlsian town *

Leonardo J. Basso[a,b], Raúl Pezoa[c], and Hugo E. Silva[d,e,b]

[a]Departamento de Ingeniería Civil, Universidad de Chile, Chile
[c]Escuela de Ingeniería Industrial, Universidad Diego Portales, Santiago, Chile.
[d]Departmento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile, Chile
[e]Instituto de Economía, Pontificia Universidad Católica de Chile, Chile
[b]Instituto Sistemas Complejos de Ingeniería, ISCI

June 18, 2021

Abstract

This paper studies the Rawlsian first-best allocation in a monocentric city model using a unifying framework of land ownership. We show that a Rawlsian planner would not choose the market outcome, except for the extreme case of public land ownership in which all the differential rent is transferred in lump-sum fashion to residents. In any other case, there is an outcome with equal utility for all city residents that brings higher welfare than the market outcome. In particular, it holds in the traditional textbook formulation with an absentee landlord that owns the land. We also show that the first-best scenario can be decentralized with a revenue-neutral combination of location-specific taxes and subsidies. This instrument may produce a Rawlsian first-best city that is more extended than the market city. Thus, depending on the structure of land ownership, in the absence of externalities, the market equilibrium city may be inefficiently compact. Then, when externalities are present, policies aimed to restrict urban sprawl should take this effect into consideration. To study the relevance of our results, we assess welfare-maximizing transport pricing policies in the absence of location-specific taxes. For public transport, we show that the fare that decentralizes the first-best scenario is below marginal cost, and thus the system should be subsidized, even in the absence of externalities. In the case of car congestion pricing, we show that the welfare-maximizing toll may be non-monotonic, yielding a city that is more extended and with more aggregated mileage than the unpriced city.

*We gratefully acknowledge financial support from ANID FONDECYT 1200801, the Center of Sustainable Urban Development CEDEUS (grant ANID/FONDAP 15110020), ANID-PFCHA/Doctorado Nacional/2018-21181528, and from the ANID PIA AFB180003.

1 Introduction

Cities are a fundamental part of economic life. According to The Economist, they occupy just 2% of the earth's land surface but are home to more than half of the world's population and generate 80% of all economic output.* In cities, many agents interact: people, who choose residence and workplace locations, consumption and transport modes; developers and landlords, who affect the size and number of dwellings as well as the rental and land prices; the government, which, via regulations, taxes, subsidies, and transport policies, shapes the playground; and service and manufacturing firms, who demand land and labor to produce. It is the interaction of all these agents and their decisions, together with economic forces such as congestion, scale and agglomeration economies, that shapes the cities' structure.

Of course, the structure of cities has changed over time. In particular, the phenomenon of urban sprawl and suburbanization has been intensifying for decades now. For instance, the size of many cities is increasing at a much faster rate than their population: a survey of 282 cities across Europe shows that, in the period 1996-2006, the urbanized area increased by 18.4%, while the population density fell by 9.4% (Oueslati et al., 2015). In the US, the average density of urbanized areas decreased from 6,160 persons per square mile in 1920 to 2,589 in 1990 (Song and Zenou, 2006). Moreover, as Nechyba and Walsh (2004) point out, in the US urban areas of over a million inhabitants in 1950, approximately 65 percent of the urbanized population lived in central cities, and by 1990 the percentage was only 35.

Indeed, urban sprawl, its causes and possible remedies have been at the forefront of the urban economics and planning debate. As Ewing and Hamidi (2015) review, some urban planners argue that authorities should implement transportation policies that lead to more compact cities, so that externalities such as pollution and congestion are reduced (see e.g., Brueckner, 2005; Dieleman and Wegener, 2004; Nechyba and Walsh, 2004). At the same time, other authors point out that concentrated settlement is costly and only worthwhile if private transport costs are high, yet they have been falling for many years (Gordon and Richardson, 1997). From an economics point of view, what is essential though, is to distinguish between urban expansion that is efficient, from urban sprawl that is excessive and caused by market failures (Brueckner and Helsley, 2011). This distinction is the primary focus of the current paper:

*https://innovationmatters.economist.com/the-future-of-cities. Accessed on August 2020.

when does the market urban equilibrium delivers a city size that is efficient?

Consider a city where everyone has the same income and preferences: a city of equals. A basic tenant of urban equilibrium is that everyone reaches the same utility level. When studying Rawlsian first-best scenarios where equals are treated equally –that is, urban equilibria that maximize the common utility– previous studies have have asserted that a social planner would choose the market equilibrium based on the result that it minimizes the resource usage or total costs (see, e.g., Duranton and Puga, 2015, for a handbook assertion). Applying this to urban sprawl analysis, Brueckner (2000) argues that city growth due to an increasing population, rising income, and decreasing commuting costs cannot be deemed as socially undesirable. He does contend, however, that three primary market failures are sources of socially inefficient sprawl: traffic congestion, failure to account for open space benefits, and unrecovered infrastructure costs associated with new low-density development. In fact, Wheaton (1998) formally shows that the internalization of the congestion externality implies a higher density than the market equilibrium density at all locations, implying a more compact city. Consistent with these results, many studies have argued that with road congestion and in the absence of car congestion tolling, urban growth boundaries are a second-best policy that increases welfare (Brueckner, 2007; Pines and Sadka, 1985).

In this paper we show that even in the absence of externalities, the urban market equilibrium delivers cities that are inefficiently compact in a Rawlsian sense, unless all land rents are fully redistributed (commonly called the *public land case*). For this, we develop a unifying framework of land rent capture and redistribution using the building block of the urban economics field, the monocentric city model (Alonso, 1964; Mills, 1967; Muth, 1969), and use it to study welfare, urban size and sprawl, and transportation pricing. Our finding may seem surprising after more than 60 years of development of the urban economics field, but the crux of the matter, we think, is that the role of land ownership, land rents capture and its distribution has not been completely clarified. On one hand, when analyzing the structure of the urban equilibrium, it is general practice –the "textbook" model– to assume that an absentee landlord owns the land, such that the urban rents vanish from the city (e.g., Alonso, 1964; Brueckner, 1987; Fujita, 1989; Mills, 1967; Muth, 1969; Wheaton, 1998). On the other hand, when focusing on welfare analyses, the standard approach is to consider that the government owns and rents the land, and then distributes the revenue to households in equal shares, a situation which we will refer to in this paper as *public ownership* (e.g., Arnott, 1979b; Brueckner, 2007; Kanemoto, 1980; Kono and Kawaguchi, 2017; Kono et al., 2019; Oron et al., 1973; Pines and Sadka, 1985; Pines and Kono, 2012; Tikoudis et al., 2015, 2018; Verhoef, 2005). Most of these last papers show or remark explicitly that the market equilibrium maximizes welfare in a Rawlsian sense, and thus, no other equilibrium involving equal treatment of equals reaches higher city resident utility.

Note the gap here: one approach that completely ignores land rents is used to study market equilibrium, but one where land rents are completely captured is used to study welfare. To connect these two approaches and to help to understand the implications of each in the answer to the primary question of city size's efficiency, we build a model where an absentee landlord owns the land and where the government captures a fraction of the excess rents, which are then distributed in equal shares among the city residents. In one extreme, when the fraction is zero, we recover the absentee landlord case. On the other extreme, when the fraction is one, we obtain the public ownership setting generally assumed in the literature to analyze welfare within the city. The framework could also be interpreted as mixed ownership, where a fraction of the land is jointly owned by city residents in equal shares, and an absentee landlord owns the remaining fraction.

We show that a Rawlsian planner would not choose the market outcome allocation, except in the extreme case of public ownership. In any other case, there is an outcome with equal utility for all city residents that brings higher welfare than the market outcome, and has a more extended city. Thus, the result that the market equilibrium maximizes welfare in a Rawlsian sense only holds for an extreme case and, in particular, does not hold for the family of models usually used to study and teach basic urban economics: the market equilibrium city in the absence of externalities is, in general, too compact from a welfare standpoint. We further show that the utility-maximizing allocation can be decentralized as an equilibrium by a planner with a revenue-neutral combination of location-specific taxes and subsidies, where people that choose to live closer to city center are taxed, while people closer to the city edge are subsidized.

Importantly, the result does not come from the fact that we are looking at residents' utility and excluding the landlord rents from the welfare function, as we show that it also holds when welfare is defined as the sum of residents' surplus and the absentee landlord rents. In other words, relative to the market outcome, there is an equilibrium that can be decentralized with taxes and subsidies, where the gain in utility measured as the compensating variation is greater than the loss in the absentee landlord's rents. This, in relation to the stream of literature that considers a welfare function that includes households' utility but also absentee landlords' rents (e.g., Kono et al., 2012; Kono and Joshi, 2012; Li et al., 2013; Li and Guo, 2017; Li and Wang, 2018).

To grasp the intuition of the main result, it is useful to identify the three forces at play when a planner taxes the central areas, subsidizes the outskirts and distributes whatever fraction of land rents she can capture. Consider for the city residents locating near the CBD; for them (i) their income is reduced, since they are taxed to finance the redistribution, (ii) their income is further reduced since the excess land rents decrease when the city grows, (iii) they face a lower housing price. The first two effects reduce residents' utility, while the third one increases it. For the residents living near the city edge, the first and third effect reverse. We show that, when considering the public ownership scenario, these three effects are perfectly balanced under the market equilibrium, leaving no incentives to redistribute income. Nonetheless, in any other case but public ownership, the second effect is partially borne by the absentee landlord, who receives a share of the excess land rents. Thus, when considering only the residents' equilibrium utility, the three effects are no longer balanced, and it is desirable to redistribute income from the city center to the outskirts to increase equilibrium utility.[†] This is also why, when we consider a welfare function composed of residents' compensating variation plus the absentee landlord rents, these three effects only cancel out if city residents and absentee landlord's marginal utilities of income are equal. This only happens for a particular class of utility functions, such as quasi-linear functions.

To further demonstrate the relevance and consequences of our finding, we study how different pricing policies in various settings are affected when they are designed to maximize equilibrium utility in the

[†]This effect has the same flavor as the so-called unequal treatment of equals paradox, which states that it is efficient to distribute income from locations close to the CBD, where the marginal utility of income is lower, to the outskirts (see Wildasin, 1986). The main difference here, is that as opposed to Wildasin (1986), we are considering only equilibrium outcomes.

absence of spatially variant taxes and subsidies, and land rents are not fully redistributed. Note here that, in most cases, urban policies designed to maximize welfare have only been addressed in what we show to be the exception; full redistribution of land rents, the only case where the market outcome is efficient. We begin by studying the Rawlsian first-best city and the optimal taxation in the presence of traffic congestion externalities, as in the pioneer work of Kanemoto (1977), subsequently followed by Arnott (1979a), Anas and Kim (1996) and Verhoef (2005) among many others. In our extended, and more general framework, the optimal tax has two components: the marginal external cost or Pigouvian toll, and the location-specific taxes and subsidies that account for the welfare improving redistribution among equals. The resulting welfare-maximizing tax turns out to be flatter than the Pigouvian toll and can even be non-monotonic on space. Moreover, the welfare maximizing tax may yield a city that is more extended and with more aggregated mileage than the unpriced congested city. Numerical simulations, using parameters widely used in the previous literature, show that only when more than 90% of the rents are captured and transferred to households, there is socially undesirable urban sprawl and that in the remaining cases the unpriced congested city is still too compact.

We also examine the welfare-maximizing spatial pricing of public transportation. To investigate the issue, we use a monocentric model where the only mode available is public transport and operators' costs are proportional to the passenger-kilometers traveled. As Brueckner (2005) shows, the minimization of resources calls for setting a fare per mile equal to the constant of proportionality (i.e., the marginal cost), so that every traveler faces the full money cost of transportation. However, when maximizing welfare in the absence of full redistribution of land rents, the optimal fare takes into account that a planner would like to redistribute income from the center to the outskirts. We show through simulations that the redistribution effect is substantial and can make the welfare-maximizing fare nearly space-invariant when the landlord is absent.[‡] In Brueckner's interpretation, the optimal subsidy financed with a head tax could be close to 100% of the operational costs due to the spatial redistribution effect.

The effect presented in this paper will have significant interactions with other largely discussed policy

[‡]The use of distance-based versus spatially flat fares is a matter of continuous debate and differs across cities. For example, in New York, Rome, and Shanghai transit fares are flat, in London, Tokyo, and Washington bus fares are flat and subway fares vary with distance, while in Amsterdam, Beijing, and Sydney, all transit fares vary with distance.

options, in addition to those shown for urban sprawl, pricing public transport and congestion externalities. For example, in light of our theoretical results, it is straightforward to conclude that in many cases, imposing an urban growth boundary is welfare reducing, since the city is already too small. Additionally, the properties of second-best congestion pricing alternatives, such as cordon charging, flat per-kilometer charges or urban growth boundaries may be quite different from what has been previously argued (Anas and Rhee, 2007; Brueckner, 2007; Mun et al., 2003, 2005; Verhoef, 2005).

We close this introduction by arguing that the first-order effect presented here is more effectively demonstrated using the simplest framework of all: the monocentric city model. On one hand, the model we propose bridges many modeling assumptions found in the literature and textbooks, and shows a fundamental new result while having previous ones as special (and rather unlikely) cases. On the other hand, we believe that the insights will remain in non-monocentric city models, since the main force behind the redistributing effect is that marginal utilities of income vary with location for identical residents, as the commuting cost changes with distance, and this effect is not balanced by the landlord rents unless fully captured and returned in a lump-sum fashion to all residents. Although the unequal marginal utility of income property does not hold in all frameworks (see e.g., Lucas and Rossi-Hansberg, 2002), this is explained because, besides location choice, some papers consider a frictionless margin of adjustment in the labor market that leads, in equilibrium, to a constant marginal utility of income everywhere. When considering a labor market that is not perfectly flexible, the marginal utility of income will not be constant within the city and the effect found in this paper will persist. In particular, it should be present in, and relevant for, the growing set of quantitative models of internal city structure, which typically feature an absentee landlord whose rents are not redistributed and marginal utilities of income that vary within the city (e.g., Ahlfeldt et al., 2015; Heblich et al., 2020).

The paper is structured as follows: Section 2 presents an overview of the land ownership frameworks considered in the literature when studying the optimal land use in the monocentric city model; this discussion helps to clarify the ground where our contributions lies. Section 3 describes the mixed-ownership framework, presenting our main results in the basic monocentric city model. Section 4 presents how our main results interact with a congestible road network. Section 5 shows how the results translate to the case where a public transport system operates, while Section 6 concludes.

2 Land ownership and the optimal land use in the monocentric city model

Early works from Alonso (1964), Mills (1967), and Muth (1969) developed the starting point of the modern urban economics literature in what is called the monocentric city model. In that model, there is a central business district (CBD) in which production takes place and that accommodates all jobs. The core of the model is that people consume a numeraire good, housing, and must commute to the CBD, where they work to obtain a fixed income. While extremely simple, the monocentric model provides the basis of our understanding of several aspects of urban economics.

When studying welfare in this model, it is a well-known result that, in general, the market outcome does not coincide with the utilitarian first-best. Wildasin (1986), based on the early work of Mirrlees (1972), shows that, in a city of equals where everyone has the same income and preferences, when the income net of transport costs decreases with distance to the CBD, the marginal utility of income increases with the distance to the CBD. This, in turn, implies that a utilitarian planner would like to redistribute income towards households further out from the center of the city. Thus, in the utilitarian optimum utilities would vary with location, producing the so-called *unequal treatment of equals* effect. Since this situation cannot arise in equilibrium, we deviate from this approach, adopting the alternative, which we believe is more policy relevant, in which identical households attain the same utility, so that we are indeed studying possible urban equilibria. Thus, we consider a Rawlsian welfare function, and therefore maximize the (common) equilibrium utility.

Restricted to equal treatment of equals, multiples points of view have been adopted to analyze the monocentric city model, depending on the study's objective, making it difficult to have a clear panorama. Consequently, we find it relevant to overview the literature at this point, emphasizing the similarities and differences between literature streams. First, there is a classic stream of papers that study and analyze the equilibrium structure of cities without referring to any welfare analysis. Since land ownership does not change the equilibrium's qualitative results, these papers assume for simplicity that an absentee landlord owns all the land so that rents vanish from the city. In this group, we find the seminal works of Alonso (1964), Mills (1967), and Muth (1969), but also many others, such as Brueckner (1987). Note that this is the textbook approach, being presented in, for example, Chapter 2 of Fujita (1989), Chapter 2 of Brueckner (2011), Section 3.3.1 of Fujita and Thisse (2013), or Section 8.2.1 of Duranton and Puga (2015). This approach has also been used to study floor-area ratio regulations (Bertaud and Brueckner, 2005), suburbanization (Baum-Snow, 2007), dynamic congestion externalities (Gubins and Verhoef, 2014), and the dynamics of slum redevelopment (Henderson et al., 2020).

Then, another stream of literature focuses on the *optimal land use* within a monocentric city model, as opposed to describing the market equilibrium. Three approaches have been usually employed in these studies. First, the most common approach is based only on the analysis of the equilibrium utility reached by households. However, these papers assume that all rents are captured and redistributed lump-sum among households. This framework has been used, among others, to analyze policies to deal directly with traffic congestion (Oron et al., 1973; Kanemoto, 1980; Verhoef, 2005; Tikoudis et al., 2015, 2018), and to study second best policies in congested cities, such as land use regulations (Brueckner, 2007; Pines and Sadka, 1985; Pines and Kono, 2012; Kono and Kawaguchi, 2017) or property taxes (Kono et al., 2019). Since these papers consider externalities, the market outcome is clearly not desirable, and some kind of spatially variable taxes or subsidies must be applied to achieve the first-best. Nevertheless, to the best of our knowledge, the first-best policies presented in all of these papers only aim to correct the model's externalities; in other words, in the absence of externalities, the first-best instruments are zero everywhere, and the market outcome maximizes equilibrium utility. Therefore, as Kanemoto (1980) shows formally, a Rawlsian planner would choose the market outcome in the absence of externalities under this approach.

