

City size and the Henry George theorem under monopolistic competition

Kristian Behrens* Yasusada Murata†

August 24, 2007

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 the city size, we present a variable elasticity of substitution (VES) case where the equilibrium markups fall with the 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 in the second best. We finally prove, within our framework, that the HGT holds in the second best if and only if: (i) the second-best allocation is first-best efficient, which turns out to be equivalent to the CES case; or (ii) a marginal change in the city size has no impact on equilibrium product diversity at the second best.

Keywords: city size; Henry George theorem; monopolistic competition; first-best and second-best allocations; variable elasticity

JEL Classification: D43; R12; R13

* *Corresponding author:* Department of Economics, Université du Québec à Montréal (UQAM), Case postale 8888, Succursale Centre-Ville, Montréal (Québec), Canada H3C 3P8; and CORE, Université catholique de Louvain, Belgium. E-mail: kristian.behrens@uclouvain.be

† Advanced Research Institute for the Sciences and Humanities (ARISH), Nihon University, 12-5, Goban-cho, Chiyoda-ku, Tokyo 102-8251, Japan. E-mail: murata@arish.nihon-u.ac.jp

1 Introduction

Most modern economies feature imperfectly competitive industries which produce differentiated varieties under increasing returns to scale. In a spatial context, such industries generate various types of 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 more 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 modeling tools drawing on monopolistic competition have become increasingly 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 and Stiglitz, 1979) holds even under the second best when firms charge prices above marginal costs. Put differently, a single confiscatory tax on land rents generates enough revenue to implement the first-best allocation. 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, the markups are independent of the city size. Second, the market provides optimum product diversity at an efficient scale of production (see Dixit and Stiglitz, 1977, Section I).

To derive more general results under monopolistic competition, we must depart from the CES model. Doing so is especially relevant for a better understanding as to when the HGT holds in the second best. 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. 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 in a second-best economy 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 in 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 monopolistic competition model with variable elasticity of substitution (henceforth, VES) by Behrens and Murata (2006a) to a monocentric city setting. Second, using this framework, we investigate the relationships among the city size, the equilibrium markups and whether or not the HGT holds in the second best, either exactly or approximately. Last, building on the result that the HGT does not hold in our VES model in the second best, we derive more general 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 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 in our VES model in the second best. We finally prove, within our framework, that the HGT holds in the second best if and only if: (i) the second-best allocation is first-best efficient, which turns out to be equivalent to the CES case; or (ii) a marginal change in the city size has no impact on equilibrium product diversity at the second best. Since most VES models are not likely to satisfy the peculiar condition (ii), the HGT will generally not hold in the second best.

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 city size. Section 4 then deals with the optimal city size in the first-best and in the second-best cases. In Section 5, we establish conditions under which the HGT holds in the second best. Finally, Section 6 concludes.

2 The model

2.1 Preferences and land rents

Consider a monocentric city endowed 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

$$s(x) = 1 - 2\theta|x|. \quad (1)$$

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 *total effective labor supply* at the CBD is given by

$$S = \int_{-L/2}^{L/2} s(x)dx = L \left(1 - \frac{\theta L}{2}\right). \quad (2)$$

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:

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

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

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 profits and 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):

$$\max_{q(i), i \in \Omega} U \equiv \int_{\Omega} u(q(i))di \quad (3)$$

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

where $E \equiv (1 - \theta L)w + (\text{ALR} + \Pi)/L$ 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

$$\text{of the CES type: } u(q(i)) \equiv q(i)^{\frac{\sigma-1}{\sigma}}, \quad \sigma > 1 \quad (5)$$

$$\text{or of the CARA type: } u(q(i)) \equiv 1 - e^{-\alpha q(i)}, \quad \alpha > 0, \quad (6)$$

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 (5):

$$q(i) = \frac{p(i)^{-\sigma}}{\int_{\Omega} p(j)^{1-\sigma} dj} E, \quad (7)$$

whereas, as shown by Behrens and Murata (2006a), the first-order conditions for an interior solution yield the following demand functions for the CARA case (6):

$$q(i) = \frac{E - \frac{1}{\alpha} \int_{\Omega} \ln \left(\frac{p(i)}{p(j)} \right) p(j) dj}{\int_{\Omega} p(j) dj}. \quad (8)$$

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

$$\epsilon(i) = -\frac{\partial q(i)}{\partial p(i)} \frac{p(i)}{q(i)} = \sigma \quad (9)$$

$$\epsilon(i) = -\frac{\partial q(i)}{\partial p(i)} \frac{p(i)}{q(i)} = \frac{1}{\alpha q(i)}. \quad (10)$$

As usual, the CES case (9) features a constant elasticity, whereas the CARA case (10) features an elasticity that falls with the quantity $q(i)$ consumed.

2.2 Technology

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 constant 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:

$$\pi(i) = Lq(i) [p(i) - cw] - Fw, \quad (11)$$

where $q(i)$ is given by (7) or by (8), depending on whether we focus on the CES or the CARA case. There is free entry in and exit from the industry, which implies that in equilibrium N is endogenously determined by the zero profit condition. Consequently, aggregate profits $\Pi \equiv \int_{\Omega} \pi(i) di$ vanish in the

free entry equilibrium, so that the expenditure E equals the wage income net of both commuting costs and land rent, plus the share of aggregate land rents:

$$E = (1 - \theta L)w + \frac{\text{ALR}}{L} = \left(1 - \frac{\theta L}{2}\right)w. \quad (12)$$

Firms maximize their profit (11) 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:

$$\frac{\partial \pi(i)}{\partial p(i)} = Lq(i) \left\{ 1 - \left[1 - \frac{cw}{p(i)} \right] \epsilon(i) \right\} = 0, \quad \forall i \in \Omega. \quad (13)$$

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 (8), $q(i)$ depends itself on two price aggregates in the latter case, condition (13) 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.

