C·I·R·P·É·E

Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

City Size and the Henry George Theorem under Monopolistic Competition

Kristian Behrens
Yasusada Murata

Octobre/October 2008

_____

Behrens: Department of Economics, Université du Québec à Montréal (UQAM), Case postale 8888, Succursale Centre-Ville, Montréal (Québec), Canada H3C 3P8 ; CIRPÉE, Canada ; and CORE, Université Catholique de Louvain, Belgium.
behrens.kristian@uqam.ca
Murata: Advanced Research Institute for the Sciences and Humanities (ARISH), Nihon University, 12-5, Gobancho, Chiyoda-ku, Tokyo 102-8251, Japan; and Nihon University Population Research Institute (NUPRI), Japan.
murata.yasusada@nihon-u.ac.jp

Abstract:
We analyze the equilibrium and the optimal resource allocations in a monocentric city under monopolistic competition. Unlike the constant elasticity of substitution (CES) case, where the equilibrium markups are independent of city size, we present a variable elasticity of substitution (VES) case where the equilibrium markups fall with city size. We then show that, due to excess entry triggered by such pro-competitive effects, the 'golden rule' of local public finance, i.e., the Henry George Theorem (HGT) , does not hold at the second best. We finally prove, within a more general framework, that the HGT holds at the second best under monopolistic competition if and only if the second-best allocation is first-best efficient, which reduces to the CES case.

1 Introduction

Most modern economies feature imperfectly competitive industries that produce differentiated varieties under increasing returns to scale. In a spatial context, such industries generate various externalities through sharing, matching, and learning, which induce urban agglomeration (see Duranton and Puga, 2004, for a recent survey). Sharing externalities of the new economic geography type, stemming either from greater product diversity in consumption (Dixit and Stiglitz, 1977) or from a wider array of differentiated intermediate inputs (Ethier, 1982) have, in particular, attracted a lot of attention in recent years. The main reasons for this seem to be that their microeconomic underpinnings are better understood, and that appropriate general equilibrium modeling tools drawing on monopolistic competition have become increasingly more popular.

As cities are the main centers of economic activity, a thorough analysis of the equilibrium and the optimal resource allocations within them is desirable. In the context of monopolistic competition with differentiated goods, these questions have been studied, among others, by Abdel-Rahman and Fujita (1990). Building on the Dixit-Stiglitz constant elasticity of substitution (henceforth, CES) model, they show that the 'golden rule' of local public finance, i.e., the *Henry George Theorem* (henceforth, HGT; Flatters *et al.*, 1974; Stiglitz, 1977; Arnott, 1979; Arnott and Stiglitz, 1979) holds even at the second best where the planner takes the equilibrium prices as given. Unfortunately, this neat result is likely to hinge on the CES specification which displays two peculiar features. First, it does not allow for pro-competitive effects and, therefore, markups are independent of city size. Second, the market provides optimum product diversity at an efficient scale of production (see Dixit and Stiglitz, 1977, Section I).

The HGT has both theoretical and empirical implications. In our monopolistic competition framework, the HGT implies that aggregate land rents equal aggregate fixed costs for producing differentiated goods. If it holds, a single confiscatory tax on land rents can raise enough revenue to implement the first-best allocation. This property is important given that deviations from optimal city sizes are quite costly in terms of productivity and welfare, as recently quantified by Au and Henderson (2006a, b). Turning to the empirical analysis, as shown by Kanemoto *et al.* (1996) and as discussed by Arnott (2004), when the HGT holds, it can be used to test whether cities are too big or too small. It is, therefore, important to know when the HGT holds as well as whether aggregate land rents exceed or fall short of aggregate fixed costs when it fails.

To derive more general results under monopolistic competition, we depart from the CES model and explore when the HGT holds at the second best where the planner takes the equilibrium prices as given. Such exploration is relevant especially because in the real world the planner can usually control only a subset of variables. As is well known, the HGT holds

in "all large economies [. . .] in which the distribution of economic activity over space is Pareto optimal" (Arnott and Stiglitz, 1979, p.472). In other words, the HGT holds in the first-best world where the planner can control all the relevant variables. In a second-best world with imperfect competition, the results are far from clear. Abdel-Rahman and Fujita (1990), for example, show that the HGT holds under CES monopolistic competition in the intermediate input sector, whereas Helsley and Strange (1990) find that it does not hold at the second best within a matching framework where firms face perfectly elastic demands and profits are distributed through wage bargaining.

To the best of our knowledge, there has been until now no attempt to verify whether or not the HGT holds at the second best within non-CES monopolistic competition frameworks encompassing pro-competitive effects, gains from product diversity, and losses from urban costs. As pointed out by Fujita *et al.* (2004, p.2934), this is

> "[. . .] a serious issue because the main part of urban agglomeration economies arises from locational externalities in an NEG type spatial economy. A market equilibrium in a model of this type is not in general Pareto optimal. As far as we know, Abdel-Rahman and Fujita (1990) is the only one that explicitly examines whether or not the Henry George Theorem holds in an NEG type model. Their result is that, in a model where the Dixit-Stiglitz type structure is assumed for intermediate products, the Henry George Theorem holds even in the second best. We do not yet know if this result is general, but it is possible that the theorem holds either exactly or approximately in a more general setting."

The main objective of this paper is threefold. First, we extend the general equilibrium model of monopolistic competition with variable elasticity of substitution (henceforth, VES) by Behrens and Murata (2007a) to a monocentric city setting. Second, using this framework, we investigate the relationships among city sizes, markups, and whether or not the HGT holds at the second best, either exactly or approximately. Last, we derive, within a more general framework, necessary and sufficient conditions for the HGT to hold in second-best economies under monopolistic competition.

Previewing the main results, we first show that a larger city has lower equilibrium markups in our VES model. This result is in accord with empirical evidence supporting the hypothesis that larger and denser urban areas are more competitive (e.g., Syverson, 2004; Campbell and Hopenhayn, 2005). We then show that, due to excess entry triggered by such pro-competitive effects, the HGT does not hold at the second best in that model. We finally prove, within a more general framework, that the HGT holds at the second best under monopolistic competition if and only if the second-best allocation is first-best efficient, which turns out to be equivalent to the CES case. Hence, the HGT does not hold at the second best in VES models of monopolistic competition.