Second, some papers base their optimal land use analysis on the concept of resource usage, defined as the sum of total non-land consumption, total commuting cost, and the opportunity cost of the urban land. This approach considers the opposite end of the spectrum regarding land ownership than those that maximize utility: an absentee landlord. The key examples of this approach are Section 3.4 of Fujita

(1989) and Section 3.3.3 of Fujita and Thisse (2013), who show that the market equilibrium minimizes resources used to achieve a given equilibrium utility. This approach has been used, for instance, to study policies in the presence of traffic congestion (De Lara et al., 2013), and the desirability of transit subsidies (Brueckner, 2005). Based on the result of Fujita (1989), Brueckner (2005) concludes that the market outcome is efficient –and subsidizing public transport is not–, while Duranton and Puga (2015) states that a Rawlsian planner would choose the market equilibrium. Note that the conclusion regarding a Rawlsian planner is not straightforward, since the result of Fujita (1989) deals with minimization of resources and not maximization of equilibrium utility. Thus, the question of whether a Rawlsian planner would choose the market outcome remains open when considering an absentee landlord.

Third, a final approach to study the optimal land use also considers an absentee landlord, but defines and aims to maximize a welfare function that includes the monetary value of households' utility and the absentee landlord rents. This approach has been used to study policies that aim to correct traffic externalities (Li and Guo, 2017; Li and Wang, 2018), land use regulations (Kono and Joshi, 2012; Kono et al., 2012), or road design (Li et al., 2013), among others. Since the models considering this welfare function deal with externalities, the market outcome is evidently not desirable. But, again, the intervention only aims to correct externalities, and they are zero in their absence. This would suggest that, in the absence of externalities, the market outcome maximizes the welfare function used.[§]

In light of this overview, it seems transparent that the question of the market outcome's desirability when considering a Rawlsian planner has not been clearly answered. Furthermore, the different approaches are vastly dissimilar regarding assumptions, so it is not very clear whether there is any connection between them or if the results obtained under one of them translate to the others. Since all of the papers presented aim to provide policy insights, we believe it is of the foremost importance to understand what are the implications, if any, of the assumptions about land rent capture and redistribution, and the welfare function used on policy results. In the next section, we aim to clarify this connection; in a nutshell, minimization of resource usage is not equivalent to maximization of equilibrium utility

[§]In this paper, we show that this result is not general and is strongly dependent on the utility function considered. Indeed, under some mild conditions over the utility function, the market outcome does not maximize a welfare function defined as the sum of households' compensating variation and landlords' rents (Proposition 3). Nonetheless, the necessary conditions for this Proposition do not hold for quasi-linear utility functions, which most of the papers of this literature stream consider.

nor maximization of a welfare function defined as the monetary value of households utility and absentee landlord rents. Furthermore, while the market equilibrium minimizes resource usage in (and only in) the absentee landlord case, it does not maximize residents' utility nor the sum of residents' surplus and land rents. The same result holds for almost any land ownership structure, and we show that the effect is sizable, strongly impacting policy conclusions.

3 A general monocentric model for land rents distribution

We consider L identical households that locate along a closed linear city parametrized by $0 \leq x \leq \bar{x}$, where the CBD and the endogenous city boundary are located at $x = 0$ and $x = \bar{x}$ respectively.[¶] All city residents work at the CBD, and commute from their residential location using a single mode (e.g., private car). Commuting costs increase linearly with the distance to the CBD and include both the time cost of the trip and the operating cost of the vehicle. Every household maximizes its utility, represented by a quasi-concave function $u(c,q)$ that depends on the consumption of a composite good c and housing q (measured by the floor size), both depending on the location. Thus, in this classical monocentric city model, the trade-off is between accessibility and dwelling size. The price of c is normalized to one irrespective of location, while the price of q is denoted by p and varies with location. At the city boundary, the residential rent $p(\bar{x})$ must be equal to the exogenous agricultural rent r_a. Finally, households' income has two components: a fixed amount y, and a share of the excess land rents, which depends –endogenously– on the land ownership structure, possible spatially variant subsidies and taxes, and the resulting equilibrium

We now refer to the land ownership structure. As Section 2 stresses, different approaches have been considered to study the optimal land use in the monocentric city model. In particular, one of the approaches considers a "public ownership" structure, where all land rents are captured, while the other two consider an absentee landlord, where all land rent vanishes from the economy. We propose a unifying framework (in line with Papageorgiou and Pines, 1999) where the government is only able to capture a fraction $\mu \in [0,1]$ of the excess land rents, which are then distributed lump-sum to all residents. Con-

[¶]Considering a circular city with a dense radial road network does not change any of the results of this Section. The only difference would be that the density function has to be multiplied by $\theta(x)$, where $\theta(x)$ is the number of radians of land available at each x.

sidering $\mu = 1$ leads to the most common setting used to study welfare in a monocentric city model, where all the rents are redistributed equally among residents, while considering $\mu = 0$ leads to the textbook treatment of the monocentric city model, where landlords are absentee. With this framework, the households' income constraint is written as follows:

$$\frac{\mu R}{L} + y = tx + c(x) + p(x)q(x) \tag{1}$$

On the left-hand side of Equation (1), R denotes the excess land rents, defined as:

$$R = \int_0^{\bar{x}} p(x) - r_a \, dx \tag{2}$$

On the right-hand side of Equation (1), t denotes the commuting cost per mile.

A natural interpretation of our framework is to consider an agent that owns the land, but that does not affect the urban equilibrium. Still, the city government is able to impose a tax rate on the land value from rents equal to μ, which is distributed equally among residents. In view of this interpretation, we believe that a more sensible approach is to consider $0 < \mu \ll 1$. For instance, in the US, rental income is subject to ordinary tax rates, ranging from 10% to 37%. Furthermore, since the US tax system treats landlords as business entities, expenses (repairs, maintenance, mortgage interest, etc.) can be used to offset the taxable rental income (Sommer and Sullivan, 2018), lowering the effective tax rate. Note that for this range of values, no formal treatment of the Rawlsian welfare function has been proposed in the literature.

To study whether the market equilibrium outcome is efficient, we consider the spatially variant taxes that decentralize the Rawlsian first-best. That is, we explore whether or not there is an equilibrium with higher equilibrium utility than the market outcome, that can be achieved using a revenue-neutral combination of taxes and subsidies, $\tau(x)$. If the optimum is achieved with $\tau \equiv 0 \; \forall \, x$, then a Rawlsian planner would choose the market outcome, implying that the city size, and the allocation of people to dwellings is the best that can be achieved in equilibrium. If $\tau \not\equiv 0$, then the market equilibrium is inefficient.

The Ralwsian Planner problem associated with this formulation is then:

$$
\begin{aligned}
(P) \qquad & \max \quad \overline{u} & & \\
& \text{s.t.} \quad u(c(x), q(x)) = \overline{u}, & 0 \le x \le \overline{x} \qquad & \text{(3a)} \\
& \int_0^{\overline{x}} \frac{1}{q(x)} dx = L & & \text{(3b)} \\
& \frac{\mu R}{L} + y - tx - \tau(x) - c(x) - p(x)q(x) = 0 & 0 \le x \le \overline{x} \qquad & \text{(3c)} \\
& R = \int_0^{\overline{x}} p(x) - r_a \, dx & & \text{(3d)} \\
& 0 = \int_0^{\overline{x}} \frac{\tau(x)}{q(x)} dx & & \text{(3e)} \\
& p(\overline{x}) = r_a & & \text{(3f)}
\end{aligned}
$$

In problem (P), (3a) restricts the outcome to be an urban equilibrium, (3b) establishes that the total urban population L fits inside the city, (3c) are the income constraints, (3d) defines the excess land rents, (3e) restricts the instrument τ to be revenue-neutral, and (3f) indicates that the residential rent equals the agricultural rent at the city boundary. It turns out the optimum of the planner's problem has a τ different from zero:

Proposition 1. *The market outcome does not achieve the Rawlsian first-best for any $\mu < 1$. The first-best Rawlsian allocation can be decentralized by a planner with a revenue-neutral combination of location-specific taxes and subsidies given by*

$$\tau(x) = (1 - \mu) \frac{p(x)q(x)}{|\sigma|} + C$$

where σ is the income-compensated price elasticity of demand for housing,[‖] *and C is a nonpositive constant that ensures that $\tau(x)$ is revenue-neutral.*

Proof. See Appendix A. □

[‖]That is $\sigma(x) = \frac{\partial p(x)}{\partial x} \bigg/ \frac{\partial q(x)}{\partial x} \cdot \frac{q(x)}{p(x)}$

Corollary 1. *A sufficient condition for $\tau'(x) < 0\ \forall x \in [0,\bar{x}]$ is that σ is a non-increasing function of x and $\sigma > -1$. In particular, this holds for Cobb-Douglas utility functions $u(c,q) = c^{1-\alpha}q^{\alpha}$, and for CES utility functions $u(c,q) = ((1-\alpha)c^{\rho} + \alpha q^{\rho})^{\frac{1}{\rho}}$ with $\rho < 0$.*

Proof. See Appendix B □

Proposition 1 shows that, except for the extreme case of public ownership, i.e. $\mu = 1$, there is a revenue-neutral location-specific tax scheme that can achieve an equilibrium with higher welfare. Additionally, under very mild conditions, the tax decreases along the city. In that case, $\tau(x)$ translates into a tax in the inner city and a subsidy in the outer city, since it was constructed as revenue-neutral.[**]

The result presented in Proposition 1 has profound implications: we now show that, if $\tau(x)$ decreases with distance from the CBD, then the market equilibrium city in the absence of externalities is too compact from a welfare standpoint. For this, we present a more general result:

Lemma 1. *Consider any revenue-neutral combination of location-specific taxes and subsidies, $\tau(x)$, that decreases with distance ($\tau'(x) < 0\ \forall x$) and that yields a higher equilibrium utility ($\tilde{u}$) than the market equilibrium utility ($\overline{u}$). The implementation of τ leads to a city that is more extended than the market equilibrium city.*

Proof. Appendix C □

The intuition behind Lemma 1 is straightforward: since we are considering a revenue-neutral instrument $\tau(x)$, with $\tau'(x) < 0$, the housing price at the CBD must decrease compared to the market city since we are providing incentives to locate in the outer city. Then, the bid rents curves of the market city and the city under the instrument τ must cross once. If they do not, then the bid rent curve of the city under τ is always below the bid rent curve of the market city, implying that the city under τ is more compact. But this also means higher consumption of housing everywhere in the city under τ, indicating that it is more extended than the market city. Using this lemma, we present the following proposition regarding the efficiency of the market equilibrium city's size:

[**]Formally, $\tau(x) > 0\ \forall x \in [0,x^*)$ and $\tau(x) < 0\ \forall x \in (x^*,\bar{x}]$ for some $x^* \in (0,\bar{x})$.

Proposition 2. *Under the conditions of Corollary 1, for any $\mu < 1$, the market equilibrium city is too compact from a Rawlsian welfare standpoint.*

Proof. The proof follows directly from Corollary 1 and Lemma 1 □

To shed light on the intuition behind Propositions 1 and 2, it is useful to identify the forces at play when a planner taxes the central areas, subsidizes the outskirts and distributes whatever fraction of land rents she can capture. We focus on the effect on the city residents located arbitrarily close to the CBD for ease of explanation, but the analysis is general as the utility is the same along the city. The impact of the tax on these residents has three components: (i) their income is reduced, since they are taxed to finance the redistribution, (ii) for any $\mu > 0$, their income is further reduced since the excess land rents decrease when the city grows, and (iii) they face a lower housing price, and therefore, increase their consumption of housing (substitution effect). The first two impacts reduce residents' utility, while the third one increases it.

To provide an intuitive yet more analytical explanation, suppose the combination of taxes and subsidies takes the linear form $\tau(x) = ax + b$, with $a < 0$, and $b > 0$. We evaluate the sign of the change in utility when implementing this linear tax, starting from the unpriced equilibrium; that is, both a and b are (close to) zero. Consider an infinitesimal change of the slope of the tax, da, together with an accompanying adjustment of b to maintain the revenue-neutrality. The utility change of city residents located at the CBD when changing a from zero to -da is given by $-d\bar{u}/da$:

$$-\frac{d\bar{u}}{da} = v_c(x=0)\left(\frac{db}{da} - \frac{\mu}{L}\frac{dR}{da} + q(x=0)\frac{dp}{da}\right) \tag{4}$$

where v is the indirect utility function, and $v_c(x=0)$ is its derivative with respect to consumption, evaluated at prices, income and consumption at the CBD. In Appendix D, we show that for $\mu = 1$, the following holds:

$$-\frac{db}{da} + \frac{1}{L}\frac{dR}{da} - q(x=0)\frac{dp}{da} = 0 \tag{5}$$

In other words, the three effects cancel out when all the excess land rents are redistributed. In view of this, it is easy to see why this redistribution is utility-increasing when $\mu < 1$: part of the second effect in equation (4) is borne by the absentee landlord, leaving an overall positive effect for the city residents.

It is important to highlight that the results presented up to this point are rooted in the spatial setting of the urban equilibrium and are not due to the absentee landlord's exclusion from the welfare function, as the previous argument may suggest. Indeed, consider the following social welfare function

$$SW = L \cdot CV(\mu) + (1-\mu)R \tag{6}$$

where $CV(\mu)$ corresponds to the compensating variation of residents with respect to the market equilibrium, when they receive a share μ of the land rents. With this welfare function, we are considering that the marginal utility of income for the absentee landlord is one. If this is also the case for the city residents (for example, when using a quasi-linear utility function), then the three effects mentioned earlier once again cancel out for any value of μ. But this is not the general case and, as Proposition 3 shows, there is a large class of utility functions where the gain in utility measured as the compensating variation is greater than the loss in the absentee landlord's rents:

Proposition 3. *Suppose $u(c,q)$ is strictly quasi-concave and $\sigma > -1$, where σ is the income-compensated price elasticity of demand for housing. Then, the market equilibrium does not maximize the social welfare function given by $SW = L \cdot CV + (1-\mu)R$ for any $\mu < 1$. In this case, a revenue-neutral redistribution of income from the city center to the outskirts leads to higher welfare in the new equilibrium.*

Proof. See Appendix D. □

Proposition 3 is important, since a number of papers in the urban economics literature study welfare in a monocentric city with an absentee landlord, considering a welfare function composed of residents' utilities in monetary terms plus the absentee landlord rents (e.g., Kono et al., 2012; Kono and Joshi, 2012; Li et al., 2013; Li and Guo, 2017; Li and Wang, 2018). Most of these papers consider for convenience, however, a quasi-linear utility function. Note that for this type of utility function, Proposition 3 does not apply, since the utility function is not strictly quasi-concave. Hence, Proposition 3 shows that the choice of utility function is not innocuous, and the use of a different utility form (for example, Cobb-Douglas or CES) would end up producing different results and policy conclusions.

3.1 Relation to minimization of resources

So far, we have focused on two of the three approaches described in Section 2 for analyzing welfare: maximization of a welfare function, defined as either the households' equilibrium utility, or the households' compensating variation plus the absentee landlord rents. In this subsection, we discuss the relationship between our results and the minimization of resources approach. For this, we generalize the efficiency result of Fujita (1989) and Fujita and Thisse (2013) that says that, in the case of an absentee landlord, the market outcome minimizes resources, understood as the sum of of total non-land consumption, total commuting cost, and the opportunity cost of the urban land. The generalization for all land ownership structures ($\mu \in [0,1]$) is:

Proposition 4. *The market equilibrium, subject to a fixed utility, minimizes a resource function, defined as the sum of total non-land consumption, total commuting cost, and the opportunity cost of the urban land, minus the share of rents received by city residents. In other words, the market equilibrium solves the following minimization problem:*

$$(P_{\min}) \qquad \min \quad \left(\int_0^{\bar{x}} \frac{tx + c(x)}{q(x)} + r_a \, dx \right) - \mu R$$

$$s.t. \quad u(c(x), q(x)) = \bar{u}, \qquad 0 \leq x \leq \bar{x} \tag{7a}$$

$$\int_0^{\bar{x}} \frac{1}{q(x)} dx = L \tag{7b}$$

$$R = \int_0^{\bar{x}} p(x) - r_a \, dx \tag{7c}$$

$$p(\bar{x}) = r_a \tag{7d}$$

Proof. See Appendix E. □

Note that when $\mu = 0$ (absentee landlord case) in Proposition 4, we recover the resource function used by Fujita (1989) and Fujita and Thisse (2013). Nonetheless, the resource function that is implicitly minimized by the market equilibrium is different than the one used by Fujita (1989) and Fujita and Thisse (2013) for any other land ownership structure. We now turn to the connection between resource minimization and the Rawlsian approach, for which we provide the following equivalence:

Proposition 5. *Consider the following Rawlsian optimization problem* $(P_{\max})$*:*

$$
\begin{aligned}
(P_{\max}) \qquad & \max \quad \bar{u} \\
& s.t. \quad u(c(x), q(x)) = \bar{u}, \qquad 0 \le x \le \bar{x} && \text{(8a)} \\
& \int_0^{\bar{x}} \frac{1}{q(x)} dx = L && \text{(8b)} \\
& \frac{\mu R}{L} + y - tx - c(x) - p(x)q(x) = 0 \qquad 0 \le x \le \bar{x} && \text{(8c)} \\
& R = \int_0^{\bar{x}} p(x) - r_a \, dx && \text{(8d)} \\
& p(\bar{x}) = r_a && \text{(8e)}
\end{aligned}
$$

Problem $(P_{\max})$ *is equivalent to the minimization of a resource function, defined as the sum of total non-land consumption, total commuting cost, the opportunity cost of the urban land, and the share of rents received by the absentee landlord. In other words, problem* $(P_{\max})$ *is equivalent to:*

$$
\begin{aligned}
(\tilde{P}_{\min}) \qquad & \min \quad \left(\int_0^{\bar{x}} \frac{tx + c(x)}{q(x)} + r_a \, dx \right) + (1-\mu)R \\
& s.t. \quad u(c(x), q(x)) = \bar{u}, \qquad 0 \le x \le \bar{x} && \text{(9a)} \\
& \int_0^{\bar{x}} \frac{1}{q(x)} dx = L && \text{(9b)} \\
& R = \int_0^{\bar{x}} p(x) - r_a \, dx && \text{(9c)} \\
& p(\bar{x}) = r_a && \text{(9d)}
\end{aligned}
$$

Proof. Appendix F. □

Corollary 2. *The sum of total non-land consumption, total commuting cost, and the opportunity cost of the urban land are minimized by the market equilibrium only when* $\mu = 0$*, while the minimization of* *those resources is equivalent to the maximization of equilibrium utility if and only if $\mu = 1$.*

As Propositions 4 and 5 show, the market equilibrium is the solution to a modified resource minimization problem, but this problem is never equivalent to the maximization of equilibrium utility. Furthermore, both problems never present the same objective function (for a given μ).[††] In view of these Propositions, Fujita (1989)'s result, which refers to Proposition 4 when $\mu = 0$, does not relate or allow us to conclude about a Rawlsian planner. Importantly, this implies that it is wrong to argue that, because in the case of the absentee landlord case resources are minimized in the market equilibrium, a planner would choose that same allocation.

