Endogenous Structure of Polycentric Urban Area I: Isolated City

Sidorov, Alexander

Sobolev Institute of Mathematics SB RAS, Novosibirsk State University, Economics Department, Higher School of Economics

16 February 2012

Endogenous Structure of Polycentric Urban Area I: Isolated City

Alexander V. Sidorov*

Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia

Abstract

The purpose of this paper is to investigate how the interplay between production, commuting and commuting and costs shapes the economy at intra-urban level. Specifically, we study how economic integration affects the internal stricture of cities and how decentralizing of production and consumption of goods in secondary employment centers allows firms located in a large city to maintain their performance. The main distinctive feature of the model is two-dimensional city structure with variable number of secondary business districts. Several new results in urban economics are established, which all agree with empirical evidence and some of them cannot be obtained in framework of the linear city model.

Keywords and Phrases: City structure, Secondary business centers, Commuting costs, Communication costs

JEL Codes: F12, F22, R12, R14

Introduction

A weakness in urban economic theory is that it has relied too heavily on the monocentric city model.[1] A single job center runs counter to the evidence that has accumulated in the empirical literature on employment subcenters. But, the main drawback of the monocentric model is that it fails to explain that job location – even in a single center – is not exogenous but depends on other determinants of urban form. Would an alternative polycentric model be too complex and intractable? A reasonably tractable polycentric model can be based on the assumption that production and residential uses can occur everywhere in an initially featureless space but become interdependent by the commuting decisions of workers and the communication linkages among firms. Producers value access to other producers, to labor, and to facilities that help to run its business. The location of production and, hence, of jobs is endogenous as is the location of residences and, hence, of labor. From this perspective the monocentric city arises as the total clustering of jobs.

We are about to explain how decentralizing the production of goods in secondary employment centers may allow large cities to retain a large share of firms and jobs in an integrating world. Our starting point is that firms' performances are affected by the level of housing and commuting costs, which we call "urban costs". This occurs through the land rent they pay to occupy central urban locations, and through the higher wages they have to pay to their workers to compensate them for their longer commutes and/or higher land rents. Hence, high urban costs render firms less competitive on local and foreign markets alike. As a result, despite scale economies arising from urban agglomeration,

*This research is supportred by Economic Education and Research Consortium grant No.11-5111 and by Russian Federation Government grant No. 11.G34.31.0059t

[1]See Wheaton (1979) and Berry and Kim (1993) for critiques of the monocentric paradigm.

increasing urban costs could shift employment from large monocentric cities either to their suburbs or to distant and smaller cities. In other words, economic integration could well challenge the supremacy of large cities in favor of small cities.

Creation of subcenters within a city, i.e. the formation of a polycentric city, appears to be a natural way to alleviate the burden of urban costs. It is, therefore, no surprise that Anas et al. (1998) observe that "polycentricity is an increasingly prominent feature of the landscape." Thus, the escalation of urban costs in large cities seems to prompt a redeployment of activities in a polycentric pattern, while smaller cities retain their monocentric shape. However, for this to happen, firms set up in the secondary centers must maintain a very good access to the main urban center, which requires low communication costs.

Comparison to other approaches

Trying to explain the emergence of cities with various sizes our framework, unlike Helpman (1998), Tabuchi (1998) and others, allows cities to be polycentric. Moreover, in contrast to A. Sullivan (1986), K. Wieand (1987), and (Helsley and Sullivan, 1991), in our treatment, there are no pre-specified locations or numbers of subcenters, and our model is a fully closed general equilibrium spatial economy. As mentioned above, emergence of additional job centers is based on the urge towards decreasing of urban costs, rather than mysterious consumer's "propensity to big malls", as suggested Anas and Kim (1996). Our approach, that takes into account various types of costs (trade, commuting, and communication) is similar to J. Cavailhès et al. (2007) with one important exception. We drop very convenient (yet usually non-realistic) assumption on "long narrow city". Our analysis is extended to the two-dimension because the geographical space in the real world is better approximated by a two-dimensional space.

1 Model overview

1.1 Spatial structure

Consider an economy with one sector and two primary goods, labor and land. The region can be urbanized by accommodating firms and workers within a city, and is formally described by a two-dimensional space $X = \mathbb{R}^2$ (or by sufficiently large area around origin). Whenever a city exists, it has a central business district (in short CBD) located at the origin $0 \in X$. Residence zone around CBD assumed to be a circle due to geographical homogeneity. One would expect us to explain why this CBD exists as well as why firms leaving the CBD want to be together and form a secondary business districts, in short SBDs (see Figure 1). Doing that would require the introduction of local spatial externalities and local public goods that would render the analysis much more involved from the technical point of view, without adding much to our results. Indeed, our model has nothing new to add to what is known in this domain. By contrast, we determine the sizes of the CBD and the SBDs, thus the structure of each city, in the presence of inter-city trade and factor mobility.

Firms are free to locate in the CBD or to set up in the suburbs of the metro where they form a SBDs. Both the CBD and SBDs are assumed to be dimensionless, while residence zones around SBDs are also circles. In what follows, the superscript C is used to describe variables related to the CBD, whereas S describes the variables associated with a SBDs. Without loss of generality, we focus on the only one of SBDs, because all SBDs are supposed to be identical. *Locations* are expressed by variable

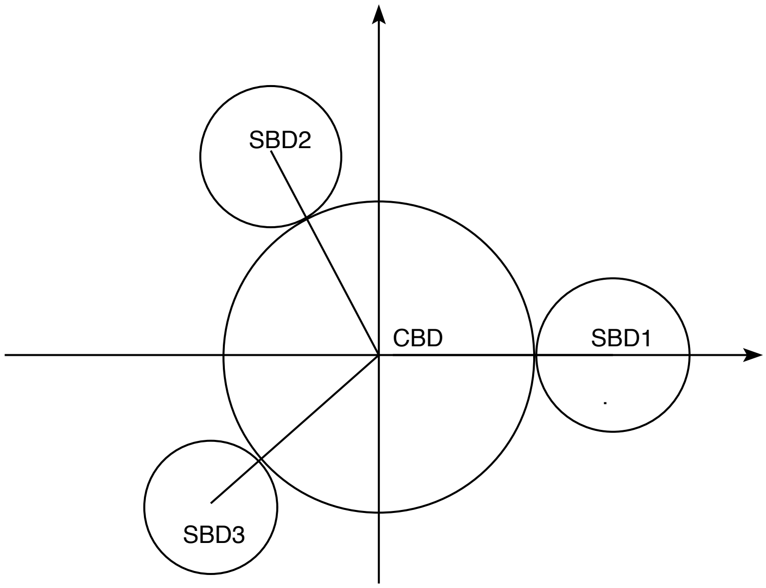


Figure 1: Polycentric city

$x \in X$ while *distances* are measured as Euclidean norm $||x||$ for CBD-zone, whereas the SBD, if any, is established at $x^S \neq 0$, which is endogenous. Even though firms consume services supplied in each SBD, the higher-order functions (specific local public goods and non-tradeable business-to-business services such as marketing, banking, insurance) are still located in the CBDs.

