TAMPERE ECONOMIC WORKING PAPERS
NET SERIES

THE CITY AS A CLUB

Hannu Laurila

Working Paper 27
December 2003
http://tampub.uta.fi/econet/wp27-2003.pdf

DEPARTMENT OF ECONOMICS
FIN-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191
ISBN 951-44-5873-7

THE CITY AS A CLUB

5.5.2004

Hannu Laurila

Department of Economics
FI-33014 University of Tampere

Abstract:

The paper presents a club theoretic model of a city. In the model welfare in a city depends on its size because of positive and negative externalities (or agglomeration economies and diseconomies) generated by close proximity of people and economic activities. Technological externalities appear in people's utility functions and pecuniary externalities are due to shortcomings of the price mechanism. Both types of externalities are internalised in the city and act as centripetal and centrifugal forces that explain migration and thus the formation and development of the city. The issue of optimal city size is studied and it is found that the individually originated formation mechanism of the city may lead to non-optimal solutions.

Key words: agglomeration, city size, club theory, externalities
JEL classification: 931

1 Introduction

Cites are spatial concentrations of people and economic activities. The dominant reasons behind the existence of modern cities are connected to the economic performance of the market. Since most market activities are very spatial in nature, the proximity of people and the functions that they engage with in their everyday life has obvious implications to the welfare that the economic environment is able to create. People choose their place of living so as to maximise their welfare and make calculations of benefits and costs attached to different sites. Observed growth of some cities is simply to say that they can create higher welfare for their citizen than those places that people are leaving from. So, the existence and development of cities must be explained by their capability to create welfare.

An economic explanation of the existence and development of cities consists of centripetal forces such that draw people closer to each other, and of centrifugal forces such that throw them apart. A standard presumption in the literature of urban economics is that there is a connection between city size and welfare. This kind of a connection can be explained by so called agglomeration economies that arise in the production and consumption of the outputs of the local private and public sectors in the city. According to the theory of agglomeration economies an increase of the size of a city yields benefits and cost savings, but eventually it may bring up also falling benefits and increasing costs. Main gains caused by agglomeration economies are experienced as higher labour incomes brought up by specialisation and increased choices in work, wider consumption facilities of private and public goods and services, cost savings in arranging the provision of collective goods, larger variety of leisure activities and better opportunities for all kinds of interaction. Among the main disadvantages caused by agglomeration diseconomies are increased housing costs, commuting costs and other costs of living, high local taxes and tariffs, increased risks of health and property damage caused by pollution and crime, aggravating social problems and other such congestion costs.

Agglomeration economies and diseconomies are due to externalities and economies of scale. Without these kinds of market failures a totally dispersed spatial pattern would sustain. In the presence of externalities, both positive and negative, people not only respond to the observed benefits and costs but also affect them and thereby the welfare anticipated by others by their choices. From club-theoretic viewpoint, the essential nature of the city is to pool up the positive and negative effects of agglomeration and to internalise the externalities into people's everyday welfare experiences. Because the externality problem is so apparent in the context of a city, city size must be a central economic variable - it is the optimal size of the city that actually facilitates full internalisation of externalities.

The paper examines the formation and development of a city based on decision-making of utility maximising people or households. Firms are ignored by assuming that they decide on their production and location so as to maximise their profits, and the provision of local collective goods is assumed to be efficient. This is to say that the decisions of firms and the local public sector, which are also affected by agglomeration economies and diseconomies, are properly reflected by the market information faced by the people. The paper proceeds as follows. Chapter 2 discusses the connection of agglomeration economies and diseconomies and city size, and chapter 3 transforms individual utility maximisation into a club-theoretic framework. Chapter 4 illustrates the model of a city, and chapter 5 compares possible solutions for city size from the point of view of efficiency. Chapter 6 concludes the findings.

2 Agglomeration economies and externalities

The centripetal and centrifugal forces that explain the formation of a city are here simply called agglomeration economies and diseconomies (Krugman, 1991; Fujita & Thisse, 2002). People gather together because of agglomeration economies, and they drift apart because of diseconomies that arise in overly concentrated surroundings. From the point of view of an individual household, the effects of agglomeration can be divided into direct and indirect effects (cf. Richardson, 1973).

The direct effects are the household effects that are experienced in the local goods and factor markets. The variety, quality and attainability of private and collective goods, occupation, housing, leisure activities, environment etc. are all locally determined and depend on agglomeration. Positive effects arise as concentration of economic activities brings up greater multiplicity of choice of private and collective goods and services, facilitates deeper specialisation and remunerating occupation, and makes it possible to enjoy a broader variety of urban amusements on one's leisure. Negative effects arise as concentration eventually causes physical and psychic congestion that depresses the enjoyment. Agglomeration diseconomies emerge in large-scale cities e.g. in the form of crowding, pollution and crime.

