DISCUSSION PAPER SERIES

No. 4440

TRADE COSTS VERSUS URBAN COSTS

Jean Cavailhès, Carl Gaigné
and Jacques-François Thisse

INTERNATIONAL TRADE

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP4440.asp**

ISSN 0265-8003

TRADE COSTS VERSUS URBAN COSTS

Jean Cavailhès, INRA-CESAER
Carl Gaigné, INRA-CESAER
Jacques-François Thisse, CERAS, Paris; CORE, Université Catholique de Louvain and CEPR

Discussion Paper No. 4440
June 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 4440

June 2004

ABSTRACT

Trade Costs versus Urban Costs*

We analyse how the interplay between urban costs, wage wedges and trade costs may affect the inter-regional location of firms as well as the intra-urban location, within the central business district or in a secondary employment centre (SEC) of the selected region. In this way we investigate, on the one hand, how trade may affect the internal structure of cities and, on the other hand, how decentralizing the production and consumption of goods to subcentres changes the intensity of trade by allowing large metropolitan areas to maintain their predominance.

We show that, despite low commuting costs, SECs may emerge when the urban population is large and communication technologies are efficient, two features that seem to characterise modern economies. Moreover, when trade costs fall from high levels, the economy moves gradually from dispersion to agglomeration, favouring the formation of SECs. In an integrating world, however, the centre of a small monocentric city could be more attractive than subcentres of large polycentric cities. Nevertheless, the core retains its predominance through the relative growth of its main centre, which occurs at the expense of its subcentres.

JEL Classification: F12, F22, R12 and R14
Keywords: city structure, commuting costs, polycentric city, relocation and trade costs

Jean Cavailhès
UMR INRA-ENESAD
26 boulevard Docteur Petitjean
21079 Dijon cedex
FRANCE
Tel: (33 3) 8077 2583
Fax: (33 3) 8077 2571
Email: jean.cavailhes@enesad.inra.fr

Carl Gaigné
INRA-CESAER
26 Bd Dr Petitjean
21079 Dijon Cedex
FRANCE
Email: gaigne@enesad.inra.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159097

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159511

Jacques-François Thisse
CORE
Université Catholique de Louvain
34 Voie du Roman Pays
B-1348 Louvain-la-Neuve
BELGIUM
Tel: (32 10) 474 312
Fax: (32 10) 474 301
Email: thisse@core.ucl.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=104425

*We wish to thank Kristian Behrens and Vernon Henderson for useful comments and suggestions. This Paper is produced as part of a CEPR Research Network on ‘Specialization Versus Diversification: The Microeconomics of Regional Development and the Spatial Propagation of Macroeconomic Shocks in Europe’, funded by the European Commission under the Research Training Network Programme (Contract No: HPRN-CT-2000-00072).

Submitted 27 April 2004

1 Introduction

Cities are major actors in the process of trade (Bairoch, 1988). It is, therefore, fundamental to understand how the way they are organized - that is, their size and internal structure - impacts on trade flows and, conversely, how interregional flows influence the organization of cities. This is what we undertake here by studying this relationship through the interplay between trade costs and urban costs. Among other things, our approach will allow us to develop a better understanding of the growing role played by large metropolitan areas in modern economies.

In most developed countries, housing and commuting costs, which we call *urban costs*, account for more than one-third and sometimes as much as one-half of households' budgets.[1] Furthermore, the proportion of households' budgets spent on housing and commuting has increased substantially over the last few decades. Firms' performances are affected too by the level of urban costs. This occurs both directly through the land rent they pay to occupy central urban locations, and indirectly through the higher wages they pay their workers, who must be compensated for longer commutes. In addition, high urban costs act as entry barriers and lead the incumbents to charge higher prices. This makes local firms less competitive on home and foreign markets alike. It is surprising, then, that the economics profession has not paid more attention to those costs.

Despite the many advantages arising from agglomeration (Duranton and Puga, 2004), high urban costs give firms incentives to leave the main urban employment centers, and this is what firms have long been doing (Hohenberg and Lees, 1985). Firms (or developers) may then choose to form secondary employment centers, enterprise zones, or edge cities (Henderson and Mitra, 1996; Lucas and Rossi-Hansberg, 2002). As they enjoy living on larger plots and/or move along with firms, workers choose to live in suburbia or exurbia (Brueckner, 2000; Glaeser and Kahn, 2004). In this way, firms are able to pay lower wages and land rents while remaining within the metropolitan area - and so benefiting from agglomeration economies. In other words, the formation of subcenters within the same city, i.e. the formation of *a polycentric city*, appears to be a natural way to alleviate the burden of urban costs.

However, secondary employment centers are dependent on the urban center proper. Indeed, the higher-order functions (business services, research, banking and insurance,...) are mainly located in urban centers of cities. These functions seek out central positions and major city centers retain specific features relative to secondary employment centers. Hence, an access cost to the main center, which we call *communication costs*, is incurred by firms located in secondary employment centers. These costs may be crucial for the choice of firms to set up in a secondary employment center.

Yet, forming subcenters is not the only way for firms and workers to escape from high urban costs: they may choose to move into smaller, but cheaper, urban (or even rural) regions (see Holmes and Stevens, 2004, for the US and Gaigné *et al.*, 2003, for France). Such relocations may be profitable once *trade costs* have declined far enough to permit large-scale exports to distant regions. When this is the case, agglomeration economies and subcenters notwithstanding, the geographical

[1] See section 2 for more details.

concentration of activities within large metropolitan areas slows down and may even be superseded by the growth of small cities.

The foregoing observations suggest that the size and the structure of a city is dependent on its exchanges with other cities, as argued by many economic historians (Bairoch, 1988). Indeed, cities exporting a large proportion of their output, both because the firms located there are efficient and because trade costs are low, are likely to attract more firms and workers from other places. Such moves trigger an escalation of urban costs, which in turn prompts a redeployment of activities in a polycentric pattern, while smaller cities maintain their monocentric shape. However, we have pointed out that the urban structure of the city in which firms operate affects their competitiveness through the level of urban costs. The causality thus also runs the other way. As a result, by focusing on urban costs, we recognize quite naturally that both *agglomeration* and *dispersion* may take two quite separate forms because they are now compounded by *centralization* or *decentralization* of activities within the same city. Such a distinction is crucial for understanding the nature and evolution of trade.

In a nutshell, we may thus conclude that the interplay between urban costs, wage wedges, and trade costs is such that *firms* (and workers) *must choose an interregional location*, which belongs to a central or peripheral region, *as well as an intraurban location*, within the central business district or in a secondary employment center of the selected region. This is precisely the process we set out to study in this paper in which we borrow at will ingredients from different fields, such as trade theory, urban economics, and industrial organization. Specifically, we want to investigate, on the one hand, *how trade may affect the internal structure of cities* and, on the other, *how decentralizing the production and consumption of goods and housing to subcenters changes the intensity of trade* by allowing large metropolitan areas to maintain their predominance. As a by-product of our analysis, we will see that the market equilibrium may lead to a *hierarchical urban structure* involving one large, polycentric city together with small, monocentric ones.

To achieve this goal, we focus on urban costs as defined above because they both explain the formation of subcenters within a city and determine the strength of firms' incentives to move to less crowded regions, where wages and land rents are lower. We also consider a monopolistic competition setting that allows us to deal with increasing returns at the plant level as well as with positive variable mark-ups that vary with the number of competitors. As trade costs are positive, geography matters in our model in that cities are not floating islands.

Our contribution may be summarized as follows. Our model sheds light on some important concrete issues that have been pretty much overlooked until now. To begin with, we show that the internal structure of a city critically depends on the size of the labor force and the efficiency of communication technologies connecting firms to the strategic services established at the city center. In particular, we will see that a polycentric city may well be the only stable configuration when commuting costs are very low, an outcome that was unexplained (Anas *et al.*, 1998). This is so when the urban population is large and communication technologies are efficient, two features that seem to characterize our modern economies. This result points to one possible implication of the communications revolution.

When cities are open to trade, the equilibrium outcome ceases to be unique and highly contrasted patterns may arise. The organization of the space-economy varies with the capability of cities to accommodate a small or a large population in the monocentric arrangement. With a small population, we obtain the bell-shaped curve of spatial development. However, transitions are still of the bang-bang type. With a large population, the picture is quite different. When trade costs fall from high levels, the economy moves gradually from dispersion to agglomeration. The growth of the core region benefits its subcenters mostly. Once agglomeration within a polycentric city has been achieved, the core maintains its primacy over a large range of trade costs, thus confirming the intuition that *the polycentric structure fosters agglomeration*. However, when trade costs decline by a sufficiently large amount, some redispersion takes place, thus suggesting that, in an integrating world, the center of a small (monocentric) city could be more attractive than subcenters of large (polycentric) cities. Nevertheless, the core retains its predominance through the relative growth of its main center, which occurs at the expense of its subcenters. All together, these reults imply that *the existence of urban subcenters may have a substantial impact on the nature and intensity of interregional trade*. Turning to the impact of another type of costs, we will also show that the fall in commuting costs within cities may also affect the interregional distribution of activities.

Thus, it seems fair to say that our paper offers a synthesis of trade and urban system models. The internal structure of cities is endogenous and determined by the interactions between different types of spatial costs. Both consumption and production activities require land, whereas firms locate in the central business district or in subcenters.[2] As trade costs evolve, individual migrations appear to be governed by incentive schemes that vary with the internal structure of cities.

In the sections that follow, we first discuss the empirical evidence that justifies our paper (section 2), and then describe the modeling strategy we use (section 3). The city equilibrium is characterized (section 4), whereas the subsequent section analyzes an urban system of two cities whose size and structure are endogenous. Two principal scenarios are considered: in the first, trade costs keep falling (section 6); in the second, we study how interactions between two types of costs (workers' commuting costs and firms' communication costs between the central business district and subcenters) shape the space-economy (section 7). Section 8 concludes.

2 Urban Costs and City Structure

2.1 Urban costs and wages

In the United States, *shelter* accounts for 32.9% of household budgets and *transport and communications* for 19.3% (excluding local taxes and opportunity costs of time spent), making a total of 52.3% in 2001 (Bureau of Labor Statistics, 2003). It is reasonable, then, to say that more than 40% of the income of American households is spent on housing and transportation. Numbers for Europe are somewhat lower, but of a comparable order of magnitude. The average for the 15 European Union

[2] In their survey, Abdel-Rahman and Anas (2004) note that there is a need for a theory of urban systems in which the structure of cities need not be monocentric.

member states in 1994 was about 40% of income for *housing* and *transport and communication* (Winqvist, 1999).[3]

However, not all of this spending corresponds to urban costs as defined above, because this expenditure includes second homes and non work-related travel, whereas another fraction is location-independent. In addition, national averages fail to separate urban areas from rural ones, where residential costs are much lower. The analysis, therefore, needs to be refined.

The 1984 and 1996 INSEE housing survey for France can be used to pin down the points that interest us. We have calculated the share of (gross) income spent on rent and commuting costs.[4] The results show that rent took up 27% of income on average and increased by 70% over the period 1984–1996. Adding our estimate of commuting costs gives *a budget share of 40%* (61% up on 1984), which is similar to the foregoing evaluations.

As predicted by urban economics, urban costs increase with city size. This is confirmed by the dotted line of Figure 1 for 20 large US cities (Bureau of Labor Statistics, 2003). The correlation coefficient between urban costs and the log of the number of consumer units is 0.54. Urban costs are less than \$15,000 per year in cities like Pittsburgh, Baltimore and Kansas City but rise to nearly \$20,000 per year in San Francisco, Los Angeles and New York.