Hence, for using such services, firms set up in a SBD must incur a communication cost, which is given by

$$\mathcal{K}(x^S) = K + k \cdot ||x^S|| \tag{1}$$

where K and k are two positive constants. Indeed, communicating requires the acquisition of specific facilities, thus explaining why communication costs have a fixed component. However, relationships between the CBD and a SBD also involves face-to-face communication. Therefore, some workers must go to the CBD, thus making communication costs dependent on the distance $||x^S||$ between the CBD and the SBD. For simplicity, we assume that this cost is linear in distance, but this does not affect the nature of our results. Both the CBD and the SBD are surrounded by residential areas occupied by workers. Furthermore, as the distance between the CBD and SBD is small compared to the intercity distance, we disregard the intra-urban transport cost of goods. Finally, we consider the case where the CBD is surrounded by $m \geq 1$ SBDs. Under those various assumptions, the location, number and size of the SBDs as well as the size of the CBD are endogenously determined. In other words, apart from the assumed existence of the CBD, the internal structure of each city is endogenous.

1.2 Workers

The economy city 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 (or produced outside of city), homogeneous and chosen as the numéraire. The second good is produced as a continuum n of varieties of a horizontally differentiated good under monopolistic competition and increasing returns, using labor as the only input. The third good is land; without loss of generality, we set the opportunity cost of land to zero. Each worker consumes a residential plot of fixed size chosen as the unit of area.

The worker also chooses a quantity $q(i)$ of variety $i \in [0,n]$, and a quantity q_0 of the numéraire. She is endowed with one unit of labor and $\bar{q}_0 > 0$ units of the numéraire. This initial endowment may be also interpreted as monetary income additional to wage. The possible sources of this endowment will be discussed in proper time. At this moment initial $\bar{q}_0$ are considered as exogenous parameters. Each worker commutes to her employment center — without cross-commuting — and bears a unit commuting cost given by $t > 0$, so that for the worker located at x the commuting cost is either $t||x||$ or $t||x - x^S||$ according to the employment center.

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

$$\int_0^n p(i)q(i)\mathrm{d}i + q_0 + R^C(||x||) + t||x|| = w^C + \bar{q}_0 \tag{2}$$

where $R^C(x)$ is the land rent prevailing at location x (in fact, it depends on distance $||x||$ from the CBD only). The budget constraint of an individual working in the SBD, located at specific place x^S is

$$\int_0^n p(i)q(i)\mathrm{d}i + q_0 + R^S(||x\text{-}x^S||) + t||x\text{-}x^S|| = w^S + \bar{q}_0. \tag{3}$$

In fact rents in suburb depend on distance to the corresponding employment center only (see subsection 2.1). Preferences over the differentiated product and the numéraire are identical across workers and represented by a utility function $U(q_0; q(i), i \in [0,n])$.

The results of this paper are universal, they do not depend on specification of utility unlike the most part of similar papers. For example, one of the popular functions is Ottaviano's quasi-linear utility function incapsulating quadratic sub-utility

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

where $\alpha > 0$, $\beta > \gamma > 0$. Comprehensive analysis of *linear city with two SBDs* with this function was carried out in Cavailhès et al. (2007). Another popular type is two-tier (Cobb-Douglas over CES) utility function

$$U(q_0; q(i), i \in [0,n]) = \left(\int_0^n [q(i)]^{\frac{\sigma-1}{\sigma}}\mathrm{d}i\right)^{\frac{\sigma\mu}{\sigma-1}} \cdot q_0^{1-\mu}$$

where $0 < \mu < 1$ is expenditure share of differentiated good, $\sigma > 1$ is constant elasticity of substitution. This one was used in Tabuchi (2009) and (Tabuchi and Thisse, 2006). In both settings we may study the indirect utility functions in analytical form.

1.3 Firms

Technology in manufacturing is such that producing $q(i)$ units of variety i requires a given number φ of labor units. 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, the total number of firms is given by $n = L/\varphi$.

Denote by Π^C (respectively Π^S) the profit of a firm set up in the CBD (respectively the SBD).

Let θ be the share of firms located in the CBD and, therefore, by $(1-\theta)/m$ the share of firms in each SBD. When the firm producing variety i is located in the CBD, its profit function is given by:

$$\Pi^C(i) = I(i) - \varphi \cdot w^C \tag{4}$$

where $I(i)$ stands for the firm's revenue earned from local sales and from exports. When the firm sets up in the SBDs of the same city, its profit function becomes:

$$\Pi_r^S(i) = I(i) - \mathcal{K}(x^S) - \varphi \cdot w^S \tag{5}$$

the firm's revenue is the same as in the CBD because shipping varieties within the city is costless so that prices and outputs do not depend on firm's location in the city. Those two expressions encapsulate the trade-off faced by firms located in city: by locating at the SBD, firms are able to pay a lower wage to workers, but must incur the communication cost $\mathcal{K}(x^S)$. At this step we don't specify exact form of revenues $I(i)$ that, in turn, requires specifying of utility $U(q_0; q(i), i \in [0, n])$.

1.4 Two-dimensional features

The short overview of the model allows to point out some differences two-dimensional model from model of "long narrow city" that could be substantial for final outcomes. Urban costs (commuting and communication) mainly depend on *distances* or *geographic size* of the city, which is proportional to *population* or *demographic size* in linear city and *less than proportional* in two-dimensional model (to be more specific, geographic size increases as *square root* of population). Moreover, an additional economy on scale in urban costs comes from possibility to allocate more than two SBDs around central zone. It means that linear model (possibly) *overestimates* dispersion forces (caused by urban costs) in comparison to agglomeration forces (related to monopolistic competition). In other words, two-dimensional model is "more favorable" to formation of larger city agglomeration.