The indirect effects arise because spatial concentration of economic activities creates general economies on the level of the local marketplace. Performing the functions of the local economy involves a wide variety of physical contacts, travel and transport in the commodity and factor markets. All these actions are costly in terms of both money and time, and the rate of spatial concentration has considerable effect on the costs. The indirect effects are called business and social agglomeration economies and diseconomies.

Business firms derive several economies from the size of local factor and product markets. They may derive economies of scale in their own production decisions, localisation economies within industries sharing the same spatial marketplace, and urbanisation economies from the overall economic activity in one area. The social agglomeration economies and diseconomies are canalised through the local public sector. The local public sector can also utilise scale economies, cost sharing etc. in its service production. But, in an overly crowded city the advantages of size become to diminish. Agglomeration diseconomies are mainly due to the fact that the city is a physical construction that involves land use. A plain fact is that the more agglomeration there is, the scarcer land becomes. Constructing and maintaining an economically, socially and environmentally sustainable technical and social city infrastructure becomes the more costly the more crowded the area is Agglomeration diseconomies arise, when congestion starts to cause extra production costs due to rising land rents and delivery costs. The business and social agglomeration economies and diseconomies are indirect in that they, through the working of the local market, eventually end up to effects on the households' welfare.

Agglomeration economies and diseconomies are essentially externalities in nature in that the choices made by individual agents affect other agent's welfare in the city. Externalities are of two kinds, technological externalities that affect the utility function, and pecuniary externalities that operate through the price mechanism (Scitovsky, 1954; Fujita & Thisse, 2002).

Technological externality is the effect of an economic activity on the consumption set and the utility function of an individual. The effect concerns an agent other than the one exerting this economic activity. Technological externalities are anticipated in everyday life and they include both physical and psychic factors. Due to technical externalities the arguments in people's welfare functions differ in cities of different size. Technological externalities have a close match to the above definition of household agglomeration economies and diseconomies. However, business and social agglomeration economies and diseconomies such that affect the consumption sets and do not work through the price system may also be regarded to be externalities of this type.

Pecuniary externalities affect people's welfare through the price system. Those business and social agglomeration economies and diseconomies that have effects on wages, on the prices of private goods and housing, and on local taxes are this kind of externalities. It is assumed here that the local business sector and the local public sector are competitive and effective in their decision-making, finance and production. This assumption implies that the pecuniary business and social agglomeration economies and diseconomies revert to unbiased market price parameters faced by the households. What is not ruled out, however, is the effects of the migrants themselves to wages, prices and taxes in the city. Migrants with market power induce pecuniary effects, which they do not take into account in their own decisions. Therefore, the set of market prices may still be biased.

In the presence of externalities, individual decisions imply market power and introduce a certain collective, or club theoretic, element into the mechanism of the formation of a city (Tiebout, 1956; Buchanan, 1965). The key thing is that, from the club theoretic viewpoint, the external effects are internalised into the households' utility within a city (Cornes & Sandler, 1986; Scotchmer,1994). It must be noted, however, that there may exist also interregional externalities that are not internalised. Agglomeration boosts up the systemic functions of big cities as a diffusion source for innovation and development impulses within regions and down the national urban hierarchy.

3 The club theoretic nature of a city

In a market economy, individual people choose their sites of residence in order to maximise utility. A standard problem of constrained utility maximisation of an individual or a household reads

$$(1)\ Max\ U(q_x,q_y,l)\ \ s.t.\ \ w(1-l) + k = p_xq_x + p_yq_y,$$

where U is the utility function, q_x and q_y are the quantities of private and collective consumption, respectively, and l is leisure. These are the endogenous variables. In the budget constraint, w is the market rent (or wage) for work time $1-l$, where the total available time is normalised to unity, k is the non-labour income, and p_x and p_y are the market prices for private and collective goods, respectively. These are the exogenous market parameters. By (1), the individual is assumed to derive utility from the amounts of private and collective consumption and from leisure. The standard properties concerning the shape of the utility function are assumed. The budget constraint says that all labour and non-labour income on the income side is spent on private and collective consumption during the considered time period. As is well seen, the problem (1) cannot be solved without further restrictions, but the formulation suffices to illustrate the main idea of the analysis.