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, each firm maximizes profits, taking (w, E, N) as given, and a *price equilibrium* $\mathbf{p}$ is determined as a distribution of prices satisfying conditions (13). Second, given a price equilibrium, firms enter in and exit from the market until they earn zero profits. This yields an *equilibrium* that is compatible with a price equilibrium, zero profits, and labor market clearing. Hence, an equilibrium is a solution to the set of conditions (13) and the following two conditions:

$$Lq(i)[p(i) - cw] - Fw = 0, \quad \forall i \in \Omega, \quad (14)$$

$$\int_{\Omega} [cLq(i) + F] di = S, \quad (15)$$

where all prices and quantities are evaluated at a price equilibrium $\mathbf{p}$. Note, finally, that we need not choose a numeraire since the model is determined in real terms.

Let us start with the well-known CES case. Inserting (9) into (13), the unique price equilibrium is trivially symmetric and given as follows:

$$p(i) = p \equiv \frac{\sigma}{\sigma - 1}cw, \quad \forall i \in \Omega. \quad (16)$$

Turning to the CARA case, the price equilibrium is determined by inserting (10) into (13). Behrens and Murata (2006a) have shown that the price equilibrium is symmetric, unique, and given by

$$p(i) = p \equiv cw + \frac{\alpha E}{N}, \quad \forall i \in \Omega. \quad (17)$$

¹It 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. First, unlike in the CES case, the markup is increasing in expenditure in the CARA case. The reason is that, as shown by expression (10), 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. Second, in the CARA case, the markup falls with the mass of competing firms in the city, i.e., there are *pro-competitive effects*.

Given the symmetry of the price equilibrium in both cases, the demands are then also symmetric so that the profit of each firm can be expressed as:

$$\pi = Lq(p - cw) - Fw. \quad (18)$$

Using the consumer's budget constraint at the symmetric price equilibrium, i.e. $E = Npq$ with E given by (12), expression (18) can be rewritten as follows:

$$\pi = pq \left[L \left(1 - \frac{cNq}{1 - \theta L/2} \right) - \frac{FN}{1 - \theta L/2} \right].$$

Zero profits then imply that the quantities must be such that

$$q = \frac{1}{c} \left(\frac{1 - \theta L/2}{N} - \frac{F}{L} \right) = \frac{1}{cL} \left(\frac{S}{N} - F \right), \quad (19)$$

where the last equality takes into account the resource constraint (15) in the symmetric case. Note that q is always positive because $NF < L(1 - \theta L/2) = S$ must hold from the resource constraint when N firms operate in the city. Note also that expression (19) holds whenever prices are symmetric and firms earn zero profit. This property will prove useful when we subsequently compare the equilibrium and the optimum allocations.

Inserting $q = E/(Np)$ into the labor market clearing condition (15), and using the relationship $LE = Sw$, we obtain the free entry mass of firms as a function of the wage-price ratio as follows:

$$N = \frac{S}{F} \left(1 - c \frac{w}{p} \right). \quad (20)$$

3.1 CES case

The equilibrium mass of firms can be determined by substituting (16) into (20). Its expression is given by:

$$N^* = \underbrace{\left(1 - \frac{\theta L}{2} \right)}_{=S/L} \frac{L}{\sigma F}. \quad (21)$$

It is worth noting that in the absence of commuting costs ($\theta = 0$), expression (21) reduces to the standard equilibrium mass of firms in the CES model, given by $L/(\sigma F)$. When there are commuting costs ($\theta > 0$), this equilibrium mass is reduced by a factor $S/L < 1$, 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 the 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 (19), yields the following indirect utility:

$$U(N) = N \left[\frac{1}{cL} \left(\frac{S}{N} - F \right) \right]^{\frac{\sigma-1}{\sigma}}. \quad (22)$$

Finally, inserting (21) into (22) yields

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

where $\kappa \equiv (F\sigma)^{-1} [c^{-1}F(\sigma - 1)]^{(\sigma-1)/\sigma} > 0$ 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 (17) into (20) and by using (12). Its expression is given by:²

$$N^* = \underbrace{\left(1 - \frac{\theta L}{2} \right)}_{=S/L} \frac{D(L) - \alpha F}{2cF}, \quad \text{where} \quad D(L) \equiv \sqrt{4\alpha cFL + (\alpha F)^2}. \quad (23)$$

It is worth pointing out that in the absence of commuting costs ($\theta = 0$), expression (23) reduces to the equilibrium mass of firms in Behrens and Murata (2006a). 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 is no longer always increasing in the city size. There exists indeed a unique threshold $\bar{L}$, given by

$$\bar{L} \equiv \frac{6c - \alpha F\theta + \sqrt{6\alpha cF\theta + (\alpha F\theta)^2}}{9c\theta} \in \left(0, \frac{1}{\theta} \right), \quad (24)$$

such that $\partial N^*/\partial L \gtrless 0$ for all $L \gtrless \bar{L}$. Put differently, when the city is large enough, an additional increase in city size may reduce the equilibrium mass of firms. The intuition for this result is that, as can be seen from (23), 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 (24), the former effect dominates the latter and the equilibrium mass of firms falls. Note that $\bar{L}$ increases in α and F , and decreases in θ . Hence, when firms have more monopoly power (α) and there are larger fixed costs (F), an increase in city size increases N^* over a larger range of city sizes; whereas higher commuting costs (θ) work in the opposite direction.

Let us summarize our findings as follows:

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

Proposition 1 (CARA equilibrium mass of firms) *The equilibrium mass of firms is increasing (resp., decreasing) in the city size when L is smaller (resp., larger) than the threshold $\bar{L}$, given by (24).*

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 simple calculation using (12), (17), and (23) shows that

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

which 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

$$Q^* = \frac{1}{c} \left[\frac{S}{N^*} - F \right] = \frac{1}{c} \left[\frac{2cFL}{D(L) - \alpha F} - F \right]. \quad (25)$$

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 (23) and (25), show that $\partial(N^*Q^*)/\partial L > 0$, $\lim_{L \rightarrow 0} \partial^2(N^*Q^*)/\partial L^2 > 0$ and $\lim_{L \rightarrow 1/\theta} \partial^2(N^*Q^*)/\partial L^2 < 0$. Hence, since $\partial^3(N^*Q^*)/\partial L^3 < 0$ there exists, by continuity, a unique threshold $\tilde{L}$ such that Q is a convex function of L for $L < \tilde{L}$ and a concave function for $L > \tilde{L}$. ■

Note that an increase in the 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 the city size.