The remainder of the paper is organized as follows. In Section 2, we present the model, whereas in Section 3 we derive the price equilibrium, the equilibrium mass of firms, and the equilibrium utility. Section 4 then deals with the optimal city size and the HGT in the first- and second-best cases. Section 5 concludes.

2   Model

Consider a monocentric city with a mass $L > 0$ of identical consumers/workers, as well as with a large amount of homogeneous land. The land stretches out along a one-dimensional space X, and the amount of land available at each location $x \in X$ is set to one. All firms in the city are set up at an exogenously given Central Business District (henceforth, CBD). In what follows, we assume that labor is the only factor of production and that land is used for housing only, i.e., firms do not consume land and the CBD is dimensionless. Without loss of generality, we label locations such that this CBD is located at $x = 0$. Each agent consumes inelastically one unit of land, supplies inelastically one unit of labor, and commutes to the CBD for work. This implies that workers are symmetrically distributed around the CBD and that the city covers the interval $[-L/2, L/2]$.

Following Murata and Thisse (2005), we assume that commuting costs are of the 'iceberg' type: the effective labor supply of a worker living at a distance $|x| \leq L/2$ from the CBD is given by

$$s(x) = 1 - 2\theta|x|. \tag{1}$$

In expression (1), the parameter $\theta > 0$ captures the efficiency loss due to commuting. For the labor supply in efficiency units to be positive regardless of the worker's location x in the city, we assume throughout the paper that $\theta < 1/L$. Consequently, the *aggregate effective labor supply* at the CBD is given by

$$S(L) = \int_{-\frac{L}{2}}^{\frac{L}{2}} s(x)\mathrm{d}x = L\left(1 - \frac{\theta L}{2}\right). \tag{2}$$

Let w stand for the wage rate paid to the workers by the firms at the CBD. Then, the wage income net of commuting costs earned by a worker residing at either city edge is such that $s(-L/2)\, w = s(L/2)\, w = (1 - \theta L)w$. Without loss of generality, we normalize the opportunity cost of land to zero. Because workers are identical, the wages net of commuting costs and land rents are equalized across all locations: $s(x)w - R(x) = s(-L/2)w = s(L/2)w$, where $R(x)$ is the land rent prevailing at x, and $R(L/2) = R(-L/2) = 0$. For a given spatial distribution of workers, the equilibrium land rent schedule in the city is therefore given by $R^*(x) = \theta(L - 2|x|)w$, which yields the following aggregate land rents:

$$\mathrm{ALR} = \int_{-\frac{L}{2}}^{\frac{L}{2}} R^*(x)\mathrm{d}x = \frac{\theta L^2 w}{2}.$$

4

In what follows, we assume that each worker owns an equal share of land in the city. Accordingly, in addition to her wage, each worker receives an equal share ALR/L of aggregate land rents from her land ownership.

There is a single monopolistically competitive industry producing a horizontally differentiated consumption good provided as a continuum of varieties. Let Ω be the set of varieties produced in the city, the mass N of which is endogenously determined. Because agents have the same claim to aggregate land rents, irrespective of their location x in the city, they make the same consumption decisions. The representative consumer solves the following utility maximization problem (which, as argued in the foregoing, is independent of x):

$$\max_{q(i),\ i\in\Omega} \quad U \equiv \int_\Omega u\big(q(i)\big)\mathrm{d}i \tag{3}$$

$$\text{s.t.} \quad \int_\Omega p(i)q(i)\mathrm{d}i = E, \tag{4}$$

where E stands for expenditure; $p(i)$ denotes the price of variety i; $q(i)$ stands for the per-capita consumption of variety i; and u is a strictly increasing and strictly concave, twice continuously differentiable sub-utility function. In what follows, we investigate two alternative models by assuming that the sub-utility u is either

$$\text{of the CES type:} \quad u\big(q(i)\big) \equiv q(i)^{\frac{\sigma-1}{\sigma}}, \qquad \sigma > 1$$

$$\text{or of the CARA type:} \quad u\big(q(i)\big) \equiv 1 - \mathrm{e}^{-\alpha q(i)}, \qquad \alpha > 0,$$

which allows us to derive closed-form solutions for the demand functions.

Maximizing utility (3), subject to the budget constraint (4), yields the following demand functions for the CES case:

$$q(i) = \frac{p(i)^{-\sigma}}{\displaystyle\int_\Omega p(j)^{1-\sigma}\mathrm{d}j}E, \tag{5}$$

whereas, as shown by Behrens and Murata (2007a), the first-order conditions for an interior solution yield the following demand functions for the CARA case:

$$q(i) = \frac{E - \dfrac{1}{\alpha}\displaystyle\int_\Omega \ln\left[\dfrac{p(i)}{p(j)}\right]p(j)\mathrm{d}j}{\displaystyle\int_\Omega p(j)\mathrm{d}j}. \tag{6}$$

Because of the continuum assumption, firms are negligible so that the price elasticities of demand in the CES and CARA cases are as follows:

$$\epsilon(i) = -\frac{\partial q(i)}{\partial p(i)}\frac{p(i)}{q(i)} = \sigma \tag{7}$$

$$\epsilon(i) = -\frac{\partial q(i)}{\partial p(i)}\frac{p(i)}{q(i)} = \frac{1}{\alpha q(i)}. \tag{8}$$

As usual, the CES case (7) features a constant elasticity, whereas the CARA case (8) features an elasticity that falls with the quantity $q(i)$ consumed.

Turning to the production side, all firms have access to the same increasing returns to scale technology. To produce $Lq(i)$ units of any variety requires $cLq(i) + F$ units of labor, where F is the fixed and c is the marginal labor requirement, respectively. We assume that firms can costlessly differentiate their products and that there are no scope economies. Thus, there is a one-to-one correspondence between firms and varieties, so that the mass of varieties N also stands for the mass of firms operating in the city. The profit of firm i is then given as follows:

$$\pi(i) = Lq(i)\left[p(i) - cw\right] - Fw, \tag{9}$$

where $q(i)$ is given by (5) or by (6), depending on whether we focus on the CES or the CARA case, respectively.

3   Equilibrium

We now solve the model for the equilibrium prices and the free entry mass of firms. To do so, we find it convenient to proceed in two steps.

