DOCUMENT DE TRAVAIL DT/2003/06

Agricultural surplus, division of labour and the emergence of cities
A spatial general equilibrium model

Gilles SPIELVOGEL

AGRICULTURAL SURPLUS, DIVISION OF LABOUR AND THE EMERGENCE OF CITIES
A spatial general equilibrium model

Gilles Spielvogel
IEP-Paris, DIAL – UR CIPRÉ de l'IRD
gilles.spielvogel@sciences-po.org

Document de travail DIAL / Unité de Recherche CIPRÉ
Juillet 2003

RESUME

Dans ce papier, nous exposons les conditions économiques d'émergence des villes dans le cadre d'un modèle d'équilibre général spatial. L'existence de rendements croissants basés sur la division du travail, de coûts de transport et la présence éventuelle d'un surplus agricole conduisent à différentes possibilités d'équilibre urbain. En raison de la contrainte de subsistance, il est possible qu'aucune ville ne soit soutenable si les coûts de transport internes sont trop élevés. D'un autre côté, la contrainte d'emploi urbain débouche sur la saturation de tout ou partie du marché du travail urbain et à la persistance d'une pression migratoire entre campagne et ville. Par ailleurs, nous étudions les conditions de stabilité du système urbain monocentrique dans les différents cas d'équilibre.

Mots clefs : Urbanisation, division du travail, surplus agricole, système urbain monocentrique.

ABSTRACT

In this paper, we expose the economic conditions of cities emergence in a spatial general equilibrium framework. The presence of increasing returns based on the division of labour, transport costs and the possible existence of an agricultural surplus are enough to generate different possible urban equilibrium. A city may not be sustainable if internal transport costs are too high. On the other hand, a persistent migratory pressure may exist between the city and the surrounding rural hinterland if the urban labour market is saturated. In addition, we study the conditions of stability of the monocentric equilibrium in the different cases.

Key Words: Urbanization, division of labour, agricultural surplus, monocentric urban system.
JEL Classification: R13, R14, O18.

Table of Contents

List of figures

1 Introduction[1]

The objective of this paper is to enlighten the economic conditions of cities' emergence and the different possible urban outcomes in a closed economy. The study of the relationship between social and economic conditions and types of urbanization has resulted in an abundant literature (e.g. Weber 1947, Castells 1972, Bairoch 1985) and it is not our purpose to present a synthesis of these works. Rather, we would like to propose a model of a central idea, common to numerous authors. At the very beginning of his book, Bairoch states:

> It is necessary to stress here this especially crucial point: the existence of true urban centres presupposes not only a surplus of agricultural produce, but also the possibility of using this surplus in trade. And the possibilities of trade are directly conditioned by the size of the surplus relative to the amount of ground that has to be covered in transporting it from one place to another, for distance reduces the economic value of the surplus.

The first point means that agricultural productivity must be high enough to allow a part of the population not to be involved in food production, and the second point that the cost of transporting the agricultural surplus from the rural area to the city must be reasonable, otherwise a city cannot be sustained. But this does not explain why cities emerge: the agricultural surplus could be consumed locally by some non-agricultural activity and no city would exist. It is well known that some kind of indivisibility is necessary to generate agglomeration effects.[2] Here, we choose to rely on increasing returns due to specialization to give rise to agglomeration economies. In this paper, we show how these three aspects – agricultural surplus, transport costs and increasing returns – can be combined within an explicitly spatial framework and may generate various types of embryonic urban systems, depending on the prevailing economic conditions.

Three earlier approaches are particularly interesting and, in a way, our model constitutes a synthesis of these.

[1]This paper has benefited from numerous fruitful discussions with Hubert Jayet. I would also like to thank Denis Cogneau, Gene Grossman, Sandrine Mesplé-Somps and Philippe de Vreyer, and conferences and seminars participants at DIAL, CERDI and Université de Pau for helpful comments and suggestions on earlier versions. Financial support from Princeton University, where part of this work was elaborated, is gratefully acknowledged. The usual disclaimer applies.

[2]See Fujita and Thisse (2000) for an extensive survey of this question.

Duranton (1998) studies precisely the link between increasing returns and transport costs in order to illustrate the mechanism by which cities emerge and the transition from pre-industrial urbanization to modern urbanization. He shows the importance of urban institutions, such as guilds in the Middle Age, for the regulation of urban population in a pre-industrial context. However, Duranton's model is a-spatial and does not allow to grasp the role of urban-rural relations. Moreover, in our opinion, the method used to account for the role of agricultural surplus is not satisfactory.

On the opposite side, the spatial general equilibrium model by Nerlove and Sadka (1991)[3] is a formalization of von Thünen's works on location of crops around a city (see for instance Huriot (1994) for a presentation of Thünen's works). This model explains how a region's population is shared between the city and the rural area on one side and among the rural area – with respect to the distance from the city – on the other side. One of its interesting features is to present the dualistic model in a spatial framework, which allows to analyse the consequences of changes in transport costs. However, the paper does not consider the question of the emergence of the city, and assumes that returns in the urban sector are constant. Still, it seems important to be able to study both the formation of the city and urban-rural linkages in a unified framework.

Fujita, Krugman and Venables (1999) propose a model assuming the existence of increasing returns in the urban sector in a von Thünen-like spatial framework.[4] But their approach seems problematic for two reasons. First, the surplus issue is not taken into account, which leads to omit an important factor in the emergence of cities. More importantly, the cultivation intensity in the rural zone is exogenous, which greatly reduces the model's interest.

Finally, Fujita and Hamaguchi (2001) propose a slightly different version of this model, introducing the role of intermediate goods in the production of the manufacturing good. While agglomeration is due to consumers' preference for variety in Fujita, Krugman and Venables, Fujita and Hamaguchi put the variety in the intermediates of the manufacturing sector, which leads to new conclusions. However, the basic structure of the model remains the same.

The geographic scale for which our model seems relevant to us is a region in which trade exists between a city and its rural hinterland. The historical context is the emergence and development of cities in a pre-industrial society – typically in a South country, even if this not explicit in the model and it could therefore be interpreted in different ways. It is useful to

[3]The method used by the authors was first exposed by Samuelson (1983).

[4]We refer here to chapters 9 and 10 of their book, which content has been previously exposed in Fujita and Krugman (1995).

keep this reference in mind since it will help to evaluate the relevance of certain hypotheses and the model's significance.

The paper is organized as follows. In section 2, we present the model and detail its characteristics. In section 3, we explain the sources and the role of the different constraints on urbanization and urban development and analyse the possible outcomes of the model. In section 4, we study the conditions of stability of the monocentric system and some conclusions are provided in section 5.

