How to make the metropolitan area work? Neither big government, nor laissez-faire.*

Carl Gaigné[†] Stéphane Riou[‡] and Jacques-François Thisse[§]

August 1, 2015

Abstract

We study how administrative boundaries and tax competition among asymmetric jurisdictions interact with the labor and land markets to determine the economic structure and performance of metropolitan areas. Contrary to general belief, cross-border commuting need not be welfare-decreasing in presence of agglomeration economies that vary with the distribution of firms within the metropolitan area. Tax competition implies that the central business district is too small and prevents public policy enhancing global productivity to produce their full impact. Although our results support the idea of decentralizing the provision of local public services by independent jurisdictions, they highlight the need of coordinating tax policies and the importance of the jurisdiction sizes within metropolitan areas.

Keywords: metropolitan area, tax competition, local labor markets, job decentralization, administrative boundary

JEL classification: H41, H71, R12

*We are grateful to K. Konrad and three referees for details comments and suggestions. We also thank M. Berliant, R. Braid, J. Hamilton, V. Henderson, R. Inman, S. Kok, F. Moizeau, H. Overman, C. Perroni, R. van der Ploeg, as well as participants to seminars at PSE, LSE, University of Warwick, and the Public Policies and Spatial Economics Workshop (Lyon-Saint-Etienne), for their comments and remarks. S. Riou and J.-F. Thisse gratefully acknowledge the financial support from, respectively, the Région Rhône-Alpes and the Government of the Russian Federation under the grant 11.G34.31.0059.

[†]INRA, UMR1302, SMART Rennes (France) and CREATE, Laval University, Quebec (Canada).

[‡]Université Jean Monnet-Saint-Etienne (France) and CNRS-GATE Lyon Saint-Etienne (France).

[§]CORE, Université catholique de Louvain (Belgium), Higher School of Economics (Russia) and CEPR.

1 Introduction

Urban sprawl and the decentralization of jobs have given rise to metropolitan areas that include a large number of political jurisdictions providing local public goods to their residents and competing in tax levels to attract jobs and residents. A few facts documented by Brülhart et al. (2015) give an idea of the magnitude of this evolution. Metropolitan areas with more than 500,000 inhabitants are divided, on average, into 74 local jurisdictions, while local governments in the OECD raise about 13 percent of total tax revenue. Empirical works by Glaeser and Kahn (2001) and Cheshire and Magrini (2009) confirm the idea that the institutional structure of a metropolitan area has a significant impact on both the efficiency of its local public services and the welfare of its residents by influencing the distribution of jobs and the level of housing costs. This difficulty has not escaped the attention of policy-makers and analysts. According to Alain Juppé, a former prime minister of France and mayor of the city of Bordeaux, "governments are too small to deal with the big problems and too big to deal with the small problems" within today's political limits. Bruce Katz, a vice president at the Brookings Institution, went one step further when he said that "metro governance is almost uniformly characterized by fragmentation and balkanization, by cultures of competition rather than one of collaboration." Since metropolitan areas also produce a sizable and growing share of the wealth of nations, we may safely conclude that there is a need for a sound economic analysis of those entities.

The complexity of the metropolitan environment has led several policy-makers to stress the need for coordinating the actions of local governments. To seriously assess the desirability and scope of such a move, we need to understand how the institutional design of metropolitan areas affects the various channels that link the local governments, the labor market and the land market. The purpose of this paper is to develop a model with one central city and several suburban jurisdictions, in which tax competition among jurisdictions interacts with those two markets to shape the economic structure and performance of the metropolitan area.

The standard approach to jurisdiction/club formation is to focus on the trade-off between the crowding effect of public services, which increases with jurisdiction size, and the unit cost of public services, which decreases with population size. We contend that the problem may be tackled from a different, but equally important, angle by recognizing that the *administrative* and *economic* boundaries of jurisdictions usually differ within metropolitan areas. Put differently, workplaces and residences do not necessarily belong to the same jurisdiction. In practice, the central city attracts a large number of workers who live in adjacent areas, thus giving rise to a substantial amount of "cross-border" commuting. Therefore, workers face the following trade-off: either they can earn a high wage in centrally located firms and bear high commuting costs, or they can receive lower pay in firms located in secondary business centers and pay less for commuting. By combining these two trade-offs within a unifying framework, we distinguish between the administrative and economic limits of the central city, a distinction that has not attracted much attention in the literature (Glaeser, 2013; Brülhart et al., 2015).

This is not yet the end of the story. It is well documented that the productivity of labor is higher in larger cities. More precisely, there is a broad consensus to recognize that, everything else being equal, the elasticity of labor productivity with respect to the current employment density is about three percent. This elasticity measures the static gains generated by a higher employment density (Combes and Gobillon, 2015). The reason for this lies in the existence of various mechanisms, generically nicknamed "agglomeration economies," which have the nature of increasing returns external to firms (Glaeser and Gottlieb, 2009). Another major trend shaping modern metropolitan areas is the decentralization of jobs toward secondary employment centers, fostered by the development of new information technologies (Glaeser and Kahn, 2001, 2004). As a result, suburban jurisdictions accommodate secondary employment centers, thus making many metropolitan areas polycentric. Any serious attempt that aims to study how the metropolitan area works must take into account these two essential parts.

To carry out our study, we develop a general equilibrium model that accounts for the following features: (i) consumers are free to choose their locations, (ii) there is cross-border commuting, (iii) local governments compete strategically, (iv) the labor and land markets determine wages and land rents, and (v) the level of agglomeration economies depends on the distribution of firms. The flip side of the coin is the need to reduce the complexity of the problem by assuming that consumers are homogeneous.

Our main findings may be organized in three distinct, but complementary, categories.

1. We study the first-best outcome, which we use later on as a benchmark. The planner, who aims to maximize welfare within the *whole* metropolitan area, determines the areas providing the public services and the employment centers by choosing where consumers live and work. First, it is not desirable to amalgamate the suburban areas with the central one. Moreover, the economic boundary of the central city always encompasses its administrative boundary. Put together, these results imply that *the optimal political and economic boundaries of the central city do not coincide*, a result that clashes with the general belief that these boundaries should be the same (OECD, 2006). This difference stems from the trade-offs that determine each type of boundary. The optimal size of jurisdictions is tightly related to the optimal degree of decentralization in the supply of local public goods. More precisely, decentralizing the provision of local public goods is socially desirable as long as the degree of increasing returns in producing these goods is not too high. By contrast, the optimal size of central and secondary business centers depends on the interplay between commuting and agglomeration economies, that is, parameters that do not enter the above trade-off. When commuting costs are not too low,

agglomeration economies are not too strong, or both, jobs are shared between the central and suburban areas.

2. We then analyze the decentralized outcome when the number of jurisdictions and their administrative boundaries are exogenously given. The game involves three types of players: a large number of consumers, a large number of firms, and a finite number of local governments. Consumers choose a residence and a workplace. Firms choose a location and the wages paid to their employees. Local governments provide local public goods and choose non-cooperatively a *business tax* paid by the firms located in their jurisdiction and a *land tax* levied on the aggregate land rent. Since agglomeration economies are unevenly distributed while workplaces and residences need not belong to the same jurisdiction, tax competition unfolds within the entire metropolitan area, which leads to a game richer than those used in standard models of fiscal competition.

We show that, under mild conditions, *the central city typically levies a higher business tax than suburban governments*, a result that is backed up by the empirical literature (see Koh et al., (2013) and Brülhart et al. (2015) for references). The reason for this is that consumers working in the central city need not reside therein, which incentivizes the central city's government to practice *tax exporting*, the extent of which depends on the relative size of jurisdictions and the intensity of agglomeration economies. Once the agglomeration economies in the central city exceeds a certain threshold, the tax policy adopted by the suburban jurisdictions is predominantly governed by the residents working in the central city, which leads these jurisdictions to set a positive business tax. Moreover, as they influence the location choices, tax policies affect the overall productivity of the metropolitan area. Specifically, we demonstrate that the equilibrium size of the central business center is too small and, therefore, the potential provided by agglomeration economies is not fully used. In other words, *tax competition reduces the gross domestic product per capita of the metropolitan area through an excessive decentralization of jobs*. This result shows the importance of using a spatial setting in which the commuting pattern is endogenous through the location choices made by firms and consumers.

What is more, under tax competition, when the population size of the central city is optimal, the central business district is too small, whereas the former is too large when the size of the latter has its optimal size. Therefore, redrawing the limit of the central city is not the remedy to correct the misallocation of jobs within the metropolitan area. This tension stems from the fact that the distribution of jobs is governed by a system of forces that overlaps imperfectly with that taken into account by the local governments. As a consequence, there is no reason to expect the two types of boundaries to coincide. It should be stressed, however, that the misallocation of jobs is exacerbated when the relative population size of the central city is small. Furthermore, although higher agglomeration economies, lower commuting costs, or both raise the global efficiency of the metropolitan area, the gap between the optimal and equilibrium size of the central business districts grows.

3. Once it is recognized that suburbanites commuting to the central business district may consume the public services supplied by the central city, the tax gap widens because the central city sets an even higher tax rate to reduce the production costs borne by its residents. All in all, *the central city residents bear higher provision costs for their public services and earn lower wages*. In other words, they are hurt twice.

Our analysis suggests that neither the amalgamation nor the decentralization among competing jurisdictions is the best way to govern large metropolitan areas. Instead, combining a multi-jurisdictional political system with an economic government of the metropolitan area or a deep inter-jurisdictional cooperation seems to be a more efficient way to solve the various distortions inherent to the working of a metropolitan economy. In particular, our findings point to the need for common governance in local tax policies within metropolitan areas. Such a recommendation has been implemented in a few European countries under the concrete form of fiscal coordination (OECD, 2006). In the United States, the tax-base sharing program implemented in Minneapolis-Saint Paul has decreased incentives for local governments to compete for a larger tax base (Inman, 2009).

A last comment is in order. The legal environment in which metropolitan areas operate vastly differs across countries. The model presented in this paper is context-free in that it focuses on (some of) the fundamental characteristics common to most metropolitan areas and disregards specific and idiosyncratic issues that are important in some countries but not in others.

Related literature. Tax competition models with jurisdictional asymmetries ignore the typical urban variables (Bucovetsky, 1991; Kanbur and Keen, 1993; Keen and Konrad, 2013), while only a handful of papers have studied the economic organization of a metropolitan area within the framework of urban economics. Hoyt (1991) and Noiset and Oakland (1995) did not account for the fact that jobs may be located outside the central city, while Braid (2002) disregarded tax competition. In some papers, when local governments compete for capital, consumers/workers live and work within the same jurisdiction in equilibrium (Braid, 1996; Epple and Zelenitz, 1981). Braid (2010) focused on the distances between jurisdictions that choose to offer, or not to offer, a local public good that may be consumed by non-residents. He did not study, however, the interactions between the provision of public services and the labor and land markets. We are closer to Janeba and Osterloh (2013) in that both papers make a clear distinction between the central city and the suburban jurisdictions. However, their purpose is different from ours. These authors study tax competition within and between metropolitan areas. While recognizing that intercity competition is an important issue, we believe that there is much to be done to understand how the political fragmentation of the metropolitan area shapes its economic structure, something that can be achieved only in a full-fledged model of urban economics, such as ours. More precisely, we account for ingredients that are crucial to understand the structure of metropolitan areas, such as land consumption, agglomeration economies and cross-border commuting.

In the next section, we present the model. In Section 3, we study the socially optimal organization of the metropolitan area, whereas Section 4 focuses on the decentralized outcome. In Section 5, we consider the second-best outcome in which a planner chooses the welfare-maximizing limit of the central city. In Section 6, we check the robustness of our main findings when (i) suburbanites working in the central business district consume the public services provided by the central city, and (ii) the central city supplies a broader array of public services than the suburban jurisdictions. Section 7 concludes.

