

CIRJE-F-773

## **The Henry George Theorem in A Second-Best World**

Kristian Behrens

Université du Québec à Montréal, CIRPÉE and CEPR

Yoshitsugu Kanemoto

University of Tokyo

Yasusada Murata

Nihon University

November 2010

CIRJE Discussion Papers can be downloaded without charge from:

<http://www.e.u-tokyo.ac.jp/cirje/research/03research02dp.html>

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

# The Henry George Theorem in a second-best world\*

Kristian Behrens<sup>†</sup>      Yoshitsugu Kanemoto<sup>‡</sup>      Yasusada Murata<sup>§</sup>

November 5, 2010

## Abstract

The Henry George Theorem (HGT), or the golden rule of local public finance, states that, in first-best economies, the fiscal surplus, defined as aggregate land rents minus aggregate losses from increasing returns to scale activities, is zero at optimal city sizes. We derive a general second-best HGT in which the fiscal surplus equals the excess burden, expressed as an extended Harberger formula. We then apply our theorem to various settings encompassing urban economics, the new economic geography and local public finance to investigate whether or not a single tax on land rents can raise enough revenue to cover aggregate losses from increasing returns to scale activities.

**Keywords:** Henry George Theorem; second-best economies; optimal city size; monopolistic competition; local public goods

**JEL Classification:** D43; R12; R13

---

\*We thank Richard Arnott and Tom Holmes, as well as participants at the Nagoya Workshop for Spatial Economics and International Trade and at the 56th Meetings of the RSAI in San Francisco for valuable comments and suggestions. Behrens is holder of the *Canada Research Chair in Regional Impacts of Globalization*. Financial support from the CRC Program of the Social Sciences and Humanities Research Council (SSHRC) of Canada is gratefully acknowledged. Behrens further gratefully acknowledges financial support from FQRSC Québec (Grant NP-127178). Kanemoto gratefully acknowledges financial support from KAKENHI (18330044). Murata gratefully acknowledges financial support from KAKENHI (17730165) and MEXT.ACADEMIC FRONTIER (2006-2010). Any remaining errors are ours.

<sup>†</sup>Canada Research Chair, Département des Sciences Economiques, Université du Québec à Montréal (UQAM), Canada; CIRPÉE, Canada; and CEPR. E-mail: [behrens.kristian@uqam.ca](mailto:behrens.kristian@uqam.ca)

<sup>‡</sup>Graduate School of Economics and Graduate School of Public Policy (GraSPP), University of Tokyo, Japan; National Graduate Institute for Policy Studies, Tokyo, Japan. E-mail: [kanemoto@e.u-tokyo.ac.jp](mailto:kanemoto@e.u-tokyo.ac.jp)

<sup>§</sup>Advanced Research Institute for the Sciences and Humanities (ARISH), Nihon University, Japan; and Nihon University Population Research Institute (NUPRI), Japan. E-mail: [murata.yasusada@nihon-u.ac.jp](mailto:murata.yasusada@nihon-u.ac.jp)

# 1 Introduction

The equilibrium sizes of agglomerations such as cities (or communities and shopping centers), are determined by the balance between increasing and decreasing returns to spatial concentration. As is well known, cities need not be optimally sized at equilibrium as the urban environment is replete with externalities. Recent empirical research suggests that any departure from optimal city size can generate sizeable economic costs, especially in terms of foregone productivity (e.g., Au and Henderson, 2006a, 2006b). Hence, elaborating efficient urban growth policies is likely to be of first-order importance to many countries, especially developing ones. An important preliminary step for devising such urban policies is to assess whether cities are too large or too small, and by how much. The ‘golden rule’ of local public finance (Flatters *et al.*, 1974), i.e., the Henry George Theorem (henceforth HGT; Stiglitz, 1977; Arnott and Stiglitz, 1979) provides a condition for the optimal size of a city or, equivalently, the optimal number of cities given a fixed total population. It is thus a potentially useful tool that can allow policy makers to assess whether cities are too large or too small.<sup>1</sup> Another application of the HGT is concerned with smaller agglomerations such as shopping centers and business subcenters that are often developed by private companies. Those developers capture increases in land prices to finance development costs. The HGT shows that free entry of developers yields an efficient allocation in a first-best world. It would be of interest for policy makers to know whether developments are too few or too many in a more realistic setting with various distortions.

The HGT may be viewed as an extension of the result on the optimal number of firms in an industry — entry into an industry is optimal when the marginal social benefit of the last entrant vanishes. In a spatial context, this optimality condition must be extended to include land rents: aggregate land rents (which capitalize agglomeration benefits) equal the aggregate losses from increasing returns activities that generate agglomeration. For example, if the driving force for agglomeration is a pure local public good, then the cost of its supply must be equal to aggregate land rents at the optimal city size. In the case of a factory town, where firm-level scale economies generate spatial concentration, aggregate land rents must be equal to the losses that the firm would incur if it were constrained to price at marginal cost. In a new economic geography (henceforth NEG) model with increasing returns and product differentiation, aggregate land rents must be equal to the subsidies paid to firms in order to achieve efficient production scale and optimum product diversity.

As is well known, the HGT holds in a first-best world without distortions but not necessarily in a second-best world (see Arnott, 2004, for a recent survey). However, as highlighted by the latter two foregoing examples, most factors that drive the concentration of economic activity involve some form of market failure. Thus, for the HGT to be of practical relevance it must be extended to cope

---

<sup>1</sup>Kanemoto *et al.* (1996, 2005) applied these ideas to empirically test the often-made claim that Tokyo, with a metropolitan population of about 30 million, is much too large. It is not easy to obtain reliable estimates of key variables such as the aggregate land rents and the aggregate Pigouvian subsidies in a city, but the HGT provides a promising theoretical framework for empirical studies. As stated by Arnott (2004, pp.1086–1087): “Does the Henry George Theorem provide a practical guide to optimal city size? The jury is not yet in, but the approach is sufficiently promising to merit further exploration.”

with settings encompassing distortions of various kinds. This has, to the best of our knowledge, not been systematically done to date. The purpose of this article is to fill that gap by identifying conditions under which the HGT holds even in a second-best economy — an economy where policy makers can implement the optimal size of an agglomeration but are constrained to take production and consumption decisions (including entry decisions of firms) and the resulting equilibrium prices as given. We also examine in which directions the theorem needs to be modified should it fail to hold in such a world.

Several related articles have examined variations on the HGT in a second-best world. First, Arnott (2004, p.1073) showed that “in distorted urbanized economies the generalized HGT continues to hold when the aggregate magnitudes are valued at shadow prices.” The article does not, however, suggest how these shadow prices can be calculated, thereby limiting the practical relevance of the generalized HGT. Second, Helsley and Strange (1990) showed that the theorem does not hold in a matching framework of urban labor markets where firms compete for workers on a circle of skill ranges and where wages are determined via Nash bargaining. Last, Behrens and Murata (2009) examined a monocentric city model with monopolistic competition and showed that the HGT holds at the second best if and only if the second-best allocation is first-best efficient. In their model, the latter holds true only in the constant elasticity of substitution (henceforth CES) case.<sup>2</sup>

As should be clear from the foregoing literature review, whether or not the HGT holds at the second best hinges on the chosen modeling framework. Unlike previous approaches that rely from the beginning on specific functional forms, our aim is to derive a more general version of a second-best HGT and apply it to various models of spatial concentration, including those of the NEG and of local public goods. Our methodology allows us to obtain general results without resorting to specific functional forms. Using the concept of *fiscal surplus*, defined as aggregate land rents minus aggregate losses from increasing returns, the HGT at the first best simply states that the fiscal surplus is zero when the cities are of optimal size (or, alternatively, when the number of cities is optimal). In a second-best world the fiscal surplus equals the *excess burden* created by increasing either city size or the number of cities. We show that the latter can be expressed as an extension of the *Harberger formula*, i.e., the weighted sum of induced changes in quantities and product diversity, with weights being the associated distortions. Our formula extends the original by adding distortions in product diversity that play an important role in monopolistic competition models.

To illustrate the usefulness of our results in a simple setting, we apply the second-best HGT to different NEG-type models where distortions originate only from the differentiated good sector. In these models, distortions take two forms: a price distortion for each variety of the differentiated good, and a distortion for the number of varieties consumed. The variety distortion works in the opposite direction of the price distortion. We show that when the utility function is additively separable with respect to varieties as in Dixit and Stiglitz (1977), the price distortion depends on the *relative risk aversion* (RRA) and the variety distortion is inversely related to the *scale elasticity of utility*. Unlike

---

<sup>2</sup>Many of the results obtained in the CES model (markups, firm size, entry etc.) are clearly knife-edge results. See Kokovin *et al.* (2010) for an overview and discussion.

the former, the latter depends on the absolute level of utility: adding a positive constant to the utility function makes the variety distortion larger. This result reveals an important difference between models with endogenous product diversity (like new trade and NEG models) and expected utility theory that has not been emphasized much until now. In expected utility theory, utility is unique up to an affine transformation and its absolute level does not matter. In monopolistic competition models, the desirability of introducing a new variety hinges on the magnitude of the utility increase it produces. The utility increase is the difference between the utility level with equilibrium consumption and that with zero consumption. Models with endogenous product diversity therefore depend on the absolute level of utility, whereas expected utility theory depends only on the marginal utility and higher-order derivatives.

Building on the foregoing results which hold for general functional forms, we turn to specific examples to derive sharper predictions. First, we examine a case where the excess burden from the variety distortion is influenced by changes in the intersectoral allocation. Using a setting *à la* Abdel-Rahman and Fujita (1990), we show that the excess burden (and hence the fiscal surplus) is positive at the second-best city size. Second, we consider a case without intersectoral distortions to focus solely on the differentiated good sector. Using the variable RRA setting of Ogaki and Zhang (2001), we show that the excess burden is positive if the RRA is increasing, and negative if it is decreasing. In the knife-edge case where the RRA is constant, the price and variety distortions just cancel out, thus yielding the result that the HGT holds even at the second best in a one-sector CES world (Duranton and Puga, 2001; Behrens and Murata, 2009).<sup>3</sup> However, this need not always be the case. For example, we show that in the case of the constant absolute risk aversion (CARA) model used by Behrens and Murata (2009), the excess burden is negative even though the RRA is increasing. In all cases, adding a constant term to the sub-utility tends to increase the excess burden, as noted above. Last, as the HGT has been mainly examined in and applied to models with local public goods (see, e.g., Ch.3 of Kanemoto, 1980), we examine a simple model with a local public good that is financed by a distortionary property tax on housing. In this setting, the excess burden is again negative.

The remainder of the paper is organized as follows. In Section 2, we develop a model of the optimal number of communities that is general enough to encompass urban models with externalities, NEG models, and local public goods models. Section 3 derives our second-best HGT using an extended Harberger formula. In Section 4, we focus on the price and variety distortions in monopolistic competition models with and without intersectoral distortions. We derive a more specific formula for the second-best HGT and illustrate the variable RRA and CARA cases. In Section 5, we turn to a local public goods model with a distortionary property tax. Finally, Section 6 concludes.

---

<sup>3</sup>Behrens *et al.* (2010) have recently shown that the HGT continues to hold in a CES model with heterogeneous agents and sorting along talent across cities. The reason is that, despite heterogeneity and sorting, the underlying tradeoff in terms of price and variety distortions is unaffected.