By the above argumentation, both the endogenous variables and the exogenous parameters in (1) are affected by externalities that are attached to agglomeration. Therefore, they can be expressed as functions of city size. Actually, it is the spatial proximity (or density) of the elements of the local market that explains agglomeration economies. However, it is simpler to use population as a measure of city size. This is reasonable, when the geographical area of the city is taken as given in the short-term.

Denote the population in the city by n, and write $q_x = q_x(n)$ and $q_y = q_y(n)$. This is to say that the sets of private and collective goods may differ between cities of different size. Moreover, $l = l(n)$ says that the amount and quality of leisure may vary between cities. This is explained by time saving due to specialisation in work, and by quality-of-life arguments. Implicitly it also includes the non-monetary emoluments of work. The connection of the endogenous variables to city size is explained by technical externalities. In the budget constraint, $w = w(n)$ refers to the wage effect of specialisation, and $p_x = p_x(n)$, $p_y = p_y(n)$ refer to the business and social effects of agglomeration on the market price of the private good and on the tax price of the collective good. Pecuniary externalities affect the formation of both prices. The non-labour income $k =$

k(n) may also be assumed to be locality-dependent as far as the required market transactions are spatial in nature.

The club-theoretic property of the city is to say that the agglomeration economies and diseconomies are transformed into locally experienced welfare. Now, since all the relevant arguments in the utility function and all the variables and parameters in the budget constraint depend on n, we can rewrite the problem (1) in a standard club theoretic form (Buchanan, 1965; Cornes & Sandler, 1986) as

(2) $Max\ W\ s.t.\ W = B(n) - C(n)$,

where W is individual welfare (or net benefit), B is the individual benefits experienced in the city and C is the individual net monetary costs of living in the city. The benefit side is refers to the utility function in (1), and the cost side refers to the budget constraint in (1). Technological externalities operate on the benefit side, and pecuniary externalities operate on the cost side.

By (2), the benefits and the costs and therefore also the welfare experienced in the city depend on city size. A conventional assumption is that agglomeration economies dominate at earlier stages of city growth, but that agglomeration diseconomies eventually start to dominate in overly crowded surroundings. On the benefit side it is quite obvious that the variety and quality of private and collective goods and leisure increase with city size, but that congestion and crowding in the streets and shops, pollution, damage, crime and other unpleasant characters eventually arise and cause the benefit to fall. On the cost side, city growth facilitates higher wages for less work, lowers prices of private goods and housing, cuts taxes and search costs and so on. But, as the geographical area of the city becomes overly congested, land rents, traffic and other transaction costs, infrastructure costs etc. start to rise up.

Expression (2) is written in terms of an individual household. In principle, the problem is now in solvable form since there is only one variable, namely n, to be solved. It cannot, however, be solved by any individual household alone, because individuals cannot choose such a variable as n, the size of a city. What they can do is to decide on whether or not to move in a city of certain size. As an optimisation problem, expression (2) can be solved only collectively - the city as a club is able to do that provided that it is provided with necessary instruments. Therefore, allowing for free mobility, expression (2) has two interpretations. On one hand, expression (2) describes how people make their location decisions by comparing all kinds of benefits and costs attached to different locations. And on the other hand, expression (2) describes the utility maximisation of an individual as a full member of the club that he resides in.

4 Illustration of the model

Assume that the geographical size of the city is set at its long-term optimum. Assume also that there are no inter-city externalities. This is to say people's perceptions about the benefits and costs associated with a city fully reflect social values. Provided that firms maximise profits and that the local public sector is efficient, agglomeration economies and diseconomies end up to be experienced by people in the city.

In order to fully understand the club theoretic nature of a city, start by examining the total welfare, which the city can produce for the whole of its citizen (Richardson, 1973; Fujita, 1989). Define $TW = nW = nB(n) - nC(n) = TB - TC$. The total benefit function $TB = nB(n)$ is assumed to be S-shaped. Taking the geographical area of the locality as fixed at long-term optimum, the function starts from the origin. On low population levels the increase in total benefits is exponential. This is because increasing population brings up a variety of agglomeration economies in goods and factor markets, including all kinds of monetary and non-monetary emoluments of urban life. However, the inevitable increase in congestion causes agglomeration diseconomies to arise. The rate of growth of the total benefits begins to fall at the point, where agglomeration diseconomies start to dominate. The total cost function $TC = nC(n)$ is assumed to be of inverted S-shape. This is reasonable because considerable agglomeration economies are due to appear, when n rises from very low levels. The cost function rises at a falling rate as agglomeration diseconomies eventually occur with increasing density. Agglomeration economies dominate until the fixed factor, the geography, becomes scarce enough so that agglomeration diseconomies start to dominate. At this point, the total cost function starts to rise at an increasing rate.