First, firms maximize their profit (9) with respect to $p(i)$, taking (w, E, N) as given since they have no influence on these variables.[1] This yields the following first-order conditions:

$$\frac{\partial \pi(i)}{\partial p(i)} = Lq(i)\left\{1 - \left[1 - \frac{cw}{p(i)}\right]\epsilon(i)\right\} = 0, \qquad \forall i \in \Omega. \tag{10}$$

A *price equilibrium* is defined as a distribution of prices satisfying conditions (10). Note that $\epsilon(i)$ does not depend on $q(i)$ in the CES case, whereas it does depend on $q(i)$ in the CARA case. Since, as shown by (6), $q(i)$ depends itself on two price aggregates in the latter case, condition (10) highlights a fundamental property of VES monopolistic competition models with a continuum of firms: although each firm is negligible to the market, it must take into account the price aggregates that enter its first-order condition.

Let us start with the well-known CES case. Inserting (7) into (10), the unique price equilibrium is trivially symmetric and given as follows:

$$p(i) = p \equiv \frac{\sigma}{\sigma - 1}cw, \qquad \forall i \in \Omega. \tag{11}$$

Turning to the CARA case, the price equilibrium is determined by inserting (8) into (10). Behrens and Murata (2007a) have shown that the price equilibrium is symmetric, unique, and given by

$$p(i) = p \equiv cw + \frac{\alpha E}{N}, \qquad \forall i \in \Omega. \tag{12}$$

---

[1]It is well known that price competition and quantity competition yield the same outcome in monopolistic competition models with a continuum of firms (see Vives, 1999, p.168).

Two comments are in order. Firstly, unlike in the CES case, the markup is increasing in expenditure in the CARA case. The reason is that, as shown by expression (8), the elasticity of demand falls with the consumption level. Stated differently, when expenditure is large, firms face less elastic demands and, therefore, charge a higher markup. Secondly, in the CARA case, the markup falls with the mass of competing firms in the city, i.e., there are *pro-competitive effects*.

Second, given a price equilibrium, firms enter in and exit from the market until they earn zero profits. Furthermore, the labor market clears. Hence, an *equilibrium* is a solution to the following two equations:

$$Lq(i)\left[p(i) - cw\right] = Fw, \qquad \forall i \in \Omega, \tag{13}$$

$$\int_\Omega \left[cLq(i) + F\right]\mathrm{d}i = S(L), \tag{14}$$

where all prices and quantities are evaluated at a price equilibrium. The budget constraint (4), the zero profit condition (13), and the labor market clearing condition (14) then yield

$$E = \frac{S(L)}{L}w = (1 - \theta L)w + \frac{\mathrm{ALR}}{L}. \tag{15}$$

Given that the price equilibrium is symmetric in both cases, the quantities are also symmetric. Inserting $q = E/(Np)$ into (14) and using (15), we obtain the free entry mass of firms as a function of the wage-price ratio as follows:

$$N = \frac{S(L)}{F}\left(1 - \frac{cw}{p}\right). \tag{16}$$

The labor market clearing condition (14) can also be rewritten as

$$q(L, N) = \frac{1}{cL}\left[\frac{S(L)}{N} - F\right] > 0, \tag{17}$$

an expression that will prove useful when comparing the equilibrium and the optimal allocations.

3.1 CES case

The equilibrium mass of firms can be determined by substituting (11) into (16). Its expression is given by:

$$N^* = \underbrace{\left(1 - \frac{\theta L}{2}\right)}_{=S(L)/L}\frac{L}{\sigma F}. \tag{18}$$

It is worth noting that in the absence of commuting costs ($\theta = 0$), expression (18) reduces to the standard equilibrium mass of firms in the CES model, given by $L/(\sigma F)$. When there are

7

commuting costs ($\theta > 0$), this equilibrium mass is reduced by a factor of $S/L < 1$, which is the *average effective labor supply* in the city. This captures the fact that higher commuting costs decrease average effective labor supply, which negatively affects the equilibrium mass of firms. It can be readily verified that the equilibrium mass of firms is nevertheless strictly increasing in the city size L and strictly decreasing in the commuting costs θ (recall that $L < 1/\theta$). Note also that the equilibrium output per firm $Q^* \equiv Lq^* = (\sigma - 1)F/c$ is independent of city size, as is the markup. Stated differently, larger cities are not more competitive and do not have larger firms producing more output.

Evaluating (3) for the CES case, using (17), yields the following indirect utility:

$$U(N) = N \left\{ \frac{1}{cL} \left[\frac{S(L)}{N} - F \right] \right\}^{\frac{\sigma-1}{\sigma}}. \tag{19}$$

Finally, inserting (18) into (19) yields

$$U(L) = \kappa \left(1 - \frac{\theta L}{2} \right) L^{\frac{1}{\sigma}},$$

where $\kappa \equiv (F\sigma)^{-1} \left[F(\sigma - 1)/c \right]^{(\sigma-1)/\sigma} > 0$ is a bundle of parameters. It is readily verified that U is a strictly concave and single-peaked function of L on the interval $(0, 1/\theta)$.

3.2   CARA case

The equilibrium mass of firms can be determined by substituting (12) into (16) and by using (15). Its expression is given by:[2]

$$N^* = \underbrace{\left(1 - \frac{\theta L}{2} \right)}_{=S(L)/L} \frac{D(L) - \alpha F}{2cF}, \quad \text{where} \quad D(L) \equiv \sqrt{4\alpha cFL + (\alpha F)^2}. \tag{20}$$

It is worth pointing out that in the absence of commuting costs ($\theta = 0$), expression (20) reduces to the equilibrium mass of firms in Behrens and Murata (2007a). When there are commuting costs ($\theta > 0$), this equilibrium mass is reduced by the average effective labor supply in the city, for the same reasons as in the CES case. However, unlike in the CES case, where an increase in L always raises N^*, the equilibrium mass of firms need not increase in city size. There exists indeed a unique threshold $\overline{L}$, given by

$$\overline{L} \equiv \frac{6c - \alpha F\theta + \sqrt{6\alpha cF\theta + (\alpha F\theta)^2}}{9c\theta} \in \left(0, \frac{1}{\theta} \right), \tag{21}$$

such that $\partial N^*/\partial L \gtreqless 0$ for all $L \lesseqgtr \overline{L}$. Put differently, when the city is large enough, an additional increase in city size may reduce the equilibrium mass of firms. The intuition for

---

[2]Note that the other root is negative and must, therefore, be ruled out.

this result is that, as can be seen from (20), the negative effect of L on average effective labor supply affects N^* linearly, whereas the positive effect affects N^* less than linearly when L gets sufficiently large, due to the presence of pro-competitive effects. When L exceeds the threshold (21), the former effect dominates the latter and the equilibrium mass of firms falls. Note that $\overline{L}$ increases in α and F, and decreases in θ. Hence, N^* increases over a larger range of city sizes when workers incur smaller efficiency losses from commuting (θ) and when firms have more monopoly power (α) and face larger fixed labor requirement (F).