2 The model

In this section, we briefly present the characteristics of the model. We consider a dualistic spatial economy. The manufacturing sector produces two distinct goods with increasing returns to scale due to specialization of workers, and the agricultural sector produces one good with decreasing returns to scale. Increasing returns provide a motive for agglomeration of the manufacturing activity which is therefore located in a unique place. This hypothesis will be reappraised later, when we study the stability of the monocentric equilibrium. The agricultural production is spread over the territory of the economy. For the sake of simplicity, we assume a unidimensional space. Thus, the economy is represented as a line segment, the length of which is determined endogenously. When a city exists, we assume that it does not occupy any space and, as long as it is unique, it is located at the centre of the segment, so as to minimize global transport costs.

All agents in the economy share the same type of preferences exhibiting a subsistence threshold in the consumption of agricultural products. This assumption is made in order to provide a rationale for the existence (or absence) of an agricultural surplus.

Total population of the economy is N and is constant. Populations of the rural area and the city are respectively denoted N_A and N_M.

2.1 The agricultural sector

We assume that agriculture requires both land (R) and labour (L), and that all productive combination of these factors must obey $y(r) = l(r)^{\gamma}$, where $y(r)$ is the output per unit of land ($q_A(r)/R(r)$) and $l(r)$ is the amount of labour per unit of land, that is the intensity of cultivation ($L(r)/R(r)$), all at distance r from the centre of the region. The elasticity of the output per unit of land with respect to l, γ, lies between 0 and 1. This specification only assumes that land yield increases with the intensity of cultivation, while labour productivity decreases, which has been explained by Boserup (1970) and is notably verified by Mazumdar (1965). Land quality is assumed to be homogenous over the territory. Since space is represented as a straight line, we can safely assume $R(r) = 1$, $\forall r \in [-f, f]$, f being the extensive margin of cultivation. The region being symmetrical, we can restrain the analysis to what happens in the segment $[0, f]$. The price of the agricultural good in the city is noted P_A.

Each farmer offers inelastically one unit of labour and we assume that the real income of a farmer located at r is made up of labour income and an equal share of land rent obtained at this distance. This is mainly assumed to avoid the presence of landowners and make the model simpler. Moreover, it seems to fit quite well what is known of land ownership in many traditional societies. However, assuming immobile landowners would not significantly change the qualitative results. Hence:

$$Y_A(r) = l(r)^{\gamma-1}. \tag{1}$$

We assume that rent reaches some exogenous minimum value at the region's extensive margin of cultivation f. Therefore, cultivation intensity at the margin amounts to $l(f) = v$.[5]

Total rural population of the region is :

$$N_A = 2\int_0^f L(r)\,dr = 2\int_0^f R(r)l(r)\,dr \tag{2}$$

and the global output of the sector is:

$$q_A = 2\int_0^f y(r)R(r)\,dr = 2\int_0^f R(r)l(r)^\gamma\,dr. \tag{3}$$

2.2 The manufacturing sector

The manufacturing sector produces two goods, M_1 and M_2. We assume that producing M_i ($i = 1, 2$) requires a certain number of inputs (or intermediate goods), elements of the set $\mathcal{D}_i = \{h_1, \ldots, h_n\}$ where $\text{card}(\mathcal{D}_i) = D_i$ is exogenous. We assume that factors are specific to each good, that is $\mathcal{D}_1 \cap \mathcal{D}_2 = \varnothing$, but $\mathcal{D}_1$ and $\mathcal{D}_2$ may be equipotent.

The only factor needed to produce intermediate goods is labour. First, each worker chooses the subsector (M_1 or M_2) which he wants to be involved in. Then, each worker has one unit of labour to allocate to the different goods of $\mathcal{D}_i$. He may decide to produce all of them, or

[5]Specifically, rent at r is $(1-\gamma)l(r)^\gamma$. If rent at f amounts to m, we get $l(f) = (m/(1-\gamma))^{1/\gamma}$. It is noticeable that it is impossible to assume a null rent at the extensive margin, for it would imply $l(f) = 0$ and $L(f) = 0$.

any subset of $\mathcal{D}_i$. The time devoted by an individual j to the production of input h is written $L_j(h)$, with $\sum_h L_j(h) = 1$. The amount of input h produced by j is then $T_j(h) = L_j(h)^{1+\alpha}$, where $\alpha > 0$ indicates the level of increasing returns provided by specialization. If all inputs have the same price, it is obvious that an agent gets an higher income as he becomes more specialized (i.e. he produces a smaller subset of $\mathcal{D}_i$). The total production of input h is then $T(h) = \sum_j^{N_i} T_j(h)$ where N_i is the number of agents producing inputs of M_i, with $N_1 + N_2 = N_M$.

Production of M_i is given by a CES aggregate of the different inputs available, belonging to $\mathcal{I}_i \subseteq \mathcal{D}_i$, with $\text{card}(\mathcal{I}_i) = I_i$:

$$q_i = \left[\sum_h^{I_i} T(h)^\rho\right]^{1/\rho} \tag{4}$$

where $1/\rho > 1+\alpha$ (we explain this hypothesis below) and the elasticity of substitution between two given inputs is $1/(1-\rho)$. Hence, profit of this sub-sector is:

$$\Pi_i = P_i\left[\sum_h^{I_i} T(h)^\rho\right]^{1/\rho} - \sum_h^{I_i} W(h)T(h) \tag{5}$$

where P_i is the nominal price of good M_i in the city and $W(h)$ is the price of input h. Maximizing the profit leads to the following first order condition:

$$\frac{\partial \Pi_i}{\partial T(h)} = 0 \Leftrightarrow P_i T(h)^{\rho-1}\left[\sum_h^{I_i} T(h)^\rho\right]^{(1/\rho)-1} = W(h). \tag{6}$$

Symmetry in the production function implies that each available input must be produced in equal amount $T(h) = T_i$, $\forall h \in \mathcal{I}_i$, which implies $L(h) = L_i = N_i/I_i$. Moreover, reward of each input must be the same at the sector equilibrium, and equal to the common marginal productivity, due to free exit and entry to the sector: $W(h) = W_i = P_i {I_i}^{(1/\rho)-1}$, $\forall h \in \mathcal{I}_i$. Finally, the production of each intermediate good is $T_i = (N_i/I_i)^{1+\alpha}$. Since each worker provides the same amount of labour, they earn the same income at the equilibrium of the sector and the subsets of $\mathcal{I}_i$ they produce are equipotent (though possibly different).