2 The model and preliminary results

The metropolitan economy is populated by L consumers who are free to choose their residential location and workplace. There are three consumption goods: (i) land, which is used as a proxy for housing, (ii) a local public good financed by jurisdictions through business and land taxes, and (iii) a homogeneous good, the numéraire, produced by profit-maximizing firms whose locations are endogenous. Our setup can easily been extended to the case of workers producing a differentiated good under monopolistic competition. Note also that the numéraire can be used to import other goods produced in other cities.

2.1 Jurisdictions and the provision of public goods

The metropolitan area (MA) is formed by $m + 1$ jurisdictions. It is endowed with a *hub-and-spoke* transportation network, which means that the $m \geq 2$ suburban jurisdictions are connected only to the central city, which has a direct access to all suburbanites.[1] Formally, the MA is described by m one-dimensional half-lines sharing the same initial point $x = 0$. Distances and locations are expressed by the same variable x measured from 0.

The central city hosts the central business district (CBD) of the MA at $x = 0$, while each suburban jurisdiction may, or may not, accommodate a secondary business district (SBD). The internal economic structure of the MA is endogenous in that the CBD and SBD sizes are

[1] We could use a more general setup in which the suburban jurisdictions are directly connected to their neighbors. However, doing this increases the complexity of the formal developments while the gains in results are limited. What matters for most of our results is that the central city has more direct connections to the other jurisdictions than each suburban jurisdiction.

variable and determined at the equilibrium. When a suburban jurisdiction accommodates a SBD, we call it an *edge city* to differentiate it from the central city. In this case, the MA is said to be *polycentric*; otherwise, it is *monocentric*. From now on, we will focus on the case of a polycentric MA. We will give simple sufficient conditions for this to hold in equilibrium. Therefore, the corresponding domain of admissible parameter values is much smaller than the actual domain for which the MA is polycentric.

The central city shares the same boundary b with each suburban jurisdiction, while all suburban jurisdictions have the same outer limit B. The economic limit y_i of the central city along the ith spoke is given by the location of the individual who is indifferent between working in the CBD or in the ith SBD located at a distance x_i^s from the CBD. Therefore, all workers located in the interval $[0, y_i]$ work in the CBD while those in $(y_i, B]$ work in the SBD. Because we focus on a symmetric outcome, we may delete the subscript i : $y_i = y$ and $x_i^s = x^s$ for $i = 1, ..., m$. Note that the economic limit y need not coincide with its administrative limit b, which implies the existence of cross-border commuting.

We treat local governments as if they were of a single type, which means that each one supplies the same amount of public good G to its residents (e.g., schools, day care centers and recreational facilities).[2] Only the local residents use the public services supplied by their jurisdictions. That G is exogenous and the same across jurisdictions is in accordance with what is observed in several European countries where the level of public good is determined by constitutional rules that impose the equal treatment of people in the access to some public services. However, this does not apply to more decentralized countries, such as the United States or Switzerland, where jurisdictions are free to choose the amount of public good they supply to their residents. As a consequence, our analysis may be viewed as a setup that focuses on the government revenue side. In Section 6, we will allow the cross-border commuters to consume the central city public services and the central city to supply a broader range of public services than the suburban jurisdictions.

When the supplied population is l, the cost of providing the public good G is given by

$$C(l) = F + \frac{c}{2} l^2. \tag{1}$$

In this expression, F stands for a jurisdiction's investment outlays and $cl^2/2 > 0$ the variable production cost that increases with the population size l of the jurisdiction. This specification features two characteristics that are often encountered in the provision of local public goods: (i) the need to build facilities having a minimum size because of scale economies in the production of public goods and (ii) the congestible nature of these facilities, so that supplying a rising

[2] This is consistent with Hochman et al. (1995) who show that the political fragmentation between different levels of governments may generate a substantial waste of resources within the MA.

number of users requires growing expenditures. In other words, (1) may be viewed as a reduced form that captures all the costs borne in the jurisdiction when the local public good is supplied to a larger population. For example, our results hold true if the variable costs of the public good are zero and the local public good is congestible with individual utility level given by $G - cl^2/2$.

2.2 Consumers and land rents

Consumers share the same preferences and are endowed with one unit of labor. So far, few papers in urban economics have been able to account for consumers endowed with different preferences. Our paper is no exception. That said, we want to stress that working with heterogeneous consumers in a setup that accounts for agglomeration economies, tax competition, workers' commuting and a land market is out of (our) reach. Assuming that consumers share the same preferences eliminates the Tiebout motive for population sorting, but allows us to insulate the various interactions between tax competition, agglomeration economies and commuting.

For simplicity, we assume that consumers use a land plot having the same fixed size; the unit of land is chosen for this parameter to be normalized to one. This assumption does not allow replicating the well-documented fact that the population density is higher in the central city than in the suburbs. However, as it accounts for the basic trade-off between long/short commutes and low/high land rents, the fixed lot size assumption is widely used in models involving an urban economics building block. This assumption implies that the central city population (ℓ_0) and a suburban jurisdiction population (ℓ) are, respectively, given by

$$\ell_0 = mb \qquad \ell = B - b,$$

where $B \equiv L/m$ is assumed to be larger than b.

Regarding land ownership, two approaches are used in the urban economics literature (Fujita, 1989). In the former, landlords are absentee and the aggregate land rent does not feed back into the city economy. In the latter, land is collectively owned and consumers pay a land rent. In this paper, we follow the second approach because it allows us to prove the existence of a tax equilibrium, which is otherwise hard to show (Laussel and Le Breton, 1998). Without loss of generality, we assume that the opportunity cost of land is zero. Since land is collectively owned within each jurisdiction, it seems natural to use the aggregate land rent to finance local public goods (Arnott and Stiglitz, 1979). The Henry George Theorem cannot be applied here because the population size is given. Nevertheless, the approach we propose is in the same spirit. To be precise, a share $\theta_i \in [0, 1]$ of the aggregate land rent ($i = 0, 1, ..., m$), called a *land tax*, is used to pay for the provision of local public goods in jurisdiction i, while the remaining share $1-\theta_i$ is equally redistributed to the residents. Since our setup is symmetric, we may set $\theta_i = \theta$ for all $i = 1, ..., m$.[3]

The unit commuting cost $\tau > 0$ borne by consumers is the same in both the central city and the suburban jurisdictions, perhaps because transportation infrastructures are planned at the level of the entire MA. Therefore, commuting costs are equal τx or $\tau |x - x^s|$ according to the location of her employment center. In what follows, we consider three types of commuters: (i) a consumer lives and works in the central city; (ii) a consumer lives in a suburban service area but works in the central city; and (iii) a consumer resides and works in the same suburban service area. Hence, we focus on the case where $y \geq b$, which means that some workers commute from their suburban residences to the central city.[4] Because of cross-border commuting, land is used on both sides of the boundary between the central and suburban jurisdictions. As a consequence, there is not vacant land within the MA.

When a consumer lives and works in the central city, her indirect utility is given by

$$V_0(x) = w_0 - R_0(x) - \tau x + G + \frac{(1-\theta_0)\, ALR_0}{\ell_0}, \tag{2}$$

where $R_0(x)$ is the land rent at a distance x from the CBD and w_0 the wage paid by the firms located in the CBD, while

$$ALR_0 = m \textstyle\int_0^b R_0(x)\mathrm{d}x$$

is the aggregate land rent in the central city.

When a consumer lives in a suburban jurisdiction and works in the central city, her indirect utility becomes

$$V_0^s(x) = w_0 - R(x) - \tau x + G + \frac{(1-\theta)\, ALR}{\ell}, \tag{3}$$

where $R(x)$ is the land rent at a distance x from a SBD and the land tax, while

$$ALR = \textstyle\int_b^B R(x)\mathrm{d}x.$$

By contrast, when a consumer lives and works in a suburban jurisdiction, her indirect utility is

$$V^s(x) = w - R(x) - \tau |x - x^s| + G + \frac{(1-\theta)\, ALR}{\ell}, \tag{4}$$

where w is the wage rate paid in a SBD.

[3] Alternatively, the local governments may levy a property tax to finance local public services and redistribute the entire aggregate land rent to the residents. Since the lot size is exogenous, the property tax is totally capitalized in the land rent. As a consequence, the two setups lead to the same equilibrium business taxes (Gaigné et al., 2014).

[4] Since reverse commuting flows from the central city to the suburbs stand for a small fraction of commuters, the case $y < b$ is empirically less relevant than $y > b$ (Baum-Snow, 2010).

Land is allocated to the highest bidder. The land rent at each location in the central city is as follows. Given $V_0(x)$, the bid rent at x of a worker living and working in the central city ($\Phi_0(x)$) must solve $\partial V_0(x)/\partial x = 0$ or, equivalently, $\partial \Phi_0/\partial x + \tau = 0$ whose solution is

$$\Phi_0(x) = r_0 - \tau x, \tag{5}$$

where r_0 is a constant that will be determined in Section 3. The land rent R_0 in the central city is given by

$$R_0(x) = \max\left\{\Phi_0(x), 0\right\}, \tag{6}$$

while the land rent R^s in a suburban jurisdiction is such that

$$R^s(x) = \max\left\{\Phi_0^s(x), \Phi^s(x), 0\right\}, \tag{7}$$

where $\Phi_0^s(x)$ ($\Phi^s(x)$) is the bid rent at x of a worker living in a suburban jurisdiction and working in the central city (an edge city). Given $V_0^s(x)$ and $V^s(x)$, the equilibrium land rent is such as $\partial V_0^s(x)/\partial x = \partial V^s(x)/\partial x = 0$. As a consequence, the bid rents are such that

$$\Phi_0^s(x) = r_0^s - \tau x \qquad \Phi^s(x) = r^s - \tau\left|x - x^s\right|, \tag{8}$$

where both r_0^s and r^s will be determined in Section 3. Thus, in each jurisdiction, the intercept of the land rent profile is endogenous.

2.3 Firms and wages

Labor is the only production factor. When a worker stands alone, she produces one unit of an homogeneous good, which is chosen as the numéraire. We may then use indifferently the terms worker and firm, so that the total number of firms N established in the MA is given by $N = L$. The economic size of the central city is equal to my, while the economic size of an edge city is $B - y$. A SBD is then located at x^s with $x^s \in [y, B]$.

In the Introduction, we have seen that agglomeration economies may stem from different sources. In what follows, we choose not to privilege one type of economies against the others, the reason being that the results we will obtain could well depend on the particular type considered. Rather, we prefer to use a formulation capturing the main features that are typically associated with the presence of agglomeration economies.

Where do agglomeration economies pop up in an MA? Whereas Glaeser and Kahn (2004) argue that, due to the development of new information and communication technologies, the scope of spillovers has spread within the MA, Baum-Snow (2014) finds evidence that "spillovers across firms within central cities represent a large fraction of metropolitan area wide agglomeration economies." We assume that a firm benefits from agglomeration economies if it chooses to set up in the CBD or in a SBD, the reason being that firms need specific local public goods and private facilities to benefit from the spillovers generated by other firms. Since firms do not consume land, firms located in the central city choose to gather in the CBD, while those located in an edge city form a SBD.

All firms (or workers) benefit from agglomeration economies regardless of the center where they are located, but they do so with different levels of intensity according to the type of center in which they are established. Let n_0 be the number of firms set up in the CBD and n_i the number of firms in the ith SBD. Since we focus on a symmetric configuration, we may set $n_i = n$. The number of firms located in the central city being equal to the number of workers commuting to the CBD, we have $n_0 = my$ and, analogously, $n = B - y$.