The curvatures of the *TB* and *TC* functions are familiar from standard cost theory. In the theory of the firm, the fixed cost determines in the short run the shape of the total cost function. Here the fixed factor is the geographical area of the city. In the long run, the fixed cost is also a choice variable in the firm's decision-making. So is also the geographical area in the case of cities - in the long run, both population and the geography are chosen optimally so as to reach to the highest possible welfare. At the long-term optimum agglomeration economies are fully utilised from the household, business and social perspectives. Figure 1 illustrates the above assumptions about the total benefit and total cost functions.

(Figure 1 here)

Figure 1 is drawn on the assumption that on very low population levels total costs of a city-like structure exceed total benefits from utilising it, but after a certain population level, namely n^o, benefits become higher than costs. The benefit side expanding in a diminishing rate, and the cost side eventually exploding, there will emerge another intersection point at *p*. When population in the city exceeds n^p, total costs will again exceed total benefits. The curve *TW*, drawn as a vertical sum of the *TB* and *TC* curves, presents the total welfare. Since the *TB* and *TC* functions include all kinds of benefits and costs perceived by the people, then *TW*, the net of *TB* and *TC* presents the attainable welfare that the locality can generate for its residents. Therefore, *TW* is the relevant measure for social welfare within the city. Without interregional externalities, the *TW*'s give the relevant measure also when considered from the viewpoint of the whole society consisting of a multiple of localities.

Total welfare in a city is an abstract concept, and it is by no means the driving force of any market-like mechanism. In their decision-making on residential location, individuals do not care about the total benefits and costs. In stead, individuals seek for their own good only, and base their migration decisions on calculations of personally perceived benefits and costs. Therefore, the concepts of total welfare presented in the upper panel of Figure 1 are not relevant for the individuals and thus not for the working of the market mechanism. The concepts must be elaborated to match to individual perceptions.

The simplest way to proceed is to assume that all individuals are alike and derive the average per capita benefit and cost, and the respective average welfare schedules from the *TB*, *TC* and *TW* curves. These concepts are presented in the lower panel of Figure 1 Note that the scale of the vertical axis is changed from the upper panel. The per capita (or average) benefit and cost functions read $B = TB/n$ and $C = TC/n$, respectively. By the properties of the *TB* and *TC* functions, *B* is inverted U shaped, and *C* is U shaped. The marginal benefit function, $MB = dTB/dn$, is a linear approximation of the curvature of the total benefit function TB. By intuition, *MB* is the effect that one immigrant has on the total benefits perceived by all members of the locality. By the properties of the total benefit function, the marginal benefit function *MB* is inverted U-shaped. Likewise, the U-shaped marginal cost function, $MC = dC/dn$, measures one immigrant's effect on total costs perceived by the whole population in the locality. Since *C* presents average numbers, adding a number smaller/higher than the average makes the average to fall/rise. Therefore, because *MC* describes the added numbers, it must strike through *C*, the average number, from below at the minimum point of *C*. By the same token, *MB* must strike through *B* from above at the maximum point of *B*.

The average concepts illustrate the individual experience, and match to the maximisation problem (2). The concepts deserve attention in two respects. First, they facilitate an accurate or even cardinal treatment of individual welfare (Ng, 2000), and second, they are straightforwardly operational in the preference-revealing mechanism of migration. The assumption of homogeneous people may sound critical, however. An alternative and less stringent interpretation might be that the average concepts are those of the representative household, and that they are eventually revealed by the systematic migration pattern. The decisions of the non-representative households thus belong to the purely stochastic element of migration.

Figure 1 is drawn on the assumption that the total benefit maximising population n^b is higher than the total cost minimising number n^a. This is based on the standard view that agglomeration diseconomies appear on the benefit side in more congested surroundings than they do on the cost side. The inverted U shaped *W* curve presents the welfare per person, drawn as a vertical sum of the *B* and *C* curves. The per capita welfare is positive between n^o and n^p, creating a motive for possible immigration from places with lower welfare. The *MW* curve presents marginal welfare and, as a presentation of marginal net benefit, it is a vertical sum of the *MB* and *MC* curves. Again by the properties of the benefit and cost functions, *MW* strikes through *W* from above at the maximum point of *W*.

