ELSEVIER Regional Science and Urban Economics 27 (1997) 373–397

regional SCIENCE & urban ECONOMICS

# More on scale economies and cities[1]

Oded Hochman*

*Department of Economics and The Monaster Center for Economic Research, Ben Gurion University, Beer Sheva, Israel*

Received 7 September 1994; revised 14 February 1996

**Abstract**

In this paper we develop a model of a closed economy with many types of traded (between cities) products of which the production functions are subject to scale economies internal to the industry as a whole. Workers are assumed to live in Muth type residential rings and commute to the center of the city where the industry is located. This setup leads to the creation of many types of cities. To each type of city, residential cost functions and output cost functions are constructed. From these cost functions all types of demand functions are derived. It is shown here that the introduction of cities into the economy leads to an equivalence between the demand for inputs and the demand for consumption goods, an equivalence which does not exist in models of economies without cities. Thus there are conditional-compensated, unconditional-compensated and Marshalian demand curves for both inputs and consumption goods. The efficiency conditions for this economy are in the short run marginal cost pricing, and for an interior solution in the long run, average cost pricing is also required. This last condition is equivalent to the Henry George rule. Marginal cost pricing remains an efficiency requirement in the case of natural monopoly cities. Next we investigate two cases of scale economies internal to the individual firms and to the industry as a whole. The first case is a model of regular products in perfectly competitive markets and the other model deals with differentiated products and monopolistic competition. These cases have been previously investigated in the literature. Here we generalize and expand previous results and gain additional insights.

*Fax: 00972 7 6472310

[1]To David: friend, teacher and colleague (FTC).

I appreciate the useful suggestions from Masahisa Fujita and other participants of the David Pines conference on an earlier draft of this paper. Ehud Zuskovitch's advice is also appreciated. I am indebted to Richard Arnott, Marvin Kraus and especially to David Pines FTC, for detailed comments and most helpful suggestions.


*PII* S0166-0462(96)02154-0

*Keywords:* Cost functions; Demand functions; Differentiated products; Monopolistic competition; Natural monopoly

*JEL classification:* R13; D50

## 1. Introduction

Scale economies and transportation costs lead to the agglomeration of industrial and residential sites in close proximity in urban areas. In this paper we construct a model of an economy in which many consumption goods and intermediate products are produced, consumed or used as inputs together with land and labor, in many types of cities. The cities are created due to scale economies that are separately internal to each industry. The layout of each of the cities is assumed to be Muthian monocentric.

Initially, cost curves are constructed for the output of each city with only economy-wide prices and utility levels as arguments. From these cost curves we obtain demand functions for the different products. It is then shown that the introduction of cities into the economy leads to an equivalence between the demand for inputs and the demand for consumption goods, an equivalence which does not exist in models of economies without cities. Thus there are conditional-compensated, unconditional-compensated and Marshalian demand curves for both inputs and consumption goods. The necessary and sufficient conditions for an efficient allocation are then derived. In the short run, when the number of cities is fixed, these conditions are the well known marginal cost pricing rule. In the long run, when free entry of cities prevail, the average cost pricing rule is added as well. This last condition is equivalent to the well-known Henry George rule (see Arnott, 1979).

The introduction of cost functions for each city with only economy-wide arguments and no local variables, allows the derivation of economy-wide supply and demand functions for consumption and intermediate goods. This in turn facilitates a relatively easy detailed investigation of the economy as a whole, and at the same time yields general and robust results. On this basis a method of analysis is available that is familiar to economists and for which they have developed an intuitive insight.

Up to this point the scale economies has been described to be internal to the individual industry as a whole. However, it has not as yet been specified how the scale economies relate to the individual firms in the industry. From here on, we concentrate on analyzing two cases of scale economies internal also to individual firms (contrary to the case of scale economies external to the individual firm). These two cases have already been investigated in the urban economics literature. By applying the model developed here we enhance and generalize the previous results and gain some additional insights.

The scale economies investigated in the recent literature are mostly external to the individual firm but internal to the industry as a whole (Henderson, 1974, 1985; Arnott, 1979; Kanemoto, 1980; Fujita, 1988; Hochman, 1981, 1990 and others). These studies show that in the case of scale economies external to individual firms, each city includes only one type of the basic-export-good-industry with many firms, all having a common interest of attracting similar firms to the city.

The first case we look at was introduced by Kanemoto, 1980 who showed that when scale economies are internal to the individual firm, there will be only one firm producing a basic export good in each city. The market allocation will be efficient if this firm also acts as the developer of the city. Otherwise the market allocation will be inefficient and the city will be smaller than its optimal size.

In this paper we show that in addition a pecuniary externality exists which may permit the firm and the city to survive inefficiently in the short run. However, the local government, using a corrective subsidy, can induce the firm to behave efficiently while acting as a self-motivated profit-maximizer. Furthermore, it is shown that a net-city-profit-maximizing local government has both the means and the inclination to achieve this goal. In the long run, when the Henry George rule prevails, this behavior of the local government is a city survival condition.

The second case we discuss was introduced to the urban literature by Henderson and Abdel-Rahman, 1991. Their model also involves the production of export goods subject to scale economies internal to the individual firm (actually the cost function of a firm includes a component of fixed cost and another component of fixed marginal cost). Moreover, the export good is a differentiated product for which individuals prefer a wider variety. Consequently, each city-firm produces a different variety of the product and hence faces a downward sloping demand curve. Like Krugman, 1991, the authors adopt simple and specific functional forms of utility and production functions. They also choose a specific transportation cost function and a specific geographic form of the city. It is therefore difficult to understand the economic intuition behind their results, and to distinguish between results due to specific choice of functions and the more general results which do not depend on particular specifications.

In this paper we synthesize the main characteristics of the specific functions in Henderson and Abdel-Rahman and implement them in the general functions of the model developed here to obtain more detailed and robust results. We show that for a fixed number of cities/varieties the short run marginal cost pricing is still the efficient condition. However, pursuing the short run efficiency may prove to be counterproductive with respect to the long run. The long run efficiency condition is somewhat similar to Samuelson's rule concerning the efficient allocation of public goods. Increasing the number of cities/varieties raises the utility of all the population, but has a cost. The long run efficiency condition states that in the optimum the sum of the rates of substitution between the number of varieties and the numeraire product should be equal to the cost of an additional city. This last cost figure is equal to the difference between the total costs of producing the variety of a city and the value of its product evaluated at its marginal cost. The laissez faire market solution is a variant of classical monopolistic competition. Local government intervention in the form of subsidies financed by land taxes improve the allocation. However, the first best can be achieved only with a federal government intervention that uses a subsidy per unit of the differentiated product. This federal subsidy is financed by a federal lump sum head tax.

The plan of the paper is as follows: In Section 2 the cost functions are introduced. Then in Section 3 the economy wide Pareto efficiency conditions are derived. In Sections 4 and 5 first the Kanemoto model and then the Henderson and Abdel-Rahman model are investigated. Some concluding remarks are presented in Section 6.

## 2. Cost functions

In this section we develop the cost functions of a single city in the economy as a function of the basic export good output of the city. This cost also includes the expenditure of providing traded consumption goods to the local residents and intermediate products that are utilized as direct inputs in the production process. In Hochman, 1990, a more intuitive diagrammatic model involving cost functions was utilized to analyze external scale economies models.

We first consider a city, one out of many in the economy, with a basic industry of a nationally traded good that is subject to scale economies internal to the industry as a whole. At this stage only the scale economies of the whole industry are given and we do not specify the relation of individual firms to the scale economies, i.e. if they are internal or external to the individual firm. The subsequent analysis is therefore general and applies to all cases of scale economies internal to the industry as a whole.

We assume that all the cities in the economy are located on a bounded but featureless plane. For simplicity, we assume the production site is located at a zero size central business district (CBD), around which the residential ring (RR) is located. The residential ring extends from the origin 0, the location of the CBD, to the city boundary $\bar{x}$, which is determined endogenously. We utilize a Muth type model of an RR with a finite number of consumption goods. For convenience, we assume only one population group in the economy. The households consume land directly, the quantity of which is denoted by $h$, and a vector of consumption goods $z=(z_1, z_2,\ldots,z_I)$. The common utility function is $U(h, z)$, where $U(\cdot)$ is quasi-concave, and the common utility level throughout the economy is $u$. Movement of population and products between the cities in the economy is free and castles. Commuting costs from location x, where the household resides, to the CBD, where the household is employed, are given by $t(x)$ with $dt/dx > 0$, $d^2t/dx^2 \leq 0$. No land is required for transportation.