2 Decentralization within a city

A city equilibrium is such that each individual maximizes her utility subject to her budget constraint, each firm maximizes its profits, and markets clear. Individuals choose their workplace (CBD or SBDs) and their residential location with respect to given wages and land rents. In each workplace (CBD or SBDs), the equilibrium wages are determined by a bidding process in which firms compete for workers by offering them higher wages until no firm can profitably enter the market. Given such equilibrium wages and the location of workers, firms choose to locate either in the CBD or in the SBDs. At the city equilibrium, 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 analyze such an equilibrium, taking as fixed the number of workers L.

2.1 Land rents, wages and workplaces

Within each city, a worker chooses her location so as to maximize her utility $U(q_0, q(i); i \in [0, n])$ under the corresponding budget constraint, (2) or (3). Let $\Psi^C(x)$ and $\Psi^S(x)$ be the bid rent at $x \in X$ of an individual working respectively in the CBD and the SBD. Land is allocated to the highest bidder. Because there is only one type of labor, at the city equilibrium it must be that

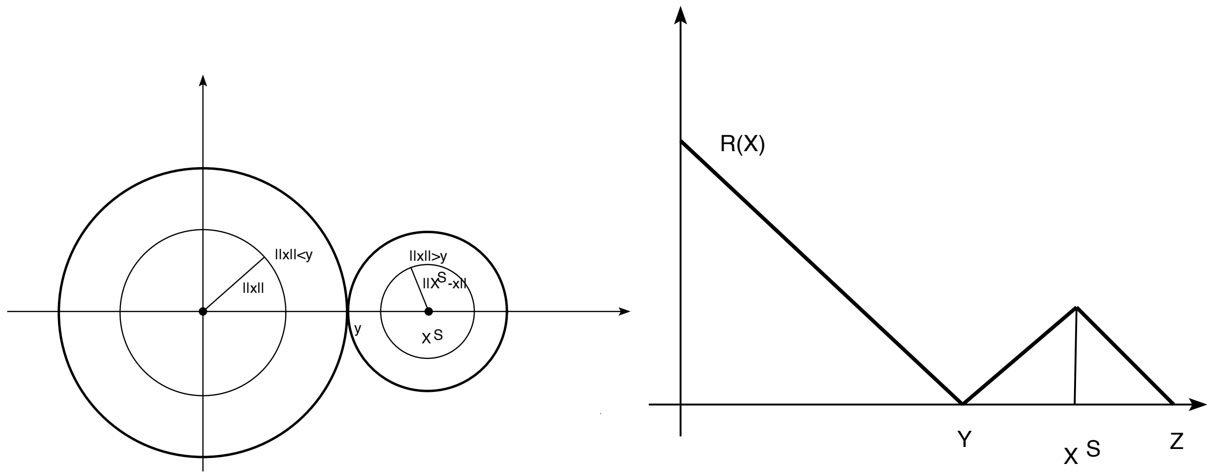


Figure 2: City with CBD and representative SBD

$R(x) = \max\left\{\Psi^C(x), \Psi^S(x), 0\right\}$ (see Figure 2). The question "Who is recipient of rent?" will be discussed in proper time. At this moment we may assume that there exists some collecting agency.

Without loss of generality we assume that SBD resides at axis of abscissas, which allows to drop Euclidean norms and consider positive scalars only. Denote by y the radius of the area formed by residents working in the CBD while z is from CBD of the most remote point of SBD residential areas. (Here and further x, y, z and so on are the positive numbers or, equivalently, points on positive part of x-axis.) Then obviously $x^S = \dfrac{y+z}{2}$. Because communication costs to the CBD increase with distance, the secondary residential areas are adjacent to central one when the city is polycentric.

Suppose for this moment that CBD's share of firms θ, number of SBDs m and city's population L are given exogenously. Then the critical points are as follows:

$$y = \sqrt{\frac{\theta L}{\pi}}, \quad z = \sqrt{\frac{\theta L}{\pi}} + 2\sqrt{\frac{(1-\theta)L}{m\pi}}, \quad x^S = \sqrt{\frac{\theta L}{\pi}} + \sqrt{\frac{(1-\theta)L}{m\pi}}. \tag{6}$$

Indeed, the equilibrium population of CBD θL is equal to area of circle πy^2, while the total population of SBDs $(1-\theta)L$ is equal to sum of SBD's areas $m\pi(x^S - y)^2$.

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

$$E = \int_0^n p(i)q(i)\mathrm{d}i + q_0$$

is the same regardless of the worker's location. Then, the budget constraint of an individual residing at x and working in the CBD implies that $w^C + \bar{q}_0 - R(x) - tx = E$, whereas the budget constraint of an individual working in the SBD is $w^S + \bar{q}_0 - R(x) - t|x - x^S| = E$. At the city equilibrium, the worker living at the border of the CBD residential area (or at the point of the SBD residential area closest to CBD) is indifferent between working in the CBD or in the SBD, which implies $w^C + \bar{q}_0 - R(y) - ty = w^S + \bar{q}_0 - R(y) - t|y - x^S|$. This equation may be considered also as a parity of *disposable incomes.*

After reduction and substituting of (6) this equation becomes

$$w^C - w^S = t\left(\sqrt{\frac{\theta L}{\pi}} - \sqrt{\frac{(1-\theta)L}{m\pi}}\right). \tag{7}$$

Thus, the difference in the wages paid in the CBD and in the SBD compensates exactly the worker for the difference in the corresponding commuting costs. The wage wedge $w^C - w^S$ is positive as long as $\theta > \dfrac{1}{1+m}$, thus implying that the size of the CBD exceeds the size of each SBD. In this case $\dfrac{\partial}{\partial L}(w^C - w^S) > 0$, i.e. rise in the population size increases the wage wedge: as the average commuting cost rises, firms located in the CBD must pay a higher wage to their workers.

Finally, it is worth noting that the equilibrium land rents for arbitrary location $x \in X$ are given by

$$R(x) = \Psi^C(x) = t \cdot \left(\sqrt{\frac{\theta L}{\pi}} - ||x||\right), \quad \text{for} \quad ||x|| \leq y = \sqrt{\frac{\theta L}{\pi}} \tag{8}$$

and

$$R(x) = \Psi^S(x) = t \cdot \max\left\{0, \sqrt{\frac{(1-\theta)L}{m\pi}} - ||x^S - x||\right\}, \quad \text{for} \quad ||x|| > y \tag{9}$$

where θ is endogenously defined in the following subsection and $x^S \in X$ is SBD closest to $x \in X$.