Evaluating (3) for the CARA case, using (19), yields the following indirect utility:

$$U(N) = N \left[1 - e^{-\frac{\alpha}{cL} \left(\frac{S}{N} - F \right)} \right]. \quad (26)$$

Finally, inserting (23) into (26) yields

$$U(L) = \left(1 - \frac{\theta L}{2}\right) \frac{D(L) - \alpha F}{2cF} \left[1 - e^{-\frac{2\alpha F}{D(L) + \alpha F}}\right]. \quad (27)$$

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 (27).*

Proof. See Appendix A. ■

3.3 Equilibrium city size

The equilibrium city size can be determined in the CES and the CARA cases as follows. Let $\bar{U}$ and $U^s \equiv \max_L U(L)$ denote the (exogenously given) national utility level and the second-best utility level in the city, respectively. Provided that there is free mobility of agents between cities, the equilibrium city sizes L then must satisfy $U(L) = \bar{U}$. Since U is strictly concave and single peaked on $(0, 1/\theta)$ for all admissible parameter values, three cases may arise: (i) at $U^s = \bar{U}$, there exists a single equilibrium city size; (ii) at $U^s > \bar{U}$, there exist two equilibrium city sizes, one being stable and the other being unstable; and (iii) at $U^s < \bar{U}$, an equilibrium city size does not exist. In what follows, we focus on the meaningful case where $U^s \geq \bar{U}$. The stable equilibrium city size, when the national utility level is $\bar{U}$, is then given by:³

$$L^* = [L^s, 1/\theta) \cap \{L \mid U(L) = \bar{U}\}, \quad \text{where } L^s = \operatorname{argmax}_L U(L).$$

4 Optimal city size

We now investigate the optimal city size by assuming that the planner is free to choose L (using, e.g., migration policies or urban growth controls). In what follows, we superscript first-best values with f and second-best values with s .

4.1 First best

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

$$\mathcal{L} = Nu(q) + \lambda \left[\left(1 - \frac{\theta L}{2}\right) - cNq - \frac{NF}{L} \right],$$

³Because U is single-peaked, an equilibrium with $L < L^s$ is unstable as it will trigger at least some migration. When all cities are of equal size $L \geq L^s$, no agent can increase his utility by migrating to another city and the equilibrium is stable.

where λ denotes the Lagrange multiplier. A first-best solution q^f , N^f , L^f and λ^f satisfies the following first-order conditions:

$$u'(q^f) - \lambda^f c = 0 \quad (28)$$

$$u(q^f) - \lambda^f \left(cq^f + \frac{F}{L^f} \right) = 0 \quad (29)$$

$$\lambda^f \left[-\frac{\theta}{2} + \frac{N^f F}{(L^f)^2} \right] = 0 \quad (30)$$

$$\left(1 - \frac{\theta L^f}{2} \right) - cN^f q^f - \frac{N^f F}{L^f} = 0. \quad (31)$$

Note that u and u' depend on whether we consider the CES or the CARA case. Let us investigate these two cases in order.

CES case. Conditions (28) and (29) are given as follows:

$$\begin{aligned} \frac{\sigma-1}{\sigma} (q^f)^{-\frac{1}{\sigma}} - \lambda^f c &= 0 \\ (q^f)^{\frac{\sigma-1}{\sigma}} - \lambda^f \left(cq^f + \frac{F}{L^f} \right) &= 0. \end{aligned}$$

Solving these conditions, together with (30) and (31), we readily obtain the unique first-best city size:

$$L^f = \frac{2}{\theta(\sigma+1)}, \quad (32)$$

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). Finally, inserting (32) into (30), one can check that $N^f = 2/[\theta F(\sigma+1)^2]$.

CARA case. Conditions (28) and (29) are given as follows:

$$\alpha e^{-\alpha q^f} - \lambda^f c = 0 \quad (33)$$

$$1 - e^{-\alpha q^f} - \lambda^f \left(cq^f + \frac{F}{L^f} \right) = 0. \quad (34)$$

Unlike in the CES case, these two conditions, together with (30) and (31), form a system of equations that cannot be solved analytically. Yet, it is straightforward to show 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 takes the equilibrium values of N and q (which already encompass the resource constraint) as given, and chooses L to maximize the utility of a representative agent. Because the market outcome is symmetric, this amounts to solving the following optimization problem:

$$\max_L U(L) = N^*(L)u(q^*(L)),$$

where N^* and q^* are the equilibrium mass of firms and the equilibrium quantity, respectively. As shown in Section 3, both in the CES and in the CARA case the indirect utility at the market outcome is strictly concave and single-peaked for $L \in (0, 1/\theta)$. Omitting arguments of functions for notational simplicity, the necessary (and sufficient) first-order condition for the second-best city size can be rewritten as

$$\frac{\partial U}{\partial L} \frac{L}{U} = \frac{\partial N^*}{\partial L} \frac{L}{N^*} + \frac{\partial u}{\partial q} \frac{q^*}{u} \left(\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} - 1 \right) = 0. \quad (35)$$

CES case. It is readily verified that

$$\frac{\partial N^*}{\partial L} \frac{L}{N^*} = \frac{1 - \theta L}{1 - \theta L/2}, \quad \frac{\partial u}{\partial q} \frac{q^*}{u} = \frac{\sigma - 1}{\sigma} \quad \text{and} \quad \frac{\partial Q^*}{\partial L} \frac{L}{Q^*} = 0, \quad (36)$$

so that the second-best city size is given by (see, e.g., Duranton and Puga, 2001, Lemma 2):

$$L^s = \frac{2}{\theta(1 + \sigma)}. \quad (37)$$

Comparing (32) with (37) we see that the second-best city size is equal to the first-best city size. Furthermore, one can check that $N^f = N^s \equiv N^*(L^s)$, which implies that $q^f = q^s \equiv q^*(L^s)$ by the resource constraint. When taken together, these conditions reveal that *the second-best allocation is first-best efficient*.