By designating the amount of land available at location x by $\theta(\mathrm{x})$, the city's population constraint is

$$\int_{0}^{\bar{\mathrm{x}}} (\theta(\mathrm{x})/\mathrm{h}(\mathrm{x}))\mathrm{dx} = N \tag{1}$$

where $N$ denotes the population size.

Let $z(R, p, u)$ be the vector of compensated demands for the consumption goods and $h(R, p, u)$ the compensated demand for housing where $R$ is land rent and $p$ is the vector of consumption prices. Then

$$\mathrm{e}(R, p, u, x) = R \cdot h(R, p, u) + p \cdot z(R, p, u) + t(\mathrm{x})$$

is the expenditure function of a household located at x, facing prices $p$ and $R(\mathrm{x})$, with utility level $u$. Let $T^{\mathrm{h}}(\mathrm{N}, p, R_{\mathrm{A}}, u)$ be the cost to a land developer of locating $N$ households in the city, given the price vector $p$, the utility level $u$ and the opportunity cost of urban residential land, $R_{\mathrm{A}}$. The opportunity cost of land is given by, $R_{\mathrm{A}} = R_{\mathrm{D}} + \sigma$, where $\sigma$ is the price of raw land and $R_{\mathrm{D}}$ is the cost of developing a unit of raw land into residential land. We shall discuss $R_{\mathrm{A}}$ and its components further in the next section, where the economy's land constraint enters into the model. Note that by development costs we do not mean construction costs, which are usually different at different locations and are assumed away in this model. What we are referring to resembles costs of residential infrastructure. Then,

$$T^{\mathrm{h}}(N, p, R_{A}, u) \overset{def}{=} \min_{R(\mathrm{x}), \bar{\mathrm{x}}} \int_{0}^{\bar{\mathrm{x}}} \frac{\theta(\mathrm{x})[p \cdot z(R(\mathrm{x}), p, u) + t(\mathrm{x}) + R_{A} \cdot h(R(\mathrm{x}), p, u)]}{h(R(\mathrm{x}), p, u)} \mathrm{dx}$$

$$\text{s.t.} \int_{0}^{\bar{\mathrm{x}}} \frac{\theta(\mathrm{x})}{h(R(\mathrm{x}), p, u)} \mathrm{dx} - N = 0 \tag{2}$$

where $R(\mathrm{x})$ now does not serve as a real market price, but rather a choice variable of the developer. This variable determines the physical quantities of the different goods, including housing, an individual will require at location x, given the other prices $p$ and $R_{\mathrm{A}}$, and the utility level $u$.

An interesting feature of the above function is

$$\frac{\partial T^{\mathrm{h}}}{\partial p_i} = \int_{0}^{\bar{\mathrm{x}}} \frac{\theta(\mathrm{x}), z_{\mathrm{i}}(N, p, R_{\mathrm{A}}, u, \mathrm{x})}{h(N, p, R_{\mathrm{A}}, u, \mathrm{x})} \mathrm{dx} \overset{\mathrm{def}}{=} Z_{\mathrm{i}}(N, p, R_{\mathrm{A}}, u) \tag{2a}$$

where $z_{\mathrm{i}}(N, p, R_{\mathrm{A}}, u, \mathrm{x})$, $h(N, p, R_{\mathrm{A}}, u, \mathrm{x})$, and $Z_{\mathrm{i}}(N, p, R_{\mathrm{A}}, u)$ are respectively, the compensated demands in a city of size $N$ for consumption good i and housing, both of an individual at location x, and consumption good i of the whole city. The above equation is obtained by differentiating $T^h$ in Eq. (2) with respect to $p_i$ and utilizing the envelope theorem.

By differentiating the Lagrangian constructed from Eq. (2) with respect to $N$ and then applying the envelope theorem, we obtain the marginal cost function,

$$\begin{aligned} \mathrm{MC}^h(N, p, R_A, u) &\overset{\text{def}}{=} \frac{\partial T^h}{\partial N} = \mathrm{e}(R(N, p, R_A, u, \mathrm{x}), p, u, \mathrm{x}) \\ &= E(N, p, R_A, u) > 0 \end{aligned} \tag{2b}$$

where $R(N, p, R_A, u, \mathrm{x})$ and $E(N, p, R_A, u)$ are the solutions of Eq. (2) for the rent and shadow price of the population constraint. The last equality in the above chain is the necessary condition resulting from differentiating the Lagrangian with respect to $R(\mathrm{x})$. The equation below is the second necessary condition for the minimization problem in Eq. (2), obtained by differentiating the Lagrangian with respect to $\bar{\mathrm{x}}$.

$$R(N, p, R_A, u, \bar{\mathrm{x}}) = R_A \tag{2c}$$

It should be noted that these conditions imply that the household expenditure in the optimum should be the same everywhere in the city. Indeed $E(N, p, R_A, u)$, the shadow price of the population constraint in Eq. (2), is this optimal household expenditure in a city of population size $N$, prices $p$, and utility level $u$. Observe that at the boundary of the city the consumption bundle is independent of $N$. Hence,

$$\frac{\partial^2 T^h}{\partial N^2} = t'(\bar{\mathrm{x}}) \frac{\partial \bar{\mathrm{x}}}{\partial N} = t'(\bar{\mathrm{x}}) \frac{h(\bar{\mathrm{x}})}{\theta(\bar{\mathrm{x}})} > 0 \tag{2d}$$

These properties of the cost function seem to be robust.

We now turn to the production of the basic good in the CBD. Let $T^p(Q, w, r)$ be the cost of producing a quantity $Q$ of the output of the basic industry, given the wage rate $w$, and the price vector, $r$, of inputs other than land or labor. From Shephard's lemma we know that

$$\partial T^p / \partial w = N^p(Q, w, r) > 0 \tag{3a}$$

where $N^p(Q, w, r)$ is the conditional demand for labor. We assume labor is a normal factor, i.e.

$$\frac{\partial N^p(Q, w, r)}{\partial Q} > 0 \tag{3b}$$

and has convex isoquants, i.e.

$$\frac{\partial N^p(Q, w, r)}{\partial w} < 0 \tag{3c}$$

The fact that economies of scale prevail is expressed in the following assumption

$$\partial(T^{\mathrm{p}}/Q)/dQ < 0 \tag{3d}$$

and derivation of the above expression yields

$$\mathrm{AC}^{\mathrm{p}} > \mathrm{MC}^{\mathrm{p}} \stackrel{\mathrm{def}}{=} \partial T^{\mathrm{p}}/\partial Q \tag{3e}$$

Now consider the following cost function

$$T^{o}(Q, r, p, R_A, \nu, u) = \min_{w}[T^{\mathrm{p}}(Q, w, r) - \{(w + \nu)N^{\mathrm{p}} - T^{\mathrm{h}}(N^{\mathrm{p}}, p, R_{\mathrm{A}}, u)\}] \tag{4}$$

where $N^{\mathrm{p}} = N^{\mathrm{p}}(Q, w, r)$ is the conditional demand for labor, and $\nu$ is the unearned income of an individual household in the economy. The underlying assumption is that each individual owns an equal share in each of the corporations in the economy and that these corporations own all the property (both land and firms) in the economy. $\nu$ is an equal share of the net profits of these corporations plus the federal lump sum transfers (positive or negative) if any.

The term in square brackets on the right hand side (r.h.s.) of Eq. (4) is the total cost to a developer of producing an output $Q$, including the cost of commuting, housing and feeding the city's laborers, provided he buys inputs and goods at market prices. In other words, we deduct the wages from the production costs and add the actual costs of maintaining $N^{\mathrm{p}}$ individuals in the city, from which we deduct the unearned income these individuals bring in. Note that the term in braces in Eq. (4) is equal to the differential residential rents (DRR) in the city. For each $w$ there exists a different cost, and for each $Q$ we choose the wage rate $w(Q, r, p, R_{\mathrm{A}}, \nu, u)$ which minimizes this cost. Thus $T^{\mathrm{o}}(Q, r, p, R_{\mathrm{A}}, \nu, u)$ is the cost to a developer of producing $Q$ units of the export good, with labor and production factors produced elsewhere as intermediate products. He buys these intermediate products in the market and in addition builds a city to provide his workers with developed land, commuting and consumption goods so that their utility level is $u$. He builds the city out of raw land which he buys for the price of $\sigma$ and develops for $R_{\mathrm{D}}$. For the workers' time, effort and unearned income, the developer provides them with consumption bundles in each location which provides them with utility level $u$. Each bundle in each location contains an optimal mix of developed land, commuting potential and consumption goods purchased at market prices, so that the developer minimizes his costs $T^{\mathrm{o}}$. Alternatively, as follows from Eq. (4), the developer may pay his employees the optimal wages and charge them $R(\mathrm{x})$ per unit of developed land, which is the maximum rents they are willing to pay for land at x. From here on we will refer to the cost function $T^{\mathrm{o}}(Q, r, p, R_{\mathrm{A}}, \nu, u)$ as the cost function of the city as a whole.