3.2 Numerical Analysis

In order to provide a more quantitative assessment of the impacts of our results, we simulate the urban equilibrium structure of a city. We base our numerical analysis on US values previously used in the literature, with parameters summarized in Table 1. We use, for practical purposes, a Cobb-Douglas utility function $u(c,q) = c^{1-\alpha}q^{\alpha}$, setting $\alpha = 0.35$. This implies that every household spends 35% of their income net of transportation costs on housing, which is consistent with the average expenditure reported by the US Department of Labor (2018)'s Consumer Expenditure Survey.[‡‡] The hourly wage is set at US$25.27 which, following Bertaud and Brueckner (2005), is obtained using the median income per household of the 2018 US census (US$63,179) and assuming 2000 work hours/year. For the agricultural rent value, we also follow Bertaud and Brueckner (2005): we consider the average US agricultural land value for the year 2019 of US$3,160/acre. Assuming a 5% discount rate, we arrive to $r_a =$ US$101,120/sq. mi. We set $L = 120,000$ households, which would be equivalent to roughly $800,000$ households in a circular (rather than linear as in our model) city.

In Figure 1 we show, for different values of μ ranging from $\mu = 0.1$ to $\mu = 0.9$, the resulting combina-

[††] Note that the objective function of the problems presented in Propositions 4 and 5 are the same when $\mu = 0$ in the first problem, while $\mu = 1$ in the second one. In this case, both problems aim to minimize the sum of total non-land consumption, total commuting cost, the opportunity cost of the urban land, as considered by Fujita (1989) and Fujita and Thisse (2013). Nonetheless, in this case both problems refer to different land ownership structures, with different underlying conditions (e.g., different households' income). Thus, even in this case they are not equivalent.

[‡‡] The US Department of Labor (2018)'s Consumer Expenditure Survey reports an average income net of taxes and transportation costs of US$57,480, of which US$20,091 is spent on housing.

Table 1: Main parameter values.

Parameter	Value
α	0.35
y [$]	63,179
r_a [$/sq. mi]	101,120
L [households]	120,000
μ	$\{0.1, 0.5, 0.9\}$

Table 2: Comparison of market equilibrium against Rawlsian optimum.

Equilibrium	Market	Optimal	Market	Optimal	Market	Optimal
μ	0.1		0.5		0.9	
Equilibrium utility $\overline{u}$	49.56	50.41	55.45	55.80	62.93	62.95
City extension [miles]	56.51	73.35	63.23	73.73	71.76	74.15
Aggregate land rents [millions US$]	$2,013.55$	$1,782.36$	$2,252.88$	$2,084.78$	$2,556.78$	$2,510.83$

tion of taxes and subsidies for every location; Table 2 presents the characteristics of the market outcome as compared to the Rawlsian first-best. It can be seen that reallocating income from the city center (imposing a tax) to the outskirts (using a subsidy) proves to be useful from a Rawlsian perspective, increasing equilibrium utility. Additionally, the lower the value of μ, the more aggressive the optimal redistribution strategy is. For example, if $\mu = 0.1$, for the people living closest to the city center, a centralized planner would like to impose a tax that represents 11.40% of their net income. On the other hand, when $\mu = 0.9$, the optimal tax is only 1.78% of their net income. This redistribution has sizable impacts on the extension of the resulting city: compared to the market equilibrium city, the Rawlsian first-best city is up to 16.8 miles more extended, representing a 29.8% increase for $\mu = 0.1$. This result shows that the effects of the land ownership structure over the Rawlsian first-best urban form are far from negligible. As demonstrated above, city size from the market outcome approaches city size in the Rawlsian first-best as the percentage of land rents that are captured and redistributed increases. But the difference becomes "small" only when the percentage of land capture is really high. Even with 50% of land rent capture, the market outcome delivers a city that should expand more than 15% in extension.

As a side effect of this redistribution, rental prices in the city center decrease (Figure 2). For instance, in the CBD, rental prices reduces by 32.61%, 21.20% and 5.11% for $\mu = 0.1$, $\mu = 0.5$ and $\mu = 0.9$, respectively. This implies in return, a lower amount of excess land rents. Finally, although the utility

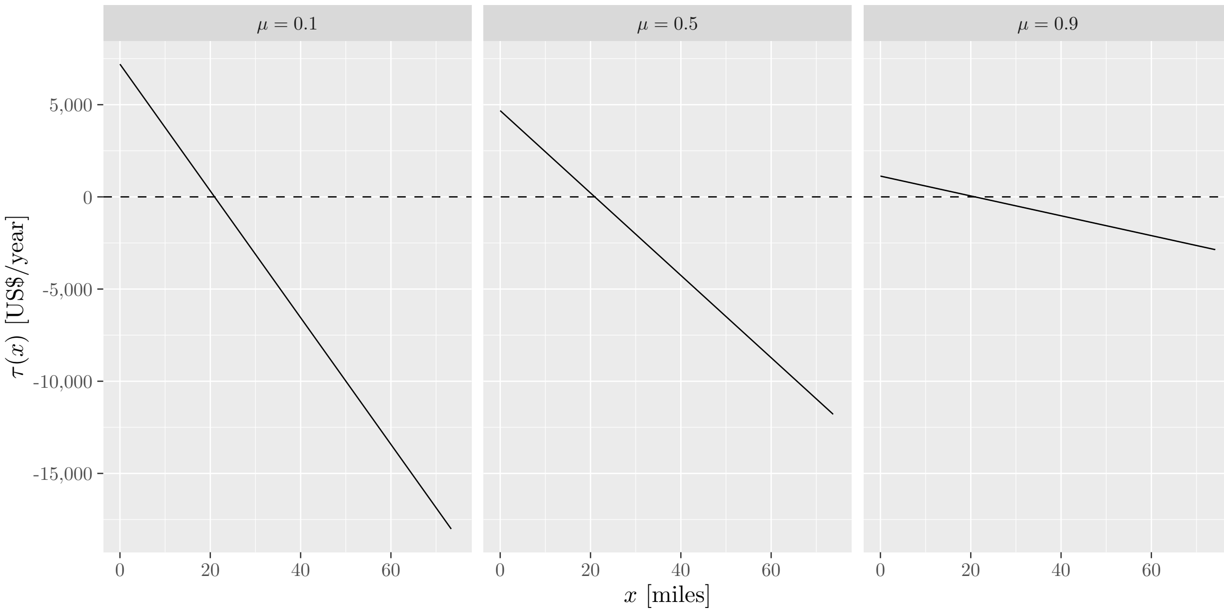


Figure 1: Optimal $\tau(x)$ for different values of μ.

gains when using the optimal redistribution policy seem modest (ranging from 0.03% to 1.71% in our example, as Table 2 shows), this is expected and in line with estimations of surplus gains when using, for example, optimal road pricing policies (Verhoef, 2005). In the following sections, we show that the redistribution scheme proposed here can be understood as part of the pricing decision of different transport systems when we introduce a more realistic setting. By studying how the optimal pricing scheme affects and interacts with optimal congestion pricing and public transport pricing, we can also shed more light on the size and relevance of the finding.

4 Congestion Pricing

For the first application of the results provided in Section 3, we now consider that residents commute from their residential location using a single mode (private car),[§§] which is the cause of a single externality: traffic congestion. That is, the commuting costs for each person depends on the number of other people using the road. In particular, we follow the traditional approach of Arnott (1979c) and assume that the

[§§]We make no assumptions on land needs for roads, which can actually be set to zero without losing generality.

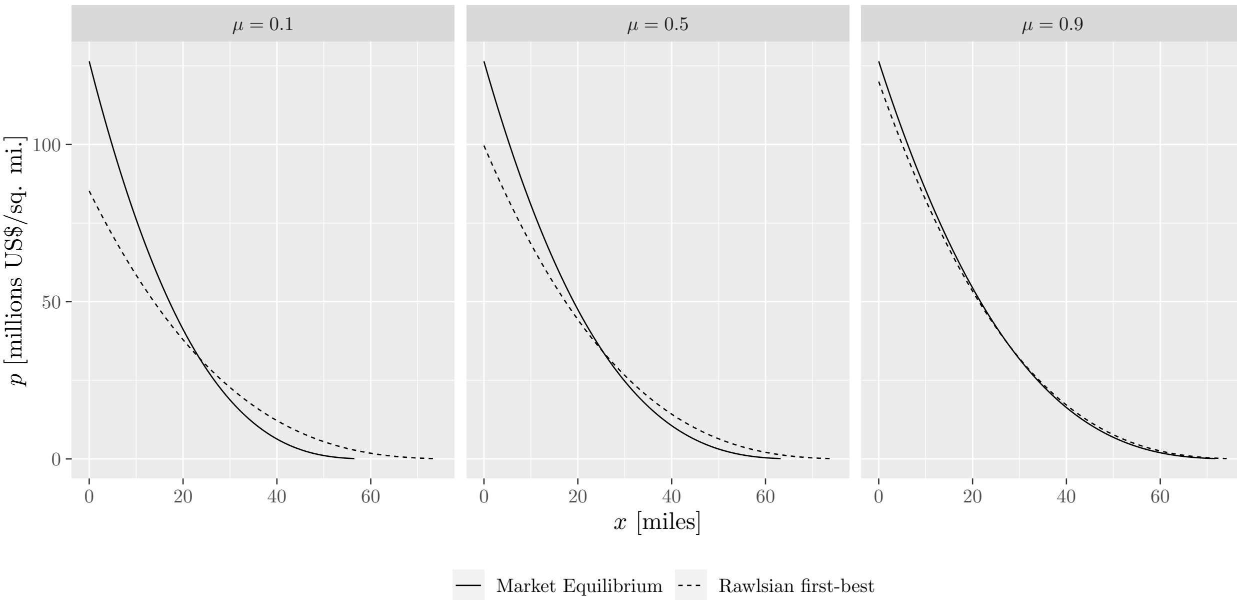


Figure 2: Rental prices - Market equilibrium versus Rawlsian first-best.

commuting costs incurred by a resident living at a distance x from the CBD is

$$t(x) = \int_0^x g(T(z))dz$$

where $T(x)$ is the total traffic flow at x, and g is such that $\frac{dg}{dT} > 0$. Since in our model $\dfrac{1}{q(x)}$ denotes the number of people located at distance x from the CBD, the total traffic flow can be written as:

$$T(x) = \int_x^{\bar{x}} \frac{1}{q(z)}dz$$

The transportation costs incurred are then the (personal) commuting time costs $t(x)$ and, if levied, a congestion toll $\tau(x)$, which is the only spatially variant charge we consider in this section. Under this setting, when considering public ownership, it is a well-known result that the first-best Rawlsian congestion toll corresponds to the marginal external cost of every trip (e.g., Kanemoto, 1980). Second-best policies, such as cordon tolls or flat pricing schemes, have been found to perform well, leading to welfare gains comparable to those of the first-best policies (e.g., Verhoef, 2005), and producing a similar effect in cities' extension: they reduce aggregate mileage by means of shrinking the city when compared to the unpriced situation. Even though some authors have argued that the optimal congestion toll should be smaller than the marginal external cost, this is only due to pre-existing distortionary taxes in the labor market (see, e.g., Parry and Bento, 2001; Tikoudis et al., 2015).

We now turn to our more general setting of land ownership, rent capture and redistribution. As before, we consider that residents earn a fixed income y and a redistribution of a fraction μ of the excess land rents, but now they also receive a lump-sum redistribution of the congestion toll revenues. For simplicity we assume that the costs of operating the tolling system is zero. We then search for the welfare-maximizing congestion toll $\tau(x)$:

$$
\begin{aligned}
(P_{\text{Congestion}})\ \max \quad & \overline{u} & & \\
\text{subject to} \quad & u(c(x),q(x)) = \overline{u}, & 0 \leq x \leq \overline{x} \quad & \text{(10a)} \\
& t(x) = \int_0^x g(T(z))dz, & 0 \leq x \leq \overline{x} \quad & \text{(10b)} \\
& T(x) = \int_x^{\overline{x}} \frac{1}{q(z)}dz, & 0 \leq x \leq \overline{x} \quad & \text{(10c)} \\
& \int_0^{\overline{x}} \frac{1}{q(x)}dx = L & & \text{(10d)} \\
& \frac{\mu R}{L} + \frac{G}{L} + y - t(x) - \tau(x) - c(x) - p(x)q(x) = 0 & 0 \leq x \leq \overline{x} \quad & \text{(10e)} \\
& R = \int_0^{\overline{x}} p(x) - r_a \, dx & & \text{(10f)} \\
& G = \int_0^{\overline{x}} \frac{\tau(x)}{q(x)}dx & & \text{(10g)} \\
& \min_{x\in[0,\overline{x}]} \tau(x) = 0 & & \text{(10h)}
\end{aligned}
$$

Equation (10g) defines the congestion toll revenue G, which is then redistributed in equal shares to city residents (Equation 10e). Equation (10h) is a border condition to guarantee a unique solution to problem ($P_{\text{Congestion}}$). Without this condition there are infinitely many solutions: since we are considering a single transport mode and the toll revenue G is redistributed lump-sum, if $\hat{\tau}(x)$ is a solution, then $\hat{\tau}(x) + C$ is also a solution, for any constant C. We choose, for convenience, border condition (10h) since this ensures that τ is indeed a toll, i.e. $\tau(x) \geq 0 \; \forall x \in [0, \bar{x}]$.

Proposition 6. *The first-best Rawlsian equilibrium for $P_{Congestion}$ can be achieved through the use of a space variant toll given by:*

$$\tau(x) = MEC(x) + (1-\mu)\frac{p(x)q(x)}{|\sigma|} + C$$

where $MEC(x)$ is the marginal external cost of a trip originated at x, σ is the income-compensated elasticity of demand for housing, and C is a nonpositive constant that ensures the fulfillment of border condition (10h), that is:

$$C = -\min_{x\in[0,\bar{x}]}\left\{MEC(x) + (1-\mu)\frac{p(x)q(x)}{|\sigma|}\right\}$$

Proof. Consider the proof presented in Appendix A, with $\tilde{\tau}(x) = \tau(x) - \frac{G}{L}$ (that is, $\tilde{\tau}$ is revenue neutral). □

Proposition 6 presents a result that goes in line with results in previous sections: in any situation other than public ownership, even in the absence of other externalities, the first-best Rawlsian congestion toll is no longer the marginal external cost, and is, under some conditions, flatter (according to Corollary 1). It is in fact the direct composition of the two effects: a Rawlsian planner would impose a congestion charge higher than the marginal external cost for people living close to the city center, and lower for the residents located in the outskirts. We now turn to a numerical analysis to illustrate the size of the effect relative to the marginal external cost of congestion.

4.1 Numerical Analysis

We use the parameters shown in Table 1. Additionally, in this case, we consider a commuting cost per mile given by the Bureau of Public Roads (BPR) function, in line with recent papers studying congestion pricing in the monocentric city model (e.g. Li et al., 2012; De Lara et al., 2013; Tikoudis et al., 2018). In particular, we consider the following BPR function

$$g(T(x)) = t_0 + c_t t_1 \left[1 + \rho_1 \left(\frac{T(x)}{K} \right)^{\rho_2} \right] \tag{11}$$

where t_0 is the per-mile monetary cost of the trip, c_t is the value of travel time, t_1 is the free-flow travel time per mile, $T(x)$ is the total traffic at x, K is the capacity of the road, and ρ_1 and ρ_2 are positive constants.

For the monetary cost, t_0, we use the current US federal allowance for business mileage of US$0.575/mile, which translates to $t_0 =$ US$359.38/mile when considering 1.25 workers/household, 250 working days/year, and two daily trips. To obtain the time cost component, we assume that the commuting time is valued at the hourly wage rate, i.e., $c_t =$ US$25.27. Then, we consider a free-flow speed of 50 miles/h,[¶¶] leading to an annualized free-flow travel time of $t_1 = 12.5$ h/mile. Finally, we assume a road capacity of $K = 54,000$ veh/h and we consider widely used parameter values for the BPR function, $\rho_1 = 0.15$, $\rho_2 = 4$ (Small and Verhoef, 2007). These values produce a city that presents an average commuting speed[***] of 22 miles/h and an average commuting time in the range [47,59] minutes. Both the average travel speed and the average commuting time are roughly consistent with those reported by the 2017 National Household Travel Survey (McGuckin and Fucci, 2018).

Table 3 provides a summary of the results, comparing the unpriced equilibrium against two congestion pricing policies: on the one hand, charging the marginal external cost of every trip (the Pigouvian toll), and on the other hand, charging the optimal toll from Proposition 6. The utility gains when charging the optimal congestion pricing toll are up to 1.1% and 1.7% with respect to the city with no pricing and to the city under marginal cost pricing, respectively. As discussed in Verhoef (2005), this is expected and in line with estimations of surplus gains from optimal road pricing in urban areas. Nevertheless, charging the optimal congestion toll has a sizable effect on the urban form. In Figure 3, we show that for the parameters considered, the redistributing component of the optimal pricing scheme offsets the

[¶¶]We arrive at this number assuming that, on average, 80% of every trip is made using highways at a free-flow speed of 60 miles/h, while the other 20% uses main roads at a free-flow speed of 20 miles/h. This gives a free-flow speed of 52 miles/h, which we round to 50 miles/h.

