Are Compact Cities Environmentally (and Socially) Desirable ?

Carl Gaigné

Stéphane Riou

Jacques-François Thisse

_____

Gaigné: Corresponding author. INRA, UMR1302 SMART, 35000 Rennes (France) and CREATE, Université Laval, Québec, Canada
carl.gaigne@rennes.inra.fr
Riou: GATE Lyon-Saint-Etienne; Université de Lyon (France); Université Jean Monnet, Saint-Etienne
stephane.riou@univ-st-etienne.fr
Thisse: CORE, Université Catholique de Louvain (Belgium), CREA, Université du Luxembourg, and CEPR
jacques.thisse@uclouvain.be

Abstract:

There is a wide consensus among international institutions and national governments in favor of compact (i.e. densely populated) cities as a way to improve the ecological performance of the transport system. Indeed, when both the intercity and intra-urban distributions of activities are given, a higher population density makes cities more environmentally friendly as the average commuting length is reduced. However, when we account for the possible relocation of activities within and between cities in response to a higher population density, the latter may cease to hold. Because changes in population density affect land rents and wages, firms and workers re-optimize and choose new locations. We show that this may reshape the urban system in a way that generates both a higher level of pollution and welfare losses. As cities become more compact, agglomeration occurs and, eventually, the secondary business centers vanish. By increasing the average commuting length, these changes in the size and structure of cities may be detrimental to both the ecological and welfare objectives even if intercity trade flows decrease. This means that compact is not always desirable, and thus an increasing-density policy should be supplemented with instruments that impact the intra- and inter-urban distributions of activities. We argue that a policy promoting the creation of secondary business centers can raise welfare and decrease emissions.

Keywords: Greenhouse gas, commuting costs, transport costs, cities

Résumé:

Il existe un large consensus parmi les institutions internationales et les gouvernements nationaux en faveur des villes compactes (c'est-à-dire densément peuplées) comme un moyen d'améliorer la performance écologique de nos systèmes de transport. En effet, lorsque la localisation des activités demeure inchangée, une forte densité de population rend les villes plus respectueuses de l'environnement car la longueur du trajet moyen est réduite. Cependant, les effets à long terme sont incertains dans la mesure où la localisation des activités au sein et entre les villes s'ajuste en réponse à une forte densité de population. En effet, le changement de densité de population, en affectant les rentes foncières et les salaires, modifie la localisation des entreprises et des travailleurs, et donc la demande de transport. Nous montrons que la densification des villes peut remodeler le système urbain d'une manière qui détériore la situation environnementale et le bien-être. Des villes plus compacts accroissent la concentration spatiale et, éventuellement, réduisent la taille des pôles d'emploi secondaires. En augmentant la longueur de trajet en moyenne, ces changements dans la taille et la structure des villes peuvent être préjudiciables pour l'environnement et le bien-être, même si le commerce interurbain de marchandises diminue. Nous soutenons qu'une politique de promotion de la création de pôles d'emploi secondaires peut augmenter le bien-être et diminuer la pollution liée au transport.

Mots clés: Gaz à effet de serre, transport de marchandises, déplacement domicile/travail, villes compactes

Classification JEL: D61, F12, Q54, Q58, R12

1 Introduction

According to Yvo de Boer, former Executive Secretary of the United Nations, "Given the role that transport plays in causing greenhouse gas emissions, any serious action on climate change will zoom in on the transport sector" (speech to the Ministerial Conference on Global Environment and Energy in Transport, January 15, 2009). The transport sector is indeed a large and growing emitter of greenhouse gases (GHG). It accounts for 30% of total GHG emissions in the US and approximately 20% of GHG emissions in the EU-15 (OECD, 2008). Within the EU-27, GHG emissions in the transport sector have increased by 28% over the period 1990-2006, whereas the average reduction of emissions across all sectors is 3%. Moreover, road-based transport accounts for a very large share of GHG emissions generated by the transport sector. For example, in the US, nearly 60% of GHG emissions stem from gasoline consumption for private vehicle use, while a share of 20% is attributed to freight trucks, with an increase of 75% from 1990 to 2006 (Environmental Protection Agency, 2011).[1]

Although new technological solutions for some transport modes might allow for substantial reductions in GHG emissions (Kahn and Schwartz, 2008), improvements in energy efficiency are likely to be insufficient to stabilize the pollution level in the transport sector (European Environment Agency, 2007). Thus, other initiatives are needed, such as mitigation policies based on the reduction of average distances travelled by people and commodities. To a large extent, this explains the remarkable consensus among international institutions as well as local and national governments to foster the development of *compact* (or densely populated) *cities* as a way of reducing the ecological impact of cities and contributing to sustainable urban development. Nevertheless, the analysis of global warming and climate change neglects the spatial organization of the economy as a whole and, therefore, its impact on transport demand and the resulting GHG emissions. It is our contention that such neglect is unwarranted.

A large body of empirical literature highlights the effect of city size and structure on GHG emissions through the level of commuting (Bento et al., 2006; Kahn, 2006; Brownstone and Golob, 2009; Glaeser and Kahn, 2010). The current trend toward increased vehicle use has been reinforced by urban sprawl, as suburbanites' trips between residences and workplaces have increased (Brueckner, 2000; Glaeser and Kahn, 2004). Kahn (2006) reports that the predicted gasoline consumption for a representative household is lowest in relatively compact cities such as New York and San Francisco and is highest in sprawling cities such as Atlanta and Houston. While the environmental costs of urban sprawl are increasingly investigated in North America, the issue is becoming important in Europe as well. For example, between 1986

---

[1]This increase is associated with an increase in the average distance per shipment. In France, from 1975 à 1995, the average kilometers per shipment has increased by 38% for all transportation modes, and by 71% for road transport only (Savin, 2000). Similar evolutions have been observed in the richer EU countries and in the USA.

and 1996 in the metropolitan area of Barcelona, the level of per capita emissions doubled, the average trip distance increased by 45%, and the proportion of trips made by car increased by 62% (Muniz and Galindo, 2005). Recognizing the environmental cost of urban sprawl, scholars and city planners alike advocate city compactness as an ideal.[2]

Starting from the prevailing view that more compact cities is the proper way to contain GHG emissions, our paper addresses the following two issues. Firstly, when assessing the merits of increasing-density policies, the existing literature disregards one major problem: *a higher population density may spark the relocation of firms and households.* Instead, a full-fledged analysis of such a policy should be conducted within a framework in which firms and households' locations are endogenously chosen in response to a higher population density. When its effects on the spatial distribution of firms and households are ignored, the environmental impact of an increasing-density policy is always positive because people commute over shorter distance. However, such a policy affects prices, wages and land rents, which lead firms and households to change places in order to re-optimize profits and utility. Accounting for these spatial effects makes the impact of higher urban density more ambiguous because their net effects depend on whether the new spatial pattern is better or worse from the environmental viewpoint.

Secondly, once it is recognized that the desirability of increasing-density policies depends on the resulting spatial pattern, another question comes to mind: which spatial distribution of firms/households minimizes transport-related GHG emissions in the space-economy as a whole? Transporting people and commodities involves environmental costs which are associated with the following fundamental trade-off: concentrating people and firms in a reduced number of large cities minimizes pollution generates by commodity shipping among urban areas but increases pollution stemming from a longer average commuting; dispersing people and firms across numerous small cities has the opposite effects. Therefore a sound environmental policy should be based upon the ecological assessment of the entire urban system. Although seemingly intuitive, this global approach has not been part of the debate on the desirability of compact cities.

That said, the above trade-off also has a monetary side, and thus an increasing-density policy has welfare implications that are often overlooked by compact cities' proponents. This should not come as a surprise because transporting people and commodities involves both economic and ecological costs. In other words, there is a tight connection between the ecological and welfare objectives. According to Stern (2008), the emissions of GHG are the biggest market failure that the public authorities have to manage. It is, therefore, tempting to argue that deadweight losses associated with market imperfections are of second order. This view is too extreme because a higher population density impacts the consumption of *all* goods, and thus

---

[2]See Dantzig and Saaty (1973) for an old but sound discussion of the advantages of compact cities. Gordon and Richardson (1997) provide a critical appraisal of this idea

changes individual welfare. Having this in mind, we show that increasing density may generate welfare losses when the urban system shifts from dispersion to agglomeration. For this reason, our paper focuses on both the *ecological* and *welfare* effects of a higher population density when firms and households are free to relocate *between* and *within* cities.

In doing so, we consider the following two urban scenarios. In the first one, cities are monocentric while consumers and firms are free to relocate *between* cities in response to a higher population density. We show that an increasing-density policy may generate a hike in global pollution when this policy leads the urban economy to shift from dispersion to agglomeration, or vice versa. For example, when both the initial population density and the unit commuting cost are low enough, an increasing-density policy incentivizes consumers and firms to concentrate within a single city. However, at this new spatial pattern, the density may remain sufficiently low for a single large city to be associated with a longer average commuting, which generates more pollutants than two small cities. Conversely, when the unit commuting cost is high, the market leads to the dispersion of activities because consumers aim to bear lower land rent and commuting costs. Yet, when the density gets sufficiently high, the average commuting is short enough for the agglomeration to be ecologically desirable because intercity transport flows vanish. Consequently, agglomeration or dispersion is not by itself the most preferable pattern from the ecological point of view. In other words, *our results question the commonly held belief of many urban planners and policy-makers that more compact cities are always desirable.* They also show that one should pay more attention to the effect of increasing-density policies on city size.

In the second scenario, we study the ecological and welfare impact of an increasing-density policy when both the city size and morphology are endogenously determined. By inducing high urban costs, a low population density leads to both the dispersion and decentralization of jobs, that is, the emergence of polycentric cities. If urban planners make the urban system more compact (i.e. raise population density), then, the secondary business centers shrink smoothly and, eventually, firms and households produce and reside in a single monocentric city. We show that these changes in the size and structure of cities may generate higher emissions from commuting. Thus, an increasing-density policy should be supplemented with instruments that influence the intra- and inter-urban distributions of households and firms. In particular, we argue that a decentralization of jobs within cities, that is, a policy promoting the creation of secondary business centers, both raises welfare and decreases GHG emissions.

In what follows, we assume that the planner chooses the same population density in all cities. Alternately, we could assume that city governments noncooperatively choose their own population density. Both approaches have merits that are likely to suit countries with different attitudes regarding major issues such as the development of more densely populated cities. Our main argument is that the planning outcome is typically used by economists when assessing the costs and benefits of a particular policy.

The remainder of the paper is organized as follows. In the next section, we present a model with two monocentric cities and discuss the main factors affecting the ecological performance of an urban system. While we acknowledge that our model uses specific functional forms, these forms are standard in economic theory and are known to generate results that are fairly robust against alternative specifications. Note also that using specific forms is not a serious issue as the main objective of the paper is not to prove a particular result, but to highlight the possible ambiguity of the desirability of more compact cities. Section 3 focuses on monocentric cities and presents the ecological and welfare assessment of an increasing-density policy. In Section 4, we extend our analysis to the more general case in which both the internal structure of cities and the intercity distribution of activities are determined endogenously by the market. The last section offers our conclusions.

2 The model

Consider an economy endowed with $L > 0$ mobile consumers/workers, two cities where city $r = 1, 2$ hosts L_r consumers (with $L_1 + L_2 = L$), one manufacturing sector, and three primary goods: labor, land, and the numéraire which is traded costlessly between the two cities. Cities are assumed to be anchored and separated by a given physical distance. In order to disentangle the various effects at work, it is convenient to distinguish between two cases: in the former, workers are immobile, i.e. L_r is exogenous; in the latter, workers are mobile, i.e. L_r is endogenous. In this section, we describe the economy for a given distribution of workers between cities.

2.1 The city

Each city, which is formally described by a one-dimensional space, can accommodate firms and workers. Whenever a city is formed, it is monocentric with a central business district (CBD) located at $x = 0$ where city r-firms are set up.[3] Without loss of generality, we focus on the right-hand side of the city, the left-hand side being perfectly symmetrical. Distances and locations are expressed by the same variable x measured from the CBD. Our purpose being to highlight the interactions between the transport sector and the location of activities, we assume that the supply of natural amenities is the same in both cities.