The first order condition of this minimization is

$$\partial T^{\mathrm{p}}/\partial w - N^{\mathrm{p}} + (\partial N^{\mathrm{p}}/\partial w)(\mathrm{MC}^{\mathrm{h}} - (w + \nu)) = 0$$

and upon substituting Eq. (3a) in the above equation, we obtain

$$\mathrm{MC}^{\mathrm{h}}(N^{\mathrm{p}}(Q, w, r), p, R_{\mathrm{A}}, u) - (w + \nu) = 0 \tag{5}$$

The second order condition for the above minimization is

$$(\partial N^{\mathrm{p}}/\partial w)[(\partial N^{\mathrm{p}}/\partial w)(\partial \mathrm{MC}^{\mathrm{h}}/\partial N) - 1] > 0$$

The above condition is satisfied if at the optimum, $\partial \mathrm{MC}^{\mathrm{h}}/\partial N$ is positive, which it is under our assumptions. If negative its multiplication with $\partial N^{\mathrm{P}}/\partial w$ should not exceed 1.

It should be noted that by substituting Eq. (2b) into Eq. (5) we obtain as a necessary condition for minimizing overall city production costs, the 'budget constraint' of a utility maximizing individual in a city with wage rate $w$ and unearned income $\nu$. From Eq. (5) we can solve for $w(Q, r, p, R_{\mathrm{A}}, \nu, u)$, and by substituting the result into $N^{\mathrm{p}}(Q, w, r)$, we obtain $N^{0}(Q, r, p, R_{\mathrm{A}}, \nu, u)$, the conditional demand for labor in the city as a function of the city output, economy-wide prices (including $u$, $R_{\mathrm{A}}$ and $\nu$) and no local prices[2].

By differentiating $T^{\mathrm{o}}(Q, r, p, R_{\mathrm{A}}, \nu, u)$ in Eq. (4) with respect to $Q$ and using the envelope theorem we obtain

$$\begin{aligned}
&\mathrm{MC}^{\mathrm{o}}(Q, r, p, R_{\mathrm{A}}, \nu, u) \overset{\mathrm{def}}{=} \frac{\partial}{\partial Q} T^{\mathrm{o}}(Q, r, p, R_{\mathrm{A}}, \nu, u) \\
&= \left.\frac{\partial T^{\mathrm{p}}(Q, w, r)}{\partial Q}\right|_{w(Q, r, p, R_{\mathrm{A}}, \nu, u)} + \left.\frac{\partial N^{\mathrm{p}}}{\partial Q}\right|_{w(Q, r, p, R_{\mathrm{A}}, \nu, u)} \cdot [w(Q, r, p, R_{\mathrm{A}}, \nu, u) \\
&\quad + \nu - \mathrm{MC}^{\mathrm{h}}(N(Q, w(Q, r, p, R_{\mathrm{A}}, \nu, u), r), p, R_{\mathrm{A}}, u)] \\
&= \mathrm{MC}^{\mathrm{p}}(Q, w(Q, r, p, R_{\mathrm{A}}, \nu, u), r)
\end{aligned} \tag{6}$$

Thus, the marginal cost of producing $Q$, for both the producer and the city as a whole is the same. However, this is not the case for the average cost

$$\begin{aligned}
\mathrm{AC}^{\mathrm{o}}(Q, r, p, R_{\mathrm{A}}, \nu, u) = \frac{1}{Q}[&T^{\mathrm{P}}(Q, w(Q, r, p, R_{\mathrm{A}}, \nu, u), r) \\
&- N^{\mathrm{p}}(Q, w(Q, r, p, R_{\mathrm{A}}, \nu, u), r)(w(Q, r, P, R_{\mathrm{A}}, \nu, u) \\
&+ \nu) + T^{\mathrm{h}}(N^{\mathrm{p}}(Q, w(Q, r, p, R_{\mathrm{A}}, \nu, u)p, R_{\mathrm{A}}, u)] \\
&< \frac{T^{\mathrm{p}}(Q, w(Q, r, p, R_{\mathrm{A}}, \nu, u), r)}{Q}
\end{aligned}$$

[2]By local prices we mean prices of factors which may have different values in different cities. In our model the only such prices are rents of developed land and wages. All other prices are economy-wide and have the same value everywhere.

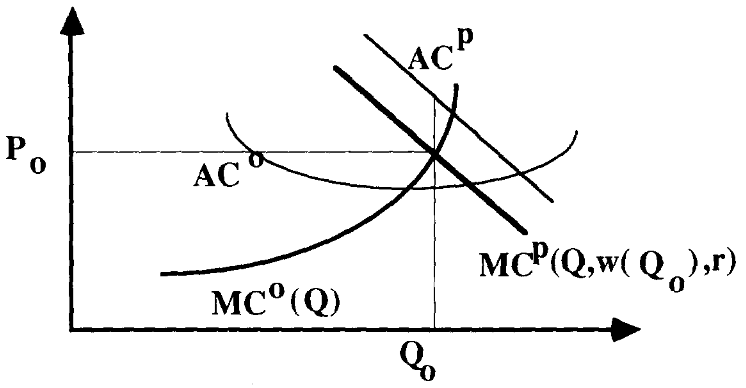


Fig. 1. Fixed wages and citywide cost curves.

$$= AC^P(Q, w(Q, r, p, R_A, \nu, u), r) \tag{7}$$

Fig. 1 depicts the relations between $AC^O$, $AC^P$, $MC^O$ and $MC^P$. It should be noted that while the $AC^P(Q, w, r)$ curves with fixed $w$ and $r$ are downward sloping by assumption, the associated $MC^P$ curves can also be upward sloping, but always below $AC^P$. $AC^O$ as a function of $Q$ can be U or L shaped or even fluctuating; $MC^O$ is shaped accordingly. Fig. 1 depicts a U-shaped $AC^O$ curve, which means that for low levels of output the economies of scale in the production of the basic industry dominates the diseconomies in the housing industry and vice versa for high levels of output. while there is only one $MC^O(Q)$ curve, there is a different $MC^P(Q, w, r)$ for each $Q$, since $w$ changes with $Q$ . The two curves intersect, as depicted, at the $Q$ where $w = w(Q)$.

The difference between the $AC^O$ and $AC^P$ curves for any given $Q$ is obtained from Eq. (7) and constitutes the differential residential rents (DRR) divided by the quantity of output, i.e.,

$$\begin{aligned} AC^P(Q, w(Q)) - AC^0(Q) &= [(w(Q) + \nu)N^P(Q) - T^h(N^P(Q))]/Q \\ &= DRR(Q)/Q \end{aligned} \tag{8}$$