[***]To obtain the travel speed in our model, we consider the time component of $g(T(x))$ to obtain the equivalent travel time, and with this, the travel speed.

Table 3: Equilibrium results - No congestion pricing vs marginal cost pricing vs optimal pricing.

Congestion Policy	Unpriced	Pigouvian	Optimal	Unpriced	Pigouvian	Optimal	Unpriced	Pigouvian	Optimal
μ		0.1			0.5			0.9	
Equilibrium utility $\overline{u}$	52.12	51.66	52.55	57.80	57.79	58.16	64.88	65.58	65.60
City extension [miles]	76.80	72.72	94.07	85.17	81.36	94.68	95.60	92.32	95.36
Aggregate land rents [millions US$]	1,862.97	2,012.29	1,780.94	2,066.04	2,251.31	2,083.11	2,318.79	2,554.77	2,508.82
Aggregate mileage [thousand miles]	2,551.57	2,216.74	2,911.85	2,829.70	2,480.04	2,912.14	3,175.91	2,814.31	2,912.30

marginal external cost, resulting in a non-monotonic toll that decreases with distance in the outer city. This result has profound implications. First, the optimal city ends up being more extended than the unpriced city for a wide range of values of μ: for $\mu = 0.1$ the city under the optimal pricing is 17.3 miles longer than the unpriced city, representing a 22.5% increase. For $\mu = 0.5$ the optimal city is 11.2% longer that the unpriced city. Only when μ reaches 0.9 is the optimal city marginally smaller than the unpriced city; note that the optimal toll is still non-always increasing. This result indicates that, from an efficiency standpoint, it is preferable to implement a toll schedule that produces urban sprawl and increases aggregated mileage, allocating money from the city center to the outskirts.

At the same time, the optimal toll reduces the excess land rents for low values of μ, by means of decreasing rental prices in the inner city. This reduction is once again sizable (4.4% for $\mu = 0.1$). A straightforward implication follows: for most values of μ, charging only the marginal external cost operates precisely in the opposite way as the optimal toll, reducing the extension of the city and the aggregate mileage, while increasing the excess land rents. Thus, in these situations –as Figure 4 depicts– it might be even welfare decreasing compared to the no pricing situation. It is only for high values of μ that the optimal toll is closer to the Pigouvian toll, producing similar effects on the urban form: for $\mu = 0.9$, the city under the optimal pricing compared to the unpriced city is smaller (0.2 miles or a 0.25% reduction), with a decreased aggregate mileage (-8.3%), and leads to an increase in excess land rents (8.2%).

5 Public Transport Pricing

We now turn to the second application: public transport pricing. We consider the same monocentric city model used in the previous section, but now residents commute to the CBD along a corridor where a public transport system is the only available transport option. This system could be, for instance, a

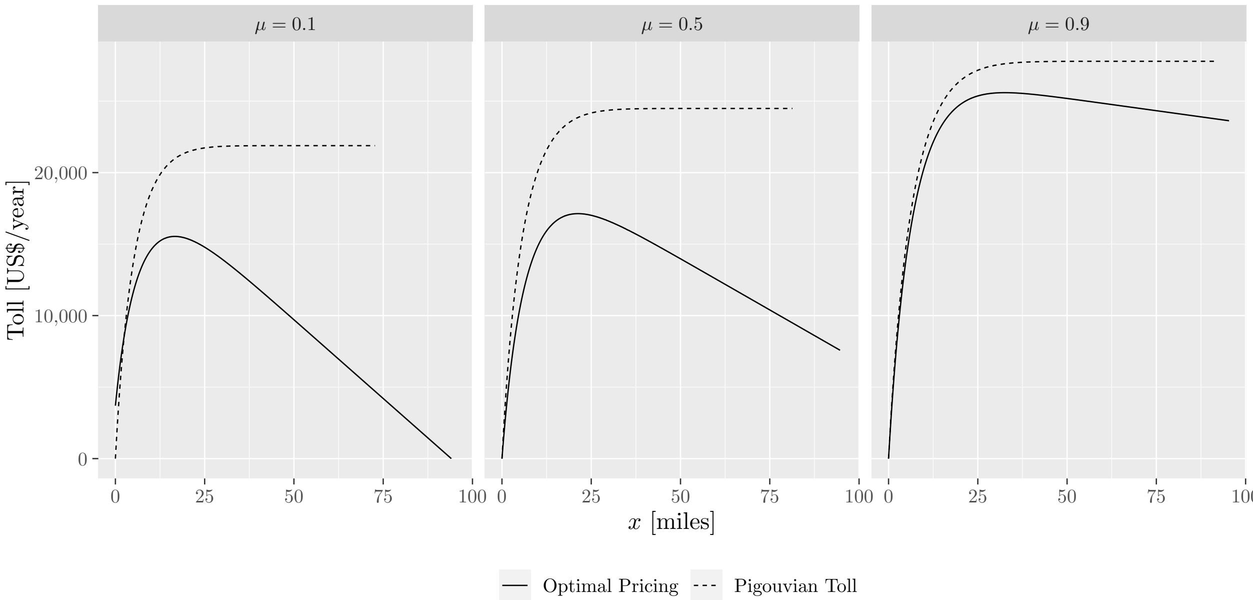


Figure 3: Pigouvian versus optimal toll.

rapid transit system, a tramway, or a subway line. Commuting costs incurred involve now the time cost of the trip and the public transport trip fare, which may be different across locations. As in Brueckner (2005), we consider a constant returns to scale production function for the transport system, with no externalities. The cost of providing public transport is therefore proportional, with constant k, to the number of passenger-miles. These simplifications imply that the social marginal cost of an additional passenger mile is given by k.

Using the same setting, Brueckner (2005) studied the desirability of public transport subsidies. For this, he considered that commuters living at x faced only a fraction $\delta \in [0,1]$ of the full money cost kx. The difference, $(1-\delta)kx$, was assumed to be subsidized and covered by general, spatially invariant, tax revenue. Thus, every resident faced a head tax H:

$$H = \frac{1}{L}\int_0^{\bar{x}} \frac{(1-\delta)kx}{q(x)}dx \tag{12}$$

Note that since there is only one mode available in both our and Brueckner's model (transit), another possible interpretation of the head tax is that commuters face a transit fare with a distance-related component (δkx) and a fixed component (H). Thus, for example, a fully subsidized public transport system

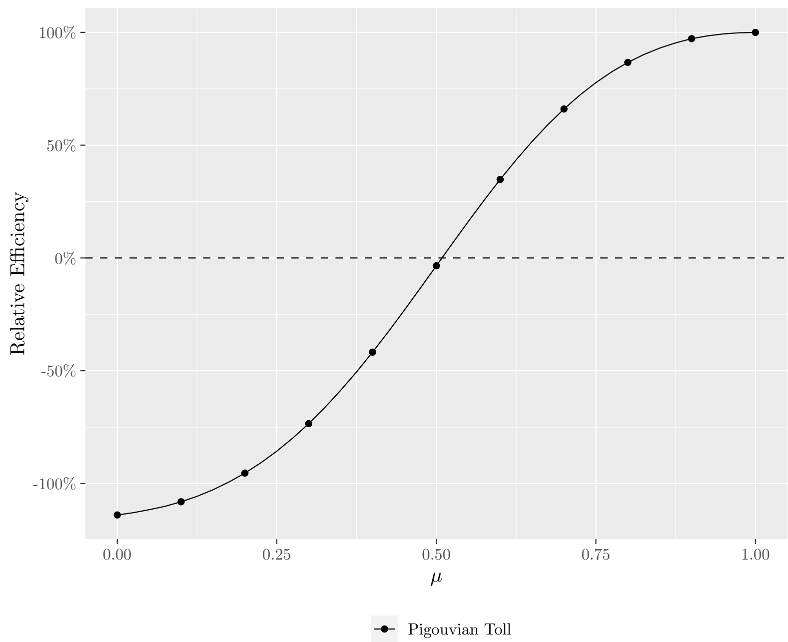


Figure 4: Relative efficiency of marginal external cost pricing (Pigouvian toll) vs no tolling as a function of land rent capture μ.

under Brueckner's interpretation translates to a public transport system with a fixed, spatially invariant, fare under the second interpretation.

Based on the result of Fujita (1989), Brueckner (2005) shows that subsidies are inefficient when minimizing total non-land consumption, total commuting cost, and the opportunity cost of the urban land, making the urban expansion they produce undesirable. Thus, every resident should face the full money cost of commuting at each location, which calls for an increasing public transport fare with distance. Even though other authors have claimed that the relationship between optimal fares and distance is not necessarily positive (e.g. Mohring, 1972; Turvey and Mohring, 1975), their argument hinges on the presence of negative externalities such as boarding delays and crowding. In this section, we show that if one considers any scenario other than public land ownership, then the public transport fare that decentralizes the first-best scenario is below marginal cost, and thus the system should be subsidized. This result does not come from externalities, which are assumed away, or differences in income, as we assume that all residents are equal.

For this, we denote by $\tau(x)$ the sum of the transit fare plus the potential head tax levied to (optimally) subsidize the transport system. We search for its utility-maximizing value, subject to self-financing of the transport system. With this, the utility-maximization optimal control problem associated with this formulation is then:

$$
\begin{array}{llll}
(P_{PT}) & \max & \overline{u} & \\
& \text{subject to} & u(c(x),q(x)) = \overline{u}, \quad 0 \leq x \leq \overline{x} & \text{(13a)} \\
& & \displaystyle\int_0^{\overline{x}} \frac{1}{q(x)} dx = L & \text{(13b)} \\
& & \displaystyle\frac{\mu R}{L} + y - tx - \tau(x) - c(x) - p(x)q(x) = 0 \quad 0 \leq x \leq \overline{x} & \text{(13c)} \\
& & \displaystyle R = \int_0^{\overline{x}} p(x) - r_a \, dx & \text{(13d)} \\
& & \displaystyle k \int_0^{\overline{x}} \frac{x}{q(x)} dx = \int_0^{\overline{x}} \frac{\tau(x)}{q(x)} dx & \text{(13e)}
\end{array}
$$

Note that this problem (P_{PT}) is very similar to the problem (P) presented before. The only difference comes from the constraint (13e), where we impose self-financing of the public transport system. Thus, we are able to obtain a result analogous to Proposition 1:

Proposition 7. *The first-best Rawlsian optimal equilibrium for (P_{PT}) can be achieved through the use of a space variant public transport fare plus a head tax for subsidization given by*

$$\tau(x) = kx + (1-\mu)\frac{p(x)q(x)}{|\sigma|} + C$$

where σ is the income-compensated elasticity of demand for housing, and C is a nonpositive constant that ensures self-financing of the transport system.

Proof. Calling $\tilde{\tau}(x) = \tau(x) - kx$, (P_{PT}) can be reduced to the problem shown in (P), with $\tilde{\tau}(x)$ instead of $\tau(x)$. Thus, the same results of Proposition 1 hold for $\tilde{\tau}(x)$ this time. From this, the results shown in Proposition 7 follow immediately. $\square$

Note that Proposition 7 establishes that, in general, the optimal public transport fare is not equal to the social marginal cost and that the difference comes exactly from the combination of taxes and subsidies presented in the previous section. In other words, in this setting, the optimal redistribution can be decentralized through the transport fare plus a head tax (more on this below) and, moreover, under some mild conditions over σ, this produces a fare that is spatially flatter (Corollary 1). This finding provides support, on efficiency grounds, for a flatter pricing of transit systems, an argument that would be additional to those based on the existence of externalities or distributional concerns. Note that we can write the optimal $\tau(x)$ described in Proposition 7 as:

$$\tau(x) = \delta(x)kx + H \tag{14}$$

where $\delta(x)$ and H are given by:

$$\delta(x) = 1 - \frac{(1-\mu)}{kx}\left(\frac{p(0)q(0)}{|\sigma(0)|} - \frac{p(x)q(x)}{|\sigma(x)|}\right) \quad \forall x \in [0,\bar{x}] \tag{15}$$

$$H = (1-\mu)\frac{p(0)q(0)}{|\sigma(0)|} + C \tag{16}$$

Then, under conditions of Corollary 1, $\dfrac{p(x)q(x)}{|\sigma|}$ is decreasing with distance to the CBD. Due to this, is clear that in this case, $\delta(x)$ given by Equation (15) fulfills $\delta(x) \leq 1$. Additionally, this ensures that H given by Equation (16) is greater than zero.[†††]

Therefore, returning to Brueckner's interpretation, under conditions of Corollary 1, Proposition 7 also calls for subsidization of the public transport system: households living at x should only face a percentage ($\delta(x)$) of the full money cost. The difference must then be covered by a head tax H imposed

[†††] If this was not the case, then we would have that:

$$(1-\mu)\frac{p(x)q(x)}{|\sigma(x)|} + C < 0 \;\; \forall x \in (0,\bar{x}]$$

This, in turn, would imply a contradiction with condition (13e)

$$\int_0^{\bar{x}} \frac{\tau(x)}{q(x)}\, dx = \int_0^{\bar{x}} \frac{1}{q(x)}\left\{kx + (1-\mu)\frac{p(x)q(x)}{|\sigma|} + C\right\} dx < \int_0^{\bar{x}} \frac{kx}{q(x)}\, dx$$

on all households. Note, however, that the optimal fraction of the marginal cost to be paid ($\delta(x)$) varies with location, implying a distance-related transport fare that may be quite complex in practice. A natural simplification would be to restrict this fraction to be constant in space, so that the transport fare is linear or proportional to distance. This can be achieved by restricting the set of feasible fares schemes to those of the form $\tau(x) = ax + b$ (thus $a = \delta k$, for some constant δ). Under this constraint on the optimal fare, we are able to prove the following result:

Proposition 8. *If we restrict the transport fare to be linear with distance to the CBD,* $\tau(x) \in \{ax + b | (a,b) \in \mathbb{R}^2\}$*, then the second-best Rawlsian optimum equilibrium is achieved using*

$$\tau(x) = (k + \tilde{a}^*)x + b^*$$

where $\tilde{a}^*$ *solves* $\tilde{a}^* = (t + k - \mu(t + k + \tilde{a}^*))\left(1 - \frac{L^2 q(x=0)(t+k+\tilde{a}^*)}{R}\right)$*, and* $b^* \geq 0$ *is a spatially invariant head tax collected to subsidize the system and ensure self-financing. Additionally, if the spending in housing pq decreases with the distance to the CBD, then*

$$\begin{cases} \tilde{a}^* < 0 \;\; \textit{and} \;\; b^* > 0 & \forall \mu \in [0,1) \\ \tilde{a}^* = 0 \;\; \textit{and} \;\; b^* = 0 & \textit{if}\; \mu = 1 \end{cases}$$

Proof. Appendix G. □

Proposition 8 states that if some mild conditions hold,[‡‡‡] then the second-best fare involves a per-distance charge of less than k dollars per mile and subsidization from general taxes in all cases but public ownership. Moreover, when the government is not able to capture all of the excess land rents, the urban sprawl induced by the optimal way to finance the transport system increases the equilibrium utility of residents.

[‡‡‡]Note that this condition is weaker than the condition of Corollary 1.

5.1 Numerical Analysis

We again consider the parameters shown in Table 1. We additionally consider a marginal cost (k in our model) of US$ 0.98/pax-mile. This number corresponds to the 2019 average of vehicle operations expenses divided by the passenger miles traveled for US public bus systems, as reported by The National Transit Database.[§§§] To obtain the value of commuting time (t in our model), we assume that the commute time is valued at the hourly wage rate[¶¶¶] and that buses move at 15 miles/h during rush hour. This allows us to get an implied value of commuting time by bus of US$1.68/mile per trip. Then, using an estimate of 1.25 workers/household, 250 working days/year, and two daily trips, we arrive to a yearly marginal cost of operation per mile for a household of US$ 612.5, and to an annual commuting cost per mile for a household of US$1052.96.

Table 4 shows the results of imposing the optimal transport fare and subsidization for different levels of land rent capture. What is most noteworthy is that, for values of μ between 0.1 and 0.5, subsidization ranges from 89% to 58%, levels that are usually observed in many cities around the world (Parry and Small, 2009; Basso and Silva, 2014). The per-mile public transport fare, in those cases, ranges from US$ 0.11 to US$ 0.41. Note also that in the extreme case of perfect land rent capture, subsidization is zero but the per-mile fare is extremely large.

The resulting city from optimal subsidization is, of course, more extense than a city with no subsidies and marginal cost pricing of public transport; this sprawl increases equilibrium utility. In line with what has been shown before, the optimal Ralwsian public transport town will have cheaper housing closer to the CBD and more aggregate mileage, differences than decrease with μ. For brevity and because they are similar to previous ones, we do not include more figures and graphs.

[§§§] https://www.transit.dot.gov/ntd/ntd-data

[¶¶¶] Note that this assumption is consistent with our model: travel time is valued at the hourly wage rate. Nevertheless, the use of any other value does not change the qualitative results of the simulations presented in this Section.

Table 4: Equilibrium results of second-best transport fare and subsidization for different levels of land rent capture.

μ	0.1	0.5	0.9
Equilibrium utility $\overline{u}$	42.94	47.54	53.63
Public transport subsidization [%]	89.18	58.16	11.43
Public transport fare per mile [US$]	0.11	0.41	0.87
City extension [miles]	47.16	47.36	47.25
Aggregate land rents [millions US$]	$1,782.68$	$2,085.23$	$2,521.76$
Aggregate mileage [thousand miles]	$1,592.80$	$1,592.72$	$1,580.59$

6 Concluding Remarks

In this paper, we developed a framework of land rent capture and redistribution using the monocentric city model to study welfare, urban sprawl and transportation pricing. We show that, even in the absence of externalities and other known urban market failures, unless all land rents are fully captured and distributed equally among residents, the urban market equilibrium is too compact from a Rawlsian welfare standpoint. Therefore, a Rawlsian planner would not choose the market outcome allocation and, the smaller the fraction of land rent captured, the larger the discrepancy between the planner and the market. We further show that the utility-maximizing city, which is more extended than the market equilibrium city, can be decentralized with a revenue-neutral combination of location-specific taxes and subsidies. To highlight the relevance of our analysis, the share of rent capture and redistribution can be approximated as either the property tax rate or the tax rate on rental income. For instance, in the US, the tax rate on rental income ranges from 10% to 37%. On the other hand, realistic values for property tax rates in the US range from approximately 0.5% to 2.5% of the property value (Song and Zenou, 2006). However, since in our model the tax is levied on rents, following Brueckner and Kim (2003) and using a 5% discount rate, realistic property tax rates translate to values of land rend capture ranging from 0.1 to 0.5.