## 2 The model

The Henry George Theorem is typically derived by optimizing the number of cities (or, equivalently, the size of a city given the total population) in a model with identical cities and homogeneous consumers. In order to keep the analysis simple, we stick to these assumptions.<sup>4</sup> Our economy consists of  $n$  cities, each endowed with the same amounts of land and other immobile factors, available in fixed supply. The total population of the economy is fixed at  $\bar{N}$ . The size of a city is then given by  $N \equiv \bar{N}/n$ . We assume that there are no moving costs, so that all consumers achieve the same level of utility in equilibrium.

Let  $\mathcal{I} = \{0, 1, \dots, I\}$  denote the set of indices of goods in the economy, where good 0 is taken as the numeraire and other goods are divided into four types. The first type  $\mathcal{M}$  is differentiated goods (either consumer goods or intermediate goods). We denote the numbers (or the masses) of varieties for differentiated goods by  $\mathbf{m} = \{m_i\}_{i \in \mathcal{M}}$ . The second type  $\mathcal{G}$  is local public goods. Both  $\mathcal{M}$  and  $\mathcal{G}$  are non-traded between cities and generate agglomeration forces. The third type  $\mathcal{L}$  consists of land and other fixed factors whose supplies are given in each city.  $\mathcal{L}$  is also non-traded between cities but generates dispersion forces. We assume that adding a new city increases the supplies of these goods, and henceforth refer to them as ‘land’. The last type  $\mathcal{H}$  may or may not be traded, and includes all remaining goods as well as housing, which we call ‘other goods’. Hence  $\mathcal{I} = \{0\} \cup \mathcal{M} \cup \mathcal{G} \cup \mathcal{L} \cup \mathcal{H}$ .

We first consider the case where only  $\mathcal{M}$  goods generate agglomeration forces. The analysis of local public goods is relegated to Section 5. The utility function of a representative consumer is given by  $U(\mathbf{x}, \mathbf{m})$ , where  $\mathbf{x} \equiv (x_0, \{x_{ij}, j \in [0, m_i]\}_{i \in \mathcal{M}}, \{x_i\}_{i \in \mathcal{L}}, \{x_i\}_{i \in \mathcal{H}})$  is a vector of goods that can be divided into our four categories: the numeraire,  $x_0$ ; differentiated goods  $\{x_{ij}, j \in [0, m_i]\}_{i \in \mathcal{M}}$ ; land  $\{x_i\}_{i \in \mathcal{L}}$ ; and other goods  $\{x_i\}_{i \in \mathcal{H}}$ .<sup>5</sup> Note that  $\mathbf{x}$  may include intermediate goods, which always take the value zero for final consumers. Each differentiated good is provided as a continuum of distinct varieties, ranging from 0 to  $m_i$ . The masses of varieties  $\mathbf{m}$  can be either fixed or endogenously determined in the model. Later, we restrict ourselves to the symmetric case where, for any  $i \in \mathcal{M}$ ,  $x_{ij} = x_i$  for all  $j \in [0, m_i]$  in equilibrium. Finally, each consumer has an initial endowment vector  $\bar{\mathbf{x}}$  where, for any  $i \in \mathcal{M}$ ,  $\bar{x}_{ij} = 0$  for all  $j \in [0, m_i]$ .

### 2.1 Consumption

Let  $\mathbf{p} = (p_0, \{p_{ij}, j \in [0, m_i]\}_{i \in \mathcal{M}}, \{p_i\}_{i \in \mathcal{L}}, \{p_i\}_{i \in \mathcal{H}})$  be the vector of consumer prices and  $\bar{U}$  be the common equilibrium utility. The compensated demand functions  $x_i(\mathbf{p}, \mathbf{m}, \bar{U})$  and  $x_{ij}(\mathbf{p}, \mathbf{m}, \bar{U})$  are the solution to the expenditure minimization problem:

$$\min \sum_{i \notin \mathcal{M}} p_i x_i + \sum_{i \in \mathcal{M}} \int_0^{m_i} p_{ij} x_{ij} dj \quad \text{s.t.} \quad U(\mathbf{x}, \mathbf{m}) \geq \bar{U}.$$

---

<sup>4</sup>It is not difficult to generalize the framework to heterogeneous cities and consumers, in which case the HGT holds approximately for the marginal city. See Mieszkowski and Zodrow (1989) and Arnott (2004) for an overview.

<sup>5</sup>Note that  $\mathbf{m}$  in the utility function indicates that utility depends directly on product diversity, as has been emphasized by, e.g., Dixit and Stiglitz (1977).

The first-order conditions for quantities are given by:

$$p_i = \frac{\partial U(\mathbf{x}, \mathbf{m})/\partial x_i}{\partial U(\mathbf{x}, \mathbf{m})/\partial x_0}, \quad i \notin \mathcal{M} \quad \text{and} \quad p_{ij} = \frac{\partial U(\mathbf{x}, \mathbf{m})/\partial x_{ij}}{\partial U(\mathbf{x}, \mathbf{m})/\partial x_0}, \quad i \in \mathcal{M}, j \in [0, m_i]. \quad (1)$$

The masses of varieties are determined by the entry decisions of producers. The consumers, however, have the option of not consuming all the varieties that are produced. In order for the varieties to be consumed, we must have:<sup>6</sup>

$$\frac{\partial U(\mathbf{x}, \mathbf{m})/\partial m_i}{\partial U(\mathbf{x}, \mathbf{m})/\partial x_0} \geq p_{ij}x_{ij}, \quad i \in \mathcal{M}, j \in [0, m_i]. \quad (2)$$

Since all  $N$  consumers in a city benefit from an additional variety, the consumer-side shadow price of product diversity (i.e., the value of increasing the mass of varieties  $m_i$ ) for good  $i \in \mathcal{M}$  satisfies

$$p_{m_i} = N \frac{\partial U(\mathbf{x}, \mathbf{m})/\partial m_i}{\partial U(\mathbf{x}, \mathbf{m})/\partial x_0} \geq N p_{ij}x_{ij}, \quad i \in \mathcal{M}, j \in [0, m_i]. \quad (3)$$

Note that when conditions (2) hold, the consumer-side shadow price of product diversity is larger than or equal to the total expenditure on a variety in a community. In this sense, varieties can be undersupplied but are never oversupplied.

## 2.2 Production

Production in a city is given by the aggregate production function  $F(\mathbf{Y}, \mathbf{m}) = 0$ , where  $\mathbf{Y} = (Y_0, \{Y_{ij}, j \in [0, m_i]\}_{i \in \mathcal{M}}, \{Y_i\}_{i \in \mathcal{L}}, \{Y_i\}_{i \in \mathcal{H}})$  is an aggregate net output vector.<sup>7</sup> For simplicity, we assume that production does not require land. The shadow prices on the production side  $\mathbf{p} - \mathbf{t}$  may differ from consumer prices  $\mathbf{p}$  in a second-best world due to price distortions  $\mathbf{t}$ , e.g., monopoly power or taxes.<sup>8</sup> Following Boadway and Bruce (1984), we assume that price distortions are given by  $\mathbf{t} = (t_0, \{t_{ij}, j \in [0, m_i]\}_{i \in \mathcal{M}}, \{t_i\}_{i \in \mathcal{L}}, \{t_i\}_{i \in \mathcal{H}})$ .<sup>9</sup> We normalize the shadow prices such that  $t_0 \equiv 0$ . The producer shadow prices satisfy

$$p_i - t_i = \frac{\partial F(\mathbf{Y}, \mathbf{m})/\partial Y_i}{\partial F(\mathbf{Y}, \mathbf{m})/\partial Y_0}, \quad i \notin \mathcal{M} \quad \text{and} \quad p_{ij} - t_{ij} = \frac{\partial F(\mathbf{Y}, \mathbf{m})/\partial Y_{ij}}{\partial F(\mathbf{Y}, \mathbf{m})/\partial Y_0}, \quad i \in \mathcal{M}, j \in [0, m_i]. \quad (4)$$

Because we examine only the marginal conditions, this representation of price distortions is quite general. The producer shadow price of increasing the mass of varieties of good  $i \in \mathcal{M}$  is given by

$$p_{m_i} - t_{m_i} = \frac{\partial F(\mathbf{Y}, \mathbf{m})/\partial m_i}{\partial F(\mathbf{Y}, \mathbf{m})/\partial Y_0}. \quad (5)$$

We write the aggregate supply functions as  $Y_i(\mathbf{p}, \mathbf{m}, \mathbf{t})$ . This formulation permits both monopolistic as well as competitive behavior of producers as long as net outputs are uniquely determined.

---

<sup>6</sup>This condition assumes that the variety index  $j$  can be ordered so that the marginal rate of substitution between variety and the numeraire is decreasing in  $j$ .

<sup>7</sup>Note that  $\mathbf{m}$  in the production function indicates that production depends directly on product diversity, as has been emphasized by, e.g., Ethier (1982), Fujita (1989) and Fujita *et al.* (1999) in the new trade, the urban economics and the NEG literature.

<sup>8</sup>In the case of intermediate goods, we use the convention that the demand-side prices are denoted by  $\mathbf{p}$ .

<sup>9</sup>In general, the vector of price distortions  $\mathbf{t}$  can depend on the number of cities  $n$ . We suppress this argument to alleviate the notational burden.

### 2.3 Market clearing conditions

Denoting the *Allais surplus* by  $A$  (Allais, 1943), we can express the market clearing conditions as follows. Market clearing for the numeraire good requires that

$$A + \bar{N} [x_0(\mathbf{p}, \mathbf{m}, \bar{U}) - \bar{x}_0] = nY_0(\mathbf{p}, \mathbf{m}, \mathbf{t}). \quad (6)$$

For the differentiated goods, which we assume to be non-tradable, market clearing in each city is given by (recall that there are no initial endowments of differentiated goods):

$$Nx_{ij}(\mathbf{p}, \mathbf{m}, \bar{U}) = Y_{ij}(\mathbf{p}, \mathbf{m}, \mathbf{t}), \quad \forall i \in \mathcal{M}, \forall j \in [0, m_i]. \quad (7)$$

Recall also that land has a special characteristic, namely that if the number of cities increases, the total available land increases by  $\bar{X}_i$  for  $i \in \mathcal{L}$  of the added city. The resource constraint for land is hence:

$$\bar{N}x_i(\mathbf{p}, \mathbf{m}, \bar{U}) - n\bar{X}_i = nY_i(\mathbf{p}, \mathbf{m}, \mathbf{t}), \quad \forall i \in \mathcal{L}. \quad (8)$$

For the other goods, which may or may not be traded, we have

$$\bar{N}[x_i(\mathbf{p}, \mathbf{m}, \bar{U}) - \bar{x}_i] = nY_i(\mathbf{p}, \mathbf{m}, \mathbf{t}), \quad \forall i \in \mathcal{H}. \quad (9)$$

Note that the equilibrium conditions for non-traded goods must hold for each city separately, but at a symmetric equilibrium, these constraints are always satisfied. We therefore use the aggregate market clearing conditions (9) in what follows, no matter whether the goods are traded or not.