Finally, let $L$ denote the land area of the city, then

$$\int_0^{\bar{x}} \theta(x)dx - L = 0 \tag{9}$$

## 3. Efficiency conditions in the economy

Consider now a closed economy located on a bounded featureless plane, with (for simplicity) one population-labor type, and many types of both intermediate and consumption, products. Each product is produced via a production function subject to scale economies internal to the individual industry. This, in turn, implies that only one industry will locate in each city. If two or more industries locate, neither benefits since the scale economies are internal and one industry does not benefit from the scale of the other. However, the fact that the laborers of all these industries are located in a single city raises the price of housing in the city. Consequently, this will increase the price of labor well above what the industry would have paid had it been the only basic industry in the city. Thus there are many cities in our economy, each producing a single export good and there may be many cities producing the same good. According to our assumptions such cities are identical. Let $J$ be the number of inputs other than labor produced and used in the economy (recall that in this model land is not a production factor). Then $I+J$ is the number of types of cities in the economy, where $I$ is the number of consumption goods. Let $m_j$, $N_j$, $L_j$, and $Q_j$ be the number of cities, the population, the land area, and output of a city of type j, respectively. We designate by $q_j$ the price of output j so that $q_j=p_j$ for $j=1,...,$I, $q_j=r_{j-I}$ for $j=I+1,...,\ J+I$, $q_{I+J+1}=R_A$, and $q_{I+J+2}=\nu$, where $R_A=R_D+\sigma$ is the price of developed land to (the developer of the city, $R_D$ is the cost of developing a unit of land, $\sigma$ is the price of a unit of raw land paid to the land holding corporations by the developer and $\nu$ is the identical share of a household in all the net profits and rents in the economy. In addition, let $z_i^j(Q_j,\ q,\ u)$, $j=1,...,I+J$; $i=1,...,I$ be the aggregate conditional compensated demand for good of type i by residents of city of type j, given the output level of the city, $Q_j$, the price vector, $(q)=(p,\ r,\ R_A,\ \nu)$, and the utility level, $u$, and let $Q_i^j(Q_j,\ q,\ u)$, $j=1,...,I+J$, $i=I+1,...,I+J$, be the aggregate conditional compensated demand of city j for input $i-I$, where $j=1,...,I$ are the types of cities producing the $I$ consumption goods and $j=I+1,...,I+J$ are the types of cities producing the $J$ factors of productions. we also add the index j to the relevant cost functions, e.g. $T_j^0$, $\mathrm{AC}_j^0$ and $\mathrm{MC}_j^0$ for $j=1,...,I+J$.

Let us now differentiate $T_j^0$ with respect to $q_{i,}$; $j,\ i=1,...,J+I$:

$$\frac{\partial(T_j^0)}{\partial q_i}=Z_i^j(Q_j,\ q,\ u);\ i,\ j=1,...,I+J \tag{10a}$$

where $z_i^j(Q_j,\ q,\ u)$ is the aggregate conditional compensated demand for consumption good i, $i=1,...,I$, in a city of type j, and the aggregate conditional compensated demand for production factor $i-I$, $i=I+1,...,I+J$, in city of type j.

$$\frac{\partial(T_j^0)}{\partial \nu}=-N_j,\ \text{and,}\ \frac{\partial(T_j^0)}{\partial R_A}=\frac{\partial(T_j^0)}{\partial \sigma}=L_j \tag{10b}$$

The above results imply the following proposition,

### 3.1. Proposition 1

The introduction of cities into the microeconomics theory leads to complete equivalence between demand functions for production factors and consumption goods, which is not the case in the general microeconomics theory (see Varian, 1992). Compensated-conditional demand functions of each city, depending on all product prices, the opportunity cost of residential land, the unearned income of an household as well as the city's output and the utility level of households, are derived by differentiating the city's cost function with respect to the factor price or the product price. Note that the price of labor is a negative price (an income ($-\nu$)), and the price of land is its opportunity cost $R_A$.

Since the economy is closed, the total quantity of a product consumed must equal the total produced. Thus,

$$m_i Q_i = \sum_{j}^{I+J} m_j Z_i^j(Q_j, q, u), \; i = 1,\ldots,I + J \tag{11a}$$

The above system of $I+J$ equations can be solved for the $I+J$ variables $Q_j$, $j=1,\ldots,I+J$, to obtain $Q_j(m, q, u)$, where $m$ is the vector of numbers of cities of each type. By substituting $Q_j(m, q, u)$ into $z_i^j(Q_j, q, u)$ we obtain $\bar{Z}_i^j(m, q, u)$ the aggregate unconditional compensated demand for product i in city j, whose arguments are the vector of numbers of cities of each type, and the prices and utility level, respectively. In addition, we substitute $Q_j(m, q, u)$ into $IN_j^0(Q_j, q, u)$, the conditional demand for labor in a city of type j, to get $N_j^u(m, q, u)$, the unconditional demand for labor in a city of type j. We now go on to substitute $N_j^u(m, q, u)$ into $E(N_j, p, R_A, u)$, the household expenditure in a city of size $N_j$ (see Eq. (2b)), to obtain, $E_j^u(m, q, u)$, the expenditure function of a household in a city of type j, given the economy wide prices, and the distribution of types of cities, i.e.,

$$E_j^u(m, q, u) = E(N_j^u(m, q, u), p, R_A, u) \tag{11b}$$

Let $I_j$ be the locally earned income required to achieve the utility level $u$. By solving for $u$ from $E_j^u(m, q, u) - \nu = I_j$, we obtain the indirect utility function of an individual in a city of type j, i.e., $V_j(m, q, I_j)$, where $\nu$ and $R_A$, are also components of $q$, The indirect utility functions fulfil,

$$E_j^u(m, q, V_j(m, q, I_j)) = I_j + \nu \tag{11c}$$

It should be noted that while $E(N_j, p, R_A, u)$ is the same for all cities because we assumed identical city layouts and the same transportation costs in all of them, $E_j^u$ and $V_j$ are each city-type specific. Thus the same $I_j$ for two different js will result in different values of utility, and the same $u$ in two different cities will result in two different expenditures. This way $I_j$, the locally earned income as defined in Eq. (11c), is in equilibrium equal to the wage rate, i.e. $I_j = w_j(m, q, u)$. By substituting $V_j$ for $u$, in the unconditional demand functions $\bar{Z}_i^j(m, q, u)$ and $N_j^u(m, q, u)$ we obtain the Marshalian demand curves for each city. Adding the Marshalian demand for a factor or a product in all the cities yields the economy-wide Marshalian demand curves as a function of the distributions of cities, the income distribution in each of the cities and the economy-wide price vector. The above analysis yields the following proposition:

### 3.2. Proposition 2

The equivalence between the conditional-compensated demand functions for inputs and the conditional-compensated demand functions for consumption goods established in proposition 1 extends to all kinds of demand functions for both inputs and consumption goods. There are both, compensated and Marshalian demand functions for inputs of a city, an industry, and the whole economy, both conditional on output and unconditional. Exactly the same types of demand functions exist for consumption goods as well, and the way to obtain any of these types of demand functions for either inputs or consumption goods, is exactly the same.

We can now formulate the balance of payments constraint of the economy,

$$\sum_{j}^{j+I} q_j m_j Q_j(m, q, u) + \sigma\bar{L} - \nu\bar{N} - \sum_{j}^{J+I} m_j T_j^0(Q_j(m, q, u), q, u) = 0 \tag{12}$$

where $\bar{L}$ and $\bar{N}$ are, respectively, the total land available in the economy and the total given population. As a reminder we point out again that in this model land and population have prices. Each city developer buys raw land at the price $\sigma$ from the land corporations, develops it at the cost of $R_D$ per unit of land, and sells to households, for their unearned income, $\nu$, and their labor, a location in the city bundled together with developed land and consumption goods which together provides the household with the economy wide utility level $u$. The net proceeds from these transactions are added to the balance of payments constraint.

The efficient allocation of resources in the economy can now be obtained by maximizing u under the balance of payments constraint (Eq. (12)). Let the short run be the length of time in which the number of cities is fixed. Then the control variables are the prices $q_k$, $k=1,\ldots,I+J+2$, (i.e. $(q)=(p, r, R_A, \nu)$) and $u$, except that in $R_A$, the only variable part is $\sigma$.

The necessary conditions obtained upon differentiation with respect to the prices $q$ are an expression of the marginal cost pricing rule, i.e.

$$\mathrm{MC}_j^0(q, u) = q_j,\ j = 1,\ldots,I + J \tag{13}$$

Differentiation with respect to $\sigma$ yields the land constraint,[3] i.e.,

$$\bar{L} - \sum_{j}^{I+J} m_j L_j = 0 \tag{14}$$

and differentiation with respect to $\nu$ yields the population constraint, i.e.,

$$\bar{N} - \sum_{j}^{I+J} m_j N_j = 0 \tag{15}$$

Eqs. (13)–(15) can now be solved for the price vector ($p$, $r$, $R_A$, $\nu$). Since $MC_j^0(q, u)$ is non-negative, it follows from Eq. (13) that $q_j$ is non-negative as well. If we constrain $\sigma$ to be non-negative, then the Kuhn-Tucker theorem implies that the left hand side of Eq. (14) can be positive when $\sigma$ equal zero, and when equality holds in Eq. (14) $\sigma$ can be positive. In this case, $R_A$ exceeds $R_D$ the per-unit land development cost. In general $R_A$ is bounded from below by $R_D$, and is therefore always positive.[4] The non-earned income term $\nu$, however, can take any sign. We discuss below some of the conditions leading to different signs of $\nu$.