To further show the relevance of our finding, we study how different transportation pricing policies are affected when they are designed to maximize equilibrium utility and in absence of other spatially-variant taxes. We begin by studying the optimal city in the presence of congestion externalities. We show that the optimal Rawlsian city is achieved using a tax that combines the marginal external cost and the redistribution component and that leads to an optimal city that is, in general, more extended and with a higher aggregated mileage than the unpriced city.

Although part of these results has been previously suggested (see the Appendix in Kanemoto, 1977, for a short discussion about the absentee landlord case), to the best of our knowledge, this is the first paper that formalizes these facts and provides very general results on how the market outcome in the absence of externalities does not maximize equilibrium utility for any case other than public ownership. We believe this formalization is fundamental to the literature since it provides a quite different starting point for teaching and studying urban economics from a welfare perspective, emphasizing that policy results in a monocentric city model might be widely different depending on the nature of land ownership and choice of utility function. This is of critical importance because, even when restricted to welfare analysis, it is common in the urban economics literature to see very different land ownership settings: a number of papers consider a welfare function that includes households' utility and absentee landlord' rents (e.g., Kono et al., 2012; Kono and Joshi, 2012; Li et al., 2013; Li and Guo, 2017; Li and Wang, 2018), while others consider a public ownership scenario (e.g., Oron et al., 1973; Arnott, 1979b; Kanemoto, 1980; Pines and Sadka, 1985; Verhoef, 2005; Brueckner, 2007; Pines and Kono, 2012; Tikoudis et al., 2015, 2018; Kono and Kawaguchi, 2017; Kono et al., 2019). These two approaches do not arrive to the same policy results in general, depending on the choice of utility function. In the former case, as this paper shows, the market equilibrium might not maximize social welfare in the absence of externalities, even if the absentee landlord's rents are included in the welfare function. Then, when externalities are present, pricing instruments should include a redistributing component.

Importantly, the issue of which model of land ownership is better suited goes well beyond the monocentric city model; in this sense, we think our findings are better demonstrated in this simple model but they obviously extend. For example, Tsivanidis (2019) studies the welfare effects of transit infrastructure using a quantitative urban model, while considering a public ownership scenario and other models of land ownership as robustness checks. Zhang and Kockelman (2016) extend the model of Lucas and Rossi-Hansberg (2002) to incorporate congestion externalities, analyzing welfare under different policies for the case where rents are uniformly redistributed to residents. Diamond and McQuade (2019) estimate spillovers of affordable housing developments in a non-parametric setting, studying the welfare impacts on homeowners, renters, and absentee landlords. Henderson et al. (2020) studies the dynamics of urban development in a monocentric city and differentiating formal and informal construction. The land rent is assumed to be received by non-resident owners. Thus, which land ownership model to consider when studying cities is not straightforward and the goal of this paper is to highlight that this decision is far from innocuous and has defining implications over policy results.

Naturally, several caveats to our findings need to be considered, although we do not believe the qualitative results will be different when they are considered. First, note that even though we show that our results hold if we include the absentee landlord in the welfare function, for a fully closed analysis of welfare, one might want to consider that landlords locate within the city. In this case, although the extent of previous implications might change, the main force behind our results would still be present: since landlords would end up with a higher income than renters, the model would become one with household heterogeneity. Regardless of the equilibrium location of different income groups, our result would hold within each group, as all residents of the same income group must achieve the same utility level and would, therefore, have different marginal utilities of income. Our result could also hold between groups if low-income groups locate in the suburbs. As they have lower marginal utilities of income, there would be gains by reallocating income between groups. Specifically, when the income elasticity of commuting costs is greater than the income elasticity of the demand for housing, landlords will occupy the inner city, while renters would locate in the outer city (Duranton and Puga, 2015). Since a Rawlsian planner would only care about the renters, and provided that the fraction μ is low enough, a redistribution from the city center to the outskirts should be welfare increasing, since marginal utilities of income would still increase with distance (within each income class). This last result is indirectly shown in Borck and Wrede (2005) for the absentee landlord case in a monocentric city with two income classes and mixed land ownership. Using a numerical example, Borck and Wrede (2005) depicts that when the rich own all the land (and thus, landowners locate within the city), and rental income is not taxed ($\theta = 1$ and $\theta_l = 0$ in the paper's notation), then, in the absence of externalities, a transport subsidy would be welfare increasing for the poor. Consequently, a Rawlsian planner would not choose the market outcome. This result is consistent with our formal treatment presented in Section 5.

Second, a monocentric city may be considered an unrealistic assumption. We believe this is not a strong limitation of our results for two reasons: (i) the monocentric structure persists in a large number of metropolitan areas (e.g., Arribas-Bel and Sanz-Gracia, 2014, for US evidence), and even in non-monocentric cities, the monocentric city model has been found to be empirically robust (e.g., McGrath, 2005); and (ii) our results are based in a natural property of spatial equilibrium, where marginal utilities of income vary with distance for identical residents, since commuting costs change with distance. In some more general frameworks, this property does not hold (e.g. Lucas and Rossi-Hansberg, 2002), but this is explained by the notion that besides location choice, there is an unrestricted margin of adjustment in the labor market that leads, in equilibrium, to a constant marginal utility of income within the city. This is achieved by imposing that the difference in wages paid between any two locations must be equal to the total commuting cost of traveling between the two locations. When considering a labor market that is not fully flexible, the marginal utility of income will not be constant within the city, and the effect found in this paper should be present to some extent.

For Online Publication

Appendices

A Proof of Proposition 1

For this proof, we consider a more general setting, where commuting costs may depend on the number of individuals using the same stretch of the road at the same time. To be more precise, the commuting cost per mile at x is equal to $g(T(x))$, where $T(x)$ is the total traffic at x, and g is such that $g > 0$, $g_T \geq 0$. The case without congestion is then a particular case of the proof presented in this Appendix, where $g_T = 0$. In consequence, some steps in this proof would not be necessary if we were to prove the Proposition for only the case without congestion, but this allow us to present only one proof instead of two. The welfare-maximization optimal control problem associated with this formulation is:

$$
\begin{aligned}
(P) \qquad & \max \quad \int_0^{\overline{x}} \frac{\overline{u}}{q(x)} dx & & \\
& \text{subject to} \quad u(c(x), q(x)) = \overline{u}, & 0 \leq x \leq \overline{x} \qquad & \text{(17a)} \\
& t(x) = \int_0^x g(T(z)) dz, & 0 \leq x \leq \overline{x} \qquad & \text{(17b)} \\
& T(x) = \int_x^{\overline{x}} \frac{1}{q(z)} dz, & 0 \leq x \leq \overline{x} \qquad & \text{(17c)} \\
& \int_0^{\overline{x}} \frac{1}{q(x)} dx = L & & \text{(17d)} \\
& \frac{\mu R}{L} + y - t(x) - \tau(x) - c(x) - p(x)q(x) = 0 & 0 \leq x \leq \overline{x} \qquad & \text{(17e)} \\
& R = \int_0^{\overline{x}} p(x) - r_a \, dx & & \text{(17f)} \\
& 0 = \int_0^{\overline{x}} \frac{\tau(x)}{q(x)} dx & & \text{(17g)}
\end{aligned}
$$

where $c(x)$, $q(x)$ and $\tau(x)$ are control variables, while $\overline{u}$ and $\overline{x}$ are control parameters. Conditions (17c) and (17d) can we written as

$$T'(x) = -\frac{1}{q(x)} \qquad 0 \leq x \leq \overline{x} \tag{18}$$

$$T(0) = L \tag{19}$$

In the same way as done in F, the budget constraint can be replaced by

$$\int_0^{\overline{x}} \frac{c(x)}{q(x)} + \frac{t(x)}{q(x)} + \mu r_a + (1-\mu)\frac{u_q}{u_c}\, dx = Ly \tag{20}$$

Now, note that the second term of (20) can be integrated by parts:

$$\int_0^{\overline{x}} \frac{t(x)}{q(x)}\, dx = \int_0^{\overline{x}} -t(x)T'(x)\, dx \tag{21}$$

$$= - t(x)T'(x)\Big|_0^{\overline{x}} + \int_0^{\overline{x}} t'(x)T(x)\, dx \tag{22}$$

$$= \int_0^{\overline{x}} T(x)g(T(x))\, dx \tag{23}$$

where we used (18) for the first equality, and the boundary conditions $t(0) = 0$, $T(\overline{x}) = 0$ and the condition $t'(x) = g(T(x))$ for the last equality. Using this last in (20), we arrive to the budget constraint we will use:

$$\int_0^{\overline{x}} \frac{c(x)}{q(x)} + T(x)g(T(x)) + \mu r_a + (1-\mu)\frac{u_q}{u_c}\, dx = Ly \tag{24}$$

The Lagrangian for this problem can then be written as

$$\Lambda(x) = \int_0^{\overline{x}} \frac{\overline{u}}{q(x)} dx + \int_0^{\overline{x}} \nu(x)(u(c(x),q(x)) - \overline{u})dx - \frac{\lambda(x)}{q(x)} \tag{25}$$

$$+ \delta\left[Ly - \int_0^{\overline{x}} \frac{c(x)}{q(x)} + T(x)g(T(x)) + \mu r_a + (1-\mu)\frac{u_q}{u_c}\, dx\right] + \gamma\left[L - \int_0^{\overline{x}} \frac{1}{q(x)}\, dx\right] \tag{26}$$

where $\nu(x)$, $\lambda(x)$ and δ are the Lagrange multipliers associated with (17a), (18) and (24) respectively. The first order conditions are then:

$$-\frac{\delta}{q(x)} - \delta\left(\frac{1-\mu}{u_c^2}\right)(u_{cq}u_c - u_q u_{cc}) + \nu(x)u_c = 0 \qquad 0 \le x \le \bar{x} \tag{27}$$

$$-\frac{\overline{u} - \lambda(x) - \delta c(x) - \gamma}{q(x)^2} - \delta\left(\frac{1-\mu}{u_c^2}\right)(u_{qq}u_c - u_q u_{cq}) + \nu(x)u_q = 0 \qquad 0 \le x \le \bar{x} \tag{28}$$

$$-\lambda'(x) = -\delta\left(g(T(x)) + T(x)\frac{dg}{dT}(T(x))\right) \qquad 0 \le x \le \bar{x} \tag{29}$$

From (29), calling $MEC(x)$ the marginal external cost at x, we can get

$$\lambda(x) = \lambda(0) + \delta\,(t(x) + MEC(x)) \tag{30}$$

On the other hand, straightforward calculations using (27) and (28) lead to

$$\frac{\nu(x)u_q}{\nu(x)u_c} = \frac{u - \lambda - \delta c - \gamma + \Lambda q^2}{q(\delta + q\Omega)} \tag{31}$$

where

$$\Lambda = \delta\left(\frac{1-\mu}{u_c^2}\right)(u_{qq}u_c - u_q u_{cq}) \tag{32}$$

$$\Omega = \delta\left(\frac{1-\mu}{u_c^2}\right)(u_{cq}u_c - u_q u_{cc}) \tag{33}$$

Then, since $\dfrac{u_q}{u_c} = p(x)$, using (30) and (31) we can obtain that, for the optimum:

$$t(x) + c(x) + p(x)q(x) = \frac{\overline{u} - \lambda(0) - \gamma}{\delta} - MEC(x) + q^2\left(\frac{\Lambda - p\Omega}{\delta}\right) \tag{34}$$

Using the budget constraint $\frac{\mu R}{L} + y = \tau(x) + t(x) + c(x) + p(x)q(x)$, we then get the optimal toll

$$\tau(x) = \left(\frac{\mu R}{L} + y - \frac{\overline{u} - \lambda(0) - \gamma}{\delta}\right) + MEC(x) + q^2\left(\frac{p\Omega - \Lambda}{\delta}\right) \tag{35}$$

Noting that the first term is a constant, and since we restrict the toll to be revenue-neutral, the following must hold:

$$C := \frac{\mu R}{L} + y - \frac{\overline{u} - \lambda(0) - \gamma}{\delta} = -\int_0^{\overline{x}} \frac{MEC(x) + q(x)^2 \left(\frac{p\Omega - \Lambda}{\delta}\right)}{q(x)} dx \tag{36}$$

Then, we can write $\tau(x)$ simply as

$$\tau(x) = MEC(x) + q^2 \left(\frac{p\Omega - \Lambda}{\delta}\right) + C \tag{37}$$

Note that if $g_T = 0$, the marginal external cost is zero. Thus, in that case, the optimal toll is given by

$$\tau(x) = q^2 \left(\frac{p\Omega - \Lambda}{\delta}\right) + C \tag{38}$$

Finally, when $\mu = 1$, $\Lambda = \Omega = 0$, so that the optimum toll in (38) is zero everywhere. For $\mu \neq 1$, it is easy to check that

$$\frac{p\Omega - \Lambda}{\delta} = \left(\frac{1-\mu}{u_c^3}\right)(2u_c u_q u_{cq} - u_q^2 u_{cc} - u_c^2 u_{qq}) \tag{39}$$

Furthermore, $2u_c u_q u_{cq} - u_q^2 u_{cc} - u_c^2 u_{qq} > 0$ if u is strictly quasi-concave, so this last expression is non-negative for all values of μ.

Now, since $p(x) = \frac{u_q}{u_c}$ $\forall\, 0 \leq x \leq \overline{x}$, differentiating this last expression with respect to x allow us to get

$$\frac{\partial p(x)}{\partial x} = \frac{u_{cq}\left(u_c \frac{\partial c(x)}{\partial x} - u_q \frac{\partial q(x)}{\partial x}\right) - u_q u_{cc} \frac{\partial c(x)}{\partial x} + u_c u_{qq} \frac{\partial q(x)}{\partial x}}{u_c^2} \qquad \forall\, 0 \leq x \leq \overline{x} \tag{40}$$

From the equilibrium condition $u(c(x), q(x)) = \overline{u}$ $\forall\, 0 \leq x \leq \overline{x}$, it is easy to obtain $u_c \frac{\partial c(x)}{\partial x} = -u_q \frac{\partial q(x)}{\partial x}$.

Using this last equality in (40) reduces to

$$\frac{\partial p(x)}{\partial x} = \frac{\frac{\partial c(x)}{\partial x}\left(2u_c u_{cq} - u_q u_{cc} - \frac{u_c^2}{u_q}u_{qq}\right)}{u_c^2} \qquad \forall\, 0 \leq x \leq \bar{x} \tag{41}$$

$$= \frac{\partial c(x)}{\partial x}\frac{u_c}{u_q}\left(\frac{2u_c u_q u_{cq} - u_q^2 u_{cc} - u_c^2 u_{qq}}{u_c^3}\right) \qquad \forall\, 0 \leq x \leq \bar{x} \tag{42}$$

Since $\frac{\partial c(x)}{\partial x} < 0$, $\frac{u_c}{u_q} > 0$, $\frac{2u_c u_q u_{cq} - u_q^2 u_{cc} - u_c^2 u_{qq}}{u_c^3} > 0$, we conclude that $\frac{\partial p(x)}{\partial x} < 0$. Using this in the Muth-Mills condition

$$\frac{\partial p(x)}{\partial x} = -\frac{t + \tau'(x)}{q(x)} \qquad \forall\, 0 \leq x \leq \bar{x} \tag{43}$$

we arrive to $\tau'(x) > -t\ \forall\, 0 \leq x \leq \bar{x}$. Returning to (42):

$$\frac{\partial p(x)}{\partial x} = \frac{\partial c(x)}{\partial x}\frac{u_c}{u_q}\left(\frac{2u_c u_q u_{cq} - u_q^2 u_{cc} - u_c^2 u_{qq}}{u_c^3}\right) \qquad \forall\, 0 \leq x \leq \bar{x} \tag{44}$$

this implies that

$$q^2\left(\frac{1-\mu}{u_c^3}\right)(2u_c u_q u_{cq} - u_q^2 u_{cc} - u_c^2 u_{qq}) = q^2(1-\mu)\frac{\partial p(x)}{\partial x} / \left(\frac{\partial c(x)}{\partial x}\frac{u_c}{u_q}\right) \tag{45}$$

$$= -q^2(1-\mu)\frac{\partial p(x)}{\partial x} / \frac{\partial q(x)}{\partial x} \tag{46}$$

$$= (1-\mu)\left(-\frac{\partial p(x)}{\partial x} / \frac{\partial q(x)}{\partial x} \cdot \frac{q(x)}{p(x)}\right) p(x)q(x) \tag{47}$$

$$= (1-\mu)\frac{p(x)q(x)}{|\sigma|} \tag{48}$$

where σ is the income-compensated elasticity of demand for housing.

B Proof of Corollary 1

Indeed, if the conditions stated in Corollary 1 hold, then $\frac{1}{|\sigma|}$ is a non-decreasing function of x. In addition, $\frac{\partial p(x)q(x)}{\partial x} = (1+\sigma)q\frac{\partial p}{\partial x}$, so that $p(x)q(x)$ –the expenditure in housing– is a decreasing function of x if $\sigma > -1$. Then, noting that $\sigma = \alpha - 1$ for a Cobb-Douglas utility function $u(c,q) = c^{1-\alpha}q^{\alpha}$, it is easy to check that both conditions hold. On the other hand, for a CES utility function, $u(c,q) = ((1-\alpha)c^{\rho} + \alpha q^{\rho})^{\frac{1}{\rho}}$, tedious yet straightforward calculations lead to

$$\sigma(x) = \sigma\left(\frac{1}{1+\left(\frac{1-\alpha}{\alpha}\right)^{\sigma} p(x)^{\sigma-1}} - 1\right)$$

where $\sigma = 1/(1-\rho)$ is the elasticity of substitution. Since the first term inside the parenthesis is positive, a sufficient condition for $\sigma > -1$ is $\sigma < 1 \Leftrightarrow \rho < 0$. In addition, since $p(x)$ is a decreasing function of x, a sufficient condition for $\sigma'(x) < 0$ is also $\sigma < 1 \Leftrightarrow \rho < 0$.