## 3 A second-best Henry George Theorem

In this section we consider the effects of changing the number  $n$  of cities (or equivalently, the size  $N$  of a city given the total population  $\bar{N}$ ). For each  $n$ , one can obtain the equilibrium values of all endogenous variables with consumers' utility levels fixed at  $\bar{U}$ . We do not specify how the equilibrium is obtained as our framework can incorporate many different models. We simply assume that, given the number of cities, a unique equilibrium  $\{\mathbf{p}(n), \mathbf{Y}(n), \mathbf{m}(n), A(n)\}$  exists.<sup>10</sup>

We now derive the welfare changes caused by a change in the number of cities,  $dA(n)/dn$ . The following theorem provides an extension of *Harberger's measure of welfare change* (Harberger, 1964), i.e., the weighted sum of induced changes with weights being the price distortions. This Harberger formula immediately yields our second-best Henry George Theorem:

**Theorem 1 (Harberger Formula and Second-best Henry George Theorem)** *An increase in the number of cities changes the Allais surplus by*

$$\frac{dA}{dn} = \text{FS} - \text{EB},$$

---

<sup>10</sup>Though this assumption is a strong one, which need not be satisfied in general, it will be satisfied in the various applications we present in the remainder of this paper.



where FS and EB are respectively the fiscal surplus of a city and the excess burden created by adding a city. The fiscal surplus is the sum of total profits and aggregate land rents in a city evaluated at producer prices, i.e.,  $FS \equiv \Pi + ALR$ , where

$$\Pi \equiv \sum_{i \in \mathcal{M}} \int_0^{m_i} (p_{ij} - t_{ij}) Y_{ij} dj + \sum_{i \notin \mathcal{M}} (p_i - t_i) Y_i \quad \text{and} \quad ALR \equiv \sum_{i \in \mathcal{L}} (p_i - t_i) \bar{X}_i.$$

The excess burden is the weighted sum of induced changes, with weights being the price and variety distortions:

$$EB \equiv -\bar{N} \left[ \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( t_{ij} \frac{dx_{ij}}{dn} \right) dj + \frac{1}{N} \sum_{i \in \mathcal{M}} t_{m_i} \frac{dm_i}{dn} + \sum_{i \notin \mathcal{M}} t_i \frac{dx_i}{dn} \right]. \quad (10)$$

When the number of cities is optimal,  $dA/dn = 0$  and the fiscal surplus equals the excess burden caused by increasing the number of cities:  $FS = EB$ . With only one type of differentiated good, the excess burden can be expressed as:

$$EB = -\bar{N} \left[ \int_0^m p_j \left( \frac{t_j}{p_j} - \frac{t_m}{p_m} \right) \frac{dx_j}{dn} dj + \sum_{i \notin \mathcal{M}} p_i \left( \frac{t_i}{p_i} - \frac{t_m}{p_m} \right) \frac{dx_i}{dn} \right], \quad (11)$$

where we may suppress subscript  $i \in \mathcal{M}$  for the differentiated good.

**Proof:** See Appendix A1. ■

The fiscal surplus of a city is the sum of total profits and aggregate land rents of that city. The excess burden is the weighted sum of induced changes in consumption and in product diversity ( $dx_i/dn$ ,  $dx_{ij}/dn$  and  $dm_i/dn$ ), with weights being the price and variety distortions ( $t_i$ ,  $t_{ij}$  and  $t_{m_i}$ ). If  $t_i$  is positive, i.e., the consumer price of good  $i$  is higher than its producer price, then an increase in the number of cities  $dn > 0$  reduces the excess burden provided that its consumption increases  $dx_i > 0$ . This is natural because the consumer price represents the marginal benefit of the good for a consumer, whereas the producer price represents its marginal social cost of production. If the former exceeds the latter, an increase in consumption results in net social benefits, i.e., a reduction in the excess burden. A similar reasoning applies to the variety distortion  $t_{m_i}$ .

The Harberger formula (10) includes induced changes in all the consumption goods and varieties, but one of them is not independent because the utility level is fixed in the Allais surplus measure. If there is only one type of differentiated good, we can exploit this constraint to eliminate the change in the mass of varieties from the formula to obtain (11). This simplified formula shows that two effects play critical roles in determining the sign of the excess burden: (i) the difference between the price margin of each variety and that of product diversity; and (ii) the difference between the price margin of each consumption good and that of product diversity.

Some additional comments are in order. First, when there are no distortions, the excess burden is zero and we obtain the well-known Henry George Theorem in the first-best world — the fiscal surplus is zero when the number of cities is optimal.

Second, Theorem 1 can be used to investigate in which direction the HGT may fail to hold. If the excess burden is negative at the second-best city size, the fiscal surplus must also be negative by Theorem 1. As the latter is the sum of total profits and aggregate land rents, the single tax on the aggregate land rents does not generate enough revenue to cover the aggregate losses (negative profits) from increasing returns to scale activities evaluated at producer prices.

Third, observe that even when distortions exist, the excess burden can be zero. Hence, the HGT may hold in a second-best economy. This can be seen by the excess burden formula (11) for the special case where there is only one differentiated good. The excess burden can be zero, for example, if  $dx_i/dn = 0$  for all  $i \notin \mathcal{M}$  and if  $t_j/p_j = t_m/p_m$  for all  $j \in [0, m]$ . As we will see later, the latter condition is satisfied in the CES model that has been widely used in the urban economics and the NEG literature. If there is no intersectoral substitution between the differentiated good and other goods, the former condition is also satisfied.

Fourth, we have so far examined the HGT using producer prices only. We can readily reformulate the theorem in terms of consumer prices as follows:

**Corollary 1 (Consumer price version of the second-best HGT)** *When the number of cities is optimal, the fiscal surplus (the sum of total profits and aggregate land rents) of a city evaluated at consumer prices*

$$FS_C \equiv \sum_{i \in \mathcal{M}} \int_0^{m_i} p_{ij} Y_{ij} dj + \sum_{i \in \mathcal{L}} p_i \bar{X}_i + \sum_{i \notin \mathcal{M}} p_i Y_i$$

*equals the excess burden caused by increasing the number of cities plus the total value of price distortions:  $FS_C = EB + T$ , where*

$$T \equiv \sum_{i \in \mathcal{M}} \int_0^{m_i} t_{ij} Y_{ij} dj + \sum_{i \in \mathcal{L}} t_i \bar{X}_i + \sum_{i \notin \mathcal{M}} t_i Y_i.$$

**Proof:** Straightforward from Theorem 1. ■

Last, our results so far have been derived without using profit maximization conditions. Theorem 1 therefore can be applied to cases where firms are not maximizing profits.

## 4 Price and variety distortions in monopolistic competition

To derive sharper results, we now examine simple monopolistic competition models with additively separable utility functions. In addition to separability, we make five assumptions. First, there exists only one type of differentiated good. Second, the differentiated good has a simple cost structure, featuring a constant marginal cost,  $c$ , and a fixed cost,  $F$ . Third, this sector is imperfectly competitive and profits are zero in equilibrium due to free entry and exit. Fourth, no distortion exists in sectors other than that producing the differentiated good. Fifth, we limit ourselves to symmetric equilibria, denoting the consumption, the price, the price distortion, and the production of each variety by  $x$ ,  $p$ ,  $t$ , and  $Y$ , respectively (without subscripts because of symmetry).

In the case of a single differentiated good, an additively separable utility function can be generally expressed as follows:

$$U(\mathbf{x}) = U(x_0, U_M, \{x_i\}_{i \in \mathcal{H}}), \quad (12)$$

where

$$U_M = \int_0^m u(x_j) dj$$

is the utility from the consumption of all varieties. Adding another variety with consumption  $x_m$  increases  $U_M$  by  $u(x_m)$ . For this to be equivalent to increasing the consumption of variety  $m$  from 0 to  $x_m$ , we must have  $u(0) = 0$ , which rules out all sub-utility functions such that  $u(0) \neq 0$ . However, this is not restrictive because the sub-utility function can be discontinuous as in

$$u(x) = \begin{cases} a + v(x) & \text{if } x > 0 \\ 0 & \text{if } x = 0, \end{cases} \quad (13)$$

where  $v$  is a well-behaved concave function that need not in general satisfy  $a + v(0) = 0$ .<sup>11</sup>

As shown by (11), the sign of the excess burden crucially depends on the price margin of a variety,  $t/p$ , and that of product diversity,  $t_m/p_m$ . In the additively separable case, the results derived in the preceding section become much simpler as we can sign both of these margins. In particular, we can prove the following theorem.

**Theorem 2 (HGT in the additively separable case)** *Assume that there is a single differentiated good and that the utility function is additively separable with respect to its varieties. Then the fiscal surplus is  $FS = \Pi + ALR$ , where  $\Pi = -mF$  and  $ALR = \sum_{i \in \mathcal{L}} p_i \bar{X}_i$ . The excess burden is*

$$EB = \bar{N} \left[ mp\theta_U(x) \frac{dx}{dn} + \frac{t_m}{p_m} \left( \frac{dx_0}{dn} + \sum_{i \in \mathcal{H}} p_i \frac{dx_i}{dn} \right) \right], \quad (14)$$

where

$$\theta_U(x) \equiv \frac{dU_R(x)}{dx} \frac{x}{U_R(x)} = \frac{t_m}{p_m} - \frac{t}{p},$$

is the elasticity of the scale elasticity of utility, defined as  $U_R(x) \equiv xu'/u$ .  $\theta_U$  depends on the difference between the price margins, both of which are nonnegative:

$$\frac{t_m}{p_m} = 1 - U_R(x) \geq 0 \quad \text{and} \quad \frac{t}{p} = R_R(x) \geq 0,$$

where  $R_R(x) \equiv -(u''x)/u'$  denotes the relative risk aversion (RRA). Furthermore,

$$\frac{dx}{dn} \begin{matrix} \geq \\ < \end{matrix} 0 \quad \text{as} \quad \theta_R + \varpi \begin{matrix} \geq \\ < \end{matrix} 0,$$

where

$$\theta_R(x) \equiv \frac{dR_R(x)}{dx} \frac{x}{R_R(x)} \quad \text{and} \quad \varpi \equiv \frac{cY}{cY + F}$$

---

<sup>11</sup>An equilibrium obtained with  $u(x)$  defined above can be replicated using another utility function without the discontinuity. As we focus on symmetric equilibria, which are bounded away from zero, there are no additional difficulties in doing so.

denote the elasticity of the RRA and the share of variable cost in the total cost of producing a variety of the differentiated good, respectively.

**Proof:** See Appendix A2. ■

Several comments are in order. First, because distortions are limited to the differentiated good sector, the price margins of all other goods disappear from the excess burden. Changes in their consumption multiplied by the price margin of product diversity, however, do affect the excess burden because these changes occur simultaneously with a change in the mass  $m$  of varieties consumed. For example, a decrease in the mass of varieties tends to increase, all else equal, the consumption of the numeraire. Because there is less than optimum product diversity ( $t_m > 0$ ), this increases the excess burden. Another implication of the assumption of no distortion in all other goods is that aggregate land rents do not include the price distortions  $t_i$ .

Second, the first term in the square bracket of (14) captures the effects of a change in the number of cities on the differentiated good sector. Its sign depends on those of  $dx/dn$  and of  $\theta_U$ . The former is positive if the sum of the elasticity of the RRA and of the variable cost share is positive. Increasing RRA (i.e.,  $\theta_R > 0$ ) is a sufficient condition for  $dx/dn$  to be positive. However, even if the RRA is decreasing, it can be positive, provided that its rate of decrease is small enough so that the sum of the elasticity of RRA and the variable cost share remains positive.<sup>12</sup> Although we cannot rule out the case where  $dx/dn < 0$ , it is positive in all the examples that are reported below.