We assume that the lot size is fixed and normalized to 1. Our policy instrument is given by the *tallness* (i.e. the number of floors) $\delta > 0$ of buildings. As a consequence, the parameter δ, which measures the city's compactness, is also the *population density*. Because δ is constant the population is uniformly distributed across the city. Although technically convenient, the assumption of a common and fixed lot size does not agree with empirical evidence: individual plots tend to be smaller in large cities than in small cities. Since the average commuting is

---

[3]See Duranton and Puga (2004) for a survey of the reasons explaining the emergence of a CBD.

typically longer in large than in small cities, we find it natural to believe that the plot size effect is dominated by the population size effect. Moreover, we test the robustness of our results in the case of nonuniform but given densities.

Because δ consumers are located at each point x, the right endpoint of the city is given by

$$y_r = \frac{L_r}{2\delta}$$

which increases (decreases) with population size (density).

2.2   Preferences and prices

Although new economic geography typically focuses on trade in differentiated products, it is convenient to assume that manufacturing firms are Cournot competitors producing a homogeneous good under increasing returns.[4] Location matters because transport costs are associated with the shipping of the manufactured good between cities. Thus each city's market may be served by local firms that produce domestically as well as by firms established in the other city. In this context, there is cross-hauling and the benefits from consuming more varieties, which are central to standard new economic geography models, are replaced by those generated by lower prices stemming from strategic competition between quantity-setting firms (Brander and Krugman, 1983).

Because the manufactured good is homogeneous, the quadratic utility proposed by Ottaviano et al. (2002) becomes

$$\max U_r = \left(a - \frac{q_r}{2} \right) q_r + q_0 \tag{1}$$

where q_r is the consumption of the manufactured good and q_0 the consumption of the numéraire. The unit of the manufactured good is chosen for $a = 1$ to hold. Each worker is endowed with one unit of labor and $\bar{q}_0 > 0$ units of the numéraire. The initial endowment $\bar{q}_0$ is supposed to be large enough for the individual consumption of the numéraire to be strictly positive at the equilibrium outcome. Each individual works at the CBD and bears a unit commuting cost given by $t > 0$, which implies that the commuting cost of a worker located at $x > 0$ is equal to tx units of the numéraire. Note that the lot size does not enter the utility because it is constant throughout our analysis.

The budget constraint of a worker residing at x in city r is given by

$$q_r p_r + q_0 + R_r(x)/\delta + tx = w_r + \bar{q}_0 \tag{2}$$

where p_r is the price of the manufactured good and w_r the wage paid by firms in city r's CBD. In this expression, $R_r(x)$ the land rent at x, and thus $R_r(x)/\delta$ is the price paid by a consumer

---

[4]The same modeling strategy is used, among others, by Gaigné and Wooton (2011), Haufler and Wooton (2010), and Thisse (2010).

to reside at x. Within each city, a worker chooses her location so as to maximize her utility (1) under the budget constraint (2).[5]

Utility maximization leads to the individual inverse demand for the manufactured good

$$p_r = \max\{1 - Q_r/L_r, 0\} \tag{3}$$

where Q_r is the total quantity of the manufactured good sold in city r.

Because of the fixed lot size assumption, the equilibrium value of urban costs, defined as the sum of commuting costs and land rent, is the same across workers' locations. The opportunity cost of land being normalized to zero, the equilibrium land rent is then given by

$$R_r^*(x) = \delta t \left(\frac{L_r}{2\delta} - x\right) \qquad \text{for } x < y_r. \tag{4}$$

Let n denote the number of operating firms and n_r the number of firms located in city r. To operate, a firm needs a fixed amount of labor $f > 0$ and m units of labor to produce one unit of the good. The unit of labor is chosen for $f = 1$. Moreover, the inverse demand functions being linear, we may normalize m to zero without loss of generality.

The manufactured good is shipped between cities at the cost of $\tau > 0$ units of the numéraire. Because they are spatially separated, the two markets are segmented (Engel and Rogers, 1996, 2001). This means that each firm chooses a specific quantity to be sold on each market; let q_{rs} be the quantity of the manufactured good that a city r-firm sells in city $s = 1, 2$.

The profits of a city r-firm are given by $\Pi_r = \pi_r - w_r$ where operating profits are defined by

$$\pi_r = q_{rr} p_r + q_{rs}(p_s - \tau) \quad \text{with } s \neq r \tag{5}$$

while w_r denotes the wage rate in city r (recall that firms use one unit of labor). Firms compete in quantity. Therefore, using (3) and $Q_r = n_r q_{rr} + n_s q_{sr}$, the equilibrium quantities sold by a city r-firm are given by $q_{rr}^* = L_r p_r^*$ and $q_{rs}^* = L_s(p_s^* - \tau)$. Substituting q_{rr}^* and q_{sr}^* into Q_r and the resulting expression in the inverse demand function, we obtain the following equilibrium prices:

$$p_r^* = \frac{1 + \tau(n - n_r)}{n + 1} \tag{6}$$

which decreases with the number of domestic firms.

Last, inserting (6) in q_{rr}^* and q_{rs}^* yields a firm's sales in each city:

$$q_{rr}^* = L_r \frac{1 + \tau(n - n_r)}{n + 1} \qquad q_{rs}^* = L_s \frac{1 - \tau(1 + n_s)}{n + 1}. \tag{7}$$

Observe that a firm exports more when transport costs decreases whereas its domestic sales decrease because competition from foreign firms in tougher. Moreover, trade between cities arises regardless of the intercity distribution of firms if and only if

$$\tau < \tau_{trade} \equiv \frac{1}{n + 1} < 1 \tag{8}$$

---

[5]For simplicity, we assume that land is owned by absentee landlords.

a condition which we assume to hold throughout the paper.

Urban labor markets are local. Labor market clearing implies

$$L_r = n_r \tag{9}$$

with $n_r + n_s = n$. The equilibrium city r-wage is determined by the zero-profit condition. In other words, the operating profits evaluated at the equilibrium prices and quantities are completely absorbed by the wage bill and no firm can profitably enter the market. Formally, this means that the equilibrium wages are determined by the conditions $\Pi_r(w_r, w_s) = 0$ and $\Pi_s(w_r, w_s) = 0$. Substituting (6) and (7) in (5), we get the equilibrium wages

$$w_r^* = \frac{L_r\left[1 + \tau(L - L_r)\right]^2}{(L+1)^2} + \frac{L_s\left[1 - \tau\left(1 + L_s\right)\right]^2}{(L+1)^2} \tag{10}$$

where $r, s = 1, 2$ and $s \neq r$. To sum up, (9) and (10) characterize labor market clearing in each city.

The indirect utility of a city r-worker is

$$V_r = S_r^* + w_r^* - UC_r + \bar{q}_0 \tag{11}$$

where S_r^* is the consumer surplus evaluated at the equilibrium prices (6):

$$S_r^* = \frac{(L - \tau L_s)^2}{2(L+1)^2} \tag{12}$$

while UC_r are the *urban costs* borne by a city r-worker, defined as the sum of land rent and commuting costs. It follows immediately from (4) that

$$UC_r = tL_r/2\delta.$$

Hence, for a given intercity distribution of workers, increasing urban population density through taller buildings leads to a higher individual welfare because urban costs are lowered.

2.3 The ecological trade-off in a space-economy

As mentioned in the introduction, goods' shipping and work-trips are the two main sources of GHG emissions generated in the transport sector. To convey our message in a simple way, the carbon footprint (E) of the urban system is obtained from the total distance travelled by commuters within cities (C) and from the total quantity of the manufactured good shipped between cities (T):

$$E = e_C C + e_T T \tag{13}$$

where e_C is the amount of carbon dioxides generated by one unit of distance travelled by a worker, while shipping one unit of the manufactured good between cities generates e_T units

of carbon dioxides. The value of e_C depends on the technology used (fuel less intensive and non-fuel vehicles, eco-driving and cycling) and on the commuting mode (public transportation versus individual cars). As for the value of e_T, it primarily depends on the distance between the two cities and the transport mode (road freight versus rail freight), but also on the technology (e.g. truck size) and the transport organization (empty running, deliveries made at night, ...).

For simplicity, we assume that e_C and e_T are given parameters which are independent from city size and compactness. Admittedly, these are strong assumptions. First, because collective forms of transport are more viable in larger and/or more compact cities, one would expect e_C to be a decreasing function of city size and/or compactness. Under these circumstances, intercity migrations increases the value of e_C in the origin city but leads to a lower e_C in the destination city. As a result, the global impact of migration would depend on a second-order effect which is hard to assess. In what follows, we treat e_C as a parameter and will discuss what our results become when e_C varies.

Denote by λ the share of workers (firms) residing (producing) in city 1. In equilibrium, consumers/workers are symmetrically distributed on each side of the CBD. Conditional upon this pattern, the value of C depends on the intercity distribution of the manufacturing sector and is given by

$$C(\lambda) = 2 \int_0^{y_1} \delta x \mathrm{d}x + 2 \int_0^{y_2} \delta x \mathrm{d}x = \frac{L^2}{4\delta}[\lambda^2 + (1 - \lambda)^2]. \tag{14}$$

Clearly, the emission of carbon dioxides stemming from commuting increases with λ for all $\lambda > 1/2$ and is minimized when workers are evenly dispersed between two cities ($\lambda = 1/2$). In addition, for any given intercity distribution of activities, the total amount of emission decreases with the population density because the distance travelled by each worker shrinks. Observe that C is independent of t when λ is fixed because the demand for commuting is perfectly inelastic.

Regarding the value of T, it is given by the sum of equilibrium trade flows, $n_1 q_{12}^* + n_2 q_{21}^*$. Using (7) and (9), we have

$$T(\lambda) = \frac{[2 - \tau(L+2)]L^2}{L+1} \lambda (1 - \lambda) \tag{15}$$

where, owing to (8), $T > 0$. As expected, T is minimized when workers and firms are agglomerated within a single city ($\lambda = 0$ or 1) and maximized when $\lambda = 1/2$. Note also that T increases when shipping goods becomes cheaper because intercity trade grows. Hence, transport policies that foster lower trading costs give rise to a larger emission of GHG.

The ecological trade-off we want to study may then be stated as follows: *a more agglomerated pattern of activity reduces pollution arising from commodity intercity shipping, but increases the GHG emissions stemming from a longer average commuting; and vice versa.*

We acknowledge with Glaeser and Kahn (2010) that most car trips, at least in the US, are not commuting trips. Car trips are also made for shopping as well as for some other activities.

Likewise, the objective function E could be augmented by introducing emissions stemming from the production of the manufactured good. As shown in Appendix A, accounting for these additional sources of emissions does not affect our results because their analytical expression behaves like C. We could similarly take into account the distribution of goods within metropolitan areas, which depends on both the city size and the consumption level. For example, the US commodity flows survey reports that more than 50% of commodities (in volume) are shipped over a distance less than 50 miles (US Census Bureau, 2007). In a nutshell, accounting for the emission of GHG stemming from additional sources such as shopping trips, production and intra-city goods transport makes the case for dispersion stronger.

3   City size and the environment

In this section, we determine the market outcome and study the ecological desirability of an increasing population density when workers and firms are free to relocate between cities.

3.1   The market outcome

As in the core-periphery model, firms and workers move hand-in-hand, which means that workers' migration drives firms' mobility. A long-run equilibrium is reached when no worker, hence firm, has an incentive to move. It arises at $0 < \lambda^* < 1$ when the utility differential between the two cities $\Delta V(\lambda^*) \equiv V_1(\lambda^*) - V_2(\lambda^*) = 0$, or at $\lambda^* = 1$ when $\Delta V(1) \geq 0$. An interior equilibrium is stable if and only if the slope of the indirect utility differential ΔV is strictly negative in a neighborhood of the equilibrium, i.e., $d\Delta V/d\lambda < 0$ at λ^*; an agglomerated equilibrium is stable whenever it exists.