Let us summarize our findings as follows:

Proposition 1 (CARA equilibrium mass of firms) *The equilibrium mass of firms is increasing (resp., decreasing) in city size when L is smaller (resp., larger) than the threshold $\overline{L}$, given by (21).*

Note that, regardless of whether N^* rises or falls with L, the following result holds.

Proposition 2 (CARA equilibrium markup) *The equilibrium markup is strictly decreasing in city size.*

Proof. Some straightforward calculation using (12), (15), and (20) yields

$$p^* = \left[1 + \frac{2\alpha F}{D(L) - \alpha F} \right] cw,$$

thus showing that the equilibrium markup is strictly decreasing in L. ∎

Proposition 3 (CARA equilibrium output per firm and total output) *The equilibrium output per firm is strictly increasing and concave in city size, whereas the total output is strictly increasing and convex-concave in city size.*

Proof. To establish the proposition, note that

$$Q^* = \frac{1}{c}\left[\frac{S(L)}{N^*} - F \right] = \frac{1}{c}\left[\frac{2cFL}{D(L) - \alpha F} - F \right]. \tag{22}$$

Therefore, $\partial Q^*/\partial L > 0$ and $\partial^2 Q^*/\partial L^2 < 0$, which yields the first result.

The second result is obtained as follows. Let $N^* Q^*$ stand for the equilibrium total output. Some calculations, using expressions (20) and (22), show that $\partial(N^* Q^*)/\partial L > 0$, $\lim_{L \to 0} \partial^2(N^* Q^*)/\partial L^2 > 0$ and $\lim_{L \to 1/\theta} \partial^2(N^* Q^*)/\partial L^2 < 0$. Hence, since $\partial^3(N^* Q^*)/\partial L^3 < 0$ there exists, by continuity, a unique threshold $\check{L}$ such that $N^* Q^*$ is a convex function of L for $L < \check{L}$ and a concave function for $L > \check{L}$. ∎

Note that an increase in city size is accompanied by a rise in the equilibrium output per firm, which leads to a better exploitation of firm-level scale economies and reduces markups. Such a finding is in accord with empirical evidence suggesting that city size positively affects establishment size, and that competition is tougher in larger and denser markets (e.g., Syverson, 2004; Campbell and Hopenhayn, 2005). It is worth pointing out that, contrary to the CES case where total output is a linear function of L, the relationship is strictly convex in the CARA case for small city sizes. This is consistent with the observation by Holmes (1999, p.317), who argues that "the empirical relationship that holds for certain industries is a convexity in the relationship between local population and production". Our results suggest that the convexity/concavity of the relationship depends, among other things: (i) on the presence/absence of pro-competitive effects; and (ii) on city size.

Evaluating (3) for the CARA case, using (17), yields the following indirect utility:

$$U(N) = N \left\{ 1 - e^{-\frac{\alpha}{cL} \left[\frac{S(L)}{N} - F \right]} \right\}. \tag{23}$$

Finally, inserting (20) into (23), we obtain

$$U(L) = \left(1 - \frac{\theta L}{2} \right) \frac{D(L) - \alpha F}{2cF} \left[1 - e^{-\frac{2\alpha F}{D(L) + \alpha F}} \right]. \tag{24}$$

Unlike in the CES case, the properties of this function are less straightforward to establish. Yet, we can prove the following result:

Proposition 4 (single-peaked CARA equilibrium utility) *For all admissible parameter values of the model, i.e., $\alpha > 0$, $c > 0$, $F > 0$, and $\theta \in (0, 1/L)$, there exists a unique city size $L \in (0, 1/\theta)$ which maximizes (24).*

Proof. See Appendix A. ∎

4   Optimal city size and the Henry George Theorem

As argued in the Introduction, the HGT has both important theoretical and empirical implications. Recall that in our monopolistic competition framework, the HGT implies that aggregate land rents equal aggregate fixed costs for producing differentiated goods (e.g., Abdel-Rahman and Fujita, 1990; Helsley and Strange, 1990).[3] If it holds, a single confiscatory tax on land

---

[3]Strictly speaking, Abdel-Rahman and Fujita (1990) do not start from the aggregate fixed costs, but from *ad valorem* sales subsidies to the increasing returns sector and they show that the total subsidies equal the aggregate fixed costs for producing differentiated goods. Note that we use the increasing returns to scale version of the HGT which is common to Abdel-Rahman and Fujita (1990) and Helsley and Strange (1990) since this allows us to compare directly their results and ours. This comparison makes sense because Kanemoto (2007) shows, in a more general framework, that the total Pigouvian subsidies required for agglomeration economies equal the aggregate fixed costs.

rents can raise enough revenue to implement the first-best allocation, thus avoiding costly deviations from optimal city sizes in terms of productivity and welfare (Au and Henderson 2006a, b). Turning to the empirical analysis, as shown by Kanemoto *et al.* (1996) and as discussed by Arnott (2004), when the HGT holds, it can be used to test whether cities are too big or too small.[4] As is well known, it holds at the first best. Whether or not the HGT holds at the second best generally depends on the nature of externalities. Once we consider pecuniary externalities, instead of technological externalities, the analysis becomes more involved. The reason is that "producers are not expected to act as price takers and second best issues that are caused by price distortions complicate the analysis" (Kanemoto *et al.*, 1996, p.398).

We now turn to the questions of optimal city size and the HGT with pecuniary externalities. After reviewing the well-known results in the first-best case, we turn to a second-best world where the planner is free to choose L and N subject to the constraint that prices be supported by a market equilibrium. In what follows, we superscript first-best values with f and second-best values with s, and we impose symmetry across varieties.