C Proof of Lemma 1

In this proof, superscripts m and τ refer to the market equilibrium, and the equilibrium when using τ, respectively. We first show that $p^\tau(0) < p^m(0)$. Indeed, the Muth-Mills conditions are

$$\frac{\partial p^m}{\partial x} = -\frac{t}{q^m} \tag{49}$$

$$\frac{\partial p^\tau}{\partial x} = -\frac{t+\tau'}{q^\tau} \tag{50}$$

Integrating (49) from 0 to $\bar{x}^m$ leads to

$$p^m(0) = r_a + Lt \tag{51}$$

while integrating (50) from 0 to $\bar{x}^\tau$ leads to

$$p^\tau(0) = r_a + Lt + \int_0^{\bar{x}^\tau} \frac{\tau'(x)}{q^\tau(x)} dx \tag{52}$$

Using (51), (52) and the fact that $\tau' < 0$, we obtain $p^\tau(0) < p^m(0)$.

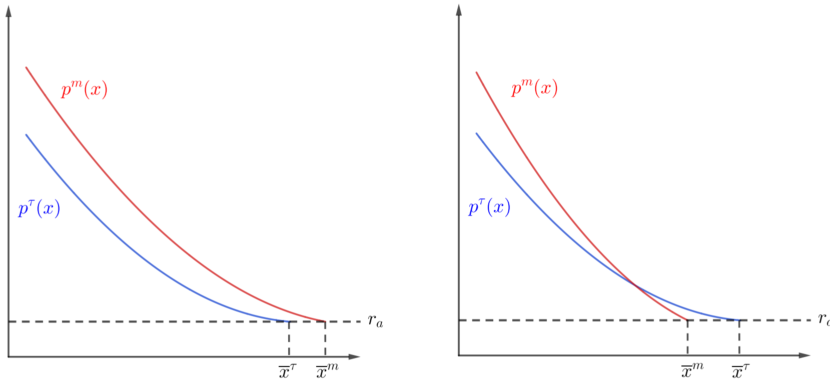
The second step of the proof is to show that the price curves have a single crossing point, which implies that the city has to be more extended (Figure 5b). There are three possible cases regarding the price curves:

Case 1. $p^\tau(x) \leq p^m(x)\ \forall x \in [0, \bar{x}^\tau]$. In this case, it follows that $\bar{x}^\tau \leq \bar{x}^m$ (Figure 5a).

Case 2. $p^\tau(x) \leq p^m(x)\ \forall x \in [0, x_c]$, $p^\tau(x) > p^m(x)\ \forall x \in (x_c, \bar{x}^m]$, for some $x_c \in (0, \bar{x}^m)$. In this case, it follows that $\bar{x}^\tau > \bar{x}^m$ (Figure 5b).

Case 3. $p^\tau(x)$ and $p^m(x)$ intersect two or more times.

We show that the cases 1 and 3 cannot occur.



(a) Case 1 (b) Case 2

Figure 5: Relationship between bid rent curves.

For the Case 1, we proceed by contradiction: suppose $p^\tau(x) \leq p^m(x)$ $\forall x \in [0, \bar{x}^\tau]$. Then, consider any $x \in [0, \bar{x}^\tau]$. We will show that $q^\tau(x) \geq q^m(x)$. Note that the income net of transportation costs and taxes/subsidies at x in the market city is $I^m(x) = y + \frac{\mu R^m}{L} - tx$, while in the city when using τ is $I^\tau(x) = y + \frac{\mu R^\tau}{L} - tx - \tau(x)$. In general, $I^\tau(x) \geq I^m(x)$ or $I^\tau(x) < I^m(x)$ may hold, since depending on the value of x, $\tau(x)$ might be a tax or a subsidy. Additionally, the excess land rents are different in both cities. Thus, we show that $q^\tau(x) \geq q^m(x)$ for both possibilities.

Case 1.a. Suppose $I^\tau(x) \geq I^m(x)$. In this case, since $p^\tau(x) \leq p^m(x)$ and housing is a normal good, it follows directly that the housing consumption at x does not decrease compared to the market city, i.e., $q^\tau(x) \geq q^m(x)$. This last inequality is strict if either $I^\tau(x) > I^m(x)$ or $p^\tau(x) < p^m(x)$.

Case 1.b. Suppose $I^\tau(x) < I^m(x)$. In this case, since $p^\tau(x) \leq p^m(x)$ and the disposable income decreases, the consumption of the composite good c decreases (due to the income and substitution effect). Then, since we assume that the equilibrium utility in the city when using τ is higher, the consumption of housing must increase, i.e. $q^\tau(x) > q^m(x)$.

To conclude, note that we proved that $q^\tau(x) \geq q^m(x)$ $\forall x \in [0, \bar{x}^\tau]$, with strict inequality if $p^\tau(x) < p^m(x)$.

Then,

$$\begin{aligned} L &= \int_0^{\bar{x}^\tau} \frac{1}{q^\tau(x)} dx \\ &< \int_0^{\bar{x}^\tau} \frac{1}{q^m(x)} dx \\ &\le L \end{aligned}$$

The first inequality is strict since $p^\tau(0) < p^m(0)$, and by continuity, $q^\tau(x) > q^m(x)$ in an interval to the right of $x = 0$. The last inequality follows from the fact that, in this case, the market city is more extended (Figure 5a). Thus, this case cannot arise in equilibrium.

For the Case 3, we once again proceed by contradiction. Suppose that p^τ and p^m intersect at least two times. It is easy to see that, since $p^\tau(0) < p^m(0)$, at the first intersection, x_1, the following inequality holds:

$$\left.\frac{\partial p^m}{\partial x}\right|_{x=x_1} < \left.\frac{\partial p^\tau}{\partial x}\right|_{x=x_1} \tag{53}$$

In words, at the first intersection, p^m is decreasing at a faster rate than p^τ. Naturally, at the second intersection the reverse is true:

$$\left.\frac{\partial p^m}{\partial x}\right|_{x=x_2} > \left.\frac{\partial p^\tau}{\partial x}\right|_{x=x_2} \tag{54}$$

Using the Muth-Mills condition, condition (54) implies

$$\frac{t}{q^m(x_2)} < \frac{t + \tau'(x_2)}{q^\tau(x_2)} \tag{55}$$

$$\Rightarrow q^\tau(x_2) < q^m(x_2) \tag{56}$$

where we used $\tau'(x_2) < 0$ for the last inequality. Then, as p^m and p^τ intersect at x_2, $p^m(x_2) = p^\tau(x_2)$ must hold. From this, again two possibilities may arise since $I^\tau(x_2) \ge I^m(x_2)$ or $I^\tau(x_2) < I^m(x_2)$ may hold:

Case 2.a $I^{\tau}(x_2) \geq I^{m}(x_2)$. Since residents living at x_2 in the city that uses the instrument τ have at least the same disposable income than the residents living at x_2 in the market city, and they face the same prices for c and q, they consume at least the same amount of housing. Thus, (56) cannot hold.

Case 2.b $I^{\tau}(x_2) < I^{m}(x_2)$. In this case, residents living at x_2 in the city that uses the instrument τ have less disposable income than the residents living at x_2 in the market city, and they face the same prices for c and q. Thus, it is impossible for them to attain a higher utility, contradicting one of our hypotheses.

D Proof of Proposition 3

In this proof, we show that if one considers a social welfare function that is the sum of the compensating variation of residents plus the share of the absentee landlord rents, then the market equilibrium does not maximize social welfare under some mild conditions for any $\mu < 1$. Those conditions are (i) an increasing marginal utility of income with distance (ii) decreasing spending in housing with distance. We show that these conditions hold if $u(c,q)$ is strictly quasi-concave. We proceed by showing that some income redistribution from the city center to the outskirts increases welfare.

We start by considering two equilibria. First, the resulting equilibrium when using a linear revenue-neutral combination of taxes and subsidies $\tau(x) = ax + b$. Second, we consider the resulting equilibrium when including a linear revenue-neutral combination of taxes and subsidies $\tau(x) = (a + da)x + b$, with $da \to 0$. That is, the second equilibrium corresponds to a marginal deviation from the first one. In this case, the compensating variation associated with this marginal deviation is defined as

$$CV = -\left.\frac{dy}{da}\right|_{u=\text{constant}} = \frac{du}{da}\bigg/\frac{du}{dy} \tag{57}$$

With this, the change in social welfare, denoted by SW, when a changes to $a + da$ is

$$\frac{dSW}{da} = L \cdot CV + (1-\mu)\frac{dR}{da} \tag{58}$$

We are interested in the sign of $\frac{dSW}{da}$ evaluated at $a = 0$. If this derivative is negative, then an income redistribution from the city center to the outskirts is welfare increasing. To find this sign, we first calculate CV, studying the initial equilibrium. The urban equilibrium condition can be stated as

$$v\left(\frac{\mu R}{L} + y - (t+a)x - b - pq, q\right) = u \qquad \forall x \in [0,\bar{x}] \tag{59}$$

with v the indirect utility function. Totally differentiating (59) with respect to a leads to

$$v_c \cdot \left(\frac{\mu}{L} \frac{dR}{da} - x - \frac{db}{da} - q\frac{dp}{da} - p\frac{dq}{da} \right) + v_q \cdot \frac{dq}{da} = \frac{du}{da} \qquad \forall x \in [0, \bar{x}] \tag{60}$$

while totally differentiating (59) with respect to y leads to

$$v_c \cdot \left(\frac{\mu}{L} \frac{dR}{dy} + 1 - \frac{db}{dy} - q\frac{dp}{dy} - p\frac{dq}{dy} \right) + v_q \cdot \frac{dq}{dy} = \frac{du}{dy} \qquad \forall x \in [0, \bar{x}] \tag{61}$$

On the other hand, the Muth-Mills condition is:

$$\frac{\partial p}{\partial x} = -\frac{t+a}{q} \quad \forall x \in [0, \bar{x}] \tag{62}$$

Integrating (62) with respect to x we get:

$$p(x=0) - r_a = L(t+a) \tag{63}$$

and thus

$$\frac{dp(x=0)}{da} = L \tag{64}$$

$$\frac{dp(x=0)}{dy} = 0 \tag{65}$$

Then, evaluating (60) and (61) in $x = 0$, using (64), (65) and the utility maximization condition $\frac{v_q}{v_c} = p$ leads to

$$\frac{du}{da} = v_c(x=0) \cdot \left(\frac{\mu}{L} \frac{dR}{da} - \frac{db}{da} - Lq(x=0) \right) \tag{66}$$

$$\frac{du}{dy} = v_c(x=0) \cdot \left(\frac{\mu}{L} \frac{dR}{dy} + 1 - \frac{db}{dy} \right) \tag{67}$$

where we denote, with some abuse of notation, $v_c\left(\frac{\mu R}{L} + y - b - p(x=0)q(x=0), q(x=0)\right)$ by $v_c(x=$

0). It follows that

$$CV = \frac{\frac{\mu}{L}\frac{dR}{da} - \frac{db}{da} - Lq(x=0)}{\frac{\mu}{L}\frac{dR}{dy} + 1 - \frac{db}{dy}} \tag{68}$$

To find expressions for $\frac{db}{da}$ and $\frac{db}{dy}$, we remember that $\tau(x)$ is revenue-neutral:

$$\int_0^{\bar{x}} \frac{ax+b}{q} dx = 0 \Rightarrow b = -\frac{a}{L}\int_0^{\bar{x}} \frac{x}{q} dx \tag{69}$$

Replacing q through the Muth-Mills condition (62)

$$b = \frac{a}{L(t+a)} \int_0^{\bar{x}} \left(\frac{\partial p}{\partial x}\right) x \, dx \tag{70}$$

$$= \frac{a}{L(t+a)} \int_0^{\bar{x}} \frac{\partial (px)}{\partial x} - p(x) \, dx \tag{71}$$

$$= \frac{a}{L(t+a)} \left(p(\bar{x})\bar{x} - \int_0^{\bar{x}} p(x) \, dx\right) \tag{72}$$

$$= \frac{a}{L(t+a)} \left(\int_0^{\bar{x}} r_a - p(x) dx\right) \tag{73}$$

$$= -\frac{a}{L(t+a)} R \tag{74}$$

Using (69) and (74), we can obtain:

$$\int_0^{\bar{x}} \frac{x}{q} dx = \frac{R}{t+a} \tag{75}$$

while totally differentiating (74) with respect to a and y, we get

$$\frac{db}{da} = -\frac{t}{L(t+a)^2} R - \frac{a}{L(t+a)} \frac{dR}{da} \tag{76}$$

$$\frac{db}{dy} = -\frac{a}{L(t+a)} \frac{dR}{dy} \tag{77}$$

Since we are interested in the sign of $\frac{dSW}{da}$ evaluated at $a=0$, we evaluate (76) and (77) at $a=0$:

$$\left.\frac{db}{da}\right|_{a=0} = -\frac{R}{Lt} \tag{78}$$

$$\left.\frac{db}{dy}\right|_{a=0} = 0 \tag{79}$$

To obtain expressions for $\dfrac{dR}{da}$ and $\dfrac{dR}{dy}$, we return to the definition of R:

$$R = \int_0^{\bar{x}} p(x) - r_a \, dx \tag{80}$$

Totally differentiating (80) with respect to y gives us

$$\begin{aligned}
\frac{dR}{dy} &= (p(\bar{x}) - r_a)\frac{d\bar{x}}{dy} + \int_0^{\bar{x}} \frac{dp}{dy} dx &(81)\\
&= \int_0^{\bar{x}} \frac{1}{q}\frac{\mu}{L}\frac{dR}{dy} + \frac{1}{q} - \frac{1}{q}\frac{db}{dy} - \frac{1}{q \cdot v_c}\frac{du}{dy} \, dx &(82)\\
&= L\left(\frac{\mu}{L}\frac{dR}{da} + 1 - \frac{db}{dy}\right) - \frac{du}{dy}\int_0^{\bar{x}} \frac{1}{q \cdot v_c} \, dx &(83)\\
&= L\left(\frac{\mu}{L}\frac{dR}{dy} + 1 - \frac{db}{dy}\right) - \left(\frac{\mu}{L}\frac{dR}{dy} + 1 - \frac{db}{dy}\right)\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c} \, dx &(84)
\end{aligned}$$

For the second equality we used $p(\bar{x}) = r_a$, and the characterization of $\frac{dp}{dy}$ given by the identity of (61). For the fourth inequality, we used (67). From (84) we can isolate $\frac{dR}{dy}$:

$$\frac{dR}{dy} = \left(\frac{L}{L(1-\mu) + \mu \displaystyle\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c} \, dx}\right)\left(L - \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c} \, dx\right)\left(1 - \frac{db}{dy}\right) \tag{85}$$

Similarly, totally differentiating (80) with respect to a allow us to get

$$\begin{aligned}
\frac{dR}{da} &= (p(\bar{x}) - r_a)\frac{d\bar{x}}{da} + \int_0^{\bar{x}} \frac{dp}{da} dx && (86)\\
&= \int_0^{\bar{x}} \frac{1}{q}\frac{\mu}{L}\frac{dR}{da} - \frac{x}{q} - \frac{1}{q}\frac{db}{da} - \frac{1}{q \cdot v_c}\frac{du}{da}\, dx && (87)\\
&= L\left(\frac{\mu}{L}\frac{dR}{da} - \frac{db}{da}\right) - \frac{R}{t+a} - \frac{du}{da}\int_0^{\bar{x}} \frac{1}{q \cdot v_c}\, dx && (88)\\
&= L\left(\frac{\mu}{L}\frac{dR}{da} - \frac{db}{da}\right) - \frac{R}{t+a} - CV\left(\frac{\mu}{L}\frac{dR}{dy} + 1\right)\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx && (89)
\end{aligned}$$

For the second equality we used $p(\bar{x}) = r_a$, and the characterization of $\frac{dp}{da}$ given by the identity of (60). For the third equality, we used (75). For the fourth inequality, we used (68) together with (66). It follows from (89) that

$$(1-\mu)\frac{dR}{da} = -L\frac{db}{da} - \frac{R}{t+a} - CV\left(\frac{\mu}{L}\frac{dR}{dy} + 1\right)\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx \tag{90}$$

Once again, since we are interested in the sign of $\frac{dSW}{da}$ evaluated at $a = 0$, we evaluate (85) and (90) at $a = 0$:

$$\begin{aligned}
\left.\frac{dR}{dy}\right|_{a=0} &= \left(\frac{L}{L(1-\mu) + \mu\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx}\right)\left(L - \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx\right) && (91)\\
(1-\mu)\left.\frac{dR}{da}\right|_{a=0} &= -CV\left(\frac{\mu}{L}\frac{dR}{dy} + 1\right)\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx && (92)
\end{aligned}$$

where we used (78) and (79) to replace $\frac{db}{da}$ and $\frac{db}{dy}$.

Note that from (91), it follows that $\left.\frac{dR}{dy}\right|_{a=0} > 0$ if the marginal utility of income is increasing with distance. Indeed, the first term on the right-hand size is clearly positive, while the second term is positive if $L > \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx$. A sufficient condition for this is precisely $\frac{\partial v_c}{\partial x} > 0$.

Now, we study the sign of $CV|_{a=0}$, starting from (68):

$$CV|_{a=0} = \frac{\frac{\mu}{L}\left.\frac{dR}{da}\right|_{a=0} + \frac{R}{Lt} - Lq(x=0)}{\frac{\mu}{L}\left.\frac{dR}{dy}\right|_{a=0} + 1} \tag{93}$$

Considering the previous comment, if the marginal utility of income is increasing with distance, then the denominator of (93) is clearly greater than zero. On the other hand, we show that if the spending in housing is decreasing with distance, then $\frac{R}{Lt} - Lq(x=0) < 0$. Indeed:

$$\frac{R}{Lt} - Lq(x=0) = \frac{R - Lt \cdot Lq(x=0)}{Lt} \tag{94}$$

$$= \frac{\int_0^{\bar{x}} p(x) - r_a \, dx - Lq(x=0)(p(x=0) - r_a)}{Lt} \tag{95}$$

$$= \frac{\int_0^{\bar{x}} \frac{p(x)q(x) - p(x=0)q(x=0) - r_a(q(x) - q(x=0))}{q(x)}}{Lt} \tag{96}$$

For the second equality we used (63) evaluated at $a = 0$. For the third equality we used $L = \int_0^{\bar{x}} \frac{1}{q} dx$.