At this moment we left behind two unanswered questions:

- What is a source of initial endowment $\bar{q}_0$ of numéraire?
- Who is/are recipient(s) of rent?

We answer them simultaneously. Assume that collecting agency, mentioned at start of this subsection, is non-profit and redistribute *uniformly* all of collected rent among all population of city. So initial endowment $\bar{q}_0$ is in fact an individual share of numéraire good, financed by the equal share of total raised rent in specific city. Due to (4)-(5) plot of the rent function is a bunch of cone surfaces: one for central zone and m for secondary residence zones (see right-hand side of Figure 2, representing the section of cone surfaces). Thus, the total rent sum is equal to the sum of *cone volumes*: one central cone and m secondary cones. From (4)-(5) heights of these cones are, respectively, $h^C = t \cdot \sqrt{\frac{\theta L}{\pi}}$ and $h^S = t \cdot \sqrt{\frac{(1-\theta)L}{m\pi}}$, while the corresponding radii are respectively, $r^C = \sqrt{\frac{\theta L}{\pi}}$ and $r^S = \sqrt{\frac{(1-\theta)L}{m\pi}}$. Using formula of cone volume ($V = \frac{1}{3}\pi r^2 h$) we obtain the total rent sum as follows

$$\frac{L \cdot t}{3}\left[\theta \cdot \sqrt{\frac{\theta \cdot L}{\pi}} + (1-\theta) \cdot \sqrt{\frac{(1-\theta) \cdot L}{m \cdot \pi}}\right]$$

while the total initial endowment is $L \cdot \bar{q}_0$. Closing our model, we obtain

$$\bar{q}_0 = \frac{t}{3} \cdot \sqrt{\frac{L}{\pi}}\left[\theta^{3/2} + \frac{(1-\theta)^{3/2}}{\sqrt{m}}\right].$$

2.2 Equilibrium wages and the city structure

Regarding the labor markets, the equilibrium wages of workers are determined by the zero-profit condition. In other words, operating profits are completely absorbed by the wage bill. Hence, the equilibrium wage rates in the CBD and in the SBDs must satisfy the conditions $\Pi^C(w^{C*}) = 0$ and $\Pi^S(w^{S*}) = 0$, respectively. Thus, setting (4) (respectively (5)) equal to zero, solving for w^{C*} (respectively w^{S*}), we get:

$$w^{C*} = \frac{I}{\varphi}, \quad w^{S*} = \frac{I - \mathcal{K}(x^S)}{\varphi} \tag{10}$$

Hence

$$w^{C*} - w^{S*} = \frac{K + kx^S}{\varphi},$$

which means that the equilibrium wage wedge is proportional to the level of the communication cost that prevails at the SBD. Substituting of (7) and (6) into previous formula yields:

$$(\varphi t - k)\sqrt{\frac{\theta L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{(1-\theta)L}{m\pi}} > 0 \tag{11}$$

which implies that inequality $\varphi t - k > 0$ is necessary condition. More exactly, the opposite inequality $k \geq \varphi t$ means that distance-sensitive communication costs are too large relative to commuting costs, so in fact we have rather *communicatively separated* cities, than connected CBD and SBD. When condition $\varphi t - k > 0$ is not satisfied, the city is monocentric regardless other values of parameters.

Assuming from now on that $\varphi t - k > 0$ holds, we have to solve this equation with respect to θ. First we consider more simple limit case $K = 0$, i.e. fixed communication costs are negligible. Then the solution is

$$\theta^* = \frac{(\varphi t + k)^2}{(\varphi t + k)^2 + (\varphi t - k)^2 m} = \frac{1}{1 + mg^2}$$

where g stands for $\dfrac{\varphi t - k}{\varphi t + k} \in (0, 1)$. Note that $\theta^* > \dfrac{1}{1+m}$, because $g^2 < 1$. The following statement is obvious.

Proposition 1. *Let $K = 0$, then polycentric city equilibrium exists if and only if $\varphi t > k$ and share of firms in CBD does not depend on city population l and decreases to zero with unlimited increasing of SBD's number m.*

Generalize this result to the case $K > 0$.

Proposition 2. *i) Let $\varphi t < k + K\sqrt{\frac{\pi}{L}}$ then equation (11) is unsolvable, i.e. city is, in fact, monocentric.*

ii) Let $\varphi t = k + K\sqrt{\frac{\pi}{L}}$ then equation (11) has unique solution $\theta^* = 1$, i.e. city is, in fact, monocentric.

iii) Let $\varphi t > k + K\sqrt{\frac{\pi}{L}}$ then equation (11) has unique solution $\theta^* \in \left(\frac{1}{1+mg^2}, 1\right)$, i.e. for each $m \geq 1$ there exists an equilibrium distribution of firms. Moreover, CBD's share of firms $\theta^*(m, l)$ decreases with respect to both population L and number of SBDs m and

$$\lim_{L\to\infty} \theta^*(m, L) = \frac{1}{1 + mg^2}, \quad \lim_{m\to\infty} \theta^*(m, L) = \frac{\pi K^2}{(\varphi t - k)^2 L} \in (0, 1). \tag{12}$$

For analytical proof see Appendix A.

Discussion

Note that $r_{\text{mono}}(L) = \sqrt{\frac{L}{\pi}}$ is in fact a radius of *monocentric* city with population size L. Therefore, would be polycentric if and only if $\varphi \cdot t \cdot r_{\text{mono}}(L) > K + kr_{\text{mono}}(L) = \mathcal{K}(r_{\text{mono}}(L))$, i.e. *commuting* costs of delivering labor to produce the next unit of output exceed corresponding *communication* costs. Thus, producing on periphery (in SBDs) is more efficient. For any given $K > 0$, k, t, φ satisfying $\varphi t > k$ we define *monocentric maximum population* $l_{\max}$ as a root of equation $\varphi t = k + K\sqrt{\dfrac{\pi}{L}}$, which implies:

$$l_{\max} = \frac{\pi K^2}{(\varphi t - k)^2}.$$

The corresponding maximum radius of monocentric city is

$$r_{\max} = \sqrt{\frac{l_{\max}}{\pi}} = \frac{K}{\varphi t - k}.$$

From perspective of production efficiency, if city population L exceeds $l_{\max}$ it give raise to SBDs and fraction $\dfrac{l_{\max}}{L}$ is an infimum of CBD firm's share under unlimited increasing of number of SBDs. What determines sufficient number of m will be considered below. The most striking implication of (12) (not quite realistic, however) is that limit *size of population* of central zone $\lim_{m\to\infty}\theta^*(m, L)\cdot L = \frac{\pi K^2}{(\varphi t-k)^2} = l_{\max}$ does not depend on L, i.e. even unlimited growth of city population in case of very large value of m will be completely absorbed by suburbia. The only difficulty is that infinite number of SBDs could not be placed around bounded CBD-zone without overlapping.