In the long run, the number of cities is not restricted and can be determined optimally. First consider the case in which the optimal value of all $m_j$ is sufficiently large to justify disregarding the divisibility problem. The necessary condition obtained by differentiating the appropriate Lagrangian with respect to $m_j$ implies average cost pricing, i.e.,

$$q_j = AC_j^0(Q_j) \textit{ for all } j = 1,\ldots,I + J \tag{16}$$

A sufficient condition for an interior solution is that $MC_j^0(Q_j)$ is increasing with $Q_j$ at the optimum. It should be noted that Eqs. (13) and (16) jointly imply that the long run optimum is attained when the product price equals the minimum average cost of production, i.e. min $AC_j^0(Q_j) = q_j$. Furthermore, if we let $\pi_j$ be the net profit function of city j, Eq. (16) implies

$$\pi_j = q_j Q_j - T_j^0(Q_j) = 0 \tag{17}$$

Eq. (17) is an expression of the well known Henry George rule (see Arnott, 1979) in terms of this model.

Upon substituting the above, together with the necessary conditions in Eq. (12), we obtain for this case,

[3] By specifying the land constraint in this way we avoid the problem of the geographical configurations of the cities and whether or not they fit together. It is assumed that the cities can always utilize all the available land efficiently.

[4] This discussion of $R_A$ and its components was inspired by a heated argument with David Pines FTC for which I am indebted to him.

$$\nu = \sigma \bar{L} / \bar{N} \geq 0, \tag{18}$$

namely, the unearned income is, in the case of an interior solution, non-negative and positive when the land constraint is binding.

If one of the export goods industries operates under constant returns to scale technology (e.g. agriculture), i.e. $AC^p = MC^p = c_0$, then the optimal city size of such an industry is one household, there are no commuting costs and the land rent is exactly $R_A$. This may explain why $R_A$ is often referred to as agricultural land rent. Observe that in the optimal solution all land in the economy would be occupied and $R_A = \sigma > 0$ when $R_D = 0$, one or more of the industries are subject to constant returns to scale, and the marginal utility of housing is always positive. Also, when $R_D = 0$, unoccupied raw land may exist only if there are no industries with constant or decreasing returns to scale, and then $R_A = R_D = \sigma = 0$. when $R_D > 0$ the optimal solution can still include undeveloped empty land in the economy even if one or more of the industries is of constant or even decreasing returns to scale. In this case $R_A = R_D > 0$, and in general, when all the land is developed $R_A \geq R_D > 0$.

In the short run, when the number of cities of each type is restricted, we are in a second best situation. Still if the condition for an interior solution holds and there is more than one city of each type, then marginal cost is higher than average cost in all the cities. Hence, $\pi_j$ in Eq. (17) is positive and so is $\nu$, that is:

$$\nu = \frac{\left( \sigma \bar{L} + \sum_{j}^{I+J} m_j \pi_j \right)}{\bar{N}} > 0 \tag{19}$$

If, however, in some types of cities the sufficient condition for an interior solution does not hold, then only one city per type exists in the optimum and $\pi_j$ is negative. Eq. (19) still holds except for the direction of the inequality which may now be reversed. This is true in particular for the case where there is only one city in each of the industries, and the land constraint is not binding. In this case $\nu$ in Eq. (19) is bound to be negative.

The discussion above implies that in the case of an interior solution, competition between profit maximizing city developers will yield the optimal solution. Moreover, the developer does not have to plan the allocation of land and consumption goods for his employees. Instead he can, as has been shown in the previous section, pay his employees the wage rate $w_j(Q_j, q, u)$ and rent the developed land to the highest bidder for $R(x)$.

In the next two sections we investigate two examples of scale economies internal to individual firms and producers owning only their plants (they own no residential land). The cities are controlled by local municipal governments which maximize net city profits.

## 4. Scale economies internal to the individual firm

When scale economies are internal to the individual firm as well as the industry as a whole, efficiency implies that only one firm of the industry should exist in a city. Obviously, if a given output is divided between two firms its production costs are higher than they would have been were it produced by a single producer. Each of the firms produces only part of the total output at a higher average cost than it would have taken to produce the whole quantity. Even if the economies of scale were exhausted by the individual firm, it still would not be worth while for another firm to enter the city because labor wages would be higher for a single city. When more than one firm produces in the same city, average labor cost is higher for both firms.

Also in competitive markets only one export good producer will locate in each city. A potential entrant will always prefer developing a city of his own to moving to an occupied city. In an existing city which already serves other firms, he will have to pay higher wages to his workers, without having any benefit from locating with other firms.

The average cost to the producer in the city when his output is $Q_0$ is $\mathrm{AC}^P(Q, w, r)|^{Q_0}_{w(Q_0,\, q,\, u)}$. The reason for the difference between $\mathrm{AC}^0(Q_0)$ and $\mathrm{AC}^\mathrm{P}(Q_0)$ is that the wages the producer pays are higher than the expenses necessary to maintain the economy-wide utility level in the city. Indeed the marginal cost of labor is the same in both cases but not the average. As shown in the previous section, the difference between the two costs is the DRR.

Given that there is a single export good producer in a city, he will necessarily notice that the wage rate increases when he increases production and hires additional workers. Were he not aware of this, he would have tried to increase his output indefinitely to reduce his marginal cost and increase his profits. But once output is raised, so are wages and consequently marginal and average costs. Therefore, if the producer does not anticipate wage increases, he learns it quickly through trial and error. That is, he will consider his average cost to be $\mathrm{AC}^\mathrm{e}(Q, q, u)$, where,

$$\mathrm{AC}^e(Q, q, u) = \mathrm{AC}^\mathrm{P}(Q, w(Q, q, u), r) \tag{20}$$

for given $q$ and $u$ (from now on we will omit the arguments $q$ and $u$ from the list of variables except for the sake of clarity).

$$\frac{\partial \mathrm{AC}^\mathrm{e}}{\partial Q} = \frac{\partial \mathrm{AC}^\mathrm{P}}{\partial Q}\bigg|^{Q}_{w(Q)} + \frac{\partial \mathrm{AC}^\mathrm{P}}{\partial w}\bigg|^{Q}_{w(Q)} \cdot \frac{\partial w(Q)}{\partial Q} \tag{21}$$

It should be noted that the relation between $\mathrm{AC}^\mathrm{e}$ and $\mathrm{AC}^\mathrm{P}$ is similar to that between $\mathrm{MC}^\mathrm{P}$ and $\mathrm{MC}^0$. In both cases we obtain one curve by substituting the optimal wage rate as a function of output.

Since we assumed $N$ to be a normal input everywhere, then $\partial w/\partial Q > 0$

everywhere. Since $\partial \mathrm{AC}^{\mathrm{P}}/\partial w > 0$, $\mathrm{AC}^{\mathrm{e}}$ can be upward sloping, and is definitely upward sloping in the range where $\mathrm{AC}^{\mathrm{O}}$ is upward sloping (see Fig. 2). The marginal cost function facing the producer is now

$$\mathrm{MC}^{\mathrm{e}}(Q) = \frac{\partial(Q \cdot \mathrm{AC}^{\mathrm{e}}(Q))}{\partial Q} \tag{22}$$

It should be noticed that in a company town where the producer of the export good also owns the land in the entire town, the costs of production are indeed given by $\mathrm{AC}^{\mathrm{e}}$, but the producer has an additional income from residential land rents. Therefore the relevant cost curves for him are $\mathrm{AC}^{0}$ and $\mathrm{MC}^{0}$, and he achieves efficiency without any outside intervention. In the case of a producer who does not own land, $\mathrm{AC}^{\mathrm{e}}$ is his relevant average cost curve since the land rents go to land owners and are therefore external to him. However, he bears the additional costs of wage increase when the city expands and land rents increase.

As depicted in Fig. 2, $\mathrm{MC}^{\mathrm{e}}$ is above $\mathrm{AC}^{\mathrm{e}}$ where $\mathrm{AC}^{\mathrm{e}}$ is upward sloping. In such a case, an equilibrium solution can exist where the market price is equal to $\mathrm{MC}^{\mathrm{e}}$ (for example, see Fig. 2 where $p = p_1$ at $Q_{\mathrm{O}}$). The equilibrium solution may be zero production (for example, when $p = p_0$ provided $\mathrm{MC}^{\mathrm{e}} > p_0$ for all $Q$). In any case, the equilibrium result is never an efficient one and city size is too small since $\mathrm{MC}^{\mathrm{e}}(Q) > \mathrm{MC}^{\mathrm{o}}(Q)$ for all $Q$.