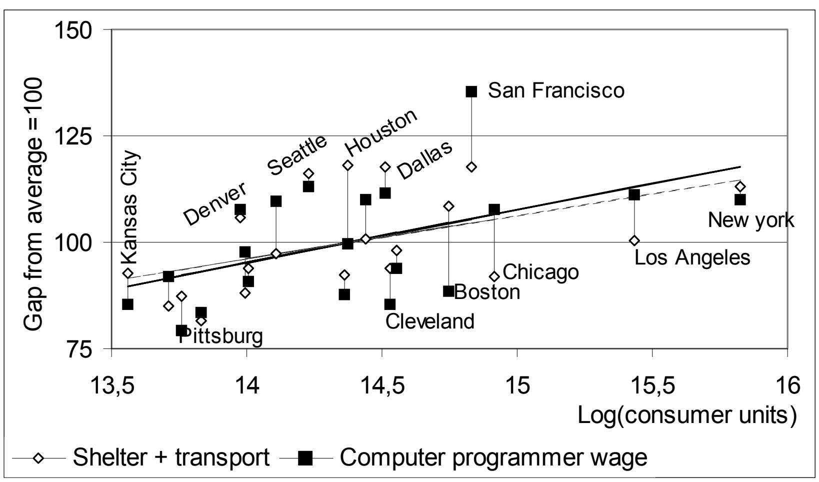

Figure 1. Urban costs and wages in the US

For France, we have used the INSEE housing survey to calculate that, in 1996, the rent per square-meter within the city of Paris was 84% higher than the national

[3] In France, housing and transport expenses have increased by 80% in household budgets between 1960 and 2000 (Rignols, 2002).

[4] The housing survey gives us the rent and annual income (only households that rent their property are included). The annual monetary cost of commuting is calculated as follows. One return trip for each of the 230 working days, times the number of workers in the household, yields the number of trips. This number is, in turn, multiplied by the crow fly distance between the residence and the center of the urban area, which is itself multiplied by 1.25 to get a more realistic estimation of the trips made by road. (In so doing, we assume that all the working population of an urban area works at its center. This entails a moderate overestimate of commuting trips.) The distance traveled is then weighted by the Revenue Service rate for 1984 and 1996 on a sliding scale with distance.

average and that rents were 28% below the national average in urban areas of cities with less than 100,000 inhabitants.

As urban economics also suggests, differences in urban costs associated with city size are offset by higher wages in big cities. Glaeser and Maré (2001, p. 317) report a close correlation between city size and wages in the US. However, when wages are set against the cost of living, this correlation collapses, which suggests that the high wages paid by firms are a form of compensation for the cost of living that is granted to workers so that their real income is the same everywhere. Figure 1 shows the wages of computer programmers, as a sample occupation, which vary in the same direction as urban costs (both variables are adjusted to a mean index of 100 to make the figure easier to read): wages rise with population at much the same rate as urban costs, the solid line regression curve being virtually superimposed on the dotted one for urban costs.

Shifting now to the intra-urban scale, using a wage equation that makes allowance for differences in human capital, Timothy and Wheaton (2001) report large variations in wages according to location: "wages are found to be as much as 15% higher in central Boston than in outlying work zones and (...) wage differentials of up 18% are found between central Minneapolis and the fringe counties" (p. 349-350). They further show that these differences are largely due to commuting costs.[5]

2.2 On polycentric cities

As said, when they grow bigger, metropolises spawn subcenters and/or sprawl out toward their hinterlands. For example, Giuliano and Small (1991) identify 29 job centers in Los Angeles, McMillen and McDonald (1998) find 15 in Chicago, and Creveso and Wu (1997) count 22 for the San Francisco Bay Area, thus leading Anas *et al.* (1998) to conclude that "polycentricity is an increasingly prominent feature of the landscape".[6]

The decentralization of jobs is a recent phenomenon in France, although somewhat older in the Ile-de-France Region, as Table 1 shows.

Table 1. Employment changes in Ile-de-France

	1968	**1979**	**1998**	**1968/1979**	**1979/1998**
Paris	1965	1957	1540	−0.4%	−21.03%
Inner belt	1468	1636	1695	+11,4%	+3.6%
Outer belt	873	1116	1486	+27.8%	+33%
Total Ile-de-France	4306	4708	4721	+9.3%	+0.2%

Source : Awada, 2001 (thousands of jobs)

After holding up through the 1970s, jobs in Paris fell by more than 20% in the 20 years that followed. Job numbers rose in the 1970s in the near suburbs and in

[5] Similar results hold for Paris. From the database built by Gaigné *et al.* (2003), it appears that wages are 12% lower in Paris fringe than in Paris for middle managers. The wage differential increases up to 30% for senior managers and takes its lowest value for manual workers (5%), which may be explained by the French law on minimum wage.

[6] Other studies of polycentric urban regions are reviewed in a special issue of *Urban Studies* edited by Kloosterman and Musterd (2001).

the more remote suburbs over both periods. The total rise was 70% in the last 30 years in the outer suburbs. However, this decentralization is not characterized by a wide scattering of firms. Indeed, some 75% of employment in the outer suburbs was concentrated in 11% of the communes, which drained most of the growth in jobs and now accommodate about one-third of the jobs of the outer suburbs, in decentralized employment centers (Lartigue and Petit, 2000). Hence, firms form local clusters when they are not in the CDB.

Moreover, as shown by Table 2, a peak in land rent around Paris confirms the existence of subcenters. The rent is about 25% lower in the belt located from 5 to 9 kilometers from Paris and it is 33% lower for the $10-14$ km belt. Interestingly, the rent rises further (from 104euro/m^2 for $10-14$ km to 122 for $15-19$ km). This higher rent is likely to be due to the presence of subcenters. Beyond, the rent falls again and reaches very lower values for remote locations.

Table 2. Housing rent according to the distance from Paris

Distance (km)	Rent (euro/m^2)
0	155
$5-10$	118
$10-15$	104
$15-20$	122
$20-24$	96
> 24	100

Source: INSEE, Housing survey (1996)

What was it that led to the emergence of polycentric urban structures? Three reasons at least can be given.

1- The increase in city population. In France, the population of the Paris Region more than doubled between 1921 and 1999. Over this period, Paris has lost inhabitants (-27%), whereas the population trebled in the remainder of the region. Suburbanization of the population in the Ile-de-France Region since 1975 has outstripped the decentralization of jobs with the suburban population increasing by 29.4% and jobs by 15.6%.

2- The increase in commuting costs. In France, for direct monetary costs, distance traveled has increased and unit cost has also risen slightly. In all, for the Ile-de-France Region, commuting costs have risen since 1976, as Table 3 shows.

Table 3. Commuting costs in the Ile-de-France Region

	1976	1991	1976-1991
Journey length (km)	7.9	9.8	+24%
Cost per km	1.77	1.97	+11%
Direct annual monetary cost	6432	8880	+38%
Journey time (minutes)	35	38	+9%

Sources: Baccaïni (1997) and French Revenue Service (FF 1997)

3- The fall in firms' communication costs. Anas *et al.* (1998) record that by the end of the 19th century telephones and trucks were making it possible for firms in the

US to decentralize. Lartigue and Petit (2002) emphasize the impact of telecommunications, where supply has exploded in the last 15 years or so: all major firms and 73% of small and medium-sized firms have (usually broadband) Internet connections thanks to the development of a dense fiber-optic network. Accordingly, communication costs have fallen dramatically, making location in the remote periphery of Paris less of a disadvantage.

However, secondary employment centers remain dependent on the urban center proper. As observed by Anas *et al.* (1998, p.1442): "subcenters have not eliminated the importance of the main center (...): downtown has the more total employment, higher employment density, and usually a larger statistical effect on surrounding densities and land prices than does any subcenter". This is confirmed by Schwartz (1993) who observes that about half of the business services consumed by US firms located in suburbia are supplied in city centers. In the case of New York, Los Angeles, Chicago and San Francisco, this figure even grows to 65%. The same is true of France, as can be seen from the distribution of higher-order metropolitan functions (executives, engineers, and business service company management jobs, research, commerce, banking and insurance, art). These functions are more common in city centers than in their periphery. For example, for the Paris urban area, they make up 19.3% of employment within Paris itself, 15.7% in the suburbs, and 6.6% in the periurban belt (Julien, 2002). These higher-order functions seek out central positions and major city centers retain specific features relative to secondary employment centers. *This means that firms in secondary employment centers incur an access cost to the main center when they resort to these higher urban functions*. Even if this cost is likely to have sharply fallen with the reduction in communication costs, allowance still has to be made for it.

To sum-up, two types of spatial structure seem to emerge. Either all firms form a central business district (in short CBD), in which case the city is monocentric, or some firms are prompted to constitute a secondary employment center, thus making the city polycentric. The market outcome depends on the population size, the level of commuting cost, and the access cost to the CBD.

3 The Model

3.1 The spatial economy

Consider an economy with *two regions*, labelled 1 and 2, *one sector* and *two factors*, labor and land. The economy is endowed with L mobile workers. The welfare of a worker depends on her consumption of the following three goods. The first good is unproduced and homogenous.[7] It is assumed to be costlessly tradeable and chosen as the numéraire. The second good is produced as a continuum of N varieties of a horizontally differentiated good under monopolistic competition and increasing returns, using labor and land as inputs and enjoying technological spillovers. Any variety of this good can be shipped from one region to the other at a unit cost of $\tau > 0$ units of the numéraire regardless of the variety, where τ accounts for all the

[7] The model can easily be extended by introducing a second sector producing the homogenous good under constant returns and perfect competition, using an immobile factor.

impediments to trade. This implies that both cities are assumed to be anchored and separated by a given physical distance.[8] The third good is land and is perfectly immobile.[9]

Ideally, we should explain why a central business district, where local public goods and business-to-business services are supplied, is formed in each region. However, this would render our analysis much more involved from the technical point of view, without adding much to our understanding of the space-economy. Indeed, the reasons for CBDs to arise have been well explored in the literature (Duranton and Puga, 2004) and we have nothing to add about this question. For the sake of simplicity, we thus choose to assume that each region has a *given* CBD located at the origin of a one-dimensional space, denoted X. The total amount of land in each region is large and its density per location is equal to one. Without loss of generality, we assume that the opportunity cost of land is zero. Finally, throughout the paper, we focus on the right-hand side of the city, the left-hand side being perfectly symmetrical.

Firms established in region r are agglomerated within clusters that form employment centers. It is well known that a prominent reason for such a concentration of producers lies in the supply of nontradeable business-to-business services provided at the center of the cluster - think of banks, insurances, law and auditing firms (Duranton and Puga, 2004). As the consumption of such services typically requires face-to-face communications, the distance to the cluster center matters to firms. However, we do not wish to introduce explicitly such services in our model because this would render its analytical treatment very cumbersome without adding much to the understanding of the phenomena we want to capture here. Hence, to keep the analysis as simple as possible, we subsume all such effects through the following (admittedly ad hoc) *spatial benefit function* that enters firms' profit functions in a way that will be described below:

$$S_r(x) = A + B\lambda_r - t_f x \tag{1}$$

where A, B and t_f are nonnegative constants, $\lambda_r \in [0, 1]$ the share of firms located in region $r = 1, 2$ (with $\lambda_1 + \lambda_2 = 1$), and x the distance to the cluster center.[10] This expression also captures the idea that the number of intermediate services varies from one region to another with the number of firms located in each city and operating in the final sector (Abdel-Rahman and Fujita, 1990).[11]

Throughout this paper, we will focus on configurations in which firms located in region r are clustered in the CBD. However, some firms might choose to set up

[8] Thus, we differ from Fujita *et al.* (1999). Note, however, that our setting also differs from theirs in two other aspects. First, our cities have a spatial extension and may be polycentric. This allows us to study how the urban structure of cities may affect trade, and vice versa. Second, we do not appeal to simulations to solve our model.