2.3 How to "endogenize" of SBD's number m

In what follows we assume that $L > l_{\max}$ (or, equivalently, $\varphi t > k + K\sqrt{\dfrac{\pi}{L}}$). Previous considerations show that then appear some number $m \geq 1$ of SBDs. What determines its number? There is no simple and unambiguous answer, because in practice it depends on many factors, that are not always economic ones.

One of the main questions is the following: "Who can afford the building of additional suburb?" If answer is "None", we find ourself in setting with predefined number of SBDs (like model of Cavailhès et al.). Let's consider the opposite situation when each firm could freely (without any costs) establish new SBD "at empty place".

2.3.1 Free foundation and maximum number of SBDs

To estimate the maximum possible number of SBDs, consider the following simple trigonometric problem. If we place m SBDs around CBD, each of them takes 2α of angular size, where $\alpha = \widehat{SOT}$ (see Figure 3). It is obvious that

$$\sin\alpha = \frac{r^S}{r^C + r^S} = \frac{\sqrt{1-\theta^*}}{\sqrt{m\theta^*} + \sqrt{1-\theta^*}},$$

where θ^* is equilibrium CBD share, defined as root of equation (11). Allocation of m SBDs without overlapping requires $2m\alpha$ of angular size, that implies inequality $m\alpha \leq \pi$. Thus the estimation of maximum number of SBDs M^* (and the corresponding θ^*) may be obtained as solution of the following system of equations

$$(\varphi t - k)\sqrt{\frac{\theta L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{(1-\theta)L}{m\pi}}$$

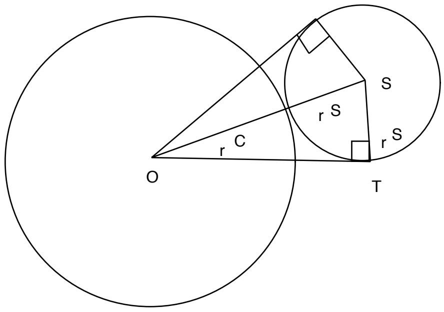


Figure 3: How many SBDs could we place around

$$m \cdot \arcsin \frac{\sqrt{1-\theta}}{\sqrt{m\theta}+\sqrt{1-\theta}} = \pi.$$

Note that for sufficiently large population size L (to be more specific, when $\dfrac{K}{\sqrt{L}}$ may be considered as negligible costs) the equilibrium share $\theta^* \approx \frac{1}{1+mg^2}$ (see Proposition 1) and

$$\frac{\sqrt{1-\theta^*}}{\sqrt{m\theta^*}+\sqrt{1-\theta^*}} \approx \frac{g}{1+g},$$

where g stands for $\dfrac{\varphi t - k}{\varphi t + k}$. Thus,

$$M^* \approx \frac{\pi}{\arcsin \frac{g}{1+g}} = \frac{\pi}{\arcsin \frac{\varphi \cdot t - k}{2\varphi \cdot t}}. \tag{13}$$

Moreover, for sufficiently small g we can approximate $\arcsin \frac{g}{1+g} \approx \frac{g}{1+g}$ and obtain the following simple (yet may be far from the exact value) estimation of maximum number of SBDs

$$M^* \approx \frac{(1+g)\pi}{g} = \frac{2\pi \cdot \varphi \cdot t}{\varphi \cdot t - k}.$$

Note that accordingly to estimation (13) (as well as its linearization) the maximum number M^* decreases with respect to marginal commuting costs t. It can be easily explained, because the larger commuting costs (in comparison to communication ones) force firms (and their workers) to locate themselves in SBDs. As result, size of each SBD zone increases, while free place around central zone decreases.

The general answer to the question that started this section is the following: “Decision on building of additional suburb is up to ‘City Developer’, who takes into account the social welfare considerations.” The very simple example of such decision-making is considered in the following subsection.

2.3.2 "People are not robots"

Usually people (and firms) cannot choose to build (or not to build) new suburb, thus for the individual or firm decision making number of SBDs m seems to be exogenous. Yet they can vote for mayor (city council) who can. For example, people are angry when commuting trip takes a lot of time. For example, if commuting takes over 1-2 hours, it could be unacceptable regardless any wage difference. It means that we should introduce in our model an *exogenous* parameter $r^S_{\max}$ of maximum SBD radius, or equivalently, maximum SBD population $l^S_{\max} = \pi \left(r^S_{\max}\right)^2$.Thus population constraint for SBD is as follows:

$$\frac{(1-\theta^*(m,L))L}{m} \leq l^S_{\max} \iff 1-\theta^*(m,L) \leq \frac{l^S_{\max}}{L}m.$$

Allowing non-integer values of m, consider equation

$$1-\theta^*(m,L) = \frac{l^S_{\max}}{L}m \tag{14}$$

with respect to m, where the positive root (if exists) define the lower bound of SBD number. Recall that $r_{\text{mono}}(L) = \sqrt{\frac{L}{\pi}}$ denotes a radius of monocentric city with population l and $r_{\max} = \sqrt{\frac{l_{\max}}{\pi}} = \frac{K}{\varphi t - k}$ is a maximum cost-efficient radius of mononocentric city.

Proposition 3. *Let* $r_{\text{mono}}(L) > r_{\max} + \dfrac{r^S_{\max}}{g}$ *then there exists the unique positive root* m^* *of equation* (14)

$$m^* = \frac{L}{l^S_{\max}} - \left[\frac{\varphi t + k + K\sqrt{\frac{\pi}{l^S_{\max}}}}{\varphi t - k}\right]^2$$

(generally, non-integer). If $r_{\text{mono}}(L) \leq r_{\max} + \dfrac{r^S_{\max}}{g}$*, then inequality* $\dfrac{(1-\theta^*(m,L))L}{m} \leq l^S_{\max}$ *holds for all admissible* m*.*

For analytical proof see Appendix B.