The sign of  $\theta_U$  may be positive or negative depending on the difference between the price margin of product diversity,  $t_m/p_m$ , and that of each variety,  $t/p$ , both of which are nonnegative. If the former is smaller than the latter, the excess burden tends to be negative. The intuition behind this is as follows. An increase in consumption,  $x$ , of each variety tends to mitigate the excess burden caused by the price distortion as there is too little consumption at equilibrium. At the same time, however, an increase in the consumption of each variety tends to reduce  $m$ . As the sign of the variety distortion implies that  $m$  is too low, this tends to exacerbate the excess burden.

Third, the price margin of product diversity depends on the level of sub-utility  $u(x)$ , unlike in expected utility theory where the utility function is unique up to an affine transformation. In our setting, an increase in  $a$  in (13) raises  $\theta_U$  when the sub-utility is positive, thereby increasing the equilibrium valuation of variety.<sup>13</sup> This result is closely linked to that of Benassy (1996), who disentangles taste for variety from market power. In the additively separable case, his measure of

---

<sup>12</sup>If marginal cost were zero, the sum of the elasticity of the RRA and the variable cost share would be negative in the decreasing RRA case. However, in that case the zero profit condition (see expression (37) in Appendix A2) cannot be satisfied unless the relative risk aversion is one. This is not possible since the price would be infinite.

<sup>13</sup>Let  $\theta_U(x, 0)$  and  $\theta_U(x, a)$  denote the elasticity of the scale elasticity of utility when  $a = 0$  and when  $a > 0$ , respectively. Without loss of generality, we can take the base utility function that satisfies  $v(x) \geq 0$ . As  $u(x) = a + v(x)$  for  $x > 0$ , we obtain

$$\theta_U(x, a) = 1 - R_R(x) - U_R(x, a) = 1 - R_R(x) - \frac{v'(x)x}{a + v(x)} = \theta_U(x, 0) + \frac{a}{a + v(x)} U_R(x, 0) \geq \theta_U(x, 0)$$

if  $a \geq 0$  (recall that  $u(x) = v(x)$  when  $a = 0$ ). Clearly, we also have  $\partial\theta_U(x, a)/\partial a \geq 0$ .

taste for variety is  $\epsilon(m) = [m\rho'(m)]/\rho(m)$ , where  $\rho(m) = [mu(x)]/u(mx)$ . This satisfies  $\epsilon(m) = 1 - [mxu'(mx)]/u(mx)$ , which is similar to our  $t_m/p_m = 1 - xu'(x)/u(x)$ . The difference is that in Benassy (1996), holding  $mx$  constant implies that a change in  $m$  has no impact on  $\epsilon(m)$ , which is not the case with our  $t_m/p_m$ .

Fourth, the total profit in a community is negative in Theorem 2 even though we have free-entry zero-profit conditions. The reason is that it is evaluated at producer prices. In Corollary 1, which evaluates the fiscal surplus using consumer prices, the fiscal surplus becomes  $FS_C = EB + T$ , where  $FS_C = \Pi_C + ALR$ ,  $\Pi_C = 0$ , and  $T = mtY = mF$ . In this version, the total profit is zero and the total value of price distortion equals  $mF$ .

Last, the first-order condition for profit maximization  $t/p = (p - c)/p = R_R(x)$  and the zero profit condition  $(p - c)(\bar{N}/n)x - F = 0$  are sufficient to determine  $p$  and  $x$  as functions of the number of cities,  $p(n)$  and  $x(n)$ . Hence, given  $n$ , these two variables are determined solely by the shape of the sub-utility function  $u(x)$  and the cost structure of the differentiated good sector,  $c$  and  $F$ . This implies that the signs of  $\theta_U(x)$  and  $dx/dn$  do not depend on sectors other than that of the differentiated good.

We now revisit some examples that have been analyzed in the literature thus far. The first example illustrates intersectoral distortions between the differentiated good and other goods assuming the CES form across varieties. In that case, price and variety distortions cancel out in the differentiated good sector. In the second example, we assume away intersectoral distortions and examine two types of utility functions — a variable RRA form which includes the CES (or constant RRA) as a special case, and a generalized CARA form.

## 4.1 Example 1: Intersectoral distortions

Theorem 2 shows that even if distortions exist only in the differentiated good sector, changes in consumption of the other goods induce changes in the excess burden through the intersectoral effects as they are associated with changes in the variety distortion. Assume, for example, that there is a single housing good in (12) with price  $p_H$ , the consumption of which is fixed to  $x_i = \bar{h}$ . In that case, the expenditure minimization problem is given by

$$\min x_0 + \int_0^m p_j x_j dj + p_H \bar{h} \quad \text{s.t.} \quad U(x_0, U_M, \bar{h}) \geq \bar{U}.$$

The first-order conditions for expenditure minimization, the first-order condition for profit maximization,  $t/p = (p - c)/p = R_R(x)$ , and the zero profit condition,  $(p - c)(\bar{N}/n)x - F = 0$ , all remain unchanged. Therefore, expression (14) applies with  $dx_i/dn = 0$  for  $i \in \mathcal{H}$ . Specifically, the excess burden is now given by

$$EB = \bar{N} \left[ mp\theta_U(x) \frac{dx}{dn} + \frac{t_m}{p_m} \frac{dx_0}{dn} \right].$$

To sign it, we require information on the signs of  $\theta_U(x)$ ,  $dx/dn$ ,  $t_m/p_m$  and  $dx_0/dn$ . These can be derived once we specify utility functions.

Let us consider a case similar to the one in Abdel-Rahman and Fujita (1990). The upper-tier utility function  $U$  is assumed to be Cobb-Douglas, whereas the lower-tier utility function  $U_M$  is a CES aggregator of the differentiated good. More precisely, we assume that:<sup>14</sup>

$$U(\mathbf{x}) = x_0^\eta \left( \int_0^m x_j^{\frac{\sigma-1}{\sigma}} dj \right)^{1-\eta}, \quad (15)$$

with  $\sigma > 1$  and  $0 < \eta < 1$ . The RRA in this case is constant and given by  $R_R(x) = 1/\sigma$ , so that  $t/p = R_R(x) = 1/\sigma$ . The scale elasticity of utility is given by  $U_R(x) = (\sigma - 1)/\sigma$ , and hence  $t_m/p_m = 1 - U_R(x) = 1/\sigma$ . Combining these two conditions yields  $\theta_U(x) = 0$ . The sign of the excess burden thus depends solely on the sign of the intersectoral distortion as captured by  $t_m/p_m$  and  $dx_0/dn$ . We can show the following result.

**Result 1 (Intersectoral distortion)** *If the utility function is given by (15), then  $t/p = t_m/p_m = 1/\sigma$  so that  $\theta_U(x) = 0$ . Furthermore,*

$$\frac{dx_0}{dn} = \frac{x_0}{\sigma n} (1 - \eta) > 0.$$

*Hence, the excess burden is positive:*

$$EB = \overline{N} \frac{t_m}{p_m} \frac{dx_0}{dn} > 0.$$

**Proof:** See Appendix A3. ■

Result 1 shows that, in the Cobb-Douglas/CES case, the fiscal surplus is positive at the second-best city size, thus implying that a single tax on aggregate land rents can raise enough revenue to implement the first-best allocation conditional on the second-best city size. Note that as  $\eta \rightarrow 0$ ,  $x_0 \rightarrow 0$  from (41) in Appendix A3. Hence,  $EB \rightarrow 0$  as  $\eta$  gets arbitrarily small. This suggests that once the intersectoral distortions disappear from the model, the HGT holds at the second best when the utility function for the differentiated good is of the CES type. This result will be established formally in the next subsection.

## 4.2 Example 2: No intersectoral distortions

Next, we consider a case where consumption of the numeraire and housing are fixed, i.e.,  $dx_0/dn = dx_i/dn = 0$  for all  $i \in \mathcal{H}$ . We can think of this case as one where labor is the numeraire and where labor supply is (exogenously) fixed. Important for our purpose is that intersectoral distortions across the numeraire good and the differentiated good vanish in that case. As shown in Theorem 2, only the signs of  $\theta_U(x)$  and  $dx/dn$  then matter for characterizing the sign of the excess burden.

---

<sup>14</sup>There are two major differences with Abdel-Rahman and Fujita (1990). First, they focus on a differentiated intermediate good, instead of a differentiated consumption good. Second, they use  $[\int_0^m x_j^{(\sigma-1)/\sigma} dj]^{\sigma/(\sigma-1)}$ , whereas we use  $\int_0^m x_j^{(\sigma-1)/\sigma} dj$  for  $U_M$ . The reason for the latter choice is that we can show that as  $\eta$  goes to zero,  $U(\mathbf{x})$  in (15) boils down to a special case of a variable RRA model that we analyze in Section 4.2.

### 4.2.1 A variable RRA form

Ogaki and Zhang (2001) proposed a family of utility functions embedding varying degrees of RRA represented by a simple extension of the CES as follows:  $v(x) = (x - \gamma)^{(\sigma-1)/\sigma}$ . Since the absolute level of utility matters for the sign of the excess burden, we add a constant term to  $v(x)$ :

$$u(x) = \begin{cases} a + (x - \gamma)^{\frac{\sigma-1}{\sigma}} & \text{if } x > 0 \\ 0 & \text{if } x = 0, \end{cases} \quad (16)$$

where  $\sigma > 1$  and  $x \geq \max\{0, \gamma\}$ . We further impose a ‘love-of-variety’ condition, which can be obtained by combining the first-order conditions for the consumer’s choice of quantity and variety, i.e., (1) and (2), as follows:

$$u(x) \geq xu'(x) \geq 0. \quad (17)$$

This condition places restrictions on the values of  $a$  and/or  $x$ .

The RRA in this case is given by

$$R_R(x) = \frac{x}{\sigma(x - \gamma)}, \quad (18)$$

and its elasticity is equal to

$$\theta_R(x) = -\frac{\gamma}{x - \gamma}. \quad (19)$$

This specification hence encapsulates three different cases: (i) CRRA (constant RRA) when  $\gamma = 0$ ; (ii) IRRA (increasing RRA) when  $\gamma < 0$ ; and (iii) DRRA (decreasing RRA) when  $\gamma > 0$ .

The shift parameter  $a$  does not matter for the RRA or its elasticity. This is not the case for the scale elasticity of utility because

$$U_R(x) = \frac{\sigma - 1}{\sigma} \frac{x}{x - \gamma} \frac{1}{1 + a(x - \gamma)^{\frac{1-\sigma}{\sigma}}}$$

is decreasing in  $a$ . The elasticity of the scale elasticity is given by

$$\theta_U(x) = -\frac{\gamma}{x - \gamma} + \frac{\sigma - 1}{\sigma} \frac{x}{x - \gamma} \frac{a(x - \gamma)^{\frac{1-\sigma}{\sigma}}}{1 + a(x - \gamma)^{\frac{1-\sigma}{\sigma}}}.$$

Comparing this expression with (19) and making use of (17) we readily see that

$$\theta_U(x) \gtrless \theta_R(x) \quad \text{as } a \gtrless 0.$$

In particular, if  $a = 0$  then  $\theta_U(x)$  coincides with the elasticity of RRA and (i)  $\theta_U(x) > 0$  in the IRRA case, (ii)  $\theta_U(x) < 0$  in the DRRA case, and (iii)  $\theta_U(x) = 0$  in the CRRA case. If  $a$  is positive (negative), then  $\theta_U$  may be positive (negative) even in the DRRA (IRRA) case.