When a firm sets up in the CBD, the production of a worker increases by the amount

$$E_0 = \alpha(n_0 + \lambda mn), \tag{9}$$

where $\alpha \geq 0$ measures the intensity of agglomeration economies and $\lambda \in (0, 1)$ that of spillovers between the CBD and a SBD. In other words, a CBD-worker produces $1 + E_0$ units of the numéraire, where E_0 rises with the number n_0 of firms located in the CBD. Such a firm also benefits from spillovers generated in the each SBD, the intensity of which depends on the number of firms located therein. Nevertheless, even at the Age of Internet information is subject to distance-decay effects, which we take into account by assuming $\lambda < 1$. As for the production of a worker in a SBD, it rises by the amount

$$E = \alpha\left(n + \lambda n_0\right), \tag{10}$$

which increases with the size n of the SBD.

As shown by (9) and (10), the central city benefits from spillovers brought about by all SBDs, whereas an edge city benefits from the spillovers generated by the CBD only. This reflects the comparative advantage of the CBD. Depending on where firms choose to set up, the sizes of the CBD and SBDs, as well as the intensity of spillovers within and between employment centers are endogenous. In accordance with the empirical evidence provided by Baum-Snow (2014), we assume that agglomeration economies in the central city are stronger than those in an edge city: $E_0 > E$ is equivalent to the following inequality:

$$\frac{n_0}{N - n_0} > \frac{\frac{1}{m} - \lambda}{1 - \lambda}.$$

This is satisfied for all values of n_0 if $\lambda > 1/m$, a condition which we assume to hold in the rest of the paper. Observe that this assumption is not restrictive as, on average, the number of jurisdictions in OECD metropolitan areas is close to 70.

In order to understand how agglomeration economies affect the production level in the MA, we find it illuminating to consider the production function of each type of center. The production function of the CBD is given by the sum of the quantities produced by the firms located therein:

$$F_0(n_0) = n_0(1 + E_0) = n_0 \left[1 + \alpha \left(n_0 + \lambda mn\right)\right].$$

This expression is increasing and convex in n_0. As for the production function of a SBD, it is similarly given by

$$F(n) = n(1 + E) = n \left[1 + \alpha(n + \lambda n_0)\right],$$

which is also increasing and convex in n. As a consequence, the marginal productivity of labor in the CBD or in a SBD increases in response to a hike in the size of the corresponding center.[5] In the presence of agglomeration economies ($\alpha > 0$), the morphology of the MA is thus the equilibrium outcome of interactions between employment centers, the productivity of which depends on the distribution of firms within the MA. In contrast, when $\alpha = 0$, the productivity of the MA has no impact on its structure as the productivity is the same regardless of the distribution of firms ($F_0 + mF = L$).

Jurisdictions compete in *business taxes* to attract firms and workers. To keep the formal analysis simple, we assume that this tax is lump-sum. Let Π_0 (Π) be the profits earned by a firm set up in the central city (an edge city). A firm located in the CBD earns profits equal to

$$\Pi_0 = 1 + E_0 - w_0 - T_0, \tag{11}$$

where $1 + E_0$ denotes the firm's revenue, w_0 the income earned by the corresponding worker and T_0 the business tax levied in the central city.

Because our setup is symmetric, all suburban jurisdictions charge the same business tax T and SBD-firms pay the same wage w. Thus, when a firm sets up in an edge city, its profit function is given by

$$\Pi = 1 + E - w - T. \tag{12}$$

In each employment center, the equilibrium wage is determined by a firm's revenue net of tax. Setting (11) and (12) equal to zero and solving, respectively, for w_0 and w, we get

$$w_0 = 1 + E_0 - T_0 \qquad w = 1 + E - T. \tag{13}$$

Thus, wages are endogenous, different between the CBD and the SBDs, and depend on the whole distribution of firms within the MA through the intensity of agglomeration economies, as well as on the levels of business taxes. Furthermore, if business taxes alleviate the land tax,

[5] Our description of agglomeration economies is not as rich as in Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002). However, we want to stress that these authors appeal to numerical methods to characterize the equilibrium outcome. In this paper, we need the explicit form of the equilibrium pattern of activities to study how tax competition affects the structure of the MA. Although simple, (9) and (10) capture the main features of the formation and diffusion of agglomeration economies studied by the above authors.

they also reduce the wages earned by workers. When consumers work and live within the same jurisdiction, both effects are washed out. This is no longer true, however, when they work and live in different jurisdictions. Indeed, both the land tax where the consumer lives and the business tax paid in the jurisdiction where she works affect her utility level, hence her residence and workplace.

3 The optimal metropolitan area

In this section, we assume that a benevolent planner maximizes the total welfare in the MA by choosing the size and number of areas supplying public services, as well as consumers' and firms' locations, hence the commuting pattern. To be precise, the planner must choose to centralize (or decentralize) the provision of publics goods with the aim of minimizing investment costs (or operating costs), as well as to concentrate firms and jobs in the CBD (disperse firms and jobs through the CBD and SBDs) with the aim of maximizing agglomeration economies (or minimizing commuting costs). Owing to symmetry, b, y, x^s, and B are the same along each spoke.

Individual utilities being linear, the total welfare W_T within the MA is defined, up to a positive constant, by the total surplus:

$$W_T = GDP - CC + LG - PC,$$

which involves two terms that depend on the marginal worker y.

(i) Using (9) and (10) shows that the gross domestic product (GDP) of the MA is given by

$$GDP = n_0(1 + E_0) + mn(1 + E).$$

Because of the firm population constraint $N = n_0 + mn$, when n_0 increases (decreases), a few firms are reallocated from the SBDs to the CBD (from the CBD to the SBDs). Since $n_0 = my$ and $n = B - y$, we may rewrite the production functions F_0 and F in terms of y only:

$$GDP = my\left[1 + \alpha\left(my + \lambda m(B - y)\right)\right] + m(B - y)\left[1 + \alpha\left(B - y + \lambda my\right)\right].$$

As expected, the metropolitan GDP increases with α and λ. Moreover, regardless the value of λ, the metropolitan GDP grows with y and reaches its maximum value when the MA is monocentric: $y = B$. Even though the SBDs are also the source of efficiency gains, the GDP of the MA is maximized when all jobs are concentrated at the CBD, that is, the MA has a unique and integrated labor market. This is because the GDP generated by the sole SBDs does not necessarily grow with the number of firms located therein. For instance, when λ is high enough, the GDP produced in the SBDs is maximized at $y = [B\alpha\left(m\lambda - 2\right) - 1]/2\alpha\left(m\lambda - 1\right)$, which is both positive and smaller than B.

(ii) Total commuting costs given by

$$CC = m\int_0^y \tau x \mathrm{d}x + m\int_y^B \tau \left| x - x^s \right| \mathrm{d}x = m\tau \left[\frac{y^2}{2} + \left(\frac{B-y}{2} \right)^2 \right].$$

Since the SBDs are optimally located at the middle point $x^s = (y+B)/2$, it is readily verified that CC is minimized when $y = B/3$, that is, the size of central city's labor market is equal to $mB/3$, while that of an edge city is $2B/3$. As a result, the population working in the central city is smaller than the total labor force in the edge-cities as a whole. This is sufficient to show the existence of a trade-off between the maximization of the metropolitan GDP and the minimization of commuting costs within the MA.

(iii) Last, the total cost of providing public goods G in the entire MA is given by

$$PC = (m+1)F + \frac{c}{2}(mb)^2 + m\frac{c}{2}(B-b)^2.$$

3.1 The optimal size of service areas and labor pools

Assume that the number of service areas is given $(m+1)$. By choosing the boundary b of the central city, the planner determines the population size in each service area. Clearly, a marginal expansion of the central city (higher b) reduces the number of residents in all suburban service areas. As a consequence, the cost of supplying public goods decreases therein, whereas it rises in the central city. Differentiating W_T with respect to b yields

$$\bar{b} = \frac{B}{m+1} < B.$$

Thus, regardless of the values of L and m *it is always optimal to decentralize the provision of public goods*. Since $L = mB$, the central city and the suburban service areas have the same population size. In other words, production costs in public goods are equalized across all service areas. Note also that, at the optimum, the total number of suburbanites exceeds the number of the central city residents.

Furthermore, by choosing y, the planner determines the size of the CBD (my), hence that of each SBD $(B-y)$. As mentioned above, total commuting costs reach their lowest value when the average traveled distance is minimized, i.e., $y = B/3 \geq \bar{b}$. Because y does not affect the production cost of public goods, the optimal economic boundary of the central city is the outcome of the trade-off between commuting costs and agglomeration economies. By implication, the optimal value of y must belong to the interval $[B/3, B]$.

When

$$\frac{\alpha}{\tau} > \frac{1}{2m(1-\lambda)}, \tag{14}$$

the function W_T is either increasing or convex over the interval $[\bar{b}, B]$ with $W_T(\bar{b}) < W_T(B)$. In this event, the optimal boundary is given by $\bar{y} = B$. Put differently, when the agglomeration economies are strong, commuting cost are low, or both, the social optimum involves the agglomeration of firms in the CBD. Thus, whereas the supply of public goods is decentralized within the MA, the labor market is integrated. To put it differently, the *decentralization* of the provision of public goods within the MA and the *agglomeration* of firms and jobs in the CBD do not necessarily conflict.

When (14) does not hold, differentiating W_T with respect to y shows that the optimal economic boundary of the central city is given by

$$\bar{y} = \frac{B\left[\tau + 4\alpha(\lambda m - 1)\right]}{3\tau - 4\alpha(m + 1 - 2\lambda m)}, \tag{15}$$

where $B/3 \leq \bar{y} < B$ because $m \geq 2$. It is readily verified that $\bar{y}$ increases in α and decreases in τ, thus confirming that agglomeration economies push toward the concentration of jobs in the CBD, whereas commuting costs act as a dispersion force that fosters the growth of the SBDs.

Proposition 1 comprises a summary.

Proposition 1. *Assume a planner maximizing total welfare within the metropolitan area. Then, the optimal metropolitan area involves several public goods providers, as well cross-border commuting from the suburban areas to the central city.*

Thus, cross-border commuting and the fragmentation of the MA into several public goods' providers need not be wasteful. Even in the absence of agglomeration economies ($\alpha = 0$), the position of the CBD in the transportation network makes cross-border commuting socially desirable when $m > 2$ because $\bar{y} = B/3 > \bar{b}$, while $\bar{y} = \bar{b}$ when $m = 2$.

3.2 The optimal number of service areas

The planner may also choose the degree of decentralization in the provision of public goods through the variable m. Since $L = mB$, choosing m amounts to choosing the spatial extent of the MA. To highlight the main forces at work, we assume that there are no agglomeration economies ($\alpha = 0$). Two cases may arise. In the first one, it is optimal to concentrate firms and jobs in the CBD. Differentiating W_T with respect to m at $\bar{b}$ and setting $\bar{y} = B$ leads to the following equilibrium condition:

$$\frac{\tau L^2}{2m^2} + \frac{c}{2}\left(\frac{L}{m+1}\right)^2 - F = 0. \tag{16}$$

Note that (16) includes the trade-off between the fixed cost of launching a new service area and the cost saved on the functioning of the existing service areas when a new one is added to the MA. Indeed, we have $\mathrm{d}\bar{m}/\mathrm{d}F < 0$ and $\mathrm{d}\bar{m}/\mathrm{d}c > 0$. It is also readily verified that

$\mathrm{d}\bar{m}/\mathrm{d}\tau > 0$ and $\mathrm{d}\bar{m}/\mathrm{d}L > 0$. Lower commuting costs, a less populated MA, or both lead to a smaller number of service areas. As a consequence, the optimal structure of the MA is governed by the trade-off between commuting costs and the cost of providing public goods. In particular, if F keeps on a higher value, the planner provides the public goods through a smaller number of service areas.