Discussion. Recall that inequality $r_{\text{mono}} \leq r_{\max}$ is exactly condition of monocentricity without restriction on suburb size, so the inequality just obtained may be interpreted as "*one suburb is quite enough*". Note also that equation (11) that determines the city equilibrium share θ^* (without any restrictions), for $m = 1$ takes the following form:

$$(\varphi t - k)\sqrt{\frac{\theta L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{(1-\theta)L}{\pi}}$$

which is equivalent to

$$r^C = r_{\max} + \frac{r^S}{g}$$

where $r^C = \sqrt{\frac{\theta L}{\pi}}$ is radius of central zone, $r^S = \sqrt{\frac{(1-\theta)L}{\pi}}$ is a radius of the single suburb. Now the endogenously defined sufficient number m of SBDs is minimum positive integer number exceeding m^*. It is obvious that m^* increases with respect to city population L, moreover it also increases with respect to commuting costs t (see Appendix B). Both theoretical conclusions are supported by empirical evidences (see MacMillen and Smith, 2003).

Note that the same considerations may be applied to CBD, which may lower the value of $l_{\max}$. This would not change our basic conclusion, so we assume that CBD is more attractive for workers than SBDs and firm's allocation in SBD's is due to production efficiency reasons.

Moreover, the above considerations imply that SBD zones have no overlapping. This assumption generates the upper bound of SBDs number, that is different from M^* due to SBD's size limitations. Nevertheless, its definition is very similar:

$$m \cdot \arcsin \frac{r^S_{\max}}{r^C + r^S_{\max}} = \pi \iff G(m) := m \cdot \arcsin \frac{1}{\sqrt{\frac{L}{l^S_{\max}} - m + 1}} = \pi. \tag{15}$$

Without loss of generality we may assume that $\frac{L}{l^S_{\max}} > 2$, otherwise SBD's size limitations are inessential. The left-hand side of equation (15) $G(m)$ is an strictly increasing function of m, while $G(2) = 2 \arcsin \frac{1}{\sqrt{\frac{L}{l^S_{\max}} - 2 + 1}} < 2 \arcsin \frac{1}{\sqrt{2-2+1}} = \pi$ and $G\left(\frac{L}{l^S_{\max}}\right) = \frac{L}{l^S_{\max}} \cdot \frac{\pi}{2} > \pi$. Thus, there exists a unique solution of this equation $M^{**} \in \left(2, \frac{L}{l^S_{\max}}\right)$ and its integer part is theoretical maximum of SBDs with size limitations.

2.3.3 Hierarchy of Business Districts.

Is the maximum number M^* (or M^{**}) the insuperable obstacle? Not yet. Let's assume that each SBD is surrounded by BDs of lesser rank: Tertiary Business Districts (TBD), possibly surrounded by Quarternary Business Districts (QBD) and so on. Here we consider model with TBDs, that could be easily extended on hierarchy of any complicacy.

Suppose that there exists m_1 of identical SBDs and each SBD is surrounded by m_2 TBDs (see, for example, left panel of Figure 5 in Appendix D). We assume that SBD has no own communication facilities and communication of any TBD with CBD is carried out via adjacent SBD. Let x^S still denote distance of any SBD from CBD, while x^T is a distance of TBD *from adjacent* SBD. Further, let θ_0 denote a share of firms, located at CBD, θ_1 is share of firms, located at each SBD and θ_2, respectively, is a share of firms, located at each TBD.

Proposition 4. *Let $\varphi t > k + K\sqrt{\frac{\pi}{L}}$ then exists unique city equilibrium and the corresponding shares are as follows:*

$$\theta_2^* = \frac{1 - \theta_0^* - m_1\theta_1^*}{m_1 m_2}, \theta_1^* = \frac{1 - \theta_0^*}{m_1(m_2 \cdot g^2 + 1)},$$

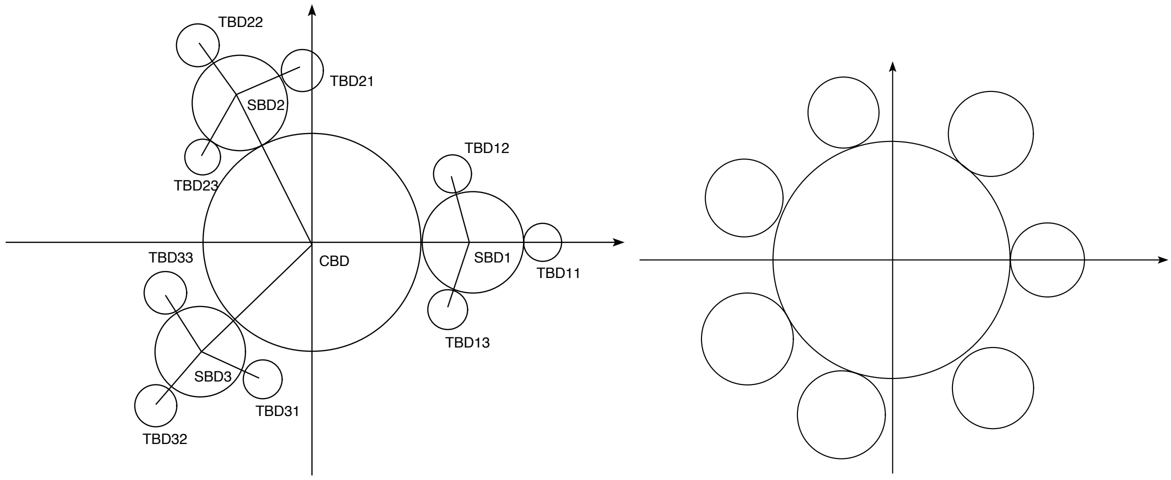
where $\theta_0^* \in \left(\frac{1}{1 + m_1 g^2 \cdot (m_2 \cdot g^2 + 1)}, 1\right)$ is a unique root of equation

$$(\varphi t - k)\sqrt{\frac{\theta_0 L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{(1-\theta_0)L}{m_1(m_2 \cdot g^2 + 1)\pi}}.$$

See Appendix C for analytical proof.