Using (11), the utility differential is given by (up to a positive and constant factor):

$$\Delta V(\lambda) \propto (\delta - \delta_m)\left(\lambda - \frac{1}{2}\right) \tag{16}$$

where

$$\delta_m \equiv \frac{t}{(\varepsilon_a - \varepsilon_b \tau)\tau} > 0$$

and $\varepsilon_a \equiv 2(2 + 3L)/(L + 1)^2 > 0$ and $\varepsilon_b \equiv (L + 2)(2L + 1)/(L + 1)^2 > 0$. Clearly, $(\varepsilon_a - \varepsilon_b \tau)\tau$ is positive and increasing with respect to τ because $\tau_{trade} < \varepsilon_a/2\varepsilon_b$. Consequently, the agglomeration of firms and workers within one monocentric city is the only stable equilibrium when $\delta > \delta_m$. In contrast, if $\delta < \delta_m$, dispersion with two identical monocentric cities is the unique stable equilibrium.

To sum up, we have:

Proposition 1 *Workers and firms are agglomerated into a monocentric city when population density is high, commuting costs are low, and transport costs are high. Otherwise, they are evenly dispersed between cities.*

3.2   Minimizing the ecological footprint

Because E is described by a concave or convex parabola in λ, the emission of GHG is minimized either at $\lambda = 1$ or at $\lambda = 1/2$. Thus, it is sufficient to evaluate the sign of $E(1) - E(1/2)$. It is readily verified that the agglomeration (dispersion) of activities is ecologically desirable if and only if $\delta > \delta_e$ $(\delta < \delta_e)$ with

$$\delta_e \equiv \frac{e_C}{2e_T} \frac{L+1}{2 - \tau(L+2)}.$$

Because $d\delta_e/d\tau > 0$ and $d\delta_e/dL > 0$, we have:

Proposition 2 *Assume that cities are monocentric. Pollution stemming from commuting and shipping is minimized under agglomeration (dispersion) when population density is high (low), transport costs are low (high), or the total population is low (high).*

Hence, agglomeration or dispersion is not by itself the most preferable pattern from the ecological point of view. A compact city is ecologically desirable only if the population density is sufficiently high for the average commuting distance to be small enough. But what do "high" and "small" mean? The answer depends on the structural parameters of the economy that determine the value of the threshold δ_e. Indeed, δ_e increases with e_C but decreases with e_T. In addition, the adoption of commuting modes with high environmental performance (low e_C) decreases the density threshold value above which agglomeration is ecologically desirable, while transport modes for commodities with high environmental performance (low e_T) increases this threshold value.

Our framework also sheds light on the effects of a carbon tax levied on the transport of commodities. The implementation of such a tax is formally equivalent to a rise in transport costs (τ). For any intercity distribution of firms, increasing transport costs reduce pollution (see (15)). However, raising transport costs fosters agglomeration (because δ_m decreases), while this spatial configuration tends to become ecologically less desirable (because δ_e increases). Therefore, the evaluation of a carbon tax should not focus only upon price signals. It should also account for its impact on the spatial pattern of activities. Finally, observe that δ_e is independent of t. Nevertheless, as shown by Proposition 1, the value of t impacts on the interregional market pattern, thus on the ecological outcome.

3.3   Are more compact cities desirable?

(i) The ecological viewpoint.   We determine the conditions under which the market yields a good or a bad outcome from the ecological viewpoint. Because $\delta_m = 0$ at $t = 0$ and increases with t, while δ_e is independent of t, there are four possible cases, which are depicted in Figure 1. In panels A and C, the market outcome minimizes pollution. In contrast, in panels B and D, the market delivers a configuration that maximizes the emissions of GHG. Consequently, *the market may yield as well as the best or the worst ecological outcome.*

Figure 1 shows that there exists a unique $\bar{t}$ such that

$$\delta_m \gtreqqless \delta_e \qquad \text{iff} \qquad t \gtreqqless \bar{t}.$$

Consider first the case where $t > \bar{t}$ (see Figure 2a). If $\delta < \delta_m$, the market outcome involves two cities. Keeping this configuration unchanged, a higher value of δ always reduces the emissions of pollutants. Note, however, that lower levels of GHG emissions could be reached under agglomeration for $\delta \in [\delta_e, \delta_m]$. Once δ exceeds δ_m, the economy gets agglomerated, thus leading to a downward jump in the GHG emissions. Further increases in δ yield lower emissions of GHG. Hence, when commuting costs are high enough, denser cities generate lower emissions of GHG.

Assume now that $t < \bar{t}$ (see Figure 2b). As in the foregoing, provided that $\delta < \delta_m$, the market outcome involves dispersion while the pollution level decreases when the cities get more compact. When δ crosses δ_m from below, the pollution now displays an upward jump. Under dispersion, however, lower levels of GHG emissions would have been sustainable over $[\delta_m, \delta_e]$. In other words, more compact cities need not be ecologically desirable because this recommendation neglects the fact that it may trigger interurban migrations. Consequently, once it is recognized that workers and firms are mobile, what matters for the total emission of GHG is the mix between the urban compactness (δ) and the interregional pattern (λ). This has the following major implication: *environmental policies should focus on the urban system as a whole and not on individual cities.*

The foregoing discussion shows how difficult it is in practice to find the optimal mix of instruments. Our model also allows us to derive some unsuspected results regarding the ability of instruments other than regulating the population density (carbon tax, low emission transport technology, ...) to reduce pollution. For example, when $t < \bar{t}$ the development of more ecological technologies in shipping goods between cities (low e_T) combined with the implementation of a carbon tax on carriers, which causes higher transport costs (high τ), lead to a higher value of δ_e and a lower value of δ_m. This makes the interval $[\delta_m, \delta_e]$ wider, while the value of $\bar{t}$ increases. Hence, the above policy mix, which seems a priori desirable, may exacerbate the discrepancy between the market outcome and the ecological optimum. Therefore, when combining different environmental policies, one must account for their impacts on the location of economic activities. Otherwise, they may result in a higher level of GHG emissions.

The conventional wisdom is that population growth is a key driver in damaging the environmental quality of cities. Restraining population growth is, therefore, often seen as a key instrument for reducing pollution. Indeed, for a given intercity pattern and a given density

level, $dE_m/dL > 0$ because a bigger population generates larger trade flows and longer commuting. Nevertheless, because firms and workers are mobile, a population hike may change the intercity pattern of the economy. For that, we must study how the corresponding increase in population size affects the greenness of the economy. In our setting, increasing L has the following two consequences. First, it raises the density threshold level $(d\delta_e/dL > 0)$ above which agglomeration is the ecological optimum. Second, dispersion becomes the market equilibrium for a larger range of density levels $(d\delta_m/dL > 0)$. What matters for our purpose is how the four domains in Figure 1 are affected by a population hike.

When $\bar{t}$ increases with L, then $\delta_m - \delta_e$ decreases with L provided that $t > \bar{t}$, whereas $\delta_e - \delta_m$ increases when $t < \bar{t}$. In this event, urban population growth decreases the occurrence of a conflict between the market and the ecological objective when commuting costs are high enough (see Figure 2a) but makes bigger the domain over which the market outcome is ecologically bad (see Figure 2b). When $\bar{t}$ decreases with L, the opposite holds. In both cases, as already noted by Kahn (2006) in a different context, *there is no univocal relationship between urban population growth and the level of pollution*. Our analysis provides a rationale for the non-monotonicity of the relationship observed between these two magnitudes.

To sum up,

Proposition 3 *Assume that cities are monocentric. If commuting costs are high, making cities more compact reduces pollution when the economy shifts from dispersion to agglomeration. However, when commuting costs are low, an increasing-density policy may be detrimental from the ecological viewpoint.*

The assumption of a constant population density is very restrictive. At the same time, it is well known that characterizing the market outcome with an endogenous determination of the population density in NEG-type models is a formidable task, which has been so far out of reach (Tabuchi, 1998). We want to take an intermediate approach in which the density is variable but exogenous. More precisely, we assume that the population density is now given by $\delta f(x)$ where $f(x)$ is a strictly decreasing function of the distance x to the CBD. In this case, a third spatial configuration spatial may emerge: the market outcome and the ecological optimum may involve a large and a small monocentric city when commuting costs take on intermediate values. However, as in the case of a uniform distribution of population density, increasing δ leads to lower land rents, which pushes toward a more concentrated pattern of activities and a higher value of the total distance travelled by commuters. Though the analytical details become much more cumbersome, Proposition 3 remains true. An example is explicitly dealt with in Appendix B.

(ii) The welfare viewpoint. As seen in the introduction, transporting people and commodities involves economic and ecological costs. It is not clear, however, what the welfare

implications of a higher population density are because the market outcome and the solution minimizing the ecological footprint involve different consumption levels of the manufactured good and of the numéraire. Therefore, it is important to figure out how social welfare is affected by changing the population density level.

For a given intercity distribution of activities, a higher population density is always welfare-enhancing because the average commuting costs are lower. However, when the population density becomes sufficiently high, firms and workers get agglomerated. In addition to increasing urban costs, this change in the spatial pattern has two effects on the utility level. First, it leads to a wage effect which is ambiguous. Indeed, agglomeration triggers fiercer competition and lowers the domestic price, which tends to reduce profits. Simultaneously, firms supply all consumers at a lower cost, thus leading to an output hike which tends to increase profits.[6] Second, as the negative effect of transport costs on consumption vanish, the consumer surplus increases (12). The total impact of denser cities on social welfare is, therefore, unclear. As a consequence, we must assess how the socially optimum configuration is affected by a higher population density. Because we study the environmental gains or losses associated with the market outcome, we find it natural to adopt a second-best approach in which the planner controls the location of firms and workers but not their production and consumption decisions.

Because utilities are quasi-linear and profits are wiped out by free entry, social welfare may be defined by the sum of indirect utilities evaluated at the equilibrium prices and incomes. In what follows, we have chosen to focus on welfare without accounting for the negative impact of pollution because weighting this externality in the social welfare function is often arbitrary. For the same reason, we do not include the various impacts that a higher population density has on the well-being of people.

Plugging the equilibrium values of S_r, w_r and UC_r into (11) for a given intercity distribution of firms and workers, the welfare function is given by

$$
\begin{aligned}
W_m(\lambda) &= \lambda L S_1^* + (1-\lambda) L S_2^* + \lambda L(w_1^* - UC_1) + (1-\lambda) L(w_2^* - UC_2) \qquad (17) \\
&= \frac{L(\varepsilon_a^o - \varepsilon_b^o \tau)\tau}{\delta} \lambda(\lambda - 1)(\delta - \delta_o) + \frac{(L+2)L}{2(L+1)^2} - \frac{tL}{2\delta}
\end{aligned}
$$

where

$$
\delta_o \equiv \frac{t}{(\varepsilon_a^o - \varepsilon_b^o \tau)\tau} > \delta_m
$$

and $\varepsilon_a^o \equiv 2(L+2)/(L+1)^2$ and $\varepsilon_b^o \equiv (2L^2 + 5L + 4)/2(L+1)^2$. Hence, agglomeration (dispersion) is welfare-maximizing when $\delta > \delta_o$ ($\delta < \delta_o$). The formal connection between the external environmental cost E and the social welfare function W_m is now clear: both are second-order polynomials in λ and the density δ governs the sign of their second derivatives.

Let λ^o be the intercity distribution maximizing W_m. The individual utilities being quasi-linear, interpersonal transfers allow consumers to reach their optimum utility level $W(\lambda^o)/L$,

---

[6]It is readily verified that $q_{rr}^*(1) + q_{rs}^*(1)$ exceeds $q_{rr}^*(1/2) + q_{rs}^*(1/2)$.

which is larger than the equilibrium utility level $V(\lambda^*)$. As a consequence, the market yields agglomeration when $\delta_o > \delta > \delta_m$, whereas dispersion is socially desirable. Otherwise, the market outcome and the social optimum are identical. However, this does not mean that a higher density is always welfare-enhancing. For example, as shown by Figure 3, when δ crosses δ_m from below, the welfare level displays a downward jump.