In the second case, it is optimal to break up the MA into several employment centers. The optimal value of m is now implicitly given by

$$\frac{\tau L^2}{6m^2} + \frac{c}{2}\left(\frac{L}{m+1}\right)^2 - F = 0 \ .$$

The impact of τ, L, F and c is the same as in the first case.

In the presence of agglomeration economies ($\alpha > 0$), (15) implies that more workers are allocated to the CBD, which increases the total commuting costs. As a consequence, the planner has an incentive to increase m to reduce these costs. However, raising m implies additional investments in transportation infrastructures, which gives rise to an additional trade-off. Therefore, apart from 5.2 where we consider the degree of fragmentation as a possible policy correcting the inefficiency of the decentralized outcome, the rest of our analysis treats m as exogenous.

4 Tax competition and the metro structure

We now consider a decentralized tax setting in which the institutional environment, i.e., the number of suburban jurisdictions, m, and the administrative boundary, b, between these jurisdictions and the central city are given. Our purpose is to find how the institutional parameters b and m, as well as the main economic parameters, such as α and τ, affect the tax policies and the location of firms and jobs.[6]

The spatial structure of the MA implies that competition among jurisdictions is strategic: each suburban jurisdiction competes directly with the central city only whereas the central city competes with every suburban jurisdiction. The interactions between local governments and market forces are described by a three-stage game that blends atomic and non-atomic players. There are three groups of players: a continuum of consumers, a continuum of firms, and $m+1$ local governments. Consumers choose where to live and where to work; firms choose where to locate and the wage to pay to their employees; and local governments choose a business tax and a land tax.

[6] Hoyt (1992) developed a setting in which the central city's government influences the land rent in suburban jurisdictions, but where the tax policy of the government of a suburban jurisdiction has no impact on the central city's land rent because its population share is negligible. Unlike us, he treated households' residential locations and workplaces as exogenous.

Once consumers are mobile, the specification of governments' objective is known to be a controversial issue (Cremer and Pestieau, 2004). We consider a sequential game that obviates this difficulty by assuming that consumers move first. Therefore, governments know who their residents are, and thus are able to determine the welfare function to maximize. Moreover, the relationship between jobs and people often has the nature of an egg-and-chicken problem. Here, firms choose their locations and consumers their workplaces simultaneously. Therefore, the timing of the game is as follows. In the first stage, consumers choose the jurisdiction they want to join and their location therein, anticipating the land tax, the land rent, and the wage they will earn. Since consumers are identical, in equilibrium they reach the same utility level. In the second stage, the population in all jurisdictions being given, local governments choose simultaneously and non-cooperatively a business tax and a land tax to maximize the welfare of their residents. Last, since business taxes are observable, firms choose their profit-maximizing locations and consumers their workplace, while land and labor markets clear. The size of the CBD and SBDs is determined when firms choose their locations in the third stage.[7]

We seek a subgame perfect Nash equilibrium. As usual, the game is solved by backward induction. In what follows, we focus on a symmetric equilibrium: $T_i = T$ and $\theta_i = \theta$ for $i = 1, ...m$. In this event, wages paid in the SBDs are the same: $w_i = w$ for $i = 1, ..., m$. Since there is no vacant land, we have $B = L/m$, while $w_i = w$ implies that $y_i = y$.

Because characterizing the equilibria of all subgames is long and tedious, we find it convenient to restrict ourselves to the equilibrium path. In particular, consumers being mobile and identical, they anticipate that they reach the same (indirect) utility level V^* at the end of the game. Thus, we have $V^* = V_0(x)$ for $0 \leq x \leq b$, $V^* = V_0^s(x)$ for $b < x \leq y$, and $V^* = V^s(x)$ for $y < x \leq B$. In what follows, we call the *economic boundary* of the central city the limit y of the area that includes all the individuals working in the CBD.

Since all tax rates are given, (2) and (13) imply that the (indirect) utility of a consumer residing in the central city ($0 \leq x \leq b$) is given by

$$V_0(x) = 1 + E_0 - T_0 - R_0(x) - \tau x + G + \frac{(1-\theta_0)\, ALR_0}{\ell_0}, \tag{17}$$

where $R_0(x)$ is obtained from (6). There are two groups of suburbanites living in a suburban jurisdiction: those who work in the CBD and pay the land rent R_0^s, and those who work in their SBD and pay the land rent R^s (see (7)). Using (3) and (13) shows that the utility of a

[7] Unlike Myers (1990), the workplaces and residences need not belong to the same jurisdiction in our setup, thereby inducing cross-border commuting. As a consequence, there is always a particular fiscal externality, that is, tax exporting. This explains why the decentralized outcome differs from the first best. Even if the local governments were to move first, as in Myers, the above fiscal externality would still be at work because the administrative and economic boundaries of jurisdictions differ. The presence of agglomeration economies exacerbates the tension between the equilibrium and optimum.

consumer belonging to the first group is

$$V_0^s(x) = 1 + E_0 - T_0 - R_0^s(x) - \tau x + G + \frac{(1-\theta)\,ALR}{\ell} \tag{18}$$

with $b < x \leq y$, while using (4) and (13) implies that the utility of a consumer belonging to the second group is

$$V^s(x) = 1 + E - T - R^s(x) - \tau\,|x - x^s| + G + \frac{(1-\theta)\,ALR}{\ell} \tag{19}$$

with $y < x \leq B$.

Since workers distribute themselves symmetrically around their employment centers, each SBD is located at the center of the interval $[y, B]$:[8]

$$x^s = y + \frac{B-y}{2}. \tag{20}$$

4.1 Labor and land market equilibrium

In the third stage, firms and consumers observe the tax rates chosen by the local governments. Then, firms select a location as well as the wage they pay while consumers choose their working places. Because consumers are mobile, they accurately anticipate that the equilibrium land rent equalizes individual utilities.

4.1.1 Job location

Wages being given by (13), it remains to determine the distribution of jobs within the MA. For this, we must find the location y of the marginal worker, which is the same along all spokes. We assume throughout this section that y exceeds b and will determine the conditions for this to hold in equilibrium.

Using (20) and (13), y is the location of the marginal worker if and only if $V_0^s(y) = V^s(y)$ or, equivalently,

$$w_0(y) - w(y) = \tau y - \tau(x^s - y) = \tau\frac{3y - B}{2}. \tag{21}$$

In other words, *CBD- and SBD-workers do not earn the same wage*, the difference between wages compensating workers for the difference between commuting costs along the same spoke. Moreover, given T_0 and T, a higher value of y strengthens the agglomeration economies arising in the CBD and weakens those in the SBDs. As a consequence, when more firms locate in the

[8] If edge cities choose non-cooperatively the location of their SBDs, they choose x^s such that the distribution of workers around x^s is symmetric. The details of calculations are available from the authors upon request. Note that the planner makes the same choice. As a consequence, the location of the SBDs generates no additional distortion.

central city, the CBD-wage and the wage gap between the CBD and the SBDs rise. This is because the central city's firms must pay a higher wage to attract more distant workers. In sum, stronger agglomeration economies allows the CBD-firms to pay a higher wage that compensates the CBD-workers for their higher commuting costs. The reverse holds in the SBD.

Substituting (13) into (21) yields the equilibrium economic boundary of the central city:

$$y(T_0, T) = \frac{B[\tau + 2\alpha(\lambda m - 1)] - 2(T_0 - T)}{3\tau - 2\alpha(m + 1 - 2\lambda m)}. \tag{22}$$

It follows from this expression that the economic boundary of the central city expands (shrinks) with T (T_0) because the CBD becomes relatively more (less) attractive. For given T_0 and T, stronger agglomeration economies, lower commuting costs, or both yield a bigger CBD. Similarly, y increases with m as long as $b < y < B$. Therefore, a more fragmented MA has smaller SBDs.

Since there are agglomeration economies, we must focus on equilibria that are stable. Observe that $V^s(y)$ is upward sloping. If

$$\frac{\alpha}{\tau} < \frac{1}{m(1-\lambda)}, \tag{23}$$

$V_0^s(y)$ is downward sloping, and thus the equilibrium of the third stage subgame is unique and stable. As expected, a polycentric MA emerges when agglomeration economies are not too strong, commuting costs are not too low, or both. Note also that the denominator of $y(T_0, T)$ is positive when (23) holds.

4.1.2 Land rent

We now turn to the determination of the equilibrium land rents. Using the equilibrium conditions $V_0 = V_0^s = V^s = V$ together with (5), (6), (7) and (8), we are equipped to determine the value of r_0 for the central city as well as the values of r_0^s and r^s for the suburban jurisdictions.

At $x = B$, the land rent equals the opportunity cost of land, which is zero. At $x = y$, the land rent must be equal to 0 for $V(y) = V(B)$ to hold. Indeed, if a consumer offers a positive bid to reside at y, her utility is given by $V(y) < V(B)$. These results in turn imply that $r^s = \tau(B - y)/2$, which yields the bid rent

$$\Phi^s(x) = \frac{\tau(B - y)}{2} - \tau|x - x^s|.$$

Since $R^s(y^*) = 0$, repeating the above argument leads to $r_0^s = \tau y$ and thus

$$\Phi_0^s(x) = \tau(y - x).$$

Using (7) and the above two expressions for the bid rents, we get the aggregate land rent in any edge city:

$$ALR = \tau\left[\frac{(y - b)^2}{2} + \frac{(B - y)^2}{4}\right]. \tag{24}$$

The equilibrium condition $V_0(b) = V_0^s(b)$ and (5) imply that

$$r_0 = \left[\tau\left(y-b\right) + \frac{\left(1-\theta_0\right)ALR_0}{\ell_0} - \frac{\left(1-\theta\right)ALR}{\ell}\right], \tag{25}$$

which shows that the central city land rent capitalizes the difference in commuting costs between the marginal worker and the border worker located at b, as well as the difference in the aggregate land rent redistributed to the residents between the central city and the edge cities. Using (6), (5), and (25), we obtain the aggregate land rent in the central city:

$$ALR_0 = m\int_0^b R_0(x)\mathrm{d}x = \frac{\ell_0}{\theta_0}\left[\tau\left(y-\frac{b}{2}\right) - \frac{\left(1-\theta\right)ALR}{\ell}\right], \tag{26}$$

which decreases with the land tax θ_0. The above expressions show that the land rent profile varies with the location of the economic boundary of the central city.

The land rent profile across the MA is thus as follows. As the distance to the CBD rises, the land decreases with a slope equal to $-\tau$ over the interval $[0, y]$. At the border $b < y$ of the central city, the land rent is discontinuous because consumers who are just inside and outside that boundary live in jurisdictions where the per capita cost of public services and the per capita aggregate land rent differ ($R_0(b) \neq R^s(b)$). Beyond y, the land rent rises with a slope equal to τ over $[y, x^s]$ and decreases with a slope $-\tau$ over $[x^s, B]$.

Because some workers move from the suburban jurisdictions to the CBD, the tax policy of the suburban jurisdictions generates a tax externality capitalized in the land rent paid in the central city. Indeed, (5) and (26) imply that $R_0\left(x\right)$ rises with θ. Note also that our framework allows determining not only the costs and benefits that are capitalized, but also *where* the capitalization arises.

4.2 Tax competition between the central and suburban jurisdictions

The business tax and the land tax allow each local government to finance the local public good provided to its residents. Hence, the budget constraint of jurisdiction $i = 0, 1, ..., m$ is given by

$$F + c\frac{\ell_i^2}{2} = T_i n_i + \theta_i ALR_i,$$

One appealing feature of our tax game is that we may determine the business tax rates independently of the land tax.