The graphs in Figure 1 must be considered from ex post perspective. The functions present the benefits and costs that appear to individual people and to the city as a whole only at a certain population level. That is, the curves can be drawn only up till that population level that the private decision-making has driven it. Anticipations about their development forwards from that point are totally irrelevant as to private decisions. Particularly, a potential newcomer encounters the situation as given by the choices made before him, and he does not take into account the effect that he may have on the situation faced by the next immigrant after him.

5 Three solutions for city size

By Figure 1, the solution to the optimal city size seems to be straightforward. Given the above assumptions about the forms of benefit and cost functions, and that the net of them measures social welfare, the question about the most preferred size of the city becomes a standard club theoretic optimisation problem. From the point of view of economic efficiency population should be chosen so that total welfare is maximised. In Figure 1 total welfare reaches its maximum at the culmination point of the *TW* curve. The optimum condition is $dTW/dn = 0$, which yields $dTB/dn = dTC/dn$ in terms of the upper panel of Figure 1. Equivalently, in terms of the *B* and *C* functions in the lower panel of Figure 1, the condition may be stated as $MB = MC$, or $MW = 0$. This is to say that marginal benefits (the slope of *TB*) must equal marginal costs (the slope of *TC*) in the optimum. The above condition gives n^* as the social welfare maximising optimal population of the city. In the lower panel of Figure 1, the product of n^* and $W(n^*)$ equals the value $TW(n^*)$ in the upper panel. Welfare is maximal because, at n^*, agglomeration economies are optimally utilised in production and consumption of private and public goods with respect to agglomeration diseconomies.

The issue of city size is, however, more complicated. A most important question is whether the socially optimal solution depicted above is actually reached in a system based on individual utility maximisation. Individual decision-making with respect to the formation of a city can be divided to the purely individual choice of residential location, and to the collective choice on city size. As a matter of fact, purely individual migration and collective optimisation may, under certain circumstances, be efficient in a general equilibrium model with a multiple of cities (Atkinson & Stiglitz, 1980, p. 533-535), but the solutions are necessarily not efficient in a partial equilibrium model of one particular city.

In order to examine the outcomes of individual decision-making and their possible deviations from the socially optimal solution, take the perspective of the individual's maximisation problem expressed in (2) and concentrate on the lower panel of Figure 1. The lower panel is redrawn and somewhat reorganised in Figure 2.

(Figure 2 here)

Recall that the socially optimal solution occurs at *E*, the intersection of *MB* and *MC* curves. The intuition is that the effect that the last immigrant has on total benefits in the locality must be equal to the respective effect on total costs. Equivalently, the optimum can be found at the point, where the *MW* curve strikes through the horizontal axis. This is the point where the opposite marginal effects, brought up by the last immigrant, just cancel out. At the optimal population level n^*, the total welfare in the city, described here by the area beneath the *MB* curve and above the *MC* curve, is maximal. By definition, this area equals the area beneath the *MW* curve. An even simpler measure of the total welfare is the area given by the product of n^* and $W(n^*)$.

What would then be the market solution given by purely individual migration? Assume that all individuals are perfectly aware about the benefits and costs associated with the sites under consideration, and that they are able to calculate and compare all physic and psychic benefits and costs commensurably in monetary terms. Assuming also that migration is costless, it is the possible welfare gains that motor up spatial evolution. People choose whether or not to immigrate into a city of certain size. The city under inspection will attract newcomers as long as it can provide higher welfare than their original sites of residence. Supposing that the initial population in the city is large enough so that welfare is positive, and that there are other localities in the economy with zero welfare, the locality will attract newcomers up to *P*, the intersection of *B* and *C* curves. The market solution is a locality of n^p residents. The individual welfare, not only for the last immigrant, but also for all previous residents becomes zero. If the welfare in the delivering places is higher than zero, migration will continue until the welfare differentials are equalised, and the population solution may be anywhere leftwards from n^p. In any case, alas with continuing immigration, all citizens become to enjoy the benefits and costs of increased population, and the external effects of immigration are internalised into local welfare as is reflected by the curvatures of *B*, *C* and *W* in Figure 2.

The market solution described above is a stable market equilibrium generated by individual gain seeking. Yet, this equilibrium is not an efficient one. This is because expansion to n^p evokes total costs up till point *e* on the *MC* curve, and total benefits up till point *e'* on the *MB* curve. The welfare loss is described by the area *Eee'*, which is equal to n^* times $W(n^*)$. The market solution based on free migration is not socially optimal because individual choices are discrete by nature - they cannot choose n as a true choice variable, and therefore do not maximise expression (2). Individual decisions tend to drive the location too big. In order to attain to higher personal welfare, true optimisation on *n* should be warranted. The optimum solution at *E* in Figure 2 is clearly a planning solution - individuals with free entry to the locality, and with being worse off in their initial sites, would rather continue immigration towards point *P*. Public intervention would be necessary in order to secure efficiency.