[9] Thus, our economy encompasses the whole range of possibilities: one good is freely mobile, one good is imperfectly mobile, and one good is immobile.

[10] Admittedly, exponential spatial decay functions are more relevant. However, the use of such functions leads to complicated expressions that are solved by means of simulations, a way out we want to avoid (Fujita and Ogawa, 1982; Lucas and Rossi-Hansberg, 2002).

[11] The total number of firms being fixed, (1) also means that the intensity of spillovers increases with the number of firms set up in region r.

in the suburbs of the metro, where they form a Secondary Employment Center (in short SEC), which is also the source of local benefits described by (1). As will be shown, the reason for the formation of a SEC is the fact that the land rent prevailing in the city outskirts is much lower than in the CBD, thus allowing firms to pay lower rents and lower wages. It remains to explain why and how a SEC is different from a CBD. We have seen in section 2 that a CBD supplies high level services that not available in a SEC. As a consequence, when some firms set up in a SEC, they must incur a cost $K > 0$ for using such services. It is natural to assume that this cost, which we call *communication cost*, decreases with the distance from the CBD. For analytical simplicity, we will assume that it is zero for the CBD-firms and positive but distance-independent when firms are outside the CBD. This is a reasonable assumption as communicating mainly requires the building of facilities which have the nature of fixed costs. We will also assume that all the land between these two centers are occupied by firms and workers, so that the CBD and the SECs interact as they form together a metropolitan area (see Figure 2 for an illustration of the land rent pattern).

Because firms use land, *each cluster of firms has a spatial extension* whose size depends on the number of firms that belong to it. In what follows, the superscript C is used to describe variables related to the CBD, whereas S describes the variables associated with the SEC (if any). Distances and locations are expressed by the same variable x measured from the center of the CBD located at $x = 0$ whereas the center of the SEC is established at $x_r^S > 0$, the location of which is endogenous. Both the CBD and the SEC are surrounded by residential areas occupied by workers. Depending on the structural parameters of the economy, the city may be monocentric or polycentric. In each city, the market for the differentiated product is open in the CBD and in the SEC, where the price $p_r(i)$ of variety $i \in [0, N]$ is the same regardless of the place. This amounts to assuming that the intra-regional transport cost is negligible with respect to the inter-regional transport cost of the differentiated product.

Under these various assumptions, the location and the size of the SEC as well as the size of the CBD are endogenously determined. In other words, apart from the assumed existence of the CBD with its high-level services, *the internal structure of each metropolitan region is endogenous.* As creating a new subcenter requires some positive fixed cost (Glaeser and Kahn, 2004), we find it convenient to restrict ourselves to the case of two subcenters. It should be clear, however, that the extension to more SECs does not generate new major insights.

3.2 Firms

Technology in manufacturing is such that producing $q(i)$ units of variety i requires a given number ϕ of labor units and a plot of fixed size s_f.[12] Clearly, such a technology exhibits scale economies. There are no scope economies so that, due to increasing returns to scale, there is a one-to-one relationship between firms and varieties. Thus, it must be that $N = L/\phi$. The number of firms located (or varieties produced) in

[12] When a second sector is considered, we may assume that the production of $q(i)$ units of variety i requires $mq(i)$ units of the immobile factor. Without loss of generality, we may then set $m = 0$.

region r is such that $n_r = \lambda_r N$.

Denote by Π_r^C (resp., Π_r^S) the profit of a firm set up in the CBD (resp., the SEC) of region r. Let θ_r be the share of region r-firms located in the CBD of region r and, therefore, by $(1-\theta_r)/2$ the share of firms in each SEC. When the firm producing variety i is located in the CBD of region r, its profit function is given by:

$$\Pi_r^C(i) = I_r(i) - \phi w_r^C - s_f R_r^C(x) + A + B\lambda_r - t_f x \tag{2}$$

where $I_r(i)$ describes the firm's revenue earned from regional sales and from exports, w_r^C the wage prevailing in the CBD in region r, and $R_r^C(x)$ the land rent to be paid in the CBD at a distance x from its center.[13]

When the firm is set up in the SEC of the same region, its profit function becomes:

$$\Pi_r^S(i) = I_r(i) - \phi w_r^S - s_f R_r^S(\left|x - x_r^S\right|) + A + B\lambda_r - t_f \left|x - x_r^S\right| - K \tag{3}$$

where w_r^S and $R_r^S(|x - x_S|)$ are respectively the wage and the land rent in the SEC, whereas the firm's revenue and lot size are the same as in the CBD.

3.3 Workers

Each worker living in region $r = 1, 2$ consumes the same, fixed amount $s_h > 0$ of land, a variable quantity $q(i)$ of variety $i \in [0, N]$, and a variable quantity q_0 of the numéraire. She is endowed with one unit of labor and $\overline{q}_0 > 0$ units of the numéraire. The initial endowment $\overline{q}_0$ is supposed to be large enough for her consumption of the numéraire to be strictly positive at the market outcome. Each worker commutes to the closer employment center and bears a unit commuting cost given by $t_h > 0$, so that a worker's commuting cost is either $t_h x$ or $t_h \left|x - x_r^S\right|$ according to the employment center. As jobs are equally distributed over each side of the CBD or of the SEC, each worker chooses her location and evaluates her commuting cost with respect to the center of the CBD or of the SEC. This implies that the wage earned by individuals working in the CBD (or in the SEC) is the same.[14]

The budget constraint of an individual residing at $x \in X$ in region $r = 1, 2$ and working in the corresponding CBD can then be written as follows:

$$\int_0^N p_r(i)q(i)di + q_0 + R_r^C(x)s_h + t_h x = w_r^C + \overline{q}_0$$

where $R_r^C(x)$ is the land rent prevailing at a distance x from the CBD. The budget constraint of an individual working in the SEC is obtained by replacing $t_h x$ by $t_h \left|x - x_r^S\right|$, $R_r^C(x)$ by $R_r^S(x)$, the land rent at a distance x of the SEC, and w_r^C by w_r^S, the wage in the SEC.

[13] In what follows, we assume that $A - n_r s_f t_f > 0$, thus allowing firms in the CBD to benefit positive spillovers regardless of its size.

[14] We see no evidence for a wage gradient defined over the CBD (or the SEC) and based on distance variations.

Preferences over the differentiated product and the numéraire are identical across workers and represented by a quasi-linear utility encapsulating a quadratic sub-utility:

$$U(q_0; q(i), i \in [0, N]) = \alpha \int_0^N q(i)di - \frac{\beta - \gamma}{2} \int_0^N [q(i)]^2 di - \frac{\gamma}{2} \left[\int_0^N q(i)di \right]^2 + q_0 \quad (4)$$

where $\alpha > 0$ and $\beta > \gamma > 0$. In this expression, α measures the intensity of preferences for the differentiated product with respect to the numéraire. The condition $\beta > \gamma$ implies that workers have a preference for variety.

3.4 Market structure

Solving the budget constraint for the numéraire consumption, plugging the corresponding expression into (4) and taking the first order conditions with respect to $q(i)$ yields

$$\alpha - (\beta - \gamma)q(i) - \gamma \int_0^N q(j)dj = p(i) \qquad i \in [0, N].$$

The demand for variety $i \in [0, N]$ by a worker living in region r can then be written as follows:

$$\begin{aligned} q_r(i) &= a - bp_r(i) + c \int_0^N [p_r(j) - p_r(i)]dj \qquad i \in [0, N] \\ &= a - (b + cN)\, p_r(i) + cP_r \end{aligned} \quad (5)$$

where $p_r(i)$ is the consumer price of variety i in region r,

$$P_r \equiv \int_0^N p_r(i)di$$

is the price index of the industrial sector in region r, and where $a \equiv \alpha/[(\beta+(N-1)\gamma]$, $b \equiv 1/[\beta + (N - 1)\gamma]$ and $c \equiv \gamma/(\beta - \gamma)[\beta + (N - 1)\gamma]$. Parameter a expresses the desirability of the differentiated product with respect to the numéraire and, therefore, measures the *total size* of this market; b gives the link between individual and industry demands: when b rises, consumers become more sensitive to price differences. Parameter c is an inverse measure of the degree of product differentiation between varieties; when $c \to \infty$, varieties are perfect substitutes, whereas they are independent for $c = 0$. These three parameters are independent of the region r.

Firm i faces a downward sloping demand in region r:

$$Q_r(i) = \lambda_r L q_r(i)$$

where $q_r(i)$ is given by (5). The consumer surplus of an individual living in region r is then as follows:

$$S_r = \frac{a^2 N}{2b} - a \int_0^N p_r(i)di + \frac{b + cN}{2} \int_0^N [p_r(i)]^2 di - \frac{c}{2} \left[\int_0^N p_r(i)di \right]^2. \quad (6)$$

As empirical evidence suggests that firms practice some form of spatial discrimination in pricing (Greenhut, 1981; Haskel and Wolf, 2001), we assume that markets are regionally segmented so that each firm chooses a delivered price which is specific to the region in which its variety is sold. By contrast, the price at which this variety is sold within region r is the same at both the CBD and the SEC (if any) of this region. Hence, the total revenue of firm i located in region r is given by

$$I_r(i) = p_{rr}(i)Q_r(i) + [p_{rs}(i) - \tau]Q_s(i).$$

Because there is a continuum of firms, each firm has a negligible impact on the market outcome in the sense that it can ignore its influence on, and hence reactions from, other firms. However, aggregate market conditions of some kind (here the price index P_r) affects any single firm. This defines a setting in which individual firms are not competitive (in the classic economic sense of having infinite demand elasticity) but, at the same time, they have no strategic interactions with one another. Because varieties are symmetric, all firms located in the same region charge the same price. As shown by Ottaviano *et al.* (2002), the equilibrium prices are as follows:

$$p^*_{rr} = \frac{1}{2}\frac{2a + c\tau(1-\lambda_r)N}{2b + cN} \tag{7}$$

$$p^*_{rs} = p^*_{ss} + \frac{\tau}{2} \qquad s \neq r. \tag{8}$$

Hence, the equilibrium price prevailing in a region decreases with the number of firms located there - the standard price competition effect of industrial organization - but it also increases with a - a result that may be interpreted as a *market size effect*, which is different from the home market effect (Ottaviano and Thisse, 2004). Note that factor costs (land rent and wage) do not enter (7)-(8) and, hence, have no direct influence the level of prices. However, they have the nature of an endogeneous fixed cost and, consequently, have a negative impact on the number of firms $\lambda_r N$ in region r, whence an indirect positive impact on equilibrium prices.

The equilibrium revenue of a firm located in r is given by

$$I_r = \lambda_r L p^2_{rr} + \lambda_s L \left(p_{rs} - \frac{\tau}{2}\right)^2$$

which in turn depends on the distribution of firms and workers between the two regions. As firms' prices net of trade costs are to be positive for any distribution of workers, we assume throughout this paper that

$$\tau < \tau_{trade} \equiv \frac{2a\phi}{2b\phi + cL}.$$

This condition also guarantees that it is always profitable for a firm to export to the other region.

To sum-up, we consider *a general equilibrium model involving labor, land as well as a differentiated product and a homogeneous good*. Increasing returns at the plant level are the agglomeration force and urban cost differential is the dispersion force, which shape the geography of the interregional economy. At the city level, spatial benefits and communication costs act as agglomeration forces, whereas commuting costs and land consumption by workers and firms are the dispersion forces. In the next section, we study the city equilibrium within one region as the outcome of the interplay between these various forces.

4 Decentralization Within a City

A *city equilibrium* is such that each worker maximizes her utility subject to her budget constraint and each firm maximizes its profits. In particular, no firm has an incentive to change place within the city, and no worker wants to change her working place and/or her residence. In this section, we characterize such an equilibrium taking as given the number of firms (n_r) and the number of workers ($l_r \equiv \lambda_r L$). For notational simplicity, we drop the subscript r because no ambiguity may arise.