Using these properties, we can show the following result.

**Result 2 (Variable RRA form)** *In the variable RRA case of (16) with  $\sigma > 1$ , we have  $dx/dn > 0$ . Hence, when  $a = 0$ ,  $EB > 0$  in the IRRA case ( $\gamma < 0$ ),  $EB < 0$  in the DRRA case ( $\gamma > 0$ ), and  $EB = 0$  in the CRRA case ( $\gamma = 0$ ). If  $a$  is positive (negative), then the excess burden may be positive (negative) even in the DRRA (IRRA) case.*

**Proof:** See Appendix A4. ■

These results are illustrated in Figure 1. Some comments are in order. First, in the ‘standard’ CES case (with  $a = \gamma = 0$ ), the excess burden is zero, even in the second best. This confirms the results obtained by Duranton and Puga (2001) and by Behrens and Murata (2009). Second, if  $a = 0$ , the sign of the excess burden depends on whether the RRA is increasing or not: the excess burden is positive in the IRRA case and negative in the DRRA case. Last, an increase in  $a$  tends to make the excess burden larger. Hence, even in the CRRA case the excess burden is positive if  $a$  is positive. When taken together, these findings show that the case of zero excess burden is a rather special one.

**Insert Figure 1 about here.**

#### 4.2.2 A generalized CARA form

In the CARA case of Behrens and Murata (2007, 2009), the sub-utility function is given by  $u(x) = 1 - e^{-\alpha x}$ , with  $\alpha > 0$ . We extend this form to

$$u(x) = \begin{cases} a - e^{-\alpha x} & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases} \quad (20)$$

This sub-utility function satisfies  $u'(x) = \alpha e^{-\alpha x}$  and  $u''(x) = -\alpha^2 e^{-\alpha x}$  for any  $x > 0$ . The RRA is then given by

$$R_R(x) = \alpha x, \quad (21)$$

which is increasing in  $x$ . The elasticity of the RRA is simply given by  $\theta_R(x) = 1$ . Some straightforward computations show that the scale elasticity of the utility function is

$$U_R(x) = \frac{\alpha x e^{-\alpha x}}{a - e^{-\alpha x}},$$

which yields

$$\theta_U(x) = 1 - \frac{a\alpha x}{a - e^{-\alpha x}}.$$

As in the variable RRA case, we impose the ‘love-of-variety’ condition (17) which restricts the values of  $a$  and/or  $x$ . First, we cannot have  $a \leq 0$  in the generalized CARA case because  $u(x) \geq xu'(x)$  cannot be satisfied for any positive  $x$ . Second, for the case of  $0 < a < 1$ , (17) will be satisfied provided that  $x$  exceeds some threshold  $\tilde{x}$ . Last, if  $a \geq 1$ , (17) is always satisfied since  $1 - (1 + \alpha x)e^{-\alpha x} \geq 0$  for all  $x \geq 0$ .

Using these properties, we can show the following result.

**Result 3 (Generalized CARA form)** *In the CARA case of (20) with  $\alpha > 0$  and  $a > 0$ , we have  $dx/dn > 0$ . If  $0 < a \leq 1$ , then  $\theta_U(x) < 0$  in equilibrium. If  $a > 1$ , then  $\theta_U(0) = 1$  and  $\theta_U(x)$  decreases monotonically to eventually become negative as  $x$  gets larger. Thus, if  $0 < a \leq 1$ , then  $EB < 0$ . If  $a > 1$ , there exists a threshold  $\hat{x}$  such that  $EB \gtrless 0$  for  $x \lesseqgtr \hat{x}$ .*



**Proof:** See Appendix A5. ■

Observe that the special case of  $a = 1$  in Result 3 coincides with the results derived by Behrens and Murata (2009), who have shown that the excess burden (and hence the fiscal surplus) is negative in the CARA case at the second best.<sup>15</sup>

As in the variable RRA case of Section 4.1, we can depict our results on the sign of the excess burden in Figure 2.<sup>16</sup>

**Insert Figure 2 about here.**

As can be seen from the figure, increasing  $a$  increases the excess burden, reflecting what we have discussed following Theorem 2. Furthermore, the excess burden is negative if  $a \leq 1$ , thus confirming Result 3. In the range of  $a > 1$ , the excess burden is negative initially but it eventually becomes positive as  $a$  becomes larger. There exists a value  $\hat{a}(\alpha)$  such that the excess burden is equal to zero and such that the associated optimal community size has zero fiscal surplus. In words, the HGT may still hold, but only for cases with zero measure in parameter space. Thus, *the HGT seems to fail generically in the CARA model*. However, it is worth noting that departures from the theorem may go in either direction, i.e., the fiscal surplus at the second-best optimum may be negative or positive. The precise condition for the direction of the failure crucially hinges on the value of  $a$  (and, thus, on the difference between the price and the variety distortions). Hence, any empirical implementation of the HGT in a second-best world with monopolistic competition will require to finely assess the relative magnitude of the price and variety distortions.

## 5 Distortionary taxation and local public goods

We now apply our second-best HGT to the case of distortionary taxes that are used to finance a local public good  $G$ . The aggregate production function is assumed to be given by  $F(\mathbf{Y}, G) = 0$ , whereas the utility function is  $U(\mathbf{x}, G)$ . For simplicity, we assume that there are no differentiated goods. Hence,  $\mathbf{m}$  no longer features in the utility and production functions. We also assume that the supply of  $G$  is fixed. The resource constraints are then as follows:

$$\begin{aligned} A + \bar{N} [x_0(\mathbf{p}, G, \bar{U}) - \bar{x}_0] &= nY_0(\mathbf{p}, G, \mathbf{t}) \\ \bar{N} [x_i(\mathbf{p}, G, \bar{U}) - \bar{x}_i] &= nY_i(\mathbf{p}, G, \mathbf{t}), \quad i \in \mathcal{H} \\ \bar{N}x_i(\mathbf{p}, G, \bar{U}) - n\bar{X}_i &= nY_i(\mathbf{p}, G, \mathbf{t}), \quad i \in \mathcal{L} \end{aligned} \tag{22}$$

---

<sup>15</sup>We also derived the results for a specification involving quadratic preferences. The results are qualitatively similar to those in the CARA case and are, therefore, not reported here. They are available upon request.

<sup>16</sup>As shown in Result 3, we need the information on  $\hat{x}$  to obtain the sign of the excess burden in the case of  $a > 1$ . Hence, we resort to a numerical investigation to distill the key insights of this case. To this end, we need to specify in more detail the production side of the model and the land and housing markets (see Appendix B for further information). The parameter values are set as follows:  $\bar{U} = 10$ ,  $\alpha = 2$ ,  $F = 0.2$ ,  $c = 0.1$  and  $b = 0.1$ .

Because we do not have differentiated goods, the change in the Allais surplus (32) in the proof of Theorem 1 becomes

$$\frac{dA}{dn} = \left[ \sum_{i \in \mathcal{I}} (p_i - t_i) Y_i + \sum_{i \in \mathcal{L}} (p_i - t_i) \bar{X}_i \right] - \left( -\bar{N} \sum_{i \neq 0} t_i \frac{dx_i}{dn} \right). \quad (23)$$

or, equivalently,  $dA/dn = \text{FS} - \text{EB}$ .

Consider a simple setting where the only distortion is a property tax on housing. There are only two consumption goods, the numeraire and housing  $H$ . There is also a single fixed factor land  $L$ , which is exclusively used to produce housing. In that case, condition (23) reduces to

$$\frac{dA}{dn} = Y_0 + (p_H - t_H) Y_H + p_L \bar{X}_L - \left( -\bar{N} t_H \frac{dx_H}{dn} \right).$$

We begin by assuming that the tax rate is exogenously fixed at  $t_H$ . The inputs to housing production are the numeraire and land. The aggregate production function for housing in a community is  $Y_H = H(H_0, H_L)$ , where  $H_0$  and  $H_L$  are the quantities of the numeraire and of land used to produce housing, respectively. Production of the numeraire requires labor only and its production function is  $F_0(N)$ , where  $N$  denotes as before the city size. Production of the local public good requires  $G_0$  units of the numeraire. Then, the net production of the numeraire in a community is  $Y_0 = F_0(N) - H_0 - G_0$ . The utility function is  $U(x_0, x_H, G)$ , where  $x_H$  is the consumption of housing. The resource constraints (22) thus boil down to

$$\begin{aligned} A + \bar{N} [x_0(1, p_H, G, \bar{U}) - \bar{x}_0] &= n [F_0(N) - H_0 - G_0] \\ Nx_H &= H(H_0, H_L) \\ \bar{X}_L &= H_L. \end{aligned}$$

We assume that the housing sector has constant returns to scale. Then, we can rewrite the production function as  $H(H_0, H_L) = H_L h(k)$ , where  $k = H_0/H_L$  and  $h(k) = H(kH_L, H_L)/H_L$  with  $h'(k) > 0$  and  $h''(k) < 0$ .

The property tax is set at  $t_H$  per unit of housing. The aggregate profit of the housing sector is then as follows:

$$\Pi_H = (p_H - t_H)H - H_0 - p_L H_L = H_L [(p_H - t_H) h(k) - k - p_L]$$

The first-order condition for  $k$  is  $(p_H - t_H) h'(k) = 1$ , which we can solve for  $k$  to obtain  $k^* = k^*(p_H - t_H) = (h')^{-1}((p_H - t_H)^{-1})$ . Letting  $k^{*'} \equiv dk^*/dp_H$  denote the change in relative inputs with respect to changes in housing prices, we then have

$$k^{*'} = -\frac{h'}{(p_H - t_H) h''} > 0.$$

Since housing is the only non-numeraire good, the excess burden reduces to

$$\text{EB} = -\bar{N} t_H \frac{dx_H}{dn} = -\bar{N} t_H \frac{\partial x_H}{\partial p_H} \frac{dp_H}{dN} \frac{dN}{dn}.$$

Observe now that the compensated demand function satisfies  $\partial x_H(1, p_H, G, \bar{U})/\partial p_H < 0$ . Furthermore, we can rewrite the resource constraint for housing as

$$Nx_H(1, p_H, G, \bar{u}) = \bar{X}_L h(k^*(p_H - t_H)),$$

which yields

$$\frac{dp_H}{dN} = \frac{x_H}{\bar{X}_L h'k^{*'} - N \frac{\partial x_H}{\partial p_H}} > 0.$$

Finally, from  $dN/dn < 0$ , we obtain  $EB < 0$ , i.e., the excess burden is negative.<sup>17</sup>

## 6 Conclusions

Most factors that cause urban agglomeration involve some form of market failure. In this article, we have identified conditions under which the ‘golden rule’ of local public finance, i.e., the Henry George Theorem holds in a second-best world and investigated in which directions the theorem needs to be modified otherwise. In so doing, we have largely focused on monopolistic competition models that are widely used in urban economics and the NEG to generate local increasing returns to scale and in which distortions in both prices and product diversity matter.

Our key findings may be summarized as follows. First, we have shown that the HGT will hold at the second best if and only if the excess burden from adding another community (or, equivalently, changing community size) equals the fiscal surplus from doing so. The excess burden can be expressed by an extended Harberger Formula — the sum of induced quantity and variety changes from adding another community, weighted by the associated price and variety distortions. In the case of a differentiated good and monopolistic competition, the sign of the excess burden is determined by the difference in the variety and price distortions, which are in turn linked to the scale elasticity of utility and the relative risk aversion, respectively. We have shown that the price distortion tends to make the excess burden negative, while the variety distortion works in the opposite direction. When the sub-utility is shifted upwards, the variety distortion becomes larger and the excess burden is more likely to be positive. Last, we have shown that the HGT only holds for a zero measure subset of cases at the second best.