Next, consider the case of collective choice, and assume a decentralised system, where cities are autonomous enough to decide on their spatial and populational characters. Taking the residents of the city as a collective - or a club - and assuming that they are equipped with preference revealing mechanisms such that facilitate the maximization of the collective's welfare, then n is a continuous choice variable in the model. Individuals alone cannot choose optimal n's, but together, as a collective they can. There are two viewpoints that can be taken into the optimal choice of the size of the city. Cornes & Sandler (1986, p. 164-167) call these the within-club approach and the total economy approach.

According to the within-club approach, the optimal size of the city is set so as to maximise the welfares of the individual residents of the city. In the simplified setting of Figure 2, this is equal to maximisation of the average welfare W in the city. Therefore, optimal population must be set to n', which is the population that maximises W. At the culmination point c, the optimum condition $dW/dn = 0$ satisfies the standard condition of a club theoretic population optimum, namely that the effect of the last newcomer on the benefits of each previous citizen must equal the respective effect on the costs side. To put it more formally, $dB/dn = dC/dn$. Implementation of the policy rule is to say that newcomers are allowed to immigrate up till the welfare of every individual citizen reaches its maximum, at which point the entrance to the city is then shut. The cities can exert this kind of power by various direct and indirect measures.

These include city planning and housing policies, limits in service provision etc. (Brueckner, 1982.)

The solution of the within-club approach, however, is not socially optimal in the partial equilibrium model of one city. By Figure 2 it is evident that the optimum given by the within-club policy rule does not correspond to the social optimum. More precisely, a within-club city remains sub-optimal in population, $n' < n^*$. Because total welfare is not maximised by the within-club rule, a welfare loss measured by the area $n'cn^*$ is caused. An equivalent measure for the welfare loss can be derived by calculating the geometric difference between the area given by the product of $n^{*\prime}$ and $W(n^*)$ and that given by the product of n' and $W(n')$.

The total economy approach says that the local government should optimise on the population of the city so as to maximise the welfare sum in the city. By this approach, it is not the average welfare W but the sum of individual welfares, namely $nW = TW$ in the present model, that is to be maximised. Therefore, the total economy approach reproduces the above efficient solution at E, and the corresponding choice of n^*. Note that since entrance to the city is closed not until after the n^*th immigrant, a fall of individual welfare is caused as compared to the above within-club case.

It is very interesting to see that the optimum of the within-club case implies a smaller city than the socially optimal solution of the total economy approach. The result is invariant of the assumption about the relative shapes of the TB and TC functions, and it only reflects the fact that total welfare is the product of population and individual welfare, $TW = nW$, and that the loss in individual welfare is more than compensated by the increasing number of residents up to n^*. This is to say that the collective as a whole benefits from extra members while the members themselves suffer.

As far as one city is concerned in a partial equilibrium context, the analysis shows that a pure market solution based on free mobility cannot end up to an efficient outcome. Furthermore, the within-club solution yielded by individual welfare maximisation remains also sub-optimal. Only a city that takes the total economy approach will end up to an efficient solution. Need for a social planner taking the total economy approach appears to be inevitable.

6 Conclusions

The paper presents a simple club theoretic model of the interdependence of city size and welfare. Agglomeration economies and diseconomies attached to city size internalise into people's benefits and costs of city life, and people respond by migration to possible differences in welfare provided by different cities.

The analysis shows that free migration of people cannot serve alone as a market mechanism such that would secure optimal formation of a city. Therefore, the city itself must act as a true market agent and choose its optimal population. In operating this function, there are two competing criteria of optimality, the within-club approach that aims to maximise the welfare of the individual members of the city, and the total economy approach that aims to maximise the welfare city as a whole. Astonishingly enough, these criteria produce different solutions: in terms of efficiency the total economy approach is superior to the within-club approach.

The results have interesting implications concerning the role of local governments. The results emphasize the fact that optimisation of population is an essential function of the local governments. This kind of a function has traditionally been banned at least in Finland. Furthermore, the results suggest that a particular city should take the total economy approach and maximise the total welfare experienced in the city even though it depresses the individually perceived welfare levels. It might, however, be argued that the within-club approach is more relevant in practice than the total economy approach. This is because of four reasons.