Insert Figure 3 about here

Observe that both the social welfare W_m and the carbon footprint E are second-degree polynomials in λ. In addition, the density δ governs the sign of their second derivatives. As a a result, social welfare and the amount of pollutants may be convex or concave functions of λ meaning that making cities more compact may or may not increase welfare and/or decrease pollution. Since δ_o exceeds δ_m, two cases may arise:

(i) When commuting costs are low $(t < \bar{t})$, our results imply that an increasing-population density policy should be accompanied by a growth control policy. Indeed, the polluting emissions in the global economy increases when δ crosses δ_m from below and takes a value in $[\delta_m, \delta_o]$ (see Figure 2b). In this case, by preventing the agglomeration of activities, the public authorities both reduce the GHG emissions and improve global welfare.

(ii) When commuting costs are high $(t > \bar{t})$, the desirability of a growth control policy is more controversial. When δ crosses δ_m from below and takes a value in $[\delta_m, \delta_o]$, such a policy yields higher welfare but washes out the environmental gains generated by the market (see Figure 2a). This is not, however, the end of the story. The conflict between the environmental and welfare criteria vanishes when $\delta > \delta_o$ because the market outcome both minimizes GHG emissions and maximizes social welfare.

To summarize,

Proposition 4 *Assume that cities are monocentric. If commuting costs are low, a higher population density may be harmful to both the environment and social welfare when the economy switches from dispersion to agglomeration. If commuting costs are high, a higher population density reduces pollution but may generate a welfare loss.*

This proposition is sufficient to show that the desirability of compact cities is more complex than suggested by their proponents, the main reason being that this recommendation disregards its impact on the location of economic activity.

4 The urban system and the environment

So far, we have treated the morphology of cities as given. In this section, we provide the ecological and welfare evaluation of the market outcome when the size and structure of each

city are endogenously determined. To reach our goal, we build on Cavailhès et al. (2007). Having done this, we show once more the possible perverse effects of city compactness and highlight the positive effects of job decentralization. Specifically, we argue that an alternative strategy could reduce the pollution emissions in the global economy: public authorities control the intra-urban distribution of firms and workers to decrease the average distance traveled by workers.

4.1 The size and structure of cities

In what follows, we determine the conditions for a city to become polycentric and, then, study how raising population density shapes the urban system.

1. The city structure. Firms are free to locate in the CBD or to form a *secondary business district* (SBD) on each side of the CBD, thus implying that a polycentric city has one CBD and two SBDs. Both the CBD and the SBDs are surrounded by residential areas occupied by workers. Although firms consume services supplied in the SBD, the higher-order functions (specific local public goods and non-tradeable business-to-business services) are still provided by the CBD. Hence, for using such services, firms established in a SBD must incur a communication cost $K > 0$. Communicating requires the acquisition of specific facilities, which explains why communication costs have a fixed component. Furthermore, as the distance between the CBD and SBDs is small compared to the intercity distance, shipping the manufactured good between the CBD and SBDs is assumed to be costless, which implies that the price of this good is the same everywhere within a city. Finally, without significant loss of generality, we restrict ourselves to the case of two SBDs. Hence, apart from the assumed existence of the CBD, the internal structure of each city is endogenous. Note that the equilibrium distribution of workers within cities depends on the distribution of workers between cities. In what follows, the superscript c is used to describe variables related to the CBD, whereas s describes the variables associated with a SBD.

At a city equilibrium, each worker maximizes her utility subject to her budget constraint, each firm maximizes its profits, and markets clear. Individuals choose their workplace (CBD or SBD) and their residential location for given wages and land rents. Given equilibrium wages and the location of workers, firms choose to locate either in the CBD or in a SBD. Or, to put it differently, no firm has an incentive to change place within the city, and no worker wants to change her working place and residence. In particular, at the city equilibrium, the distribution of workers is such that $V_r^c(\lambda) = V_r^s(\lambda) \equiv V_r(\lambda)$. Likewise, firms are distributed at the city equilibrium such that $\Pi_r^c(\lambda) = \Pi_r^s(\lambda)$.

Denote by y_r the right endpoint of the area formed by residents working in the CBD and by z_r the right endpoint of the residential area on the right-hand side of the SBD, which is

also the outer limit of city r. Let x_r^s be the center of the SBD in city r. Therefore, the critical points for city r are as follows:

$$y_r = \frac{\theta_r L_r}{2\delta} \qquad x_r^s = \frac{(1+\theta_r) L_r}{4\delta} \qquad z_r = \frac{L_r}{2\delta} \qquad (18)$$

where $\theta_r < 1$ is the share of city r-firms located in the CBD. Observe that the bid rents at y_r and z_r are equal to zero because the lot size is fixed and the opportunity cost of land is zero.

At the city equilibrium, the budget constraint implies that

$$w_r^c - R_r^c(x)/\delta - tx = w_r^s - R_r^s(x)/\delta - t\,|x - x_r^s|$$

where R_r^c and R_r^s denote the land rent around the CBD and the SBD, respectively. Moreover, the worker living at y_r is indifferent between working in the CBD or in the SBD, which implies

$$w_r^c - R_r^c(y_r)/\delta - ty_r = w_r^s - R_r^s(y_r)/\delta - t(x_r^s - y_r).$$

It then follows from $R_r^c(y_r) = R_r^s(y_r) = 0$ that

$$w_r^c - w_r^s = t(2y_r - x_r^s) = t\frac{3\theta_r - 1}{4\delta}L_r \qquad (19)$$

where we have used the expressions of y_r and x_r^s given in (18).

In each workplace (CBD or SBD), the equilibrium wages are determined by the zero-profit condition. As a result, the equilibrium wage rates in the CBD and in the SBD must satisfy the conditions $\Pi_r^c(w_r^c, w_r^s) = \Pi_r^s(w_r^c, w_r^s) = 0$. Solving these expressions for w_r^c and w_r^s, we get:

$$w_r^{c*} = \pi_r^* \qquad w_r^{s*} = \pi_r^* - K \qquad (20)$$

which shows that the wage wedge $w_r^{c*} - w_r^{s*}$ is positive.

Finally, the equilibrium land rent in the area occupied by the CBD-workers is given by

$$R_r(x) = R_r^c(x) = t\delta\left(\frac{\theta_r L_r}{2\delta} - x\right) \qquad \text{for } x < y_r \qquad (21)$$

where we have used the expression of y_r and the condition $R^C(y_r) = 0$, while the equilibrium land rent in the area occupied by the SBD-workers is as follows:[7]

$$R_r(x) = R_r^s(x) = t\delta\left[\frac{(1 - \theta_r)L_r}{4\delta} + x_r^s - x\right] \qquad \text{for } x_r^s < x < z_r. \qquad (22)$$

Substituting (10) and (20) into (19) and solving with respect to θ yields:

$$\theta_r^* = \min\left\{\frac{1}{3} + \frac{4\delta K}{3tL_r}, 1\right\} \qquad (23)$$

---

[7]In this expression, we do not account for the fact that transport modes may not be the same in these different areas of the metropolis. Our results remain valid as long as individual worktrips to a SBD do not generate much higher pollutants than those to the CBD.

which always exceeds 1/3: the CBD is always larger than each SBD. It is readily verified that city r is polycentric ($\theta_r^* < 1$) if and only if

$$\delta < \delta_r \equiv \frac{tL_r}{2K}. \tag{24}$$

Observe also that, when $\theta_r^* < 1$, a larger population L_r leads to a decrease in the relative size of the CBD, though its absolute size rises, whereas both the relative and absolute sizes of the SBDs rise. Indeed, increasing $\lambda_r L$ leads to a more than proportionate hike in the wage rate prevailing in the CBD because of the rise in the average commuting cost (see (19)). Moreover, because $\theta_r^* < 1$, the higher the city compactness, the larger the CBD; the lower the commuting cost, the larger the CBD. In short, *when city compactness steadily rises, both SBDs shrink smoothly and, eventually, the city becomes monocentric.*

2. The urban system. The utility differential between cities now depends on the degree of job decentralization within each city. The indirect utility of an individual working in the CBD is still given by (11) in which the urban costs she bears are now given by[8]

$$UC_r^c \equiv \theta_r^* \frac{tL_r}{2\delta} < UC_r.$$

It follows from (24) that

$$\delta_1 \equiv \frac{\lambda Lt}{2K} \qquad \delta_2 \equiv \frac{(1-\lambda)Lt}{2K} \tag{25}$$

where $\delta_1 \geq \delta_2$ because we focus on the domain $\lambda \geq 1/2$. The following three patterns may emerge: (i) when $\delta > \delta_1$, both cities are monocentric ($\theta_1^* = \theta_2^* = 1$), (ii) when $\delta_1 > \delta > \delta_2$, city 1 is polycentric and city 2 is monocentric ($\theta_1^* < \theta_2^* = 1$) (iii) when $\delta_2 > \delta$, both cities are polycentric ($\theta_r^* < 1$).

In order to determine the equilibrium outcome, we must consider the utility differential corresponding to each of these three patterns. In Appendix D, we show the existence and stability of five equilibrium configurations: (i) dispersion with two monocentric cities having the same size (m, m); (ii) agglomeration within a single monocentric city (m, 0);(iii) partial agglomeration with one large polycentric city and a small monocentric city (p, m); (iv) agglomeration within a single polycentric city (p, 0) and (v) dispersion with two polycentric cities having the same size (p, p). In Figure 4, the domains of the positive quadrant (K, δ) in which each of these configurations is a market outcome are depicted.

Insert Figure 4 about here

The implications of city compactness depend on the level of communication costs. We focus here on the today relevant case of low communication costs (see Appendix D), i.e.

$$K < \bar{K} \equiv \frac{L(\varepsilon_a - \varepsilon_b \tau)\tau}{4}.$$

---

[8]We may disregard the case of SBD-workers because, at the city equilibrium, they reach the same utility level as the CBD-workers.

In this event, the economy traces out the following path when the population density δ steadily increases from very small to very large values: we have (p, p) or (p, m) when $\delta < \delta_m/3$, then (p, m) when $\delta_m/3 < \delta < \delta_{pm}$, further (p, 0) when $\delta_{pm} < \delta < 2\delta_p$ with $\delta_p \equiv Lt/4K$ (see Appendix D for more details) and (m, 0) when $2\delta_p < \delta$, with

$$\delta_{pm} \equiv \frac{t}{3(\varepsilon_a - \varepsilon_b\tau)\tau - 4K/L}$$

which is positive because $K < \bar{K}$. This may be explained as follows. By inducing high urban costs, a low δ-value leads to both the dispersion and decentralization of jobs, that is, the emergence of two polycentric cities. When cities' population density gets higher, urban costs decrease sufficiently for the centralization of jobs within one city to become the equilibrium outcome; however, they remain high enough for the equilibrium to involve two cities of different sizes and structures. Last, for very high δ-values, urban costs become almost negligible, thus allowing one to save the cost of shipping the manufactured good through the emergence of a single city.

4.2 How the structure and size of cities impact on the environment?

We now determine whether more compact cities lead to lower GHG emissions when firms and workers are free to locate *between* and *within* cities.

The total level of emissions of GHG corresponding to the spatial structure $(\lambda^*, \theta_1^*, \theta_2^*)$ is given by

$$E(\lambda^*, \theta_1^*, \theta_2^*) = e_C C(\lambda^*, \theta_1^*, \theta_2^*) + e_T T(\lambda^*).$$

Note first that the value of T is still given by (15) because it does not depend on the city structure. In contrast, the total distance travelled by commuters depends on the internal structure of each city (θ_1 and θ_2) as well as on the distribution of workers/firms between cities:

$$C(\lambda, \theta_1, \theta_2) = \frac{\lambda^2 L^2}{4\delta} \left[\theta_1^2 + \frac{1}{2}(1 - \theta_1)^2 \right] + \frac{(1 - \lambda)^2 L^2}{4\delta} \left[\theta_2^2 + \frac{1}{2}(1 - \theta_2)^2 \right] \qquad (26)$$

which boils down to (14) when the two cities are monocentric ($\theta_1 = \theta_2 = 1$). It is straightforward to check that the GHG emissions increase when the CBDs grow. However, the strength of this effect decreases when cities become more compact.