4.2.1 Business tax

Local governments set non-cooperatively their business tax rates or subsidies with the aim of maximizing the welfare of their residents. Specifically, the central city maximizes W_0 with respect to $T_0 \in \mathbb{R}$, while every suburban jurisdiction maximizes W with respect to $T \in \mathbb{R}$. Since firms choose their locations in the third stage, governments anticipate the consequences of their choices on the size of their business districts.

The social welfare functions of the local governments. (i) The total welfare in the central city is given by

$$W_0 = m\int_0^b V(x)\mathrm{d}x = \ell_0(1 + E_0 - T_0 + G) - \theta_0 ALR_0 - \tau\frac{mb^2}{2},$$

where we have used (17). Substituting the budget constraint $F + c\ell_0^2/2 = T_0 n_0 + \theta_0 ALR_0$ and the labor market balance condition $n_0 = \ell_0 + m(y - b)$ into the above expression, we obtain

$$W_0 = \ell_0 \left(1 + E_0 + G - T_0\right) + T_0 my - F - c\frac{\ell_0^2}{2} - \frac{\tau}{2}\frac{\ell_0^2}{m}, \tag{27}$$

where y is given by (22).

Raising the business tax T_0 gives rise to two opposing effects. First, a higher business tax tends to generate higher tax revenue, but it yields a smaller fiscal basis. Moreover, $y > b$ implies that there is cross-border commuting toward the central city. Therefore, because a fraction of the CBD-workers lives outside the central city, there is *tax exporting* and its extent is equal to $T_0 my - T_0\ell_0 = T_0 m(y - b)$.

Second, a higher T_0 reduces the wage paid to the CBD-workers (see (13)). In addition to this direct effect, (22) shows that fewer individuals work in the CBD, which further lowers the net wage paid to the CBD-workers through weaker CBD-agglomeration economies (lower E_0). The equilibrium business tax in the central city is the outcome to this trade-off.

Note that a marginal increase in T_0 has no impact on the commuting costs within its jurisdiction because all the residents work in the CBD. However, by reducing the number of CBD firms, increasing T_0 affects the commuting costs paid by the suburban consumers, an effect not internalized by the central city government.

Maximizing W_0 with respect to T_0 yields:

$$\frac{\mathrm{d}W_0}{\mathrm{d}T_0} = m\left[y - b + T_0\frac{\partial y}{\partial T_0} + \alpha mb(1 - \lambda)\frac{\partial y}{\partial T_0}\right] = 0, \tag{28}$$

where $\mathrm{d}^2W_0/\mathrm{d}T_0^2 < 0$ holds for any polycentric configuration. The first three terms between brackets capture the tax exporting effect described above while the last term pins down the negative impact of the business taxation on the CBD-agglomeration economies capitalized into the wage of the central city's residents. As expected, the capitalization effect decreases in presence of stronger spillovers (higher λ) because wages become less dependent on the spatial distribution of firms. As noticed above, this is at the origin of a new externality because the central city's government does not internalize the impact of their tax policy on the wage earned by the suburbanites working in the CBD.

(ii) Recall that an edge city has two types of residents, those who work in the CBD and those who work in their own SBD. Using (18), (19) and the budget constraint $F + c\ell^2/2 = Tn + \theta ALR$ where $n = B - y$, the total welfare in an edge city is given by

$$\begin{aligned} W &= (y-b)(1+E_0+G-T_0) + (B-y)(1+E+G-T) + T(B-y) \\ & \quad - \left(\int_b^y \tau x \mathrm{d}x + \int_y^B \tau \left| x - x^s \right| \mathrm{d}x \right) - F - c\frac{\ell^2}{2}, \end{aligned} \tag{29}$$

while the total commuting costs borne by the edge city's residents are given by

$$\int_b^y \tau x \mathrm{d}x + \int_y^B \tau \left| x - x^s \right| \mathrm{d}x = \tau \left[\frac{y^2 - b^2}{2} + \frac{(B-y)^2}{4} \right].$$

The total welfare in an edge city can then be rewritten as follows:

$$W = (y-b)(E_0 - T_0) + (B-y)E - \tau \left(\frac{3y^2}{4} - \frac{By}{2} \right) + \text{constant}. \tag{30}$$

Business taxes affect the distribution of firms within the MA. Therefore, they also affect the level of the agglomeration economies in both the CBD and the SBDs, and thus the welfare of the two groups of workers residing in the suburban jurisdiction. In a polycentric MA, it is readily verified that the increased agglomeration in the CBD that follows a hike in T boosts the productivity in the CBD ($\mathrm{d}E_0/\mathrm{d}T > 0$) while being damaging to the SBD ($\mathrm{d}E/\mathrm{d}T < 0$). In addition, a higher business tax lowers the share of jobs in the SBD, and thus increases commuting costs borne by the edge city's residents. Unlike the central city government, a suburban government cares about the trade-off between commuting costs and agglomeration economies, but it does so within its own jurisdiction only.

We consider a symmetric outcome, but focus on the choice made by a single suburban jurisdiction. Differentiating W with respect to T yields the equilibrium condition:

$$\frac{\mathrm{d}W}{\mathrm{d}T} = -T\frac{\mathrm{d}y}{\mathrm{d}T} + (y-b)\frac{\mathrm{d}E_0}{\mathrm{d}y}\frac{\partial y}{\partial T} + (B-y)\frac{\mathrm{d}E}{\mathrm{d}y}\frac{\partial y}{\partial T} = 0, \tag{31}$$

where we have used (21). Moreover, $\mathrm{d}^2W/\mathrm{d}T^2 < 0$ for all $\alpha/\tau < 1/m\,(1-\lambda)$.

To disentangle the various effects at work, we first discuss the case in which there are no agglomeration economies ($\alpha = 0$). We then turn our attention to the more relevant case of positive agglomeration economies.

The case of no agglomeration economies. When $\alpha = 0$, we have $\mathrm{d}E_0/\mathrm{d}y = \mathrm{d}E/\mathrm{d}y = 0$. It then follows from (31) that $T^* = 0$. Plugging this value into (28) and using (22), we obtain the equilibrium business tax set in the central city:

$$T_0^* = \frac{\tau}{4}(B - 3b) \quad \text{and} \quad T^* = 0. \tag{32}$$

Thus, the suburban governments neither tax nor subsidize firms at the equilibrium, while the business tax set by the central city is positive as long as B exceeds $3b$. At the tax rates (32), the economic boundary of the central city is given by

$$y^* = \frac{B}{6} + \frac{b}{2}. \tag{33}$$

In equilibrium, the MA is always polycentric ($y^* < B$) because $b < B$. As for $y^* > b$, it holds if and only if

$$b < \frac{B}{3}, \tag{34}$$

that is, the population of the central city (mb) is less than one third of the total population (L) of the MA, a condition which holds in many MAs. In this case, we have $T_0^* > T^* = 0$. Thus, even in the absence of agglomeration economies, the central city's government sets a business tax rate higher than the edge cities. This is due to the central position of the CBD in the transportation network. The intuition behind this finding is easy to grasp: a locale with a comparative advantage can set a higher business tax rate because more firms want to locate there. The following proposition summarizes this result:

Proposition 2. *Assume a polycentric metropolitan area and no agglomeration economies. The central city sets a higher business tax than the edge cities.*

When $b < B/3$, there is tax exporting. Observe that T_0^* decreases with b. When the limit of the central city expands, (22) implies that y^* remains the same as long as T_0 and T are given. Therefore, the extent of tax exporting $y^* - b$ shrinks. This incentivizes the central city government to lower its tax rate to restore its fiscal base.[9]

Finally, as in Hoyt (1991), but in a different context, the equilibrium tax rate T_0^* decreases with the number of suburban jurisdictions. Indeed, reducing the number of suburban jurisdictions softens tax competition and allows the central city to set a higher business tax. Similarly, lowering commuting costs leads to a smaller business tax rate in the central city because the location of jobs is more sensitive to a change in the spatial difference in business tax rates for low commuting costs. Put differently, the elasticity of the tax base with respect to the tax rate increases when τ falls.

The case of positive agglomeration economies. Assuming $\alpha > 0$ renders the analysis more involved. Using (22) and solving (28) with respect to T_0 yields the best-reply function of the central city's government:

$$T_0^*(T) = \frac{T}{2} + \frac{\alpha}{2}(\lambda m - 1)(B - b) + \frac{\tau}{4}(B - 3b), \tag{35}$$

[9] When b exceeds $B/3$, (27) and (30) no longer describe the social welfare functions of the local governments, the reason being that $b > y^*$. In this case, it can be shown that all jurisdictions set a positive business tax. Since the suburban jurisdictions now benefit from tax exporting, they levy a higher tax than the central city.

which is linear and upward sloping. In this expression, everything being equal the second term of the right-hand side shows how the central city's government capitalizes on the agglomeration economies to charge a higher tax rate when α rises. However, as will be seen below, T^* depends on α in a non-monotone way, thus suggesting that the global impact of α on T_0^* is a priori ambiguous.

Similarly, solving (31) with respect to T, we obtain the best reply of a suburban government:

$$T^*(T_0) = -\frac{4\alpha}{D}T_0 + \frac{\alpha}{D}\left[2\alpha m(B - (2\lambda - 1)b) - 2\alpha(B - b) - \tau(B + 3b)\right], \tag{36}$$

where we have used (9), (10) and (22), while

$$D \equiv \frac{3\tau - 2\alpha(m + 3 - 2\lambda m - 2\lambda)}{1 - \lambda}$$

is positive for all $\alpha/\tau < 1/m\,(1 - \lambda)$ (see (23)). Therefore, the best reply function of a suburban government is linear and downward sloping. When $\alpha = 0$, the foregoing expressions boil down to (32).

Since the slopes of the best reply curves have opposite signs, the two curves intersect in the plane, and there always exists a unique tax equilibrium. Given that the central city's government sets a positive tax when $\alpha = 0$, it is reasonable to expect T_0^* to be positive when there are agglomeration economies. We show in the Appendix that T_0^* is positive for all $\alpha \geq 0$. Thus, regardless of the intensity of agglomeration economies, *the central city's government always sets a positive business tax.*

Last, as shown by (36), the intensity of the agglomeration economies has an impact on the extent to which the tax policy in the suburbs reacts to the central city's tax policy. More precisely, stronger agglomeration economies makes T^* more sensitive to T_0^*. As a consequence, the presence of agglomeration economies erodes the fiscal autonomy of the suburban jurisdictions.

Regarding T^*, observe that (31) may be rewritten as follows:

$$T^* = \alpha(1 - \lambda)\left[(y^* - b) - (B - y^*)\right]. \tag{37}$$

Therefore, the sign of T^* depends on the relative numbers of individuals working in the CBD. If $y^* - b > B - y^*$, i.e., a majority of the suburban residents work in the CBD, then $T^* > 0$. This is because the overall welfare of the suburban residents is less dependent on the attractiveness of their employment centers. As consequence, the suburban governments have less incentives to grant subsidies to firms while the positive business CBD-tax rate allows the cross-border commuters to enjoy a higher wage through stronger agglomeration economies in the CBD. In contrast, if the majority of suburban residents work in their SBD, the suburban governments subsidize firms to boost the productivity of their employment centers, which in turn increases the wage of the residents working therein. In sum, the sign of T^* depends on where the median resident of a suburban jurisdiction works.