CARA case. It is readily verified that

$$\frac{\partial N^*}{\partial L} \frac{L}{N^*} = \frac{1}{2} \left[3 + \frac{\alpha F}{D(L)} - \frac{4}{2 - \theta L} \right] \quad (38)$$

$$\frac{\partial u}{\partial q} \frac{q^*}{u} = \frac{2\alpha F}{\alpha F + D(L)} \frac{e^{-\frac{2\alpha F}{\alpha F + D(L)}}}{1 - e^{-\frac{2\alpha F}{\alpha F + D(L)}}} > 0 \quad (39)$$

$$\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} = \frac{1}{2} \left[1 + \frac{\alpha F}{D(L)} \right] \in (0, 1). \quad (40)$$

It thus turns out from (35) that $(\partial N^*/\partial L)(L/N^*)$ must be positive at the second-best, because $(\partial u/\partial q)(q^*/u)[(\partial Q^*/\partial L)(L/Q^*) - 1]$ is negative when evaluated at that allocation. Hence, from Proposition 1, we have the following result:

Proposition 5 (second-best city size) *The second-best city size in the CARA model is such that $L^s < \bar{L}$, where $\bar{L}$ is given by (24).*

As in the first best, we cannot derive the second-best city size explicitly because condition (35), when evaluated at expressions (38)–(40), cannot be solved analytically for L . Yet, because the second-best city size exists and is uniquely determined, the model can readily be solved numerically. We exploit this property in the subsequent analysis to illustrate several interesting features.

5 Henry George theorem in the second best

As shown by Kanemoto *et al.* (1996) in a framework with technological externalities, and as discussed by Arnott (2004), when the Henry George theorem holds, it can be used to test whether cities are too big or too small.⁴ It is therefore important to know when the HGT holds. As is well known, it holds in the first best. Whether or not the HGT holds in the second best generally depends on the nature of externalities. Unfortunately, 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).

Despite this difficulty, we now prove, within our framework, that the HGT holds in second-best economies under monopolistic competition if and only if: (i) the second-best allocation is first-best efficient, which is shown to be equivalent to the CES case; or (ii) a marginal change in the city size has no impact on equilibrium product diversity at the second best. Recall that the increasing returns to scale version of the Henry George theorem states that, at the optimal city size, aggregate land rents equal aggregate fixed costs in the increasing returns sector (e.g., Abdel-Rahman and Fujita, 1990; Helsley and Strange, 1990).⁵ Using our notation, the HGT holds in the first best because

$$\frac{\theta(L^f)^2}{2} = N^f F, \quad (41)$$

whereas it holds in the second best if and only if

$$\frac{\theta(L^s)^2}{2} = N^*(L^s)F. \quad (42)$$

In both cases, the left-hand side stands for aggregate land rents, and the right-hand side stands for aggregate fixed costs. Using condition (35), which must hold in the second best, we can show the following result:

⁴Kanemoto *et al.* (1996) apply the theorem to examine whether Japanese cities, in particular, Tokyo, are too big or too small. To do so, they compare the ratio of aggregate land values to the aggregate Pigouvian subsidy (i.e., a measure of external economies) across cities.

⁵Strictly 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. Note that we use the increasing returns to scale version of the Henry George theorem 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 subsidy required for agglomeration economies equals the aggregate fixed costs.

Proposition 6 (HGT in the second best) *Let $(L^s, N^*(L^s), q^*(L^s))$ be a second-best allocation. Then the Henry George theorem holds in the second best if and only if the second-best allocation is first-best efficient or $\partial N^*/\partial L|_{L=L^s} = 0$.*

Proof. See Appendix C. ■

Note that the case $\partial N^*/\partial L|_{L=L^s} = 0$ can be ruled out in the CES model since $\partial N^*/\partial L > 0$ for all values of L . It can also be ruled out in the CARA model because, as shown by Proposition 5, $\partial N^*/\partial L > 0$ at the second best.

CES case. Proposition 6 shows that, when $\partial N^*/\partial L|_{L=L^s} \neq 0$, the HGT holds in the second best if and only if the second best is first-best efficient. This is true in the CES case and, therefore, the HGT holds even in the second-best economy (Abdel-Rahman and Fujita, 1990). To check this, note that conditions (35) and (42) must be *simultaneously* satisfied for the HGT to hold in the second best. Whenever (42) is satisfied in the CES case, expression (35) holds regardless of the value of L because of (36), $N^*F = S/\sigma$ and $(\partial u/\partial q)(q^*/u) = (\sigma - 1)/\sigma$. This is not in general the case for non-CES functions. More concretely, assuming that $\partial N^*/\partial L|_{L=L^s} \neq 0$, we can show that *the Henry George theorem holds in a second-best economy if and only if the sub-utility is of the CES type*.

Proposition 7 (equivalence) *Let $(L^s, N^*(L^s), q^*(L^s))$ be a second-best allocation and assume that $\partial N^*/\partial L|_{L=L^s} \neq 0$. Then the Henry George theorem holds in the second best if and only if the sub-utility is of the CES type.*

Proof. See Appendix D. ■

CARA case. The CARA first-best allocation differs in general from the second-best allocation due to the presence of pro-competitive effects and therefore, the HGT will not hold. To see this, note again that, for the HGT to hold in the second best, conditions (35) and (42) must be *simultaneously* satisfied. Substituting (38)–(40) into (35) and using (42) yield a system of two equations with one unknown for which generically no solution L^s exists.

[Insert Figures 2, 3 and 4 about here]

To derive additional insights, we compare the first and the second best using numerical examples. Figures 2 and 3 depict the differences between the first and the second best as the value of commuting costs θ varies. First, Figure 2 plots the ratio L^s/L^f and shows that the second-best city size exceeds the first-best city size. Second, Figure 3 plots the ratio $N^*(L^s)/N^f$. As one can see, the second-best mass of firms exceeds the first-best one and the ratio rises as θ falls. Put differently, the second-best city is larger and also has a larger mass of firms. Figure 4 depicts the difference between aggregate land rents and aggregate fixed costs as a function of the commuting costs θ .