To achieve efficiency the local government has to internalize this externality by subsidizing the basic industry which in this case consists of a single firm. The average subsidy per unit output when the market price is $p_{\mathrm{O}}$ is $p_1 - p_0$ where $p_1$ is the value of $\mathrm{MC}^{\mathrm{e}}$ at $Q_{\mathrm{O}}$, and $Q_{\mathrm{O}}$ is the quantity of output where $\mathrm{MC}^{\mathrm{O}} = p_{\mathrm{O}}$ (see Fig. 2).

The external effect here is of the pecuniary type, resulting from the fact that the producer takes into account the change in the wage rate due to his activity. Since the externality is due to changing labor prices, the subsidy should be given to the firm per worker employed. Otherwise the allocation of production factors will be

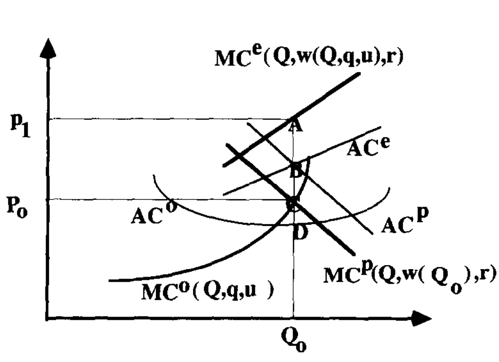


Fig. 2. Fixed wages, city and industry cost curves.

distorted —not enough labor will be employed, and consequently the first best will not be attained.

The subsidy required can be financed fully by the local government from taxes on land rents. The total subsidy is $Q_0(p_1 - p_0)$. Let

$$\begin{aligned} \mathrm{MC}^{\mathrm{e}}(Q_0) &= \mathrm{AC}^{\mathrm{e}}(Q_0) + Q_0 \frac{\partial \mathrm{AC}^{\mathrm{e}}}{\partial Q}\bigg|_{Q_0} \\ &= \mathrm{AC}^{\mathrm{P}}(Q_0, w(Q_0)) + Q_0\left(\frac{\partial \mathrm{AC}^{\mathrm{P}}}{\partial Q}\bigg|_{w(Q_0)}^{Q_0} + \frac{\partial \mathrm{AC}^{\mathrm{P}}}{\partial w}\bigg|_{w(Q_0)}^{Q_0} \cdot \frac{\partial w}{\partial Q}\right) \\ &= \mathrm{MC}^{\mathrm{P}}\bigg|_{w(Q_0)}^{Q_0} + N(Q_0, w(Q_0)) \cdot \frac{\partial w}{\partial Q}\bigg|_{Q_0} = \left[\mathrm{MC}^{0} + N\frac{\partial w}{\partial Q}\right]_{Q_0} \end{aligned}$$

in which the equality before the last is obtained after substituting Eq. (3a) in the equation.

$N(dw/dQ)|_{Q_0}$ is indeed the average (per unit output) pecuniary externality and, in turn, the above chain proves that it should be equal to the average subsidy. The optimal subsidy per laborer is consequently $Q(dw/dQ)|_{Q_0}$. This subsidy (per unit output) is equal to segment CA in Fig. 2. The average rent in the city is depicted by the segment AD so that the rents are sufficient to finance the subsidy. The average profits of the producer in the CBD are equal to AB, and they are never enough to cover the whole subsidy. Consequently, residential rents must be used as well. It should be noted that the optimum depicted in Fig. 2 at $Q_0$ is a short run equilibrium with positive net profits represented by $\pi/Q$ being equal to segment CD in Fig. 2. In fact, if the city is operating at the point of min $\mathrm{AC}^{\mathrm{O}}$, the long run first best with free entry of cities, then all the residential rents are needed to cover the subsidy. This is the well known Henry George rule (e.g. see Arnott, 1979).

We also note in passing that in this case a city government should be able to tax all the land on which households employed in the city reside. Otherwise, the city will not be able to obtain funds efficiently to finance its activity (see Hochman, 1990, for details).

Consider the case of a corner solution that occurs when the export good demand curve of the whole economy intersects the $\mathrm{AC}^{\mathrm{o}}$ curve of a single city where it is downward sloping, and there are no imports of this good to the economy. Then we are dealing with a natural monopoly town. In this case, the local government does not have the means or the inclination to achieve efficiency. In the first place, at the point of efficiency, average cost, $\mathrm{AC}^{\mathrm{o}}$, is above marginal cost, $\mathrm{MC}^{\mathrm{o}}$, so the required subsidy is higher than the city's aggregate land rents. Second, since net city profits are maximized where the marginal revenue intersects the $\mathrm{MC}^{\mathrm{o}}$ curve and not at the efficiency point (the intersection of the demand curve with $\mathrm{MC}^{\mathrm{o}}$), the local government has no incentive to reach efficiency. Federal government intervention in the form of a corrective federal subsidy is needed to achieve the desired solution in this case. This subsidy has to be financed by a federal head tax.

The unearned income $\nu$ of the previous section consists in this case of a share of land rents and profits, if any, minus this federal tax. Thus in this case $\nu$ can be negative. A detailed discussion of such a case follows in the next section.

## 5. Variety and monopolistic competition

We now re investigate the model developed by Henderson and Abdel-Rahman (subsequently referred to as HAR) and see if by using the tools developed here we can gain some more insight. In what follows we initially identify the cost functions of the HAR model and their special traits. Then we incorporate the essence of these traits into our general model.

First consider the simpler case of HAR where the only products are different varieties of a single differentiated product. The utility function is then $U(z_1,...,z_m)=(\Sigma_{i=1}^{m} z_i^{\rho})^{1/\rho}$, where $z_i$ is the $i$th variety of a differentiated product, with $i=1,...,m$ and $m$, the number of varieties, is a choice variable of the economy as a whole and an upper bound for the number of varieties chosen by an individual household. It is further assumed that the parameter $\rho$ is a constant and $0<\rho<1$. The production function of each variety is assumed to be the same, utilizing only labor with a fixed marginal productivity and a fixed labor input component. Thus the production function displays increasing returns to scale internal to the individual firm at all levels of output. The production relations are given by $N_i^p=f+cQ_i$, where $N_i^p$ is the labor involved in the production of $Q_i$ units of variety i, where both $f$, the fixed labor input, and $1/c$, the marginal productivity of labor, are the same for all varieties. Symmetry implies that all $Q_i$ will have the same price and be produced in equal quantities. If we choose the common price of the varieties of the differentiated product to be one, then $1/c$ is the value of the marginal product of labor in all cities of all sizes. Let $z$, $z=\Sigma_{i=1}^{m} z_i$, be the total quantity of the differentiated product consumed by the household. Then the utility function can be written as $u=U(z, m)=m^{(1/\rho)-1}z$, and since the exponent of $m$ is positive, and $m$, unlike $z$, costs nothing to the individual, households will consume all there is of $m$ in the market. In short, $m$ behaves like a public good. Then $z(m, u)$ is the compensated demand of an individual, given $m$ and $u$.

It is assumed that each household consumes one unit of land independent of the location of the household and that this unit of land does not enter the utility function. Commuting costs, $t$(x), are specified in terms of labor units.

The arguments of the previous section apply to this case as well and each variety, under the above assumptions, should be produced in a separate city. Thus $m$ is also the number of cities in this economy. Let TCC($N$) stand for total commuting costs in a city with a population of $N$; then

$$\mathrm{TCC}(N) = \frac{1}{c}\int_0^{\bar{x}} \theta(\mathrm{x})t(\mathrm{x})\mathrm{dx}$$

$$\text{s.t.} \int_0^{\bar{x}} \theta(x)dx - N = 0 \tag{23}$$

where the constraint in Eq. (23) is both the city population constraint and the land constraint. However, the total land in this model is not restricted and its alternative price is therefore zero. By differentiation we obtain

$$\frac{\partial \text{TCC}}{\partial N} = \frac{t(\bar{x})}{c} > 0, \text{ and } \frac{\partial^2 \text{TCC}}{\partial N^2} = \frac{t'(\bar{x})}{c\theta(\bar{x})} > 0$$

We can now calculate the total cost of housing $N$ households in a city, $T^{h}(N, m, u)$, with utility level, $u$, and variety $m$,[5]

$$T^{h}(N, m, u) = N \cdot (z(m, u) - \nu) + \text{TCC}(N) \tag{24}$$

By solving the following equation for $N(Q)$,

$$N - c \cdot \text{TCC}(N) = f + cQ$$

and substituting the result into Eq. (24), we obtain $T^{0}(Q, m, u)$.