Since $q(x) - q(x=0) \geq 0$, from (96) it follows that a sufficient condition for $\frac{R}{Lt} - Lq(x=0) < 0$ is indeed $\frac{\partial p(x)q(x)}{\partial x} < 0$. Finally, note that if this holds, then $\frac{dR}{da}\big|_{a=0}$ cannot be less or equal than zero. We proceed by contradiction. If $\frac{dR}{da}\big|_{a=0} \leq 0$, then $CV|_{a=0} < 0$ under the previous assumptions, since the denominator in (93) is positive, and $\frac{R}{Lt} - Lq(x=0) < 0$. But, from (92), $\frac{dR}{da}\big|_{a=0}$ and $CV|_{a=0}$ have opposite signs. Thus, $\frac{dR}{da}\big|_{a=0} > 0$, and then $CV|_{a=0} < 0$. The only exception for this is when $\mu = 1$. In this case, from (92), $CV|_{a=0} = 0$.

Finally, we study the sign of $\frac{dSW}{da}\big|_{a=0}$:

$$\left.\frac{dSW}{da}\right|_{a=0} = L \cdot CV|_{a=0} + (1-\mu)\left.\frac{dR}{da}\right|_{a=0} \tag{97}$$

$$= CV|_{a=0}\left(L - \left(\frac{\mu}{L}\left.\frac{dR}{dy}\right|_{a=0} + 1\right)\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\right) \tag{98}$$

where we used (92) for the last equality. Clearly, $\frac{dSW}{da}\Big|_{a=0} = 0$ for $\mu = 1$, since $CV|_{a=0} = 0$. This is consistent with the well-known fact that the market equilibrium maximizes welfare under a public-ownership setting. We focus now on the case $\mu \neq 1$, where we have $CV|_{a=0} < 0$. Note that if the marginal utility of income is increasing with distance, then $L > \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx$ and the following holds from (91):

$$\frac{dR}{dy}\Big|_{a=0} = \left(\frac{L}{L(1-\mu) + \mu \displaystyle\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx} \right) \left(L - \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx \right) \tag{99}$$

$$< \frac{L}{\displaystyle\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx} \left(L - \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx \right) \tag{100}$$

Using (100) in (98)

$$\frac{dSW}{da}\Big|_{a=0} \leq CV|_{a=0} \left[L - \left\{ \frac{\mu}{\displaystyle\int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx} \left(L - \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx \right) + 1 \right\} \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c} \right] \tag{101}$$

$$= CV|_{a=0} (1-\mu) \left(L - \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c} dx \right) \tag{102}$$

$$< 0 \tag{103}$$

(101) holds with equality for $\mu = 0$. For the last inequality, we used $CV|_{a=0} < 0$ and $L > \int_0^{\bar{x}} \frac{v_c(x=0)}{q \cdot v_c}\, dx$ under our assumptions. Thus, a marginal resdistribution of income from the city center to the outskirts is welfare improving for any $\mu \in [0, 1)$, provided two assumptions used in this proof: (i) $\frac{\partial v_c}{\partial x} > 0$ (ii) $\frac{\partial (pq)}{\partial x} < 0$. We now show that these conditions hold if $u(c, q)$ is strictly quasi-concave and $\sigma > -1$:

(i)

$$\frac{\partial v_c}{\partial x} = v_{cc} \frac{\partial c}{\partial x} + v_{cq} \frac{\partial q}{x} \tag{104}$$

Then, since the utility is the equilibrium utility is constant throughout the city, it is easy to obtain

$v_c \dfrac{\partial c}{\partial x} = -v_q \dfrac{\partial q}{\partial x}$. Using this in (104), we obtain:

$$\frac{\partial v_c}{\partial x} = \frac{1}{v_c} \frac{\partial q}{\partial x} \left(v_c v_{cq} - v_q v_{cc} \right) \tag{105}$$

We conclude using two well-known results: if $u(c,q)$ is strictly quasi-concave, then $\frac{\partial q}{\partial x} > 0$ (e.g., Brueckner, 1987) and $v_c v_{cq} - v_q v_{cc} > 0$ (e.g., Section 2.1 of Kanemoto, 1980).

(ii) As Brueckner (1987) shows

$$\frac{\partial (pq)}{\partial x} = (1+\sigma) q \frac{\partial p}{\partial x} \tag{106}$$

with σ the income-compensated price elasticity of demand for housing. Using the fact that $\frac{\partial p}{\partial x} < 0$, it follows directly that $\frac{\partial (pq)}{\partial x}$ if $\sigma > -1$.

E Proof of Proposition 4

Let $C(q(x),\overline{u})$ be the quantity of the composite good for which $u(C(q(x),\overline{u}),q(x)) = \overline{u}$. The problem stated in Proposition 4 can be restated as

$$(P_{\min}) \qquad \max \quad \left(\int_0^{\overline{x}} \frac{\overline{y} - tx - C(q(x),\overline{u})}{q(x)} - r_a \, dx \right) \tag{107a}$$

$$\int_0^{\overline{x}} \frac{1}{q(x)} dx = L \tag{107b}$$

Where $\overline{y} = y + \frac{\mu R}{L}$. From this, it follows directly from the proof of Section 3.3.3 of Fujita and Thisse (2013) that the market equilibrium solves this problem.

F Proof of Proposition 5

The utility maximization problem under our setting can be stated as

$$
\begin{aligned}
(P_{\max}) \qquad & \max && \int_0^{\overline{x}} \frac{\overline{u}}{q(x)} dx && && (108a)\\
& \text{subject to} && u(c(x), q(x)) = \overline{u}, && 0 \le x \le \overline{x} && (108b)\\
& && \int_0^{\overline{x}} \frac{1}{q(x)} dx = L && && (108c)\\
& && \frac{\mu R}{L} + y - t(x) - c(x) - p(x)q(x) = 0 && 0 \le x \le \overline{x} && (108d)\\
& && R = \int_0^{\overline{x}} p(x) - r_a \, dx && && (108e)
\end{aligned}
$$

with $\mu \in [0,1]$. Define M as the total budget of a household:

$$
\begin{aligned}
M &= y + \frac{\mu R}{L} && (109)\\
&= t(x) + c(x) + p(x)q(x) && (110)
\end{aligned}
$$

Using (110) we obtain

$$
p(x) = \frac{M - t(x) - c(x)}{q(x)} \qquad (111)
$$

Then, replacing (108e) in (109), we get

$$
\begin{aligned}
M &= y + \frac{\mu}{L} \int_0^{\overline{x}} p(x) - r_a \, dx\\
&= y + \frac{1}{L} \int_0^{\overline{x}} p(x) - \mu r_a - (1-\mu) p(x) \, dx
\end{aligned}
$$

Then, using (111)

$$
\begin{aligned}
M &= y + \frac{1}{L}\int_0^{\overline{x}} \frac{M - t(x) - c(x)}{q(x)} - \mu r_a - (1-\mu)p(x)\; dx \\
&= y + M + \frac{1}{L}\int_0^{\overline{x}} -\frac{t(x)+c(x)}{q(x)} - r_a - (1-\mu)\frac{u_q}{u_c}\; dx
\end{aligned}
$$

Where we used the individual utility maximization condition $p(x) = \dfrac{u_q(c(x),q(x))}{u_c(c(x),q(x)}$. The budget constraint (108d) can then be replaced by

$$
\int_0^{\overline{x}} \frac{t(x)+c(x)}{q(x)} + \mu r_a + (1-\mu)\frac{u_q}{u_c}\; dx = Ly \tag{112}
$$

The Lagrangian for this problem can then be written as

$$
\Lambda_1(x) = \lambda_1 \int_0^{\overline{x}} \frac{\overline{u}}{q(x)} dx + \int_0^{\overline{x}} \nu(x)(u(c(x),q(x)) - \overline{u})dx + \delta\left[Ly - \int_0^{\overline{x}} \frac{t(x)+c(x)}{q(x)} + \mu r_a + (1-\mu)\frac{u_q}{u_c}\; dx\right] + \gamma\left[L - \int_0^{\overline{x}} \frac{1}{q(x)}\; dx\right] \tag{113}
$$

where $\nu(x), \delta$ and γ are respectively the Lagrange multipliers associated with (108b), (112) and (108c). Applying the maximum principle, the first order conditions are (for the maximum principle used in the following, see Kanemoto, 1980):

$$
\nu(x)u_c - \frac{\delta}{q(x)} - \delta\left(\frac{1-\mu}{u_c^2}\right)(u_{cq}u_c - u_q u_{cc}) = 0 \qquad 0 \le x \le \overline{x} \tag{114}
$$

$$
\nu(x)u_q - \frac{\lambda_1\overline{u} - \delta t(x) - \delta c(x) - \gamma}{q(x)^2} - \delta\left(\frac{1-\mu}{u_c^2}\right)(u_{qq}u_c - u_q u_{cq}) = 0 \qquad 0 \le x \le \overline{x} \tag{115}
$$

$$
\int_0^{\overline{x}} \frac{1}{q(x)} - \int_0^{\overline{x}} \nu(x) = 0 \tag{116}
$$

$$
\frac{\lambda_1\overline{u}}{q(\overline{x})} - \delta\left(\frac{t(\overline{x})+c(\overline{x})}{q(\overline{x})} + r_a\right) - \frac{\gamma}{q(\overline{x})} = 0 \tag{117}
$$

On the other hand, consider the following problem, where we seek to minimize resource usage plus a share of excess land rents:

$$
\begin{aligned}
(\tilde{P}_{\min}) \qquad & \min & & \int_0^{\overline{x}} \frac{t(x)+c(x)}{q(x)} + r_a \, dx + (1-\mu)R & & & (118a)\\
& \text{subject to} & & u(c(x),q(x)) = \overline{u}, & & 0 \leq x \leq \overline{x} & (118b)\\
& & & \int_0^{\overline{x}} \frac{1}{q(x)} dx = L & & & (118c)
\end{aligned}
$$

Note that the objective function can be written as

$$
\int_0^{\overline{x}} \frac{t(x)+c(x)}{q(x)} + r_a \, dx + (1-\mu)R = \int_0^{\overline{x}} \frac{t(x)+c(x)}{q(x)} + \mu r_a + (1-\mu)\frac{u_q}{u_c} \, dx \tag{119}
$$

Then, the Lagrangian is

$$
\begin{aligned}
\Lambda_2(x) = & -\lambda_2 \int_0^{\overline{x}} \frac{t(x)+c(x)}{q(x)} + \mu r_a + (1-\mu)\frac{u_q}{u_c} \, dx \\
& + \int_0^{\overline{x}} \kappa(x)(u(c(x),q(x)) - \overline{u})dx + \rho \left[L - \int_0^{\overline{x}} \frac{1}{q(x)} \, dx \right]
\end{aligned} \tag{120}
$$

where $\kappa(x)$ and ρ are respectively the Lagrange multipliers associated with (118b) and (118c). The first order conditions are

$$
\kappa(x)u_c - \frac{\lambda_2}{q(x)} - \lambda_2 \left(\frac{1-\mu}{u_c^2}\right)(u_{cq}u_c - u_q u_{cc}) = 0 \qquad 0 \leq x \leq \overline{x} \tag{121}
$$

$$
\kappa(x)u_q - \frac{-\lambda_2 t(x) - \lambda_2 c(x) - \rho}{q(x)^2} - \lambda_2 \left(\frac{1-\mu}{u_c^2}\right)(u_{qq}u_c - u_q u_{cq}) = 0 \qquad 0 \leq x \leq \overline{x} \tag{122}
$$

$$
\int_0^{\overline{x}} \frac{1}{q(x)} - \int_0^{\overline{x}} \kappa(x) = 0 \tag{123}
$$

$$
-\lambda_2 \left(\frac{t(\overline{x})+c(\overline{x})}{q(\overline{x})} + r_a\right) - \frac{\rho}{q(\overline{x})} = 0 \tag{124}
$$

Finally, suppose that for some feasible control variables and parameters $(c^*(x), q^*(x), \overline{x}^*, \overline{u}^*)$ there exists a set of multipliers $(\nu^*(x), \lambda_1^*, \delta^*, \gamma^*)$ such that (114)-(117) hold. Then, note that there also exists a set of multipliers $(\kappa^*(x), \lambda_2^*, \rho^*)$ such that (121)-(124) also hold for $(c^*(x), q^*(x), \overline{x}^*, \overline{u}^*)$. Indeed, consider

$$(\kappa^*(x), \lambda_2^*, \rho^*) = (\nu^*(x), \delta^*, \gamma^* - \lambda_1^* \overline{u}) \tag{125}$$

Conversely, suppose that for some feasible control variables and parameters $(c^*(x), q^*(x), \overline{x}^*, \overline{u}^*)$ there exists a set of multipliers $(\kappa^*(x), \lambda_2^*, \rho^*)$ such that (121)-(124) hold. Then, note that there also exists a set of multipliers $(\nu^*(x), \lambda_1^*, \delta^*, \gamma^*)$ such that (114)-(117) also hold for $(c^*(x), q^*(x), \overline{x}^*, \overline{u}^*)$. Indeed, consider

$$(\nu^*(x), \lambda_1^*, \delta^*, \gamma^*) = (\kappa^*(x), m, \lambda_2^*, \rho^* + m\overline{u}) \tag{126}$$

for any $m \geq 0$.

G Proof of Proposition 8

We consider that the cost of providing public transport is proportional, with constant k, to the number of passenger-miles. Define $\tilde{a} = a - k$, then the self-financing of the system can we written as

$$\int_0^{\overline{x}} \frac{ax+b}{q} dx = \int_0^{\overline{x}} \frac{kx}{q} dx \Leftrightarrow \int_0^{\overline{x}} \frac{\tilde{a}x+e}{q} dx = 0 \tag{127}$$

Additionally, define $\tilde{t} = t + k$ and note that the budget constraint for the households located at x is

$$\mu \frac{R}{L} + y - (t+a)x - b - c(x) - p(x)q(x) = 0 \tag{128}$$

$$\Leftrightarrow \mu \frac{R}{L} + y - (\tilde{t}+\tilde{a})x - b - c(x) - p(x)q(x) = 0 \tag{129}$$

Thus, considering a city with commuting time cost t, with a public transport system with cost k per passenger-mile and a fare $ax+b$ is equivalent to considering a city with commuting time cost $\tilde{t}$, with a public transport system with zero cost per passenger-mile and a fare $\tilde{a}x+b$. This is the setting we study in this proof.

The equilibrium condition can be stated as

$$v\left(\frac{\mu R}{L} + y - (\tilde{t}+\tilde{a})x - b - pq, q\right) = \overline{u} \qquad \forall x \in [0, \overline{x}] \tag{130}$$

with v the indirect utility. Totally differentiating (130) with respect to $\tilde{a}$ we obtain

$$v_c\left(\frac{\mu}{L}\frac{dR}{d\tilde{a}} - x - \frac{db}{d\tilde{a}} - q\frac{dp}{d\tilde{a}} - \cancel{p\frac{dq}{d\tilde{a}}}\right) + \cancel{v_q\frac{dq}{d\tilde{a}}} = \frac{d\overline{u}}{d\tilde{a}} \qquad \forall x \in [0, \overline{x}] \tag{131}$$

where the canceled terms follows from the utility maximization condition:

$$\frac{v_q}{v_c} = p \quad \forall x \in [0, \overline{x}] \tag{132}$$

On the other hand, the Muth-Mills condition is:

$$\frac{\partial p}{\partial x} = -\frac{\tilde{t}+\tilde{a}}{q} \quad \forall x \in [0,\bar{x}] \tag{133}$$

Integrating (133) with respect to x we get:

$$p(x=0) - r_a = L(\tilde{t}+\tilde{a}) \Rightarrow \frac{dp(x=0)}{d\tilde{a}} = L \tag{134}$$

Note that from the first equality of (134), it follows that for any feasible $\tilde{a}$, $\tilde{a} > -\tilde{t}$. Then, evaluating (131) at $x=0$, using (132) and (134), we can obtain

$$\frac{\mu}{L}\frac{dR}{d\tilde{a}} - \frac{db}{d\tilde{a}} - Lq(x=0) = \frac{1}{v_c}\frac{d\bar{u}}{d\tilde{a}} \tag{135}$$

Since we are interested in maximizing the aggregated utility over the city under an equilibrium, equivalently we can maximize $\bar{u}$. Thus, to find the optimal $\tilde{a}$, we focus on the critical points of $\bar{u}$ as a function of $\tilde{a}$, i.e., the slopes $\tilde{a}^*$ such that $\left.\frac{d\bar{u}}{d\tilde{a}}\right|_{\tilde{a}^*} = 0$. From (135) is easy to note that the set of critical $\tilde{a}^*$ fulfills:

$$\frac{\mu}{L}\left.\frac{dR}{d\tilde{a}}\right|_{\tilde{a}^*} - \left.\frac{db}{d\tilde{a}}\right|_{\tilde{a}^*} = Lq(x=0)|_{\tilde{a}^*} \tag{136}$$

For the computation of $\left.\frac{dR}{d\tilde{a}}\right|_{\tilde{a}^*}$, we start from the definition of R:

$$R = \int_0^{\bar{x}} p(x) - r_a \, dx \Rightarrow \frac{dR}{d\tilde{a}} = (p(\bar{x}) - r_a)\frac{d\bar{x}}{d\tilde{a}} + \int_0^{\bar{x}} \frac{dp}{d\tilde{a}} dx \tag{137}$$

$$= \int_0^{\bar{x}} \frac{1}{q}\frac{\mu}{L}\frac{dR}{d\tilde{a}} - \frac{x}{q} - \frac{1}{q}\frac{db}{d\tilde{a}} - \frac{1}{q \cdot v_c}\frac{d\bar{u}}{d\tilde{a}} dx \tag{138}$$