In the first best, the HGT holds, i.e., $\theta(L^f)^2/2 = N^f F$. As shown in Appendix E, there is excess entry in equilibrium due to pro-competitive effects for any given city size, so that $N^*(L^f)F > N^f F$.

Therefore, $N^*(L^f)F > \theta(L^f)^2/2$ must hold. In words, holding the city size constant, when firms' entry is unrestricted, aggregate fixed costs exceed aggregate land rents, so that the HGT fails to hold. Note that this is true even when the city size is endogenously determined in the second best, as can be seen from Figure 4. Finally, one can see that the *gap increases as θ decreases*, which is in accord with the numerical results of Helsley and Strange (1990, p.209).

[Insert Figure 5 about here]

Figure 5 summarizes the difference between the CES and the CARA cases. Let $N^f(L)$ be the solution of equations (28), (29) and (31). Note first that $N^*(L) = N^f(L)$ for all L in the CES case, i.e., the two curves coincide. Hence, if $ALR/w > N^f F$ (resp., $ALR/w < N^f F$) at the equilibrium city size, we know that the city is too big (resp., too small). However, this does not apply to the CARA case, as $N^*(L) > N^f(L)$ because of excess entry for any given L (see Appendix E). Furthermore, comparing $N^f(L)F$ and ALR/w at the second-best city size provides no information on whether the city is too big or too small. As pointed out by Kanemoto *et al.* (1996), their method may therefore not be directly applicable to the second-best case in which price distortions cause excess (or insufficient) entry of firms.

6 Conclusion

We have analyzed the equilibrium and the optimal resource allocations in a monocentric city under monopolistic competition focusing on the CES and the CARA cases. Summarizing our key insights, we have shown that a larger city has lower equilibrium markups in the CARA model. Furthermore, we have shown that the HGT does not hold in the CARA model at the second best: aggregate fixed costs exceed aggregate land rents because of excess entry triggered by pro-competitive effects. Because the gap between aggregate land rents and aggregate fixed costs increases as 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: (i) the second-best allocation is first-best efficient, which turns out to be equivalent to the CES case; or (ii) a marginal change in the city size has no impact on equilibrium product diversity at the second best. Since most VES models are not likely to satisfy the peculiar condition (ii), the HGT will generally not hold in second-best economies under monopolistic competition with variable elasticities.

Acknowledgements. We thank Yoshitsugu Kanemoto, Giordano Mion, Pierre M. Picard and Jacques Thisse for helpful comments and suggestions. Kristian Behrens gratefully acknowledges financial support from the European Commission under the Marie Curie Fellowship MEIF-CT-2005-024266. Yasusada Murata gratefully acknowledges financial support from Japan Society for the Promotion of Science (17730165). Part of this paper was written while Yasusada Murata was visiting CORE (Université catholique de Louvain, Belgium), and while Kristian Behrens was visiting ARISH

(Nihon University, Japan). We gratefully acknowledge the hospitality of these institutions. The usual disclaimer applies.

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. (2004) Does the Henry George Theorem provide a practical guide to optimal city size?, *American Journal of Economics and Sociology* 63, 1057-1090.
- [3] 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.
- [4] Behrens, K. and Y. Murata (2006a) General equilibrium models of monopolistic competition: a new approach, *Journal of Economic Theory*, forthcoming (doi: 10.1016/j.jet.2006.10.001).
- [5] Behrens, K. and Y. Murata (2006b) Gains from trade and efficiency under monopolistic competition: a variable elasticity case, CORE Discussion Paper #2006/49, Univ. catholique de Louvain.
- [6] Benassy, J.-P. (1996) Taste for variety and optimum production patterns in monopolistic competition, *Economics Letters* 52, 41-47.
- [7] Campbell, J.R. and H.A. Hopenhayn (2005) Market size matters, *Journal of Industrial Economics* 53, 1-25.
- [8] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297-308.
- [9] Duranton, G. and D. Puga (2001) Nursery cities: urban diversity, process innovation, and the life cycle of products, *American Economic Review* 91, 1454-1477.
- [10] 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.
- [11] Ethier, W.J. (1982) National and international returns to scale in the modern theory of international trade, *American Economic Review* 72, 389-405.
- [12] Flatters, F., J.V. Henderson and P. Mieszkowski (1974) Public goods, efficiency, and regional fiscal equalization, *Journal of Public Economics* 3, 99-112.

- [13] 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.
- [14] 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.
- [15] Holmes, T.J. (1999) Scale of local production and city size, *American Economic Review Papers and Proceedings* 89, 317-320.
- [16] Kanemoto, Y. (2007) The Henry George theorem in a second best economy: a note, University of Tokyo, mimeographed.
- [17] 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.
- [18] Mankiw, N.G. and M.D. Whinston (1986) Free entry and social inefficiency, *RAND Journal of Economics* 17, 48-58.
- [19] Murata, Y. and J.-F. Thisse (2005) A simple model of economic geography à la Helpman-Tabuchi, *Journal of Urban Economics* 58, 137-155.
- [20] 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.
- [21] Syverson, C. (2004) Market structure and productivity: a concrete example, *Journal of Political Economy* 112, 1181-1222.
- [22] Vives, X. (1999) *Oligopoly Pricing: Old Ideas and New Tools*. Cambridge, MA: MIT Press.

Appendix A: Proof of Proposition 4

We first show that (27) is a strictly concave function of the population size L for all admissible parameter values. Applying (17) and $E = (S/L)w$ to $q = E/Np$, we get $q = 1/[\alpha + c\tilde{N}(L)]$, where $\tilde{N}(L) \equiv (L/S)N^* = [D(L) - \alpha F]/2cF$. Let

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

Note that $\tilde{U}$ is strictly increasing and strictly concave in L because $\tilde{U}$ is a strictly increasing and strictly concave function of $\tilde{N}$, 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

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

and

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

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 (27) on $(0, 1/\theta)$.