The model depicted so far is somewhat degenerate since $m$ is the only choice variable. Once $m$ is given there is no choice — the population size is $N = \bar{N}/m$, where $\bar{N}$ is the total population in the economy.

The total output of a city $Q$ is also uniquely determined: $Q = [\bar{N}/m - c \cdot \text{TCC}(\bar{N}/m) - f]/c$, and the household consumption is $z(m) = [\bar{N} - m \cdot c \cdot \text{TCC}(\bar{N}/m) - m \cdot f]/c \cdot \bar{N}$. It is obvious that marginal cost pricing for all $m$ is maintained in a trivial way in the short run since[6] $MC^{P} = w/(\partial Q/\partial N) = (1/c)/(1/c) = 1$. The long run optimum is attained by solving the following maximization problem:

$$\max_{m} U(z(m), m) = U(\{[\bar{N} - m \cdot c \cdot \text{TCC}(\bar{N}/m) - m \cdot f]/c \cdot \bar{N}\}, m)$$

The necessary condition is then,

$$\bar{N} \frac{U_m}{U_z} = \frac{c \cdot \text{TCC} + f}{c} - \frac{\bar{N}}{m} \frac{t(\bar{x})}{c} \tag{25}$$

The above condition is not the same as the average cost pricing found in the previous section. It is rather a version of Samuelson's rule regarding public goods: on the l.h.s of Eq. (25) are given the marginal benefits for the whole population from an additional variety, and on the r.h.s are given the marginal costs of an

[5]The assumption is that the price of the differentiated product is 1. In HAR labor is the numeraire.

[6]It should be noted that in the simple HAR model, for any given $m$ there is only one feasible solution which is obviously both the optimal and competitive solution. Therefore, unlike the more general models in the previous sections, no intervention of a local government is required to achieve efficiency.

additional city. The first term in the r.h.s is the output lost due to fixed commuting and production costs, and the second term is the reduction in commuting costs due to shortened travel distances. Observe that when the marginal utility of $m$ is zero, i.e. there is no benefit from an increased variety of products, the l.h.s of Eq. (25) disappears and the optimality condition becomes an equality between the two terms on the r.h.s of Eq. (25). Under these circumstances, one would expect equality between marginal and average cost of production to be the condition, and indeed if we add $\bar{N}z/m$ to both terms and divide them by $Q$, we do obtain $\mathrm{AC}^0 = \mathrm{MC}^0$ as expected. This argument also implies that when the marginal utility of $m$ is positive, $\mathrm{AC}^0 > \mathrm{MC}^0$. Consequently, in the case of a differentiated product, the optimum is attained where $\mathrm{AC}^0$ is downward sloping. The equilibrium solution to the simple HAR model is attained when each city equates its marginal revenue to its marginal costs, and then equates total city profits to zero. HAR show that in this simple model the optimal allocation is also the equilibrium one.

The main problem of the specific example above is whether its results are robust. So instead of exploring this model further, we consider a more general case which should yield more robust results, and provide additional and more intuitive results as well.

Consider the model investigated in the previous section, but assume that $z_1$ is a differentiated product; that is, $m_1$, the number of cities producing the differentiated product which is also the number of varieties, enters the utility function. The short run efficient solution in which the number of cities and varieties are fixed is the same as in the previous section, namely the necessary conditions are all expressions of the marginal cost pricing rule. In the long run, for cities of type i, $i > 1$, average cost pricing is again the efficiency rule. The only difference is in the differentiation with respect to $m_1$. It should be noted that since $m_1$ appears in the utility function it will also appear with non-zero derivatives in all the compensated demand functions for the consumption goods. Consequently, $m_1$ will also be an argument of $T^{\mathrm{h}}$ and $T^0$. By differentiation of Eqs. (2) and (4) with respect to $m_1$ we obtain

$$\partial T^0 / \partial m_1 = \partial T^{\mathrm{h}} / \partial m_1 = -\int_0^{\bar{x}} \frac{\theta(\mathrm{x}) \cdot U_{m_1}(\mathrm{x})}{h(\mathrm{x}) \cdot U_{z_{\mathrm{n}}}(\mathrm{x})} \mathrm{dx} \tag{26}$$

where $U_{z_n}$ is the marginal utility of $z_n$, the numeraire (a consumption good with price $p_n = 1$, $n \in (1,...,I)$).

We can now perform the differentiation with respect to $m_1$ of Eq. (12) in which $Q_j$ is substituted by the compensated demand functions from Eq. (11a). The necessary condition is

$$T_1^0(q, m_1, u) - Q_1 \cdot \mathrm{MC}_1^0(q, m_1, u) = \sum_{i=1}^{I+J} m_i \int_0^{\bar{x}_i} \frac{\theta_i(x) \cdot U_{m_i}(\mathrm{i}, \mathrm{x})}{h_i(\mathrm{x}) \cdot U_{z_{\mathrm{n}}}(\mathrm{i}, \mathrm{x})} \mathrm{dx} \tag{27}$$

Once more, this time on the r.h.s of Eq. (27), we have the benefits of an additional city, and on the l.h.s the costs. with marginal cost pricing, the l.h.s of Eq. (27) is the city's losses, which is the cost society pays for an additional city. Since the r.h.s is positive, so is the l.h.s, which implies that at the optimum $AC^0 > MC^0$. This in turn implies that we are in the range where the city as a whole still displays scale economies.

Let us now turn to the market allocation. For cities of type i, $i>1$, and assuming an internal solution, i.e. $m_i>1$ for all $i>1$, the results of the previous section apply, namely only local government intervention is required in order to achieve the optimality conditions in both the short and long runs. Furthermore, each net-land-rent-maximizing local government has an incentive to behave optimally.

Next we consider the case of cities of type 1 which produce the differentiated product. For a given $m_1$, which is both the number of cities of type 1 and the number of varieties, each of these cities has a unique product and therefore faces a downward sloping Marshalian demand curve $z_1^M(m, q, I)$, given by

$$Z_1^M(m, q, I) = \frac{1}{m_1} \sum_{i=1}^{I+J} m_i \bar{Z}_1^i(m, q, V_i(m, q, I_i)) \tag{28}$$

where $I$ is the vector of local incomes, and in the r.h.s the indirect utility functions $V_i(m, q, I_i)$ are replacing the utility variable in the unconditional compensated demand curves, $z_1^i(m, q, u)$.[7] Differentiation of Eq. (28) with respect to $p_1$ yields

$$\frac{\partial Z_1^M}{\partial p_1} = \frac{1}{m_1} \sum_{i=1}^{I+J} m_i \left\{ \frac{\partial \bar{Z}_1^i(m, q, u)}{\partial p_1} + \frac{\partial \bar{Z}_1^i(m, q, u)}{\partial u} \frac{\partial V_i(m, q, I_i)}{\partial p_1} \right\} \Bigg|_{u=V_i(m, q, \nu, I_i)} < 0 \tag{29}$$

where the prices $\nu$ and $R_A$ are included in the vector $q$.

The sign in Eq. (29) follows upon assuming normality of $z_1$, i.e. $\partial \bar{Z}_1^i / \partial u > 0$. we can now calculate the marginal revenue function,

$$MR_1(m, q, 1) = p_1 \left( 1 + \frac{Z_1^M}{p_1(\partial Z_1^M / \partial p_1)} \right) < p_1 \tag{30}$$

The producer in the city maximizes his profits upon equating $MR_1$ to $MC_1^e$, which is far from the optimum, but as in the previous section, a net-city-profit-maximizing local government will steer him by subsidizing his labor, a subsidy financed by land taxes, to the point where

$$MR_1(m, q, I) = MC_1^0(Z_1^M(m, q, I), q, u) \tag{31}$$

It should be noted that all the arguments in the above equation are taken as given

[7]One should remember that $z_1$ was defined to be the sum of all the varieties of the differentiated products. Hence the optimal quantity of a single variety is $z_1^* / m_1$.

parameters by both the firm and the city except for $p_1$, which is taken to be a function of the output.

The above result does not imply marginal cost pricing and therefore is definitely not the first-best. It is a second-best solution compared to the outcome when there is no intervention, not even local, since while keeping utility constant, city profits and the accompanying net economic surplus are necessarily higher under the city profit maximizing regime. It is conceivable that in the short run, local governments might be induced to attain marginal cost pricing even without federal financial aid, if the federal government enforces price regulations. However such a trend may be counterproductive in the long run, since the number of cities is well under the long run optimal number when all costs are covered under marginal cost pricing. Moreover, under these conditions there is no incentive to additional cities of type 1 to enter so that the long run optimum cannot be attained.