4.1 First best

In the first best, the planner chooses L and N to maximize the utility of the representative agent under the economy's resource constraint. The optimization problem is given by:

$$\max_{L,N} \ Nu(q(L,N)),$$

where $q(L, N)$ is given by (17).[5] The optimality conditions can be expressed as follows:

$$Nu'(q)\frac{\partial q}{\partial L} = 0 \tag{25}$$

$$u(q) + Nu'(q)\frac{\partial q}{\partial N} = 0. \tag{26}$$

Equation (25) is the HGT because, from (2) and (17), it can be rewritten as the equality between the aggregate land rents and the aggregate fixed costs, i.e., ALR $= NFw$.

CES case. Some straightforward calculation shows that (26) implies $N = S(L)/(\sigma F)$. Plugging this into ALR $= NFw$, we readily obtain the unique first-best city size:

$$L^f = \frac{2}{\theta(\sigma+1)}, \tag{27}$$

---

[4]Kanemoto *et al.* (1996) use the HGT to investigate whether Japanese cities, in particular, Tokyo, are too big or too small. To do so, they compare the ratios of aggregate land values to the aggregate Pigouvian subsidies across cities.

[5]In what follows, we assume that $u(q) \geq 0$ and $u(0) = 0$ for maximization to make sense and to yield interior solutions. It is readily verified that both the CES and the CARA cases satisfy these conditions.

which is always smaller than $1/\theta$ since $\sigma > 1$. As expected, the optimal city size decreases with commuting costs θ and when varieties become less differentiated (larger value of σ). Note that the optimal mass of firms is the same as the equilibrium mass given by (18).

CARA case. Unlike in the CES case, the first-best city size cannot be solved analytically. Yet, it is straightforward to prove the uniqueness of the first-best solution (see Appendix B) and, therefore, to solve the model numerically. Figure 1 reveals that, as in the CES case, the optimal city size falls with commuting costs θ.

[**Insert Figure** 1 **about here**]

4.2 Second best

In the second best, the planner chooses L and N to maximize the utility of the representative agent subject to the resource constraint, the price equilibrium, and the zero profit condition. The second-best problem is given by:

$$\max_{L,N} \ Nu(q(L,N)) \qquad \text{s.t.} \quad \text{SBC} \equiv NLq(L,N) - \frac{S}{c}\left[1 - \frac{1}{\epsilon(q(L,N))}\right] = 0,$$

where $q(L,N)$ is defined as in (17) and SBC is the second-best constraint.[6] Letting μ be the Lagrange multiplier, the optimality conditions are given by

$$Nu'(q)\frac{\partial q}{\partial L} + \mu\frac{\partial \text{SBC}}{\partial L} = 0 \tag{28}$$

$$u(q) + Nu'(q)\frac{\partial q}{\partial N} + \mu\frac{\partial \text{SBC}}{\partial N} = 0. \tag{29}$$

Comparing (25) and (26) with (28) and (29), we obtain the following result.

Proposition 5 (HGT at the second best) *The Henry George Theorem holds at the second best if and only if the second-best allocation is first-best efficient.*

Proof. The HGT holds at the second best if and only if the second term of the left-hand side in (28) vanishes. There are two possibilities: either $\mu = 0$ or $\partial \text{SBC}/\partial L = 0$. In the former case, comparing (25) and (26) with (28) and (29), we immediately see that the second-best allocation is first-best efficient. The latter case never occurs at the second best. This can be shown as follows. Suppose that

$$\frac{\partial \text{SBC}}{\partial L} = 0 \quad \Longleftrightarrow \quad cN\left(q + L\frac{\partial q}{\partial L}\right) = \left(1 - \frac{1}{\epsilon}\right)\frac{\mathrm{d}S}{\mathrm{d}L} + \frac{S}{\epsilon^2}\frac{\mathrm{d}\epsilon}{\mathrm{d}q}\frac{\partial q}{\partial L}. \tag{30}$$

---

[6]Behrens and Murata (2007b) consider an alternative second-best problem in which the planner controls only city size L and takes the equilibrium mass of firms N^* as given. Note that both problems yield the same result because the second-best constraint SBC $= 0$ can be solved for N^* as a function of L.

Then, equation (28), together with $Nu'(q) > 0$, implies that $\partial q/\partial L = 0$. It then follows from (30) and (17) that $dS/dL = cqN[\epsilon/(\epsilon - 1)] = (1/L)(S - NF)[\epsilon/(\epsilon - 1)]$. Noting that (16) implies $S - NF = cw/p$ and that $\epsilon/(\epsilon - 1) = p/(cw)$, we obtain $dS/dL = 1/L$.

Furthermore, using (17), $\partial q/\partial L = 0$ yields $dS/dL = (1/L)(S-NF) = (1/L)(cw/p)$. Since marginal cost pricing violates the zero profit condition in the presence of fixed costs, we may conclude that $\partial \mathrm{SBC}/\partial L = 0$ cannot occur at the second best.

Finally, the converse is trivially true: if the second-best allocation is first-best efficient, the HGT holds. ∎

When is the second-best allocation first-best efficient? To answer this question, we differentiate expression (17) with respect to N and use $\mathrm{SBC} = 0$ to obtain

$$N\frac{\partial q}{\partial N} = -\frac{S}{cLN} = -q\frac{\epsilon}{\epsilon - 1}. \tag{31}$$

We know from the proof of Proposition 5 that the HGT holds at the second best if and only if $\mu = 0$. Plugging (31) into (29), it is readily verified that $\mu = 0$ holds if and only if

$$u(q) - u'(q)q\frac{\epsilon}{\epsilon - 1} = 0 \tag{32}$$

when evaluated at $q^s = q(L^s, N^s)$. Substituting $\epsilon = -u'(q)/[u''(q)q] > 1$ into (32) yields:[7]

$$\frac{u''(q)}{u'(q)} = \frac{u'(q)}{u(q)} - \frac{1}{q}. \tag{33}$$

Using expression (33), we obtain the following proposition.