For any given λ, the expression (26) shows that the decentralization of jobs away from the CBD leads to less GHG emissions through a shorter average commuting. Regarding the impact of a higher population density, it is a priori ambiguous. Indeed, for a given degree of job decentralization, a higher population density induces shorter commuting distances and, therefore, lower emissions. However, (23) shows that a rising δ also leads to a higher number of jobs in the CBD at the expense of the SBDs, which increases the emission of GHG. Plugging (23) into (26), we readily verify that the former effect dominates the latter. Thus, for any given λ, more compact cities always generate lower GHG emissions once the city equilibrium is reached. However, this result may be reverse once workers/firms can relocate between cities.

(i) Commuting. In order to disentangle the various effects at work, we begin by focusing on pollution stemming from commuting. For any given location pattern, a higher δ leads to a lower level of pollution. However, the impact of such a change in population density on the total distance travelled by commuters is not clear when firms and workers may change places within and between cities. In addition, one may wonder what happens when the economy shifts from one pattern to another.

To illustrate, we assume that the initial market outcome is given by (p, p). The corresponding GHG emissions generated by commuting are then given by

$$C_{\rm pp} \equiv \delta \left(\frac{L^2}{24\delta^2} + \frac{4K^2}{3t^2} \right).$$

As long as this urban configuration prevails, a higher population density reduces commuting pollution.[9] However, once δ crosses $\delta_m/3$ from below, the economy shifts to the configuration (p, m) (see Figure 4). At $\delta = \delta_m/3$, the level of pollution exhibits an upward jump.[10] This is because city 1, which remains polycentric, becomes larger while city 2, which now accommodates fewer workers, becomes monocentric.

At the equilibrium configuration (p, m), λ_{pm}^* increases with δ whenever $K < \bar{K}$.[11] In this case, (23) and (26) show that the level of pollution C_{pm} unambiguously decreases with δ. However, at $\delta = \delta_{pm}$, the economy moves from (p, m) to $(p, 0)$, which implies that the level of GHG emissions due to commuting is given by

$$C_{\rm po} = \delta \left(\frac{L^2}{12\delta^2} + \frac{2K^2}{3t^2} \right).$$

Once more, a change in the intercity structure generates an upward jump in commuting pollution.[12]

When δ keeps rising, the CBD grows at the expense of the SBDs. Under these circumstances, $C_{\rm po}$ decreases for $\delta < \sqrt{2}\delta_p$ but increases when δ rises from $\sqrt{2}\delta_p$ to $2\delta_p$. Because jobs relocate in the CBD at the expense of the SBDs in response to higher population density, the average distance traveled by commuters may increase. This result might come as something of a surprise because it says that making a polycentric city more compact need not be good for the environment, *even when the city remains polycentric*. We are thus far from the famous "compact is always better."

Finally, when δ reaches the threshold $2\delta_p$, the SBDs vanish, meaning that city 1 becomes monocentric. At $\delta = 2\delta_p$, we have $C_{\rm po} = C_{\rm mo}$ where

$$C_{\rm mo} = \frac{L^2}{4\delta}.$$

---

[9]Indeed, $dC_{\rm pp}/d\delta < 0$ if and only if $\delta \leq \delta_p/\sqrt{2}$. Because $\delta_p/\sqrt{2} > \delta_m/3$ when $K < \bar{K}$, we have $dC_{\rm pp}/d\delta < 0$ as long as the economy involves two identical polycentric cities $(\delta < \delta_m/3)$.

[10]Indeed, we have $C_{\rm pp} < C_{pm}$ when $\delta \leq \delta_m/3$.

[11]The value of λ_{pm}^* can be determined from case (iii) in the Appendix B by solving $\Delta_{pm}V(\lambda) = 0$.

[12]This is because $C_{pm} < C_{\rm po}$ over the interval $\delta_m/3 \leq \delta \leq \delta_{pm}$.

In this case, increasing further the population density leads to lower pollution.

The entire equilibrium path is described in Figure 5. It reveals an interesting and new result: although increasing the population density is likely to reduce GHG emissions when the city size remains unchanged, *the resulting change in urban structure might well raise the GHG emissions stemming from commuting.* In particular, because the minimum value of C_{pm} over $(\delta_m/3, \delta_{pm})$ exceeds the maximum value of C_{pp} over $(\delta_p, \delta_m/3)$, moving from (p,p) to $(\mathrm{p},0)$ through (p,m) leads to higher levels of commuting pollution. In other words, by affecting the urban system, a higher population density may have undesirable effects from the environmental viewpoint.

Insert Figure 5 about here

(ii) Shipping. Consider now GHG emissions generated by the transport of goods. Dispersion $(\lambda = 1/2)$ is the worst and agglomeration $(\lambda = 1)$ the best configuration: $T(1/2) > T(\lambda_{pm}^*) > T(1)$. Consequently, for $K < \bar{K}$, the recommendations based on commuting (C) and interregional shipping (T) do not point to the same direction. Specifically, when the city structure shifts from (p,m) to (p,0), the pollution generated by workers' commuting jumps upward, whereas the pollution stemming from shipping goods vanishes. In this event, it is a priori impossible to compare the various market outcomes, hence to determine the best ecological configuration. Yet, given the relative importance of commuting and other within-city trips in the global emission of carbon dioxides, we believe that the conclusions derived for the former case are empirically relevant.

4.3 Welfare and the environment

The above results suggest that the decentralization of jobs within cities should supplement a higher population density from the ecological standpoint. One may wonder what this recommendation becomes when it is evaluated at the light of a second-best approach in which the planner chooses the number and structure of cities $(\lambda^o, \theta_1^o, \theta_2^o)$.

At any given intercity distribution of firms (λ), the intra-urban allocation of firms maximizing global welfare is given by:

$$\theta_r^o = \frac{1}{3} + \frac{2\delta K}{3t L_r} < \theta_r^*. \tag{27}$$

Hence, starting from the market equilibrium (θ_r^*), *a coordinated decentralization of jobs within cities both raises welfare and decreases GHG emissions.* It is readily verified that socially optimal outcome implies that city r is polycentric if

$$\delta < \delta_r^o \equiv \frac{t L_r}{K}. \tag{28}$$

Let us now turn to the case where the intercity distribution of activities and the city structure are both endogenous. As in the foregoing, we restrict ourselves to the case of low communication

costs ($K < \bar{K}$). It is shown in Appendix E that the welfare optimum is given by (i) two identical polycentric cities when $\delta_o/3 > \delta$, (ii) two asymmetric cities when $\delta_{pm}^o > \delta > \delta_o/3$, (iii) one single polycentric city when $4\delta_p > \delta > \delta_{pm}^o$, and (iv) one single monocentric when $\delta > 4\delta_p$ (the expression of δ_{pm}^o is given in Appendix E). Because $\delta_o > \delta_m$ and $\delta_{pm}^o > \delta_{pm}$, the market need not deliver the optimal configuration. For example, the market sustains two asymmetric cities when $\delta_o/3 > \delta > \delta_m/3$ while the second-best optimum involves two identical polycentric cities. In addition, when $\delta_{pm}^o > \delta > \delta_o/3$, a single polycentric city is the equilibrium spatial configuration; the second-best optimum is given by a large polycentric city with a small monocentric city.

Consequently, as in Section 3, when firms and workers' locations are given, a marginal increase in δ is always ecologically and socially desirable. However, when the population density hike generates a new pattern of activities, the move may be detrimental to both objectives. For instance, when δ crosses $\delta_m/3$ from below and takes a value in $[\delta_m/3, \delta_o/3]$, pollution from commuting exhibits an upward jump (see Figure 5). Moreover, the market outcome involves two asymmetric cities (p,m), while the second-best optimum involves two identical polycentric cities. This shows that what we have seen in Section 3 remains valid when the morphology of the urban system is endogenous. Though incomplete, our analysis suggests that there is no systematic conflict between welfare and environmental objectives.

5 Conclusion

This paper has focused on a single facet of compact cities: the transport demand. Observe, however, that trips related to activities such as recreation and schooling have a less direct relation to the city structure than commuting, thus blurring the connection between compactness and GHG emissions. Our model, therefore, should be extended to account for the location of such facilities. Furthermore, we have left aside the role of population density in the emissions of carbon dioxides generated by home heating and air conditioning. For example, residential energy use accounts for another 20% of America's GHG emissions. Therefore, a housing sector should be grafted onto our setting to capture this additional facet of the problem. In the same vein, it should be recognized that high population densities generate congestion and other negative externalities that are likely to clash with the social norms prevailing in many developed countries. Another limit of our approach is the implicit assumption of "liquid housing" in that the population density may be increased at no cost. Accounting for adjustment costs in housing size would make the case for compact cities weaker. Finally, our planning approach should be compared to a decentralized mechanism in which cities are free to choose their land-use policies. Due to the lack of coordination between jurisdictions, one may expect more tension to occur

between the ecological and social welfare objectives.[13]

To sum up, our work is too preliminary for strong and specific policy recommendations. This work must be viewed as a first step toward a theory of an ecologically and socially desirable urban system. We believe, however, that our results are sufficiently convincing to encourage city planners and policy-makers alike to pay more attention to the various implications of urban compactness. Unless modal changes lead workers to use mass transport systems, compact and monocentric cities may generate more pollution than an urban system with polycentric dispersed cities. In addition, by lowering urban costs without reducing markedly the benefits generated by large urban agglomerations, the creation of secondary business centers may allow large cities to reduce GHG emissions while enjoying agglomeration economies. The future of China and India, among others, will be urban, and the land-use rules they choose will have a considerable impact on the world carbon footprint (Glaeser, 2011). Building tall cities is clearly part of the answer, but we contend that policy-makers should also pay attention to the structure and number of the megacities that will emerge.

References

[1] A. Bento, S. Franco, D. Kaffine, The efficiency and distributional impacts of alternative anti-sprawl policies, Journal of Urban Economics 59 (2006) 121-141.

[2] J. Brander, P. R. Krugman, A "reciprocal dumping" model of international trade, Journal of International Economics 15 (1983) 313-321.

[3] D. Brownstone, T. Golob, The impact of residential density on vehicle usage and energy consumption, Journal of Urban Economics 65 (2009) 91-98.

[4] J. K. Brueckner, Urban sprawl: Diagnosis and remedies, International Regional Science Review 23 (2000) 160-171.

[5] J. Cavailhès, C. Gaigné, T. Tabuchi, and J.-F. Thisse, Trade and the structure of cities, Journal of Urban Economics 62 (2007) 383-404.

[6] B. Copeland, S. Taylor, Trade and the Environment: Theory and Evidence, Princeton Univ. Press, Princeton, 2003.

[7] G.B. Dantzig, T.L Saaty, Compact city. A plan for a liveable urban environment, W.H. Freeman, San Francisco, 1973.

---

[13]In particular, as observed by Glaeser and Kahn (2010), more empirical work is needed to determine whether this lack of coordination pushes people from more to less regulated urban areas.

[8] G. Duranton, D. Puga, Micro-foundations of urban increasing returns: theory, in J.V. Henderson and J.-F. Thisse (Eds.), Handbook of Regional and Urban Economics. Volume 4, North Holland, Amsterdam, 2004, pp. 2063-117.

[9] C. Engel, J. Rogers, How wide is the border? American Economic Review 86 (1996) 1112–1125.

[10] C. Engel, J. Rogers, Deviations from purchasing power parity: Causes and welfare costs, Journal of International Economics 55 (2001) 29–57.

[11] European Environment Agency, Greenhouse gas emission trends and projections in Europe 2007. European Environment Agency Report No 5, COPOCE, European Union, 2007.

[12] Environmental Protection Agency, US Greenhouse Gas Inventory Report, 2011.

[13] Gaigné C., S. Riou and J.-F. Thisse, Are compact cities environmentally friendly? Discussion Paper N°8297, CEPR, 2011.

[14] C. Gaigné, I. Wooton, The gains from preferential tax regimes reconsidered, Regional Science and Urban Economics 41 (2011) 59-66.

[15] E.L. Glaeser, Triumph of the City, Macmillan, London, 2011.