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

Taking the ratio of (33) and (34), we have

$$\frac{\alpha e^{-\alpha q^f}}{1 - e^{-\alpha q^f}} = \frac{c}{c q^f + (F/L^f)}. \quad (\text{B.1})$$

From (30) and (31), we get

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

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

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

Differentiating LHS with respect to q^f , we have

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

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

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

which ensure the uniqueness of q^f . Finally, from expressions (30) and (B.2), we have the uniqueness of the first-best city size as well as of the mass of firms.

Appendix C: Proof of Proposition 6

Using expression (19), one can check that

$$\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} = \frac{S}{S - N^* F} \left(\frac{\partial S}{\partial L} \frac{L}{S} - \frac{\partial N^*}{\partial L} \frac{L}{N^*} \right).$$

Condition (35) can then be rewritten as follows:

$$\frac{\partial N^*}{\partial L} \frac{L}{N^*} \left[1 - \left(\frac{S}{S - N^* F} \right) \frac{\partial u}{\partial q} \frac{q^*}{u} \right] - \frac{\partial u}{\partial q} \frac{q^*}{u} \left[1 - \left(\frac{S}{S - N^* F} \right) \frac{\partial S}{\partial L} \frac{L}{S} \right] = 0. \quad (\text{C.1})$$

Assume now that the HGT holds in the second best. Whenever the HGT holds in the second best, the second term of (C.1) vanishes because

$$\left(\frac{S}{S - N^*(L^s)F} \right) \frac{\partial S}{\partial L} \frac{L}{S} \Big|_{L=L^s, N^*(L^s)F = \frac{\theta(L^s)^2}{2}} = 1.$$

Therefore, either $\partial N^*/\partial L|_{L=L^s} = 0$ or

$$\frac{\partial u}{\partial q} \frac{q^*}{u} \Big|_{L=L^s} = \frac{S^s - N^*(L^s)F}{S^s}, \quad (\text{C.2})$$

must hold, where $S^s = L^s(1 - \theta L^s/2)$ is the second-best total effective labor supply.

We now show that the second-best allocation is first-best efficient in the latter case. Using the labor market clearing condition $N^*(L^s)(cq^s L^s + F) = S^s$, expression (C.2) can be rewritten as

$$\frac{\partial u}{\partial q} \frac{q^*}{u} = \frac{cq^s}{cq^s + F/L^s}.$$

Taking the ratio of (28) and (29), the first-best allocation satisfies $\lambda^f \neq 0$,

$$\frac{\partial u}{\partial q} \frac{q^f}{u} = \frac{cq^f}{cq^f + F/L^f}, \quad (\text{C.3})$$

(30) and (31). As shown in Appendix B, the first-best solution (satisfying (30), (31) and (C.3)) is unique. A second best allocation that satisfies the HGT also satisfies conditions (30), (31) and (C.3). We may, therefore, conclude that the second-best allocation is first-best efficient.

Conversely, it is readily verified that $\partial N^*/\partial L|_{L=L^s} = 0$ and (C.1) imply that the HGT holds in the second best. If the second-best allocation is first-best efficient, (30) must be satisfied, which implies that the HGT holds in the second best.

Appendix D: Proof of Proposition 7

Taking the rate of change in the resource constraint $N^*(cQ^* + F) = S$, we have

$$\frac{\partial N^*}{\partial L} \frac{L}{N^*} + \frac{cQ^*}{cQ^* + F} \frac{\partial Q^*}{\partial L} \frac{L}{Q^*} = \frac{\partial S}{\partial L} \frac{L}{S}. \quad (\text{D.1})$$

From (35) and (D.1), we have

$$-\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} \left(\frac{\partial u}{\partial q} \frac{q^*}{u} - \frac{cQ^*}{cQ^* + F} \right) + \left(\frac{\partial u}{\partial q} \frac{q^*}{u} - \frac{\partial S}{\partial L} \frac{L}{S} \right) = 0. \quad (\text{D.2})$$

Assume now that the HGT holds in the second best. Then, using (2) and (25) it is readily verified that

$$\frac{cQ^*}{cQ^* + F} \Big|_{N^*(L^s)F = \frac{\theta(L^s)^2}{2}} = \frac{\partial S}{\partial L} \frac{L}{S} \Big|_{L=L^s} = \frac{1 - \theta L^s}{1 - \theta L^s/2}. \quad (\text{D.3})$$

Expressions (D.2) and (D.3) then yield

$$\left(1 - \frac{\partial Q^*}{\partial L} \frac{L}{Q^*}\right) \left(\frac{\partial u}{\partial q} \frac{q^*}{u} - \frac{1 - \theta L^s}{1 - \theta L^s/2}\right) \Big|_{L=L^s} = 0.$$

Note that we can disregard the case where $(\partial Q^*/\partial L)(L/Q^*)|_{L=L^s} = 1$ because, in this case, we have $\partial N^*/\partial L|_{L=L^s} = 0$ by (35), which is excluded by the assumption of Proposition 7. Hence, we have

$$\frac{\partial u}{\partial q} \frac{q^*}{u} \Big|_{L=L^s} = \frac{\partial S}{\partial L} \frac{L}{S} \Big|_{L=L^s} = \frac{1 - \theta L^s}{1 - \theta L^s/2}. \quad (\text{D.4})$$

Next, take the rate of change in the budget constraint $N^*p^*q^* = (S/L)w^*$ (or $N^*p^*Q^* = Sw^*$) as follows:

$$\frac{\partial N^*}{\partial L} \frac{L}{N^*} + \left(\frac{\partial p^*}{\partial L} \frac{L}{p^*} - \frac{\partial w^*}{\partial L} \frac{L}{w^*}\right) + \frac{\partial Q^*}{\partial L} \frac{L}{Q^*} = \frac{\partial S}{\partial L} \frac{L}{S}. \quad (\text{D.5})$$