Discussion. Note that city structure with m_1 SBDs, each surrounded by m_2 TBDs, is *technically equivalent* to city structure with $m_1(m_2 \cdot g^2 + 1)$ SBDs. It is easy to understand that city structure with m_1 SBDs, each surrounded by m_2 TBDs, each surrounded by m_3 QBDs, is also technically equivalent to city with non-integer "number" of SBDs $m = m_1(m_2 \cdot g^2 + 1)(m_3 \cdot g^2 + 1)$, and so on. Note that sequence of numbers

$m_1(m_2 \cdot g^2 + 1)(m_3 \cdot g^2 + 1) \ldots (m_n \cdot g^2 + 1)$ increases boundlessly with n even if all m_i are bounded.



Three-tier city with $m_1 = m_2 = 3$ "Equivalent" two-tier city for $g = \dfrac{2}{3}$.

Figure 4: City with TBDs and "equivalent" two-tier city

For example, suppose that $m_1 = m_2 = 3$ and $k = 1, \varphi = 1, t = 5$, then $g = \dfrac{\varphi t - k}{\varphi t + k} = \dfrac{2}{3}$ and "effective" number of SBDs is $m = 3 \cdot \left(3 \cdot \left(\dfrac{2}{3}\right)^2 + 1\right) = 7$ (see right panel of Figure 4). Note that the similar considerations on multi-level system of marketplaces with Christaller-Lösch system of hexagonal market areas was used by Tabuchi (2009). Unlike this we do not use any pre-defined location grid.

References

[1] A. Anas, R. Arnott, K.A. Small, (1998) Urban spatial structure, Journal of Economic Literature, 36, pp. 1426–1464.

[2] A. Anas, I. Kim, (1996) General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration, Journal of Urban Economics , 40: 217–32.

[3] J. L. Berry, H. M. Kim, (1993) Challenges to the monocentric model, Geographical Analysis, 25, pp. 1-5 .

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

[5] E. Helpman, (1998) The size of regions, in: D. Pines, E. Sadka, I. Zilcha (Eds.), Topics in Public Economics. Theoretical and Applied Analysis, Cambridge Univ. Press, Cambridge, pp. 33–54.

[6] R. W. Helsley, A. M. Sullivan, (1991) Urban subcenter formation, Regional Science and Urban Economics, 21, 255-275

[7] D.P. MacMillen, S. Smith, (2003) The number of subcenters in large urban areas, Journal of Urban Economics 53 321–338.

[8] A. Sullivan, (1986) A general equilibrium model with agglomerative economies and decentralized employment, Journal of Urban Economics, 20, 55-74.

[9] T. Tabuchi, (1998) Urban agglomeration and dispersion: a synthesis of Alonso and Krugman, Journal of Urban Economics, 44, pp. 333–351.

[10] T. Tabuchi, (2009) Self-organizing marketplaces, Journal of Urban Economics, Elsevier, vol. 66(3), pp. 179-185.

[11] T. Tabuchi, J.-F. Thisse, (2006) Self-organizing Urban Hierarchy, CIRJE F-Series CIRJE-F-414, CIRJE, Faculty of Economics, University of Tokyo.

[12] W. C. Wheaton, (1979) Monocentric models of urban land use: Contributions and criticism, in "Current Issues in Urban Economics" (P. Mieszkowski and M. Straszheim, Eds., Johns Hopkins Press, Baltimore)

[13] K. Wieand, (1987) An extension of the monocentric urban spatial equilibrium model to a multicenter setting: The case of the two-center city, Journal of Urban Economics, 21, 259-271 .

APPENDIX

A. Proof of Proposition 2

Raising to the second power both sides of equation

$$(\varphi t - k)\sqrt{\frac{\theta L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{(1-\theta)L}{m\pi}}$$

we obtain, after simple transformations, the following one

$$2f\sqrt{1-\theta} = -(1+mg^2)(1-\theta) + [mg^2 - f^2], \tag{16}$$

where $g = \dfrac{\varphi t - k}{\varphi t + k} \in (0,1)$ and $f = \dfrac{K\sqrt{m\pi}}{(\varphi t + k)\sqrt{L}} > 0$. Note that,

$$[mg^2 - f^2] > 0 \iff \varphi t > k + K\sqrt{\frac{\pi}{L}}.$$

Consider three possible cases:

i) $\varphi t < k + K\sqrt{\frac{\pi}{L}}$ then $[mg^2 - f^2] < 0$ and equation (16) has no real roots.

ii) $\varphi t = k + K\sqrt{\frac{\pi}{L}}$ then $[mg^2 - f^2] = 0$ and equation (16) has unique positive root $\theta^* = 1$.

iii) $\varphi t > k + K\sqrt{\frac{\pi}{L}}$ then $[mg^2 - f^2] > 0$. To obtain the admissible root of equation (16), let's substitute $1-\theta$ with v^2 for $v \in (0,1)$. Then this equation transforms into

$$F(v) = (1+mg^2)v^2 + 2fv - [mg^2 - f^2] = 0.$$

This quadratic equation has two real roots, positive and negative. Moreover, $F(0) = -[mg^2 - f^2] < 0$ while $F(1) = 1+2f+f^2 > 0$, thus there is unique root $v^* \in (0,1)$, and, therefore $\theta^* = 1-(v^*)^2 \in (0,1)$ is uniquely defined.

Being a root of quadratic equation, term v^*, as well as θ^*, could be represented in closed form in terms of the model parameters, yet is is more convenient consider v^* as *implicit function* of parameters, in particular, of population L and number of SBDs m. Accordingly to Theorem on Implicit Function Derivative, we obtain

$$\frac{\partial v^*}{\partial L} = -\frac{\left(\frac{\partial F}{\partial L}\right)}{\left(\frac{\partial F}{\partial v}\right)} = -\frac{(v+f)\cdot\frac{\partial f}{\partial l}}{(1+mg^2)v + f} > 0,$$

because obviously $\dfrac{\partial f}{\partial l} = -\dfrac{K\sqrt{m\pi}}{2(\varphi t + k)\sqrt{L^3}} < 0$. It implies that $\theta^*(l) = 1-(v^*(l))^2$ is *decreasing* function. Moreover, $f \to 0$ when $l \to +\infty$, thus $(v^*)^2 \to \dfrac{mg^2}{1+mg^2}$ and $\theta^* = 1-(v^*)^2 \to \dfrac{1}{1+mg^2}$. It means that for all $l < +\infty$ inequality $\theta^* > \dfrac{1}{1+mg^2} > \dfrac{1}{1+m}$ holds. Note that, $m\dfrac{\partial f}{\partial m} = \dfrac{K\sqrt{m\pi}}{2(\varphi t + k)\sqrt{L}} = \dfrac{f}{2}$, thus

$$m\cdot\frac{\partial F}{\partial m} = mg^2v^2 + 2v\cdot m\frac{\partial f}{\partial m} + 2f\cdot m\frac{\partial f}{\partial m} - mg^2 = mg^2v^2 + vf + f^2 - mg^2 = F(v) - v^2 - fv.$$