The production of good M_i can be rewritten as a function of the number of workers involved in the production of inputs for this good: $q_i = I^\eta {N_i}^{1+\alpha}$, where $\eta = (1/\rho) - 1 - \alpha$. Moreover, the condition $1/\rho > 1 + \alpha$ (i.e. $\eta > 0$) stated above implies that advantages related to inputs complementarity in production outweigh those linked to increasing returns to specialization. This ensures that all possible inputs are produced at the equilibrium (we therefore have $\mathcal{I}_i = \mathcal{D}_i$): if this was not true, a non-produced input would potentially have an infinite reward. The total production of good M_i is thus:

$$q_i = D_i{}^{\eta} N_i{}^{1+\alpha} \tag{7}$$

while the real income of a worker in the sub-sector i is:

$$Y_i = D_i{}^{\eta} N_i{}^{\alpha}. \tag{8}$$

The resulting production function exhibits increasing returns at the sector level, which makes it very similar to other examples in the literature (e.g. Henderson, 1974). More precisely, the production of each sub-sector is an increasing function of the specialization level of workers (N_i/D_i). When the number of workers employed in sector M_i increases, the level of specialization rises and the production increases consequently. But whenever N_i reaches D_i, employment and production can no longer increase, therefore $D_i{}^{1/\rho}$ is the maximum feasible level of production. This feature will have crucial consequences on the ability of a city to expand and on the relations between a city and its hinterland, since it will impact both on terms of trade and on migration outcome.

2.3 Prices and individual behaviour

In order to reflect transportation costs for goods, we assume that if one unit of good A (M) is shipped a distance s, $e^{-t_A s}$ ($e^{-t_M s}$) units actually arrive. Since the manufacturing good is produced in the city and the agricultural good is produced in the surrounding rural area, the nominal price of the agricultural good at a distance s from the centre is $P_A(s) = P_A e^{-t_A s}$, while the nominal price of the manufacturing good M_i is $P_i(s) = P_i e^{t_M s}$.

Since we want to exhibit the role of agricultural surplus in the emergence of cities, it seems appropriate to assume that each agent must consume at least a certain quantity of agricultural good: if food production is too small to guarantee enough food to each farmer, no surplus can exist, and therefore, no city will emerge. On the contrary, when global production exceeds global basic needs of farmers, a city can be sustained. (The precise conditions for a city to exist will be stated below.) It is possible to obtain such a pattern by assuming that food consumption does not increase beyond some level and all income exceeding this level is devoted to the manufacturing good (Duranton, 1998, 1999). However, we will rather adopt a more flexible form allowing the demand for food to increase with income. Moreover, we may also want to consider what happens when no subsistence threshold is required. Hence, the common utility function is:

$$U = Cq_1{}^{\beta/2} q_2{}^{\beta/2} \left[q_A - z\right]^{1-\beta} \tag{9}$$

where $C = (\beta/2)^{-\beta}(1-\beta)^{\beta-1}$.

3 The isolated region, with or without a city

In this section, we present the method used to determine the possible equilibriums of the economy and the different possible outcomes of the model. We illustrate the different cases with numerical simulations.

3.1 Migration process

Utility maximization by all agents leads to the following set of demand functions, where q_{ij} is the demand of good j by an agent of the sector i:

$$\begin{array}{lll}
q_{11} = \frac{\beta}{2}\left[Y_1 - z\frac{P_A}{P_1}\right] & q_{12} = \frac{\beta}{2}\left[Y_1\frac{P_1}{P_2} - z\frac{P_A}{P_2}\right] & q_{1A} = (1-\beta)Y_1\frac{P_1}{P_A} + \beta z \\
q_{21} = \frac{\beta}{2}\left[Y_2\frac{P_2}{P_1} - z\frac{P_A}{P_1}\right] & q_{22} = \frac{\beta}{2}\left[Y_2 - z\frac{P_A}{P_2}\right] & q_{2A} = (1-\beta)Y_2\frac{P_2}{P_A} + \beta z \\
q_{A1} = \frac{\beta}{2}\left[Y_A(r) - z\right]\frac{P_A(r)}{P_1(r)} & q_{A2} = \frac{\beta}{2}\left[Y_A(r) - z\right]\frac{P_A(r)}{P_2(r)} & q_{AA} = (1-\beta)Y_A(r) + \beta z
\end{array}$$

As a consequence, indirect utilities of the different agents are:

$$U_1 = \left[Y_1P_1 - zP_A\right] P_1{}^{-\beta/2} P_2{}^{-\beta/2} P_A{}^{\beta-1} \tag{10}$$

$$U_2 = \left[Y_2P_2 - zP_A\right] P_1{}^{-\beta/2} P_2{}^{-\beta/2} P_A{}^{\beta-1} \tag{11}$$

$$U_A = \left[Y_A(r)P_A(r) - zP_A(r)\right] P_1(r)^{-\beta/2} P_2(r)^{-\beta/2} P_A(r)^{\beta-1} \tag{12}$$

These equations have important consequences. The real income of manufacturing workers should not be lower than z, otherwise they could not even consume the subsistence minimum. Since the same condition applies for farmers, the following two inequalities must be satisfied:

$$Y_i\frac{P_i}{P_A} \geq z \Leftrightarrow D_i{}^{\eta} N_i{}^{\alpha}\frac{P_i}{P_A} \geq z \tag{13}$$

$$Y_A(r) \geq z \Leftrightarrow l(r)^{\gamma-1} \geq z, \forall r \in [0, f]. \tag{14}$$

Agents wish to have the highest possible utility level. If an agent observes that it is possible to have a better situation in some other location or sector, he instantly and costlessly migrates to this place or switches to this sector. But we assume that this migration process is sequential and operates first within sectors or locations. As a consequence, we assume that – when the structure of the equilibrium makes it possible – utility is always equal among urban workers and among farmers. Considering manufacturing workers, this leads to $U_1 = U_2$, that is:

$$D_1{}^{\eta} N_1{}^{\alpha} P_1 = D_2{}^{\eta} N_2{}^{\alpha} P_2. \tag{15}$$

A similar condition for farmers is $U_A(s_1) = U_A(s_2), \forall (s_1, s_2) \in]0, f]$. This leads to the following endogenous intensity of cultivation in any location r of the rural area:

$$l^*(r) = \left[\left(v^{\gamma-1} - z \right) e^{\beta(t_A+t_M)(r-f)} + z \right]^{1/(\gamma-1)}. \tag{16}$$

As is customary in this type of models, we note that the intensity of cultivation is strictly decreasing with the distance from the centre of the region, provided that $v^{\gamma-1} > z$, which is always true if condition (14) holds.

3.2 Constraints on urbanization and determination of the equilibrium

Equilibrium is characterized by three conditions or equalities. First, good market equilibrium, that is global supply equals global demand for goods A, M_1 and M_2. Second, intrasectoral and intersectoral migratory equilibrium, which means that each worker in the economy has the same level of utility. Third, population equilibrium implies that employment in all sectors equals the total population of the economy, i.e. the population is constant and no migration is possible from or to the rest of the world. The variables to be determined are the real prices of goods and the way the population is allocated among the sectors and locations of the economy.

Since we are primarily interested in prices determination and intersectoral population allocation, we assume that population equilibrium is always met: $N = N_M + N_A$. Since we also assume that intrasectoral migratory equilibrium is always met before intersectoral equilibrium, we are left with the good market equilibrium and the intersectoral migratory equilibrium.