[16] E.L. Glaeser, M.E. Kahn, Sprawl and urban growth, in J.V. Henderson and J.-F. Thisse (Eds.), Handbook of Regional and Urban Economics. Volume 4, North Holland, Amsterdam, 2004, pp. 2481-527.

[17] E.L. Glaeser, M.E. Kahn, The greenness of cities: carbon dioxide emissions and urban development, Journal of Urban Economics 67 (2010) 404-418.

[18] P. Gordon, H.W. Richardson, Are compact cities a desirable planning goal? Journal of the American Planning Association 63 (1997) 1-12.

[19] A. Haufler, I. Wooton, Competition for firms in an oligopolistic industry: The impact of economic integration, Journal of International Economics 80 (2010) 239-248.

[20] Kahn, M.E., Green Cities: Urban Growth and the Environment, Brookings Institution Press, Washington, DC, 2006.

[21] M.E Kahn, J. Schwartz, Urban air pollution progress despite sprawl: the "greening" of the vehicle fleet, Journal of Urban Economics 63 (2008) 775-787.

[22] I. Muniz, A. Galindo, Urban form and the ecological footprint of commuting. The case of Barcelona, Ecological Economics 55 (2005) 499-514.

[23] OECD, Highlights of the international transport forum 2008: transport and energy. The challenge of climate change. OECD Publishing, Paris, 2008.

[24] G.I.P. Ottaviano, T. Tabuchi, J.-F. Thisse, Agglomeration and trade revisited, International Economic Review 43 (2002) 409-436.

[25] J.M. Savin, L'évolution des distances moyennes de transport des marchandises. Note de synthèse du SES, Logistique, May-June 2000.

[26] L. Schipper, L. Fulton, Carbon dioxide emissions from transportation: Trends, driving forces and forces for change, in D.A. Hensher and K.J. Button (Eds.), Handbook of Transport and the Environment, Elsevier, Amsterdam, 2003, pp. 203-226.

[27] N. Stern, The economics of climate change, American Economic Review 98 (2002) 1-37.

[28] T. Tabuchi, Agglomeration and dispersion: A synthesis of Alonso and Krugman, Journal of Urban Economics 44 (1998) 333-51.

[29] D. Timothy, W. Wheaton, Intra-urban wage variation, employment location, and commuting times, Journal of Urban Economics 50 (2001) 338-366.

[30] J.-F. Thisse, Toward a unified theory of economic geography and urban economics, Golden Issue of Journal of Regional Science 50 (2010) 281-296.

[31] US Census Bureau, US Commodity flow survey, 2007.

6 Appendix A

Accounting for emissions stemming from shopping-trips and manufacturing production do not change the qualitative properties of our fundamental trade-off.

- If shopping malls, say, are located at the city outskirts, this means that a consumer located at x has to travel the distance $y_r - x = L_r/2\delta - x$ to go to the mall. If the number of shopping trips is given by $\alpha > 0$, $C(\lambda)$ must be supplemented by

$$M(\lambda) = 2\alpha \int_0^{y_1} \delta(y_1 - x)\mathrm{d}x + 2\alpha \int_0^{y_2} \delta(y_2 - x)\mathrm{d}x = \alpha C(\lambda)$$

and thus $C(\lambda) + M(\lambda) = (\alpha + 1)C(\lambda)$. As a result, accounting for shopping trips in (13) amounts to giving a higher weight to the component capturing the within-city emissions. Of course, this argument does not account for localized trips, such as those made to schools. As shown by Glaeser and Kahn, the length of these trips also depends on the density. Their total length decreases with δ but increases with λ, very much like C.

- The carbon footprint E of the urban system can be augmented by introducing emissions stemming from production. The total output is given by

$$Q_1^*(\lambda) + Q_2^*(\lambda) = \frac{L^2}{L+1}[1 - 2\tau\lambda(1-\lambda)].$$

Thus, production behaves like C and is minimized (maximized) when $\lambda = 1/2$ ($\lambda = 1$). It suggests that accounting for production in the carbon footprint of cities makes the case for dispersion stronger.

Appendix B

To see how things work, assume that $f(x) = 1 - kx$ with $k < \delta/L$ (hence, regardless the city size, we have $1 - k\bar{y}_r > 0$). The meaning of δ is as before while a low value of k means that the distribution is dispersed; it is uniform when $k = 0$.

The right endpoint of city r is now such that

$$\int_0^{\bar{y}_r} \delta(1 - kx)\mathrm{d}x = \frac{L_r}{2}$$

the solution of which is

$$\bar{y}_r = \frac{1 - \sqrt{1 - kL_r/\delta}}{k}.$$

It is readily verified that $\mathrm{d}\bar{y}_r/\mathrm{d}L_r > 0$ and $\mathrm{d}^2\bar{y}_r/\mathrm{d}L_r^2 > 0$ as well as $\mathrm{d}\bar{y}_r/\mathrm{d}\delta < 0$ and $\mathrm{d}\bar{y}_r/\mathrm{d}k > 0$. In other words, the spatial extension of city r increases with its population size and decreases with its population density, that is, a higher δ and a smaller k. Note also that $\bar{y}_r = L_r/2\delta$ when $k = 0$.

The urban costs are now given by

$$UC_r = \frac{t}{k}\left(1 - \sqrt{1 - kL_r/\delta}\right)$$

while the equilibrium wages and surplus are not affected by k. Standard calculations show that full agglomeration arises if and only if

$$\delta_m < \frac{kL}{2\left(1 - \sqrt{1 - kL/\delta}\right)} \equiv \Omega_A$$

and full dispersion if and only if

$$\delta_m > \delta\sqrt{1 - \frac{kL}{2\delta}} \equiv \Omega_D.$$

Note that $\Omega_D = \Omega_A = \delta$ when $k = 0$. In addition, since $k < \delta/L$ we have $\Omega_D > \Omega_A$. As a consequence, we have full dispersion when $\delta < \delta_m^D$ where δ_m^D is the solution to $\Omega_D(\delta) = \delta_m$ with

$$\delta_m^D = \max\left\{\left(kL + \sqrt{k^2L^2 + 16\delta_m^2}\right)/4, kL\right\},$$

partial agglomeration when $\delta_m^A > \delta > \delta_m^D$ where δ_m^A is the solution to $\Omega_A(\delta) = \delta_m$ with

$$\delta_m^A = 4\delta_m^2/(4\delta_m - kL),$$

and full agglomeration when $\delta > \delta_m^A$. It is straightforward to check that $\delta_m^A = \delta_m^D = \delta_m$ when $k = 0$. In addition, we have $\delta_m^A = \delta_m^D = 0$ when $t = 0$ and $\delta_m^A \to \infty$ and $\delta_m^D \to \infty$ when $t \to \infty$.

In the case of partial agglomeration, the spatial equilibrium is given by

$$\lambda^* = \frac{1}{2} + \frac{2\delta_m\sqrt{\Omega_D - \delta_m}}{L\delta k}.$$

The sum of distance traveled by workers is now given by

$$C(\lambda) = 2\sum_r \int_0^{\bar{y}_r} x(\delta - kx)\mathrm{d}x = \left(\frac{1}{2} - \frac{k\bar{y}_r}{3}\right)\delta\bar{y}_r^2.$$

This function reaches its minimum at $\lambda = 1/2$ while $C(\lambda) = L_r^2/4\delta$ when $k = 0$. It is straightforward to check that E is minimized at

$$\lambda^e = \frac{1}{2} + \frac{e_C}{e_T}\frac{\sqrt{Y}}{kLZ}$$

where

$$Y \equiv 1 - \frac{Lk}{2\delta} - \frac{e_C^2}{4e_T^2\delta^2 Z^2} \qquad Z \equiv [2 - \tau(L+2)]/(L+1).$$

It is readily verified that $\lambda^e = 1/2$ ($\lambda^e = 1$) if and only if δ is sufficiently small (large), whereas λ^e belongs to the interval $(1/2, 1)$ for intermediate values of δ. Consequently, the positive quadrant of the (t, δ)-space is now divided into several domains in which the market delivers either a good or a bad ecological outcome, very much as in the case of uniform densities. This is illustrated in Figure 1b where the market outcome minimizes pollution in Panels A and C while in panels B', D' and E, the market delivers a bad ecological outcome.

Figure 1b

Appendix C

When cities can be monocentric or polycentric, the welfare function becomes

$$
\begin{aligned}
W(\theta_1, \theta_2, \lambda) &= \lambda L S_1^* + L_2 S_2^* + \theta_1\lambda L(w_1^C - UC_1^C) + \theta_2(1-\lambda)L(w_2^C - UC_2^C) \\
&\quad + (1-\theta_1)\lambda L(w_1^S - UC_1^S) + (1-\theta_2)(1-\lambda)L(w_2^S - UC_2^S).
\end{aligned}
$$

Plugging (27) into this expression, we get the following cases.

(i) If $\delta > \delta_1^o$ where δ_1^o is given by (28), both cities must be monocentric and the social optimum is given by the maximizer of (17).

(ii) If $\delta_1^o > \delta > \delta_2^o$, city 1 must be polycentric and city 2 must be monocentric, which implies that W is given by

$$W_{pm}(\lambda) \equiv \left[(\varepsilon_a^o - \varepsilon_b^o \tau)\tau - \frac{2t}{3\delta}\right]\lambda^2 L - \left[(\varepsilon_a^o - \varepsilon_b^o \tau)\tau - \frac{t}{\delta} + \frac{2K}{3L}\right]\lambda L$$
$$+ \frac{(L+2)L}{2(L+1)^2} - \frac{tL}{2\delta} + \frac{\delta K^2}{3tL}.$$

The social optimum now involves an interior configuration (λ_{pm}^o) when $\delta < 2\delta_m^o/3$ and $\delta < \delta_{pm}^o$ with

$$\delta_{pm}^o \equiv \frac{t}{3(\varepsilon_a^o - \varepsilon_b^0 \tau)\tau - 2K/L}.$$

Note that $W_m(1/2) = W_{pm}(\lambda_{pm}^o)$ at $\delta = 2\delta_m^o/3$, whereas $W_{pm}(\lambda_{pm}^o) < W_m(1/2)$ when $\delta < 2\delta_m^o/3$.

(iii) If $\delta_2^o > \delta$, both cities are polycentric. Social welfare is thus given by

$$W_{pp}(\lambda) = \frac{3L(\varepsilon_a^o - \varepsilon_b^o \tau)\tau}{\delta}\lambda(\lambda - 1)(\delta - \delta_m^o/3) - \frac{2KL + L^2 t\delta}{6}.$$

Accordingly, dispersion maximizes social welfare when $\delta < \delta_m^o/3$. Note that $W_{pp}(1/2) = W_{pm}(\lambda_{pm}^o)$ at $\delta = \delta_m^o$ and $W_p(1/2) > W_{pm}(\lambda_{pm}^o)$ when $\delta < \delta_m^o$.

If dispersion $(\lambda = 1/2)$ is socially desirable from the welfare viewpoint, we have $\delta_1^o = \delta_2^o = 2\delta_p$ so that the two cities must be monocentric if $\delta > 2\delta_p$ and polycentric if $\delta < 2\delta_p$. Similarly, under agglomeration $(\lambda = 1)$, $\delta_1^o = 4\delta_p$ while $\delta_2^o = 0$. Thus, agglomeration arises as a monocentric city when $\delta > 4\delta_p$ or as a polycentric city when $\delta < 4\delta_p$. Last, $\delta_1^o > \delta > \delta_2^o$ holds if and only if $1/2 < \lambda^o < 1$. Consequently, welfare is maximized when the economy is characterized by (i) a single monocentric city if $\delta > \max\{\delta_m^o, 4\delta_p\}$; (ii) a single polycentric city if $\delta_{pm}^o < \delta < 4\delta_p$; (iii) two identical monocentric cities if $2\delta_p < \delta < \delta_m^0$; (iv) two identical polycentric cities if $\delta < \min\{\delta_m^o/3, 2\delta_p\}$; (v) one large polycentric city and one small monocentric city if $\delta_m^o/3 < \delta < \min\{2\delta_p, \delta_{pm}^o\}$.