4.1 Equilibrium land rents

Each worker chooses her location so as to maximize utility given her wage and the land rent in the metro; let $\Psi_h^C(x)$ and $\Psi_h^S(x)$ be the bid rent at $x \in X$ of a worker located respectively in the CBD and the SEC. Similarly, $\Psi_f^C(x)$ and $\Psi_f^S(x)$ stand for a firm's bid rents. Land is owned by absentee landlords who allocate their lot to the highest bidder. Because there is only one type of firms and one type of labor, at the city equilibrium it must be that

$$R(x) = \max\left\{\Psi_f^C(x), \Psi_h^C(x), \Psi_h^S(x), \Psi_f^S(x)\right\}.$$

These functions are illustrated in Figure 2, where the following notation is used (the value of these points is given in Appendix A.1 and A.2):

- x_f^C the right endpoint of the CBD;
- x_h^C the right endpoint of the residential area formed by those working in the CBD
- x_h^S the left endpoint of the residential area formed by those working in the SEC;
- x^S the center of the SEC;
- x_f^S the outer limit of the SEC;
- z_h the outer limit of the residential area formed by those working in the SEC. Workers' bid rent at z_h is equal to zero.

Because of the fixed lot size assumption, at the city equilibrium the value of the equilibrium consumption of the nonspatial goods

$$\mathbf{E} = \int_0^N p(i)q(i)di + q_0$$

is the same regardless of the worker's location (Fujita, 1989). Then, the budget constraint of an individual residing at $x \in X$ and working in the CBD may be rewritten as follows:

$$w^C - s_h R(x) - t_h x = \mathbf{E}$$

Similarly, the budget constraint of an individual working in the SEC becomes

$$w^S - s_h R(x) - t_h \left| x - x^S \right| = \mathbf{E}.$$

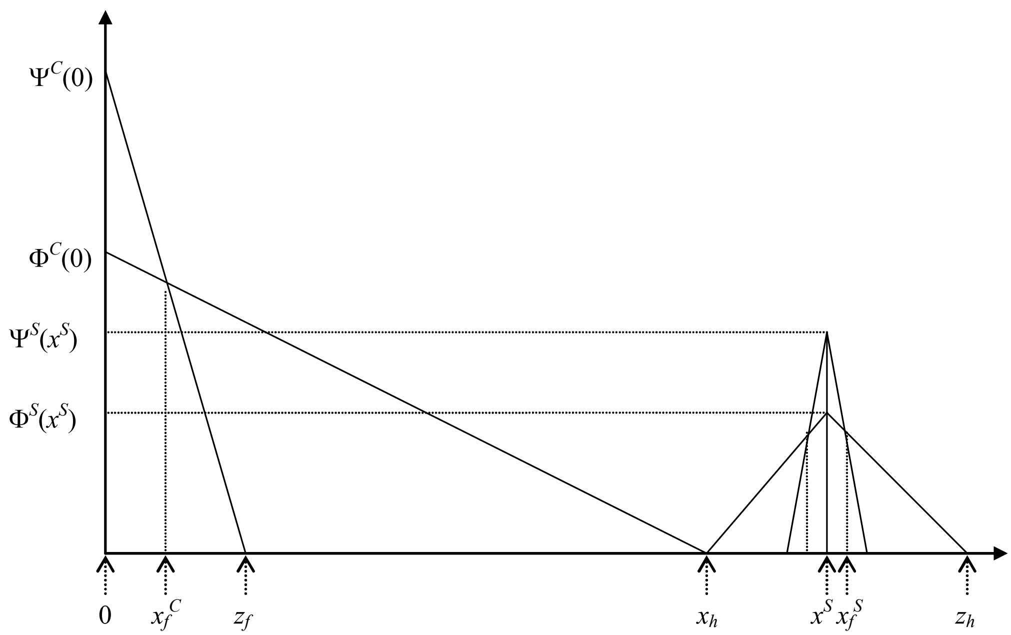

Figure 2. Urban equilibrium

As will be seen in section 4.2, the level of these wages is determined at the free entry equilibrium in which profits are zero, i.e. $\Pi^C(i) = \Pi^S(i) = 0$ for each firm located in region $r = 1, 2$.

At the city equilibrium, the CBD and SEC residential zones being contiguous, it must be that $x_h \equiv x_h^C = x_h^S$. Hence, the worker living at $x_h \equiv x_h^C = x_h^S$ is indifferent between working in the CBD or in the SEC and the workers' bid rents are just equal to the opportunity cost of land ($R_A = 0$) at x_h:

$$w^C - s_h R(x_h^C) - t_h x_h^C = w^S - s_h R(x_h^S) - t_h(x^S - x_h^S).$$

As $R(x_h^C) = R(x_h^S) = 0$ at x_h, we get:

$$w^C - w^S = t_h(2x_h - x^S). \tag{9}$$

Thus, *the difference in the wages paid in the CBD and in the SEC must compensate the worker for the difference in the corresponding commuting costs.* This wage wedge is always positive.

Using the expression of x_h (see Appendix A.1) and the fact that the bid rent of workers must be equal to zero at this point, we get:

$$\Psi_h^C(x) = \frac{1}{s_h}\left[\frac{\theta t_h}{2}(ns_f + ls_h) - t_h x\right] \tag{10}$$

where θ is the share of firms located in the CBD. The bid rent of the firms Ψ_f^C in the CBD is obtained as follows. The slope of this curve is $-t_f$ whereas Ψ_f^C must be equal to Ψ_h^C at x_f^C:

$$\Psi_f^C(x) = \frac{1}{s_f}\left[\frac{\theta s_f}{2}(nt_f + t_h l) - t_f x\right]. \tag{11}$$

Similarly, workers' and firms' bid rents around the SEC are maximized at x^S and intersect each other at x_f^S while the workers' bid rent is zero at z_h (see Appendix A.2 and Figure 2). Starting from $\Psi_h^S(z_h) = 0$, for $x > x^S$ we obtain:

$$\Psi_h^S(x) = \frac{1}{s_h}\left[\frac{(1-\theta)t_h}{4}(ns_f + ls_h) - t_h(x - x^S)\right]. \tag{12}$$

The bid rent of the firms set up in the SEC has a slope equal to $-t_f$ and is determined by solving $\Psi_h^S(x_f^S) = \Psi_f^S(x_f^S)$, thus yielding for $x > x^S$:

$$\Psi_f^S(x) = \frac{1}{s_f}\left[\frac{(1-\theta)s_f}{4}(nt_f + lt_h) - t_f(x - x^S)\right]. \tag{13}$$

It follows from (10) and (11) that the area around $x = 0$ is occupied by firms if and only if

$$\frac{t_f}{s_f} > \frac{t_h}{s_h}. \tag{14}$$

In other words, urban space is segregated with a "thick" CBD surrounded by a residential area. When this inequality does not hold, the equilibrium is given by a mixed configuration, as in Ogawa and Fujita (1980). Throughout this paper, we assume that (14) holds.

Note that the households' (resp., firms') bid rents $\Psi_h^C(x)$ and $\Psi_h^S(x)$ (resp., $\Psi_f^C(x)$ and $\Psi_f^S(x)$) in the CBD and the SEC are identical once the employment centers have the same size. As there are two SECs around the CBD, this arises at $\theta = 1/3$. The gap $\Psi_h^C - \Psi_h^S$ (resp., $\Psi_h^C - \Psi_h^S$) at any given x rises as the relative size of the CBD increases.

4.2 Equilibrium wages and the city structure

Wages are determined by the zero-profit condition associated with free entry and exit of firms in the city. Thus, setting (2) (resp., (3)) equal to zero, solving for w^C (resp., w^S) and plugging the corresponding expression into (9), we get:

$$w^C = \frac{1}{\phi}\left[I + A + B\lambda - \frac{\theta s_f}{2}(nt_f + lt_h)\right] \tag{15}$$

and

$$w^S = \frac{1}{\phi}\left[I + A + B\lambda - \frac{(1-\theta)s_f}{4}(nt_f + lt_h) - K\right]. \tag{16}$$

Substituting (15) and (16) into (9) and solving with respect to θ yields:

$$\theta^* = \min\left\{1, \frac{1}{3} + \frac{4K}{3\Lambda}\frac{\phi}{l}\right\} \tag{17}$$

where we have set

$$\Lambda \equiv s_f t_f + (s_h\phi + 2s_f)t_h\phi$$

which increases in each of its parameters. [15]

Observe that, in the limit case where $K = 0$, we have $\theta^* = 1/3$ because the region is formed by three identical employment centers. In addition, everything else being equal, *when the number of firms rises, the relative size of the CBD decreases* (but its absolute size rises) *whereas both the relative and absolute sizes of the SEC rises.*

As long as $\theta^* < 1$, the higher the communication costs, the larger the CBD. In the same way, the lower the commuting costs (t_h) or the spatial decay parameter affecting firms (t_f), the larger the CBD size. In the limit, the configuration becomes monocentric when $\theta^* = 1$. This arises if and only if the size of the labor force l does not exceed the threshold:

$$\bar{l} \equiv \frac{2\phi}{\Lambda}K. \tag{18}$$

The lower the communication costs, the lower this critical value. Likewise, the larger the land consumption by firms and workers, or the higher commuting or spatial decay parameter, the lower the population size for which the city is monocentric.

Another way to obtain a monocentric city is to bound from above the commuting cost: the city is always monocentric if and only if t_h does not exceed the threshold

$$\underline{t} \equiv \frac{2K\phi - s_f t_f l}{l(s_h\phi + 2s_f)\phi}. \tag{19}$$

Let

$$\overline{K} \equiv \frac{s_f t_f L}{2\phi} \tag{20}$$

so that $\underline{t}$ is strictly positive if and only if $K > \overline{K}$. Hence, the city is always monocentric when communication costs are sufficiently large, thus implying that *low commuting costs do not necessarily lead to a monocentric city.*

We summarize the main results of that analysis in the following proposition.

Proposition 1 *Assume that $t_f/s_f > t_h/s_h$. When $K/l > \Lambda/2\phi$, the city is monocentric regardless of the value of commuting costs. When $\Lambda/2\phi > K/l > s_f t_f/2\phi$, the city is monocentric if and only if $t_h \leq \underline{t}$. Otherwise, the city is polycentric.*

[15] In particular, we see from (17) that increasing returns are needed ($\phi > 0$) for a CBD to be larger than each SEC.

Observe that the urban equilibrium is always unique for a given population size. This property of uniqueness will vanish when mobility between cities is accounted for.

Computing the wage gap between the CBD and the SEC at the city equilibrium from (9) yields

$$w^C - w^S = t_h(ns_f + ls_h)\frac{3\theta^* - 1}{4} \geq 0.$$

This wage gap decreases as the commuting costs fall; in the limit, wages are equal when $t_h = 0$. The way workers split themselves between the CBD and the SEC is reminiscent of a spatial competition model à la Hotelling in which the two centers compete to attract workers instead of customers.