The system of three equations describing the good market equilibrium is easily built from the set of demand functions displayed above. Due to Walras law, one of these equations is redundant – say the one relating to M_2 – and we therefore choose the agricultural good as a numeraire. Combining the two remaining equations leads to an equality between real outputs of the two manufacturing sub-sectors : $D_1{}^{\eta} N_1{}^{1+\alpha} P_1 = D_2{}^{\eta} N_2{}^{1+\alpha} P_2$.[6] Together with intra-urban migratory equilibrium described in equation (15), this leads to $N_1 = N_2$. This result,

[6] Details on the determination of the good market equilibrium may be found in appendix A.

stating that the two manufacturing sub-sectors have the same size at the equilibrium is clearly due to the equal weighting of M_1 and M_2 in the utility function.[7] Using these results and the population equilibrium, the market equilibrium now reduces to the following equation:

$$P_1 = \frac{\beta}{1-\beta} \frac{\int_0^f \left(l^*(r)^\gamma - l^*(r)z\right) e^{-t_A r}\, dr - \left(\frac{N}{2} - \int_0^f l^*(r)\, dr\right) z}{D_1{}^\eta \left(\frac{N}{2} - \int_0^f l^*(r)\, dr\right)^{1+\alpha}}. \tag{17}$$

Intersectoral migratory equilibrium implies that a farmer gets the same utility as a manufacturing worker ($U_i = U_A$). This is rewritten as:

$$P_1 = \frac{(v^{\gamma-1} - z)e^{-(t_A+t_M)\beta f} + z}{D_1{}^\eta \left(\frac{N}{2} - \int_0^f l^*(r)\, dr\right)^{\alpha}}. \tag{18}$$

These two equations determine jointly the real price of the manufacturing good M_1 (P_1) and the extensive margin of cultivation (f), i.e. the size of the region. The price of M_2 is then deduced from either the intra-urban migratory equilibrium or the good market equilibrium: $P_2 = (D_1/D_2)^\eta P_1$. The sectoral allocation of population is also easily determined: the rural population is calculated using equation (2) and the level of specialization in the two urban sub-sectors can then be computed.

Several comments may be made at this stage. Due to nonlinearity, the system (17)–(18) cannot be solved analytically. It is nevertheless possible to obtain some interesting results from numerical simulations. Before carrying out these analyses, it is however essential to fully describe and understand the functioning of the model. One of the main characteristics of this model is to underline two potential limits to the emergence and development of a city: one related to food production and the other linked to urban employment.

The first constraint derives from the subsistence threshold and has several implications. There can exist economic conditions where the existence of a city is impossible. A minimum city, inhabited by one individual requires two conditions.[8] First, this individual must have means of subsistence, which implies that the global agricultural surplus in the city is large enough to meet his needs, that is :

$$2\beta \int_0^f \left(l^*(r)^\gamma - l^*(r)z\right) e^{-t_A r}\, dr \geq z. \tag{19}$$

[7] But whenever D_1 and D_2 are different, the level of specialization of workers will not be identical in the two sub-sectors, which will lead to interesting features.

[8] To be fully rigorous, there should actually be two urban workers in the minimal city: one in each sub-sector.

Second, this lonely urban worker must earn enough money to purchase the minimum amount of food, that is ${D_i}^{\eta} P_i > z$. In addition, once the city exists, its population cannot exceed a certain limit, since migration from the agricultural sector to the city can potentially reduce food production. In the extreme case where the city population reaches N_{Mmax}, the agricultural sector generates a minimum surplus, only providing a subsistence level food consumption to urban workers:

$$N_{Mmax} = \frac{2\beta \int_0^f \left(l^*(r)^\gamma - l^*(r) z\right) e^{-t_A r}\, dr}{z}. \tag{20}$$

Urban population should never reach this level since it implies a null utility in the city and a positive utility for farmers ; therefore this situation cannot be a equilibrium. However, N_M / N_{Mmax} can be used as an indicator of urban saturation.

The second constraint comes from the maximum number of workers in each sub-sector (D_1 and D_2). Since nothing *a priori* prevents the model to lead to $N_i > D_i$ at the equilibrium, there must exist some urban institution preventing people to try and enter an already crowded sub-sector. Whenever N_i reaches D_i, the behaviour of the model is therefore deeply modified and corner solutions appear. Let us begin with the case where only one of the two sub-sectors is saturated, while workers in the other are not fully specialized. For this purpose, let us assume for instance $D_2 < D_1$, which implies that the sub-sector M_2 will always be saturated before the other. First, the equality of real output between the two sub-sectors must hold : ${D_1}^{\eta} {N_1}^{1+\alpha} P_1 = {D_2}^{1/\rho} P_2$. The market equilibrium holds as well:

$$P_1 = \frac{\beta}{2(1-\beta)} \frac{2\int_0^f \left(l^*(r)^\gamma - l^*(r) z\right) e^{-t_A r}\, dr - \left(N - 2\int_0^f l^*(r)\, dr\right) z}{{D_1}^{\eta} \left(N - D_2 - 2\int_0^f l^*(r)\, dr\right)^{1+\alpha}}. \tag{21}$$

If there was no labour market constraint N_2 would tend to increase beyond D_2. Therefore, we have $U_2 > U_1$: there is no intra-urban migratory equilibrium in this situation. However, intersectoral migratory equilibrium between sub-sector M_1 and the agricultural sector does hold:

$$P_1 = \frac{(v^{\gamma-1} - z) e^{-(t_A + t_M)\beta f} + z}{{D_1}^{\eta} \left(N - D_2 - 2\int_0^f l^*(r)\, dr\right)^{\alpha}}. \tag{22}$$

Equations (21)–(22) fully determine the equilibrium of the model in this case.

Now, let us turn to the case where both manufacturing sub-sectors are saturated ($N_1 = D_1$ and $N_2 = D_2$). In this case, the agricultural sector must accommodate the rest of the population, which implies that f is a solution of $\int_0^f l^*(r)\, dr = (N - D_1 - D_2)/2$ and that the intersectoral migratory equilibrium (equation (18)) cannot hold any longer. The equality between real outputs of the manufacturing sub-sectors remains valid and determines the price of one good as a function of the other: $P_2 = (D_1/D_2)^{1/\rho} P_1$. As a consequence, intra-city equilibrium is also violated whenever $D_1 \neq D_2$. Finally, the remaining real price, here P_1, is determined using the good market equilibrium (equation (17)):

$$P_1 = \frac{\beta}{2(1-\beta)} \frac{2\int_0^f \left(l^*(r)^\gamma - l^*(r)z\right) e^{-t_A r}\, dr - (D_1 + D_2)z}{{D_1}^{1/\rho}}. \tag{23}$$

It is important to underline that these constrained equilibriums are situations where some agents are better off than some others and where some migratory pressure persists. For instance, in the case were the sub-sector M_2 is saturated, workers in M_2 have a higher level of utility than those in M_1 because the former are more specialized than the latter. But M_1 workers cannot switch to M_2 since they would end up unemployed. These situations may possibly lead to conflicts or other strategies to deal with these inequalities. It is also possible to refine the model in order to introduce urban unemployment, but this is beyond the scope of this paper.