Symmetrizing (31) and plugging (28)-(31) into (22) implies

$$y^* = \frac{1}{2}\frac{\tau\,(B+3b) + 2\alpha[\lambda(m+1)-2](B+b)}{3\tau - 2\alpha(m+2+\lambda-2\lambda m)}, \tag{38}$$

which is equal to (33) when $\alpha = 0$. Since y^* increases with α and $y^*(0) > b$ when $b < B/3$, it must be that $y^*(\alpha) > b$, that is, there is cross-border commuting from the suburban jurisdictions to the central city at the tax equilibrium. Furthermore, it is readily verified that y^* always decreases with τ. In sum, *at the equilibrium tax rates, stronger agglomeration economies, lower commuting costs, or both yield a bigger CBD*, which is in accordance with what we saw in Section 3 about the optimum.

We first check when the MA is polycentric ($y^* < B$) at the tax equilibrium. This is so if and only if

$$\frac{\alpha}{\tau} < \frac{1}{2}\frac{5B-3b}{2[B(m+1)-b] - \lambda[2mB+(m+1)(B-b)]},$$

which is positive if

$$\lambda < \frac{2(mB+B-b)}{2mB+(m+1)(B-b)} < 1.$$

As suggested by the foregoing discussion, jobs are decentralized at the tax competition outcome if the intensity of agglomeration economies is not too high, commuting costs are not too low, or both.

We now discuss the equilibrium business taxes. In presence of agglomeration economies, both T_0^* and T^* depend on the central city size. How the tax policy of the suburban jurisdictions is affected by b is a priori ambiguous. Indeed, T^* depends on the relative numbers of suburbanites working in the CBD, which varies with b. Since y^* first increases with b, the number of suburbanites working in the SBD ($B - y^*$) fall when the limit of the central city expands. On the other hand, the effect on the number of suburbanites working in the CBD ($y^* - b$) is a priori unclear because it depends on the intensity of the agglomeration economies. Using (38), it is straightforward to show that

$$\frac{\partial\,(y^*-b)}{\partial b} \gtrless 0 \quad \text{when} \quad \alpha \gtrless \frac{3\tau}{2\,(2m-\lambda-3m\lambda+2)}.$$

Therefore, *when the intensity of agglomeration economies is strong enough, expanding the limit of the central city yields an even bigger expansion of its economic limit.* In this case, T^* unambiguously increases with b. Otherwise, the numbers of the two types of suburbanite workers shrink, so that the net effect of b on T^* is a priori ambiguous. Nevertheless, using the expression given in the Appendix for T^*, we obtain

$$\frac{\partial T^*}{\partial b} > 0.$$

Hence, a bigger central city always leads the suburban jurisdictions to set a higher business tax.

As for the tax differential, it is given by

$$T_0^* - T^* = -\frac{T^*}{2} + \frac{\alpha}{2}(\lambda m - 1)(B - b) + \frac{\tau}{4}(B - 3b), \tag{39}$$

which is positive for all α and τ if the right-hand side of this expression is positive when T^* takes on its highest value. It follows from (37) that the largest value of T^* is reached when $y^* = B$, so that $T^*_{\max} \equiv \alpha(1 - \lambda)(B - b)$. Plugging $T^*_{\max}$ into (39) shows that $T_0^* > T^*$ if $\lambda > 2/(m + 1)$ and $B > 3b$. The former inequality holds when the number of suburban jurisdictions is large, while the latter is (34).

The next proposition comprises a summary.

Proposition 3. *Assume a polycentric metropolitan area. Then, the central city's government always sets a positive business tax, while the suburban governments tax (subsidize) firms if the median resident works in the CBD (SBD). Furthermore, the business tax set in the central city exceeds that set in the edge cities if* $\lambda > 2/(m + 1)$.

4.2.2 Land tax

It remains to determine the equilibrium land taxes. Once T^* is determined, the equilibrium land tax of a suburban jurisdiction must be such that

$$F + c\frac{\ell^2}{2} = T^* \cdot (B - y^*) + \theta^* \cdot ALR.$$

Using this expression and (24), we obtain

$$\theta^* = \frac{F + c\frac{\ell^2}{2} - T^*(B - y^*)}{\tau}\left[\frac{(y^* - b)^2}{2} + \frac{(B - y^*)^2}{4}\right]^{-1}.$$

This expression shows how the land tax set in the suburban jurisdictions depends on firms' locations. Specifically, the value of θ^* is affected by a higher α through three channels: (i) negatively through the aggregate land rent, (ii) positively through the business tax basis, and (iii) positively or negatively through the level of business tax. For this reason, it is hard to make any prediction about the impact of agglomeration economies on θ^*.

The suburban land tax may increase or decrease with the central city size. Indeed, whereas θ^* may increase with b because the business tax basis of a suburban jurisdiction shrinks, the opposite effect may arise because, the suburban jurisdiction becoming less populated, the public services are less costly to provide. Obviously, the latter effect dominates the former when c is high enough. Moreover, as demonstrated above, the number of cross-commuters from the suburbs $(y^* - b)$ and, then, the aggregate land rent may increase or decrease with b depending on the value of α.

As for the central city, its budget constraint implies that the equilibrium land tax satisfies the relationship:

$$\theta_0^* = \frac{F + c\ell_0^2/2 - T_0^* m y^*}{ALR_0},$$

where ALR_0 depends on θ_0^*. There is no need to solve this equation, however. Indeed, once $T_0^* m y^*$is determined, the value of $\theta_0 ALR_0$ is constant regardless of the value of θ_0 (see (26)). Put differently, once the business tax is chosen, the budget constraint is satisfied through the adjustment of the land tax base ALR_0. In other words, there is a continuum of land tax equilibria but the revenue is the same.

5 On the efficiency of the decentralized outcome

Being given by a Nash equilibrium, the tax competition outcome is inefficient. However, we do not know how this inefficiency manifests itself and how to correct it.

5.1 Is the CBD too large or too small?

Consider first the case of no agglomeration economies. Comparing (15) and (38) when $\alpha = 0$ yields

$$y^* - \bar{y} = -\frac{B - 3b}{6} < 0.$$

Thus, *the inefficiency of tax competition takes the form of too small a CBD.* In other words, the business tax rate set in the central city is too high when compared with the rate set in the edge cities. Any common business tax rate $T_0 = T = \hat{T}$ set by *all* local governments would imply that $y^* = \bar{y}$. In this case, the central city would benefit from a high business rate because the tax exporting effect $\hat{T} m(\bar{y} - b)$ increases linearly with $\hat{T}$. In this case, it follows from (27) that consumers in the central city are better-off when $\hat{T}$ is larger. Conversely, because (29) decreases with $\hat{T}$, residents in the suburban jurisdictions are worse-off. As a consequence, even in the absence of agglomeration economies, *the central city and the edge cities have opposing interests when they strive to choose a common tax policy.*

In the presence of agglomeration economies, tax harmonization ceases to be optimal. To show it, assume that all jurisdictions set the common tax rate $\hat{T}$. The economic boundary of the central city is then given by

$$y^*(\hat{T}) = \frac{B[\tau + 2\alpha(\lambda m - 1)]}{3\tau - 2\alpha(m + 1 - 2\lambda m)}.$$

Since $y^*(\hat{T}) < \bar{y}$, tax harmonization yields a spatial distribution of firms that remains inefficiently dispersed when $\alpha > 0$. In other words, the equilibrium size of the CBD remains smaller than its optimal size, whereas the SBDs are too large.

For $y^*(T_0, T) = \bar{y}$ to hold, the tax differential must be equal to

$$\Delta^O = \frac{\alpha\tau B\,(m + \lambda m - 2)}{4\alpha[1 + m(1 - 2\lambda)] - 3\tau}. \tag{40}$$

It is readily verified that $\Delta^O < 0$ for all $\bar{y} < B$. Hence, in a polycentric MA, for the distribution of jobs to be socially optimal the business tax rate in the suburban jurisdictions must exceed the rate set in the central city. Recall that tax competition leads to the opposite outcome.

To summarize, we have:

Proposition 4. *Assume a polycentric metropolitan area. Then, business tax competition leads to an insufficient concentration of jobs in the CBD.*

Thus, tax competition between jurisdictions makes the MA less productive by preventing a better exploitation of agglomeration economies through a higher concentration of firms and jobs. Furthermore, besides the laboriousness of gathering the information needed to compute Δ^O, the central city and the suburban jurisdictions still have conflicting interests in choosing specific tax rates. This points once again to the difficulty of making the MA work when this one is fragmented in competing jurisdictions.

To illustrate the nature of the difficulty, let us consider the following argument. Differentiating $W_0 + mW$ with respect to T_0, using the first-order condition (31) and simplifying yields the following expression:

$$m\frac{\mathrm{d}W}{\mathrm{d}T_0}\bigg|_{T_0=T_0^*,T_i=T^*} = \alpha m\,(m-1)\,[(y^*-b)\,(1-\lambda) + (B-y^*)\,\lambda]\,\frac{\mathrm{d}y^*}{\mathrm{d}T_0} - m(y^*-b) < 0.$$

Since W_0 is concave, the above expression implies that, given T^*, the equilibrium tax T_0^* exceeds the tax rate that maximizes $W_0 + mW$. Similarly, it follows from (27) that the derivative of $W_0 + mW$ with respect to T_i is equal to $m\,(y^* - b) > 0$. Given T_0^*, this implies that T^* is lower than the tax rate maximizing total welfare. Since a fraction of the suburbanites work in the central city, suburban governments internalize the impact of their tax policy on the agglomeration economies at work in the CBD.

However, there exist instruments other than business tax coordination to correct the sub-optimal distribution of jobs within the MA. One natural instrument is to subsidize the parking fees paid by the cross-border commuters. To illustrate, consider a pricing system that takes the form of a fixed subsidy s paid to every suburbanite working in the CBD. The value of the subsidy at which the equilibrium and optimal economic sizes of the central city coincide is given by:

$$s^O = (T_0^* - T^*) - \Delta^O.$$

Since (40) implies $\Delta^O < 0$ while $T_0^* > T^*$ holds under mild conditions, we may conclude that s^O is positive. Moreover, given that $T_0^* - T^*$ decreases with b by (39), the optimal subsidy is lower for bigger central cities.

It should be clear that subsidizing cross-border commuters may serve to boost the productivity of the MA by fostering a better exploitation of agglomeration economies (Borck and Wrede, 2009). However, in a politically fragmented MA, it is not clear how such subsidies may be financed. What is more, standard road pricing techniques have the nature of a tax, not that of a subsidy. Indeed, road pricing policies aim to reduce the discrepancy between the social and the private costs of commuters' trips, the former being typically higher than the latter because of congestion or excessive emissions, a problem not addressed here. Our model thus highlights the difficulty to find an optimal road pricing system, as improving the overall productivity of the MA through a stronger concentration of firms and jobs clashes with the objective of reducing congestion and the emission of greenhouse gases in the central city. That the application of such instruments is currently limited to a handful of cities strongly suggests that their political acceptability is pretty low. However, the recommendation of a bigger CBD need not conflict with the objective of lower emissions within the MA. Indeed, when the central city provides a denser and a more energy-efficient public transport network than the SBDs, increasing the attractiveness of the CBD may be justified both for economic and environmental reasons. In addition, car use typically increases at the expense of public transportation when SBDs grow.

5.2 Does redrawing the central city limit remedy the misallocation of jobs?

Since tax coordination is hard to implement and since the central city's population size influences the extent of the distortion created by tax competition within the metro, it is natural to ask whether changing the central city limit prior to the tax game is a desirable strategy. We have seen that $T_0^* - T^*$ decreases with b, whereas $\bar{y}$ given by (15) is independent of b. As a consequence, productive efficiency loss generated by the misallocation of jobs decreases when the central city population rises ($\mathrm{d}(\bar{y} - y^*)/\mathrm{d}b < 0$).

The following proposition is a summary.