Proposition 6 (equivalence) *The Henry George Theorem holds at the second best if and only if the sub-utility is of the CES type.*

Proof. Assume that the HGT holds at the second best. Then, by the proof of Proposition 5, $\mu = 0$ and thus (33) must hold. Equation (33) then implies $\ln u'(q) = \ln u(q) - \ln q + \text{const.}$, which implies $u'(q)/u(q) = \text{const.} \times (1/q)$. This in turn yields $\ln u(q) = \text{const.} \times \ln q + \text{const.}$, thus showing that u is of the CES type.

Conversely, if the sub-utility is of the CES type, equation (32) is satisfied. Plugging (32) into (29) yields $\mu = 0$ because

$$\frac{\partial \mathrm{SBC}}{\partial N} = L\left(q + N\frac{\partial q}{\partial N}\right) - \frac{S}{c\epsilon^2}\frac{d\epsilon}{dq}\frac{\partial q}{\partial N} = -\frac{Lq}{\epsilon - 1} < 0,$$

where we have used (31) and the property that $d\epsilon/dq = 0$. Applying $\mu = 0$ to (28) then shows that the HGT holds at the second best. ∎

---

[7]From the first-order conditions of the Lagrangian $\int_\Omega u(q(i))di + \lambda\left[E - \int_\Omega p(i)q(i)di\right]$, we have $u'(q(j)) = p(j)\int_\Omega q(i)u'(q(i))di/E$. Differentiating this expression with respect to $p(j)$ and imposing symmetry across varieties, the price elasticity of demand is given by $\epsilon = -u'(q)/[qu''(q)]$ because there is a continuum of firms.

The intuition underlying Proposition 6 can be understood in terms of entry. Recall that entry at the first best is given by (26). Entry at the second best is first-best efficient if and only if (32) holds. Using the definition of relative risk aversion, given by $r_R(q) \equiv 1/\epsilon < 1$, the right-hand side of (32) can be rewritten as

$$\varphi(q) \equiv u(q) - u'(q)q \frac{1}{1 - r_R(q)}. \tag{34}$$

Noting that excess entry occurs at the second best if and only if $\varphi(q) < 0$, we obtain the following proposition.

Proposition 7 (excess entry) *Excess entry occurs at the second best if the relative risk aversion r_R is increasing in q.*

Proof. First, $u(0) = 0$ implies $\varphi(0) \leq 0$. Differentiating φ, we obtain

$$\varphi'(q) = -\frac{u'(q)q r_R'(q)}{[1 - r_R(q)]^2},$$

thus showing that $\varphi < 0$ when $r_R' > 0$. $\blacksquare$

The excess entry result is reminiscent of Mankiw and Whinston (1987) and Vives (1999). Unlike in the basic CES model, where the relative risk aversion is constant, pro-competitive effects trigger excessive entry as $r_R' > 0$.[8]

CES case. Proposition 5 shows that the HGT holds at the second best if and only if the second best is first-best efficient, which turns out to be equivalent to the CES case by Proposition 6. Hence, the first-best and the second-best city sizes coincide and the HGT holds even in the second-best economy (Abdel-Rahman and Fujita, 1990). This peculiar result arises because entry is first-best efficient, i.e., $\varphi(q) = 0$ regardless of q, in the CES case.

CARA case. By Propositions 5 and 6, the HGT does not hold at the second best. The reason is that entry is excessive due to the presence of pro-competitive effects. This can be confirmed by Proposition 7. Indeed, in the CARA case, $r_R(q)$ is given by αq and is increasing in q. Excess entry suggests that there are losses from excessive fixed costs, which is likely to make aggregate land rents fall short of aggregate fixed costs. This intuitive result can readily be confirmed numerically.

[**Insert Figure** 2 **about here**]

---

[8]In a more general CES model, where market power and taste for variety are disentangled, the equilibrium mass of firms can be greater or smaller than the optimal one (Benassy, 1996). It should be noted, however, that this discrepancy is not due to pro-competitive effects because the model displays constant markups.

As in the first best, we cannot derive the second-best city size explicitly. Yet, because the second-best city size exists and is uniquely determined, the model can readily be solved numerically. Figure 2 depicts the difference between aggregate land rents and aggregate fixed costs as a function of commuting costs θ. As can be seen, aggregate fixed costs exceed aggregate land rents due to excess entry triggered by pro-competitive effects. One can further see that the *gap between aggregate land rents and aggregate fixed costs increases as θ decreases*, which is in accord with the numerical results obtained by Helsley and Strange (1990, p.209) in a matching model. In words, the HGT holds approximately in cities with high commuting costs, whereas it does not hold even approximately in cities with low commuting costs.

5   Conclusion

We have analyzed the equilibrium and the optimal resource allocations in a monocentric city under monopolistic competition. Summarizing our key insights, we have shown that a larger city has lower equilibrium markups in the CARA model. Furthermore, the HGT does not hold in that model at the second best: aggregate fixed costs exceed aggregate land rents because of excess entry triggered by pro-competitive effects. As the gap between aggregate land rents and aggregate fixed costs increases when commuting costs fall, the HGT is not likely to hold, even approximately, in cities with low commuting costs. More generally, we have proved that the HGT holds in our second-best economies under monopolistic competition if and only if the second-best allocation is first-best efficient, which is equivalent to the CES case. Therefore, the HGT does not hold in second-best economies under monopolistic competition with variable elasticities.

In this paper, we have focused entirely on the simple case of a monocentric city in order to derive the necessary and sufficient conditions for the HGT to hold at the second best in general equilibrium models of monopolistic competition. Nevertheless, we believe that the basic mechnism of our VES model carries over to various spatial settings such as a classical system of cities (Henderson, 1974), a core-periphery economy (Krugman, 1991) and a hierarchical urban system (Fujita *et al.*, 1999). The reason is that, as shown in the paper, excess entry due to pro-competitive effects occurs at the second best. To derive the exact necessary and sufficient conditions for each case is left for future research.

References

[1] Abdel-Rahman, H. and M. Fujita (1990) Product variety, Marshallian externalities, and city sizes, *Journal of Regional Science* 30, 165-184.

[2] Arnott, R.J. (1979) Optimal city size in a spatial economy, *Journal of Urban Economics* 6, 65-89.

[3] Arnott, R.J. (2004) Does the Henry George Theorem provide a practical guide to optimal city size?, *American Journal of Economics and Sociology* 63, 1057-1090.