The first reason in favour of the within-club approach concerns the availability of data. Operating the total-economy viewpoint necessitates calculations of total welfare and/or marginal benefits and costs. These concepts are far more abstract than the average welfare and average benefits and costs, which are observed by the citizen and revealed by elections and polls. The average concepts are observable also from the systematic patterns of migration. It is easy to conclude that, since the operation of the planning-oriented total economy approach suffers from severe data problems whereas the within-club approach can be based simply on everyday experience, it is quite obvious that the latter would be chosen in practice.

The second evidence in favour of the within-club approach is given by the elegance of the modelling technique. The formation of a city as a club is motored by individual utility maximisation, and migration between cities derives from selfish economic reasons. It would be odd to assume that, after people have entered the city club, they would change their perspective from this selfish orientation towards a more altruistic total economy approach. This would imply that while they make their migration decisions so as to maximise on $W(n) = B(n) – C(n)$ according to the optimisation problem (2), they, at the same time, take part in the collective decision-making in their present site of residence by maximising a different utility function, namely $TW(n) = nW(n) = n(B(n)-C(n))$. Since the solutions to these functions clearly are different, this kind of a change in behaviour seems too inconsistent to be accepted without further reasoning. It is much more elegant to assume that people will continue to be selfish in the collective decision-making and apply the within-club approach.

Third, there is the question about the construction of the local government. In practice, local governance is usually constructed on the basis of representative democracy. Taken that people are selfish in their utility maximisation, they will vote for those decisions/electorates that promote their personal welfare. Electoral candidates who declare the total-economy approach would simply not be elected, or they would be replaced in the following election. The within-club approach is implicit in the system of representative democracy.

And fourth, as it seems that representative democracy may induce inefficient solutions, the results seem to imply that collective decision-making mechanisms should be constructed on social-planner–based maximisation of the total welfare of the collective irrespective of the individual welfares of the voters. This kind of a reform could quite plausibly be implemented, but it would not necessarily solve the problem if the households' utility maximisation pattern remained unchanged. The oppression of the voters' will would induce them to vote with their feet. If there were other locations with higher welfare than $W(n^*)$, the induced emigration would eventually nullify the social planner's optimisation. The major problem with a total economy type solution is that it is not necessarily stable.

To sum up, it seems plausible that cities either grow too big (because of purely individualistic migration) or remain too small (because of within-club type optimisation) with reference to the socially optimal size. However, the evidence of the model of an isolated city is less rigorous in practice than in theory. On one hand, it must be noted that the problems of reaching to the optimal size appears on the falling regime of the welfare curve. This is to say that the city is big enough so that the net effect of agglomeration economies and diseconomies is negative. In practice, most cities are on the rising part of their welfare curves, and they are not at all in the position to consider optimisation by welfare grounds. On the second hand, while the solutions are unambiguously inefficient in the partial equilibrium model of one city, both the purely migration-driven market solution and the within-club type collective solution can, under certain conditions, be socially efficient, when the general equilibrium of multiple cities is considered.

References:

Atkinson, A. & J. Stiglitz (1980): *Lectures on Public Economics.* McGraw-Hill, London.

Brueckner, J. (1982): A Test for Allocative Efficiency in the Local Public Sector. *Journal of Public Economics,* 19, p. 311-331.

Buchanan, J. (1965): An Economic Theory of Clubs. *Economica,* vol. 32.

Cornes, R. & T. Sandler (1986): *The Theory of Externalities, Public Goods, and Club Goods.* Cambridge University Press, Cambridge.

Cullis, J. & P. Jones (1998): *Public Finance and Public Choice.* 2nd ed. Oxford University Press Inc., New York.

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

Fujita, M. & J-F. Thisse (2002): *Economics of Agglomeration. Cities, Industrial Location, and regional Growth*. Cambridge University Press, Cambridge.

Krugman, P. (1991): *Geography and Trade.* MIT Press, Cambridge Ma.

Mills, E. & B. Hamilton (1984): *Urban Economics*. Harper Collins College Publishers, New York.

Ng, Y-K. (1973): The Economic Theory of Clubs: Pareto Optimality Conditions. *Economica*, vol. 4, no. 159, August.

Ng, Y-K. (2000): *A Case for Happiness, Cardinalism, and Interpersonal Comparabiliry.* In: H. Dixon (ed.), Controversies in Macroeconomics, Growth, Trade and Policy. Blackwell Publishers Ltd., Oxford.

Richardson, H. (1973): *The Economics of Urban Size.* Saxon House/Lexington Books, Westmead.