Proposition 5. *Assume a polycentric metropolitan area. Then, the productive efficiency loss decreases when the relative size of the central city population rises.*

This proposition suggests the following question: how can the planner choose the central city limit? The boundary b at which the misallocation of jobs vanishes under tax competition is given by the solution $\tilde{b}$ to the equation $y^*(b) = \bar{y}$. (i) When $\alpha = 0$, it is readily verified that the solution to $y^*(b) = \bar{y}$ is equal to $\tilde{b} = B/3 = \bar{y}$. In this case, there is no cross-border commuting, hence no tax exporting. However, in doing so, the planner does not choose the optimal boundary because $\tilde{b} = \bar{y} > \bar{b}$. (ii) When α becomes positive, the solution to $y^*(b) = \bar{y}$ is given by $\tilde{b}(\alpha)$ which increases with α. As a result, we have $\tilde{b}(\alpha) > \tilde{b}(0) > \bar{b}$ because $\bar{b}$ does not depend on α. Thus, for all $\alpha \geq 0$ the central city size exceeds its optimal size, while the suburban jurisdictions are too small. Conversely, choosing $\bar{b}$ for the limit of the central city leads to an insufficient concentration of jobs in the CBD because $y^*(\bar{b}) < \bar{y}$.

In sum, under business tax competition, we have:

Proposition 6. *Assume a polycentric metropolitan area. When the population size of the central city is optimal the CBD is too small, whereas the central city is too large when the size of the CBD is optimal.*

The above proposition suggests a second-best approach in which the planner chooses the central city limit that maximizes the total welfare within the MA. In other words, the planner first chooses the welfare-maximizing value of b, and then lets consumers, firms and local governments to pursue their own interest. In what follows, we focus on the case where $\alpha = 0$.

The central city limit maximizing total welfare when T_0 and T are given by the corresponding equilibrium tax rates is given by

$$b^* = \frac{B\,(8c + \tau)}{3\tau + 8c\,(m+1)}$$

where $y^*\,(b^*) > b^* > \bar{b}$ for $m \geq 2$ and $B > b^*$ if $B > y^*$. Hence, as in the first-best solution, the second-best outcome under tax competition involves political fragmentation and a decentralized supply of public services. However, unlike the first-best solution, suburban jurisdictions are now smaller than the central city. Because tax competition favors the edge cities at the expense of the central city, the second-best approach aims to reduce this distortion by launching a bigger central city. To sum up, we present the next proposition.

Proposition 7. *If the planner chooses the total welfare-maximizing limit of the central city, the second-best central city is bigger than the first-best central city. Furthermore, the second-best metropolitan area involves several public goods providers, as well cross-border commuting from the suburban areas to the central city.*

As discussed in 3.2, the planner may also determine the degree m of fragmentation of the MA maximizing the total welfare $W_T^s = W_0 + mW$, which is implicitly given by the equilibrium condition:

$$\frac{\partial W_T^s}{\partial m} = -F + \frac{3\tau}{16}\left[\left(\frac{L}{m}\right)^2 - b^2\right] + \frac{c}{2}\left[\left(\frac{L}{m}\right)^2 - b^2\,(2m+1)\right] = 0 \qquad (41)$$

with $\partial^2 W_T^s / \partial m^2 < 0$. As in the first-best analysis, for a given administrative boundary b, the second-best number of suburban jurisdictions increases with the population size and the cost parameter capturing the crowding effect of public services, whereas it decreases with investment outlays.

Tax competition incentivizes the planner to establish a more fragmented MA than in the first-best configuration. Indeed, the optimality condition (41) evaluated at $b = \bar{b}$ and $m = \overline{m}$

implies that

$$\left.\frac{\partial W_T^s}{\partial m}\right|_{m=\overline{m},b=\bar{b}} = \frac{3\tau}{16}\left[\left(\frac{L}{3\overline{m}}\right)^2 - \bar{b}^2\right],$$

which is positive for all $\overline{m} \geq 2$. Therefore, the planner raises the number of edge cities, that is, reduces their population size, to alleviate the efficiency loss associated with the insufficient concentration of firms in the CBD.

Last, when $\alpha = 0$, it is straightforward to show that a larger number of suburban jurisdictions leads the central city government to decrease its tax rate in higher proportion than the authorities of the edge cities, thus reducing the misallocation of jobs. In this case, the central city competes with more rivals. Therefore, like in standard oligopoly theory, a larger number of suburban jurisdictions exacerbates tax competition and reduces the misallocation of jobs.

6 The consumption of public goods

Thus far, we have neglected (i) the possibility for the suburbanites working in the CBD to consume the public services provided in the central city and (ii) the fact that the central city provides a broader range of public services than the suburban jurisdictions. In this section, we discuss what our main findings become in each of these two cases.

6.1 Public good spillovers

That suburbanites consume the public services supplied in the central city is known to be a major source of distortion in the allocation of public resources within a MA. For this reason, we assume here that suburbanites working in the CBD benefit from both the public services provided in their own jurisdiction and in the central city. These consumers enjoy a utility level equal to $2G$ from consuming all the public services. What do our main findings become in the presence of those spillovers? For conciseness, we consider the simple case where $\alpha = 0$.

The optimal metropolitan area. In presence of public good spillovers, the analysis of Section 3 is no longer valid because the planner faces an additional trade-off. On the one hand, the suburbanites enjoy an additional utility gain given by $m\,(y - b)\,G$; on the other hand, the cost of the public services provided by the central city is raised by $c\,(my)^2\,/2 - c\,(mb)^2\,/2$. If the total utility change is non-positive, that is, $G \leq cm\,(b + y)\,/2$, the analysis of Section 3 holds true because consuming the public services supplied in the central city makes the suburbanites worse-off. When $G > cm\,(b+y)\,/2$, the social welfare function becomes

$$\begin{aligned} W_T &= mBG + m\,(y-b)\,G - \frac{m\tau}{2}y^2 - m\tau\left(\frac{B-y}{2}\right)^2 \\ &\quad -\frac{c}{2}\,(my)^2 - \frac{cm}{2}\,(B-b)^2 - (m+1)F. \end{aligned} \tag{42}$$

Solving the first-order conditions for welfare maximization with respect to b and y yields

$$\bar{b}^{sp} = B - \frac{G}{c} < B \qquad \bar{y}^{sp} = \frac{2G + B\tau}{3\tau + 2cm}, \tag{43}$$

where $\bar{b}^{sp} > 0$ if and only if cB exceeds G; otherwise, we have $\bar{b}^{sp} = 0$.

The question addressed here is to figure out how public good spillovers affect the optimal organization of the MA. One solution consists in comparing $\bar{y}^{sp}$ with $\bar{y}$ and $\bar{b}^{sp}$ with $\bar{b}$. The expression (43) shows that the central city population size shrinks whereas its economic size expands with G. More precisely, when G is large enough for $G > cm\left(\bar{b}^{sp} + \bar{y}^{sp}\right)/2$ to hold, we always have $\bar{b}^{sp} < \bar{b}$ while $\bar{y}^{sp} > \bar{y}$ if and only if $G > cmB/3$. Under these circumstances, the presence of public good spillovers leads the planner to choose a smaller central city but a larger CBD, the reason being that the value of $m\,(y-b)\,G$ in (42) is high. As a result, the discrepancy between $\bar{y}^{sp}$ and $\bar{b}^{sp}$ widens. In sum, when the utility of the public services is sufficiently high, *cross-border consumption is always socially desirable.*

The decentralized outcome. The indirect utility of a consumer living and working in an edge city remains unchanged, but the indirect utility of a suburbanite working in the central city is now given by $v_i^0(x) \equiv V_i^0(x) + G$. Equalizing the indirect utilities yields the equilibrium economic boundary of the central city:

$$y^{sp}(T_0, T) = \frac{B}{3} + \frac{2[G - (T_0 - T)]}{3\tau},$$

which increases with G because the central city is more attractive to the suburbanites. The objective function of the central city government is given by $W_0^{sp} = W_0(y^{sp}) - c(my^{sp})^2/2 + c\ell_0^2/2$, while the suburban governments maximize $W^{sp} = W(y^{sp}) + (y^{sp} - b)\,G$. The first order conditions are given, respectively, by

$$\frac{\mathrm{d}W_0^{sp}}{\mathrm{d}T_0} = m\left[\left(y^{sp} - b\right) - \frac{2T_0}{3\tau} + \frac{2cm}{3\tau}\right] \qquad \frac{\mathrm{d}W^{sp}}{\mathrm{d}T} = -\frac{2T}{3\tau},$$

which implies

$$T_0^{sp} = \frac{(3\tau + 2cm)\,(B\tau + 2G) - 9b\tau^2}{4(3\tau + cm)} \qquad T^{sp} = 0. \tag{44}$$

Standard calculations show that $\mathrm{d}W_0^{sp}/\mathrm{d}T_0 > 0$ when $T_0 = T_0^*$. As a result, *the business tax set by the central city is higher in the presence of public good spillovers than in the absence of* *such spillovers*. The reason is easy to grasp. More workers lure the CBD because they can enjoy the public services provided by the central city. This entices more firms to set up there. The extent of tax exporting thus increases, whereas the provision cost of public services is higher. Both effects lead the central city to increase its business tax. In the suburban jurisdictions, there are two opposite effects. On the one hand, fewer firms locate in a SBD, so that the corresponding local government collects less tax income from firms. On the other hand, a few suburbanites consume more public services. The two effects cancel each other, so that the tax rate set by the suburban jurisdictions does not change ($T^{sp} = T^{*} = 0$).

The existence of public good spillovers has both expected and unexpected distributional implications for consumers living in the central and peripheral jurisdictions. First, the out-commuting suburbanites benefit from more public services, whereas the central city residents bear a higher provision cost for the public good. What is more, because the business tax paid by the CBD firms is higher ($T_0^{sp} > T_0^{*}$), the central city workers get a lower pay. In contrast, the SBD-workers earn the same wage because $T^{sp} = T^{*} = 0$. As a consequence, *the central city residents are hurt twice through an externality effect and an income effect*. Therefore, the free-riding problem between the central city and the suburban jurisdictions has implications that go beyond the standard consumption effects generated by spillovers. This makes the cooperation between the central and edge cities even more compelling for the MA to be efficient.

The economic boundary of the central city y^{sp} is now such that

$$\bar{y}^{sp} - y^{sp} = \frac{3\tau + cm}{3\tau} \left(y^{sp} - b \right).$$

Hence, when $y^{sp} - b > 0$, there is again insufficient concentration of jobs in the CBD ($\bar{y}^{sp} > y^{sp}$). However, it is not clear whether the presence of public good spillovers exacerbates the misallocation of jobs within the MA. Indeed, although the consumption of the central city public services by suburbanites tends to generate more jobs in the CBD, the higher business tax set by the central city deters firms to locate in the CBD. The ultimate impact depends on the parameter values.

To summarize,

Proposition 8. *Assume a polycentric metropolitan area. If the suburbanites working in the CBD consume the central city's public goods, the tax differential between the central city and the edge cities rises. Furthermore, the size of CBD remains too small.*

6.2 The central city as a bigger provider of public goods

In the real world, the central city often supplies a broader range of public services than the suburban jurisdictions. It is, therefore, legitimate to ask what the above findings become in such a context. To show it, we assume that the central city provides a public good of size βG with $\beta > 1$.

The optimal population size of the central city increases with β because of the higher utility stemming from the consumption of the public services:

$$\bar{b}^{\beta} = \bar{b} + \frac{\beta (G-1)}{c(m+1)} > \bar{b}.$$

As a result, the institutional structure of the MA now depends on the *relative* provision of public services between the central city and the suburban jurisdictions, β. The optimal size of the central city increases with the range of public services it provides, whereas the suburban jurisdictions shrink.