A first natural extension of the present work would be to follow Kanemoto *et al.* (1996) and to use our more general conditions to evaluate whether observed city sizes are likely to be associated with positive or negative excess burdens. A second natural extension would be to apply our approach to the evaluation of transport policies, as they are likely to significantly affect city sizes and the spatial distribution of economic activity (Venables, 2007). We leave these avenues open for future research.

---

<sup>17</sup>We have assumed so far that the property tax rate is fixed. We also extended the analysis to the case where the local government sets a tax rate subject to its budget constraint. In that case, the excess burden remains negative provided the cost of the local public good is small enough. Results are available upon request.

## References

- [1] Abdel-Rahman, H. and M. Fujita (1990) Product variety, Marshallian externalities, and city sizes, *Journal of Regional Science* 30, 165–184.
- [2] Allais, M. (1943) *A la Recherche d'une Discipline Économique, vol. I*. Paris: Imprimerie Nationale.
- [3] Arnott, R. (2004) Does the Henry George Theorem provide a practical guide to optimal city size?, *The American Journal of Economics and Sociology* 63, 1057–1090.
- [4] Arnott, R. 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) Are Chinese cities too small?, *Review of Economic Studies* 73, 549–576.
- [6] Au, C.-C. and J.V. Henderson (2006b) How migration restrictions limit agglomeration and productivity in China, *Journal of Development Economics* 80, 350–388.
- [7] Behrens, K., G. Duranton and F.L. Robert-Nicoud (2010) Productive cities: Sorting, selection and agglomeration. CEPR Discussion Paper #7922.
- [8] Behrens, K. and Y. Murata (2007) General equilibrium models of monopolistic competition: A new approach, *Journal of Economic Theory* 136, 776–787.
- [9] Behrens, K. and Y. Murata (2009) City size and the Henry George Theorem under monopolistic competition, *Journal of Urban Economics* 65, 228–235.
- [10] Benassy, J.-P. (1996) Taste for variety and optimum production patterns in monopolistic competition, *Economics Letters* 52, 41–47.
- [11] Boadway, R. and N. Bruce (1984) *Welfare Economics*. Oxford: Basil-Blackwell.
- [12] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297–308.
- [13] Duranton, G. and D. Puga (2001) Nursery cities: Urban diversity, process innovation, and the life cycle of products, *American Economic Review* 91, 1454–1477.
- [14] Ethier, W.J. (1982) National and international returns to scale in the modern theory of international trade, *American Economic Review* 72, 389–405.
- [15] Flatters, F., J.V. Henderson and P. Mieszkowski (1974) Public goods, efficiency, and regional fiscal equalization, *Journal of Public Economics* 3, 99–112.

- [16] Fujita, M. (1989) *Urban Economic Theory: Land Use and City Size*. Cambridge, MA: Cambridge Univ. Press.
- [17] Fujita, M., P. Krugman, and A. Venables (1999) *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Massachusetts.
- [18] Harberger, A.C. (1964) The measurement of waste, *American Economic Review* 54, 58–76.
- [19] 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.
- [20] Kanemoto, Y. (1980) *Theories of Urban Externalities*. Amsterdam: North-Holland.
- [21] Kanemoto, Y., T. Kitagawa, H. Saito and E. Shioji (2005) Estimating urban agglomeration economies for Japanese metropolitan areas: Is Tokyo too large? In: A. Okabe (ed.) *GIS-Based Studies in the Humanities and Social Sciences*. Taylor&Francis, Boca Raton, pp. 229–241.
- [22] 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.
- [23] Kokovin, S., J.-F. Thisse and E. Zhelobodko (2010) Monopolistic competition: Beyond the CES. CEPR Discussion Paper #7947.
- [24] Mieszkowski, P. and G.R. Zodrow (1989) Taxation and the Tiebout model: The differential effects of head taxes, taxes on land rents, and property taxes, *Journal of Economic Literature* 27, 1098–1146.
- [25] Ogaki, M. and Q. Zhang (2001) Decreasing risk aversion and tests of risk sharing, *Econometrica* 69, 515–526.
- [26] 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, pp. 274–333.
- [27] Venables, A.J. (2007) Evaluating urban transport improvements: Cost-benefit analysis in the presence of agglomeration and income taxation, *Journal of Transport Economics and Policy* 41, 173–188.

## Appendix A: Proofs

### A1. Proof of Theorem 1

Given  $n$ , an equilibrium satisfies (6)–(9). Differentiating (6) with respect to  $n$  yields

$$\frac{dA}{dn} = Y_0 + n \frac{dY_0}{dn} - \bar{N} \frac{dx_0}{dn}. \quad (24)$$

We modify this expression by using the condition that the equilibrium values satisfy the production constraint  $F(\mathbf{Y}(n), \mathbf{m}(n)) = 0$  and the compensated demand function  $U(\mathbf{x}(n), \mathbf{m}(n)) = \bar{U}$  for any  $n$ .

First, differentiating the production function with respect to  $n$  and using the first-order conditions for optimal production choices (4) and (5) yields

$$\begin{aligned} \sum_{i \notin \mathcal{M}} \frac{\partial F}{\partial Y_i} \frac{dY_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( \frac{\partial F}{\partial Y_{ij}} \frac{dY_{ij}}{dn} \right) dj + \sum_{i \in \mathcal{M}} \frac{\partial F}{\partial m_i} \frac{dm_i}{dn} \\ = - \frac{\partial F}{\partial Y_0} \left\{ \sum_{i \notin \mathcal{M}} (p_i - t_i) \frac{dY_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left[ (p_{ij} - t_{ij}) \frac{dY_{ij}}{dn} \right] dj + \sum_{i \in \mathcal{M}} (p_{m_i} - t_{m_i}) \frac{dm_i}{dn} \right\} = 0. \end{aligned}$$

Recall that  $p_i - t_i$  (respectively,  $p_{ij} - t_{ij}$ ) is the production side shadow price of good  $i$  (respectively, of variety  $ij$ ) and that  $m_i$  is the number of varieties of the differentiated consumer good  $i \in \mathcal{M}$ . The latter may of course be an endogenous variable. Because of our normalizations  $p_0 \equiv 1$  and  $t_0 \equiv 0$ , we can rewrite the foregoing relationship as:

$$\frac{dY_0}{dn} = - \sum_{i \neq 0, i \notin \mathcal{M}} (p_i - t_i) \frac{dY_i}{dn} - \sum_{i \in \mathcal{M}} \int_0^{m_i} \left[ (p_{ij} - t_{ij}) \frac{dY_{ij}}{dn} \right] dj - \sum_{i \in \mathcal{M}} (p_{m_i} - t_{m_i}) \frac{dm_i}{dn}. \quad (25)$$

Next, since the compensated demand satisfies by definition  $U(\mathbf{x}(n), \mathbf{m}(n)) = \bar{U}$ , for a fixed utility level  $\bar{U}$ , we have

$$\frac{dU}{dn} = \sum_{i \notin \mathcal{M}} \frac{\partial U}{\partial x_i} \frac{dx_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( \frac{\partial U}{\partial x_{ij}} \frac{dx_{ij}}{dn} \right) dj + \sum_{i \in \mathcal{M}} \frac{\partial U}{\partial m_i} \frac{dm_i}{dn} = 0.$$

Using the first-order conditions for optimal consumption choice (1) and the definition of the consumer side shadow price of product diversity (3) then yields:

$$\sum_{i \notin \mathcal{M}} p_i \frac{dx_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( p_{ij} \frac{dx_{ij}}{dn} \right) dj + \frac{1}{N} \sum_{i \in \mathcal{M}} p_{m_i} \frac{dm_i}{dn} = 0. \quad (26)$$

Note that condition (26) is an extension of the homogeneity property of the expenditure function to include the varieties of differentiated commodities,  $\mathbf{m}$ . Since good 0 is the numeraire, we can further rewrite this as

$$\frac{dx_0}{dn} = - \sum_{i \neq 0, i \notin \mathcal{M}} p_i \frac{dx_i}{dn} - \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( p_{ij} \frac{dx_{ij}}{dn} \right) dj - \frac{1}{N} \sum_{i \in \mathcal{M}} p_{m_i} \frac{dm_i}{dn}. \quad (27)$$

Substituting (25) and (27) into (24) yields

$$\begin{aligned} \frac{dA}{dn} = Y_0 - n \left\{ \sum_{i \neq 0, i \notin \mathcal{M}} (p_i - t_i) \frac{dY_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left[ (p_{ij} - t_{ij}) \frac{dY_{ij}}{dn} \right] dj + \sum_{i \in \mathcal{M}} (p_{m_i} - t_{m_i}) \frac{dm_i}{dn} \right\} \\ + \bar{N} \left[ \sum_{i \neq 0, i \notin \mathcal{M}} p_i \frac{dx_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( p_{ij} \frac{dx_{ij}}{dn} \right) dj + \frac{1}{N} \sum_{i \in \mathcal{M}} p_{m_i} \frac{dm_i}{dn} \right]. \quad (28) \end{aligned}$$

Furthermore, by differentiating the market clearing conditions (7)–(9), with respect to  $n$ , we obtain

$$\bar{N} \frac{dx_i}{dn} = \bar{X}_i + Y_i + n \frac{dY_i}{dn}, \quad i \in \mathcal{L}, \quad (29)$$

$$\bar{N} \frac{dx_{ij}}{dn} = Y_{ij} + n \frac{dY_{ij}}{dn}, \quad i \in \mathcal{M}, \quad j \in [0, m_i]. \quad (30)$$

$$\bar{N} \frac{dx_i}{dn} = Y_i + n \frac{dY_i}{dn}, \quad i \in \mathcal{H}. \quad (31)$$

Substituting (29)–(31) into (28) and using  $t_0 = 0$  finally yields:

$$\begin{aligned} \frac{dA}{dn} &= \sum_{i \notin \mathcal{M}} (p_i - t_i) Y_i + \sum_{i \in \mathcal{M}} \int_0^{m_i} (p_{ij} - t_{ij}) Y_{ij} dj + \sum_{i \in \mathcal{L}} (p_i - t_i) \bar{X}_i \\ &+ \bar{N} \left[ \sum_{i \notin \mathcal{M}} t_i \frac{dx_i}{dn} + \sum_{i \in \mathcal{M}} \int_0^{m_i} \left( t_{ij} \frac{dx_{ij}}{dn} \right) dj \right] + \frac{\bar{N}}{N} \sum_{i \in \mathcal{M}} t_{m_i} \frac{dm_i}{dn} = \text{FS} - \text{EB}, \end{aligned} \quad (32)$$

which proves the theorem. The simplified formula in the theorem for the case of a single differentiated good is finally obtained by solving (26) for  $dm/dn$  and by substituting the result into EB.