Finally, the indirect utility of an individual residing in region r and working in the corresponding CBD is given by

$$V_r^C(\lambda_r) = S_r + w_r^C - s_h R(x) - t_h x + \overline{q}_0 \tag{21}$$

where S_r is evaluated at the equilibrium prices (7)-(8). If she works in the SEC, her indirect utility becomes:

$$V_r^S(\lambda_r) = S_r + w_r^S - s_h R(x) - t_h \left| x - x^S \right| + \overline{q}_0.$$

By assumption, when $\lambda_r L$ workers live in region r, workers are distributed at the city equilibrium such that

$$V_r^C(\lambda_r) = V_r^S(\lambda_r) \equiv V_r(\lambda_r)$$

Likewise, when $\lambda_r N$ firms are established in region r, firms are distributed at the city equilibrium such that

$$\Pi_r^C(\lambda_r) = \Pi_r^S(\lambda_r) = 0.$$

5 Decentralization in a System of Cities

Consider now our two-region setting in which workers (and firms) are free to choose the region in which they want to be located. Let $\lambda \equiv \lambda_1$ so that $\lambda_2 = 1 - \lambda$. An *intercity equilibrium* arises at $0 < \lambda^* < 1$ when the utility differential $\Delta V(\lambda^*) \equiv V_1(\lambda^*) - V_2(\lambda^*) = 0$, or at $\lambda^* = 1$ when $\Delta V(1) \geq 0$, or at $\lambda^* = 0$ when $\Delta V(0) \leq 0$. By convention, when we consider an asymmetric equilibrium ($\lambda \neq 1/2$), we assume that region 1 is the larger region, which means that λ takes its values in $[1/2, 1]$.

When deciding whether or not to move, workers know if the cities of origin and destination are monocentric or polycentric; they also know the land rent that prevails in each one of them. Hence, in order to determine what an intercity equilibrium is, we must consider the three forms that the utility differential governing migrations may take. In the first one, both cities are polycentric. In the second, one city is monocentric whereas the other is polycentric. In the last one, both cities are monocentric. It is easy to see that the value of

$$\overline{\omega} = \frac{\overline{l}}{L} = \frac{2\phi K}{\Lambda L}$$

is critical, where $\bar{l}$ is given by (18). Indeed, each of the above-mentioned configurations is associated with a specific domain for the values of ϖ: (i) $\varpi < 1/2$, (ii) $1/2 \le \varpi < 1$, and (iii) $\varpi \ge 1$.

The most interesting situation arises in case (i) because it involves at least one polycentric city. It is, therefore, fully developed in this section. By contrast, in case (ii), at most one city is polycentric. Its detailed analysis is provided in Appendix B. Case (iii) means that both cities are always monocentric, which has been considered by Ottaviano *et al.* (2002).

In what follows, we determine the two forms that the utility differential $\Delta V(\lambda)$ may take when $\varpi < 1/2$ (see Appendix B for the case $\varpi > 1/2$). In other words, two types of spatial structure may prevail: both cities are polycentric (case A) and one city is polycentric whereas the second is monocentric (case B).

Case A. Assume that both regions are polycentric ($\theta_1^* < 1$ and $\theta_2^* < 1$), so that $\varpi < 1/2$ must hold. The corresponding equilibrium wages (w_r^C and w_r^S for $r = 1, 2$) are such that all firms, located either in the CBD or in the SEC of each region, earn zero profits (given, respectively, by (2) and (3)). More precisely, the equilibrium wages in region 1 are given by

$$w_1^C(\lambda) = \frac{1}{\phi}\left[I_1(\lambda) + A + B\lambda - (\lambda N t_f + \lambda L t_h)\left(\frac{s_f}{6} + \frac{2K}{3\Lambda}\frac{\phi}{\lambda L}s_f\right)\right] \tag{22}$$

and

$$w_1^S(\lambda) = \frac{1}{\phi}\left[I_1(\lambda) + A + B\lambda - K - (\lambda N t_f + \lambda L t_h)\left(\frac{s_f}{6} - \frac{K}{3\Lambda}\frac{\phi}{\lambda L}s_f\right)\right] \tag{23}$$

where

$$I_1(\lambda) = \frac{(b\phi + cL)L}{4(2b\phi + cL)^2\phi^2}\left\{[2a\phi + \tau cL(1-\lambda)]^2\lambda + [2a\phi - 2\tau b\phi - \tau cL(1-\lambda)]^2(1-\lambda)\right\}$$

is quadratic in λ; a similar expression holds for $w_2^C(\lambda)$ and $w_2^S(\lambda)$. In this case, the utility differential is as follows:[16]

$$\Delta_{pp}V(\lambda) = \delta_{pp}(\tau)(\lambda - 1/2) \tag{24}$$

where

$$\delta_{pp}(\tau) \equiv \varepsilon_1\tau^2 - \varepsilon_2\tau + \frac{\Omega}{3} + 12(2b\phi + cL)^2 B \tag{25}$$

with

$$\begin{aligned}
\varepsilon_1 &\equiv -3(b\phi + cL)(6b^2\phi^2 + 6b\phi cL + c^2L^2) < 0 \\
\varepsilon_2 &\equiv -12a\phi(b\phi + cL)(3b\phi + 2cL) < 0 \\
\varepsilon_3 &\equiv -6(2b\phi + cL)^2(\phi s_h + s_f + 1)\phi < 0 \\
\varepsilon_4 &\equiv -6(2b\phi + cL)^2 < 0
\end{aligned}$$

[16] Up to the multiplicative constant $L/[6\phi^2(2b\phi + cL)^2]$ which is the same in the three equations of motion.

and

$$\Omega \equiv \varepsilon_3 t_h + \varepsilon_4 t_f < 0.$$

which are all negative and independent of the communication costs K.

Remark. Note that, throughout this paper, we will encounter quadratic equations that differ from (25) only in the independent term. They all have two real and positive roots when the desirability of the differentiated good, measured by a, is above a certain threshold. Unless explicitly mentioned, we restrict ourselves to such cases.

Clearly, $\lambda^* = 1/2$ is always an equilibrium, which is stable if and only if $\delta_{pp}(\tau) < 0$. This is so when τ falls outside the two real and positive roots τ_{pp}^- and τ_{pp}^+ of the equation $\delta_{pp}(\tau) = 0$ given in Appendix A.3 for the case where $B = 0$.[17]

Lemma 1 *Assume that migration is governed by the utility differential $\Delta_{pp}V(\lambda)$. Then, the configuration formed by two polycentric cities is a stable equilibrium if and only if $\tau < \tau_{pp}^-$ or $\tau_{pp}^+ < \tau$. Otherwise, the economy is formed by a single polycentric city.*

Consider now the case where B is positive. Formally, this is equivalent to assuming that the independent term in (25) increases. Hence, as B rises, it is clear that τ_{pp}^+ increases and τ_{pp}^- decreases, thus implying that a single polycentric city emerges as the unique stable outcome for a wider range of trade cost values. In addition, when B takes a sufficiently large value, τ_{pp}^- becomes negative so that the economy never returns to dispersion as trade costs fall. This is because the gains generated by the spatial benefits always dominate those obtained by reducing urban costs through the existence of a second city. All of this agrees with the intuitive idea that *agglomeration is more likely when the intensity of spatial benefits augments with the size of the agglomeration.*

Case B. Assume now that city 1 is polycentric whereas city 2 is monocentric ($\theta_1^* < 1$ and $\theta_2^* = 1$).[18] This is consistent with both $\overline{\omega} < 1/2$ and $\overline{\omega} > 1/2$ (see Appendix B for the case $\overline{\omega} > 1/2$). The equilibrium wages (w_1^C, w_1^S and w_2^C) are such that no firm established in region 1 and located either in the CBD or in the SEC of this region, or established in the CBD of region 2 is able to make positive profits. More precisely, they are given by (22) and (23) for region 1 and by

$$w_2^C(\lambda) = \frac{1}{\phi}\left[I_2(\lambda) + A + B(1-\lambda) - \frac{s_f}{2}\left(t_f(1-\lambda)N + t_h(1-\lambda)L\right)\right] \tag{26}$$

for region 2, where

$$I_2(\lambda) = \frac{(b\phi + cL)L}{4(2b\phi + cL)^2\phi^2}\left\{(2a\phi + \tau cL\lambda)^2(1-\lambda) + (2a\phi - 2\tau b\phi - \tau cL\lambda)^2\lambda\right\}$$

is quadratic in λ. We now have:

$$\Delta_{pm}V(\lambda) \equiv \delta_1(\tau)\lambda + \delta_2(\tau) = \delta_1(\tau)\left[\lambda - \frac{-\delta_2(\tau)}{\delta_1(\tau)}\right] \tag{27}$$

[17] Those two roots are real and positive as long as $\varepsilon_2^2 - 4\varepsilon_1\Omega > 0$.

[18] In order to keep the analysis short, we focus on the sole case where $B = 0$.

with

$$\delta_1(\tau) \equiv \varepsilon_1\tau^2 - \varepsilon_2\tau + \frac{2}{3}\Omega \tag{28}$$

$$\delta_2(\tau) \equiv -\frac{1}{2}\left[\varepsilon_1\tau^2 - \varepsilon_2\tau + \Omega\left(1 - \frac{2}{3}\overline{\varpi}\right)\right]. \tag{29}$$

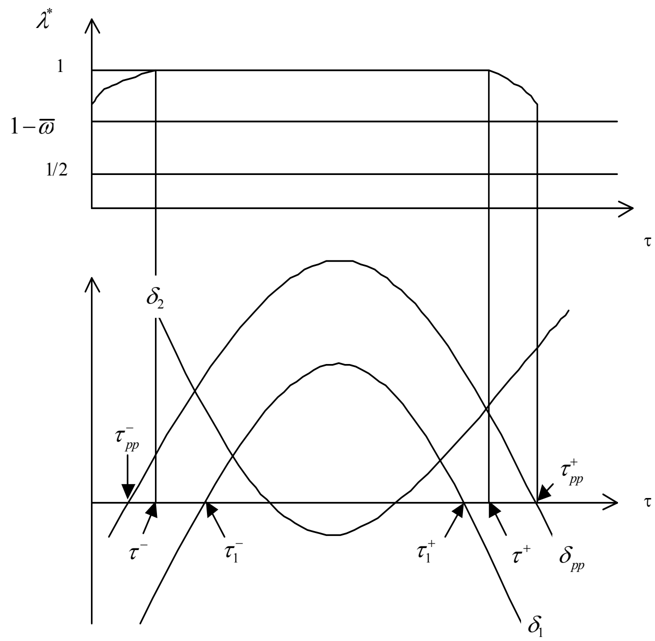

Figure 3. The case where $\overline{\varpi} < 1/2$.

The equation (27) has a steady-state given by $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$ if and only if $0 \leq -\delta_2(\tau)/\delta_1(\tau) \leq 1$. The condition $0 \leq -\delta_2(\tau)/\delta_1(\tau)$ means $\delta_1(\tau)$ and $\delta_2(\tau)$ have opposite signs. The condition $-\delta_2(\tau)/\delta_1(\tau) \leq 1$ has two different implications according to the sign of $\delta_1(\tau)$. First, when $\delta_1(\tau) > 0$, $-\delta_2(\tau)/\delta_1(\tau) \leq 1$ holds if and only if τ fall inside the two positive roots τ^- and τ^+ of $\delta_1(\tau) + \delta_2(\tau) = 0$ (which are given in Appendix A.3). Second, when $\delta_1(\tau) < 0$, $-\delta_2(\tau)/\delta_1(\tau) \leq 1$ holds if and only if τ fall outside the two positive roots τ^- and τ^+. In both cases, $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$ is an equilibrium, which is stable if and only if $\delta_1(\tau) < 0$. This in turn implies that $\delta_2(\tau)$ must be positive. Accordingly, τ must fall outside the two positive roots τ_1^- and τ_1^+ of $\delta_1(\tau) = 0$ and outside the two positive roots τ_2^- and τ_2^+ of $\delta_2(\tau) = 0$ (which are given in Appendix A.3 and illustrated in Figure 3).