Therefore,

$$\frac{\partial v^*}{\partial m} = -\frac{\left(\frac{\partial F}{\partial m}\right)}{\left(\frac{\partial F}{\partial v}\right)} = -\frac{F(v^*) - (v^*)^2 - fv^*}{((1+mg^2)v^* + f)\cdot m} = \frac{(v^*)^2 + fv^*}{((1+mg^2)v^* + f)\cdot m} > 0$$

and $\dfrac{\partial \theta^*}{\partial m} < 0$.

Consider equation

$$\frac{F(v)}{m} = (\frac{1}{m} + g^2)v^2 + 2\frac{f}{m}v - \left[g^2 - \left(\frac{f}{\sqrt{m}}\right)^2\right] = 0$$

that is equivalent to (16). Note that, $\dfrac{f}{m} \to 0$, when $m \to +\infty$ and $\dfrac{f}{\sqrt{m}} \equiv \dfrac{K\sqrt{\pi}}{(\varphi t + k)\sqrt{L}}$, thus

$$(v^*)^2 \to 1 - \frac{K^2\pi}{g^2(\varphi t + k)^2 L} = 1 - \frac{K^2\pi}{(\varphi t - k)^2 L}$$

and $\theta^* = 1 - (v^*)^2 \to \dfrac{K^2\pi}{(\varphi t - k)^2 L}$.

B. Proof of Proposition 3

Substituting $1 - \theta^*(m) = \frac{l^S_{\max}}{L} m$ into equation (16)

$$2f\sqrt{1-\theta} = -(1 + mg^2)(1-\theta) + [mg^2 - f^2],$$

where $g = \dfrac{\varphi t - k}{\varphi t + k}$ and $f = \dfrac{K\sqrt{m\pi}}{(\varphi t + k)\sqrt{L}}$ we obtain

$$2\tilde{f}\sqrt{\frac{l^S_{\max}}{L}} \cdot m = -(1 + mg^2)\frac{l^S_{\max}}{L}m + [g^2 - \tilde{f}^2]m,$$

where $\tilde{f} = \dfrac{K\sqrt{\pi}}{(\varphi t + k)\sqrt{L}} > 0$. Reducing improper root $m = 0$ we obtain, after obvious transformations, that the unique root is

$$m^* = \frac{L}{l^S_{\max}} - \left[\frac{\varphi t + k + K\sqrt{\frac{\pi}{l^S_{\max}}}}{\varphi t - k}\right]^2.$$

Direct calculations show that $m^* > 0$ if and only if $r_{\text{mono}}(L) > r_{\max} + \dfrac{r^S_{\max}}{g}$. Unsolvability of equation $1 - \theta^*(m) = \frac{l^S_{\max}}{L} m$ means that inequality $1 - \theta^*(m) < \frac{l^S_{\max}}{l} m$ holds for all $m > 0$, because left-hand side is bounded while right-hand side is not. It is obvious that $\dfrac{\partial m^*}{\partial l} > 0$, moreover fraction $\frac{\varphi t + k + K\sqrt{\frac{\pi}{l^S_{\max}}}}{\varphi t - k} = 1 + \frac{2k + K\sqrt{\frac{\pi}{l^S_{\max}}}}{\varphi t - k}$ decreases with respect to t, thus $\dfrac{\partial m^*}{\partial t} > 0$.

C. Proof of Proposition 4

Obviously, share of firms at each TBD may be calculated as follows:

$$\theta_2 = \frac{1 - \theta_0 - m_1\theta_1}{m_1 m_2}.$$

Considerations similar to ones from subsection 2.2 show that on production side wage differences are

$$w^C - w^S = \frac{K + kx^S}{\varphi}, \quad w^S - w^T = \frac{kx^T}{\varphi},$$

where w^T is wage at TBD, while $w^C$and w^S are the same as earlier. For given θ_0 and θ_1 equilibrium distances of SBDs and TBDs are as follows

$$x^S = \sqrt{\frac{\theta_0 L}{\pi}} + \sqrt{\frac{\theta_1 L}{\pi}},$$

$$x^T = \sqrt{\frac{\theta_1 L}{\pi}} + \sqrt{\frac{(1-\theta_0-m_1\theta_1)L}{m_1 m_2 \pi}}.$$

(Recall that x^T denote distance of TBD from *adjacent* SBD. On the other hand, considerations on rents and wages, similar to subsection 2.1, show that

$$w^C - w^S = t\left(\sqrt{\frac{\theta_0 L}{\pi}} - \sqrt{\frac{\theta_1 L}{\pi}}\right),$$

$$w^S - w^T = t\left(\sqrt{\frac{\theta_1 L}{\pi}} - \sqrt{\frac{(1-\theta_0-m_1\theta_1)L}{m_1 m_2 \pi}}\right).$$

Putting altogether we obtain system of equations

$$(\varphi t - k)\sqrt{\frac{\theta_0 L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{\theta_1 L}{\pi}}$$

$$(\varphi t - k)\sqrt{\frac{\theta_1 L}{\pi}} = (\varphi t + k)\sqrt{\frac{(1-\theta_0-m_1\theta_1)L}{m_1 m_2 \pi}}.$$

From latter equation follows that

$$\theta_1 = \frac{1-\theta_0}{m_1(m_2 \cdot g^2 + 1)},$$

where g, as usual, stands for $\dfrac{\varphi t - k}{\varphi t + k}$. Thus, the first equation of this system turns into

$$(\varphi t - k)\sqrt{\frac{\theta_0 L}{\pi}} = K + (\varphi t + k)\sqrt{\frac{(1-\theta_0)L}{m_1(m_2 \cdot g^2 + 1)\pi}}.$$

Note that substituting m for $m_1(m_2 \cdot g^2 + 1)$ we obtain exactly equation (11), thus we may apply considerations from proof of Proposition 2 (see Appendix A).