We now consider the long run market allocation. Cities will enter until the net city profits, $\pi_j$, for all j, vanishes. For all types of cities other than $j=1$, this condition coincides with optimal condition (Eq. (17)); however this is not the case for cities of type 1. First it should be noted that when both Eqs. (17) and (31) hold, average cost pricing prevails; i.e. $p_1 = AC^0$ in the range where average cost is downward sloping. This is the classical monopolistic competition solution. Observe that the optimum is also reached in the same range of $AC^0$, as follows from Eq. (27).

There are two distortive effects which act in opposite directions and thus tend to cancel each other out, at least with respect to the number of varieties/cities. One of these effects is the distortion caused by monopolistic competition which acts to increase the number of cities compared to the optimum. The other effect is caused when cities ignore the advantages of an increase in the number of varieties. Such an increase is a benefit to the general public and ignoring it leads to a reduction in the number of cities compared to the optimum.[8] Therefore, like in the traditional I.O. literature (see Lancaster, 1979) the number of cities in the market allocation, which also equals the number of varieties, may exceed, coincide with, or be less then the optimal number, depending on the particular production and utility functions involved. It should be noted that even when the number of cities in the optimal and market allocations coincide, the two allocations differ, since the optimum involves marginal cost pricing, contrary to the average cost pricing of the market allocation. This implies that the optimal cities will be larger and each will produce more of its product than in the market allocation because the downward sloping demand curve for each of the differentiated product varieties meets its respective marginal cost curve at a higher level of output than it meets the respective average cost curve.

To achieve efficiency the federal government, and only the federal government, has to subsidize the differentiated product. The subsidy has to be equal to the

[8]This argument is one of the suggestions made by David Pines.

difference between the market price and MR at the optimum, so that the price producers get is $p_1^{p*} = AC^0(Z_1^*)$. However, consumers pay $p_1^* = P_1^{p*} - s = MR_1(Z_1^*) = MC^0(Z_1^*)$, where the asterisks denote optimal values, and $p_1^p$ is the price facing producers. Thus the subsidy, $s$, given below is

$$s = -\frac{Z_1^*}{(\partial Z_1^M(p_1^*)/\partial p_1)} = \frac{1}{Z_1^*}\sum_{i=1}^{I+J} m_i \int_0^{\bar{x}_i} \frac{\theta_i(x)\cdot U_{m_i}(i, x)}{h_i(x)\cdot U_{z_n}(i, x)}\,dx \tag{32}$$

The last equality in Eq. (32) follows from Eq. (27) and the fact that the difference between the demand price and marginal revenue in the optimum is also the difference between $AC^0$ and $MC^0$. This subsidy causes the marginal revenue curve facing the cities of type 1 to intersect the demand curve at the optimum output level. Local governments do not have the incentive nor the means to finance such a subsidy so the federal government has to use a lump sum head tax through $\nu$ to finance this subsidy efficiently.

## 6. Concluding remarks

In this paper a general and comprehensive model of an economy with industries subject to scale economies was constructed and used to reexamine two models from the urbaneconomics literature in which scale economies are internal to he individual firm. The first model (Kanemoto, 1980) is a general equilibrium model of cities, each with a single firm representing an export good industry with scale economies internal to the firm. The model implies that if the firm is also the developer of the city, then the equilibrium allocation is efficient. If, however, the firm owns only its industrial production site, then the equilibrium is inefficient.

These results are shown in this paper to be robust. In addition, we demonstrate that a net-city-profit-maximizing local government (PMLG) will use an employment subsidy to the local export firm, financed by local land taxes, to maximize net city profits and thus achieve efficiency.

The second model ((Henderson and Abdel-Rahman, 1991) (HAR)) deals with a differentiated product of which each variety is produced via a production function with a fixed labor component and another component of fixed marginal productivity of labor. Labor, the only input, is also used in commuting. Each household consumes a unit of land which does not contribute to utility. There are two versions of the utility function: a simple one, where the utility is given by a CES function of only the varieties of the differentiated product, and a slightly more complex version in which a regular product is added as a consumption good and an output of an industry.

The main problem with the simple version of this model is not just that it uses specific functions, but that it lacks degrees of freedom. Once the number of cities is determined, all other factors of this simple model are determined as well.

Consider, for example, the result of our paper which states that if the number of cities in the optimal allocation and in the market allocation, with only PMLG intervention, are the same, then the cities producing the differentiated product would, in the optimal allocation, be larger and would be producing more of the differentiated product. Such a result could not be attained in the HAR simple model since once the number of cities is the same, all other factors are determined uniquely. Hence, an equal number of cities implies identical allocations. All three propositions of the HAR study deal with the simple model. In fact, when HAR examine their more general model, they themselves show that most of the results of their simple model are not robust.

Only one of their results is robust and holds in the more general model of our paper. This is when the solution brought about by local government intervention, in the form of an employment subsidy in the differentiated product industry, is a second best solution when the federal government is prohibited from intervening. This result is quite important and has practical implications. In a numerical example of their expanded model, HAR show that the deviation from optimal city size in the second best is only 6%, compared to a deviation of about 100% in the laissez faire case.

In our model, we show that for a fixed number of cities/varieties, short run marginal cost pricing is still the efficient condition. However, pursuing short run efficiency may prove to be counterproductive with respect to the long run. The long run efficiency condition is somewhat similar to Samuelson's rule concerning the efficient allocation of public goods. Increasing the number of cities/varieties raises the utility of the entire population, but has a cost. The long run efficiency condition states that in the optimum the sum of the rates of substitution between the number of varieties and the numeraire product should be equal to the cost of an additional city. This cost of an additional city is the difference between the total costs of producing the variety of the city in question and the value of that product evaluated by its marginal cost. The laissez faire market solution is a variant of classical monopolistic competition. Similar to the results of HAR, local government intervention in the form of subsidies financed by land taxes improve the allocation. But unlike HAR who argue in favor of federal price regulation and lump sum subsidies to cities, the first best can also be achieved by a federal subsidy per unit of the differentiated product. This federal subsidy is financed by a federal lump sum head tax.

We have also shown that the reason competition between PMLGs yields a second best solution is due to the fact that there are two distortions that act in opposite directions and therefore tend to cancel each other out. One is the distortion caused by monopolistic competition between the PMLGs which tends to increase the number of cities/varieties. The second is due to the fact that the benefits of additional varieties are external to the local governments and therefore do not take them into account. This distortion acts to reduce the number of cities/varieties.

In summary, monopolistic competition of local governments may yield more, fewer, or an equal number of cities/varieties as the optimum. But, as stated above, even if the number of cities were the same, the size and output of the differentiated product industry is below its optimal level.

It is also argued that to achieve the first-best, federal government intervention is needed, in addition to the PMLG intervention, in the form of subsidies per unit of differentiated product. The federal intervention can also be in terms of price regulations and a lump sum subsidy to the firm. But price regulations, contrary to HAR's reasoning without federal subsidy will not do the job since the local government does not have sufficient funds at its disposal.

## References

Arnott, R.J., 1979, Optimal City Size in a Spatial Economy, Journal of Urban Economics 6, 65–89.

Fujita, M., 1988, Urban Economic Theory, (Cambridge University Press, Cambridge).

Henderson, J.V., 1974, The Size and Types of Cities, American Economic Review, 64.

Henderson, J.V., 1985, Economic Theory and the Cities, 2nd ed., (Academic Press, NY).

Henderson, J.V. and H. Abdel-Rahman, 1991, Urban Diversity and Fiscal Decentralization, Regional Science and Urban Economics 21, 491–509.

Hochman, O., 1981, Land Rents, Optimal Taxation and Local Fiscal Independence in an Economy with Local Public Goods, Journal of Public Economics 15, 290–310.

Hochman, O., 1990 Cities, Scale Economics, Local Goods and Local Governments, Urban Studies 27, 45–64.

Kanemoto, Y., 1980, Theories of Urban Externalities, (North Holland, Amsterdam).

Krugman, P., 1991, Increasing Returns and the Economic Geography, Journal of Political Economy 99,483–499.

Lancaster, K., 1979, Variety, Equity and Efficiency, (Columbia University Press, New York).

Varian, Hal R., 1992, Microeconomic Analysis (3rd edition), (Norton, New York).