When $\overline{\varpi} < 1/2$, the ranking of all these six roots is as shown in Figure 3. As a consequence, for $\overline{\varpi} < 1/2$, $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$ is a stable equilibrium if and only if $\tau > \tau^+$or $\tau < \tau^-$. Stated differently, when $\tau^- < \tau < \tau^+$ for $\overline{\varpi} < 1/2$, $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$ is not a stable equilibrium. When $0 < -\delta_2(\tau)/\delta_1(\tau) < 1$ is an interior stable equilibrium, the economy involves a polycentric city of size $-\delta_2(\tau)/\delta_1(\tau) > 1 - \overline{\varpi}$ as well as a monocentric city of size smaller than $\overline{\varpi}$.[19] Indeed, (28) and (29) implies

[19] Observe that this equilibrium changes with the value of trade costs.

that $-\delta_2(\tau)/\delta_1(\tau) > 1-\overline{\omega}$, which is in turn equivalent to $(1-2\overline{\omega})\delta_1(\tau) < 0$. Finally, this inequality holds when $\tau > \tau^+$or $\tau < \tau^-$. Hence, we have:

Lemma 2 *Assume that migration is governed by the utility differential $\Delta_{pm}V(\lambda)$ and that $\overline{\omega} < 1/2$. The configuration formed by a polycentric city and a monocentric city is a stable equilibrium if and only if $\tau < \tau^-$ or $\tau^+ < \tau$. Otherwise, the economy is formed by a single polycentric city.*

Observe, finally, that $\delta_1(\tau)$, $\delta_2(\tau)$ and $\delta_1(\tau)+\delta_2(\tau)$ are also described by parabolas, which have the same axis of symmetry at $\tau = \varepsilon_2/2\varepsilon_1 > 0$ as (25) (see Figure 3). In particular, $\delta_1(\tau)$ and $\delta_1(\tau)+\delta_2(\tau)$ are both represented by concave parabolas with a maximum at $\tau = \varepsilon_2/2\varepsilon_1$, whereas $\delta_2(\tau)$ is represented by a convex parabola with a minimum at $\tau = \varepsilon_2/2\varepsilon_1$. It is readily verified that the following inequalities always hold: $\tau^+_{mm} < \tau^+_1 < \tau^+_{pp}$ and $\tau^+_{mm} < \tau^+_2 < \tau^+_{pp}$. Consequently, the *pp*-parabola includes the parabola corresponding to (28).

The intercity equilibrium depends on the relative position of these four parabolas through their respective roots. More precisely, when $\overline{\omega} < 1/2$, we have the following ranking:

$$\tau^-_{pp} < \tau^- < \tau^-_1 < \tau^-_2 < \tau^+_2 < \tau^+_1 < \tau^+ < \tau^+_{pp}$$

These inequalities will be used below to describe the path followed by the economy when trade costs decrease.

6 The Impact of Trade Costs

In this section, we consider the usual thought experiment of economic geography, namely the impact of falling trade costs on the interregional distribution of activities. In our setting, we may also consider the impact of decreasing trade costs on the distribution of firms and workers *within each city* as the regional population changes.

6.1 The case for polycentricity

Consider first the case where $\overline{\omega} < 1/2$ and assume that $B = 0$ in (1).As long as trade costs exceeds τ^+_{pp}, $\lambda^* = 1/2$ is a stable equilibrium, thus implying by Lemma 1 that *the economy is formed by two identical polycentric cities.*

When τ falls just below τ^+_{pp}, the equilibrium $\lambda^* = 1/2$ ceases to be stable. As a result, λ rises up to $1-\overline{\omega}$, which implies that region 1 is larger than region 2. As soon as λ exceeds $1-\overline{\omega}$, the smaller city becomes monocentric and the equation of motion is now given by (27). Since $\tau^+_2 < \tau^+_1 < \tau$, we have $\delta_1(\tau) < 0$ and $\delta_2(\tau) > 0$ so that $-\delta_2(\tau)/\delta_1(\tau)$ is positive but not necessarily smaller than 1. Then, the new stable equilibrium is given by

$$\lambda^* = \min\left\{1; \frac{-\delta_2(\tau)}{\delta_1(\tau)}\right\}$$

For τ belonging to the interval (τ^+, τ^+_{pp}), the equilibrium is such that $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$ (Lemma 2). Furthermore, as τ decreases within this interval, $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$ increases because

$$\frac{\partial(-\delta_2/\delta_1)}{\partial\tau} = \frac{2}{3\delta_1^2}(2\varepsilon_1\tau - \varepsilon_2)\Omega(\overline{\omega} - 1) \tag{30}$$

is negative as long as $\tau > \varepsilon_2/2\varepsilon_1$. Hence, as trade costs fall in the interval (τ^+, τ^+_{pp}), *the polycentric city grows whereas the monocentric city shrinks.* When τ tends to τ^+, λ^* tends to 1. Furthermore, over the whole interval (τ^-, τ^+), we always have $\delta_1(\tau) + \delta_2(\tau) > 0$, which implies that $\lambda^* = 1$ is a stable equilibrium. This means that *the entire population is concentrated within a single city, which is polycentric.*

Similarly, when τ belongs to (τ^-_{pp}, τ^-) the equation of motion remains (27) and the equilibrium is given by $\lambda^* = -\delta_2(\tau)/\delta_1(\tau)$. Using (30), we see that λ^* decreases as τ falls, but remains larger than $1 - \overline{\omega}$. When $\tau < \tau^-_{pp}$, city 2 being monocentric, we still have partial agglomeration within city 1, which is polycentric.

In the limit, when $\tau = 0$ we have

$$\lambda^*(0) = -\frac{\delta_2(0)}{\delta_1(0)} = \frac{3}{4}\left(1 - \frac{4\phi K}{3\Lambda L}\right)$$

which is larger than $1 - \overline{\omega}$. *The asymmetric urban structure remains a stable equilibrium because firms and workers save on the communications that are associated with the existence of two polycentric cities.*

Proposition 2 *Assume that $\overline{\omega} < 1/2$ and that the two cities initially have the same size. When trade costs exceed τ^+_{pp}, the economy involves two polycentric cities of equal size. Below this level but above τ^+, agglomeration in city 1, which is polycentric, is partial but grows as τ falls, whereas city 2 is monocentric. When trade costs take their values in the interval (τ^-,τ^+), there is complete agglomeration in city 1. Below τ^-, agglomeration in city 1 is again partial but it now shrinks as τ decreases but the small city remains monocentric until $\tau = 0$.*

This proposition, illustrated in Figure 3, has several main implications. To start with, it suggests that a gradual fall in trade costs gives rise to a bell-shaped curve of spatial development. However, the path followed by the economy is such that symmetry does not re-emerge (unlike what we will observe in the case where $\overline{\omega} > 1/2$). Specifically, agglomeration occurs over a wider range of values for the trade costs. *This triggers a hysteresis effect that now prevents the small city to recoup all the activities it had when trade costs were high.* Yet, the large city becomes smaller when trade costs fall by a large amount, but retains a sufficiently large share of the industry for the other city to keep its monocentric structure. Interestingly, our equilibrium path does not exhibit any discontinuity: *once the large city exists, its size grows and shrinks smoothly while remaining polycentric.* It is also worth noting that *an urban hierarchy emerges as an equilibrium outcome*, whereas (partial) agglomeration in one region is more likely to arise when the monocentric structure is more difficult to sustain as an equilibrium.

In addition to the standard description of the interregional distribution of firms and workers, our setting allows us to say how the internal structure of cities is affected by the fall in trade costs. First, when the large city grows, the share of firms located in its CBD decreases whereas both the absolute and relative sizes of the SEC rise (see (17)). In other words, *secondary employment centers expand when trade costs take intermediate values*. Second, when the large city loses some of its activities, its SEC appear to be more affected than its CBD, regardless of the value of the communication costs.

For large values of trade costs, the symmetric configuration is not the only stable equilibrium. There is another stable equilibrium in which one city is larger than the other. As trade costs start falling from large values, the initially large city grows and accommodates all activities once trade costs reach the value τ_{pp}^{+} because $\overline{\omega} < 1/2$ (see Lemma 1). *This multiplicity of equilibria arises because a city can become polycentric*, thus allowing firms and workers to pay lower land rents.

6.2 The case for monocentricity

The argument to follow when $\overline{\omega} > 1/2$ is almost the same as in the foregoing. It is developed in Appendix B and leads to:

Proposition 3 *Assume that $\overline{\omega} > 1/2$ and that the two cities initially have the same size. When trade costs exceeds τ_{mm}^{+}, the economy involves two monocentric cities of equal size. Below this level but above τ^{-}, the industry is fully agglomerated and the economy has a single polycentric city. When trade costs are lower than τ^{-}, the two cities are monocentric and have the same size.*

This proposition, illustrated in Figure 4, has two major implications. First, a gradual fall in trade costs gives rise also to a bell-shaped curve of spatial development, as in the foregoing. *This process is triggered by a change in the urban structure of one city from monocentric to polycentric, and vice versa*. Second, agglomeration arises over a domain larger than the interval $(\tau_{mm}^{-}, \tau_{mm}^{+})$ for which the space-economy with two monocentric cities is an unstable equilibrium. Indeed, our proposition shows that agglomeration is sustained over $(\tau^{-}, \tau_{mm}^{+})$. *The large city is able to maintain its primacy because some of its activities are decentralized in SECs*. Stated differently, we have *hysteresis* in urban structure for trade costs that are not too small. However, it vanishes when trade costs decrease by a sufficiently large amount: the large region cannot keep its dominant position and complete re-dispersion toward the smaller city occurs, the two cities becoming identical.

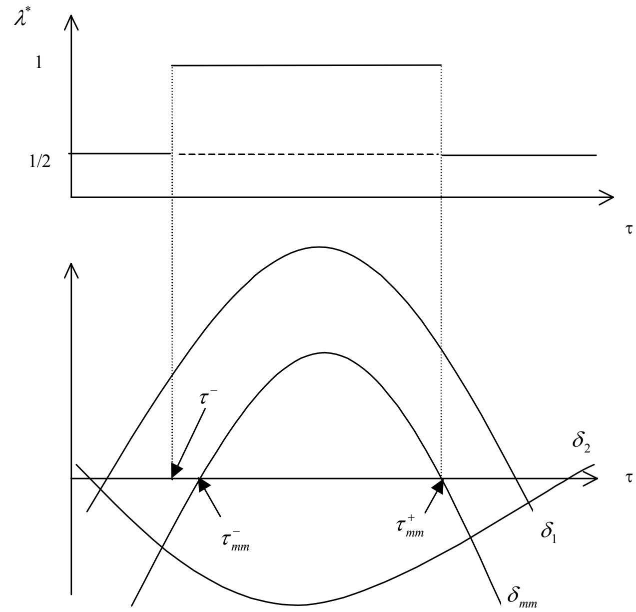

Figure 4. The case where $\overline{\omega} > 1/2$.

7 The Interaction between Commuting and Communication Costs

As discussed in section 2, the existence of high commuting and/or communication costs is one of the main reasons for the existence of polycentric cities as well as for the relocation of firms away from large cities. It is, indeed, very costly for firms to establish themselves within the CBD, so that firms (and workers) prefer to move away either in the suburbs when communication costs have decreased enough, or in a smaller city where wages are lower once trade costs have sufficiently declined.

In what follows, we restrict ourselves to the two polar cases in which communication costs are sufficiently high or low for each city to be monocentric or polycentric. A full analysis requires long algebraic developments that would not add much to our results. In particular, we show that *identical variations in cities' commuting costs may drastically change the interregional distribution of activities.* More precisely, as long as communication costs are sufficiently high (or low) and/or the labor force is sufficiently small (or large), we will see that the monocentric (or polycentric) structure of cities is unaffected by steadily decreases in commuting costs. Thus, when changes arise in the space-economy, what is affected is the interregional distribution between cities, not their intraurban structure.