Appendix D

Under dispersion $(\lambda = 1/2)$, we have $\delta_1 = \delta_2 = \delta_p$ where

$$\delta_p \equiv Lt/4K$$

so that the two cities are monocentric if $\delta > \delta_p$ and polycentric if $\delta < \delta_p$. Similarly, under agglomeration $(\lambda = 1)$, $\delta_1 = 2\delta_p$ while $\delta_2 = 0$. Thus, agglomeration arises within a monocentric city when $\delta > 2\delta_p$ or within a polycentric city when $\delta < 2\delta_p$. Last, $\delta_1 > \delta > \delta_2$ holds if and only if $1/2 < \lambda < 1$.

Case (i). Dispersion with two monocentric cities.

When $\delta < \delta_m$, Proposition 1 implies that $\lambda = 1/2$ is an equilibrium outcome once we restrict ourselves to monocentric cities. Note further that the condition $\delta > \delta_p$ also prevents a marginal

deviation to a polycentric city to occur because, in the neighborhood of $\lambda = 1/2$, city r remains monocentric. Hence, the market equilibrium involves two monocentric cities having the same size if and only if $\delta_p < \delta < \delta_m$. For such a configuration to arise, it must be that $\delta_p < \delta_m$, i.e. $K > \bar{K}$ with:

$$\bar{K} \equiv \frac{L\left(\varepsilon_a - \varepsilon_b \tau\right) \tau}{4}.$$

Case (ii). Agglomeration within a single monocentric city.

Consider now the case of agglomeration in a monocentric city ($\lambda = 1$). For this to arise, it must be that $\delta > 2\delta_p$. In this case, when some workers leave city 2 to city 1, the latter must be monocentric. Because $\Delta V(1) > 0$ when $\delta > \delta_m$, $\lambda^* = 1$ is a stable equilibrium if and only if $\delta > \delta_m$ and $\delta > 2\delta_p$.

Case (iii). Dispersion with one polycentric city and one monocentric city.

When $\delta_1 > \delta > \delta_2$, the utility differential with $\theta_1^* < 1$ and $\theta_2^* = 1$ is given by

$$\Delta_{pm} V(\lambda) \equiv 2\left[(\varepsilon_a - \varepsilon_b \tau)\tau - \frac{2t}{3\delta}\right]\lambda + \left[-(\varepsilon_a - \varepsilon_b \tau)\tau + \frac{t}{\delta} - \frac{4K}{3L}\right].$$

Because $\Delta_{pm} V(\lambda)$ is linear in λ, the equilibrium $1/2 < \lambda_{pm}^* < 1$ is stable if and only if $\Delta_{pm} V(1/2) > 0$ and $\Delta_{pm} V(1) < 0$ hold. The first condition is equivalent to $\delta < \delta_p$ whereas the second condition amounts to $\delta < \delta_{pm}$.

Case (iv). Agglomeration within a single polycentric city.

Agglomeration ($\lambda = 1$) in the polycentric city occurs if and only if $\delta_{pm} < \delta < 2\delta_p$. Note that $\delta_{pm} < 2\delta_p$ if and only if $K < 2\bar{K}$, which holds when communication costs are low, transport costs are high, or both. Otherwise, even though agglomeration in a monocentric city remains a possible outcome, agglomeration in a polycentric city is not an equilibrium.

Case (v). Dispersion with two polycentric cities.

When $\delta < \delta_2$, the corresponding utility differential, which requires $\theta_1^* < 1$ and $\theta_2^* < 1$, is given by

$$\Delta_{pp} V(\lambda) \equiv \frac{L(\varepsilon_a - \varepsilon_b \tau)\tau}{\delta}\left(\delta - \frac{\delta_m}{3}\right)\left(\lambda - \frac{1}{2}\right). \tag{D.1}$$

Dispersion with two polycentric cities is an equilibrium if $\delta < \delta_2$, which becomes $\delta < \delta_p$ when $\lambda = 1/2$. It remains to show that this configuration is stable. First, it must that the coefficient of λ is negative in (D.1), which amounts to $\delta < \delta_m/3$. Second, this configuration is stable against a marginal deviation to a monocentric city in, say, city 2 because, in the neighborhood of $\lambda = 1/2$, city 2 is polycentric because $\delta < \delta_p$. Therefore, the dispersed configuration with two polycentric cities is a stable equilibrium if and only if $\delta < \delta_m/3$ and $\delta < \delta_p$.

These results are summarized as follows. There exist five stable spatial configurations: (i) a single monocentric city when $\delta > \max\{\delta_m, 2\delta_p\}$; (ii) a single polycentric city when $\delta_{pm} < \delta < 2\delta_p$; (iii) two identical monocentric cities when $\delta_p < \delta < \delta_m$; (iv) two identical polycentric cities when $\delta < \min\{\delta_m/3, \delta_p\}$; (v) one large polycentric city and one small monocentric city when $\delta < \min\{\delta_p, \delta_{pm}\}$.

Appendix E

When cities can be monocentric or polycentric the second best allocation is the solution of the following program:

$$\text{Max } W(\theta_1, \theta_2, \lambda) = L_1 S_1^* + L_2 S_2^* + \theta_1 L_1(w_1^c - UC_1^c) + \theta_2 L_2(w_2^c - UC_2^c)$$
$$+ (1 - \theta_1) L_1(w_1^s - UC_1^s) + (1 - \theta_2) L_2(w_2^s - UC_2^s).$$

Plugging (27) into this expression, we get:

(i) if $\delta > \delta_1^o$ where δ_1^o is given by (28), both cities must be monocentric and the social optimum is given by the solution to (17);

(ii) if $\delta_1^o > \delta > \delta_2^o$, city 1 must be polycentric and city 2 must be monocentric, which implies that W is given by

$$W_{pm}(\lambda) \equiv \left[(\varepsilon_2^o - \varepsilon_1^o \tau)\tau - \frac{2t}{3\delta} \right] \lambda^2 L - \left[(\varepsilon_2^o - \varepsilon_1^o \tau)\tau - \frac{t}{\delta} + \frac{2K}{3L} \right] \lambda L$$
$$+ \frac{(L+2)L}{2(L+1)^2} - \frac{tL}{2\delta} + \frac{\delta K^2}{3tL}.$$

The optimum now involves an interior configuration (λ_{pm}^o) when $\delta < 2\delta_m^o/3$ and $\delta < \delta_{pm}^o$ with

$$\delta_{pm}^o \equiv \frac{t}{3(\varepsilon_2^o - \varepsilon_1^0 \tau)\tau - 2K/L}.$$

Note that $W_m(1/2) = W_{pm}(\lambda_{pm}^o)$ at $\delta = 2\delta_m^o/3$, whereas $W_{pm}(\lambda_{pm}^o) < W_m(1/2)$ when $\delta < 2\delta_m^o/3$.

(iii) if $\delta_2^o > \delta$, both cities must be polycentric, so that W is now given by

$$W_p = \frac{3L(\varepsilon_2^o - \varepsilon_1^o \tau)\tau}{\delta} \lambda (\lambda - 1)(\delta - \delta_m^o/3) - \frac{2KL + L^2 t\delta}{6}$$

Accordingly, dispersion maximizes social welfare when $\delta < \delta_m^o/3$. Note that $W_p(1/2) = W_{pm}(\lambda_{pm}^o)$ at $\delta = \delta_m^o$ and $W_p(1/2) > W_{pm}(\lambda_{pm}^o)$ when $\delta < \delta_m^o$.

If dispersion $(\lambda = 1/2)$ is socially desirable from the welfare viewpoint, we have $\delta_1^o = \delta_2^o = 2\delta_p$ so that the two cities must be monocentric if $\delta > 2\delta_p$ and polycentric if $\delta < 2\delta_p$. Similarly, under agglomeration $(\lambda = 1)$, $\delta_1^o = 4\delta_p$ while $\delta_2^o = 0$. Thus, agglomeration must arise within a monocentric city when $\delta > 4\delta_p$ or within a polycentric city when $\delta < 4\delta_p$. Last, $\delta_1^o > \delta > \delta_2^o$ holds if and only if $1/2 < \lambda^o < 1$. Consequently, welfare is maximized when the economy is characterized by (i) a single monocentric city when $\delta > \max\{\delta_m^o, 4\delta_p\}$; (ii) a single polycentric city when $\delta_{pm}^o < \delta < 4\delta_p$; (iii) two identical monocentric cities when $2\delta_p < \delta < \delta_m^0$; (iv) two identical polycentric cities when $\delta < \min\{\delta_m^o/3, 2\delta_p\}$; (v) one large polycentric city and one small monocentric city when $\delta_m^o/3 < \delta < \min\{2\delta_p, \delta_{pm}^o\}$.
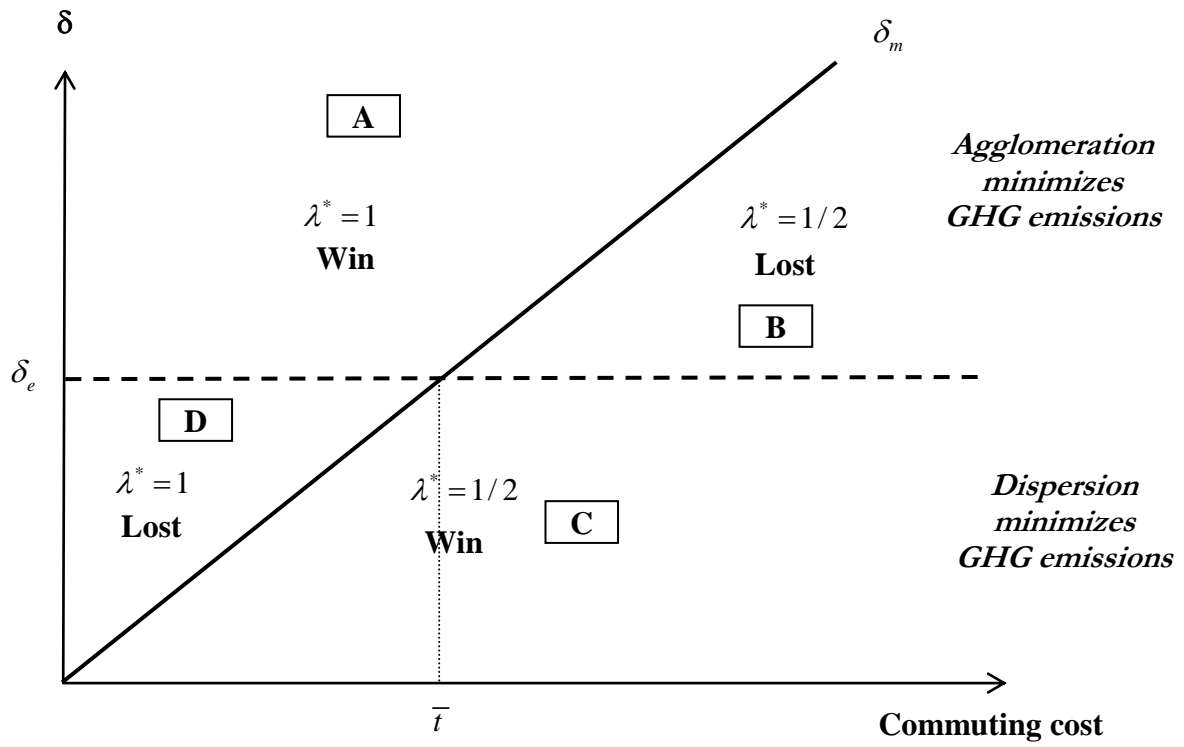
Figure 1.