## A2. Proof of Theorem 2

The fiscal surplus consists of the total profit and the aggregate land rents. Since we assume that all goods other than the differentiated good have no price distortion, the profits evaluated at producer prices are zero except for the differentiated good. The producer price of the differentiated good is the marginal cost,  $c$ , and the profit of a differentiated good producer (evaluated at producer prices) is  $-F$ . Hence, the total profit in the differentiated sector (evaluated at producer prices) is  $\Pi = -mF < 0$ . Since the land rent is, by assumption, not distorted, the fiscal surplus is then simply given by:

$$\text{FS} = -mF + \sum_{i \in \mathcal{L}} p_i \bar{X}_i.$$

The excess burden can be obtained as follows. Using (11) from Theorem 1 and imposing symmetry for the varieties of the differentiated good, the excess burden in the present case is given by:

$$\text{EB} = -\bar{N} \left[ mp \left( \frac{t}{p} - \frac{t_m}{p_m} \right) \frac{dx}{dn} - \frac{t_m}{p_m} \left( \frac{dx_0}{dn} + \sum_{i \in \mathcal{H}} p_i \frac{dx_i}{dn} \right) \right].$$

We now investigate the signs of  $t/p$  and of  $t_m/p_m$ . First, let  $\delta > 0$  denote the Lagrange multiplier. Then, under the regularity assumptions, the first-order condition for optimal consumption can be uniquely solved to yield the compensated demand function

$$x = d(p, \delta, \mathbf{p}, \bar{U}), \quad \text{with} \quad \frac{\partial d(p, \delta)}{\partial p} = \frac{1}{\delta u''} < 0.$$

Furthermore, from the first-order conditions (4) for producers,  $t/p = (p - c)/p$  is the gross price margin satisfying

$$\frac{t}{p} = \frac{1}{\zeta}, \quad \text{where} \quad \zeta \equiv -\frac{\partial d}{\partial p} \frac{p}{d} = -\frac{u'}{u''x} = \frac{1}{R_R(x)} > 0$$

is the price elasticity of demand or, equivalently, the inverse of the RRA. Since the demand curve is downward sloping, the price distortion  $t$  is always positive.

Second, from (5),  $p_m - t_m = cY + F$  so that  $t_m/p_m$  satisfies

$$\frac{t_m}{p_m} = \frac{p_m - (cY + F)}{p_m}. \quad (33)$$

In the second-best with free entry, the zero-profit condition implies that  $(p - c)Y = F$ . From this relationship and  $Y = Nx$ , expression (33) becomes

$$\frac{t_m}{p_m} = \frac{p_m - Npx}{p_m} \geq 0, \quad (34)$$

where the inequality follows from (3).

Using the first-order conditions for the consumer in (1) and (3), the consumer side shadow price of product diversity can be written as

$$p_m = Np \frac{\partial U / \partial m}{\partial U / \partial x_i}. \quad (35)$$

The additively separable utility function satisfies  $\partial U / \partial m = u(x)$  and  $\partial U / \partial x_i = u'(x)$  for any variety with positive consumption. Hence, we obtain  $(\partial U / \partial m) / (\partial U / \partial x_i) = u(x) / u'(x)$  which allows to rewrite (35) as

$$p_m = Np \frac{u(x)}{u'(x)}.$$

Combining this expression with (34) yields

$$\frac{t_m}{p_m} = 1 - \frac{xu'}{u} = 1 - U_R(x), \quad (36)$$

where  $U_R(x)$  denotes the scale elasticity of utility. The excess burden can therefore be expressed as in equation (14).

Next, the sign of  $dx/dn$  can be obtained as follows. From the zero-profit condition and from  $p = c/(1 - R_R(x))$ , we readily obtain

$$\pi = (p - c)Nx - F = \frac{R_R(x)}{1 - R_R(x)} cx \frac{\bar{N}}{n} - F = 0. \quad (37)$$

For convenience, let us rewrite this expression as follows:

$$xR_R(x) + \phi n [R_R(x) - 1] = 0, \quad (38)$$

where  $\phi \equiv F/(c\bar{N}) > 0$  is a bundle of parameters. Differentiating (38) with respect to  $x$  and  $n$ , and using the zero-profit condition, we obtain

$$\frac{dx}{dn} = \phi \frac{1 - R_R(x)}{R_R(x)} \varpi \frac{1}{\theta_R + \varpi}, \quad \text{where } \varpi = \frac{cY}{cY + F}$$

denotes the share of the variable cost in the total cost of producing a variety of the differentiated good. Since the profit margin in (37) must be positive, we have  $1 - R_R(x) > 0$ . Hence,

$$\frac{dx}{dn} \gtrless 0 \quad \text{as } \theta_R + \varpi \gtrless 0.$$



### A3. Proof of Result 1

The only step that has not been proved is the sign of  $dx_0/dn$ . In order to sign this derivative, observe that the first-order conditions for expenditure minimization with respect to  $x_0$  and  $x_j$  imply that

$$\frac{1-\eta}{\eta} \frac{x_0}{\int_0^m x_j^{\frac{\sigma-1}{\sigma}} dj} \frac{\sigma-1}{\sigma} x_j^{-\frac{1}{\sigma}} = p_j.$$

At a symmetric equilibrium, this expression reduces to

$$\frac{1-\eta}{\eta} \frac{x_0}{mx} \frac{\sigma-1}{\sigma} = p. \quad (39)$$

From the zero-profit condition  $(p-c)Nx - F = 0$  and the profit maximizing price  $p = \sigma c/(\sigma-1)$ , we obtain the following relationship between the number of communities  $n$  and consumption of the differentiated good  $x$ :

$$x = (\sigma-1) \frac{F}{c\bar{N}} n. \quad (40)$$

Plugging this into (39) then yields a relationship between  $x_0$  and  $n$  as follows:

$$x_0 = \frac{\eta}{1-\eta} \frac{\sigma^2}{\sigma-1} \phi c m n, \quad (41)$$

where  $\phi \equiv F/(c\bar{N})$  is a constant term. Substituting both (40) and (41) into the utility function (15) and setting  $U(\mathbf{x}) = \bar{U}$  then yield:

$$m = \frac{\bar{U}}{\left[ \frac{\eta}{1-\eta} \left( \frac{\sigma}{\sigma-1} \right)^2 c \right]^\eta [(\sigma-1)\phi]^{n+\frac{\sigma-1}{\sigma}(1-\eta)}} n^{-\eta-\frac{\sigma-1}{\sigma}(1-\eta)}. \quad (42)$$

Since  $\eta \in (0, 1)$  and  $\sigma > 1$ , it is readily verified that

$$\frac{dm}{dn} = - \left[ \eta + \frac{\sigma-1}{\sigma}(1-\eta) \right] \frac{m}{n} < 0. \quad (43)$$

From (41) we have

$$\frac{dx_0}{dn} = \frac{x_0}{n} + \frac{x_0}{m} \frac{dm}{dn}, \quad (44)$$

which, using (43), can finally be expressed as follows:

$$\frac{dx_0}{dn} = x_0 \left\{ \frac{1}{n} - \frac{1}{n} \left[ \eta + \frac{\sigma-1}{\sigma}(1-\eta) \right] \right\} = \frac{x_0}{\sigma n} (1-\eta) > 0.$$

### A4. Proof of Result 2

It remains to be proved that  $dx/dn > 0$ . As shown in Appendix A2, the zero-profit condition can be expressed as in (38). Substituting (18) into that condition and recalling that  $\phi \equiv F/(c\bar{N})$  yield

$$x^2 - (\sigma-1)\phi n x + \sigma \gamma \phi n = 0. \quad (45)$$

Note that the smaller root of this quadratic equation is either negative or unstable so that we can concentrate on the larger root. To see this, note that the profit of a producer can be expressed as follows:

$$\pi(x) = \frac{c\bar{N}}{n} \left[ \frac{xR_R(x)}{1 - R_R(x)} - \phi n \right] = \frac{c\bar{N}}{n} \left[ \frac{x^2}{(\sigma - 1)x - \sigma\gamma} - \phi n \right],$$

so that

$$\frac{d\pi}{dx} = x \frac{c\bar{N}}{n} \frac{(\sigma - 1)x - 2\sigma\gamma}{[(\sigma - 1)x - \sigma\gamma]^2}.$$

When  $\gamma \leq 0$ , the profit is increasing in  $x$  for non-negative values of  $x$ , as depicted in the top panel of Figure 3. In this case, the smaller root is negative and the larger root is the only possible solution.

**Insert Figure 3 about here.**

In the DRRA case where  $\gamma > 0$ , the profit is increasing in  $x$  when  $x > 2\sigma\gamma/(\sigma - 1)$ , but decreasing when  $x < 2\sigma\gamma/(\sigma - 1)$ . Furthermore, the profit tends to  $+\infty$  when  $x$  approaches  $\sigma\gamma/(\sigma - 1)$  from above, and to  $-\infty$  when approached from below. As depicted in the bottom panel of Figure 3, if real roots exist, both of them are positive in this case. It is easy to see that the equilibrium associated with a smaller  $x$  is unstable.<sup>18</sup>

The larger root of the zero profit condition is computed as follows:

$$x = \frac{(\sigma - 1)\phi n + \sqrt{\phi n[(\sigma - 1)^2\phi n - 4\sigma\gamma]}}{2},$$

which yields

$$\frac{dx}{dn} = \frac{1}{2} \left[ (\sigma - 1) + \frac{(\sigma - 1)^2\phi n - 2\sigma\gamma}{\sqrt{\phi n[(\sigma - 1)^2\phi n - 4\sigma\gamma]}} \right] \phi.$$

If  $\gamma \leq 0$ , we readily obtain  $dx/dn > 0$ . When  $\gamma > 0$ , we require that  $(\sigma - 1)^2\phi n - 4\sigma\gamma > 0$  for real roots to exist. In that case, we also have  $(\sigma - 1)^2\phi n - 2\sigma\gamma > 0$ , thus showing that  $dx/dn > 0$ .

## A5. Proof of Result 3

Let us examine the sign of  $\theta_U$  for the case of  $a > 0$ . First, it is obvious that  $\theta_U(0) = 1$  when  $a \neq 1$ . When  $a = 1$ , by l'Hôpital's rule, we have

$$\theta_U(0) = \lim_{x \rightarrow 0^+} \theta_U(x) = \lim_{x \rightarrow 0^+} \left( 1 - \frac{\alpha x}{1 - e^{-\alpha x}} \right) = \lim_{x \rightarrow 0^+} \left( 1 - \frac{\alpha}{\alpha e^{-\alpha x}} \right) = 0.$$

Next, the derivative of  $\theta_U(x)$  with respect to  $x$  satisfies

$$\frac{\partial \theta_U(x)}{\partial x} = -a\alpha \frac{a - (1 + \alpha x)e^{-\alpha x}}{(a - e^{-\alpha x})^2} \leq 0,$$

---

<sup>18</sup>Since a firm exits when it makes losses, if  $x$  gets slightly larger than the smaller root  $x_L$ , the number of firms decreases, which increases consumption  $x$  of each variety and thus moves the equilibrium towards the larger root  $x_H$ . If  $x$  happens to become smaller than  $x_L$ , the number of firms increases, which makes  $x$  even smaller and so on. Using a similar reasoning, we can show that the larger root  $x_H$  is stable.

where the inequality is again obtained from (17). Hence, if  $a = 1$ , then  $\theta_U(0) = 0$  and it decreases as  $x$  gets larger. In the case of  $a > 1$ ,  $\theta_U(0) = 1$  and  $\theta_U(x)$  decreases monotonically as  $x$  gets larger. In the case of  $0 < a < 1$ ,  $\theta_U(0) = 1$  but  $x = 0$  does not satisfy the inequality in (17). For the smallest  $x$  that satisfies the inequality in (17), we have  $a - e^{-\alpha\tilde{x}} = \alpha\tilde{x}e^{-\alpha\tilde{x}}$ , which implies  $\theta_U(\tilde{x}) = -\alpha\tilde{x} < 0$ . From this point on,  $\theta_U(x)$  decreases monotonically as  $x$  gets larger.