[4] Arnott, R.J. and J.E. Stiglitz (1979) Aggregate land rents, expenditure on public goods, and optimal city size, *Quarterly Journal of Economics* 93, 471-500.

[5] Au, C.C. and J.V. Henderson (2006a) How migration restrictions limit agglomeration and productivity in China, *Journal of Development Economics* 80, 350-388.

[6] Au, C.C. and J.V. Henderson (2006b) Are Chinese cities too small, *Review of Economic Studies* 73, 549-576.

[7] Behrens, K. and Y. Murata (2007a) General equilibrium models of monopolistic competition: a new approach, *Journal of Economic Theory* 136, 776-787.

[8] Behrens, K. and Y. Murata (2007b) City size and the Henry George theorem under monopolistic competition, CORE Discussion Paper #2007/63, Univ. catholique de Louvain.

[9] Benassy, J.-P. (1996) Taste for variety and optimum production patterns in monopolistic competition, *Economics Letters* 52, 41-47.

[10] Campbell, J.R. and H.A. Hopenhayn (2005) Market size matters, *Journal of Industrial Economics* 53, 1-25.

[11] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297-308.

[12] Duranton, G. and D. Puga (2004) Micro-foundations of urban agglomeration economies. In: Henderson, J.V. and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics, vol. 4.* Elsevier, North-Holland, 2063-2117.

[13] Ethier, W.J. (1982) National and international returns to scale in the modern theory of international trade, *American Economic Review* 72, 389-405.

[14] Flatters, F., J.V. Henderson and P. Mieszkowski (1974) Public goods, efficiency, and regional fiscal equalization, *Journal of Public Economics* 3, 99-112.

[15] Fujita, M., P. Krugman, and T. Mori (1999) On the evolution of hierarchical urban systems, *European Economic Review* 43, 209-251.

[16] Fujita, M., T. Mori, J.V. Henderson, and Y. Kanemoto (2004) Spatial distribution of economic activities in Japan and China. In: Henderson, J.V. and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics, vol. 4.* Elsevier, North-Holland, 2911-2977.

[17] Helsley, R.W. and W.C. Strange (1990) Matching and agglomeration economies in a system of cities, *Regional Science and Urban Economics* 20, 189-212.

[18] Henderson, V.J. (1974) The sizes and types of cities, *American Economic Review* 64, 640-656.

[19] Holmes, T.J. (1999) Scale of local production and city size, *American Economic Review Papers and Proceedings* 89, 317-320.

[20] Kanemoto, Y. (2007) The Henry George Theorem in a second best economy: a note, University of Tokyo, mimeographed.

[21] Kanemoto, Y., T. Ohkawara and T. Suzuki (1996) Agglomeration economies and a test for optimal city sizes in Japan, *Journal of the Japanese and International Economies* 10, 379-398.

[22] Krugman, P. (1991) Increasing returns and economic geography, *Journal of Political Economy* 99, 483-499.

[23] Mankiw, N.G. and M.D. Whinston (1986) Free entry and social inefficiency, *RAND Journal of Economics* 17, 48-58.

[24] Murata, Y. and J.-F. Thisse (2005) A simple model of economic geography à la Helpman-Tabuchi, *Journal of Urban Economics* 58, 137-155.

[25] Stiglitz, J.E. (1977) The theory of local public goods. In: Feldstein, M.S. and R.P. Inman (eds.) *The Economics of Public Services.* MacMillan: London, 274-333.

[26] Syverson, C. (2004) Market structure and productivity: a concrete example, *Journal of Political Economy* 112, 1181-1222.

[27] Vives, X. (1999) *Oligopoly Pricing: Old Ideas and New Tools.* Cambridge, MA: MIT Press.

Appendix A: Proof of Proposition 4

We first show that (24) is a strictly concave function of the population size L for all admissible parameter values. Applying (12) and $E = [S(L)/L]w$ to $q = E/Np$, we get $q = 1/[\alpha + c\widetilde{N}(L)]$, where $\widetilde{N}(L) \equiv [L/S(L)]N^* = [D(L) - \alpha F]/2cF$. Let

$$U(L) = \left(1 - \frac{\theta L}{2}\right)\widetilde{U}(\widetilde{N}(L)), \quad \text{where} \quad \widetilde{U}(\widetilde{N}(L)) \equiv \widetilde{N}(L)\left[1 - \mathrm{e}^{-\frac{\alpha}{\alpha + c\widetilde{N}(L)}}\right].$$

Note that $\widetilde{U}$ is strictly increasing and strictly concave in L because $\widetilde{U}$ is a strictly increasing and strictly concave function of $\widetilde{N}$, which is itself strictly increasing and strictly concave in L. Premultiplying it by the affine and decreasing function $1 - \theta L/2 \geq 0$ preserves this concavity (yet, it does not in general preserve the monotonicity). Next, some longer but relatively straightforward computations show that

$$\left.\frac{\partial U}{\partial L}\right|_{L=0} = \frac{1 - \mathrm{e}^{-1}}{F} > 0$$

and

$$\mathrm{sgn}\left\{\left.\frac{\partial U}{\partial L}\right|_{L=\frac{1}{\theta}}\right\} = \mathrm{sgn}\left\{\theta\left[\alpha F - D(1/\theta)\right] - 2c\left(\mathrm{e}^{\frac{2\alpha F}{\alpha F + D(1/\theta)}} - 1\right)\right\} < 0.$$

When combined with the strict concavity of U and the continuity of $\partial U/\partial L$, these last two results prove that there is a unique value of L that maximizes (24) on $(0, 1/\theta)$.