The last inequality folllows from $p(\bar{x}) = r_a$, and the characterization of $\frac{dp}{d\tilde{a}}$ given by (131). Then, it follows from (138) that for the set of critical $\tilde{a}^*$:

$$\left.\frac{dR}{d\tilde{a}}\right|_{\tilde{a}^*} = L\left(\frac{\mu}{L}\left.\frac{dR}{d\tilde{a}}\right|_{\tilde{a}^*} - \left.\frac{db}{d\tilde{a}}\right|_{\tilde{a}^*}\right) - \int_0^{\overline{x}} \frac{x}{q} dx \tag{139}$$

$$= L^2 q(x=0)\big|_{\tilde{a}^*} - \int_0^{\overline{x}} \frac{x}{q} dx \tag{140}$$

where the last inequality follows from (136). Then, remember the self-financing condition:

$$\int_0^{\overline{x}} \frac{\tilde{a}x+b}{q} dx = 0 \Rightarrow b = -\frac{\tilde{a}}{L}\int_0^{\overline{x}} \frac{x}{q} dx \tag{141}$$

Replacing q through the Muth-Mills condition (133)

$$b = \frac{\tilde{a}}{L(\tilde{t}+\tilde{a})}\int_0^{\overline{x}} \left(\frac{\partial p}{\partial x}\right) x\, dx \tag{142}$$

$$= \frac{\tilde{a}}{L(\tilde{t}+\tilde{a})}\int_0^{\overline{x}} \frac{\partial (px)}{\partial x} - p(x)\, dx \tag{143}$$

$$= \frac{\tilde{a}}{L(\tilde{t}+\tilde{a})}\left(p(\overline{x})\overline{x} - \int_0^{\overline{x}} p(x)\, dx\right) \tag{144}$$

$$= \frac{\tilde{a}}{L(\tilde{t}+\tilde{a})}\left(\int_0^{\overline{x}} r_a - p(x) dx\right) \tag{145}$$

$$= -\frac{\tilde{a}}{L(\tilde{t}+\tilde{a})} R \tag{146}$$

Using (141) and (146), we can obtain:

$$\int_0^{\overline{x}} \frac{x}{q} dx = \frac{R}{\tilde{t}+\tilde{a}} \tag{147}$$

while totally differentiating (146) we get

$$\frac{db}{d\tilde{a}} = -\frac{t}{L(\tilde{t}+\tilde{a})^2} R - \frac{\tilde{a}}{L(\tilde{t}+\tilde{a})}\frac{dR}{d\tilde{a}} \tag{148}$$

Finally, using (140), (147) and (148) in (136), we find that the set of critical $\tilde{a}^*$ is characterized by

$$\left[\frac{\mu}{L}\left(L^2 q(x=0)\big|_{\tilde{a}^*} - \frac{R}{\tilde{t}+\tilde{a}^*}\right)\right] - \left[-\frac{t}{L(\tilde{t}+\tilde{a}^*)^2}R - \frac{\tilde{a}^*}{L(\tilde{t}+\tilde{a}^*)}\left(L^2 q(x=0)\big|_{\tilde{a}^*} - \frac{R}{\tilde{t}+\tilde{a}^*}\right)\right] = Lq(x=0)\big|_{\tilde{a}^*} \quad (149)$$

Reordering (149), the set of critical $\tilde{a}^*$ is then characterized by

$$\tilde{a}^* = (t - \mu(\tilde{t}+\tilde{a}^*))\left(1 - \frac{L^2 q(x=0)(\tilde{t}+\tilde{a}^*)}{R}\right) \quad (150)$$

where both $q(x=0)$ and R depend on $\tilde{a}^*$. We will show that

$$1 < \frac{L^2 q(x=0)(\tilde{t}+\tilde{a}^*)}{R} < 2 \quad \text{if} \quad \frac{\partial(pq)}{\partial x} < 0 \quad (151)$$

Indeed, using the first equality of (134), we get

$$\begin{aligned} L^2 q(x=0)(\tilde{t}+\tilde{a}^*) - R &= Lq(x=0)(p(x=0) - r_a) - \int_0^{\bar{x}} p(x) - r_a \, dx & (152)\\ &= r_a\left(\int_0^{\bar{x}} 1 - \frac{q(0)}{q(x)} dx\right) + \int_0^{\bar{x}} \frac{1}{q(x)}\left(p(0)q(0) - p(x)q(x)\right) dx & (153) \end{aligned}$$

Since $q(x) > q(x=0)$ $\forall x \in (0,\bar{x}]$, the first term of (153) is positive. Additionally, if $\frac{\partial(pq)}{\partial x} < 0$, the second term is positive aswell, and we get the left-hand size inequality of (151). On the other hand,

$$\begin{aligned} L^2 q(x=0)(\tilde{t}+\tilde{a}^*) - R &= \int_0^{\bar{x}} r_a \frac{1}{q(x)}(q(x) - q(0))dx + \int_0^{\bar{x}} \frac{1}{q(x)}(p(0)q(0) - p(x)q(x))dx & (154)\\ &< \int_0^{\bar{x}} \frac{1}{q(x)}(p(x)q(x) - p(x)q(0)) + \frac{1}{q(x)}\left(p(0)q(0) - p(x)q(x)\right) dx & (155)\\ &= \int_0^{\bar{x}} \frac{q(0)}{q(x)}(p(0) - p(x))dx & (156) \end{aligned}$$

But since in this case p is a convex function of x (Brueckner, 1987), the following holds

$$p(x) \geq p(0) + x \cdot \left.\frac{\partial p}{\partial x}\right|_{x=0} \tag{157}$$

$$\Rightarrow p(0) - p(x) \leq -x \cdot \left.\frac{\partial p}{\partial x}\right|_{x=0} = (\tilde{t} + \tilde{a}^*)\frac{x}{q(0)} \tag{158}$$

where we used the Muth-Mills condition (133) for the last equality. Using (158) in (156) we have

$$L^2 q(x=0)(\tilde{t} + \tilde{a}^*) - R < (\tilde{t} + \tilde{a}^*)\int_0^{\overline{x}} \frac{x}{q(x)} dx \tag{159}$$

$$= R \tag{160}$$

and the right-hand size inequality of (151) follows. We now show that both inequalities of (151) imply that $\tilde{a}^* = 0$ if $\mu = 1$, and that $\tilde{a}^*$ cannot be positive for $\mu \in [0,1)$. For this, we study three cases:

a) $t - \mu(\tilde{t} + \tilde{a}^*) > 0$. Note that this case always hold for $\mu = 0$, and for $\tilde{a}^* < \left(\frac{1-\mu}{\mu}\right) t$ if $\mu \neq 0$. Then, using (151) in (150) we have

$$-(t - \mu(\tilde{t} + \tilde{a}^*)) < \tilde{a}^* < 0 \tag{161}$$

$$\Rightarrow \begin{cases} -t < \tilde{a}^* < 0 & \text{if } \mu \in [0,1) \\ \tilde{a}^* < \tilde{a}^* < 0 & \text{if } \mu = 1 \end{cases} \tag{162}$$

Thus, this case is infeasible for $\mu = 1$, and allows only negative values of $\tilde{a}^*$ if $\mu \neq 1$.

b) $t - \mu(\tilde{t} + \tilde{a}^*) < 0$. Note that this case never holds for $\mu = 0$, and for $\tilde{a}^* \geq \left(\frac{1-\mu}{\mu}\right) t$ if $\mu \neq 0$. Then, using (151) in (150) we have

$$0 < \tilde{a}^* < -(t - \mu(\tilde{t} + \tilde{a}^*)) \tag{163}$$

$$\Rightarrow \begin{cases} 0 < \tilde{a}^* < -t & \text{if } \mu \in [0,1) \\ 0 < \tilde{a}^* < \tilde{a}^* & \text{if } \mu = 1 \end{cases} \tag{164}$$

Thus, this case is infeasible for every $\mu \in [0,1]$.

c) $t - \mu(\tilde{t} + \tilde{a}^*) = 0$. This case holds if and only if $\tilde{a}^* = t(1-\mu)$. But using $t - \mu(\tilde{t} + \tilde{a}^*) = 0$ in (150), this case only admits solution $\tilde{a}^* = 0$. Thus, we must have $0 = t(1-\mu)$ which only happens if $\mu = 1$.

The conclusions follows when we return to the initial notation, $a = \tilde{a} + k$, $t = \tilde{t} - k$.

References

Ahlfeldt, G.M., Redding, S.J., Sturm, D.M., Wolf, N., 2015. The economics of density: Evidence from the berlin wall. Econometrica 83, 2127–2189.

Alonso, W., 1964. Location and land use. Toward a general theory of land rent. Cambridge, Mass.: Harvard University Press.

Anas, A., Kim, I., 1996. General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. Journal of Urban Economics 40, 232–256.

Anas, A., Rhee, H.J., 2007. When are urban growth boundaries not second-best policies to congestion tolls? Journal of Urban Economics 61, 263–286.

Arnott, R., 1979a. Optimal city size in a spatial economy. Journal of Urban Economics 6, 65–89.

Arnott, R., 1979b. Optimal taxation in a spatial economy with transport costs. Journal of Public Economics 11, 307–334.

Arnott, R.J., 1979c. Unpriced transport congestion. Journal of Economic Theory 21, 294–316.

Arribas-Bel, D., Sanz-Gracia, F., 2014. The validity of the monocentric city model in a polycentric age: Us metropolitan areas in 1990, 2000 and 2010. Urban Geography 35, 980–997.

Basso, L.J., Silva, H.E., 2014. Efficiency and substitutability of transit subsidies and other urban transport policies. American Economic Journal: Economic Policy 6, 1–33.

Baum-Snow, N., 2007. Suburbanization and transportation in the monocentric model. Journal of Urban Economics 62, 405–423.

Bertaud, A., Brueckner, J.K., 2005. Analyzing building-height restrictions: predicted impacts and welfare costs. Regional Science and Urban Economics 35, 109–125.

Borck, R., Wrede, M., 2005. Political economy of commuting subsidies. Journal of Urban Economics 57, 478–499.

Brueckner, J.K., 1987. The structure of urban equilibria: A unified treatment of the muth-mills model. Handbook of regional and urban economics 2, 821–845.

Brueckner, J.K., 2000. Urban sprawl: diagnosis and remedies. International regional science review 23, 160–171.

Brueckner, J.K., 2005. Transport subsidies, system choice, and urban sprawl. Regional Science and Urban Economics 35, 715–733.

Brueckner, J.K., 2007. Urban growth boundaries: An effective second-best remedy for unpriced traffic congestion? Journal of Housing Economics 16, 263–273.

Brueckner, J.K., 2011. Lectures on urban economics. MIT press.

Brueckner, J.K., Helsley, R.W., 2011. Sprawl and blight. Journal of Urban Economics 69, 205–213.

Brueckner, J.K., Kim, H.A., 2003. Urban sprawl and the property tax. International Tax and Public Finance 10, 5–23.

De Lara, M., De Palma, A., Kilani, M., Piperno, S., 2013. Congestion pricing and long term urban form: Application to paris region. Regional Science and Urban Economics 43, 282–295.

Diamond, R., McQuade, T., 2019. Who wants affordable housing in their backyard? an equilibrium analysis of low-income property development. Journal of Political Economy 127, 1063–1117.

Dieleman, F., Wegener, M., 2004. Compact city and urban sprawl. Built Environment 30, 308–323.

Duranton, G., Puga, D., 2015. Urban land use, in: Handbook of regional and urban economics. Elsevier. volume 5, pp. 467–560.

Ewing, R., Hamidi, S., 2015. Compactness versus sprawl: A review of recent evidence from the united states. Journal of Planning Literature 30, 413–432.

Fujita, M., 1989. Urban economic theory: land use and city size. Cambridge University Press.

Fujita, M., Thisse, J.F., 2013. Economics of Agglomeration: Cities, Industrial Location, and Globalization. Cambridge University Press.

Gordon, P., Richardson, H.W., 1997. Are compact cities a desirable planning goal? Journal of the American planning association 63, 95–106.

Gubins, S., Verhoef, E.T., 2014. Dynamic bottleneck congestion and residential land use in the monocentric city. Journal of Urban Economics 80, 51–61.

Heblich, S., Redding, S.J., Sturm, D.M., 2020. The making of the modern metropolis: evidence from london. The Quarterly Journal of Economics 135, 2059–2133.

Henderson, J.V., Regan, T., Venables, A.J., 2020. Building the city: from slums to a modern metropolis. Review of Economic Studies. In Press .

Kanemoto, Y., 1977. Cost-benefit analysis and the second best land use for transportation. Journal of Urban Economics 4, 483–503.

Kanemoto, Y., 1980. Theories of urban externalities. Amsterdam: North-Holland.

Kono, T., Joshi, K.K., 2012. A new interpretation on the optimal density regulations: Closed and open city. Journal of Housing Economics 21, 223–234.

Kono, T., Joshi, K.K., Kato, T., Yokoi, T., 2012. Optimal regulation on building size and city boundary: An effective second-best remedy for traffic congestion externality. Regional Science and Urban Economics 42, 619–630.

Kono, T., Kawaguchi, H., 2017. Cordon pricing and land-use regulation. The Scandinavian Journal of Economics 119, 405–434.

Kono, T., Pines, D., Yokoi, T., 2019. Spatially-variable property tax and optimal tax composition in congested monocentric cities: George, pigou, ramsey and strotz unified. Journal of Urban Economics 112, 122–132.

US Department of Labor, B.o.L.S., 2018. Consumer expenditure survey.

Li, Z.C., Chen, Y.J., Wang, Y.D., Lam, W.H., Wong, S., 2013. Optimal density of radial major roads in a two-dimensional monocentric city with endogenous residential distribution and housing prices. Regional Science and Urban Economics 43, 927–937.

Li, Z.C., Guo, Q.W., 2017. Optimal time for implementing cordon toll pricing scheme in a monocentric city. Papers in Regional Science 96, 163–190.

Li, Z.C., Lam, W.H., Wong, S., 2012. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. Transportation Research Part B: Methodological 46, 30–49.

Li, Z.C., Wang, Y.D., 2018. Analysis of multimodal two-dimensional urban system equilibrium for cordon toll pricing and bus service design. Transportation Research Part B: Methodological 111, 244–265.

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

McGrath, D.T., 2005. More evidence on the spatial scale of cities. Journal of Urban Economics 58, 1–10.

McGuckin, N., Fucci, A., 2018. Summary of travel trends: 2017 national household travel survey. Washington, DC: Federal Highway Administration .

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

Mirrlees, J.A., 1972. The optimum town. The Swedish Journal of Economics , 114–135.

Mohring, H., 1972. Optimization and scale economies in urban bus transportation. The American Economic Review 62, 591–604.

Mun, S.i., Konishi, K.j., Yoshikawa, K., 2003. Optimal cordon pricing. Journal of Urban Economics 54, 21–38.

Mun, S.i., Konishi, K.j., Yoshikawa, K., 2005. Optimal cordon pricing in a non-monocentric city. Transportation Research Part A: Policy and Practice 39, 723–736.

Muth, R., 1969. Cities and housing: The spatial patterns of urban residential land use. University of Chicago, Chicago, 4, 114-123.

Nechyba, T.J., Walsh, R.P., 2004. Urban sprawl. Journal of economic perspectives 18, 177–200.

Oron, Y., Pines, D., Sheshinski, E., 1973. Optimum vs. equilibrium land use pattern and congestion toll. The Bell Journal of Economics and Management Science , 619–636.

Oueslati, W., Alvanides, S., Garrod, G., 2015. Determinants of urban sprawl in european cities. Urban Studies 52, 1594–1614.

Papageorgiou, Y.Y., Pines, D., 1999. An essay on urban economic theory. volume 1. Springer Science & Business Media.

Parry, I.W., Bento, A., 2001. Revenue recycling and the welfare effects of road pricing. Scandinavian Journal of Economics 103, 645–671.

Parry, I.W., Small, K.A., 2009. Should urban transit subsidies be reduced? American Economic Review 99, 700–724.

Pines, D., Kono, T., 2012. Far regulations and unpriced transport congestion. Regional Science and Urban Economics 42, 931–937.

Pines, D., Sadka, E., 1985. Zoning, first-best, second-best, and third-best criteria for allocating land for roads. Journal of Urban Economics 17, 167–183.

Small, K.A., Verhoef, E.T., 2007. The economics of urban transportation. Routledge.

Sommer, K., Sullivan, P., 2018. Implications of us tax policy for house prices, rents, and homeownership. American Economic Review 108, 241–74.

Song, Y., Zenou, Y., 2006. Property tax and urban sprawl: Theory and implications for us cities. Journal of urban economics 60, 519–534.

Tikoudis, I., Verhoef, E.T., van Ommeren, J.N., 2015. On revenue recycling and the welfare effects of second-best congestion pricing in a monocentric city. Journal of Urban Economics 89, 32–47.

Tikoudis, I., Verhoef, E.T., van Ommeren, J.N., 2018. Second-best urban tolls in a monocentric city with housing market regulations. Transportation Research Part B: Methodological 117, 342–359.

Tsivanidis, N., 2019. Evaluating the impact of urban transit infrastructure: Evidence from bogotá's transmilenio. Unpublished manuscript .

Turvey, R., Mohring, H., 1975. Optimal bus fares. Journal of Transport Economics and Policy , 280–286.

Verhoef, E.T., 2005. Second-best congestion pricing schemes in the monocentric city. Journal of Urban Economics 58, 367–388.

Wheaton, W.C., 1998. Land use and density in cities with congestion. Journal of urban economics 43, 258–272.

Wildasin, D.E., 1986. Spatial variation of the marginal utility of income and unequal treatment of equals. Journal of Urban Economics 19, 125–129.

Zhang, W., Kockelman, K.M., 2016. Optimal policies in cities with congestion and agglomeration externalities: Congestion tolls, labor subsidies, and place-based strategies. Journal of Urban Economics 95, 64–86.