δ

δ_m

A

$\lambda^* = 1$
Win

$\lambda^* = 1/2$
Lost

*Agglomeration
minimizes
GHG emissions*

B

δ_e

D

$\lambda^* = 1$
Lost

$\lambda^* = 1/2$
Win

C

*Dispersion
minimizes
GHG emissions*

$\bar{t}$

Commuting cost

Figure 1. Inter-city distribution and ecological outcome with monocentric cities

δ

δ_m^A

δ_m^D

A

Win

$\lambda^* = 1$

$1 > \lambda^* \geq 1/2$

*Agglomeration
minimizes
GHG emissions*

B'

*1 > λ^e > 1/2
minimizes
GHG emissions*

E

$1 > \lambda^* \neq \lambda^e > 1/2$

D'

$1 \geq \lambda^* > 1/2$

$\lambda^* = 1/2$

C

Win

*Dispersion
minimizes
GHG emissions*

Commuting cost
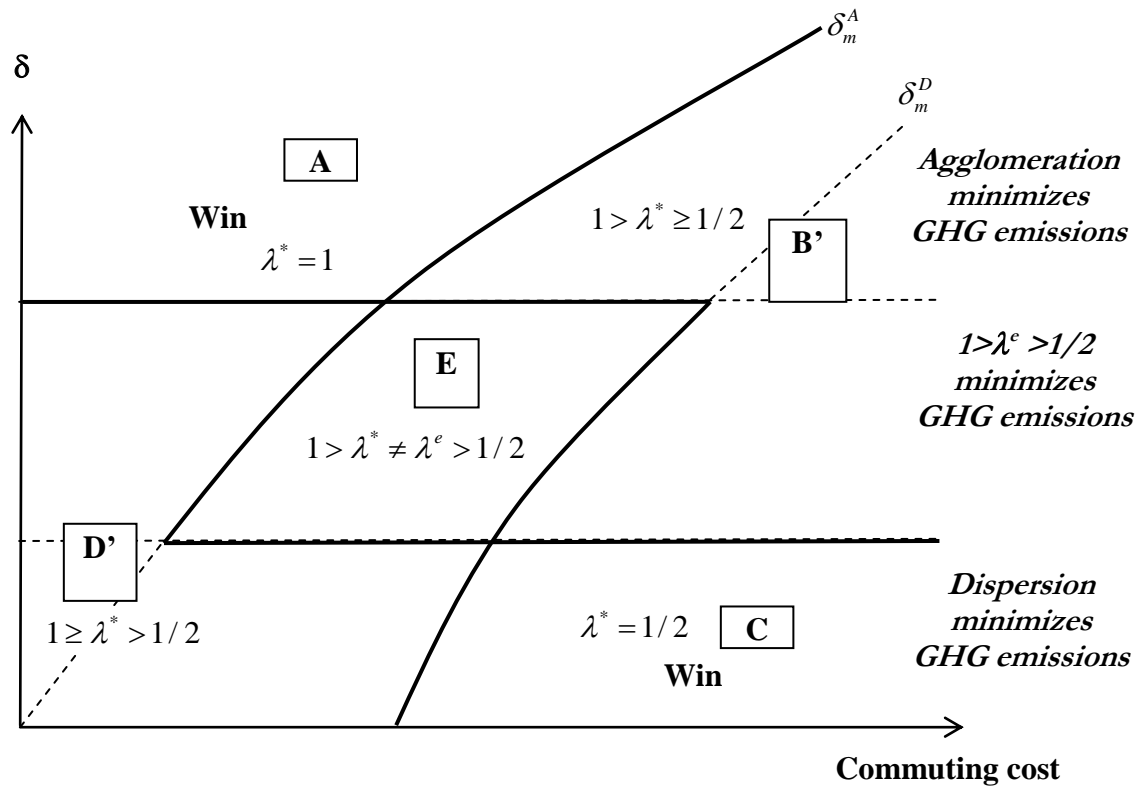
Figure 1bis. Inter-city distribution and ecological outcome with non-uniform density
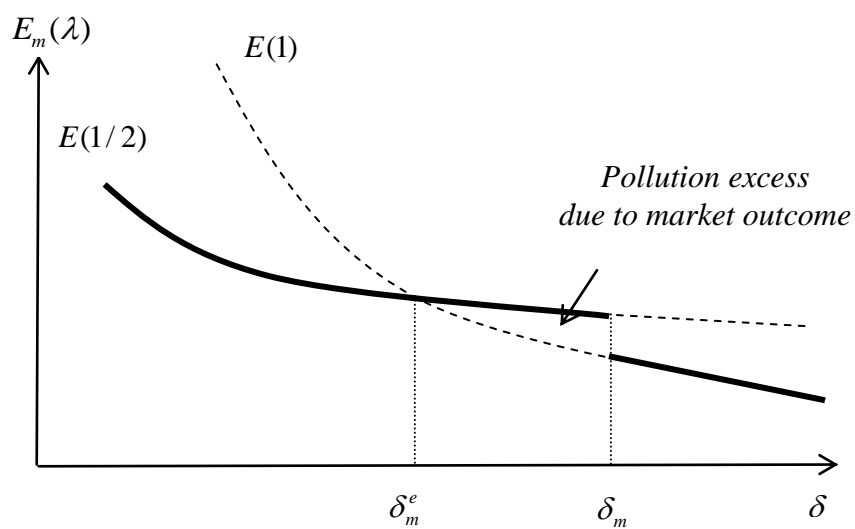
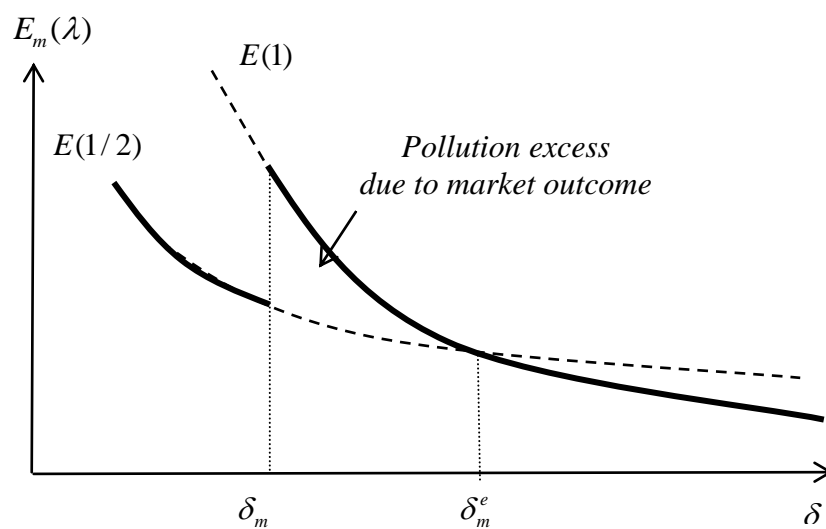**Figure 2a. Ecological and market outcomes when** $t > \bar{t}$

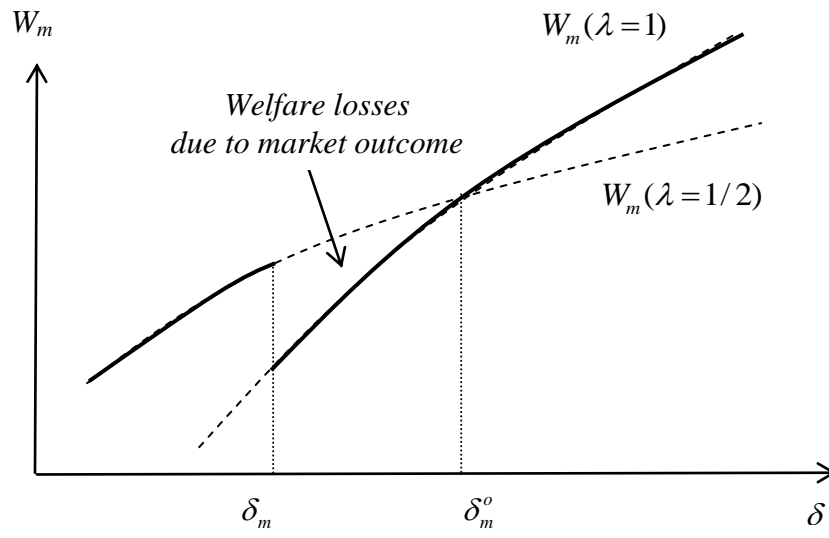**Figure 2b. Ecological and market outcomes when** $t < \bar{t}$

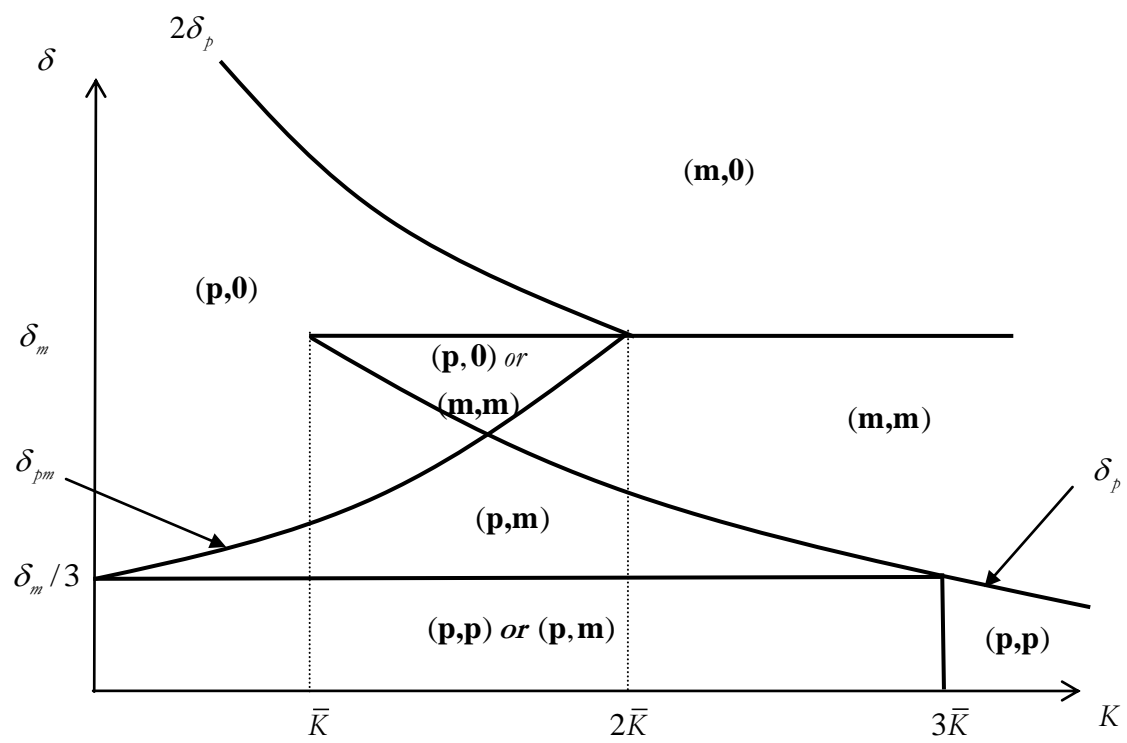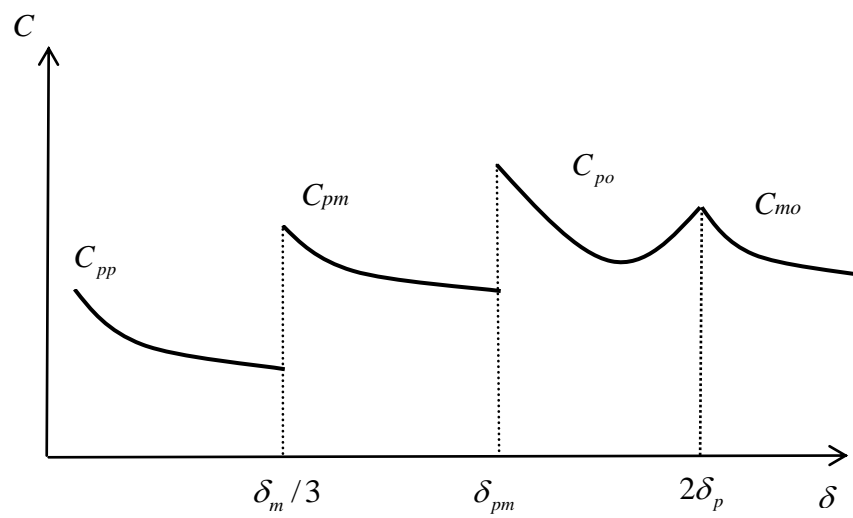**Figure 3. Market outcome and welfare**

Figure 4. The set of equilibria

Figure 5. Commuting pollution when $K < \bar{K}$