Next, let us examine the sign of  $dx/dn$ . As shown in Appendix A2, the zero-profit condition can be expressed as in (38). Substituting (21) into that condition and recalling that  $\phi \equiv F/(c\bar{N})$  yield

$$\alpha x^2 + \phi n(\alpha x - 1) = 0. \quad (46)$$

This equation always has one positive root, given by

$$x = \frac{1}{2\alpha} \left[ -\alpha\phi n + \sqrt{\alpha\phi n(4 + \alpha\phi n)} \right]. \quad (47)$$

It is readily verified that this solution satisfies

$$\frac{dx}{dn} = \frac{1}{2} \left[ -1 + \frac{2 + \alpha\phi n}{\sqrt{\alpha\phi n(4 + \alpha\phi n)}} \right] \phi > 0. \quad (48)$$

## Appendix B: Additional expressions for simulations

Assume, for simplicity, that there are four goods: labor (the numeraire), one differentiated good, housing, and land. The labeling of these four goods is as follows:  $x_0$  denotes the numeraire;  $x_i$ ,  $i \in [0, m]$  denotes variety  $i$  of the differentiated good;  $x_H$  stands for housing; and  $x_L$  denotes land.

Each consumer has one unit of labor and supplies it inelastically to producers of consumption goods and housing. The total labor endowment in the economy is  $\bar{N}$ , whereas the labor endowment of each community is  $N = \bar{N}/n$ . We assume that there is a one-to-one relationship between varieties and firms, and that labor is the only factor of production for the differentiated good. Production of  $Y_i$  units of any variety requires  $cY_i + F$  units of labor. Since we focus on symmetric equilibria only ( $p_i = p$  and  $x_i = x$  for all varieties), the profit of a firm is given by  $\pi = (p - c)Y - F$ , and it is zero at equilibrium because of free entry and exit.

Housing is produced by combining labor and land, and housing production in any community is denoted by  $Y_H$ . The supply of land is exogenously fixed at  $\bar{S}$  for all communities. To produce  $Y_H = N\bar{h}$ , i.e.,  $N$  lots of a fixed quality  $\bar{h}$ , requires  $-Y_L = \bar{S}$  units of land and  $bN^2$  units of labor (remember that negative outputs denote inputs). We may view the problem as one where fixed quality housing  $\bar{h}$  is produced by combining  $s$  units of land and  $b\bar{S}/s$  units of labor. The cost of supplying housing of quality  $\bar{h}$  (given the choice of  $s$ ) is then  $c^H(s) = p_L s + b\bar{S}/s$ , where  $p_L$  is the price of land. Cost minimization in the housing sector yields

$$\frac{\partial c^H}{\partial s} = r - \frac{b\bar{S}}{s^2} = 0, \quad \text{and hence} \quad r = \frac{b\bar{S}}{s^2}.$$

Since by definition  $s = \bar{S}/N$ , we have  $p_L = (bN^2)/\bar{S}$ . Hence, the aggregate land rents in the community are  $ALR \equiv bN^2$ . The total labor input in housing production is given by  $-N^2b$ .<sup>19</sup>

Assuming constant returns to scale in the housing sector, the profit in that sector is given by

$$\pi_H = p_H N \bar{h} - N^2 b - p_L \bar{S},$$

which must be zero at equilibrium. Hence,  $p_H = (2bN)/\bar{h}$ . Last, the fiscal surplus is just given by the difference between aggregate land rents and aggregate fixed costs in the differentiated sector, as shown by Theorem 2:  $FS = bN^2 - mF$ .

Observe that the market clearing condition for labor requires that

$$A = \bar{N} - n \left\{ m \left[ cx(n) \frac{\bar{N}}{n} + F \right] + b \left( \frac{\bar{N}}{n} \right)^2 \right\},$$

where the first term in braces denotes the labor used for the production of differentiated varieties, and where the second term denotes the labor used to produce housing. As before,  $A$  denotes the Allais surplus. Since  $U = mu(x(n)) = \bar{U}$  at any symmetric equilibrium, we have  $m = \bar{U}/u(x(n))$ , which allows us to rewrite the foregoing expression as follows:

$$A = \bar{N} - n \left\{ \frac{\bar{U}}{u(x(n))} \left[ cx(n) \frac{\bar{N}}{n} + F \right] + b \left( \frac{\bar{N}}{n} \right)^2 \right\}.$$

The optimal number of communities  $n$  satisfies  $dA/dn = 0$ . It is slightly more convenient for numerical purposes to rewrite the foregoing problem in terms of the optimal community size  $N$ :

$$\frac{A}{\bar{N}} = 1 - \frac{\bar{U}}{u(x(N))} \left[ cx(N) + \frac{F}{N} \right] - bN,$$

where we may view  $x$  as a function of  $N$ .

Given the CARA sub-utility used in Section 4.2.2, we can then easily run numerical simulations and draw pictures like Figure 2.

---

<sup>19</sup>Note that similar expressions obtain in standard monocentric city models with linear commuting costs and fixed lot size (Fujita, 1989).

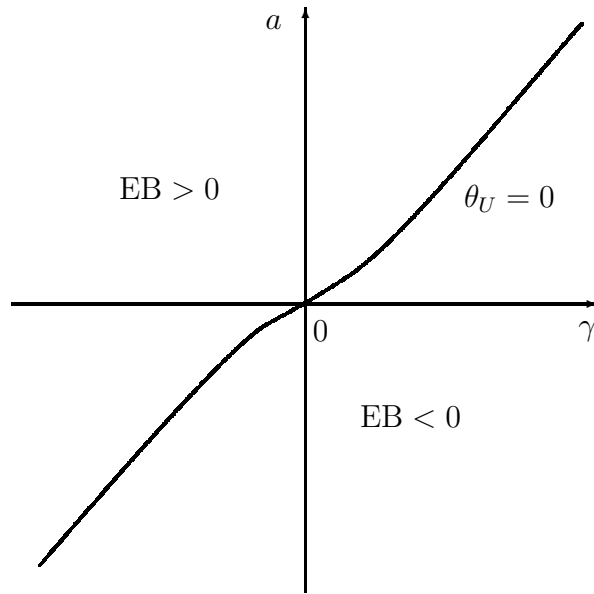


Figure 1. Variable RRA case – sign of excess burden in  $(\gamma, a)$ -space

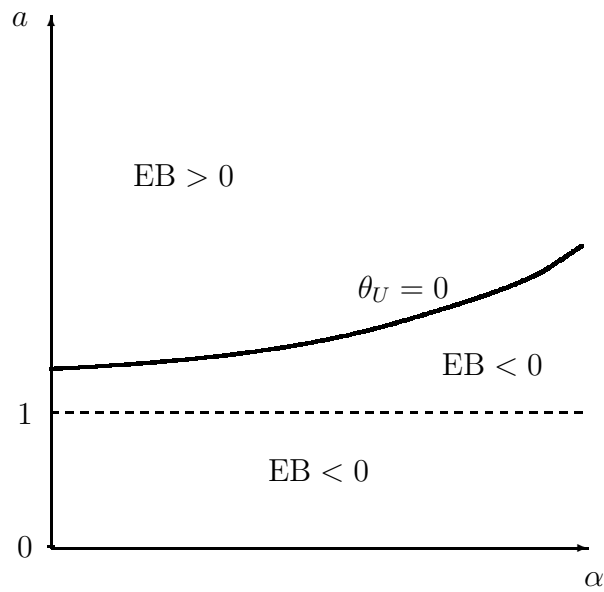


Figure 2. Generalized CARA case – sign of excess burden in  $(\alpha, a)$ -space

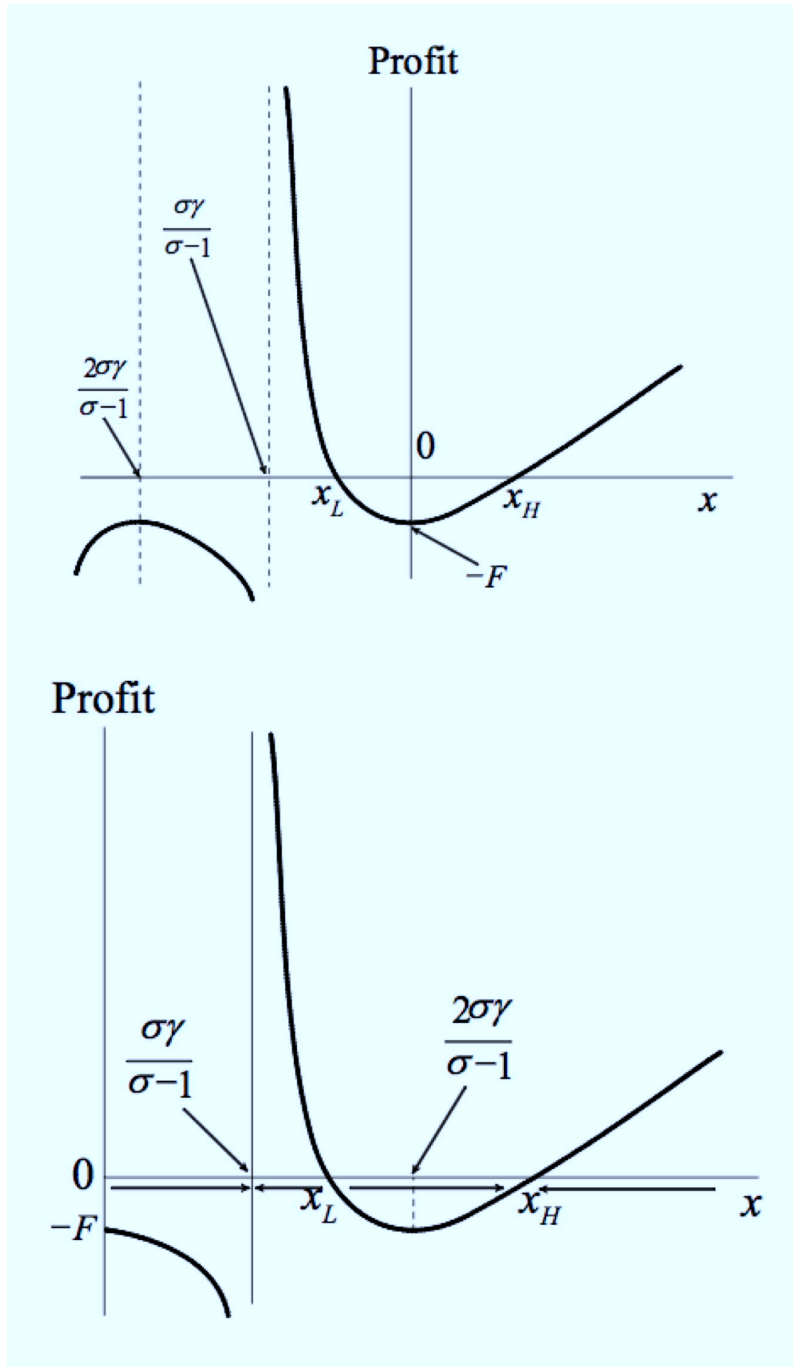


Figure 3. Equilibrium  $x$  in the variable RRA case