7.1 The case of high communication costs

As a start, we consider the case in which communication costs are so high that both cities remain monocentric regardless of the interregional distribution of the industry, that is, when λ varies from 1/2 to 1. In other words, we have $\theta^*(\lambda, t_h) = 1$ for all $\lambda \in [1/2, 1]$, where θ^* is given by (17). The most stringent case arises when $\lambda = 1$. Cities are always monocentric when $t_h \leq \underline{t}$, where $\underline{t}$ is defined by (19) in which l has been replaced by L. For that to be possible, we need $K > \overline{K}$ where $\overline{K}$ is defined by (20).

When the two cities are monocentric, the utility differential expressed in terms of the commuting costs is as follows:

$$\Delta_{mm} V(\lambda) \equiv \delta_{mm}(t_h)(\lambda - 1/2)$$

where

$$\delta_{mm}(t_h) \equiv \varepsilon_3 t_h + (\varepsilon_1 \tau^2 - \varepsilon_2 \tau + \varepsilon_4 t_f)$$

As usual, $\lambda = 1/2$ is an equilibrium, which is stable when $\delta_{mm}(t_h) < 0$. Since $\varepsilon_3 < 0$, the stability of this equilibrium depends on the constant

$$\Theta_m \equiv \varepsilon_1 \tau^2 - \varepsilon_2 \tau + \varepsilon_4 t_f.$$

As in the foregoing, when $\Theta_m \leq 0$, that is, when τ is large or small, we have $\delta_{mm}(t_h) < 0$, so that dispersion with two monocentric cities prevails. By contrast, when $\Theta_m > 0$, that is, when trade costs take intermediate values, there exists a positive threshold

$$t_m \equiv \frac{\Theta_m}{-\varepsilon_3} > 0$$

such that $\delta_{mm}(t_h) < 0$ (resp., $\delta_{mm}(t_h) > 0$) if and only if $t_h > t_m$ (resp., $t_h < t_m$). We focus on the meaningful case of low commuting costs in which $t_m < \underline{t}$. As t_m increases with a whereas $\underline{t}$ is independent of a, this condition is satisfied when the parameter a does not exceed the unique value of the threshold a_m for which

$$\varepsilon_2 \tau = \varepsilon_1 \tau^2 + \varepsilon_3 \underline{t} + \varepsilon_4 t_f$$

holds:[20] the market size parameter cannot be too large to sustain a system of monocentric cities. Otherwise, when $a > a_m$ - hence $t_m > \underline{t}$ - the industry is always agglomerated in a single monocentric city. This form of extreme agglomeration arises because *the intensification of price competition that agglomeration brings about is itself lessened by a sufficiently large market size effect* ($a > a_m$).

The discussion above may be summarized as follows.

Proposition 4 *Assume that $t_h \leq \underline{t}$. When trade costs take intermediate values, the economy involves the agglomeration of the whole industry into a monocentric city if one of the following two conditions holds: (i) $t_h < t_m$ or (ii) $a > a_m$. Otherwise, the economy is formed by two monocentric cities.*

Thus, if trade costs take intermediate values, the industry is agglomerated in a single city when either commuting costs are very small ($t_h < t_m$) or the market size is sufficiently large ($a > a_m$). Note that decentralization within this city does not even arise. In such a context, rising commuting costs or reducing trade costs leads to the dispersion of industry under the form of two monocentric cities.

[20] Recall that ε_2 is linear in a whereas ε_1, ε_3 and ε_4 are independent of a.

7.2 The case of low communication costs

Assume now that communication costs are sufficiently low for both cities to be polycentric whatever the interregional distribution of the industry. In other words, we have $\theta^*(\lambda, t_h) < 1$ for all $\lambda \in [1/2, 1]$. The most stringent case arises at $\lambda = 1/2$. Hence, if $K \leq \overline{K}/2$, then cities are always polycentric. Observe that the existence of polycentric cities does not require any condition on the value of commuting costs.

In this case, the equation of motion (24) may be rewritten as follows:

$$\Delta_{pp} V(\lambda) = \delta_{pp}(t_h)(\lambda - 1/2)$$

where

$$\delta_{pp}(t_h) \equiv \frac{\varepsilon_3}{3} t_h + (\varepsilon_1 \tau^2 - \varepsilon_2 \tau + \frac{\varepsilon_4 t_f}{3}).$$

As $\delta_{mm}(t_h)$ and $\delta_{pp}(t_h)$ are almost identical, the argument is very similar to the one developed in the foregoing section. The results are not necessarily symmetric, however. Because the slope of $\delta_{pp}(t_h)$ is negative, the equilibrium distribution depends on the sign of the constant

$$\Theta_p \equiv \varepsilon_1 \tau^2 - \varepsilon_2 \tau + \frac{\varepsilon_4}{3} t_f$$

which exceeds Θ_m since $\varepsilon_4 < 0$. Clearly, when $\delta_{pp}(t_h) < 0$, dispersion with two polycentric cities prevails. By contrast, when $\Theta_p > 0$ there exists a positive threshold

$$t_p \equiv \frac{3\Theta_p}{-\varepsilon_3} > t_m > 0$$

such that $\delta_{pp}(t_h) < 0$ (resp., $\delta_{pp}(t_h) > 0$) if and only if $t_h > t_p$ (resp., $t_h < t_p$). Hence, the stable equilibrium is such that the industry is evenly distributed between two polycentric cities of equal size when $t_h > t_p$. As in the foregoing, when trade costs take intermediate values, the entire industry is concentrated in a single polycentric city if and only if $t_p > t_h$.

Hence, we may state:

Proposition 5 *When trade costs take intermediate values, the economy involves the agglomeration of the whole industry into a polycentric city if $t_h < t_p$. Otherwise, the economy is formed by two polycentric cities.*

When commuting costs are *large* and communication costs sufficiently low, firms and workers alleviate the burden of urban costs by having two polycentric cities: there is both dispersion and decentralization. In other words, the spatial organization of the production of the differentiated good leads to the lowest level of urban costs in the global economy. Here also, reducing trade costs leads to the dispersion of industry, which now takes the form of two polycentric cities.

However, when commuting costs are not "too" large ($t_h < t_p$) and trade costs take intermediate values, the industry is agglomerated in a single polycentric city. As in the foregoing, decreasing commuting costs fosters agglomeration. Yet, as t_p is larger than t_m, agglomeration can be sustained as an equilibrium over a larger domain of commuting costs. This means that *agglomeration is more likely to occur when decentralization through the creation of secondary employment centers is taken into account.*

Before concluding, observe that Propositions 4 and 5 have two other important implications. First, they imply that a fall in commuting costs in both cities fosters agglomeration in a single city. To show it, define

$$
\begin{aligned}
C_h^C(t_h) &\equiv s_h \Psi_h^C(t_h) + t_h x \\
C_h^S(t_h) &\equiv s_h \Psi_h^S(t_h) + t_h (x - x^s)
\end{aligned}
$$

as the urban costs borne by workers when they are respectively located around the CBD and the SEC. When cities are monocentric, it is readily verified that

$$
\left.\frac{d\Psi_f^C(t_h)}{dt_h}\right|_{\theta^*=1} = \frac{L}{2} > 0 \qquad \left.\frac{dC_h^C(t_h)}{dt_h}\right|_{\theta^*=1} = \frac{L(s_f + \phi s_h)}{2\phi} > 0
$$

when cities are polycentric ($K < \overline{K}$), some tedious calculations also show that

$$
\left.\frac{d\Psi_f^j(t_h)}{dt_h}\right|_{\theta^* \in [1/3,1)} > 0 \qquad \left.\frac{dC_h^j(t_h)}{dt_h}\right|_{\theta^* \in [1/3,1)} > 0 \qquad j = C, S
$$

Hence, in both cases, when commuting costs fall, the land rents paid by firms and the urban costs incurred by workers decrease. As a result, net wages increase in both types of configurations because firms are able to offer higher wages, whereas workers bear lower urban costs. Agglomeration in a single city may then arise for the following reason. For a given utility level, as commuting costs decrease, more workers and firms are willing to choose to locate in a given city. This larger concentration of workers and firms makes the agglomeration forces stronger, which in turn increases workers' utility so that all workers may end up in the same city.

Second, Propositions 4 and 5 reveal that *dispersion may prevail for intermediate values of trade costs, provided that commuting costs are sufficiently large.* Although economic geography models typically predict the agglomeration of activities when trade costs take intermediate values (Ottaviano and Thisse, 2004), this result tells us how the level of commuting cost may change the entire space-economy, thus showing how the local may affect the global.

8 Concluding Remarks

We have presented a model that shows how urban costs and trade costs may affect the location of economic activities *between and within* cities. Historical evidence shows that both trade and commuting costs have been decreasing since the beginning of the Industrial Revolution. Thus, what matters for the organization of the space-economy is the relative evolution of these two costs.

Our main conclusions may be organized around the following two ideas: (i) global forces may affect the local organization of production and employment, whereas (ii) local factors may well change the global organization of the economy. This interaction arises because the spatial decentralization of production and employment may take different forms as either a single polycentric city or two monocentric cities may emerge, thus yielding very contrasted economic landscapes.

Regarding the first idea, our thought experiment is about trade costs. When trade costs decrease steadily, agglomeration emerges gradually and becomes the more likely outcome. This is because the decentralization of jobs in a polycentric city allows firms to pay lower rents and wages. This result agrees with the existence of a huge megalopolis in which employment is decentralized in several centers that all belong to the same urban region. Interestingly, such an outcome may still prevail even when full dispersion is also a stable equilibrium, thus showing in a very neat way the prevalence of hysteresis in urban structures. Yet, eventually, the agglomeration becomes partial in that it loses jobs to the benefit of smaller cities. During this process, the internal structure of the megalopolis changes gradually in that the SECs first gain more jobs, whereas the CBD recovers some of its importance when trade costs take very low values.

Concerning the second idea, we focus on commuting costs. When these costs are high, the economic landscape is likely to be formed by cities having the same size and structure, but trading different varieties. This may be so even when trade costs take intermediate values. Unsuspectedly, when commuting costs decrease in parallel, eventually one city becomes dominant. This is because urban costs become sufficiently low for the benefits of agglomeration to be the stronger force. Once communication costs are sufficiently low, a single large polycentric city may then arise.

The interaction between the local and the global could be made more general by embedding our model within a broader setting in which firms also trade with the rest of the world, as in Krugman and Livas Elizondo (1996). Another line of research worth investigating is to introduce different commuting and communication costs across cities in order to study how local policies affect the global pattern of location and trade.

References

[1] Abdel-Rahman, H. and A. Anas (2004) Theories of Systems of Cities. In J.V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*. Amsterdam, North Holland, 2293-339.

[2] Abdel-Rahman, H. and M. Fujita (1990) Product Variety, Marshallian Externalities, and City Sizes. *Journal of Regional Science* 30, 165-183.

[3] Anas, A., R. Arnott and K.A. Small (1998) Urban Spatial Structure. *Journal of Economic Literature* 36, 1426-64.

[4] Awada, F., ed., (2001) *40 ans en Ile-de-France*. Paris, IAURIF.

[5] Baccaïni, B. (1997) Commuting and Residential Strategies in the Ile-de-France: Individual Behaviour and Spatial Constraints. *Environment and Planning A* 29,1801-29.

[6] Bairoch, P. (1988) *Cities and Economic Development: From the Dawn of History to the Present*, Chicago, University of Chicago Press.

[7] Brueckner, J. K. (2000) Urban Sprawl: Diagnosis and Remedies. *International Regional Science Review* 23, 160-71.

[8] Bureau of Labor Statistics (2003) *2001 Metropolitan Area Occupational Employment and Wage Estimates*. US Department of Labor, Occupational Employment Statistics. www.bls.gov/oes/2001.

[9] Cerveso, R. and K.L. Wu (1997) Commuting and Residential Location in the San Francisco Bay Area. *Environment and Planning A* 29, 865-86.