Supplying a wider range of public services in the central city does not affect its optimal economic boundary as long as the suburbanites do not consume these services. Thus, the optimal MA may be institutionally fragmented while having an integrated labor market or may involve a single jurisdiction together with several employment centers. Moreover, since the business tax competition process does not depend on G, the asymmetry in the provision of public services has no impact on the jurisdictions' business taxes, and thus Propositions 2 and 3 remain valid.

When suburbanites working in the CBD consume the central city public services, the central city becomes more attractive. In this case, the discrepancy between the optimal administrative and economic boundaries is exacerbated. Indeed, the limit given by (43) is unaffected because the planner has no reason to differentiate across CBD-workers. On the contrary, $\bar{y}^{sp}$ increases with β because more consumers are able to enjoy the wider array of public services provided by the central city. Furthermore, as shown by (44) in which G is replaced with βG, the central city's government increases its business tax. As a consequence, the tax differential widens and Proposition 3 holds true.

7 Concluding remarks

Metropolitan areas are non-legal entities that play a key role in the economic development of emerging and developed countries alike. This probably explains why political scientists have long been interested in issues related to metropolitan governance. The earliest approach that we are aware of - the regionalism approach that continues to shape the political debates - views the multiplicity of political jurisdictions as inherently inefficient. Political fragmentation would limit the ability to deal with area-wide urban problems that transcend local jurisdictions. The prescription is then to promote metropolitan governments and a better correspondence between political and functional or economic areas. In contrast to this view, the public choice approach, based on Ostrom et al. (1961), does not see systematic inefficiency in the polycentric political organization of a MA.

Our model delivers a clear-cut message that strongly suggests an intermediate approach. Indeed, both the first-best and second-best solutions involve a certain degree of decentralization of public services within the MA, as well as an economic boundary of the central city that is larger than its administrative boundary. In addition, redrawing the boundary between the central and suburban jurisdictions does not allow one to reach the first best, but such a policy may dampen the inefficient allocation of firms and jobs across employment centers. This points to the need for multifunctional governance: "small" things should be managed by local jurisdictions, and "big" things by a metropolitan government. Labor and transport issues in particular should be handled at the metropolitan level. Although derived from a simple model, these conclusions are sufficient to show that policy recommendations based on the regionalism and public choice approaches are unwarranted.

Although we recognize that political fragmentation is not bad per se, the tax competition process, in which central-city business taxes are inefficiently high relative to suburban business tax rates, leads to an inefficient distribution of firms and jobs. This leads us to formulate some policy recommendations in the spirit of what is known as "New Regionalism" (Savitch and Vogel, 2000) - mixing a polycentric political system with inter-municipal cooperation to solve mutual problems. Our analysis shows that some policies must be conducted at the level of the MA as a whole. Our framework can also serve to address more controversial issues in local public finance and transportation economics. According to Inman (2009), rethinking the governance of the MAs through Business Improvement Districts (BIDs) could be a way to improve the fiscal performance of a large MA. BIDs are business associations and can be considered self-financing private governments that offer supplemental services to their members. By restoring market-driven incentives in location choices, the development of BIDs within the central city would make it more attractive, thus strengthening agglomeration economies. Helsley and Strange (1998) analyze such organizations and show that their welfare effects on consumers are ambiguous and complex. However, their analysis should be extended to the case of a genuine urban framework with the aim of determining the impact of private governments on the spatial distribution of firms and consumers.

Our analysis supports what seems to be the minimal set of requirements needed to promote more efficient MAs. Yet, our findings have been obtained under several simplifying assumptions; thus care is needed in interpreting them. First, we did not address competition in public goods, an issue that is notoriously difficult, especially because many models are plagued with the non-existence of a Nash equilibrium. In this respect, it is worth noting that our model may be interpreted as one in which jurisdictions avoid the damaging effects of a race to the bottom by coordinating their supply of public services. Therefore, even in the absence of such distortions, our analysis has unveiled new sources of market failure. Moreover, it is well known that one political and social difficulty encountered within a MA stems from the heterogeneity of households that cluster in specific neighborhoods, which in turn generates spatial discrimination across socioeconomic groups. This issue has been tackled in the monocentric city model of urban economics (de Bartolome and Ross, 2004), but has not been explored in the context of a polycentric MA. Lastly, we did not allow consumers to choose a variable lot size by trading the homogeneous good against land. Several of our results remain valid when the population density is not uniform anymore but the determination of the equilibrium land rent within each jurisdiction is a more delicate issue.

References

Arnott, R. J. and J. E. Stiglitz (1979) Aggregate land rents, expenditure on public goods, and optimal city size. *Quarterly Journal of Economics* 93: 471-500.

Baum-Snow, N. (2010) Changes in transportation infrastructure and commuting patterns in U.S. metropolitan areas, 1960-2000. *American Economic Review. Papers & Proceedings* 100: 378-82.

Baum-Snow, N. (2014) Urban transport expansions, employment decentralization, and the spatial scope of agglomeration economies. Brown University, mimeo.

Borck, R. and M. Wrede (2009) Subsidies for intracity and intercity commuting. *Journal of Urban Economics* 66: 25-32.

Braid, R. (1996) Symmetric tax competition with multiple jurisdictions in each metropolitan area. *American Economic Review* 86: 1279-90.

Braid, R. (2002) The spatial effects of wage or property tax differentials, and local government choice between tax instruments. *Journal of Urban Economics* 51: 429-45.

Braid, R. (2010) Provision of a pure local public good in a spatial model with many jurisdictions. *Journal of Public Economics* 94: 890-7.

Brülhart, M., S. Bucovetsky and K. Schmidheiny (2015) Taxes in cities. In: G. Duranton, J.V. Henderson and W. Strange (eds.) *Handbook of Regional and Urban Economics. Volume 5.* Amsterdam: Elsevier, 1123-96.

Bucovetsky, S. (1991) Asymmetric tax competition. *Journal of Urban Economics* 30: 167-81.

Cheshire, P. and S. Magrini (2009) Urban growth drivers in a Europe of sticky people and implicit boundaries. *Journal of Economic Geography* 9: 85-116.

Combes, P.-P., L. Gobillon (2015) The empirics of agglomeration economies. In: G. Duranton, J.V. Henderson and W. Strange (eds.) *Handbook of Regional and Urban Economics. Volume 5.* Amsterdam: Elsevier, 247-348.

Cremer, H. and P. Pestieau (2004) Factor mobility and redistribution. In: J.V. Henderson and J.-F. Thisse (eds.). *Handbook of Regional and Urban Economics. Volume 4.* Amsterdam: Elsevier, 2529-60.

de Bartolome, C. and S. Ross (2004) Who's in charge of the central city? The conflict between efficiency and equity in the design of a metropolitan area. *Journal of Urban Economics* 56: 458-83.

Epple, D. and A. Zelenitz (1981) The implications of competition among jurisdictions: Does Tiebout need politics? *Journal of Political Economy* 89: 1197-217.

Fujita, M. (1989) *Urban Economic Theory. Land Use and City Size.* Cambridge: Cambridge University Press.

Fujita, M. and H. Ogawa (1982) Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics* 12: 161-96.

Glaeser, E. L. (2013) Urban public finance. In: A.J. Auerbach, R. Chetty, M. and E. Saez (eds.). *Handbook of Public Economics. Volume 5.* Amsterdam: Elsevier, 195-256.

Glaeser, E.L. and J.D. Gottlieb (2009) The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature* XLVII: 983-1028.

Glaeser, E.L. and M.F. Kahn (2001) Decentralized employment and the transformation of the American city. *Brookings-Wharton Papers on Urban Affairs* (2): 1-47.

Glaeser, E.L. and M.E. Kahn (2004) Sprawl and urban growth. In: J.V. Henderson and J.-F. Thisse (eds.). *Handbook of Regional and Urban Economics. Volume 4.* Amsterdam: Elsevier, 2481-527.

Haughwout, A. and R. Inman (2009) How should the suburbs help their central city? Growth- and welfare-enhancing intrametropolitan fiscal distributions. *Annals of the American Academy of Political and Social Science* 621: 39-52.

Helsley, R.W. and W.C. Strange (1998) Private government. *Journal of Public Economics* 69: 281-304.

Hochman, O., D. Pines and J.-F. Thisse (1995) On the optimal structure of local governments. *American Economic Review* 85: 1224-40.

Hoyt, W.H. (1991) Property taxation, Nash equilibrium and market power. *Journal of Urban Economics* 30: 123-31.

Hoyt, W.H. (1992) Market power of large cities and policy differences in metropolitan areas. *Regional Science and Urban Economics* 22: 539-58.

Inman, R. (2009) *Making Cities Work.* Princeton (NJ): Princeton University Press.

Janeba, E. and S. Osterloh (2013) Tax and the city - A theory of local tax competition.

Journal of Public Economics 106: 89-100.

Kanbur, R. and M. Keen (1993) Jeux sans frontières: Tax competition and tax coordination when countries differ in size. *American Economic Review* 83: 877-92.

Keen, M. and K.A. Konrad (2013) The theory of international tax competition and coordination. In: A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.). *Handbook of Public Economics. Volume 5.* Amsterdam: Elsevier, 257-328.

Koh, H.-J., N. Riedel and T. Böhm (2013) Do governments tax agglomeration rents?. *Journal of Urban Economics* 75: 92-106.

Laussel, D. and M. Le Breton (1998) Existence of Nash equilibria in fiscal competition games. *Regional Science and Urban Economics* 28, 283-96.

Lucas, R.E. and E. Rossi-Hansberg (2002) On the internal structure of cities. *Econometrica* 70: 1445-76.

Myers, G. (1990) Optimality, free mobility, and the regional authority in a federation. *Journal of Public Economics* 43: 107-21.

Noiset, L. and W. Oakland (1995) The taxation of mobile capital by central cities. *Journal of Public Economics* 57: 297-316.

OECD (2006) *Competitive Cities in the Global Economy.* OECD Territorial Reviews.

Ostrom, V., C. Tiebout and R. Warren (1961) The organization of government in metropolitan areas: A theoretical inquiry. *American Political Science Review* 55: 831-42.

Savitch H.V and R. Vogel (2000) Paths to New Regionalism. *State and Local Government Review* 32: 158-68.

Appendix

The Nash tax equilibrium is derived from the two reaction functions (35) and (36). Solving these two equations yields

$$\begin{aligned} T_0^* &= \frac{B}{4}\frac{2\alpha\left[2\alpha(m-1)-\tau\right]+D\left[\tau+2\alpha(\lambda m-1)\right]}{2\alpha+D} \\ &\quad -\frac{b}{4}\frac{2\alpha\left[3\tau-2\alpha m(m+1)+4\alpha\lambda m\right]+D\left[3\tau+2\alpha(\lambda m-1)\right]}{2\alpha+D}, \end{aligned}$$

and

$$T^* = -\frac{2\alpha B\left[\tau-\alpha m(1-\lambda)\right]}{2\alpha+D}+\frac{(1-\lambda)\,2\alpha^2 bm}{2\alpha+D}.$$

To show that $T_0^* > 0$, consider T as the abscise and T_0 as the ordinate of the plane. The expression (35) shows that $T_0 = 0$ when $T_1 = -\alpha(\lambda m-1)(B-b)-\tau(B-3b)/2 < 0$. The best reply of a suburban government (36) yields

$$T_2 = \frac{8\alpha\left[\alpha m(B-(2\lambda-1)b\right]-2\alpha(B-b)-\tau(B+3b)}{D}$$

when $T_0 = 0$. Since $T_2 > T_1$ when $\alpha = 0$ and T_2 grows while T_1 decrease with α, it is straightforward that $T_0^*(T)$ and $T^*(T_0)$ always intersect in the positive half-plane $T_0 > 0$. Q.E.D.