3.3 Some results

Now we have explored the different possible outcomes of the model, let us give a more detailed account of the role of some key parameters of the model. First, as the total population increases, both rural and urban populations increase. However, due to the subsistence constraint, the rural population increases more than the urban population and the urbanisation rate is therefore decreasing. The evolution of prices is more complex, since it is modified by the nature of the equilibrium. Starting from a low level of total population, relative prices of the manufacturing goods begin to fall because of increased efficiency in the urban sub-sectors. If one of the sub-sectors happens to be saturated for some level of the total population, the price of this good rises. This is exemplified in figure 1 where the sub-sector M_2 is saturated before M_1.[9] Welfare levels are also affected when the total population grows. The general pattern when no corner equilibriums appear is the following : utility of all agents first increases thanks to beneficial effects of a bigger and more efficient manufacturing sector and then declines when disadvantages due to a farther extensive margin and increasing global transport costs outweigh the advantages. Figure 2 presents an example of such a case, where there exists

[9] Figure 1 uses the following the following parameters: $D_1 = 15$, $D_2 = 10$, $\beta = 0.6$, $\gamma = 0.7$, $\alpha = 0.3$, $\rho = 0.5$, $v = 0.02$, $z = 0.2$, $t_A = t_M = 0.05$.

an optimum city size.[10] When corner solutions appear, utility levels of the different agents diverge: urban workers in a saturated sub-sector are better off than workers in a non-saturated sub-sector or farmers. Depending on the parameters of the model, and particularly transport costs, the pattern of utility as a function of total population may be very different, as reveals the comparison between figures 2 and 3.[11]

What happens to the sectoral allocation of the population when transport costs decrease? A decrease in the transport cost of the agricultural good will unambiguously increase the global agricultural surplus in the city and therefore the potential size of the city. This effect works perfectly when the economy is in interior equilibriums: the urbanisation rate increases as t_A decreases. But improvements in transport of good A have only a modest impact on urbanisation when one urban sub-sector is already saturated – because this part of the city can no longer expand – and have no effect at all if both sub-sectors are saturated. Transport cost of the manufacturing goods does not affect the global surplus and its decrease generally induce a very modest concave evolution of the urbanisation rate in the case of interior equilibriums and no evolution at all in the corner equilibrium where both sub-sectors are saturated. Relative prices of the manufacturing goods in the city always increase when transport costs decrease mainly because the extensive margin of cultivation moves away and the agricultural good becomes cheaper. Utility levels increase when transport costs decrease since global surplus increase and real prices are higher (this is beneficial to urban workers because their real income rise). Farmers enjoy a lower intensity of cultivation everywhere and their real income increase. As we already know, when corner solutions appear, utility levels diverge and urban workers benefit more than farmers from the decrease of transport costs. In these cases, further reductions in transport costs entail more and more urban-rural, and possibly intra-urban, inequality.

4 Stability of the monocentric equilibrium

Until now, we have assumed that manufacturing activities only take place in the city. If we give up this assumption, a more flexible spatial distribution of economic activity can emerge, where manufacturing production may take place in any location. However, even

[10]Figure 2 uses the following the following parameters: $D_1 = 15$, $D_2 = 10$, $\beta = 0.6$, $\gamma = 0.7$, $\alpha = 0.7$, $\rho = 0.5$, $v = 0.2$, $z = 0.2$, $t_A = t_M = 0.75$.

[11]Figure 3 uses the following the following parameters: $D_1 = 15$, $D_2 = 10$, $\beta = 0.6$, $\gamma = 0.7$, $\alpha = 0.7$, $\rho = 0.5$, $v = 0.2$, $z = 0.2$, $t_A = t_M = 0.05$.

if production of goods M is possible in the rural zone, it does not imply that it actually happens, since the central city may generate enough agglomeration economies to concentrate all the manufacturing activities. In order to know whether such a situation is likely, that is to say whether the monocentric equilibrium is stable, we need to compare the utility of an agent in location $s \in [0, f]$ and sector X ($X = M_1, M_2, A$) with the utility he would obtain in a different location and/or sector. For instance, if a manufacturing worker of the central city finds it profitable to migrate to some location s and settle a new manufacturing production site there, the monocentric equilibrium is not stable. It is important to distinguish this kind of analysis from the process in which new cities emerge. While instability of the monocentric system is a prerequisite for having new cities in the region, studying their emergence requires to make further assumptions about the migration process. Furthermore, we believe that the unidimensional spatial framework is not well suited to analyse the emergence of new cities in a region and that this should only be done in a bidimensional framework.[12]

Depending on which type of monocentric equilibrium is prevailing, the condition for an agent to deviate – and therefore for the monocentric system to be unstable – is different. Let us begin with the interior equilibrium. Since all agents have the same level of utility at the equilibrium in this case, the location and activity of the deviating agent does not matter. For instance, let us consider the deviation of a farmer located in s. He will choose to deviate to activity M_1 or M_2 if he can enjoy an higher level of utility. We define the deviation potential $\Gamma(X_r, Y_s)$ of a given agent as the ratio between the utility he would obtain in sector X and location r and his current utility, in sector Y and location s. If an agent is in Y_s with a potential higher than 1, he should move to X_r. A farmer located in s switches to the manufacturing sub-sector M_i if $\Gamma(M_{i,s}, A_s) > 1$, that is:

$$\frac{{D_i}^{\eta} P_i e^{(t_A+t_M)s} - z}{v^{\gamma-1} - z} e^{(t_A+t_M)\beta(f-s)} > 1. \tag{24}$$

Distance from the centre plays a crucial role: the farther an agent is from the city, the easier it is for him to deviate. Employment of the sub-sector M_i in the central city has a negative effect on the deviation potential since higher employment generates increasing returns and hence higher utility level for workers of the central city. The subsistence threshold has a positive effect on Γ whenever $e^{(t_A+t_M)s} > {N_1}^{\alpha}$. Since ${D_1}^{\eta} P_1 = {D_2}^{\eta} P_2$ in this monocentric interior equilibrium, agents have exactly the same incentive to deviate to M_1 or M_2.

[12]In a unidimensional economy, new cities can only appear symmetrically on each side of the central city, while in a bidimensional economy, more than two cities can emergence simultaneously.