[10] Duranton, G. and D. Puga (2004) Micro-foundations of urban increasing returns: theory. In J.V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*. Amsterdam, North Holland, 2063-117.

[11] Fujita, M. (1989) *Urban Economic Theory. Land Use and City Size*. New York, Cambridge University Press.

[12] Fujita, M., P. Krugman and T. Mori (1999) On the Evolution of Hierarchical Urban Systems. *European Economic Review* 43, 209-51.

[13] Fujita, M. and H. Ogawa (1982) Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations. *Regional Science and Urban Economics* 12, 161-96.

[14] Gaigné, C., V. Piguet and B. Schmitt (2003) Changes in Rural vs. Urban Manufacturing Employment. A Shift-Share Analysis on French Data. CESAER Working Paper.

[15] Giuliano, G. and A. Small (1991). Subcenters in the Los Angeles Region. *Regional Science and Urban Economics* 21, 163-82.

[16] 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*. Amsterdam, North Holland, 2481-527.

[17] Glaeser, E.L. and D.C. Maré (2001) Cities and Skills. *Journal of Labor Economics* 19, 316-42.

[18] Greenhut, M.L. (1981) Spatial Pricing in the United States, West Germany and Japan. *Economica* 48, 79-86.

[19] Haskel, J. and H. Wolf (2001) The law of one price - A case study. *Scandinavian Journal of Economics* 103, 545-58.

[20] Henderson, V. and A. Mitra (1996) New Urban Landscape: Developers and Edge Cities. *Regional Science and Urban Economics* 26, 613-43.

[21] Hohenberg, P. and L.H. Lees (1985) *The Making of Urban Europe (1000-1950)*. Cambridge, MA, Harvard University Press.

[22] Holmes, T. and J. Stevens (2004) Spatial Distribution of Economic Activities in North America. In J.V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*. Amsterdam, North Holland, 2797-843.

[23] Julien, P. (2002) Onze fonctions pour qualifier les grandes villes. *INSEE Première* N°840.

[24] Kloosterman, C. and S. Musterd (2001) The Polycentric Urban Region: Towards a Research Agenda. *Urban Studies* 38, 623-33.

[25] Krugman, P. and R. Livas Elizondo (1996) Trade Policy and the Third World Metropolis. *Journal of Development Economics* 49, 137-150.

[26] Lartigue, S. and T. Petit (2002) *Dispersion ou polarisation de l'emploi dans la grande couronne d'Ile-de-France, 1979-1998*. Paris, IAURIF.

[27] Lucas, R.E. and E. Rossi-Hansberg (2002) On the Internal Structure of Cities. *Econometrica* 70, 1445-76.

[28] McMillen, D.P. and J. F. McDonald (1998) Population Density in Suburban Chicago: A Bid-Rent Approach. *Urban Studies* 35, 1119-30.

[29] Ogawa, H. and Fujita, M. (1980) Equilibrium Land Use Patterns in a Non-monocentric City, *Journal of Regional Science* 20, 455-75.

[30] Ottaviano G.I.P., T. Tabuchi and J.-F. Thisse (2002) Agglomeration and Trade Revisited. *International Economic Review* 43, 409-36.

[31] Ottaviano G.I.P. and J.-F. Thisse (2004) Agglomeration and Economic Geography. In J.V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*. Amsterdam, North Holland, 2563-608.

[32] Rignols, E. (2002) La consommation des ménages depuis quarante ans. *INSEE Première* N°832.

[33] Schwartz, A. (1993) Subservient Suburbia. *Journal of the American Planning Association* 59, 288-305.

[34] Timothy, D. and W.C. Wheaton (2001) Intra-Urban Wage Variation, Employment Location and Commuting Times. *Journal of Urban Economics* 50, 338-66.

[35] Winqvist, K. (1999) Le consommateur européen en 1994. *Eurostat, Statistique en bref, population et conditions sociales.*

Appendix A

1. Given the share θ of firms located in the CBD as well as their land consumption s_f, it must be that:

$$x_f^C = \frac{\theta n s_f}{2}$$

The value of x_h may then be obtained by adding to x_f^C the size of the residential area formed by the $\theta l/2$ individuals working in the CBD:

$$x_h = \frac{\theta n s_f}{2} + \frac{\theta l s_h}{2}$$

2. We have:

$$x^S = \frac{1+\theta}{4}(ns_f + ls_h) \qquad x_f^S = \frac{1}{2}ns_f + \frac{1+\theta}{4}ls_h \qquad z_h = \frac{1}{2}(ns_f + ls_h)$$

3. The equation $\delta_{pp}(\tau) = 0$ has two positive roots given by

$$\tau_{pp}^- = \frac{-\varepsilon_2 - \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega/3}}{-2\varepsilon_1} \quad \text{and} \quad \tau_{pp}^+ = \frac{-\varepsilon_2 + \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega/3}}{-2\varepsilon_1}$$

Likewise, the equation $\delta_1(\tau) = 0$ has two positive roots given by

$$\overline{\tau}_1 = \frac{-\varepsilon_2 - \sqrt{\varepsilon_2^2 - 8\varepsilon_1\Omega/3}}{-2\varepsilon_1} \quad \text{and} \quad \overline{\overline{\tau}}_1 = \frac{-\varepsilon_2 + \sqrt{\varepsilon_2^2 - 8\varepsilon_1\Omega/3}}{-2\varepsilon_1}$$

whereas $\delta_2(\tau) = 0$ has two positive roots given by

$$\tau_2^- = \frac{-\varepsilon_2 - \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega\left(1 - \frac{4\phi K}{\Lambda L}\right)}}{-2\varepsilon_1} \quad \text{and} \quad \tau_2^+ = \frac{-\varepsilon_2 + \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega\left(1 - \frac{4\phi K}{\Lambda L}\right)}}{-2\varepsilon_1}$$

and $\delta_1(\tau) + \delta_2(\tau) = 0$ has two positive roots given by

$$\tau^- = \frac{\varepsilon_2 - \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega\left(\frac{1}{3} + \frac{4\phi K}{\Lambda L}\right)}}{-2\varepsilon_1} \quad \text{and} \quad \tau^+ = \frac{\varepsilon_2 + \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega\left(\frac{1}{3} + \frac{4\phi K}{\Lambda L}\right)}}{-2\varepsilon_1}$$

Appendix B. Proof of Proposition 3

Migration dynamics when both cities are monocentric. Assume that both regions are monocentric ($\theta_r^* = 1$) so that $1/2 \leq \overline{\omega} < 1$ must hold. The equilibrium wages are given by a bidding process in which firms compete for workers by offering them higher wages until no firm can earn positive profits, given by (2), in the CBD of either region. Formally, the equilibrium wage w_1^C is given by (26). It is then readily verified that the utility differential is now as follows:

$$\Delta_{mm}V(\lambda) \equiv \delta_{mm}(\tau)(\lambda - 1/2) \tag{31}$$

where

$$\delta_{mm}(\tau) \equiv \varepsilon_1\tau^2 - \varepsilon_2\tau + \Omega + 12(2b\phi + cL)^2 B. \tag{32}$$

Clearly, $\lambda = 1/2$ is always an intercity equilibrium. As usual, this equilibrium is stable if and only if $\delta_{mm}(\tau) < 0$. Assuming that $B = 0$, this is so when τ falls outside the two real and positive roots of the equation $\delta_{mm}(\tau) = 0$ given by

$$\tau_{mm}^{-} = \frac{-\varepsilon_2 - \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega}}{-2\varepsilon_1} \quad \text{and} \quad \tau{+}_{mm} = \frac{-\varepsilon_2 + \sqrt{\varepsilon_2^2 - 4\varepsilon_1\Omega}}{-2\varepsilon_1}$$

which exist in so far as a, is sufficiently large for $\varepsilon_2^2 - 4\varepsilon_1\Omega > 0$ to be satisfied.

It is straightforward to check that the expressions (32) and (25) are described by two concave parabolas, which have the same axis of symmetry at $\tau = \varepsilon_2/2\varepsilon_1 > 0$ (see Figures 3 and 4). As the *pp*-parabola corresponding to (25) includes the *mm*-parabola corresponding to (32), the domain of τ-values for which dispersion prevails is wider when both cities are monocentric than when both are polycentric. This already points to the fact that *allowing for the decentralization of activities within the same city fosters the agglomeration of firms in this city.*

To sum-up:

Lemma 3 *Assume that migration is governed by the utility differential $\Delta_{mm}V(\lambda)$. Then, the configuration formed by two monocentric cities is a stable equilibrium if and only if $\tau < \tau_{mm}^{-}$ or $\tau_{mm}^{+} < \tau$. Otherwise, the economy is formed by a single monocentric city.*

As in the section 5, when B rises, τ_{mm}^{+} increases and τ_{mm}^{-} decreases so that a single monocentric city emerges as the unique stable outcome for a wider range of trade cost values. Note also that the same increase in B makes the case of agglomeration lesser than in Case 1 because a polycentric city allows workers and firms to bear lower urban costs.

Again, the intercity equilibrium depends on the relative position of these parabolas through their respective roots. It is readily verified that the following inequalities always hold: $\tau_{mm}^{+} < \tau_1^{+} < \tau_{pp}^{+}$ and $\tau_{mm}^{+} < \tau_2^{+} < \tau_{pp}^{+}$. Consequently, the *pp*-parabola includes the parabola corresponding to (28), which itself includes the *mm*-parabola. More precisely, when $\overline{\omega} > 1/2$, we have:

$$\tau_2^{-} < \tau_1^{-} < \tau^{-} < \tau_{mm}^{-} < \tau_{mm}^{+} < \tau^{+} < \tau_1^{+} < \tau_2^{+}$$

For $\overline{\omega} > 1/2$, $\lambda = -\delta_2(\tau)/\delta_1(\tau)$ is a stable equilibrium if and only if $\tau > \tau_2^{+}$ or $\tau < \tau_2^{-}$. Stated differently, when $\tau_2^{-} < \tau < \tau_2^{+}$ for $\overline{\omega} > 1/2$, $\lambda = -\delta_2(\tau)/\delta_1(\tau)$ is not a stable equilibrium. When $0 < -\delta_2(\tau)/\delta_1(\tau) < 1$ is an interior stable equilibrium, the economy involves a polycentric city of size $-\delta_2/\delta_1 > \overline{\omega}$ and a monocentric city of size smaller than $\overline{\omega}$.

The impact of trade costs when $\overline{\omega} > 1/2$. When trade costs are above τ_{mm}^{+}, the equilibrium $\lambda^* = 1/2$ is stable and the two cities are monocentric and have the same size (Lemma 3). Once τ falls below τ_{mm}^{+} but remains above τ_{mm}^{-} (see Figure 4), $\Delta_{mm}V(\lambda)$ becomes positive so that λ rises because $\lambda^* = 1/2$ is now unstable. As soon as the value $\overline{\omega}$ is reached, the utility differential changes and is now given by $\Delta_{pm}V(\lambda)$, which is also positive so that λ keeps rising. Hence, $\lambda^* = 1$ is the only stable equilibrium (Lemma 2).

When trade costs are below τ^- but above τ_1^-, we have $\delta_1(\tau) + \delta_2(\tau) < 0$ and $\delta_1(\tau) > 0$ so that $\Delta_{pm}V(1) < 0$. This implies that $\lambda^* = 1$ is no longer an equilibrium. Because $\delta_1(\tau) > 0$, region 1 keeps shrinking until the value $\overline{\omega}$ is reached. Below this value, the two cities become monocentric so that the equation of motion is given by $\Delta_{mm}V(\lambda)$, which is also negative. This implies that the new stable equilibrium is given by dispersion (Lemma 3). Hence, the economy involves two monocentric cities once trade costs fall below the level τ^-, as stated in Proposition 3.