Rubinfeld, D. (1987): *The Economics of the Local Public Sector*. In: Handbook of Public Economics, A. Auerbach and M. Feldstein (eds.). Elsevier Science Publishers B.V. (North Holland)

Scitovsky, T. (1954): Two Concepts of External Economies. *Journal of Political Economy* 62, April.

Scotchmer, S. (1994): *Public Goods and the Invisible Hand*. In: Modern Public Finance, J. Quigley and E. Smolensky (eds.) Harvard University Press, Cambridge Ma.

Tiebout, C. (1956): A pure theory of local expenditures. *Journal of Political Economy* 64, October.

Figure 1: Benefits, costs and welfare generated in a city

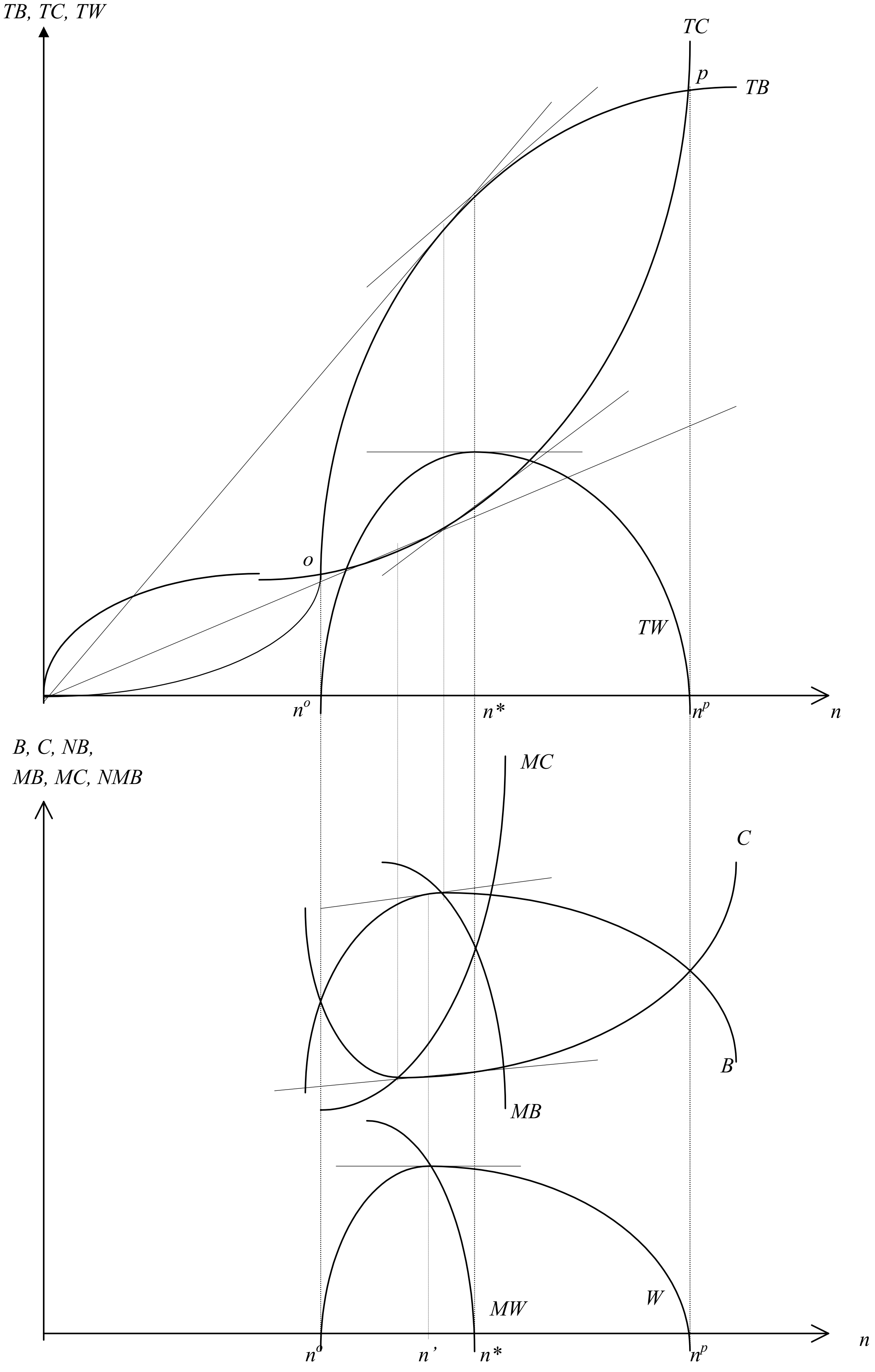

Figure 2: Three possible solutions for city size