If the economy is in the first corner equilibrium, i.e. the case where only one of the two sub-sectors is saturated (say, M_2), two types of deviation are possible. First, farmers or workers of M_1 may switch to M_1 or M_2 in some location of the rural area. Second, workers of M_2 may also want to deviate. In the first case, the deviation potential is similar to (24) and it is highly probable that deviating to M_2 is more profitable than deviating to M_1, since the constraint on the urban labour market induces an important distortion on P_2, favourable to producers of M_2. In the second case, the condition $\Gamma(M_{i,s}, M_{2,0})$ is rewritten:

$$\frac{{D_i}^{\eta} P_i e^{(t_A+t_M)s} - z}{{D_2}^{\eta} {N_2}^{\alpha} P_2 - z} e^{-(t_A+t_M)\beta s} > 1. \tag{25}$$

If the economy is in the second corner equilibrium, where both sub-sectors are saturated, farmers will deviate first because they have a lower utility level than urban workers. The deviation condition is again the same as in the interior equilibrium, (24). In this type of corner equilibrium, the monocentric system is all the more unstable as the manufacturing activity is profitable and the agricultural sector is repulsive. Hence, the potential extent of increasing returns (D_i) and the price of the manufacturing good (P_i) have a positive effect on Γ, while v has a negative effect. Distance still has a positive effect: it is easier to deviate when one is farther from the central city.

5 Conclusion

In this paper, we propose a spatial general equilibrium model of cities' formation in which the source of agglomeration is the division of labour in the manufacturing sector, while two constraints limit this effect: the subsistence constraint and the urban employment constraint.

Four types of equilibrium are highlighted: when the agricultural productivity is insufficient or the subsistence threshold too high, it is possible that no city emerge and the economy does not produce manufacturing goods. However, in many cases, a city can exist. If the two urban subsectors are not saturated, all agents enjoy the same level of welfare and no persistent migratory pressure exists at the equilibrium. On the contrary, if one or both sub-sectors are saturated, utility levels diverge and rural-urban and/or intra-urban inequalities appear. This kind of situations cannot be solved unless some technological progress in the manufacturing sector allows to relax the migratory pressure. Indeed, a deepening in the division of labour would permit to increase urban employment and therefore to go back to an interior equilibrium. If the economy is in a corner equilibrium, a decrease in transport costs may increase the rural-urban welfare gap and therefore aggravate the possible conflicts between urban workers and farmers.

In all the cases where a central city exists, it is possible to determine under which conditions the monocentric system is stable and "where" a settler needs to go in the rural area to defeat the central city.

This model may have several developments. First, it may be interesting to introduce urban unemployment to get a more realistic account of what happens in the city when one subsector (or both) becomes saturated, and how farmers react to the level or urban unemployment (though this would probably be done along the lines of Harrod and Todaro (1970) or their followers). Second, analysing trade and migrations between two regions of this type is certainly worthwhile since it would permit to understand both what happens at the local level and at the interregional level, notably in terms of specializations. Third, it is also probably possible to describe how new cities can emerge when the monocentric system and to compare the results with those of Fujita, Krugman and Venables to see if our hypotheses lead to significantly different urban systems.

A Details on the good market equilibrium

The market equilibrium for the manufacturing good i ($i = 1, 2$) is written:

$$\frac{2-\beta}{\beta} {D_i}^{\eta} {N_i}^{1+\alpha} P_i = {D_j}^{\eta} {N_j}^{1+\alpha} P_j - (N_i + N_j) z P_A + 2P_A \int_0^f (l(r)^{\gamma} - l(r)z) e^{-t_A r}\, dr \quad (26)$$

where the index j represents the other manufacturing good. The market equilibrium for the agricultural good is:

$$\frac{1-\beta}{\beta} ({D_i}^{\eta} {N_i}^{1+\alpha} P_i + {D_j}^{\eta} {N_j}^{1+\alpha} P_j) + (N_i + N_j) z P_A = 2P_A \int_0^f (l(r)^{\gamma} - l(r)z) e^{-t_A r}\, dr. \quad (27)$$

Combining the two equations for the manufacturing goods leads to the equality between the real outputs of the two sub-sectors: ${D_1}^{\eta} {N_1}^{1+\alpha} P_1 = {D_2}^{\eta} {N_2}^{1+\alpha} P_2$. Then, due to Walras law, one of the three equations is redundant and we can choose a numeraire. We choose to drop the equation for M_2 and to set $P_A = 1$. After adequate substitutions, we find:

$$\frac{1-\beta}{\beta} {D_1}^{\eta} {N_1}^{1+\alpha} P_1 + N_1 z = \int_0^f (l(r)^{\gamma} - l(r)z) e^{-t_A r}\, dr. \quad (28)$$

B Solving the general equilibrium

Eliminating P_1 from the system (17)–(18) leads to the following equation in f:

$$V(f) = A(f)B(f) - C(f) \quad (29)$$

where

$$A(f) = \frac{N}{2} - \int_0^f l^*(r)\, dr \quad (30)$$

$$B(f) = (1-\beta)(v^{\gamma-1} - z) e^{-\beta(t_A + t_M)f} + z \quad (31)$$

$$C(f) = \beta(v^{\gamma-1} - z) e^{-\beta(t_A+t_M)f} \int_0^f l^*(r) e^{\beta(t_A+t_M)r} e^{-t_A r}\, dr. \quad (32)$$

Solving $V(f) = 0$ leads to the optimal value of the extensive margin of cultivation we are looking for. Unfortunately, it is impossible to find an analytical solution to this equation and we have to resort to numerical methods. In the next appendix, we show that $V(f)$ is a monotonic function, so that the equilibrium is unique when it exists. The equilibrium value of f is then used to calculate N_1 (and N_2, which is equal) with $N_1 = A(f^*)$. Apart from the case where no city exists, there are three possibilities, depending on the value taken by N_1

and N_2. If we obtain $N_1 \leq D_1$ and $N_2 \leq D_2$, the economy is in the interior equilibrium case and prices are computed using for instance equation (17).

If we have $N_1 \leq D_1$ and $N_2 > D_2$, we are confronted with the first corner equilibrium, where the smallest sub-sector is saturated. We then solve numerically the system (22)–(21) and obtain new values for f, N_1 and the prices, while N_2 is set equal to D_2.

If we obtain $N_1 > D_1$ and $N_2 > D_2$, the economy is in the second type of corner equilibrium, where both sub-sectors are saturated. Thus f is determined by $\int_0^f l^*(r)\,dr = (N - D_1 - D_2)/2$ and P_1 by equation (23).

C Uniqueness of the monocentric interior equilibrium

In order to prove the uniqueness of the monocentric interior equilibrium, we only need to demonstrate that $V(f)$ is strictly monotonic, that is $V'(f) = A'(f)B(f) + B'(f)A(f) - C'(f)$ is strictly positive or strictly negative. For relevant values of the parameters, $A(f) > 0$ and $B(f) > 0$. Therefore, determining the signs of A', B' and C' allows to know the sign of V'.