Appendix B: Uniqueness of the first-best allocation in the CARA case

In this appendix, we prove that the first-best allocation in the CARA case is uniquely determined. A first-best solution q^f, N^f and L^f satisfies the first-order conditions (25) and (26). Expression (25), together with (17), immediately yields the HGT, i.e., $N^f = \theta(L^f)^2/(2F)$. Plugging this into (17), we obtain the quadratic equation $c\theta q^f(L^f)^2 + 2\theta F L^f - 2F = 0$ with respect to L^f. The positive solution is then given by

$$L^f = \frac{\sqrt{2c\theta F q^f + (\theta F)^2} - \theta F}{c\theta q^f}. \tag{B.1}$$

Turning to the first-order condition with respect to N, equation (26) can be rewritten as

$$\frac{\alpha q^f e^{-\alpha q^f}}{1 - e^{-\alpha q^f}} = cL^f q^f \frac{N^f}{S(L^f)} = \frac{cL^f q^f}{cL^f q^f + F}, \tag{B.2}$$

where we use (14). Inserting (B.1) into (B.2), we obtain a single equation with respect to q^f as follows:

$$\text{LHS} \equiv \frac{\alpha q^f e^{-\alpha q^f}}{1 - e^{-\alpha q^f}} = 1 - \frac{\theta F}{\sqrt{2c\theta F q^f + (\theta F)^2}} \equiv \text{RHS}. \tag{B.3}$$

We next show that equation (B.3) has a unique solution for q^f since LHS is decreasing and RHS is increasing in q^f. To see this, differentiate LHS with respect to q^f to obtain

$$\frac{\partial \text{LHS}}{\partial q^f} = -\frac{\alpha \left[1 + e^{\alpha q^f}(\alpha q^f - 1) \right]}{(1 - e^{\alpha q^f})^2} < 0.$$

Note that the last inequality is obtained as follows. Let $g(q^f) \equiv 1 + e^{\alpha q^f}(\alpha q^f - 1)$. We then get $g(q^f) > 0$ for all $q^f > 0$ because $g(0) = 0$ and $g'(q^f) > 0$ for all $q^f > 0$. Note also that RHS is increasing in q^f and that

$$\lim_{q^f \to 0} \text{LHS} = 1, \quad \lim_{q^f \to \infty} \text{LHS} = 0, \quad \lim_{q^f \to 0} \text{RHS} = 0, \quad \lim_{q^f \to \infty} \text{RHS} = 1,$$

which ensure the uniqueness of q^f. Finally, from expressions (B.1) and (B.2), we have the uniqueness of the first-best city size L^f as well as of the mass of firms N^f.
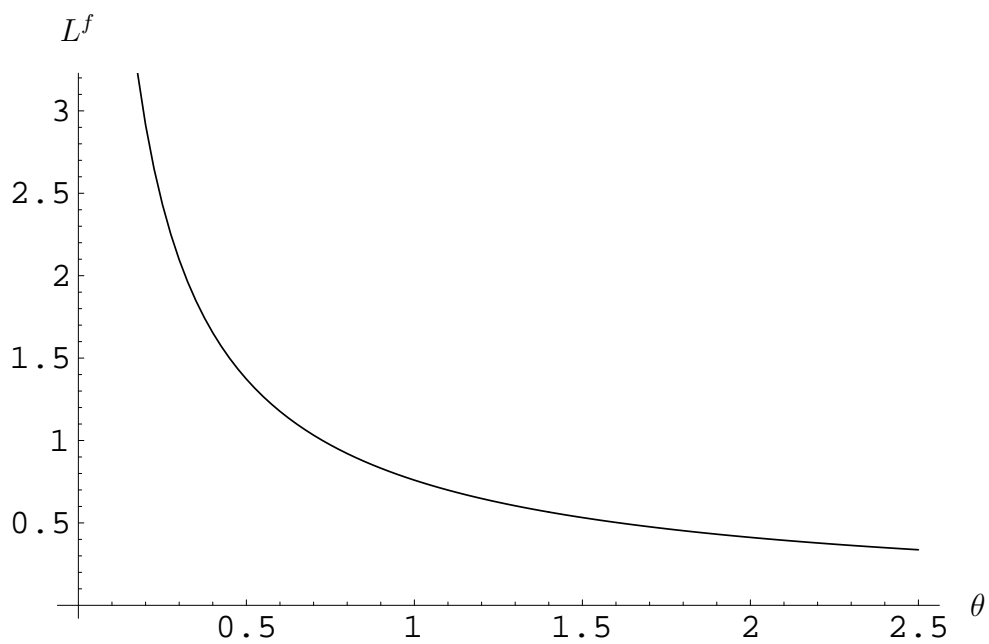
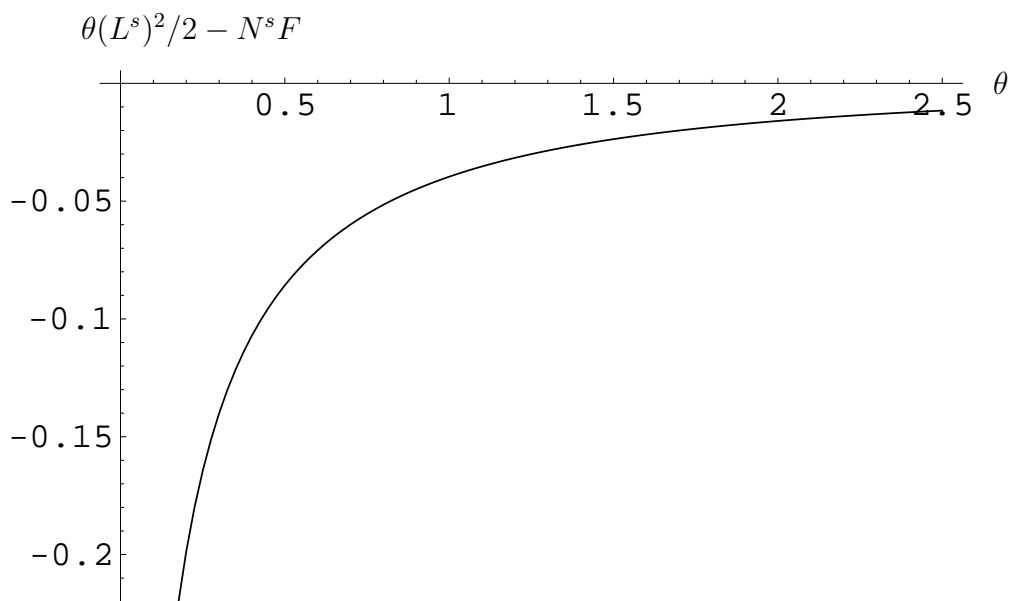Figure 1: First-best city size and commuting costs ($\alpha = 1.2$, $c = 0.3$, $F = 0.5$)

Figure 2: ALR and aggregate fixed costs ($\alpha = 1.2$, $c = 0.3$, $F = 0.5$)