From (35) and (D.5), together with (D.4), we have

$$\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} \left(\frac{\partial u}{\partial q} \frac{q^*}{u} - 1\right) \Big|_{L=L^s} = \left(\frac{\partial p^*}{\partial L} \frac{L}{p^*} - \frac{\partial w^*}{\partial L} \frac{L}{w^*}\right) \Big|_{L=L^s}. \quad (\text{D.6})$$

Finally, take the rate of change in the zero profit condition $Q^*(p^* - cw^*) = Fw^*$ as follows:

$$\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} + \frac{p^*}{p^* - cw^*} \left(\frac{\partial p^*}{\partial L} \frac{L}{p^*} - \frac{\partial w^*}{\partial L} \frac{L}{w^*}\right) = 0. \quad (\text{D.7})$$

From (D.6) and (D.7), we obtain

$$\frac{\partial Q^*}{\partial L} \frac{L}{Q^*} \left(\frac{\partial u}{\partial q} \frac{q^*}{u} - \frac{cw^*}{p^*}\right) \Big|_{L=L^s} = 0.$$

Two cases may arise:

Case 1: $(\partial Q^*/\partial L)(L/Q^*)|_{L=L^s} = 0$.

In this case, expression (D.7) implies $(\partial p^*/\partial L)(L/p^*)|_{L=L^s} = (\partial w^*/\partial L)(L/w^*)|_{L=L^s}$, thus leading to constant markups, i.e., $p^* = \kappa cw^*$, where κ is a constant. From the first-order conditions of the Lagrangian $\int_{\Omega} u(q(i))di + \mu [E - \int_{\Omega} p(i)q(i)di]$, we have $u'(q(i)) = p(i) \int_{\Omega} q(j)u'(q(j))dj/E$. Differentiating this expression with respect to $p(i)$ we have:

$$-\frac{\partial q(i)}{\partial p(i)} \frac{p(i)}{q(i)} = -\frac{u'(q(i))}{q(i)u''(q(i))}, \quad (\text{D.8})$$

because of a continuum of firms. The profit-maximizing price is generally written as

$$p(i) = \left[\frac{(\partial q(i)/\partial p(i))(p(i)/q(i))}{(\partial q(i)/\partial p(i))(p(i)/q(i)) + 1} \right] cw. \quad (\text{D.9})$$

For the term in the square brackets to be a constant regardless of the value of $q(i)$, $(\partial q(i)/\partial p(i))(p(i)/q(i))$ and hence $q(i)u''(q(i))/u'(q(i))$ must be a constant, thus implying that u must be of the CES type.

Case 2: $(\partial u/\partial q)(q^*/u)|_{L=L^s} = cw^*/p^*|_{L=L^s}$.

In this case, the markup is given by $u(q(i))/[q(i)u'(q(i))]$. Setting this equal to the terms in the square brackets of (D.9), we have

$$-\frac{\partial q(i)}{\partial p(i)} \frac{p(i)}{q(i)} = \frac{u(q(i))}{u(q(i)) - q(i)u'(q(i))}. \quad (\text{D.10})$$

Hence, in this case (D.8) and (D.10) must be simultaneously satisfied, which implies $q(i)(u'(i))^2 - u(q(i))[u'(q(i)) + q(u)u''(q(i))] = 0$ on the one hand. On the other hand, differentiating the markup $u(q(i))/[q(i)u'(q(i))]$ with respect to $q(i)$ yields $\{q(i)(u'(i))^2 - u(q(i))[u'(q(i)) + q(u)u''(q(i))]\}/[q(i)u'(q(i))]^2$. These two results finally lead to a constant markup, thus implying that u must be of the CES type.

Conversely, if the sub-utility is of the CES type, plugging (21) and (37) into (42) shows that the HGT holds in the second best.

Appendix E: Excess entry for any given city size

In this appendix, we show that excess entry occurs in the CARA case. Unlike in the basic CES model, this arises because of pro-competitive effects.⁶ To establish the result, we determine the first-best mass of firms for any *given* value of L . The planner maximizes the utility of a representative agent, as given by (3), subject to the technology and resource constraint (15). The first-order conditions with respect to $q(i)$ for this problem show that the quantities must be symmetric. This result, together with (15), implies that they are given by $q = (1/cL)(S/N - F)$. Plugging this expression into the utility function, the planner chooses

$$N^f(L) = \operatorname{argmax}_N N \left[1 - e^{-\frac{\alpha}{cL} \left(\frac{S}{N} - F \right)} \right],$$

taking L as given. Standard calculations show that

$$\frac{\partial U}{\partial N} = 1 - \left(1 + \frac{\alpha S}{cNL} \right) e^{-\frac{\alpha}{cL} \left(\frac{S}{N} - F \right)} \quad (\text{E.1})$$

and $\partial^2 U / \partial (N)^2 < 0$, i.e., U is a strictly concave function of N which has a unique maximum. Equating (E.1) to zero, letting $\tilde{\alpha} \equiv (S/L)\alpha$, and rearranging, the optimal mass of firms solves the following first-order condition:

$$\frac{cN^f(L)}{\tilde{\alpha} + cN^f(L)} = e^{-\frac{\tilde{\alpha}}{c} \left(\frac{1}{N^f(L)} - \frac{F}{S} \right)}.$$

We can then prove the following result:

⁶In 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. See Mankiw and Whinston (1986) and Vives (1999) for more general results on excess entry.