Sign of $A'(f)$

The function $A'(f)$ is written:

$$\frac{dA(f)}{df} = -\frac{d}{df}\left[\int_0^f l^*(r)\,dr\right] \tag{33}$$

$$= -\left[l^*(f) + \int_0^f \frac{\partial l^*(r)}{\partial f}\,dr\right]. \tag{34}$$

We have $l^*(f) = v > 0$. Moreover,

$$\frac{\partial l^*(r)}{\partial f} = \frac{-\beta(t_A + t_M)}{\gamma - 1} l^*(r)^{2-\gamma}(l^*(r)^{\gamma-1} - z) \quad \geq 0, \forall r > 0. \tag{35}$$

Therefore

$$\int_0^f \frac{\partial l^*(r)}{\partial f}\,dr \quad \geq 0, \forall f > 0 \tag{36}$$

which allows to conclude that $A'(f) \leq 0$, $\forall f > 0$.

Sign of $B'(f)$

The derivative of B

$$\frac{dB(f)}{df} = -\beta(1-\beta)(v^{\gamma-1} - z)(t_A + t_M)e^{-\beta(t_A+t_M)f} \tag{37}$$

is strictly negative for relevant values of the parameters.

Sign of $C'(f)$

The function $C'(f)$ is:

$$\frac{dC(f)}{df} = \beta \frac{d}{df} \left[\int_0^f (l^*(r)^\gamma - zl^*(r)) e^{-t_A r} \, dr \right] \tag{38}$$

$$= \beta(v^\gamma - vz)e^{-t_A f} + \beta \int_0^f \frac{\partial l^*(r)}{\partial f} (\gamma l^*(r)^{\gamma-1} - z) e^{-t_A r} \, dr. \tag{39}$$

The sign of $C'(f)$ seems indeterminate since it is difficult to know if the integral is negative or positive. However, it is obvious that the value of this integral is all the more important that the agricultural surplus is big. In the worst possible case, surplus is zero everywhere ($l^*(r)^{\gamma-1} = z$) and it appears that $C'(f) = 0$ (the derivative of $l^*(r)$ with respect to f is zero in this case). Therefore, we can legitimately argue that $C'(f) \geq 0$, $\forall f > 0$.

Sign of $V'(f)$

As a consequence, $V'(f) < 0$, $\forall f > 0$ and $V(f)$ is a strictly decreasing function, which implies that $V(f) = 0$ has at most one solution. Since $V(0) = (N/2)[(1-\beta)v^{\gamma-1} + \beta z]$ is strictly positive, $V(f) = 0$ has exactly one solution if $\lim_{f \to +\infty} V(f) < 0$.

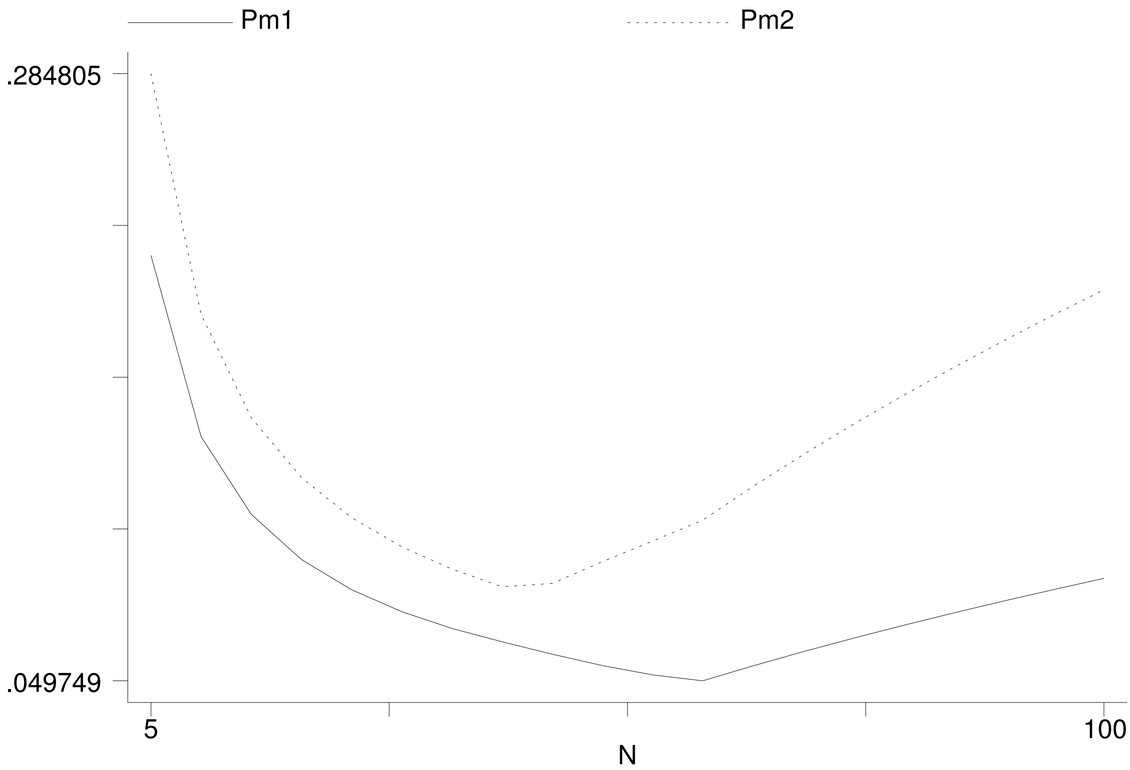

Figure 1: Evolution of relative prices when total population grows

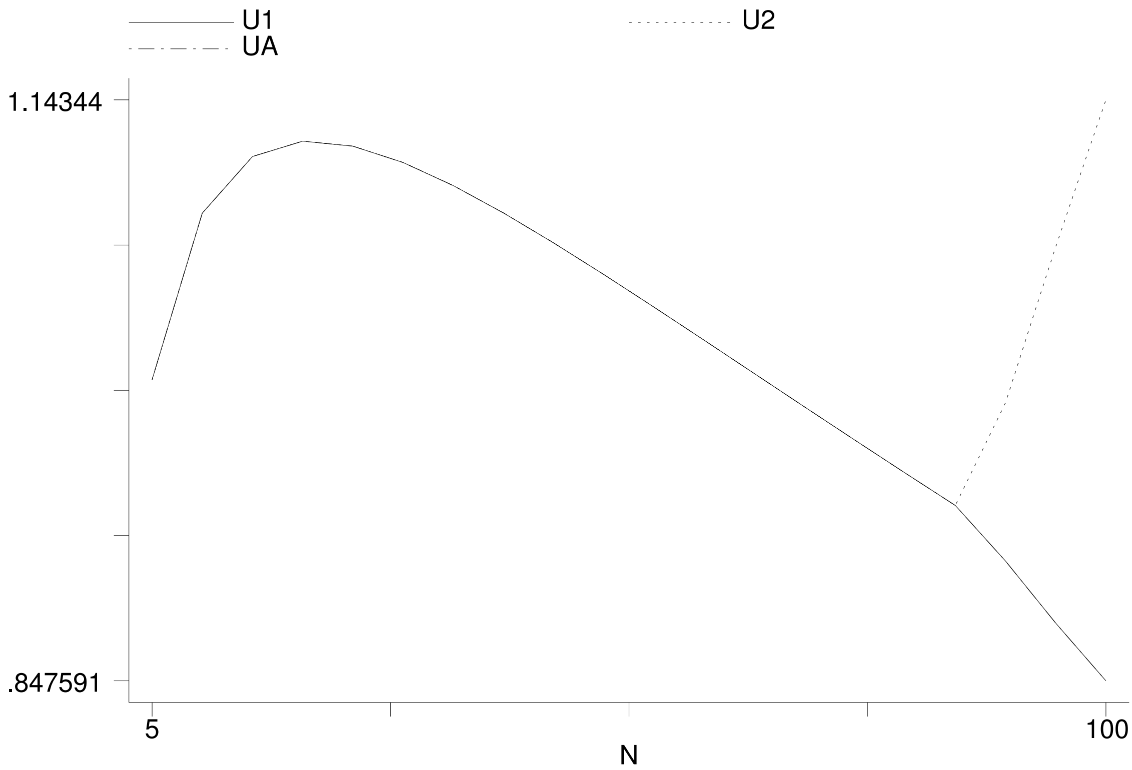

Figure 2: Evolution of utility levels when total population grows: Optimum city size

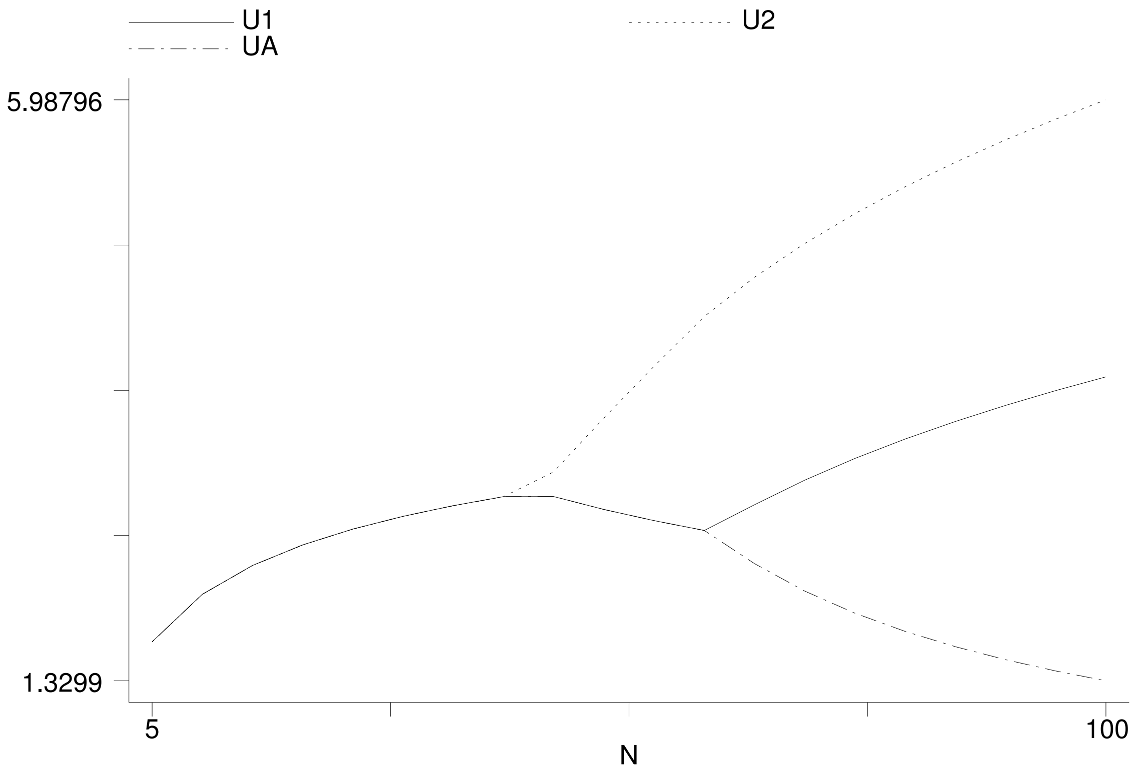

Figure 3: Evolution of utility levels when total population grows: Rising inequalities

References

[1] Bairoch, P., 1985. *De Jéricho à Mexico. Villes et économie dans l'histoire*. Paris: Gallimard. English translation (1988): *Cities and economic development. From the dawn of History to the present*. University of Chicago Press.

[2] Beckmann, M. J., 1972. Von Thünen revisited: A neoclassical land use model. *Swedish Journal of Economics*, 74, 1–7.

[3] Boserup, E., 1970. *The conditions of agricultural growth : the economics of agrarian change under population pressure*. London: Allen and Unwin.

[4] Castells, M., 1972. *La question urbaine*. Paris: Maspero. English translation (1977): *The urban question : a Marxist approach*. Cambridge, Mass.: MIT Press.

[5] Dixit, A. K. and J. E. Stiglitz, 1977. Monopolisitc competition and optimum product diversity. *American Economic Review*, 67, 297–308.

[6] Duranton, G., 1998. Labor specialization, transport costs, and city size. *Journal of Regional Science*, 38, 553–573.

[7] Fujita, M. and N. Hamaguchi, 2001. Intermediate goods and the spatial structure of an economy. *Regional Science and Urban Economics*, 31, 79–100.

[8] Fujita, M. and P. Krugman, 1995. When is the economy monocentric ?: von Thünen and Chamberlin unified. *Regional Science and Urban Economics*, 25, 505–528.

[9] Fujita, M., P. Krugman and A. Venables, 1999. *The spatial economy*. Cambridge, Mass.: MIT Press.

[10] Fujita, M. and J.-F. Thisse, 2000. The formation of economic agglomerations: Old problems and new perspectives. In: J.-M. Huriot and J.-F. Thisse, *Economics of cities*, pp.3–73. Cambridge: Cambridge University Press.

[11] Harris, J. R. and M. P. Todaro, 1970. Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60, 126–142.

[12] Henderson, J. V., 1974. The sizes and types of cities. *American Economic Review*, 64, 640–656.

[13] Huriot, J.-M., 1994. *Von Thünen. Économie et espace*, Bibliothèque de Science Régionale. Paris: Economica.

[14] Mazumdar, D., 1965. Size of farm and productivity: A problem of Indian peasant agriculture. *Economica*, 32, 161–173.

[15] Nerlove, M. and E. Sadka, 1991. Von Thünen's model of the dual economy. *Journal of Economics*, 54, 97–123.

[16] Samuelson, P., 1983. Thünen at two hundred. *Journal of Economic Literature*, 21, 1468–1488.

[17] Weber, M., 1947. *Die Stadt.* Tübingen: J.C.B. Mohr. English translation (1958): *The City.* Glencoe, Ill.: Free Press.