Proposition 8 (CARA excess entry) *For any given city size, there exists a unique optimal mass of firms $N^f(L)$ such that the equilibrium mass $N^*(L)$ exceeds the optimal mass ($N^*(L) > N^f(L)$). Hence, there are too many firms operating at an inefficiently small scale.*

Proof. Let $\tilde{N}$ stand for the equilibrium mass of firms when $\theta = 0$. Behrens and Murata (2006b, Proposition 1) have shown that

$$\frac{c\tilde{N}}{\alpha + c\tilde{N}} < e^{-\frac{\alpha}{c}(\frac{1}{\tilde{N}} - \frac{F}{L})} \quad (\text{E.2})$$

holds for all values of $\alpha > 0$, $c > 0$, $F > 0$ and $L > 0$, which implies that there is excess entry in the special case where $S = L$ (i.e., $\theta = 0$). Excess entry when $\theta > 0$ occurs when

$$\frac{cN^*(L)}{\tilde{\alpha} + cN^*(L)} < e^{-\frac{\tilde{\alpha}}{c}(\frac{1}{N^*(L)} - \frac{F}{S})}. \quad (\text{E.3})$$

Since $N^*(L) = (S/L)\tilde{N}$, condition (E.3) reduces to (E.2), which yields the result. ■

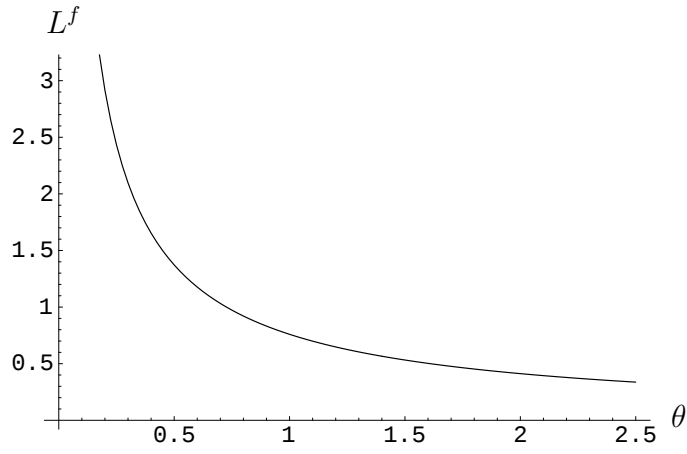


Figure 1: First-best city size and commuting costs ($\alpha = 1.2$, $c = 0.3$, $F = 0.5$)

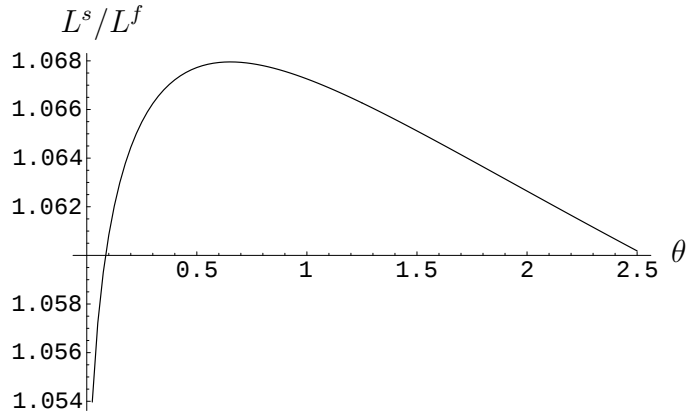


Figure 2: Second-best to first-best city size ratio ($\alpha = 1.2$, $c = 0.3$, $F = 0.5$)

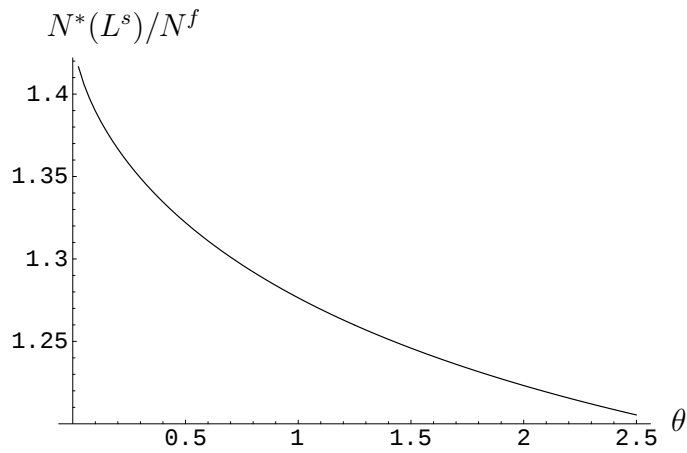


Figure 3: Second-best to first-best mass of firms ($\alpha = 1.2$, $c = 0.3$, $F = 0.5$)

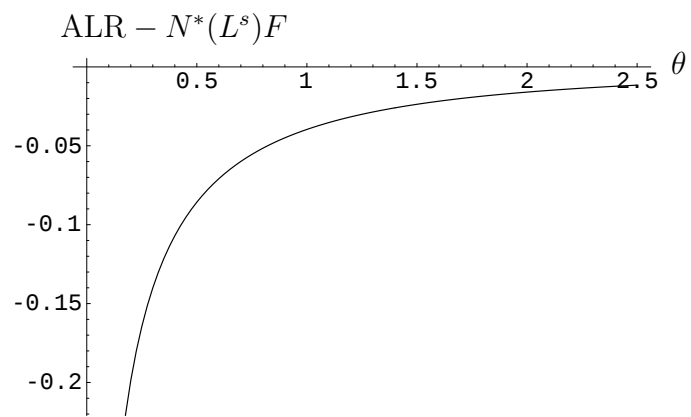


Figure 4: ALR and aggregate fixed costs ($\alpha = 1.2$, $c = 0.3$, $F = 0.5$)

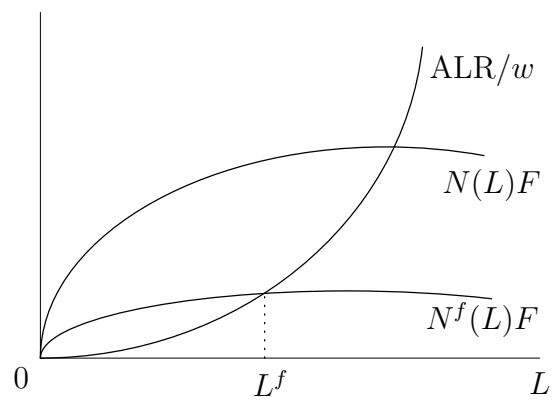


Figure 5: Excess entry and the Henry George theorem